

Universidade Estadual de Campinas
Faculdade de Engenharia Elétrica e de Computação

Gestão de conteúdos em um ambiente de ensino de línguas baseado em Sistema Tutor Inteligente com agentes-lexemas e semantização de imagens

Autor: Ismael Mattos Andrade Ávila

Orientador: Prof. Dr. Ricardo Ribeiro Gudwin

Dissertação de Mestrado apresentada à Faculdade
de Engenharia Elétrica e de Computação como parte
dos requisitos para obtenção do título de Mestre em
Engenharia Elétrica. Área de concentração:
Engenharia de Computação.

Banca Examinadora

Prof. Edson Françaço.....IEL
Prof. Marco Aurélio A. Henriques.....FEEC
Prof. Ricardo R. Gudwin.....FEEC

Campinas, SP
Novembro/2010

FICHA CATALOGRÁFICA ELABORADA PELA
BIBLIOTECA DA ÁREA DE ENGENHARIA E ARQUITETURA - BAE -
UNICAMP

Av55g Ávila, Ismael Mattos Andrade
Gestão de conteúdos em um ambiente de ensino de
línguas baseado em sistema tutor inteligente com
agentes-lexemas e semantização de imagens / Ismael
Mattos Andrade Ávila. --Campinas, SP: [s.n.], 2010.

Orientador: Ricardo Ribeiro Gudwin.
Dissertação de Mestrado - Universidade Estadual de
Campinas, Faculdade de Engenharia Elétrica e de
Computação.

1. Sistemas tutoriais inteligentes. 2. Sistemas
multiagentes. 3. Semiótica e computação. 4. Ensino a
distância. 5. Metadados. I. Gudwin, Ricardo Ribeiro. II.
Universidade Estadual de Campinas. Faculdade de
Engenharia Elétrica e de Computação. III. Título.

Título em Inglês: Content management in a language teaching environment
based on intelligent tutoring system with lexeme-agents
and image semantization

Palavras-chave em Inglês: Intelligent tutoring systems, Multiagent systems,
Semiotic and computing, Distance learning,
Metadata

Área de concentração: Engenharia de Computação

Titulação: Mestre em Engenharia Elétrica

Banca examinadora: Edson Françaço, Marco Aurélio Amaral Henriques

Data da defesa: 29/11/2010

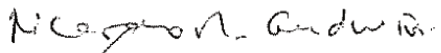
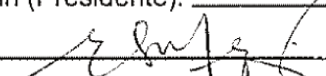
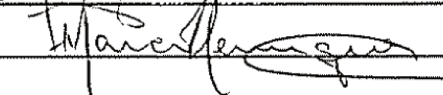
Programa de Pós Graduação: Engenharia Elétrica

COMISSÃO JULGADORA - TESE DE MESTRADO

Candidato: Ismael Mattos Andrade Ávila

Data da Defesa: 29 de novembro de 2010

Título da Tese: "Gestão de conteúdos em um ambiente de ensino de línguas baseado em Sistema Tutor Inteligente com agentes-lexemas e semantização de imagens"

Prof. Dr. Ricardo Ribeiro Gudwin (Presidente): 
Prof. Dr. Edson Françaço: 
Prof. Dr. Marco Aurélio Amaral Henriques: 

Resumo

Este trabalho propõe uma arquitetura baseada em um sistema tutor inteligente (STI) para ambientes de ensino de línguas. A principal contribuição da proposta está no mecanismo de seleção e sequenciamento dos conteúdos, e baseia-se na estratégia de tratar cada lexema da língua-alvo (L2) como uma unidade pedagógica autônoma, cujo objetivo é ser aprendida instanciada na forma de um agente de software. Essa estratégia cria um arranjo *bottom-up* que é flexível o bastante para adaptar-se a mudanças nos objetivos pedagógicos e para reproduzir a natureza gradual, contextualizada e em grande medida inconsciente da aquisição da língua materna (L1). A arquitetura proposta é baseada em três principais pilares: (i) os agentes-lexemas da língua-alvo (L2) competem por um recurso limitado, a interface do curso; (ii) as imagens usadas nas cenas do curso são associadas a metadados que as descrevem semanticamente, influenciando no comportamento dos agentes e (iii) uma ontologia da L2 define as relações semânticas e sintáticas entre seus lexemas e assim embasa as inferências feitas pelo sistema multiagente.

Palavras-chave: Semiótica computacional, Inteligência Artificial, Sistemas Tutores Inteligentes, Sistemas Multiagentes, Ontologias, Metadados, Web Semântica.

Abstract

This work proposes an architecture based on an Intelligent Tutoring System (ITS) for language teaching environments. The main contribution of the proposal is in its content selection and sequencing mechanism, and is based on the strategy of treating each lexeme of the target language (L2) as an autonomous pedagogical unit, whose goal is to be learned instantiated as a software agent. This strategy creates a bottom-up approach that is flexible enough to adapt to changes in the pedagogical objectives and to reproduce the gradual, contextualized and somewhat unconscious nature of the first language (L1) acquisition. The proposed architecture is based on three main pillars: (i) the lexeme-agents of the target language (L2) compete for a limited resource, the course interface; (ii) the images used in the course scenes are associated to metadata that describe them semantically, influencing the behavior of the agents and (iii) an ontology of L2 defines the semantic and syntactic relations of its lexemes and then underpins the inferences made by the multiagent system.

Keywords: Computational Semiotics, Artificial Intelligence, Intelligent Tutoring Systems, Multiagent Systems, Ontologies, Metadata, Semantic Web.

Aos meus familiares

Agradecimentos

Ao meu orientador, Prof. Ricardo Ribeiro Gudwin, sou grato pela orientação, pelo apoio, e pela paciência.

A minha esposa Marilda pelo apoio e pelos incentivos constantes.

Aos meus filhos Helena e Pedro pela compreensão por minha indisponibilidade frequente.

A minha mãe Eliana e a meu pai Antônio (*in memoriam*).

À minha família pelo apoio durante esta jornada.

Aos queridos amigos Rosely e José Carlos.

Aos colegas do grupo de discussão sobre consciência artificial, Elisa, Vera, André, Klaus e Anderson pelas valorosas discussões.

Aos colegas de trabalho pelo companheirismo.

Aos membros da banca examinadora: Prof. Dr. Ricardo Ribeiro Gudwin, Prof. Dr. Edson Françaço, Prof. Dr. Marco Aurélio A. Henriques, e aos membros suplentes Prof. Antonio A. C. Gomide e Prof. Dr. David Bianchini.

Aos membros da banca de qualificação: Prof. Dr. Ricardo Ribeiro Gudwin, Prof. Dr. Ivan Ricarte e Prof. Dr. Edson Françaço.

Sumário

1	Introdução	1
1.1	Estudo de caso para o ambiente de ensino.....	7
1.2	Organização da dissertação.....	8
2	Ensino mediado pelo computador.....	9
2.1	Aprendizado de línguas assistido por computador – CALL.....	9
2.2	Adaptatividade e adaptabilidade em ambientes educacionais.....	11
2.3	A IA na educação: Sistemas Tutores Inteligentes – ITS.....	13
2.4	Abordagens dos ambientes computacionais de aprendizagem.....	15
2.5	Taxonomia dos ambientes computacionais de aprendizagem.....	16
2.6	Ontologias em sistemas educacionais e no ensino de uma L2.....	19
2.7	Modelagem do domínio baseada em agentes.....	20
3	O aprendizado de uma língua.....	25
3.1	A aquisição da língua materna.....	25
3.2	O ritmo da aquisição das palavras.....	27
3.3	As bases da aquisição das palavras.....	28
3.4	A teoria ecológica da aquisição de uma língua.....	31
3.5	O aprendizado de uma L2 segundo a teoria Semiótica.....	32
3.6	Fatores relevantes no ensino de uma L2.....	34
4	O processo de interpretação de imagens	39
4.1	Considerações iniciais sobre a percepção de imagens.....	39
4.2	A interpretação de imagens segundo a Semiótica de Peirce.....	42
4.3	A inter-relação de imagens visuais e mentais.....	46
4.4	A inter-relação de imagens e textos em uma cena.....	50
4.5	A influência de fatores culturais na interpretação de imagens.....	51
4.6	Iconicidade intersimbólica e iconicidade intersintática.....	53
5	Arquitetura ILTS e a gestão dos conteúdos	55
5.1	Modelagem de domínio proposta.....	55
5.2	Uma ontologia do léxico da L2.....	56
5.3	Gestão do léxico baseada em agentes	62
5.4	Semantização das imagens de apoio do ILTS.....	64
5.5	Arquiteturas de referência para o ILTS aqui proposto	68
5.7	Transposição da arquitetura CTS para o ILTS.....	74
5.8	Configuração esquemática da arquitetura proposta.....	78
6	Simulação ABMS da gestão de conteúdos.....	83
6.1	Condições de realização das simulações.....	83
6.2	Uso de imagens de apoio baseadas em desenhos.....	88
7	Discussões e considerações finais.....	91
7.1	Papéis e atores para a viabilização da arquitetura.....	93
7.2	Desenvolvimentos futuros.....	93
7.3	Perspectivas da arquitetura.....	94

7.4	Considerações finais.....	94
8	Apêndice 1.....	95
	Resumo das principais teorias sobre aprendizado linguístico.....	95
	História dos métodos de ensino de uma segunda língua.....	107
	Os métodos da leitura e oral.....	109
	Os métodos audiolingual e audiovisual.....	110
	Método do código cognitivo.....	110
	Método comunicativo.....	111
	Métodos humanistas.....	112
9	Apêndice 2.....	115
	Aspectos principais da teoria Semiótica.....	115
10	Apêndice 3.....	121
	Resumo das principais características do esperanto.....	121
	A estruturação do esperanto.....	127
11	Referências bibliográficas.....	133

Lista de Figuras

Figura 1.1: Vários possíveis contextos de ensino de línguas.....	2
Figura 1.2: Cenários de ensino a distância de línguas mediado por TICs.....	3
Figura 1.3: Disciplinas relevantes para o EaD de línguas mediado por computadores.....	4
Figura 1.4: Disciplinas relevantes para o EaD de línguas via Web.....	5
Figura 1.5: Foco de contribuição deste trabalho.....	6
Figura 2.1: a) Áreas de conhecimento em um ITS; b) Foco do estudo.....	13
Figura 2.2: Relação entre aprendiz e sistema tutor.....	21
Figura 4.1: Agrupamentos por similaridade.....	38
Figura 4.2: Agrupamentos por proximidade.....	38
Figura 4.3: Agrupamentos por continuidade.....	38
Figura 4.4: Escada de Schröder – exemplo de ilusão de ótica.....	40
Figura 4.5: Exemplo de arranjo visual no curso.....	46
Figura 4.6: Sequência da interpretação visual da cena.....	47
Figura 5.1: Superclasses da ontologia para seres vivos e conceitos relacionados.....	56
Figura 5.2: Classes de adjetivos dentro da superclasse “eco” (qualidade).....	57
Figura 5.3: Classes de verbos dentro da superclasse “ago” (ação).....	57
Figura 5.4: Classe “rinocero” dentro da superclasse “mamulo” (mamífero).....	58
Figura 5.5: Regras de relacionamento entre classes.....	59
Figura 5.6: Exemplo de atributos para o conceito “rinocero”.....	59
Figura 5.8: Exemplo de anotação de metadado em imagem.....	65
Figura 5.8: Exemplos de anotações do lexema principal e dos relacionados.....	65
Figura 5.9: Braço robótico da Estação Espacial Internacional.....	69
Figura 5.10: A metáfora do teatro da consciência.....	70
Figura 5.11: Segmentação dos codelets na implementação.....	75
Figura 5.12: Rede de comportamento e mecanismo de consciência.....	76
Figura 5.13: Síntese da arquitetura proposta para o ILTS.....	79
Figura 5.14: Diagrama de blocos do ILTS e sua relação com conteúdos da Web.....	80
Figura 6.1: Condição inicial: pano de fundo define o contexto da cena.....	84
Figura 6.2: Agentes importados permanecem ativos mas ocultos.....	85
Figura 6.3: Agentes importados deslocam-se aos elementos indicados pelos metadados.....	85
Figura 6.4: Processo de rotulagem dos elementos, destaque para o lexema-foco.....	86
Figura 6.5: Lexema-foco inicia sentença e um dos lexemas correlatos junta-se a ela.....	86
Figura 6.6: Nível de ativação do agente-lexema vencedor da primeira rodada.....	87
Figura 6.7: Nível de ativação do agente-lexema vencedor da última rodada.....	87
Figura 6.8: Exemplo de cena baseada em desenhos - destacado o agente “lito”.....	88
Figura 6.9: Exemplo de cena baseada em desenhos - destacado o agente “fenestro”.....	89
Figura A.2.1: As três dimensões do signo na Semiótica de Peirce.....	116
Figura A.3.1: Exemplo de interônimo - a palavra kato.....	123

Lista de Tabelas

Tabela 2.1: Disciplinas e áreas relevantes para o estudo da CALL.....	10
Tabela 2.2: Identificação das características do ambiente de ensino proposto.....	18
Tabela 3.1: Fatores relevantes no ambiente de ensino aqui proposto.....	35
Tabela 5.1: Exemplo dos campos de um agente-lexema.....	65
Tabela 5.2: Fontes de léxico para a formação das sentenças-exemplo.....	77
Tabela A.1.1: Teorias de aquisição linguística.....	101
Tabela A.1.2: Incompatibilidade de esquemas mentais de L1-L2.....	103
Tabela A.1.3: Incompatibilidade de esquemas mentais de L1-L2.....	103
Tabela A.3.1: Exemplos de modificação de raízes.....	122
Tabela A.3.2: Formação dos adjetivos em português e em esperanto.....	125
Tabela A.3.3: Formação dos substantivos abstratos em português e em esperanto.....	130

Glossário

ABMS - Agent-Based Modeling Systems
ACA - Ambientes Computacionais de Aprendizado
ALE - Adaptative Learning Environment
BS - Buffer Sensorial
BN - Behavior Network (Rede de Comportamentos)
CAI - Computer Assisted Instruction
CALL - Computer-Assisted Language Learning
CBM - Constraint-Based Modeling
CLN - Compreensão de Linguagem Natural
CTS - Conscious Tutoring System
EaD - Ensino a Distância
EBW - Educação Baseada na Web
EH - Educational Hypermedia
EIBW - Educação Inteligente Baseada na Web
GU - Gramática Universal
GW - Global Workspace
IA - Inteligência Artificial
IAD - Inteligência Artificial Distribuída
IAED - Inteligência Artificial na Educação
IC - Inteligência Computacional
IDA - Intelligent Distribution Agent
ILE - Interactive Learning Environment
ILTS - Intelligent Language Tutoring Systems
ITS - Intelligent Tutoring Systems (ver STI)
L1 - Primeira língua ou língua materna
L2 - Segunda língua ou língua-alvo
LN - Língua natural
MMCALL- Multimedia Computer-Aided Language Learning
MT - Memória de Trabalho
ODL - Open Distance Learning
RB - Redes de Bayes
SA - Seleção de Ação
STI - Sistemas Tutores Inteligentes (ver ITS)
TIC - Tecnologias de Informação e Comunicação
TFIDF - Term Frequency Inverse Document Frequency
UP - Unidade Pedagógica

Trabalhos Publicados pelo Autor

Ávila, I.; Gudwin, R. (2010) *Geração e sequenciamento automáticos de lexemas e sentenças em Sistema Tutor Inteligente para ensino de línguas*. In: Anais do Simpósio Brasileiro de Informática na Educação – SBIE 2010, João Pessoa, Brasil.

Ávila, I.; Osorio, A.; Rolim, L. (2010) *Personas na interação com serviços de e-gov via Web: resultados qualitativos*. In: Anais do Simpósio Brasileiro de Sistemas Multimídia e Web – WebMedia 2010, Belo Horizonte, Brasil.

Rolim, L.; Osorio, A.; Ávila, I. (2010) *Ambiente Colaborativo para Anotação Semântica de Conteúdos Audiovisuais*. In: Anais do Simpósio Brasileiro de Sistemas Colaborativos – SBSC 2010, Belo Horizonte, Brasil.

Ávila, I.; Gudwin, R. (2010) *Image semantization with MPEG-7 descriptors in a Web language course*. Poster apresentado na Summer School on Multimedia Semantics – SSMS 2010, Amsterdã, Holanda.

Ávila, I.; Osorio, A.; Rolim, L. (2010) *From collaborative annotation to collaborative production of audiovisual contents*. Poster apresentado na Summer School on Multimedia Semantics – SSMS 2010, Amsterdã, Holanda.

Ávila, I.; Gudwin, R. (2010) *An ITS CALL system based on active lexeme-agents*. III Encontro dos Alunos e Docentes do Departamento de Engenharia de Computação e Automação Industrial (EADCA), Vol. CD-ROM, pp.25-27, Campinas, Brasil.

Ávila, I.; Gudwin, R. (2009) *Semiotics-Based Learner Modeling in Distance Language Courses*, VIII Encontro Brasileiro Internacional de Ciência Cognitiva (EBICC), Vol. CD-ROM, pp.1-1, Campinas, Brasil.

Ávila, I.; Costa, R.G. (2009) *Ícones como facilitadores da interação de usuários iletrados com interfaces computacionais*. In: Cadernos CPqD Tecnologia, v. 3, p. 15-36.

Ávila, I.; Rolim, L.; Holanda, G.M. (2009) *Alternative Neighborhood Configurations in an ABMS Model to Estimate the Adoption of Telecenters in Brazil*. In: IADIS Computer Science and Information Systems, v. V IV, p. 51-65.

Ávila, I.; Gudwin, R. (2009) *Icons as Helpers in the Interaction of Illiterate Users with computers*. In: Proc. of the IADIS Intl. Conf. on Interfaces & Human Computer Interaction 2009, Carvoeiro, Portugal.

Ávila, I.; Pinto, J.C.L.; Rolim, L. (2009) *Screen Reader for Illiterate Users*. In: Proc. of the IADIS International Conf. on Interfaces and Human Computer Interact., Algarve, Portugal.

Ávila, I.; Gudwin, R. (2009) *Lexical similarity metrics for vocabulary learning modeling in Computer-Assisted Language Learning (CALL)*. In: Proc. of the AIED Workshop on Natural Language Processing in support of Learning, v. 10. p. 9-17, Brighton, Reino Unido.

Ávila, I.; Costa, R.G. (2009) *Desenho Participativo de Ícones para Interfaces Computacionais Voltadas a Usuários Analfabetos*. In: Proc. of the Interaction '09 - South America, São Paulo, Brasil.

Ávila, I.; Piccolo, L.; Carvalho, R.F. ; Zago, A. (2009) *Personas como Facilitadoras da Interação com Serviços de Governo Eletrônico*. In: Proc. of the Latin American Conf. on Human-Computer Interaction CLIHC 2009, Mérida, México.

Tambascia, C.A.; Ávila, I.; Holanda, G.M. (2008) *Digital Accessibility for illiterates: a novel interface and interaction model*. In: Proc. of the Applied Human Factors and Ergonomics, Las Vegas, EUA.

Ávila, I.; Menezes, E.; Piccolo, L.; Tambascia, C.A. (2007) *Interaction Models for Digital Inclusion of Low-literacy, Aged and Impaired Users in Brazil*. In: Proc. of the Latin American Conference on Human-Computer Interaction CLIHC2007, Rio de Janeiro, Brasil.

Bonadia, G.C.; Ávila, I.; Ogushi, C.M.; Holanda, G.M. (2007) *Modeling the Acceptance and Use of Telecenters in Brazil*. In: Journal of Technology Management and Innovation, v. 2, pp. 86-97.

Holanda, G.M.; Ávila, I.; Martins, R.B. (2006) *Mapping users' perspectives and outlining social impacts from digitalization of terrestrial TV in Brazil*. In: Telematics and Informatics, DOI: 10.1016/j.tele.2006.01.001. Disponível em: www.sciencedirect.com

Ávila, I.; Pinto, J.C.L.; Lemos, L.M.; Holanda, G.M. (2006) *Redes sociais e disseminação de inovações tecnológicas: uma modelagem por agentes aplicada ao caso da TV digital brasileira*. In: Cadernos CPqD Tecnologia, v. 2, p. 27-38.

Holanda, G.M.; Ávila, I. (2006) *Sociedade, tecnologia e exclusão*. In: Souto, Á.A.; Dall'Antonia, J.C.; Holanda, G.M. (Org.). Capítulo do livro: As cidades digitais no mapa do Brasil: uma rota para a inclusão social. 1ª ed. Brasília, DF: MiniCom, v. 1, p. 1-11.

Ávila, I.; Holanda, G.M. (2006) *Inclusão digital no Brasil: uma perspectiva sociotécnica*. In: Souto, Á.A.; Dall'Antonia, J.C.; Holanda, G.M. (Org.). Capítulo do livro: As cidades digitais no mapa do Brasil: uma rota para a inclusão social. 1ª ed. Brasília, DF: Ministério das Comunicações, v. 1, p. 13-60.

Lemos, L.M.; Pinto, J.C.L.; Ávila, I.; Bonadia, G.C.; Dall'Antonia, J.C. (2005) *Broadband Internet over Digital TV Platform: a Techno-Economic Analysis*. 16th European Regional Conf. – ITS, Porto, Portugal.

Capítulo 1

Introdução

A real essência da memória humana, segundo Lev Vigotsky (Vigotsky, 1978), é que os seres humanos se lembram ativamente por meio de símbolos. Essas palavras são, para Danesi (1997), a razão de se estabelecer uma conexão entre a Semiótica, a ciência dos signos, e a Educação, a ciência que trata de como ensinar os indivíduos a manejar signos.

O vínculo entre Semiótica e Educação se estende naturalmente à área da Semiótica Computacional, sobretudo quando esta é aplicada na concepção de ambientes de ensino assistidos por computador, pois tais ambientes se propõem, em última análise, a ampliar o repertório sógnico dos alunos.

É inquestionável a enorme relevância dos signos para a espécie humana, em especial dos símbolos, signos arbitrariamente definidos pelo contexto cultural. Para Deacon (1997), a capacidade simbólica, isto é, a habilidade de processar símbolos, é o traço que distingue a linguagem humana das demais modalidades de comunicação observadas na natureza. O processamento de símbolos, acrescenta Deacon, permite economia de referência, variedade infinita de novas representações, capacidade de predição, planejamento e organização da memória. Mas, por ser convencionado e arbitrário, e não inato, todo sistema simbólico deve ser assimilado antes de poder ser usado. Isso se aplica também à língua materna, ainda que tenhamos pouca consciência dessa aquisição como uma atividade fundamental para a nossa natureza humana. Se ela não ocorre, ou ocorre de maneira incompleta, como no caso das crianças autistas ou surdas, por exemplo, uma situação de exclusão desses indivíduos em relação àquele sistema simbólico e à sociedade que dele se serve é praticamente inevitável.

Na esfera internacional, todavia, não são só as deficiências sensoriais e intelectuais que causam exclusão linguística. Essa resulta também da falta de domínio de uma segunda língua (L2). Isso se tornou mais claro com o advento das novas Tecnologias de Informação e Comunicação – TICs. As novas TICs criaram na sociedade atual um novo paradigma na natureza temporal e espacial da produção e recepção de conteúdos informacionais que, nas palavras de (Thompson, 1995), corresponde a uma distanciação espaço-tempo entre a forma simbólica e seu contexto de produção, de tal maneira que o intercâmbio simbólico pode ocorrer entre pessoas que ocupam diferentes posições no tempo e no espaço. Com isso, a necessidade de domínio de símbolos se estende agora muito além das fronteiras da língua materna, tendo a sociedade em rede eliminado a noção de distância geográfica e exposto cada um de nós a um universo de interlocutores e de informações frente ao qual o nosso repertório local de símbolos se mostra insuficiente. Como o intercâmbio simbólico aludido por Thompson supõe uma base linguística comum a todos os interlocutores, as novas TICs fizeram com que a principal barreira entre os indivíduos conectados passasse a ser a da distância cultural, sobretudo na forma da barreira dos diferentes idiomas.

Assim, enquanto a efetiva participação na cibersociedade exige domínio de línguas, o ensino de línguas permanece sendo um desafio para a maior parte dos sistemas de ensino em todo o mundo, dada a complexidade de se aprender uma segunda língua (L2) em sala de

aula. Para compensar parte das limitações do ensino escolar de línguas, entre as quais a falta de efetiva exposição dos alunos à L2 em seu uso real, o acesso à Rede e o uso das TICs como instrumentos de ensino podem trazer ganhos e vantagens, como o contato com outras culturas e a percepção, pelos estudantes, de que a L2 já não é uma realidade tão distante geograficamente, o que pode fazer surgir uma nova motivação para seu aprendizado.

Esse foi então um dos principais motivadores do presente trabalho, cujo objetivo é contribuir para a área de ensino a distância – EaD de línguas mediado e assistido por TICs propondo uma nova solução de gestão de conteúdos que alie flexibilidade, individualização e escalabilidade. Esse tripé é crucial porque, embora o avanço da informática e da Internet nos sistemas educacionais de muitos países tenha dado novo impulso aos ambientes de ensino assistido por computador, essa evolução só trará ganhos de escala e de portabilidade se um mesmo recurso pedagógico puder ser usado por um número expressivo de alunos, dispersos geograficamente e culturalmente, na forma de objetos de aprendizagem. Entretanto, isso requer modelos de domínio capazes de mapear os conteúdos ensinados a fim de que os objetos de aprendizagem resultantes sejam acrescidos de metadados que os identifiquem e categorizem, e requer modelos de aprendiz capazes de diagnosticar o aprendizado de forma individualizada, com razoável acuidade e em larga escala, o que implica entender a essência do processo de aprendizado.

Em geral, nas situações de ensino presencial, como as ilustradas nas Figuras 1.1a a 1.1e, o professor tem o papel de especialista no domínio, deve estar habilitado para utilizar diversas estratégias e métodos de ensino e, principalmente, deve ser capaz de avaliar o aprendizado de cada aluno individualmente para detectar quando um tópico deve ser reforçado ou contrastado em resposta a hesitações, dúvidas e erros do aluno. Isso significa que o tutor humano deve tanto perseguir um modelo pedagógico, quanto manter um modelo do domínio e vários modelos sobre seus alunos.



Figura 1.1: Vários possíveis contextos de ensino de línguas

Mas essa supervisão individualizada da aprendizagem de cada aluno se torna mais difícil para turmas maiores, como no caso mostrado na Figura 1.1b, pois o professor tem mais alunos para acompanhar. Além disso, novos desafios se somam a esse, com a inclusão de recursos audiovisuais (Figura 1.1c) ou de TICs (Figuras 1.1d a 1.1f), pois isso exige do professor destreza no manejo desses recursos e impõe às modelagens de domínio e de aprendiz considerações sobre a interação humano-computador – IHC. Mas tais dificuldades crescem sobretudo nos contextos de EaD, como o ilustrado na Figura 1.1f, em que, sem contato direto com os alunos e diante da possibilidade de que esses sejam muito numerosos e diversificados, cresce o desafio de modelar individualmente os aprendizados e de manter um modelo do domínio que permita a individualização da experiência pedagógica do aluno.

Assim, um aspecto relevante em ambientes de EaD mediados por computador é a forma como os tutores humanos acompanharão o processo de aprendizado. Se o número de alunos é pequeno o bastante, tem-se o Cenário 1, ilustrado na Figura 1.2, que a despeito do contexto EaD e do uso de TICs é relativamente simples por permitir uma supervisão mais individualizada por parte dos tutores humanos (não pressupondo o uso de sistemas tutores inteligentes - ITS). Mas se o curso é dirigido a um público-alvo amplo e diversificado, como em aplicações baseadas na Web, os desafios tornam-se maiores, em vista da dificuldade de se prover um apoio a cada aluno e de manter para ele um modelo atualizado de sua aprendizagem e de sua interação com as interfaces e seus conteúdos. Na prática o Cenário 2 implica o uso de sistemas tutores inteligentes, como ilustrado na Figura 1.2, como solução mais viável para que o ensino tenha ao menos parte do que se observa em tutorias humanas presenciais.

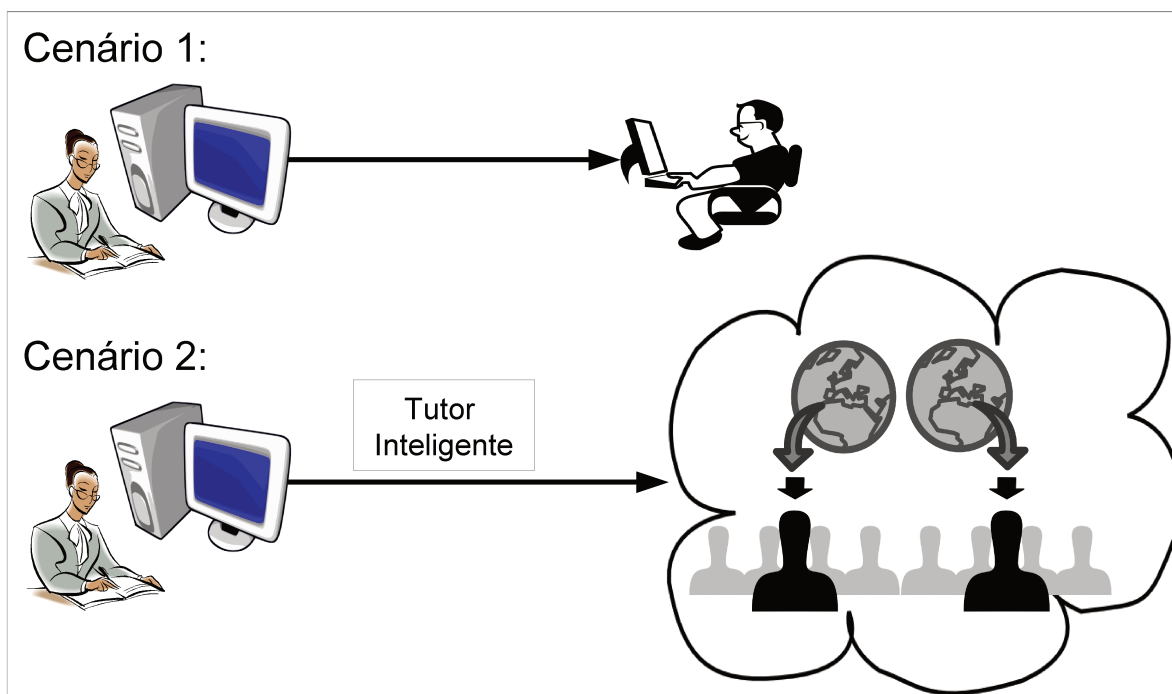


Figura 1.2: Cenários de ensino a distância de línguas mediado por TICs

É por isso que nas mais recentes arquiteturas tutoriais para o EaD o tutor inteligente busca refletir uma teoria pedagógica, manter uma representação do domínio a ser ensinado e, idealmente, ser também capaz de modelar o processo de aprendizado de cada aluno. No contexto do ensino de línguas isso significa que o tutor inteligente deve possuir um modelo de domínio que capture a riqueza e a sutileza da linguagem humana, além de um modelo do aprendiz que correlacione as ações e respostas deste a aspectos da L2 e de seu aprendizado, estando ambos os modelos alinhados a uma teoria pedagógica do aprendizado de uma L2.

Mas mesmo no Cenário 1 o desenvolvimento de aplicações de EaD de línguas mediado por computador deve considerar várias disciplinas e áreas de conhecimento, como visto na Figura 1.3. Os campos de interesse incluem a psicologia e pedagogia, em particular as teorias de cognição e de aprendizado de línguas, além das áreas de semiótica, linguística e iconologia, e naturalmente as competências tecnológicas da engenharia e da computação. Essas muitas disciplinas (cujos tamanhos dos círculos na Figura 1.3 não representam suas respectivas importâncias) devem dialogar para, por exemplo, produzir avanços nos campos da IHC e da linguística computacional, essenciais a aplicações neste contexto. Assim, o estudo dedica o Capítulo 2 a descrever a evolução e o estado da arte das implementações de ensino assistidas por computador para, em seguida, no Capítulo 3, discutir aspectos da aquisição da língua materna e, com base nisso, buscar teorias e modelos para o aprendizado de uma L2 que possam orientar a especificação do ambiente de ensino aqui proposto. Avançando na revisão das áreas pertinentes, no Capítulo 4 discute-se a visão da semiótica sobre os processos de percepção e interpretação de imagens e sua inter-relação com textos, tendo em vista o fato de que um curso de língua para um público-alvo diversificado deve se apoiar no uso de imagens para com isso fundamentar o aprendizado da língua-alvo.

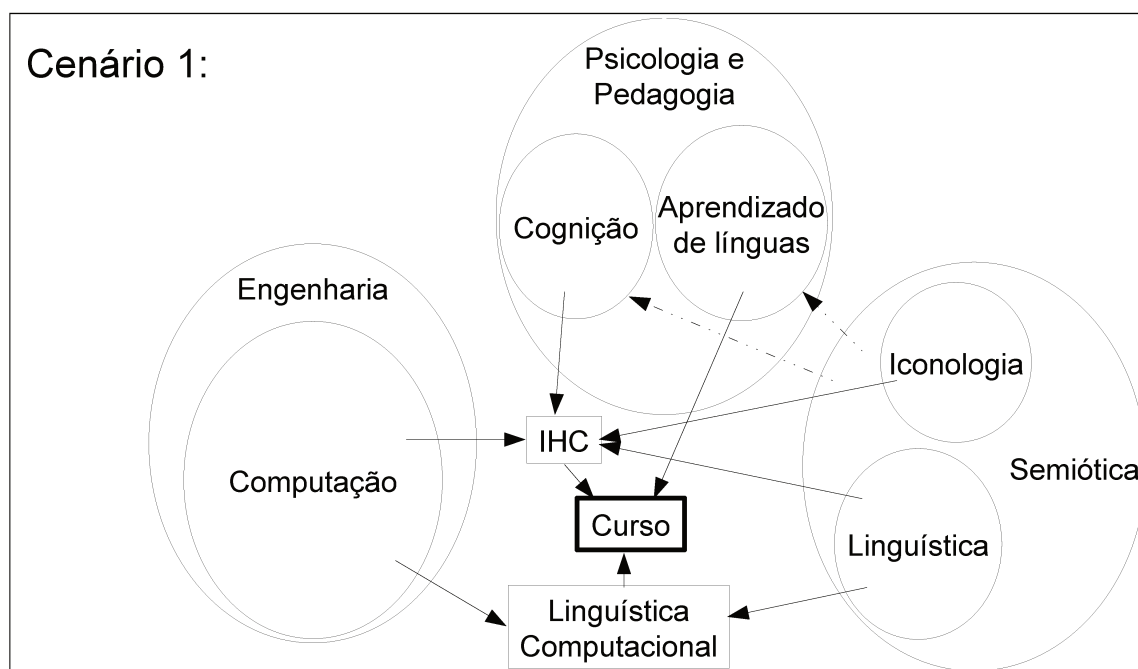


Figura 1.3: Disciplinas relevantes para o EaD de línguas mediado por computadores

Mas como um dos objetivos do ambiente de ensino aqui proposto é ser acessível a um grande número de usuários via Web, o mais pertinente ao presente estudo é o Cenário 2, que aponta na direção dos sistemas tutoriais inteligentes. E para que o sistema tutorial possa acompanhar em profundidade o aprendizado de uma língua, ele deve ser capaz de mapear os processos inferenciais envolvidos nesse aprendizado, e também de ter um modelo de domínio que reflita a complexidade da linguagem humana. Portanto, e como ilustrado na Figura 1.4, o estudo recorreu à área de sistemas inteligentes para avaliar as ferramentas e técnicas da Inteligência Artificial (IA) e identificar aquelas mais adequadas a este contexto.

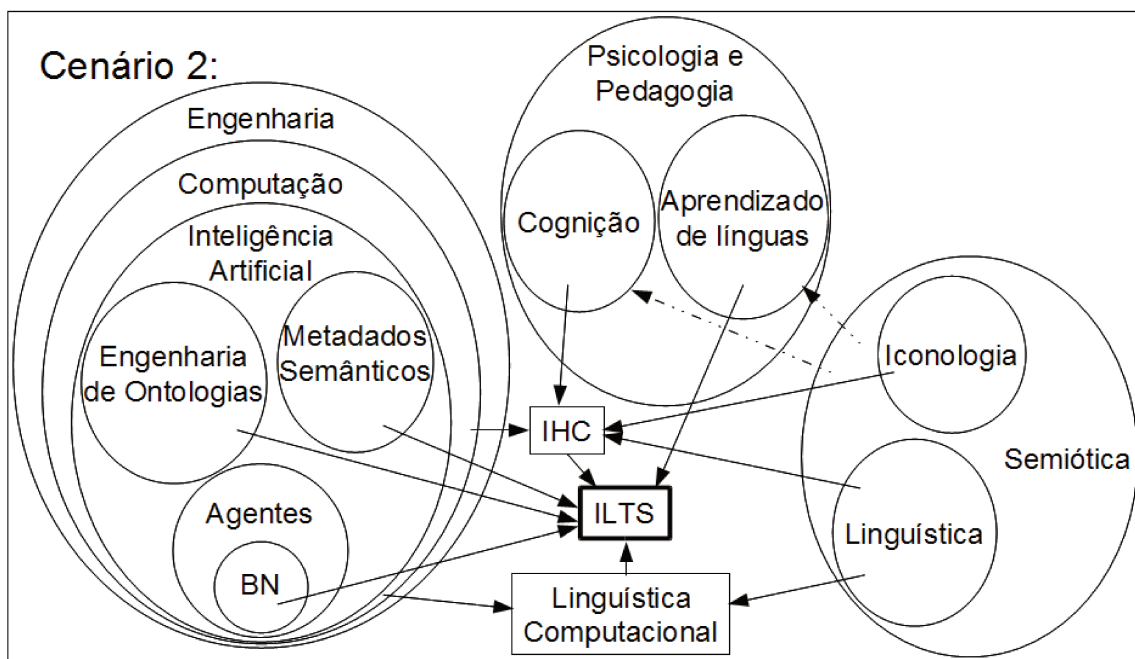


Figura 1.4: Disciplinas relevantes para o EaD de línguas via Web

É possível observar que a principal diferença entre os cenários 1 e 2 é que este último, por suas características específicas, exige mais soluções tecnológicas de apoio ao ensino. Isso significa que para modelar o aprendizado e o domínio de ensino o Cenário 2 demanda, além dos requisitos discutidos para o Cenário 1, algumas ferramentas adicionais da engenharia e da computação de forma a se obter uma visão completa da língua-alvo e do processo de aprendizado de cada aluno. Em vista disso, ferramentas e técnicas da IA são discutidas no Capítulo 5, em particular as áreas de ontologias do domínio, de modelagem baseada em multiagentes e de uso de metadados semânticos em imagens.

Todas essas áreas de conhecimento foram então consideradas na especificação de uma arquitetura para o ambiente de ensino. Mas, tendo em vista o grande número de áreas de estudo pertinentes ao ensino de línguas assistido por computadores, o estudo foi dividido em duas etapas:

1- Na primeira etapa, que corresponde ao trabalho de mestrado aqui descrito, o tema de ensino e aprendizado de línguas foi estudado e contextualizado para situações assistidas por computador, em particular aquelas baseadas em tutores inteligentes. Foi então proposta,

como contribuição principal da etapa uma nova estratégia de modelagem de domínio, a qual é realizada segundo uma abordagem baseada no uso de agentes, e que se soma à proposta de semantização de imagens e de uso de uma ontologia da L2. Foi então especificada a arquitetura que articula esses três elementos, e que é baseada no conceito de consciência artificial. Essa arquitetura, que também é uma das contribuições deste estudo, busca dar ao tutor inteligente autonomia e flexibilidade para responder e adaptar-se a necessidades específicas de cada aluno e ao, mesmo tempo, pautar-se por informações extraídas do modelo do domínio (que em última instância reflete a gramática da língua-alvo) ou por parâmetros didáticos definidos pelos tutores humanos. Por fim, para ilustrar a técnica proposta para o curso de línguas assistido por computador foram feitas algumas simulações ABMS iniciais do módulo de gestão de conteúdos didáticos. Esse recorte da contribuição da primeira etapa do trabalho está ilustrado na Figura 1.5.

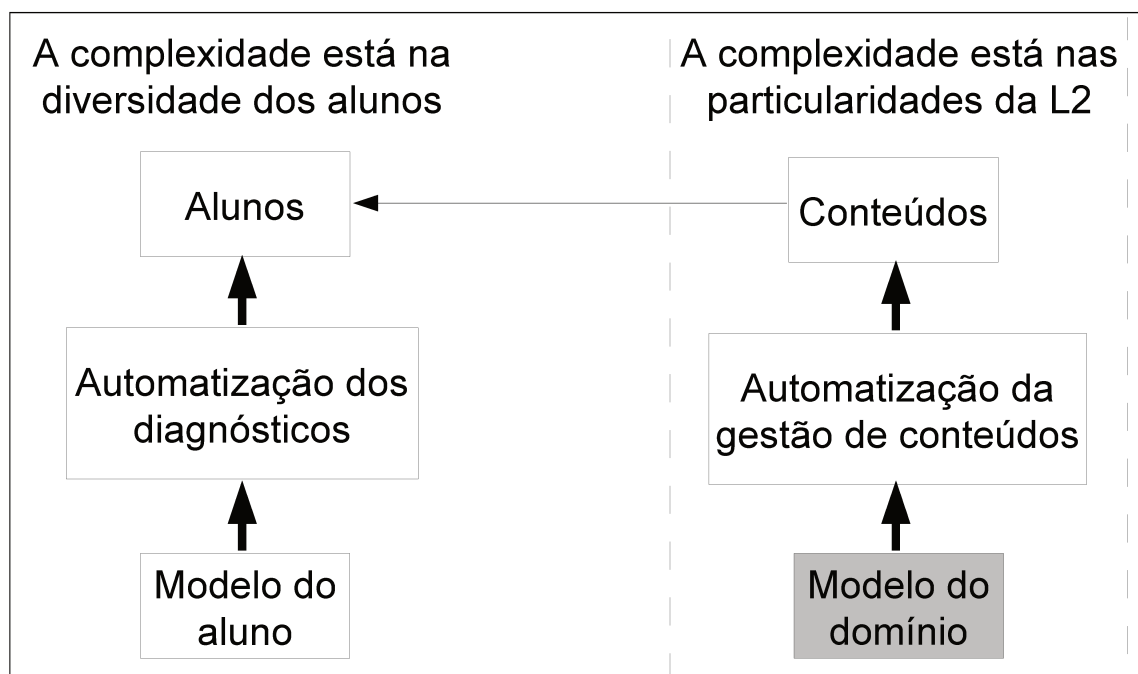


Figura 1.5: Foco de contribuição deste trabalho

2- Na segunda etapa, que corresponderá ao subsequente trabalho de doutorado, será feita a implementação completa da arquitetura proposta na primeira etapa, com a integração do ambiente de interação multiagentes a um módulo de ontologia e de inferência, e a um aplicativo de manipulação de metadados de imagens. Será também proposta uma solução de modelagem dos aprendizes que tenha a precisão e a escalabilidade desejadas e que, além disso, esteja alinhada com a modelagem do domínio aqui proposta e especificada.

Após implementada, a arquitetura será testada com usuários reais, em situações de EaD, permitindo avaliar suas vantagens e eventuais limitações e também possibilitando estimar o esforço de sua extensão para outros idiomas além do utilizado no estudo de caso adotado, o qual está descrito na Seção 1.1.

1.1 Estudo de caso para o ambiente de ensino

A primeira implementação da arquitetura enfoca o ensino do idioma esperanto¹, que é uma língua internacional que se propõe exatamente a facilitar os contatos e intercâmbios entre pessoas de origens culturais diferentes, e cujo valor nesse aspecto foi reconhecido pela Unesco em duas resoluções (1954 e 1986).

Mas as razões principais para essa escolha desse estudo de caso do ponto de vista de suas vantagens para a pesquisa foram:

- (i) Por ser essa uma língua que não dispõe de um base significativa de falantes nativos, o uso do ambiente torna-se pertinente a usuários dos mais diversos perfis culturais, sem excluir qualquer nacionalidade pelo fato de esta já ser formada por “falantes nativos”;
- (ii) Por ter léxico baseado em raízes internacionais, possibilita avaliar os efeitos da maior ou menor proximidade etimológica de cada unidade pedagógica em relação à L1 de cada aluno (e, como será discutido no Capítulo 4, seu efeito na interação com as imagens em que o curso se baseia);
- (iii) Por ser “econômica” em termos da amplitude de seu léxico básico, a língua-alvo escolhida torna possível um caso de estudo com um número relativamente pequeno de lexemas, algo útil para um protótipo, mas que não impede a avaliação da escalabilidade da solução para línguas com léxicos mais vastos;
- (iv) Por não ser esse aprendizado excessivamente motivado por fatores externos (como exigências profissionais), é mais fácil avaliar a agradabilidade geral do ambiente na retenção desses usuários “automotivados”;
- (v) Por não ter uma norma fundada na figura do “falante nativo ideal”, a língua-alvo possibilita o uso de conteúdos instrucionais (imagens e textos) provenientes das diversas regiões do mundo;
- (vi) Por ter uma gramática relativamente simples e regular, o aprendizado da língua-alvo baseia-se em grande medida na aquisição do léxico, e um dos objetivos do ILTS aqui proposto é facilitar essa aquisição.

Entre as possíveis desvantagens dessa escolha pode-se mencionar o fato de que, por conta de sua gramática simples e sem exceções, o aprendizado de esperanto baseia-se muito mais no aprendizado de seu léxico. Isso difere do aprendizado de línguas com estruturas mais irregulares, como o inglês, o espanhol e o francês. Contudo, muito embora o ensino dessas línguas deva tratar dessas irregularidades gramaticais, ele não prescinde de modo algum do ensino de seu vocabulário, e este é o principal objetivo da arquitetura proposta.

Em outras palavras, as especificidades do estudo de caso adotado não impedem que o ambiente seja avaliado na sua possível extensão ao ensino de outras línguas. A avaliação desse aspecto, como discutido acima, será feita na segunda etapa do estudo, e sob a ótica das exigências específicas de cada língua (por exemplo a dimensão do léxico básico) tendo em conta as dificuldades, oportunidades e ganhos de escala observados no desenvolvimento do estudo de caso escolhido para o ambiente piloto.

¹ Informações mais detalhadas sobre o esperanto podem ser encontradas no Apêndice 3.

1.2 Organização da dissertação

No Capítulo 1 procurou-se introduzir e sintetizar os objetivos gerais do estudo em termos de sua pertinência e oportunidade no âmbito da chamada “sociedade informacional”.

Para começar a situar a contribuição tecnológica e a motivação acadêmica deste trabalho foi necessário rever a evolução, as várias segmentações dessa área de estudo e identificar as principais tendências das pesquisas correlatas. Isso é feito no Capítulo 2, onde são discutidos aspectos relacionados ao ensino mediado por computador, suas implicações e limitações, em particular para o ensino de línguas.

Dando sequência ao estudo, no Capítulo 3 são discutidos tanto a aquisição da primeira língua (língua materna), quanto o aprendizado e o ensino de uma segunda língua, buscando, desse modo, dimensionar e qualificar a tarefa para a qual a arquitetura proposta de sistema tutor inteligente e, em particular, o seu módulo de gestão de conteúdos, se propõem a contribuir.

Tendo em vista a proposta de se desenvolver um ambiente computacional de ensino de línguas apoiado em imagens, no Capítulo 4 são estudados os processos de percepção e de interpretação de imagens, das interações entre imagens visuais e mentais, e também entre imagens e textos (em uma cena).

Feitas essas primeiras revisões de áreas pertinentes ao estudo, a arquitetura proposta para o ambiente de ensino, principal contribuição do trabalho, é detalhada no Capítulo 5 juntamente com a estratégia de desenvolvimento do mecanismo de gestão de vocabulário em um ambiente computacional de ensino de línguas com o uso de multiagentes, ontologias e metadados para as imagens. No capítulo 6 são feitas algumas simulações de cenários do módulo de gestão de conteúdos.

Esse conjunto de conhecimentos permite então resumir, no capítulo 7, as principais conclusões do estudo e identificar seus possíveis desdobramentos futuros.

Capítulo 2

Ensino mediado pelo computador

Para darmos início ao estudo, neste capítulo tratamos do uso do computador como ferramenta de ensino de línguas e dos principais conceitos dessa área de estudo, enfocando em seguida o uso de um Sistema Tutor Inteligente (ITS), que é o foco do trabalho.

2.1 Aprendizado de línguas assistido por computador – CALL

Datam da década de 60 as primeiras propostas de aprendizado de línguas assistido por computador (*Computer-Assisted Language Learning* – CALL), sendo esse um ramo do ensino assistido por computador. Até os anos 70, os sistemas de CALL estavam fortemente ligados à instrução programada, área essa conhecida pela sigla CALI (*Computer-Assisted Language Instruction*). Foi somente a partir dos anos 80 que a sigla CALL tornou-se predominante. Embora no início da década de 90 uma sigla alternativa, TELL (*Technology-Enhanced Language Learning*), tenha sido proposta por representar melhor as atividades normalmente associadas a essa área de estudo, o uso dessa nova denominação não se tornou difundido, e a sigla CALL ainda hoje é a mais empregada.

A evolução da área de CALL na busca por meios de melhorar o ensino e o aprendizado de línguas pelo uso de meios tecnológicos foi impulsionada pela evolução dos hardwares, pelos avanços da tecnologia, pelo desenvolvimento de áreas da IA, pelo advento da Internet e da Web e, mais recentemente, pelo interesse de alguns governos pela área. Cameron (1999: 2) salienta que tal progresso não ocorreu sem críticas, seja em relação à eficácia, seja em relação aos efeitos do CALL na demanda por professores humanos. Mas ele discorda da visão de que o ensino assistido por computador seja uma panaceia ou um substituto dos professores, podendo entretanto complementar o ensino, facilitando o ensino a distância, e, no contexto da Internet, dando aos usuários acesso a uma enorme quantidade de materiais culturais que podem estimular o aprendizado.

Segundo Levy (1997), o surgimento dos sistemas CALL aconteceu em um período (início dos anos 60) em que prevalecia na linguística a teoria behaviorista de B. F. Skinner, proposta em sua obra *Verbal Behaviour*, de 1957, e cujos pontos centrais eram as ideias de estímulo, resposta e reforço. O principal reflexo do behaviorismo no ensino de línguas foi o método audiolingual, que propunha um aprendizado baseado na formação de hábitos, isto é, na prática da língua-alvo através da imitação e repetição dos modelos apresentados pelo professor. Os modelos deveriam ser padrões-estruturais-chave extraídos de diálogos. Além disso, respostas corretas eram reforçadas positivamente. Como comenta Levy, o próprio Skinner defendia o uso de “máquinas de ensino” para uma instrução individualizada, capaz de respeitar o ritmo dos aprendizes, cujas respostas seriam seguidas de retornos imediatos.

As primeiras implementações de CALL adotavam sequenciamento gramatical, numa clara influência do conceito de instrução programada. Os desenvolvedores de tais sistemas se deram conta de que a estratégia baseada em exercícios de prática e repetição podiam ser prontamente programados em um computador, em razão de seu caráter “sistemático e de rotina” (Kenning e Kenning, 1990: 53, *apud* Levy, 1997).

Mas novas correntes de pensamento em diversas disciplinas relevantes para a área de CALL, incluindo a linguística e a psicologia, favoreceram o aparecimento de propostas novas nas décadas seguintes. Até mesmo a revolução da computação, com o advento da Internet e dos recursos multimídia, propiciou novas possibilidades de uso dos computadores no ensino, como as aplicações MMCALL (multimedia CALL), com a adoção de CD-ROMs e DVDs e o surgimento de implementações CALL baseadas na Web, que incluem o uso de blogs, wikis, podcasts e outras ferramentas de redes sociais, e até salas de aula virtuais.

Levy (1997) salienta ainda que a fatura de acrônimos para o uso de computadores na educação se deveu aos diferentes focos que cada autor deu a sua própria implementação, ou até mesmo à origem geográfica dos pesquisadores, com uma certa preferência nos EUA pelo foco instrucional e no Reino Unido pelo de aprendizado. No caso particular dos termos ICAL (Intelligent Computer-Assisted Instruction) e ITS (Intelligent Tutoring Systems) a nomenclatura pode ser explicada pela mudança do foco, do papel do computador para o do programa, visto que atualmente essas áreas fazem usos variados de técnicas da IA. Esses usos podem incluir a modelagem de aprendiz, isto é, a construção de um modelo daquilo que o aprendiz sabe em um dado momento, para que, com base nessa informação, uma instrução adequada lhe seja oferecida.

Um outro aspecto da área de CALL é seu escopo amplo e interdisciplinar. Por isso, seu campo de atuação permeia muitas fronteiras com outras disciplinas e linhas de pesquisa. Assim, os pesquisadores e desenvolvedores que atuam nesse ramo do conhecimento devem estar cientes de que avanços nessas áreas afins podem ter consequências importantes nas aplicações de CALL. É portanto necessário acompanhar essas outras teorias, disciplinas ou linhas de estudo, muito embora essa não seja uma tarefa fácil, dado o grande número de áreas e subáreas afins, como ilustrado na Tabela 2.1.

Tabela 2.1: Disciplinas e áreas relevantes para o estudo da CALL (Levy, 1997)

Categorias centrais de pesquisa	Disciplinas, teorias e áreas abarcadas pela categoria
Psicologia	Instrução programada, aquisição de L2, psicologia cognitiva (incluindo processamento de informação)
Inteligência artificial	Sistemas especialistas
Linguística computacional	Tradução por máquina, processamento de linguagem natural, análise sintática e processamento de dados linguísticos
Tecnologia e <i>design</i> instrucional	Tecnologia instrucional e educacional, <i>design</i> instrucional e abordagem sistêmica
Interação humano-computador	—

Embora muitos pesquisadores compartilhem a crença de que toda implementação CALL é intrinsecamente tutorial, caracterizando-se por interações um-para-um em que o computador avalia as respostas do aprendiz e, em função delas, adapta a apresentação de novos conteúdos, Levy (1997) comenta que essa suposição desconsidera temas importantes, como avaliação, metodologia e os papéis do professor e do aluno, o que, segundo ele, pode-se corrigir pelo entendimento dos papéis do tutor e da ferramenta num *framework* próprio.

2.2 Adaptatividade e adaptabilidade em ambientes educacionais

Ambientes adaptativos de aprendizado (ALE) englobam, segundo Magoulas et al. (2001), duas diferentes tecnologias de CAI: a de Sistemas Tutores Inteligentes (ITS) e a de Hiperídia Educacional (EH) (Brusilovsky, 1996). Isso resulta da combinação de duas abordagens opostas de sistemas de aprendizado assistido por computador: os sistemas baseados na IA tradicional, de estilo mais diretivo e centrados em tutores, e os sistemas EH, com sua abordagem de navegação mais flexível e centrada no aprendiz (Davidson, 1999).

Embora os chamados sistemas abertos de ensino a distância (ODL) venham tendo um papel crescente no treinamento efetivo de pessoas, para muitos pesquisadores da área (Brusilovsky, 1996), o principal problema dos sistemas abertos é que eles oferecem uma educação uniforme (e monótona) para todos. E com o aumento de seu alcance somado ao advento do princípio da educação continuada, surge a necessidade de lidar com diferenças culturais, linguísticas e de gênero. Em razão disso, a importância da análise dos perfis e atitudes dos aprendizes cresce (Cristea e De Bra, 2002). Assim, o novo paradigma da interação humano-computador inteligente vai na direção da individualização do processo de aprendizado.

A individualização das experiências pedagógicas baseia-se nos conceitos de adaptabilidade e adaptatividade. O primeiro se refere à possibilidade de o usuário/aprendiz fazer alterações no aspecto, nos conteúdos ou em funcionalidades do material instrucional. O segundo pressupõe a real capacidade de o sistema adaptar-se automaticamente a novas condições, em geral deduzidas do modelo de aprendiz, o que representa um passo mais avançado em direção à Inteligência Artificial. Embora tanto a adaptabilidade quanto a adaptatividade possam prescindir dos modelos de aprendiz, quanto mais personalizado um ambiente pretende ser, mais fundamentais se tornam esses modelos.

Segundo Sison e Shimura (1998), a modelagem do aprendiz envolve a interpretação do comportamento dos usuários do ambiente educacional como sendo uma representação qualitativa das suas intenções. A construção do modelo do aprendiz pode ainda levar em conta fatores ambientais (horário, local, ambiente computacional, rede, deficiências físicas, etc.) e o comportamento do usuário em sua navegação pelo ambiente de ensino.

Segundo Magoulas et al. (2001), a avaliação do conhecimento de um aprendiz baseia-se em dois tipos de informação: respostas a questões que testam a parte cognitiva do conhecimento (Nkambou, 1999), e medições que avaliam a parte comportamental, que trata de fatores como atenção, interesse, responsabilidade, concentração (Embenon, 1990). Nos dois casos, muitos fatores contribuem para a incerteza no procedimento de avaliação, entre eles os erros por displicência ou acertos por “chute” nas respostas que o aprendiz fornece, mudança em seus conhecimentos em razão de aprendizado ou esquecimento, e padrões de resposta não previstas pelo projetista do modelo de aprendiz.

Fazer a modelagem de um aprendiz é algo que envolve a inferência de informações não observáveis a partir de informações observáveis sobre ele, como, por exemplo, suas ações e elocuições. Para desempenhar essa tarefa, um sistema de modelagem deve lidar com a incerteza quando tenta inferir sobre o aprendiz com base em informações incompletas. A área de reconhecimento de planos tem se ocupado de inferências acerca de preferências, metas, ações e situações futuras dos usuários (Zukerman e Albrecht, 2001).

Pesquisas sobre o aprendizado demonstraram que a tutoria individualizada é mais efetiva do que outras formas de instrução (Core et al. 2000), esse resultado levou alguns pesquisadores (Merrill et al., 1992; Fox, 1993; Graesser et al., 1995) a afirmarem que o que promove o aprendizado é o diálogo colaborativo entre aprendiz e tutor, por meio do qual o tutor pode intervir corrigindo erros e evitando impasses por parte do aluno. Os tutores humanos experientes são capazes de manter um delicado equilíbrio entre autonomia dada aos aprendizes para trabalhar e sentirem-se no controle e a necessária condução do processo para evitar que eles se confundam ou se frustrem.

Segundo (Vasilakos et al., 2004), a Educação Baseada na Web (EBW) tornou-se um ramo muito importante das tecnologias educacionais, na medida em que ela oferece acesso a fontes de informação e conhecimento que são praticamente ilimitadas, permitindo um número de oportunidades para aprendizado personalizado, a distância e colaborativo, com as claras vantagens de independência em relação à sala de aula e a plataformas. Contudo, com relação aos esforços despendidos nos últimos anos por inúmeros grupos de pesquisa no desenvolvimento de metodologias e abordagens, ferramentas e sistemas educacionais baseados na Web, teria havido comparativamente pouca pesquisa voltada à aplicação dos princípios e resultados da área de Inteligência Computacional (IC): sistemas nebulosos, redes neurais, computação granular e evolutiva. As tecnologias da IC, quando aplicadas ao desenvolvimento de sistemas EBW, trazem importantes melhorias que tornam os sistemas mais flexíveis, amigáveis e intuitivos. Para isso, a IC tem de ser capaz de lidar com vários aspectos desafiadores, tais como representação do conhecimento, propriedades adaptativas, habilidades de aprendizagem e desenvolvimentos estruturais.

Com relação às propriedades dos sistemas inteligentes suportados pela IC, (Vasilakos et al., 2004) destacam duas perspectivas gerais para distingui-las: intrínsecas ou extrínsecas ao sistema inteligente. As primeiras, embora cruciais ao desenvolvimento dos sistemas, não se manifestam externamente. Já as segundas são dominantes e tornam-se de extrema relevância quando se trata da comunicação dos SIs e da facilitação de uma efetiva interação com os usuários humanos, a fim de dar a sensação de capacidades inteligentes e amigáveis. Isso exige as capacidades de lidar com informações heterogêneas, de estabelecer comunicação efetiva e personalizada com o usuário-final, o que pressupõe generalidade (abstração) e granularidade de informação. Essas são condições para se oferecer interfaces centradas no usuário, essenciais a muitos sistemas interativos, incluindo ITS. Vasilakos et al. (2004) comentam ainda que a Educação Inteligente Baseada na Web (EIBW) resulta da aplicação das tecnologias inteligentes à EBW. Nesse particular, uma das tecnologias-chave é a de agentes pedagógicos inteligentes, responsáveis por prover a infraestrutura para o fluxo de informação e conhecimento entre clientes e servidores na EBW.

Todo sistema EIBW usa várias técnicas de representação de conhecimento derivadas da IA. O conteúdo instrucional (definindo o que ensinar) é chamado de módulo especialista, enquanto que as diferentes estratégias de ensino (especificando como ensinar) são ditadas pelo módulo pedagógico. O conhecimento que o sistema tem da proficiência do aprendiz nos tópicos ensinados é representado pelo modelo do aprendiz, e o processamento dessa informação é a capacidade de ter alguma forma de raciocínio baseado em conhecimento no sequenciamento do currículo, na análise das respostas do aprendiz e na ajuda dada a ele.

2.3 A IA na educação: Sistemas Tutores Inteligentes – ITS

Datam da década de 70 as primeiras iniciativas de aliar métodos e técnicas da IA (Inteligência Artificial) à concepção de ambientes de ensino e aprendizagem assistidos por computador. A área foi de início chamada Instrução Inteligente Auxiliada por Computador (Intelligent Computer-Assisted Instruction – ICAI), e posteriormente denominada Sistemas Tutores Inteligentes (STI ou ITS) ou de Inteligência Artificial em Educação (IAED).

A inteligência computacional na educação

Dada a natureza muito dinâmica da aprendizagem, mecanismos de aprendizado são necessários para capturar as características desse processo. Entre as ferramentas e técnicas da Inteligência Computacional (IC), algumas como os sistemas nebulosos, as redes neurais, os sistemas multiagentes e a computação evolutiva são aplicáveis a ambientes educacionais. As redes neurais (RNs), por exemplo, podem ser um veículo apropriado a um formalismo de representação de conhecimentos para a CALL. Uma RN com nós que correspondam a conceitos de um domínio e conexões cujos pesos reflitam as relações entre os conceitos pode ser usada para modelar o componente de domínio do conhecimento. Uma RN pode também ser usada para criar um modelo do aprendiz ou um modelo pedagógico. Por sua vez, os sistemas multiagentes podem ser uma alternativa interessante para a modelagem do domínio, em particular, como será visto neste trabalho, se esse domínio é uma língua, tendo em vista a natureza eminentemente granular do léxico e a complexa interação dos lexemas.

Sistemas Tutores Inteligentes (STI ou ITS) são sistemas computacionais que buscam reproduzir o comportamento de um professor humano de um dado domínio. Isso pressupõe a capacidade de prover instrução individualizada e dinamicamente adaptável em forma e conteúdo. Um ITS deve possuir conhecimentos especializados para tratar as três questões clássicas: *O que ensinar?*, *A quem ensinar?* e *Como ensinar?*. A primeira diz respeito à modelagem do conhecimento de um dado domínio, a segunda trata do estado cognitivo do aprendiz e a última concerne à estratégia pedagógica. Para isso, ele deve recorrer a várias áreas de estudo, entre elas a Educação (Pedagogia), a Psicologia Cognitiva e a Inteligência Artificial, como ilustrado na Figura 2.1.a.

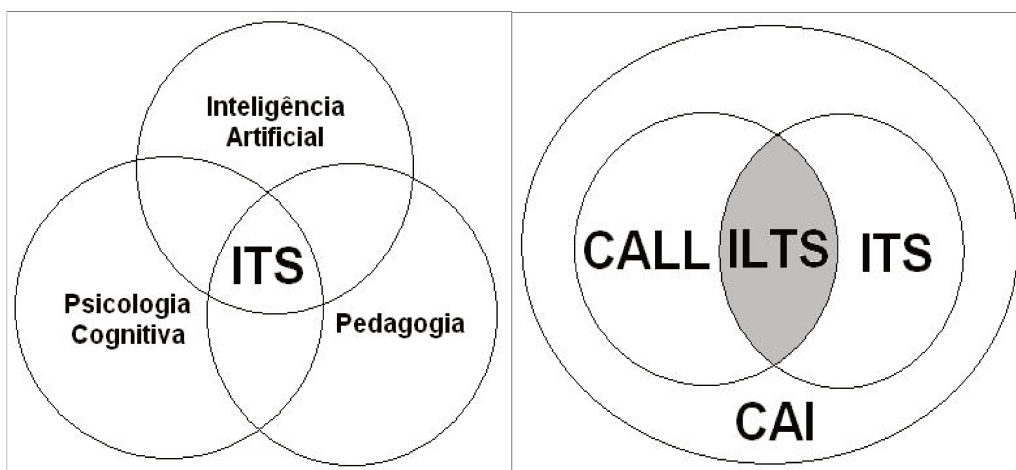


Figura 2.1: a) Áreas de conhecimento em um ITS; b) Foco do estudo

Módulos principais de um ITS

Um ITS é um sistema especialista desenvolvido com vistas a articular um grande volume de informações distintas mas inter-relacionadas. Para lograr esse objetivo, os ITSs em geral contam com módulos especializados. São eles:

- **Modelo do domínio:** Trata da questão “*O que ensinar?*”, e representa o conhecimento especializado do domínio de aplicação. Fundamenta as funções pedagógicas.
- **Modelo do aprendiz:** Trata da questão “*A quem ensinar?*”, e é a representação abstrata que o sistema constrói sobre o conhecimento e o desempenho do aprendiz na interação dinâmica com o sistema. Visa a determinar o estado cognitivo do aluno: o que ele sabe ou não sabe, o que entendeu mal. Esse conhecimento fundamentará as escolhas de estratégias de ensino adequadas a um dado aprendiz em uma dada situação.
- **Modelo pedagógico:** Trata da questão “*Como ensinar?*”, e relaciona-se às estratégias de ensino-aprendizagem adotadas, incluindo seleção e planejamento de atividades propostas e que atividade pedagógica apresentar, com base nos modelos do aprendiz e do domínio.

Assim, nessa visão a psicologia cognitiva traz um entendimento sobre indivíduos (alunos), a IA tem a função de representar e manipular os conhecimentos envolvidos, e a pedagogia oferece os métodos de ensino. No caso particular de um ITS para ensino de línguas, os métodos de ensino disponíveis são vários e cada um deles se fundamenta em uma linha específica da linguística, como tratado no Apêndice 1.

É importante ressaltar também que assim como nem toda implementação CALL é um ITS, nem todo ITS é voltado para o aprendizado de línguas, não havendo portanto uma absoluta superposição entre as duas áreas. Um ITS em um ambiente de ensino de línguas é uma interseção das áreas de CALL e ITS, como ilustrado na Figura 2.1.b. Essa interseção de áreas tem sido denominada por alguns autores como ILTS (Intelligent Language Tutoring Systems). Embora bastante apropriado, esse acrônimo ainda não se estabeleceu amplamente na comunidade. Questões de nomenclatura à parte, essa área de superposição representadas pela sigla ILTS é então o foco de interesse da implementação aqui descrita.

A partir dos anos 90 as pesquisas em ITS vêm buscando um estilo de interação mais flexível, tratando o aluno como elemento ativo, dotado de certa autonomia no aprendizado. Isso tem levado a novas abordagens na concepção de ambientes computacionais de ensino, resultando em ambientes de aprendizado interativos (ILEs) ou ITS cooperativos.

Um ILE é um tipo de software educacional que coloca o aprendiz em uma situação de descoberta, em geral por meio de atividades de resolução de problemas. Durante esse processo o tutor artificial assiste o aprendiz em sua tarefa e monitora seu aprendizado. Assim, enquanto um ITS tradicional é baseado em um estilo rígido de interação, em que o sistema está sempre no controle, um ILE é flexível a ponto de permitir que a iniciativa da interação seja compartilhada entre sistema e aprendiz.

No caso de um ILTS o tutor deve ter seu comportamento embasado por uma teoria de aprendizado linguístico, e por essa razão no Capítulo 3 será estudado o processo segundo o qual adquirimos a língua materna e serão discutidas algumas teorias que tentam explicar esse processo em termos que possam ser refletidos no ambiente de ensino do ILTS.

2.4 Abordagens dos ambientes computacionais de aprendizagem

Grande parte da pesquisa em Sistemas IAED está relacionada à busca de soluções cada vez mais efetivas para o problema da adaptação, e, como comenta Costa (1997), tem três principais abordagens: canônicas, mentalistas e especializadas, cujas características são:

Abordagens canônicas

As abordagens ditas canônicas caracterizam-se por organizar a arquitetura do ITS de forma que cada agente seja encarregado de dar suporte a uma das especialidades requeridas. Assim, por exemplo, um agente trata do domínio da aplicação, um segundo captura o estado cognitivo do aprendiz, um terceiro trata do planejamento pedagógico e supervisiona os demais, etc. O conceito de agente é estendido, em alguns casos, a fim de incluir agentes humanos (alunos) e agentes comunicadores (interfaces), além dos outros descritos acima.

Abordagem mentalista

Abordagem de IA-ED que busca conceber uma arquitetura que comporte a noção de estado mental. Assim, em uma interação ensino/aprendizagem visando à solução de um problema, dois agentes cognitivos autônomos são capazes de, em um diálogo, simular mudanças de seus estados mentais (crença, desejo, intenção e expectativa). Nesse diálogo os dois agentes podem alcançar uma das seguintes situações:

- Concordam a respeito do conhecimento envolvido mas discutem o método;
- Têm conhecimentos diferentes sobre o problema e devem entrar em acordo sobre ele;
- Um agente quer conversar sobre um problema e o outro não quer ou não pode cooperar.

Abordagens especializadas

São abordagens que enfatizam a modelagem de um dos aspectos: domínio, aprendiz ou pedagogia. No caso dos domínios, é introduzida a noção de pontos de vista (múltiplas representações de um certo objeto de conhecimento) e é estudada a capacidade de interação cooperativa entre sistema tutor e aprendiz, por exemplo. No caso do aprendiz, estuda-se a possibilidade de este participar da elaboração do modelo, sendo capaz de inspecioná-lo para não somente saber o que o modelo “sabe” sobre ele, mas até mesmo modificar o modelo, retificando eventuais erros de diagnóstico (de seu ponto de vista). Essa estratégia tem a vantagem de dar ao aprendiz uma noção sobre seus conhecimentos e ao sistema uma visão sobre a qualidade de seu diagnóstico. Há ainda algumas vertentes na modelagem do aprendiz que fazem uso de diagnósticos distribuídos baseados em multi-agentes, tentando com isso evitar o problema dos falsos entendimentos (*misconceptions*).

Em vista da natureza bastante particular do aprendizado e do ensino de línguas, uma abordagem especializada parece preferível se isso significar que nas interações do aluno com o ambiente de ensino os conteúdos propostos pelo sistema respeitarão a lógica da L2 (especialização no domínio). Isso evitará a proposição de sentenças e textos estereotipados, sem correspondência com o uso real da L2, e, como será discutido no Capítulo 3, tratará a língua como um “sistema emergente e evolutivo de atos públicos coordenados, preexistente e externo a todo agente individual” (Fettes, 1999).

Mas restringir a especialização do ILTS apenas à lógica da L2 equivaleria a tratar os aprendizes de uma língua como *tabulas rasas*, e, como será visto no Capítulo 3, eles são na verdade “sistemas complexos altamente adaptados” cuja negociação com o processo de aprendizado será fortemente influenciada por seus **estilos individuais**. Além disso, eles nunca realmente “adquirirão a língua” no sentido de se tornarem aptos a reproduzir por completo o sistema (agrupamentos e subsistemas de dispositivos públicos de linguagem) na sua complexidade dinâmica, mas aprenderão a reproduzir **elementos** desse sistema.

Assim, na abordagem aqui proposta esse compromisso entre uma especialização do ILTS em torno do domínio (língua-alvo) e a natureza individualizada e contextualizada do aprendizado linguístico foi atingido por meio da estratégia de atomização dos conteúdos instrucionais na forma de lexemas dotados de autonomia para interagir com outros lexemas e, de maneira emergente, para se ordenar e organizar de maneira totalmente compatível com os dispositivos públicos da L2. E sua assimilação pelo aprendiz será pautada pela interação deste com as cenas do curso, por sua exploração dos elementos nessas cenas em **busca do significado** de sentenças e textos que essas ensajem. Essa nova abordagem será descrita em mais detalhe nos próximos capítulos.

2.5 Taxonomia dos ambientes computacionais de aprendizagem

Os Ambientes Computacionais de Aprendizagem (ACAs) são aqueles que oferecem recursos para que os usuários possam adquirir ou construir conhecimentos por meio do computador. Eles podem ser classificados conforme quatro critérios (Gonçalves, 2003):

Arquitetura de ensino

A arquitetura de ensino trata da maneira como se dá o ensino. As principais são:

- **Aprender fazendo:** partindo do pressuposto de que para aprender a realizar uma tarefa é necessário executá-la, essa estratégia se concretiza na forma de simulações que criam contextos favoráveis ao aprendizado de um domínio de conhecimento, induzindo os aprendizes a aplicar hipóteses acerca de um determinado conhecimento e, dada o retorno recebido, a constatar, refletir e modificar seu pensamento e conhecimento na ação.
- **Aprendizado incidental:** método que, por meio de tarefas interessantes, facilita o aprendizado de informações enfadonhas, visto que nem todo aprendizado é divertido.
- **Aprendizado por reflexão:** o ambiente permite ao aluno criar representações detalhadas de seu pensamento, em um processo reflexivo. A função do tutor é ajudá-lo a perceber deficiências e erros em suas ideias e hipóteses e sugerir procedimentos e caminhos.
- **Ensino baseado em casos:** inclui o aprendizado através dos erros e visa a indicar aos aprendizes exatamente o que eles precisam saber, no momento adequado, tomando por base experiências similares anteriores, inclusive as vivenciadas pelos alunos. Alguns ambientes permitem que os alunos criem e avaliem hipóteses examinando perguntas e pratiquem o raciocínio sobre os casos.
- **Aprendizado por exploração:** durante o processo de aprendizado os aprendizes podem propor questões a serem respondidas pelo sistema, em linguagem natural. Assim, o aprendiz influencia no fluxo das questões e o sistema influi no das respostas.

Processo de aprendizagem

Esta classificação trata do processo de aprendizagem e da lógica de concatenação dos estímulos que possibilitam a aprendizagem. Existem dois enfoques principais:

- **Algorítmico:** visa à transmissão de conhecimento do professor/especialista para o aprendiz. O programa decide (de modo inteligente ou não) as tarefas para o aprendiz, e em razão disso o percurso do aprendiz deve ser modelado *a priori* pelo desenvolvedor do software. É adequado a conteúdos sistematizados, ao treinamento de habilidades específicas e a critérios de avaliação quantitativos, como em software tutoriais e CAI.
- **Heurístico:** valoriza o aprendizado experimental, propondo cenários que favoreçam a criação de hipóteses e de conflitos cognitivos, sem atividades programadas *a priori*. Pertencem a esta abordagem os ambientes de simulação, jogos e sistemas especialistas, que ensejem abstrações, estratégias, raciocínio e processos reflexivos.

Comando

Esta classificação se refere à precedência de comando sobre as interações entre computador e usuário, no âmbito do processo de ensino/aprendizagem. Em alguns casos ou situações esse comando pode partir do aluno, do sistema ou de ambos, conforme classificação abaixo:

- **Aprendiz no comando:** estabelece-se uma relação unilateral em que ao computador somente compete responder aos comandos do aprendiz.
- **Cooperativismo:** pressupõe a divisão do trabalho entre vários aprendizes, mediante tarefas desenvolvidas de forma independente por cada um deles.
- **Colaborativismo:** os alunos trabalham juntos para solucionar problemas e atingir metas comuns, compartilhar resultados coordenadamente, influenciando uns nos outros.
- **O tutor no comando:** nesta abordagem um sistema tutor, que pode ser implementado por meio de ferramentas da Inteligência Artificial, atua na orientação dos aprendizes.

Conteúdo do aprendizado

Esta classificação diz respeito à coerência entre o conteúdo a ser aprendido e as tarefas propostas. Schar e Krueger (2000) sugerem sete diferentes abordagens de ACAs:

- **Baseados em simulação:** ambientes nos quais podem ser reproduzidas condições do tipo ação/reação relacionadas a eventos do mundo real. Neles os aprendizes podem criar e testar hipóteses que levem à ampliação de seus conhecimentos, inclusive por tentativa e erro. Está relacionado com o paradigma do “aprender fazendo”, descrito acima.
- **Sistemas tutores:** ambientes nos quais o aprendizado ocorre de forma modularizada e objetiva, com níveis, passos e etapas bem definidos, e testes que habilitam o acesso aos módulos subsequentes. O sistema tutor oferece também correções e sugestões.
- **Hipertexto e hipermídia:** nestes ambientes os conteúdos, textos, desenhos, vídeos ou áudios são organizados em nós e encadeados como hiperdocumentos. O aprendizado se dá no hiperespaço e o aluno determina a estratégia de acesso conforme suas necessidades individuais. Nesse sentido, esta categoria se assemelha às simulações.
- **Prática por repetição:** ambientes em que o computador propõe problemas específicos, nos quais os aprendizes trabalham até atingirem um dado patamar de proficiência. O aprendizado se dá por perguntas e respostas, com estímulos de recompensa por acertos.

- **Sistemas informativos:** ambientes cujo objetivo é propiciar aprendizado informativo, por meio de consultas a bases de dados, tais como em dicionários ou atlas digitais.
- **Aprendizagem a distância:** ambientes em que o aprendizado pode se dar remotamente, pelo acesso a recursos como apresentações multimídia, e-mail, videoconferência, etc.
- **Sistemas híbridos:** combinam características descritas acima, de forma complementar.

Levando-se em conta as possíveis abordagens e as quatro classificações principais e suas subcategorias, pode-se situar a abordagem do ambiente aqui proposto como sendo especializada, como discutido na Seção 2.4, embora também guarde algumas características canônicas, como a atribuição de agentes específicos para as especialidades requeridas.

A arquitetura do ambiente situa-se entre aquelas em que o aluno aprende fazendo, isto é, utilizando a L2, e, futuramente, entre as que propiciam aprendizados incidentais, com conteúdos apresentados na forma de jogos, enigmas, etc., e, em certa medida, com recursos para livre exploração. Essa combinação busca reproduzir em parte a natureza gradual da aquisição da L1 e criar estados afetivos que favoreçam o engajamento dos alunos.

No que trata do processo, propõe-se aqui a uma implementação que embora mescle elementos das estratégias algorítmica e heurística, é caracterizado pelo modelagem multi-agente dos conteúdos (vocabulário) e pela formação espontânea e emergente de situações e exemplos, podendo assim ser classificada em uma nova categoria como **emergente**. Isso será tratado mais extensamente ao longo do trabalho e discutido em mais profundidade nas conclusões por ser uma de suas contribuições principais. A natureza particular do processo de aprendizagem conduzido pelas interações multiagentes merece então uma terceira categoria na taxonomia, como mostrado na terceira coluna da Tabela 2.2.

O comando será de certo modo compartilhado entre aluno, que pode iniciar alguns modos de interação com o ambiente, e o ILTS, que constantemente organizará os conteúdos segundo as relações de precedência e dependência próprias do aprendizado de uma língua.

Por fim, o ambiente proposto é voltado a um aprendizado a distância, assistido por um tutor inteligente e, em menor grau, também baseado em simulações (cenários e situações de uso da língua) e em práticas de repetição das estruturas, tendo portanto um conteúdo híbrido, mas com ênfase no contexto de tutoria a distância.

O ambiente apresentará assim a combinação ilustrada na Tabela 2.2 (as células com fundo cinza indicam características presentes, tendo as predominantes tonalidade escura):

Tabela 2.2: Identificação das características do ambiente de ensino proposto

Tabela 2.2.1: Taxonomização das características de diferentes de ensino propostos				
Abordagem	Classificação			
	Arquitetura	Processo	Comando	Conteúdo
Mentalista	Casos	Heurístico	Cooperativo	Hipertexto-hipermídia
	Reflexão		Colaborativo	Sistema informativo
Canônica	Exploração	Algoritmo	Aprendiz	Prática por repetição
	Incidental			Simulação
Especializada	Fazendo	Emergente	Tutor	Aprendizado a distância

2.6 Ontologias em sistemas educacionais e no ensino de uma L2

Um dos fundamentos mais importantes de um projeto de software é o conhecimento sobre o domínio de aplicação. Esse conhecimento é em geral obtido dos especialistas da área e deve ser transformado em algum tipo de modelo do domínio. Segundo (Knublauch, 2003), nos casos mais simples de projetos pequenos, esses modelos podem ser rabiscados em uma folha de papel, mas o ideal é ter modelos que possam ser traduzidos diretamente para um programa em Java. Para esse fim pode-se usar, por exemplo, a Unified Modeling Language (UML), pois esta permite rascunhar os modelos de domínio com diagramas de classe e casos de uso. Ainda segundo (Knublauch, 2003), a dificuldade da UML está no fato de ela ser uma linguagem de programação (orientada a objetos) e haver poucos especialistas de domínio [fora da computação] capacitados a utilizá-la com desenvoltura. Por isso, seria recomendável usar alguma ferramenta mais simples de construção de modelos de domínio.

Muitos pesquisadores (Mizoguchi et al., 1996a 1996b; Bourdeau e Mizoguchi, 1999; Aroyo et al., 2002) utilizaram metodologias de projeto instrucional para construir sistemas educacionais orientados a ontologias ou exploraram a engenharia do conhecimento em tais sistemas baseados em ontologias de tarefas ou em aplicações informacionais baseadas na Web. Chen e Mizoguchi (2004) definem a ontologia como a quintessência da pesquisa orientada a conteúdos, e uma ontologia educacional é definida em termos de como os vocabulários usados no modelo de aprendiz e as atividades de ensino são organizados.

Segundo (Kumar et al., 2004), a sofisticação exigida em uma ontologia de ensino se justificaria pelo contraste entre as estratégias de ensino empregadas em aplicações recentes de ITS e LMS e aquelas usadas por professores humanos, as quais incluem as capacidades de: i) reconhecer princípios instrucionais e modificá-los ou reformulá-los em tempo real; ii) fazer associações precisas entre domínios/tópicos, estratégias de ensino e experiências de aprendizado; iii) adaptar estratégias de ensino para que elas se adequem a necessidades individuais; iv) personalizá-las para ajustarem-se às habilidades dos aprendizes e ao ambiente de ensino; v) empregar estratégias múltiplas para atender as necessidades dos aprendizes; vi) alternar entre representações diferentes das tarefas de aprendizado e vii) planejar um aprendizado/ensino sustentável, enquanto que um ITS seria parcialmente ou totalmente incapaz de fazê-lo. Essa limitada adaptabilidade das aplicações ITS se deve, ainda segundo (Kumar et al., 2004), à “falta de entendimento formal de como várias estratégias de ensino se relacionam umas às outras, ao conteúdo, aos aprendizes, aos professores e ao restante do ambiente de ensino”.

Assim, a busca por uma modelagem de domínio, no caso de um ILTS, deve considerar a natureza particular do aprendizado de uma L2, que é o fato de que a língua é, ao mesmo tempo, o objeto do aprendizado, a principal ferramenta de ensino e o mais importante aspecto a definir a bagagem de conhecimentos do aprendiz. Isso contrasta nitidamente com sistemas de ensino para outras disciplinas, como física, química, geografia ou matemática. Logo, a modelagem do domínio no ensino de línguas demanda uma ontologia própria que, aliada a um modelo pedagógico que represente bem o aprendizado de uma L2, consiga recriar a riqueza e a naturalidade das situações de aquisição da L1.

As ferramentas mais adequadas à construção de modelos de domínio são os editores de ontologias. Um dos editores mais conhecidos é o Protégé, que, embora desenvolvido

originalmente para a aquisição de conhecimentos em sistemas especialistas da área médica, vem sendo utilizado em muitas outras áreas. Entre seus recursos estão a interface de usuário que permite modelar classes (conceitos do domínio), seus atributos e relacionamentos; a geração automática de formulários interativos por meio dos quais os especialistas podem preencher instâncias válidas; uma biblioteca de *plug-ins* que auxilia na definição de uma semântica, em consultas e na definição da lógica de comportamento. Os modelos (classes e instâncias) resultantes podem ser salvos em formatos diversos, como XML, UML e RDF (Resource Description Framework) (Knublauch, 2003). No Protégé classes representam conceitos do domínio que podem ter atributos e relações. Essas classes assemelham-se às de Java e da UML, exceto pelo fato de não admitirem métodos atrelados. As classes podem ser abstratas ou concretas, mas assim como em Java somente as concretas podem ter instâncias. A construção da ontologia de domínio do ILTS é descrita no Capítulo 5.

2.7 Modelagem do domínio baseada em agentes

Grande parte da pesquisa na concepção de IAED prevê interações tutoriais entre duas entidades principais, uma máquina que desempenha o papel de tutor e um aprendiz humano, como ilustrado na Figura 2.2. O objetivo essencial das interações é facilitar, por meio de atividades pedagógicas, a aquisição de conhecimentos por parte do aprendiz. Para isso o sistema tutor se vale de bases de conhecimentos e mecanismos de raciocínio que avaliam as ações do aprendiz e se adaptam às necessidades individuais deste. Nesse contexto, a IA pode prover técnicas de representação e manipulação do conhecimento envolvido. Um dos desafios está em como lidar com a complexidade do conhecimento que o sistema deve ter sobre um domínio que, como no caso aqui tratado, é a língua-alvo, além de considerar aspectos da pedagogia e do aprendiz, organizados em termos da concepção de aprendizagem adotada *a priori*. Os problemas decorrentes dessa abordagem dizem respeito à engenharia do conhecimento.

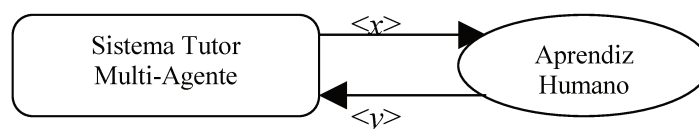


Figura 2.2: Relação entre aprendiz e sistema tutor

Modelagens baseadas em sistemas multiagentes (ABMS) apresentam características muito interessantes quando se buscam técnicas computacionais capazes de tratar fenômenos de alta complexidade por meio do uso de regras simples. Entre os fenômenos que sistemas baseados em agentes são capazes de modelar podem ser citados os de simulação de tráfego e congestionamento, os de logística e otimização de cadeias de suprimentos em linhas de produção, os que simulam efeitos de redes sociais e redes comunicacionais em que ocorra comunicação boca a boca, sistemas de computação distribuída, modelos de gestão de força de trabalho, simulações de comportamento de consumidores, sistemas ecológicos, tráfego em centrais de *callcenter*, entre outros. Em todos esses casos os resultados globais resultam de comportamentos individuais de cada agente e do modo como esses agentes influenciam-

se em função de relações de vizinhança, de “efeito de mundo pequeno” (*small world effect*) ou de seus comportamentos a cada iteração. Esses fatores dão origem às três principais características dos sistemas multiagentes: a atribuição de um agente (objeto) para cada ente de um sistema social/funcional/ecológico, a complexidade que isso enseja e a emergência de fenômenos não previstos. Implementações ABMS permitem, por exemplo, simular os efeitos de formação de revoadas de pássaros ou de enxames de insetos sem que para isso se tenha que recorrer a regras gerais, só aos comportamentos individuais de cada agente, que é uma entidade inteligente, com objetivos próprios e situada no tempo e no espaço.

O uso de sistemas multiagentes como forma de inteligência distribuída desenvolveu-se a partir de meados dos anos oitenta, com a teoria da Sociedade da Mente (Minsky, 1986), que defende a construção de sistemas inteligentes por meio da interação de uma sociedade de “agentes sem mente”, tendo cada um suas competências específicas. A ideia de que esses módulos de competência cooperem localmente de maneira que a sociedade como um todo funcione adequadamente e cumpra seus objetivos, é, segundo Maes (1990), muito atraente por suas características de descentralização, modularidade, robustez e por possibilitar a emergência de uma funcionalidade global. Ainda segundo Maes (1990), esse paradigma de “inteligência emergente” não se enquadra por completo na abordagem tradicional da IA (na qual a competência é programada), nem na do conexionismo (no qual a competência é o resultado de um aprendizado do tipo “*tabula rasa*”). Tampouco esse se apresenta como um sistema híbrido, com componentes simbólicos e subsimbólicos distintos. Em lugar disso, o algoritmo integra por completo características de ambas as abordagens por meio do uso de um modelo computacional conexionista dentro de uma representação simbólica estruturada (Maes, 1990). Assim, ele combina o melhor dos dois mundos, herdando do conexionismo as interessantes propriedades de paralelismo intrínseco, de tolerância a falhas, capacidades sofisticadas de recuperação e *matching*, continuidade e emergência de resultados globais a partir de regras uniformes de interação local, sem todavia depender somente de aprendizado e de classificação. Da IA simbólica, por sua vez, ele adota princípios de representação e de estruturação, nos quais a rede é conhecida *a priori* e seus *links* têm significados específicos e inteligíveis e cujos nós são unidades grandes e significativas (Maes, 1990).

No paradigma de modelagem baseada em agentes o fenômeno real é decomposto em um conjunto de elementos (agentes) e em suas interações, e o modelo geral é o resultado dessas interações entre os agentes. Esse artifício facilita a construção de modelos de simulação, sobretudo em se tratando de simular sistemas complexos, como é o caso da seleção de vocabulário em um ILTS, e atende a necessidade de um paradigma de programação que contemple distribuição, autonomia, flexibilidade e heterogeneidade. Esse paradigma deve permitir uma fácil adição ou remoção de funções, de forma extensível e flexível, requisitos que são preenchidos pelo conceito de agente.

Do ponto de vista conceitual, não há consenso quanto à definição exata do que seja um agente. Algumas definições são encontradas na literatura, por exemplo, a de que:

“Um agente é uma entidade à qual se pode associar uma identidade única, e que é capaz de realizar cálculos formais. Um agente pode ser considerado como um meio que produz um certo número de ações a partir dos conhecimentos e mecanismos que lhe são próprios” (Gasser e Hoffmann, 1992).

Segundo Klüegl *et al.* (2001), os ambientes devem ter algumas propriedades para que a simulação baseada em agentes possa ocorrer de forma adequada. São elas:

- Identificar claramente os agentes, seus componentes e atividades no ambiente;
- Agentes têm um comportamento não-trivial, e os mais simples são implementados de forma mais eficiente;
- Quanto mais simples a estrutura do agente, melhor para o sistema, pois o foco da modelagem deve ser o comportamento e as interações dos agentes, e não suas habilidades internas;
- É necessária uma descrição do sistema e do ambiente em que o agente se insere, de forma a facilitar a implementação e a solução de possíveis erros.

Fica claro que a caracterização do agente é um aspecto importante da modelagem de sistemas complexos. Caracterizar o agente envolve levantar e dimensionar, por exemplo, seus atributos, atitudes, crenças e preferências. Uma definição correta das características de cada agente torna o modelo mais robusto e é fundamental para a qualidade dos resultados obtidos, visto que o modelo consiste nas funcionalidades e interações dos agentes.

Mas para que o comportamento combinado dos agentes conduza a estados que façam sentido do ponto de vista didático é necessário implementar algum mecanismo que selecione as ações, isto é, que organize o comportamento dos agentes. Mecanismos de seleção de ação (SA) têm o objetivo de solucionar uma das questões mais básicas dos sistemas inteligentes: decidir o que fazer em seguida. Na área da IA a seleção de ação está em geral associada a implementações baseadas em agentes, que, como visto, são sistemas artificiais capazes de demonstrar comportamentos complexos dentro de um ambiente no qual diversos agentes interagem e influenciam uns aos outros.

A especificação completa de um mecanismo de seleção de ação pressupõe a prévia definição dos tipos de ação que se pretende selecionar, sendo esse um conjunto predefinido e fixo, na maior parte dos casos. Os tipos de ação podem ser, no caso de um autômato ou robô, movimentar um membro em uma direção, girar uma junta, ou ficar imóvel.

Em geral as implementações buscam mecanismos de seleção de ação para ambientes dinâmicos e imprevisíveis, com agentes que desempenham várias tarefas diferentes, operam em tempo real e tomam decisões imediatas. Uma característica recorrente são os conflitos por alocação de recursos limitados, como, no caso que é objeto do presente estudo, o acesso à interface do curso. Em alguns casos, os ambientes em que os agentes atuam podem incluir usuários humanos, o que eleva a complexidade para as decisões dos agentes autômatos. Esse é também o contexto do ILTS aqui descrito, no qual os agentes do tutor inteligente devem estar atentos a ações do aprendiz e também ao modelo de domínio.

Na programação de um agente inteligente a inclusão de um mecanismo de seleção de ação se mostra essencial. Sua realização, todavia, pode se dar de forma distribuída ou ficar a cargo de um módulo especializado. Os mecanismos de seleção de ação podem ser divididos em algumas categorias:

- Sistemas reativos (ou de planejamento dinâmico): computam uma única ação a cada instante com base no contexto momentâneo e em planos previamente codificados.

- Sistemas simbólicos (conhecidos como de planejamento clássico): baseiam-se na crença de que a melhor estratégia de seleção é baseada em planos otimizados e de que, para tanto, agentes capazes de manipular símbolos são necessários e suficientes para se obter um comportamento inteligente.
- Sistemas distribuídos: ao contrário dos simbólicos, não têm um módulo único que decide a próxima ação, mas muitos módulos distribuídos, operando em paralelo e levando em conta informações locais para determinar a melhor ação, a qual, se espera, emergirá das interações.

Existem ainda alguns tipos de SA que não se encaixam perfeitamente nessas três categorias.

Uma arquitetura que se mostrou muito influente foi proposta por Maes (1989). Ela emprega dois tipos de módulos de competência, que equivalem ao que em algumas outras arquiteturas se denomina módulos de comportamento. O primeiro tipo é o de *Ação*, e quando ativado perfaz uma ação física. O segundo, de *Crença*, tem a função de, quando ativado, fazer o robô adotar uma dada crença. Todo módulo de competência tem um nível de ativação que é um número real e que é usado para determinar qual módulo de competência controla o agente autônomo global (ou quais ações e crenças ele adotará). Essa arquitetura propõe para o mecanismo de seleção dos módulos o uso de uma rede de espalhamento de ativação, formada por ligações entre módulos. Nessa rede algumas das ligações são predecessoras, outras são sucessoras, e outras são inibidoras. A dinâmica de troca de ativação por essas ligações faz com que seja selecionado o módulo mais indicado a um dado contexto. Assim, em comparação com outras arquiteturas baseadas em agentes, ela mostra maior integração de funções cognitivas diferentes, com o uso de um mesmo mecanismo para o planejamento (seleção de ação) e para inferência (seleção de crença). A sua rede não só produz seleção de ação orientada a meta e a contexto, mas também inferências e percepção orientadas a meta e a contexto. Ela foi proposta para ambientes ou tarefas de difícil caracterização, e que exigem do agente experimentação e aprendizado.

Segundo Dubois (2007), um aspecto interessante dessa arquitetura é que ela não requer um mecanismo de unificação que a cada ciclo cognitivo percorra toda a base de regras. Tudo o que há é um processo leve de adição que soma todas as fontes de energia para cada nó de competência. Com isso, o processo de raciocínio da rede de comportamento é diferencial, cumulativo, com a energia encontrando seu caminho pela rede a cada ciclo cognitivo. Ele acrescenta que, graças a esse fluxo de energia e à acumulação de ativação nos nós sucessores, soluções alternativas estão sempre sendo preparadas (caso disponíveis). Essa preparação natural de alternativas é também útil à adaptabilidade do agente quando a solução inicialmente favorecida fica travada no processo de seleção porque uma condição permanece ausente do seu contexto necessário (pré-condição).

A versatilidade das implementações baseadas em agentes com mecanismos de SA foi então levada em consideração no estudo aqui descrito como uma solução de engenharia do conhecimento para a modelagem do domínio no ILTS. Isso permite modelar a língua-alvo em termos de seus elementos constituintes, e daí derivar os efeitos (previsíveis ou emergentes) da interação desses elementos na interface do ambiente de ensino, tanto no contexto de sequenciamento de vocabulário quanto na formação de sintagmas e sentenças válidos na L2. Os detalhes dessa estratégia serão tratados no Capítulo 5.

Capítulo 3

O aprendizado de uma língua

Como discutido no Capítulo 2, este estudo explora o ensino de línguas assistido por computador. É assim importante em primeiro lugar entender a aquisição de nossa primeira língua (L1), a língua materna, estudando como se dá esse salto cognitivo e quais são as suas implicações para o posterior aprendizado de uma segunda língua (L2), processo este muito distinto em diversos aspectos, mas que é bastante influenciado pelos conhecimentos da L1. Para tanto, nos referenciamos ao livro “Como as crianças aprendem os sentidos das palavras” (Bloom, 2000), em razão de sua rica análise desse tema. Além dessa, outras obras pertinentes serão citadas ao longo do capítulo.

3.1 A aquisição da língua materna

A aquisição da língua materna (L1) é em grande medida determinada pelo processo de assimilação, pela criança, das palavras que compõem o sistema de comunicação ao seu redor. Para Bloom (2000, p.15), o estudo dessa aquisição requer, todavia, uma definição apropriada do que é uma palavra. Uma definição comum poderia ser a de que a palavra é um átomo sintático (verbo, adjetivo, etc.), que pode ser o produto de regras morfológicas (Pinker, 1994, *apud* Bloom, 2000) e que representa um som escrito ou uma sequência de letras. Assim, segundo Bloom, por essa definição, a sentença *John stayed in the poker game until he got cleaned out* (João continuou no jogo de pôquer até ser depenado) teria 11 palavras. Mas Bloom salienta que essa definição é inadequada quando se deseja estudar a assimilação do léxico pelas crianças durante a aquisição da L1. Por exemplo, na sentença acima, a palavra *stayed* (ficou) não precisará ser aprendida se já forem conhecidos o verbo *stay* (ficar) e a regra morfológica que adiciona *-ed* para indicar que o verbo está no tempo pretérito. Isso significa então que somos capazes de usar um número de palavras (no sentido morfológico) muito maior do que as que tivemos efetivamente que aprender. No português, por exemplo, assim que aprendemos o verbo *falar*, as formas *falei*, *falou*, *falará*, etc. vêm de graça, contanto que as regras morfológicas para formá-las já tenham sido aprendidas.

Por outro lado, Bloom (2000, p.16) chama a atenção para o fato de que no caso das expressões idiomáticas, como *cleaned out* (depenado), da sentença exemplificada acima, não basta saber o sentido do verbo *clean* e da preposição *out* para ser capaz de entender o significado da expressão; é necessário aprender algo mais: seu sentido idiomático, que na maioria dos casos deve ser aprendido como se se tratasse de um novo item lexical. Esse é também o caso das palavras homógrafas, como *poker*, do exemplo citado, que em inglês pode significar tanto o jogo de cartas quanto um instrumento para uso em lareiras, razão pela qual cada um desses sentidos deve ser aprendido separadamente.

Assim, Bloom conclui que, do ponto de vista da aquisição linguística, a definição de palavra deve incorporar todas as formas (e nada além delas) cujos significados devem ser aprendidos, o que corresponde a signos Saussurianos, como cachorro, cair fora, ganhos de capital, e Cidadão Kane, mas exclui termos como *cachorros*, *cachorrinho*, etc.

Segundo Bloom (2000, p.18), o aprendizado e a representação de significados são ainda um tema controverso. Na visão tradicional, por exemplo, de inspiração aristotélica, o significado de uma palavra é o que determina sua referência, e cada palavra tem uma extensão (a que entidades ela se refere) e uma intenção (o que essas entidades têm em comum). Assim, o significado de *cachorro* determina as coisas que são e as que não são cachorros. Bloom, citando Murphy (1991), comenta que essa visão está implícita em quase todas as discussões psicológicas sobre aprendizado e representação. Uma visão tradicional na psicologia é a de que significados são desenhos: o significado de *cachorro* é o desenho de um cachorro, e sabemos esse significado quando temos uma representação mental daquele desenho que nos permita diferenciar cachorros de não-cachorros. Uma outra noção muito popular é a de que significados são conjuntos de atributos ponderados, de tal forma que conhecer o significado de *cachorro* é possuir uma representação mental do conjunto apropriado de atributos que nos permitirá determinar em que medida os diversos objetos no mundo são cachorros. Outras visões incluem a ideia de que significados são modelos mentais, ou nós de uma rede semântica, ou ainda conjuntos de exemplares específicos. Bloom (2000, p.19) comenta que essas visões partilham da suposição de que conhecer o significado de algo envolve a capacidade de diferenciar essa entidade de todas as outras, e que a despeito das diferentes teorias para explicar o conhecimento das palavras, todos seus defensores concordam que é essencial que a teoria seja capaz de categorizar os conceitos, o que reflete o senso comum a respeito dos significados de uma palavra. Logo, se traduzimos do grego antigo a palavra *estrela*, pouco importa que os gregos pensassem que as estrelas fossem buracos no céu; basta que utilizemos a palavra para nos referirmos às mesmas coisas; qualquer sobreposição cognitiva adicional é desnecessária (Bloom, 2000, p.19). Mas aqueles casos em que uma mesma palavra evoca diferentes significados para diferentes pessoas parecem sugerir que o problema ultrapassa a esfera cognitiva ou neurológica, tornando-se aí uma questão sociológica, relacionada, entre outros, a autoridade, experiência, padrões comunitários (Chomsky, 1995, *apud* Bloom, 2000).

Parece haver então, segundo Bloom, duas linhas concorrentes de pensamento acerca do conceito de significado, uma psicológica, que recorre às representações mentais, e uma filosófica, que se apoia na noção de referência. E ainda que alguns autores desaconselhem a fusão dessas visões, Bloom afirma que é impossível para cada uma dessas visões ignorar por completo a outra, e que em razão disso e da falta de uma alternativa completa, seria aceitável recorrer a ambas. Com isso, pode-se assumir que há dois aspectos (determinantes) do significado de uma palavra – um aspecto psicológico interno, às vezes chamado de conteúdo estrito, e um aspecto contextual e social externo, às vezes chamado de conteúdo amplo. Em conjunto eles determinam a que uma palavra se refere. Na discussão que segue, a expressão *significado de uma palavra* será usada no sentido de conteúdo estrito.

Por fim, cabe lembrar que a associação da aquisição do léxico com as referências e categorizações funciona melhor para substantivos comuns, nomes próprios, e pode também se estender a alguns adjetivos e verbos. O problema está em aplicar esse raciocínio a determinantes, preposições e modais, cuja semântica é bem diferente, e cujo significado não é função da referência, mas dos papéis que eles desempenham na modulação do significado de outros termos, referenciais (Bloom, 2000, p.22).

3.2 O ritmo da aquisição das palavras

Segundo alguns autores (Aithcinson, 1994; Miller, 1996; Pinker, 1994 *apud* Bloom 2000), ao final do ensino médio cada jovem nos Estados Unidos e no Reino Unido conhece, em média, 60 mil palavras. Dado que a aquisição de palavras se inicia em torno dos doze meses de idade, isso significa ter adquirido 3750 palavras por ano, dez palavras por dia, uma palavra a cada 90 minutos. Por outro lado, pondera Bloom (Bloom, 2000, p.25), esses números não devem ser tomados em termos absolutos, pois adquirir o significado de certas palavras, sobretudo verbos, pode ser um processo longo, que exige tentativas repetidas. Então pode ser que a criança adquira não dez palavras por dia, mas que ela diariamente adquira um centésimo de cada uma de mil palavras diferentes (Bloom, 2000, p.25). Tampouco o ritmo da aquisição é constante: as crianças aos três anos de idade não adquirem dez palavras por dia, mas não mais que dez por semana, o que significa que esse início lento necessariamente implica uma posterior aceleração da aquisição, para em torno de 15 palavras por dia. Em qualquer dos casos esse é um feito notável, visto que adquirir uma palavra exige a memorização da relação arbitrária entre uma forma e um significado, e essa associação é notoriamente lenta e difícil, como bem ilustra, por exemplo, o esforço necessário para memorizarmos os nomes das capitais de cada país (Bloom, 2000, p.25). Some-se a isso o fato de que a imensa maioria dessas 60 mil palavras não é encontrada com frequência no cotidiano. Esse fato desconcertante enseja assim algumas importantes indagações: Qual a natureza dessa aquisição? Quais as suas fases? Quanta exposição é necessária para que uma criança adquira uma nova palavra?

Bloom relata resultados empíricos que comprovam a existência de um mecanismo de mapeamento rápido, tanto em crianças, quanto em adultos, por meio do qual somos capazes de reter o nome de um objeto que nos foi apresentado em um contexto incidental. Esse mecanismo não se limita à aquisição de palavras, e se estende a fatos arbitrários que tenham sido apresentados linguisticamente, relativos a uma dada entidade, mas não se aplica a toda e qualquer tarefa arbitrária de memorização. Uma explicação possível para essa distinção é a de que o mapeamento rápido só se aplique a fatos sociais e históricos, que só são acessíveis através da atenção ao que é dito em nosso entorno, informação essa que, caso seja esquecida, pode nunca mais voltar a estar acessível. Em contraste, informações que são percebidas como estando “armazenadas no mundo”, tendem a não ser memorizadas rapidamente, pois supostamente estarão disponíveis no ambiente de forma mais perene, bastando ao observador dar uma nova olhada para recuperar a informação.

Assim, os resultados de experimentos controlados da aquisição de palavras pelas crianças mostram que o mapeamento rápido emerge de uma capacidade geral de aprender informações transmitidas socialmente – incluindo o, mas não limitadas ao, significado de palavras. Isso reforça o papel da cognição social nessa aquisição (Bloom, 2000, p.34-35).

Mas voltando à questão da velocidade de aquisição léxica pelas crianças, Bloom comenta que diversos estudos sugerem que a taxa diária de palavras aprendidas é de 0,3 dos 12 aos 16 meses de idade; 0,8 dos 16 aos 23 meses; 1,6 dos 23 aos 30 meses; 3,6 dos 30 meses aos seis anos; 6,6 dos seis aos oito anos; e 12,1 dos oito aos dez anos. As possíveis razões para o início aos 12 meses e para o aumento da velocidade de aquisição até os dez anos incluem a gradual adequação do conhecimento fonológico, a maturação da memória e

da atenção, o afloramento das habilidades conceituais que permitem às crianças codificar as noções às quais as palavras correspondem e que, embora já estejam latentes nos primeiros meses de vida, só a partir do primeiro ano passam a se associar a um entendimento adequado de tipos e indivíduos. De resto, os resultados de vários estudos sugerem que só a partir dos 18 meses as crianças compreendem que apenas palavras fonológicas são usadas normalmente como signos (no sentido Saussuriano) e que as pessoas nomeiam as coisas com palavras fonológicas, não com gestos ou guinchos. Só então o entendimento que elas têm da comunicação se integra plenamente ao entendimento da linguagem.

Um outro fator que vem contribuir para paulatinamente acelerar a aquisição é o ganho de experiência com o idioma, que pouco a pouco habilita a criança a se valer de informações sintáticas para elucidar a semântica (significado) de algumas das palavras novas com as quais ela se depara. O fato de que esse mecanismo, conhecido como pistas linguísticas (*linguistic cues*), não se manifesta no início da aquisição nada tem a ver com a tenra idade dos sujeitos, mas com o simples fato de que as pistas não estão disponíveis quando os conhecimentos são incipientes. Isso faz com que as primeiras aquisições léxicas se deem sem a ajuda de pistas sintáticas.

Com o início da alfabetização, em torno dos sete anos, a criança é exposta a mais palavras, agora também em registros de linguagem escrita. Isso a levará a acumular um vocabulário ainda maior e, por vezes, bastante impressionante como nos casos de adultos cultos. Por fim, em um dado momento a taxa de aquisição léxica começa a se reduzir, não por perda da capacidade de assimilação, mas porque na idade adulta já teremos adquirido a maior parte das palavras pertinentes ao nosso ambiente imediato (Bloom, 2000, p.47).

3.3 As bases da aquisição das palavras

É certo que a aquisição do vocabulário é um processo social, visto que palavras são convenções arbitrárias estabelecidas pela comunidade à qual a criança pertence. Mas no que diz respeito ao desenvolvimento cognitivo subjacente a essa aquisição, não há ainda, segundo Bloom, consenso entre os autores. Uma linha teórica tradicional, denominada associacionismo, partilhada entre outros pelos filósofos empiricistas John Stuart Mill, David Hume e John Locke, e defendida também por autores contemporâneos, propõe que essa aquisição decorre da sensibilidade à covariância de ocorrência: se frequentemente o nome de um objeto é pronunciado quando a criança o está vendo ou está pensando nele, a palavra e o pensamento se tornam associados, e em razão disso se pode dizer que a criança aprendeu o significado da palavra (Bloom, 2000, p. 56). Essa teoria inspirou Skinner na explicação behaviorista do aprendizado baseada no estabelecimento de conexões entre estímulos e respostas por meio de reforço e punição, e mais recentemente ensinou teorias connexionistas, inclusive com implementações computacionais. Em uma delas (Plunkett *et al.*, 1992, *apud* Bloom, 2000), uma rede é alimentada com rótulos e imagens (através de caminhos sensoriais distintos) e treinada a associá-los. Quando a rede se torna capaz de gerar os rótulos corretos em resposta a uma imagem, considera-se que se trata de uma produção léxica correta, e quando uma imagem correta é produzida em resposta a um rótulo, considera-se que se trata de uma compreensão correta.

Todavia, Bloom (2000, p.58) adverte que, embora o associacionismo tenha o mérito de oferecer uma explicação parcimoniosa para a aquisição, ele falha quando confrontado com observações reais, que mostram que em grande parte das situações em que as crianças são expostas a novas palavras não existe uma relação direta clara entre a elocução do nome e a simultânea presença do objeto correspondente no campo de visão da criança, o que pela teoria associativista deveria implicar um período inicial de associações incorretas, que duraria até que as associações corretas se tornassem estatisticamente preponderantes. O fato de que a aquisição da relação nome-objeto é em geral rápida e sem erros sugere que não se trata de uma forma de aprendizado estatístico (Bloom, 2000, p.59), o que não significa uma rejeição absoluta do connexionismo como explicação do mecanismo subjacente à aquisição, mas tão somente a constatação de que a sensibilidade à covariância estatística não é a chave da aquisição do léxico. Essa chave parece estar, segundo Bloom, no papel desempenhado pela teoria da mente (ou cognição social), isto é, na capacidade que as crianças demonstram de “ler a mente” ou, em outras palavras, de perceber as intenções e o foco de atenção dos adultos. Essa capacidade parece habilitar a criança a compreender a que os adultos se referem quando usam uma dada palavra, o que, somado às demais capacidades perceptivas da criança, permite a ela estabelecer os relacionamentos entre nome e objeto. Assim, as dificuldades de aquisição do léxico por parte das crianças autistas podem ser atribuídas ao seu déficit de cognição social, pois lhes faltam as capacidades inferenciais de que as demais crianças dispõem mesmo antes de dois anos de idade (Bloom, 2000, p.80).

O que a teoria da mente faz é habilitar a criança a estabelecer o mapeamento entre uma palavra e um conceito, mas isso pressupõe a disponibilidade do conceito. Não importa com que frequência uma criança ouça palavras como *hipoteca* ou *debênture*, ela não será capaz de adquirir um termo cujo conceito lhe esteja inacessível aos dois anos (Bloom, 2000, p.86). Em síntese, segundo Bloom (2000, p.87), o papel da teoria da mente na aquisição das palavras é o de que as inferências que as crianças fazem a respeito das intenções referenciais das outras pessoas são usadas para criar setas ou ponteiros que ligam as palavras ao mundo. Mas a ponta da seta não toca os objetos no mundo. Cabe à criança descobrir, a partir disso, a que é que a palavra se refere, se à categoria (coelho), se a um indivíduo (Sansão), ou a uma propriedade (branco). Bloom conclui então que, mesmo com uma teoria da mente em plena atividade, os problemas da criança apenas começaram.

Muitos resultados experimentais sugerem a existência de uma tendência de que a criança interprete novas palavras como sendo nomes de objetos. Na verdade, essa tendência se manifesta também em adultos. Quando apresentados a um novo objeto e a uma palavra que se refere a esse objeto, tomamos a palavra como sendo um nome do objeto. Mas isso leva a outra questão pertinente: o que é um objeto? O relacionamento entre partes e todos é uma dificuldade em particular. Pode ser que qualquer coisa possa ser considerada um objeto se quisermos pensar nela dessa forma. Mas, segundo Bloom (2000, p. 93), é importante adotar uma definição compatível com as tendências de aquisição léxica das crianças, e o conceito de *objetos-Spelke* atende esse requisito, sobretudo com o princípio da coesão: ser um objeto é ser uma região de matéria conexa e limitada que mantém essas características mesmo quando em movimento. Porém, não é preciso ver o movimento para compreender que algo é um objeto. Isso significa que uma cena estacionária pode ser decomposta em

objetos distintos, pois as descontinuidades entre as entidades implicam que elas não se moveriam juntas. E enquanto o princípio de coesão revela a nossa compreensão do que significa *ser um objeto*, os princípios de solidez, continuidade e contato descrevem nosso entendimento de como objetos devem se comportar (Bloom, 2000, p. 94-95).

Como comenta ainda Bloom (2000, p. 108), muito embora a propriedade de ser um objeto-Spelke possa ser condição suficiente para definir um indivíduo nomeável, ela não é em absoluto necessária, como nos casos de partes (dedo, olho, ponta), ações (tossir, sorrir), períodos de tempo (minuto, hora), espaços negativos (buraco), sons (barulho), coletivos (família, floresta). Neste último caso, por exemplo, para aprender tais palavras a criança tem que de algum modo superar a tendência de focalizar objetos inteiros e em lugar disso passar a enfocar grupos (Bloom, 2000, p.110-111). Uma possibilidade é a de que algo possa ser visto como um coletivo se for tratado como um indivíduo único por alguém ou algo (Bloom, 2000, p. 113). Uma pista para a individualização de coletivos pode ser o movimento comum, mas ainda não se sabe se as crianças são capazes de seguir essa pista.

Assim, o que foi discutido acima sugere que os possíveis referentes para a atribuição de nomes (comuns ou próprios) a entidades do mundo emergem de dois sistemas cognitivos distintos: um deles relativo à contiguidade de matéria que satisfaça os princípios-Spelke, sobretudo a coesão, e outro baseado na teoria da mente que analisa movimento e matéria por meio de uma compreensão de objetivo, função e intenção (Bloom, 2000, p.114).

Por outro lado, o significado de algumas palavras, como artigos definidos, não pode ser aprendido por meio da audição da palavra isoladamente. A criança deve se valer de expressões em que essas palavras aparecem, como “o cachorro”, “a baleia” para inferir o significado desses determinantes. Mas esses determinantes correspondem, segundo Bloom, a contrastes que são menos salientes em termos semânticos do que no caso de substantivos: a diferença entre *cachorros* e *cadeiras* é mais forte do que entre *uma cadeira* e *várias cadeiras*, e algumas das noções semânticas codificadas por determinantes são bastante sutis. A aquisição desses itens lexicais então exige mais do que uma simples referência do tipo *token-type*, seja na aquisição da L1, seja em um posterior aprendizado de uma L2.

Os pronomes, por sua vez, pertencem a uma classe linguística conhecida como indicial (ou deítica, para usar uma denominação proposta por Peirce), cuja interpretação muda radicalmente conforme o contexto em que são usados, como na sentença “Traga-me o copo da esquerda agora”, na qual a interpretação das palavras *me*, *esquerda* e *agora* (deíticos pessoal, espacial e temporal, respectivamente) depende de quem seja o falante e de onde ele esteja quando fala (Bloom, 2000, p.122). E dado que a aquisição dos pronomes pessoais requer foco de atenção na (e compreensão da) conversa de outras pessoas, tanto crianças autistas quanto cegas têm dificuldade nessa aquisição, possivelmente em razão do déficit de atenção das primeiras, e da falta de orientação visual das últimas.

Assim, a aquisição da L1 é um processo longo que exige da criança algumas faculdades mentais e sensoriais específicas, que a habilitem a: i) capturar auditivamente os signos (palavras); ii) perceber visualmente o contexto em que esses signos são utilizados; iii) conceber um modelo mental das intenções das demais pessoas quando estas se utilizam dos signos; iv) memorizar os signos e sua relação com o contexto. Mas isso não significa que exista uma teoria consensual sobre tal processo (no Apêndice 1 são resumidas teorias).

3.4 A teoria ecológica da aquisição de uma língua

A busca por uma compreensão da aquisição linguística conduziu a várias teorias concorrentes, algumas delas seguindo uma corrente empiricista de pensamento que credita a aquisição linguística à interação com fatores externos, enquanto outras são nativistas, isto é, defendem a existência de habilidades inatas para a aquisição linguística. A análise dessas teorias revela conflitos e também pontos convergentes quanto à ênfase dada aos fatores que explicariam a aquisição linguística (ver mais detalhes no Apêndice 1).

Por sua vez, a teoria ecológica da linguagem propõe uma explicação alternativa da aquisição linguística em que habilidades inatas e fatores ambientais interagem para que nosso aparato cognitivo apreenda o mundo a nossa volta.

A teoria ecológica descreve os efeitos mentais/cerebrais da aquisição com base no *paradigma dos esquemas*, admitindo com isso a existência de atributos inatos que habilitam o ser humano a construir significados e desenvolvê-los na forma de sistemas de linguagem. Mas a constatação de uma predisposição congênita para a linguagem simbólica é entendida como uma instância de capacidades cognitivas mais genéricas, e portanto não significa a concessão de um status de gramática universal aos fatores inatos, que, embora não estejam explicitados em várias das teorias baseadas na educação, não são incompatíveis com elas.

Segundo Fettes (1999), o conceito de *aquisição* linguística sugere um processo no qual um objeto externo, “língua”, é capturado e guardado pela mente do aprendiz. Um dos inconvenientes desse paradigma é que ele despreza o conhecimento individual, tratando toda pessoa como *tabula rasa* ou como se trouxesse faculdades inscritas geneticamente para a construção de significados. A teoria ecológica (Reed, 1996) propõe que a evolução da vida e a seleção natural asseguraram o desenvolvimento de mecanismos de percepção e de resposta a **informações persistentemente disponíveis** no ambiente denominados *esforço em busca do significado*. Ela recorre ao *paradigma dos esquemas*, em que esses são redes neurais hierárquicas complexas associadas a estados particulares de consciência e ação. A partir desse ponto de vista, o conhecimento não é “armazenado”, isto é, guardado de forma permanente e segura, mas provisoriamente retido, como qualquer outra capacidade que enfraquece quando não usada regularmente (Ben-Ze'ev, 1995). O paradigma dos esquemas enfatiza a natureza holística da percepção e do aprendizado. Novas capacidades adquiridas não são armazenadas em “prateleiras” ou “depósitos” separados, mas na forma de modificações no sistema cognitivo como um todo: estrutura e conteúdo são a mesma e única coisa. Duas consequências desse modelo são o fato de que o aprendizado é algo mais profundo do que em geral se supõe, pois ele implica mudar o que o organismo é, e envolve alterações em um sistema cognitivo que já é altamente interconectado. Nesse sentido, aprendizes, seja de uma língua, seja de outras habilidades, nada têm de *tabula rasa*: eles são sistemas complexos altamente adaptados à sua maneira, e cuja negociação com o processo de aprendizado será fortemente influenciada por seus **estilos individuais** frente ao mundo.

Ainda segundo Fettes (1999), o paradigma ecológico, em contraste com o idealismo Cartesiano, implica que o conhecimento é **ativo**, resultado de um esforço que dura toda a vida na busca pelo significado, que ele é **realista**, derivado do contato direto com o mundo.

Segundo Fettes, a visão ecológica vê a linguagem como um sistema emergente e evolutivo de atos públicos coordenados, preexistente e externo a todo agente individual:

Isso significa que os indivíduos nunca realmente “adquirem uma língua” no sentido de serem capazes de reproduzir o sistema como um todo em toda sua complexidade dinâmica. Os indivíduos aprendem a reproduzir **elementos** do sistema, que compreende uma miríade de agrupamentos naturais e sub-sistemas de dispositivos públicos de linguagem, para coordenar suas próprias ações e consciências, assim como a de outros de maneiras que sejam úteis a eles. É verdade que o estoque de capacidades de um indivíduo para interpretar e produzir dispositivos de linguagem constitui ele próprio um sistema, mas a ontologia e a epistemologia desse sistema privado difere nitidamente do **fenômeno emergente da linguagem pública** (realce nosso).

Ainda segundo Fettes, os *tokens* linguísticos não são automaticamente interpretados, simplesmente em razão de sua disponibilidade no ambiente: eles devem ser percebidos como relevantes ao foco de atenção atual. Por isso, a comunicação linguística se baseia no **atingimento de uma consciência compartilhada** entre os participantes. Essa consciência compartilhada, como discutido na seção anterior, significa termos um modelo da mente de nossos interlocutores. Assim, no âmbito da psicologia ecológica, aprender quando e como atingir a consciência compartilhada com outra pessoa é uma capacidade vital que a criança começa a adquirir logo após o nascimento, para então, por volta de nove a doze meses, se tornar apta a tomar parte em quadros de interação triádica que envolvem, além dela própria, uma pessoa adulta e algum objeto ou evento (Fettes, 1999). À medida que a criança cresce, ela aprende a participar de quadros de interação mais amplos, inclusive quando o interlocutor não está presente fisicamente, e na comunicação impressa e de gêneros textuais diversos. Esses quadros usam uma variedade de signos linguísticos, miméticos, etc., e novos quadros específicos podem derivar de outros mais genéricos, e nosso conhecimento da língua materna não resulta de dicionários e gramáticas, mas de elocuições concretas que ouvimos e reproduzimos na comunicação com pessoas ao nosso redor.

No que trata do aprendizado de uma L2, as linguísticas moderna e aplicada sofrem, segundo Fettes, forte influência da razão normativa, que trata dicionários e gramáticas como definidores mais válidos da língua do que os comportamentos reais observáveis dos falantes, o que dá lugar à noção de “falante nativo ideal” em contraste com os falantes reais.

3.5 O aprendizado de uma L2 segundo a teoria Semiótica¹

Peirce considerava o aprendizado um processo essencialmente semiótico, ligado intimamente à arte de prestar atenção em princípios gerais e deixá-los permear para dentro do raciocínio. Se o aprender equivale a aumentar a habilidade de compreender signos, isso começa com a capacidade de compreender ícones e, assim que possível, compreender índices a partir de *simplex* icônico, e símbolos a partir de um complexo icônico-indicial. Um índice sem um ícone é cego, um símbolo sem um índice é vazio (De Tienne, 2002).

Segundo Peirce, ícones trazem em si uma capacidade para experiência, mas essa capacidade só pode ser explorada dentro dos limites da real experiência que o intérprete já teve em relação ao mundo em que o signo aparece. Para serem reconhecidos, ícones muitas

¹ Mais detalhes sobre a teoria Semiótica são fornecidos no Apêndice 2.

vezes exigem poder de discriminação consideravelmente sutil. A iconicidade aumenta com a experiência, assim como a simbolicidade. O processo representacional se inicia com a iconicidade, pois não há signo que não contenha, minimamente, ícones em sua realização.

Aprender é então uma questão de ampliar o campo de interpretância por meio do teste da experiência. Aprendizado é virtualmente raciocínio, o qual conecta algo que acaba de ser aprendido com conhecimentos já adquiridos para que desse modo aprendamos o que era desconhecido. O raciocínio é uma nova experiência que envolve algo velho e algo até então desconhecido. Como visto na teoria ecológica, o aprendizado implica modificações no sistema cognitivo como um todo, não sendo possível separar estrutura e conteúdo.

De Tienne comenta ainda que, para Peirce, aprender consiste na apreensão de um *continuum*, que é uma ideia geral que permanece como tal se não há reflexão ou análise. O poder da generalidade está nesse controle das conexões, que são a costura do *continuum*. Aprender é em essência apreender as leis que governam as conexões.

Segundo Danesi (1997), a Educação seria em essência aprender como representar o mundo e, assim, interpretá-lo de formas diferentes. Quando a criança entra em contato com um objeto desconhecido, sua primeira reação é explorá-lo com os sentidos, manipulá-lo, cheirá-lo, sentir seu paladar, escutar seus sons e examiná-lo visualmente. Essa fase pode ser chamada de *cognição sensorial*. Os estudos da Psicologia mostram que, ao preencher sua memória com esses estímulos, a criança se habilita a posteriormente *reconhecer* o mesmo objeto sem ter que examiná-lo novamente.

Ao se desenvolver, a criança gradualmente transcende a fase sensório-cognitiva, passando ao comportamento representacional, no qual ela começa a apontar para o objeto, imitar seus sons, etc. Nesse estágio o objeto passa a assumir uma forma representacional de existência, que Peirce denominou *representamen*, e na qual a imagem do objeto pode ser evocada ainda que o próprio objeto não esteja ao alcance dos sentidos da criança. Esses signos, cunhados culturalmente e adquiridos pela criança inserida naquele contexto, juntamente com o significado que eles carregam, permitem à criança aceder à ordem cultural. Por meio de seu uso continuado, os signos adquiridos no contexto vão gradualmente ganhar dominância cognitiva na criança e, por fim, vão mediar e regular os pensamentos, ações e comportamentos dela (Danesi, 1997). O aprendizado da L1 parte então de uma forte fundamentação nos objetos e chega a uma autonomia plena dos signos.

A perspectiva semiótica teria também implicações pedagógicas. O aprendizado de novas tarefas (eventos cognitivos) tende a ser mais fácil se essas são de início apresentadas de forma icônica, enquanto que a prática de conhecimentos já adquiridos (eventos de reconhecimento) traz provavelmente uma forma de ensino mais simbólica, convencionada. A respeito do processo segundo o qual a criança adquire o ferramental sógnico de que necessita para interagir em seu meio sociocultural, Cunningham (2001) comenta:

Uma criança nasce em um meio sociocultural que funciona com base em alguns processos de semiose organizados socialmente. À medida que a criança age nesse contexto, ela é exposta aos meios pelos quais a comunidade media suas atividades. As pessoas que cuidam da criança se valem de ferramentas e signos para interagir com ela e intuitivamente coordenar essas ações com o

comportamento da criança. A criança então se apropria dessas ferramentas e signos para adquirir influência e controle sobre suas interações sociais, e, por conseguinte, transforma sua habilidade de influenciar suas próprias ações no âmbito de suas crescentes esferas de ação. Incorporar meios de mediação não só facilita as ações que poderiam ocorrer sem esses recursos, mas sim, como notou Vygotsky, “ao ser incluída no processo de comportamento, a ferramenta psicológica altera todo o fluxo e toda a estrutura das funções mentais (...)”

Em razão dos conceitos discutidos até aqui, e dado o nosso escopo, abordaremos o aprendizado de uma L2 em termos de um processo ecológico de interação de um aprendiz com um ambiente (que para nossos propósitos será a interface computacional do ambiente de ensino) com o objetivo de buscar um sentido nos signos ali apresentados e no código que eles se propõem a transmitir, representado pela L2. O aprendizado desse código é entendido como uma adaptação dos esquemas cerebrais, tendo por objetivo criar automatismos que habilitem uma proficiência no uso da L2. Isso então significa que o ambiente de ensino deve oferecer uma diversidade de conteúdos e estímulos, em contextos que possibilitem relacionamentos e gêneros tais como os necessários à aquisição da L1. Isso se refere, por exemplo, às nuances semânticas de cada lexema, às suas interdependências em relação a outros lexemas e a aspectos pragmáticos do uso de cada vocábulo ou expressão idiomática. Esses conhecimentos, que nos falantes da língua já se encontram internalizados, devem ser expostos e exemplificados aos alunos em contextos válidos, variados e tantas vezes quanto necessário para com isso permitir sua assimilação como segunda língua (L2).

3.6 Fatores relevantes no ensino de uma L2

Tendo em conta tudo o que foi discutido até este ponto, é necessário agora começar a delinear as características que o ambiente de ensino deverá ter para poder contribuir para o processo de aprendizado de uma L2 de forma tão efetiva quanto possível. Um passo para entender esses fatores seria examinar todos os meios e recurso já propostos e testados no ensino presencial assistido por tutores humanos. Uma revisão dos métodos e estratégias para esse contexto é feita no Apêndice 1. Embora essa revisão permita extrair linhas de reflexão sobre as diversas implicações do ensino de uma L2, Szanser (1992: p.5) salienta que não existe um método absoluto, válido sempre e em todo lugar e para todo e qualquer aluno e acrescenta que, segundo alguns autores, só existem dois métodos verdadeiros de ensino de línguas, o de tradução da gramática e o direto, sendo todos os demais uma combinação de princípios e procedimentos desses dois. E tendo em vista que o objetivo deste estudo não é propor um novo método mas contribuir para ambientes mais eficazes de ensino de línguas a distância, já é possível sintetizar os principais requisitos do ambiente.

Segundo Szanser (1992: p.5), qualquer que seja o método é necessário escolher um conceito e os meios, técnicas e procedimentos adequados a ele, dependendo dos aspectos supracitados, isto é, contexto e público-alvo.

Assim, tendo em conta os meios, as técnicas e os procedimentos a empregar, o método dependerá de um grande número de fatores:

- Objetivos: o que se quer atingir com o ensino
- Objeto do ensino: a(s) L2(s) que se quer ensinar
- Características dos alunos: L1 e outros conhecimentos prévios
- Percurso didático: meios, procedimentos, técnicas e formas de ensino adotados
- Situação concreta
- Conhecimentos linguísticos gerais prévios

Para o **contexto de ensino** a que se propõe o ambiente aqui descrito, o objetivo é o ensino da gramática e do léxico básicos da L2, sendo léxico e gramática interdependentes quando se almeja um domínio ativo da L2. Como discutido, o **objeto do ensino** no protótipo será o idioma esperanto, sendo porém possível a extensão do ambiente para qualquer outra língua. Por se tratar de um curso internacional, isto é, um curso que não se baseia no uso da L1 para ensinar a L2, e por ser ele baseado na Web, admite-se que as **características dos alunos** serão marcadas por variados perfis e origens culturais. Já o **percurso didático** proposto envolve a exposição do aluno a cenas e a interação com seus elementos. No que se refere à **situação concreta de ensino**, este será mediado por um sistema tutor inteligente e gerido por tutores humanos. Por fim, considera-se que os **conhecimentos linguísticos prévios** dos alunos incluem o domínio de uma L1 qualquer e possíveis noções de outras L2.

No que diz respeito à condição do ensino, deve-se considerar se esse é:

- Presencial ou a distância
- Obrigatório ou opcional

A **condição de ensino** prevista para o ambiente é *a distância e opcional*, o que não exclui seu eventual uso em outras situações, caso, por exemplo, a L2 seja de ensino obrigatório ou o uso do ambiente seja um complemento a práticas presenciais de ensino da L2.

No que toca à teoria de ensino empregada, essa pode ser, por exemplo:

- Construtivista
- Instrutivista

A **teoria de ensino** em que o ambiente proposto se apoia é o construtivismo, pois ele parte do princípio de que o desenvolvimento da habilidade linguística do aprendiz em relação à L2 será determinado pelas interações entre o aprendiz e o meio, isto é, entre o aluno e o ambiente, e que ambas as partes terão autonomia para responder às ações da outra parte.

No que diz respeito ao material instrucional, este pode enfatizar:

- Vocabulário
- Regras de gramática
- Exercícios

O **material instrucional** do curso dará ênfase à aquisição do vocabulário, mas as sentenças formadas a partir dos lexemas já apresentados possibilitarão também ao aluno aprender elementos da gramática da L2. Exercícios serão usados para fixar e avaliar o aprendizado. Dos possíveis tipos de exercício (compreensão oral, compreensão da leitura, de escrita, de automatização, de fixação consciente, de gramática e de fala produtiva e de conversação) os que mais se adequam ao ambiente em EaD são os seis primeiros, restando aos dois últimos a possibilidade de uso em cursos presenciais que usem o ambiente como complemento.

No que trata do sequenciamento do material instrucional, podemos ter:

- Do mais fácil ao mais difícil
- Do mais simples ao mais complexo
- Do mais concreto ao mais abstrato
- Do mais frequente ao mais raro
- Do mais conhecido ao menos conhecido
- Do mais comum ao menos comum

O ambiente de ensino foi projetado para combinar flexibilidade de baixo nível (*bottom up*) e adaptabilidade de alto nível (*top down*), e esta última significa que os tutores humanos poderão definir a(s) regra(s) que vão ditar a dinâmica de **sequenciamento do material**, que pode seguir uma das regras listadas acima, ou a combinação de algumas delas.

Quanto às alternativas para o ensino do léxico, temos:

- Demonstração real/icônica
- Descrição sem tradução
- Descrição com tradução
- Apoio do contexto e da situação
- Apoio de tópicos de gramática
- Por meio de sinônimos
- Por meio de palavras de sentido parecido
- Apresentação da definição na L2
- Uso de provérbios, canções, versos, etc.

Entre as alternativas para **ensino de léxico** as que se adequam melhor aos objetivos, objetos e modelos de interação do ambiente de ensino são as de demonstração icônica, com as imagens de apoio possibilitando aos alunos a apreensão dos significados dos lexemas, mas também as alternativas que envolvem apoio do contexto e da situação, por meio de sinônimos, apresentação da definição na L2 e, em fases mais avançadas do aprendizado, uso de provérbios, canções, etc.

Por sua vez, os erros podem ser classificados segundo suas causas como:

- Por engano: distração ou displicência, não por má compreensão
- Conceitual: má compreensão de regras linguísticas
- Intralinguístico: generalizações indevidas de regras
- Interlinguístico: interferências da L1 ou de outras línguas
- Por esquecimento: falhas de memória

Os **tipos de erro** que os alunos cometem durante o aprendizado podem ter todas as causas listadas acima, mas com exceção dos erros por engano, de modo geral, todos os demais podem ser classificados como devidos a fatores intralinguísticos, por particularidades da L2 (são mais prováveis em línguas com estruturas irregulares), ou a fatores interlinguísticos, nos casos em que a L1 interfere seja na entendimento conceitual dos recortes semânticos da L2, seja favorecendo os esquecimentos, nos casos em que as formas na L1 sejam muito diferentes das da L2 (sendo mais prováveis entre línguas não aparentadas), eliminando com isso qualquer facilidade mnemônica que poderia decorrer de semelhanças etimológicas.

Todos esses fatores referentes ao ambiente de ensino estão sintetizados na Tabela 3.1:

Tabela 3.1: Fatores relevantes ao ambiente de ensino aqui proposto

Fator	Valor
Objetivos	Aprendizado da gramática e do léxico básicos da L2
Objeto de ensino	Protótipo: esperanto; Extensão: qualquer língua
Características dos alunos	Origens variadas (curso internacional baseado na Web)
Percurso didático	Exposição a cenas e interação com seus elementos
Situação concreta de ensino	Mediado por tutor inteligente e gerido por tutor humano
Vivência linguística prévia	Domínio de uma L1 e possíveis noções de outra L2
Condição do ensino	Opcional e a distância
Teoria do ensino	Construtivista (resultante da interação com as cenas)
Material instrucional	Ênfase na aquisição do vocabulário
Programa	Léxico fundamental
Plano de ensino	Tempo do curso definido pelo aluno
Sequenciamento do material	Automático pelo ITS e configurável pelo tutor humano
Ensino do léxico	Demonstração icônica, apoio do contexto e da situação, por meio de sinônimos, apresentação da definição na L2, uso de provérbios, canções, etc.
Exercícios	Compreensão da leitura e capacidade de escrita.
Tipos de erro	Intralinguísticos e interlinguísticos

No que concerne às características do aluno, as variáveis incluem:

- Idade
- Sexo
- Personalidade
- Orientação política
- Tipo de memória (auditiva, visual, motora)
- Técnica de aprendizado (raciocinar, imitar, criar, traduzir, escutar, escrever)
- Conhecimentos gerais (ensino fundamental, médio ou superior)
- Conhecimentos linguísticos (uma ou mais línguas)
- Interesses (temas culturais, brincadeiras, distração, música, experiências de vida, atualidades sociais ou políticas, turismo, etc.)

Características adicionais de alunos de curso baseado na Web

- Nome
- Endereço IP
- Data de início
- Último acesso

Diante de todos esses aspectos, fica claro que a arquitetura de ensino ideal deve ser flexível a fim de tratar de modo individual os alunos e suas necessidades específicas, os elementos da L2, com suas inter-relações e particularidades, e os possíveis métodos de ensino, com suas táticas e recursos específicos.

Tendo em vista que o objetivo central do estudo é propor e especificar um ambiente computacional de ensino de línguas com as características resumidas na Tabela 3.1, foi necessário primeiramente discutir, no Capítulo 2, os principais conceitos da área de ensino assistido por computador, em especial o ensino de línguas. Em seguida, no Capítulo 3, foi necessário explorar o tema da aquisição e do aprendizado de uma língua, para entendermos quais são os requisitos de um ambiente de ensino de línguas. Mas considerando-se que o ambiente de ensino proposto fará uso intenso de imagens, o Capítulo 4 tratará do processo de interpretação de imagens, a fim de compreendermos melhor como os alunos serão capazes de inferir a semântica dos lexemas da L2 tendo por referência as imagens de apoio. Em seguida, no Capítulo 5, será finalmente descrita a arquitetura do ambiente de ensino, tendo como referência todos os aspectos já tratados até aquele ponto.



“– Esse país principia justamente ali onde o ar começa a zumbir. Os sons espalhados pelo ar, e que são representados por letras, fundem-se logo adiante em SÍLABAS, e essas Sílabas formam PALAVRAS – as tais palavras que constituem a população da cidade onde vamos.” (Rinoceronte Quindim, em Emília no País da Gramática)

Capítulo 4

O processo de interpretação de imagens

Como visto no Capítulo 3, a visão tem um papel muito importante no processo de aprendizado de uma língua, pois ela facilita a cognição social, capacidade que as crianças demonstram de perceber as intenções e o foco de atenção dos adultos e assim de perceber a que eles se referem quando usam determinada palavra. Isso, somado a outras capacidades perceptivas da criança, permite a ela estabelecer os relacionamentos entre nomes e objetos.

A visão é também fundamental no caso do ambiente computacional de ensino aqui tratado, pois é em extensa medida por meio da percepção que os alunos obterão as pistas visuais necessárias às inferências nas quais se baseará o aprendizado da L2. É portanto necessário entender os processos de percepção e de interpretação de imagens, para assim analisar como, no contexto do ambiente de ensino aqui proposto, os alunos interpretarão as cenas apresentadas na interface do curso e se valerão delas no aprendizado da L2. Este capítulo discute então a interpretação de imagens e sua interação com conteúdos textuais.

4.1 Considerações iniciais sobre a percepção de imagens

A percepção visual é um processo muito mais complexo do que em geral se supõe. Não se assemelha ao mero registro de uma imagem, como quando tiramos uma foto com uma câmera digital. A esse respeito, Baars (2010, p.158) comenta que a câmera é capaz de capturar a imagem projetada do ambiente e armazená-la como uma sequência de pixels, mas não é capaz de fazer muito mais com a imagem, como reconhecer o que ela retrata, que objetos ela mostra e onde eles estão. Baars acrescenta que se quisermos saber o que aparece na foto o modo mais fácil é pedir a alguém para olhar para ela e interpretá-la com seu cérebro. E a percepção visual é o que ocorre após a imagem atingir os olhos: ela forma um padrão de ativação na rede de receptores dentro do olho, e o padrão detalhado é analisado pelos centros visuais do cérebro, desse modo revelando o que é, e onde está.

Muito embora a área de reconhecimento e recuperação automáticos de imagens tenha experimentado notáveis avanços na última década, com um aumento tanto no número de conceitos identificáveis, quanto no índice de acertos na recuperação de imagens a partir de buscas textuais, em geral esses sistemas estão muito longe de reproduzir toda a riqueza da interpretação humana. Uma das capacidades que faltam aos sistemas automáticos é a de correlacionar os diversos elementos presentes em uma imagem com conhecimentos que o observador tenha do mundo real a respeito daquele ser ou objeto. Assim, a identificação de uma flor já leva o sistema perceptual a estar atento a elementos como folhas, caules, etc.

Segundo Baars (2010, p.159), os estudos sobre a percepção humana e neurociência sugerem que há muitos níveis de percepção. No nível mais fundamental, o cérebro humano processa atributos visuais básicos como cor, orientação, movimento, textura e profundidade estereoscópica. Muitos neurônios nas áreas visuais primárias são altamente especializados em determinados atributos - alguns disparam intensamente em resposta a uma linha em um certo ângulo, ou a uma cor específica, ou a uma determinada direção de movimento. Ainda

segundo Baars, de algum modo o cérebro é capaz de organizar esses atributos elementares ou básicos em grupos perceptuais organizados, ou, como defendido pela psicologia Gestalt, a percepção não pode ser entendida simplesmente pelo estudo dos elementos básicos da percepção (Wertheimer, 1912; Koffka, 1935), e o todo é maior que a soma das partes. Algumas das leis de agrupamentos visuais da Gestalt são as de similaridade, proximidade, continuidade, destino comum. Elas sugerem que elementos que mais se assemelham em cor ou formato tendem a ser percebidos como um grupo, e se eles são arranjados de modo a ficarem mais próximos em colunas ou em linha eles são percebidos como agrupamentos de linhas ou de colunas. Isso é ilustrado nas Figuras 4.1 a 4.3:

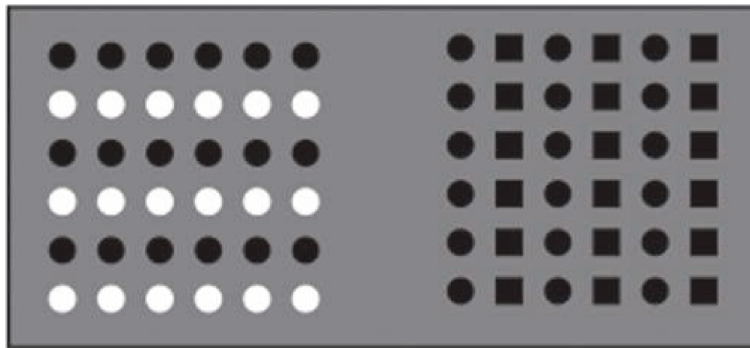


Figura 4.1: Agrupamentos por similaridade (Baars, 2010)

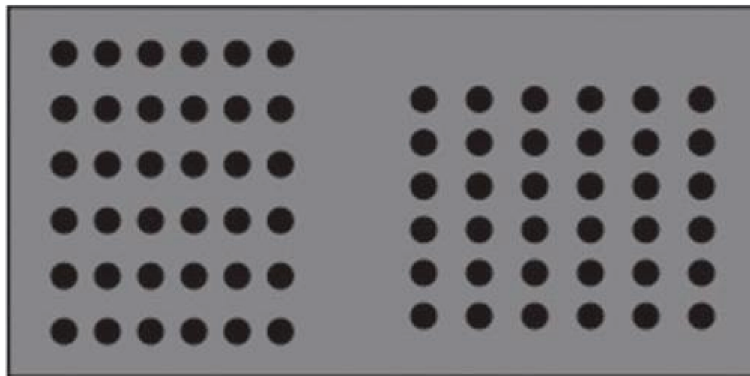


Figura 4.2: Agrupamentos por proximidade (Baars, 2010)

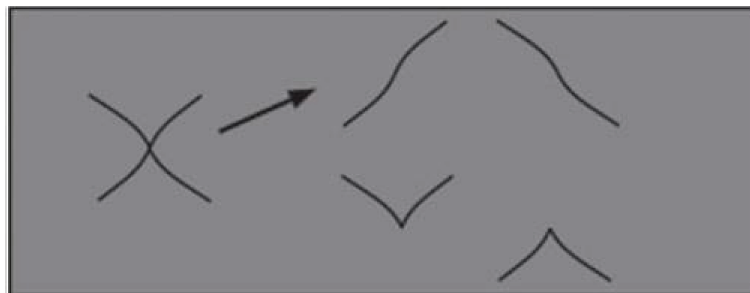


Figura 4.3: Agrupamentos por continuidade (Baars, 2010)

Os agrupamentos visuais são importantes na percepção de quais características pertencem a um possível objeto, ajudando a distinguir um objeto do fundo da imagem.

Baars conclui lembrando que somos capazes de perceber a forma de objetos inteiros e associar essas representações de formas aos objetos que conhecemos de nossa experiência prévia. A percepção de um objeto pelo cérebro passa por muitos estágios de processamento visual, do processamento dos atributos básicos do objeto, à combinação desses elementos em grupos coerentes, para, por fim, descobrir como esses elementos produzem uma forma organizada coerente que pode então ser comparada às representações guardadas na memória a fim de achar a correta. Mas o grande desafio do reconhecimento de objetos reside no fato de que existem milhares de objetos em nosso ambiente e de que a projeção bidimensional de qualquer objeto projetada no fundo dos olhos varia dependendo do ponto de vista, da iluminação, das condições de visibilidade. Baars salienta que de algum modo o cérebro deve abstrair as propriedades estáveis e invariantes de um objeto e desconsiderar todos os modos superficiais como a imagem 2D de um objeto pode variar (Baars, 2010, p.160).

Ciências cognitivas e psicologia definem a percepção como o processo de aquisição, interpretação, seleção e organização de informações sensoriais. Mas, apesar dessa definição comum, as teorias para explicar os processos subjacentes à percepção são inúmeras, algumas com ênfase mais biológica, outras mais psicológicas e outras mais filosóficas.

Psicólogos cognitivos sustentam que, ao nos movermos no mundo, criamos um modelo de como ele funciona. As sensações que captamos do mundo são mapeadas em perceptos, mas estes são provisórios, estando sujeitos a posteriores correções à medida que adquirirmos novas informações. Há ainda debate sobre se a percepção se caracteriza pela entrada de dados sensoriais não organizados, ou se de informações que representam objetos do mundo. Mas a Teoria do Percepto postula que na percepção comum somos apresentados a “perceptos de objetos” e não a meros dados sensoriais. Quando necessário podemos até nos referir mais especificamente a “perceptos de gato”, “perceptos de montanha”, etc. Até mesmo as qualidades das quais temos consciência na percepção quase sempre nos são apresentadas como qualidades de objetos físicos: liquidez da água, frio do gelo, solidez das rochas. De certo modo somos conscientes de um objeto físico que nos é apresentado completamente vestido de qualidades sensoriais. Essas qualidades se apresentam como qualidades do objeto. Em outras palavras, o objeto sensorialmente vestido é o único conteúdo sensorial de consciência durante a percepção comum, e é justamente a observação de objetos físicos que se pode chamar de “percepção”. Objetos físicos são percebidos mas nunca são objeto de “consciência direta”.

O procedimento que é usualmente visto como a confirmação de uma crença sobre um objeto físico envolve várias operações, por exemplo a manipulação do objeto, mas não envolve a operação de redução introspectiva. O procedimento usual é aquele de especificar o não especificado. Nós olhamos para um tomate, por exemplo, a partir de diferentes pontos de vista, o giramos em nossas mãos, o apertamos, etc., e dessa forma produzimos uma série de estados perceptivos. O objeto físico ostensivo permanece um tomate durante todo o processo, mas se torna progressivamente mais e mais especificado, e assim, ao mesmo tempo nos tornamos mais e mais convencidos de que nossa crença se tornou justificada.

4.2 A interpretação de imagens segundo a Semiótica de Peirce

Visto que neste trabalho adotamos a Semiótica Peirceana como uma das referências para modelar o aprendizado, é necessário investigar se Peirce propôs alguma explicação baseada signos e em inferências para explicar a percepção. Segundo Rösler (2009), o fato de existir interpretação é uma prova de que a percepção tem algo a ver com signos. Ilusões são um efeito de um processo de signos; nós não percebemos coisas diretamente, mas signos que pensamos representar a realidade. Quando essa representação dá errado, nos damos conta de que os objetos que pensamos ver são somente signos (de uma realidade que difere de nossa percepção) interpretados de maneira errada.

A emergência do código de percepção segundo Peirce

Assim, ainda segundo Rösler, a percepção é um processo de signos em que uma imagem toma um significado. O fato de que a imagem é imediatamente conectada a esse significado demonstra que deve haver um processo que governa a atribuição de significados às imagens, uma regra de conexão entre certas imagens – signos – com certos significados. O percepto – que é realmente o que percebemos – deve ser diferente do juízo perceptivo, como provam as ilusões. Enquanto o percepto é concreto, o juízo perceptivo deve ser abstrato. O juízo perceptivo sempre generaliza um percepto ao compará-lo com outros para dizer o que têm em comum. Nas palavras do próprio Peirce:

Daí vem então que nossos juízos perceptivos são as primeiras premissas de todo nosso raciocínio e que eles não podem ser postos em dúvida. (C.P. 5.116)

A questão agora é como essas duas partes da percepção são reunidas. A resposta de Peirce é que os juízos perceptivos devem ser vistos como “um caso extremo de inferência abdutiva”. (C.P. 5.181) A ideia de que a percepção tem algo a ver com a abdução veio a Peirce quando ele analisou as ilusões de ótica, como a escada de Schröder (ver Figura 4.4). Em um momento vemos os degraus a partir de cima, e no momento seguinte a partir de baixo, mas eles continuam sendo o mesmo objeto. A abdução nesse caso consiste em inferir duas hipóteses diferentes a respeito do ponto de vista. Dado que há duas maneiras de ver a escada, cada uma dessas maneiras só pode ser uma hipótese. Isso, segundo Rösler, convenceu Peirce do papel da interpretação no processo de percepção.

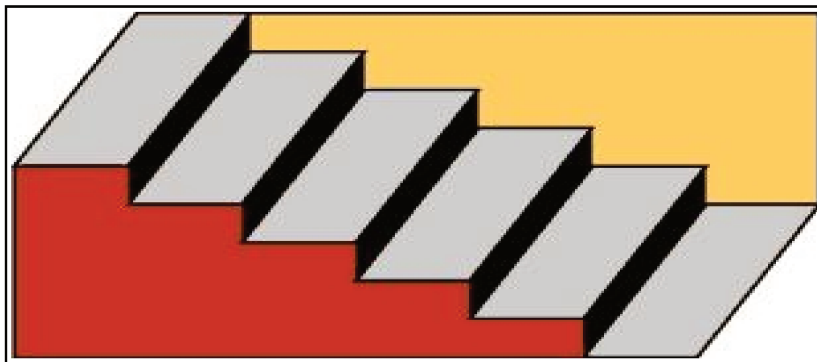


Figura 4.4: Escada de Schröder – exemplo de ilusão de ótica

Mas a abdução, no caso da percepção, difere daquela em uma pesquisa científica. Por isso Peirce fala dos juízos perceptivos como um caso limítrofe (C.P. 5.186) de abdução. Há uma diferença que consiste no fato de que não se pode negar um juízo perceptivo.

A percepção, segundo Peirce, é então a interpretação imediata de uma imagem, um percepto, em um juízo perceptivo pelo modo lógico da abdução, a proposição de uma hipótese que explique o fato que é o percepto, fazendo-o assim inteligível para nós. Temos que entender esses juízos como o limite último do raciocínio e do conhecimento. Eles são o nosso ponto de partida da cognição.

Como explicar as ilusões? O juízo perceptivo é a conclusão de uma abdução, que é uma forma muito fraca de inferência lógica, baseado em uma hipótese, e assim falível.

Segundo Peirce, se cabe algum questionamento em relação ao juízo perceptivo, esse questionamento se limitaria a refazer o processo para ver se, com atenção redobrada, obtemos o mesmo resultado. Mas ao fazê-lo o percepto supostamente não é o mesmo que o anterior (C.P. 5.115)

Para Rösler, um código estável é então possível porque pensamos na natureza e na experiência como algo regular e sem sobressaltos. Cada conclusão abductiva tem que ser corroborada em posteriores experiências. Quanto mais determinados perceptos similares causam o mesmo juízo perceptivo, tanto mais certeza adquirimos sobre nossas conclusões abductivas na percepção. Quanto mais experiências validam os juízos perceptivos, mais estável se torna o nosso código de percepção.

O *percipuum* e a questão dos fundamentos

Para Peirce, fatos perceptivos em sua essência mais primordial emergem não só da mente, nem só da realidade dinâmica do universo, mas da interação dos dois que constitui a experiência (Rosenthal, 2001). O que nos habilita a reconhecer objetos que já tenham sido percebidos em outras ocasiões anteriores é a **recorrência** de aspectos da experiência perceptiva que se assemelham à experiência perceptiva anterior. Se esses aspectos são suficientes para validar a crença, o objeto é reconhecido. A essas lembranças de perceptos anteriores Peirce chamou de *poneceptos*.

De acordo com (Rosenthal, 2001), Peirce considerava o *ponecepto* uma abstração para fins de análise, e que uma síntese de *poneceptos* (perceptos semelhantes passados), reunidos para o reconhecimento de um percepto presente como um “critério na mente” bastante primitivo, é um *ponecippuum*. Assim, um *ponecippuum*, como precedente lógico de qualquer percepto presente, é necessário à interpretação do percepto presente através do juízo perceptivo. Ele fornece o critério sintético para apreender uma apresentação como um conteúdo repetível que pode ativar um hábito. Na verdade, é a própria apreensão pelo *ponecippuum* que torna o conteúdo repetível. O *ponecippuum* serviria como o veículo entre impressões em transformação e sensações reconhecíveis e diferenciadas que ajudam a constituir o interpretante imediato.

O que instiga atividade no curso presente de experiência são perspectivas de objetos, fatos aparentes, não são aparências, mas objetos aparentes, mas isso só é possível porque a estrutura de significado como hábito inclui o *ponecippuum*, ou cerne de conteúdo sensorial generalizado (Rosenthal, 2001).

O percepto presente, interpretado à luz do *ponecipuum*, é o *percipuum* no sentido estrito, resultado do juízo perceptivo, e que fornece “conteúdo repetível” que serve para ativar um hábito. O juízo perceptivo em seu sentido estrito é a hipótese abdutiva primitiva de uma repetição presente de conteúdo experiencial passado, e o conteúdo de fato se torna uma repetição de conteúdos previamente experienciados assim como o juízo perceptivo o assemelha àqueles conteúdos no processo de reconhecimento abdutivo. O percepto por si mesmo, então, não afirma nada, enquanto que o juízo perceptivo afirma a presença do percepto como um conteúdo reconhecido (Rosenthal, 2001).

Juízos perceptivos no sentido estrito têm um “elemento hipotético”. Para Peirce, tanto concepções quanto sensações envolvem inferências hipotéticas, pois o julgamento é uma hipótese de que um conteúdo é “o mesmo que” o que foi visto antes ou “parece ser”. A não ser que seja dado algum espaço para dúvida na situação perceptiva, juízos perceptivos devem ser tomadas como indubitáveis, pois embora eminentemente falíveis não podem ser seriamente questionados, visto que estão sujeitos ao teste de experiências futuras. Assim, ao mesmo tempo que não podemos ter um controle crítico sobre o juízo, podemos criticar seus resultados e concluir, com base em futuras experiências, que ele é falso (Rosenthal, 2001).

Segundo Santaella e Nöth (2005: 85), a forma como exploramos imagens é feita não de modo global, mas por rápidas fixações sucessivas que se limitam estritamente às partes da imagem mais providas de informação. Isso determina o nível fisiológico da percepção do tempo em imagens, que portanto não é instantânea. Já o tempo biológico da percepção baseia-se, segundo eles, numa capacidade inata a todos mamíferos de “apreender o sentido imediato de padrões de espaço e tempo muito complicados para um processo inteligente consciente”. Isso nos habilita a usar os sentidos inconscientemente para mensurar intervalos de tempo, distâncias e suas diversas relações. E a essa capacidade biológica os seres humanos adicionam a habilidade de perceber dentro de esquemas lógicos, de criar símbolos para objetos, o espaço e o tempo (Santaella e Nöth, 2005: 85).

Santaella e Nöth (2005: 86) comentam que aquilo que é comumente chamado de estímulo corresponde a um *percepto*, ou, nos termos definidos por Peirce, o elemento de compulsão e insistência na percepção e que “corresponde à teimosia com que o percepto (...), apresentando-se à porta dos sentidos, insiste na sua singularidade, compelindo-nos a atentar para ele”. Tão logo o percepto, ou feixe de perceptos, atinge os nossos sentidos, ele é imediatamente convertido em *percipuum*, ou seja, o modo como o percepto se apresenta àquele que percebe, ao ser filtrado pelos sentidos, traduzido na forma e de acordo com os limites e determinações que nossos sensores lhe impõem.

Ainda segundo Santaella e Nöth (2005: 85), recebemos continuamente uma chuva de perceptos que, como *percipua*, fluem dentro de nós, são imediatamente colhidos e absorvidos nas malhas dos nossos esquemas interpretativos, convertendo-se assim em juízo perceptivo, isto é, sendo identificados e reconhecidos. Além desse *percipuum*, há outros elementos temporais que Peirce denominou de *antecipuum* e *ponecipuum*. O primeiro é o ingrediente antecipatório da percepção, enquanto o segundo é o ingrediente de memória infinitesimal que participa em todo ato perceptivo. Esses ingredientes são essenciais, pois sem memória e antecipação não pode haver reconhecimento e identificação, os quais por sua vez são a base da percepção (Santaella & Nöth, 2005: 87).

O reconhecimento icônico

A produção de um juízo perceptivo é condição necessária mas não suficiente para que o observador seja capaz de interpretar um signo. Mesmo no caso dos ícones, que são as peças básicas na construção de um sistema simbólico, o processo interpretativo requer algo mais que o mero juízo perceptivo. Para tentar explicar mais profundamente a maneira como um ícone é interpretado, Sowa (1990) propôs uma teoria para o reconhecimento icônico:

- 1) Informações sensoriais entram no cérebro na forma icônica. Um ícone não precisa ter a riqueza de uma fotografia, mas o mapeamento entre uma imagem projetada na retina e um ícone no cérebro deve preservar relações espaciais essenciais (...).
- 2) As pessoas são capazes de criar e transformar “imagens mentais” que parecem ser da mesma natureza de ícones sensoriais.
- 3) Fragmentos de ícones, denominados perceptos (outros autores preferem o termo *iconemas*), são armazenados no cérebro com o propósito de interpretar ícones recebidos pelos sentidos ou de criar novos ícones no imaginário mental. O processo de percepção compara um ícone com uma construção feita a partir de um ou mais perceptos que se encaixam em alguma parte do ícone.
- 4) Alguns perceptos estão associados a conceitos, os quais são unidades simbólicas adequadas para a linguagem e a lógica. Um ícone que encaixa em uma construção de perceptos pode ser interpretado por um grafo dos conceitos associados com aqueles perceptos.
- 5) Ícones usados no imaginário mental podem ser gerados por grafos conceituais: perceptos são ajuntados de acordo com os relacionamentos descritos pelos grafos.
- 6) Esse mapeamento não é unívoco, um conceito pode ser associado a diferentes perceptos (para diferentes tipos daquele conceito, como por exemplo, diferentes tipos de gato) e a conceitos abstratos, como Justiça, não correspondem perceptos.
- 7) A interpretação de um ícone não é única: diferentes pessoas podem ter estoques diferentes de perceptos que podem combinar melhor ou pior; uma mesma pessoa pode ter diferentes perceptos para encaixar parcial ou totalmente um mesmo ícone.

Segundo Sowa, os três primeiros pontos da hipótese acima nada mais são do que uma reformulação das propostas feitas por Neisser (1967) as quais têm sido corroboradas por várias décadas de provas psicológicas. O item 4 é necessário para explicar como as pessoas são capazes de olhar para uma cena e descrevê-la por meio de proposições feitas com base em símbolos discretos. O item 5 considera o modo como proposições são capazes de estimular o imaginário mental. Por fim, os itens 6 e 7 levam em conta os diferentes pontos de vista, as mudanças na percepção como resultado da experiência, e o fato de que muitos conceitos têm relações remotas com qualquer coisa que possua imagem.

É esclarecedora, para os propósitos do ambiente aqui proposto, a hipótese de que os ícones apresentam relacionamentos espaciais essenciais e que em razão disso podem ser reconhecidos por qualquer indivíduo que já possua em seu cérebro um certo repertório de perceptos. Esses perceptos, por sua vez, teriam sido colecionados (na forma de *poneceptos*) à medida que os ícones correspondentes foram sendo interpretados pelo sujeito a partir de imagens recebidas de seu contexto sociocultural.

O fato de que ícones são reconhecidos por crianças antes mesmo de elas aprenderem os símbolos (palavras) correspondentes na L1 parece reforçar a noção de que a interpretação icônica é uma capacidade mais básica, inata e universal do que o domínio dos símbolos linguísticos, cuja interpretação é própria de cada contexto cultural. Nesse sentido, atende-se a condição inicial de que o público-alvo do ILTS já terá vivenciado experiências visuais suficientes e terá um *ponecipuum* rico o bastante para que as imagens utilizadas na interface façam sentido, uma vez que estas remetem a objetos comuns em sociedades modernas, como “carros”, “casas”, “ruas”, etc. Isso assegura a inteligibilidade das cenas do curso, a despeito das diferentes origens culturais dos aprendizes.

Alguns autores propõem também uma escala de iconicidade segundo a qual o nível de fidelidade de um ícone em relação ao seu objeto representado varia de um grau máximo, correspondente a um holograma, e decresce gradualmente, passando por uma fotografia, uma ilustração, um desenho esquemático, um pictograma, um diagrama, um ideograma, uma onomatopeia, uma palavra arbitrária (símbolo linguístico), um símbolo abstrato, e por fim um símbolo tácito (uma mímica do objeto em questão). O ILTS aqui tratado usa em sua interface ícones baseados em fotos e desenhos (ou arranjos espaciais desses elementos a fim de estabelecer as inter-relações funcionais entre as partes). Isso assegura um nível aceitável de iconicidade para fundamentar a aquisição dos símbolos, dados os objetivos do ILTS.

4.3 A inter-relação de imagens visuais e mentais

Santaella e Nöth (2005: 15) comentam que o mundo das imagens divide-se em dois domínios, o das imagens como representações visuais (desenhos, pinturas, fotos, imagens cinematográficas, televisivas, holo e infográficas), em que são “objetos materiais, signos que representam o nosso meio ambiente visual”, e o domínio imaterial das imagens mentais (visões, fantasias, imaginações, esquemas, modelos e representações mentais). Os dois domínios são interdependentes, pois estão “inextricavelmente ligados já na sua gênese” e não há imagem como representação visual que não resulte de imagens mentais daquele que a produziu, nem imagem mental que não tenha alguma origem no mundo concreto dos objetos visuais. Os conceitos de signo e representação², objetos de estudo da Semiótica, unificam os dois domínios da imagem (seus lados perceptível e mental) em algo terceiro.

No que importa para um ambiente computacional de ensino de línguas, como o aqui tratado, a ideia central é usar imagens como apoio icônico para o ensino do léxico da L2. Dito de outra forma, considera-se que a apresentação de uma imagem na interface do curso irá produzir no aluno uma imagem mental que conduzirá o pensamento do aluno àquele objeto ou conceito. Uma pré-condição para isso é que ele já tenha experiências anteriores com aquele objeto, ou ao menos representações dele (via TV, livros, etc.). A interpretação da imagem de início produzirá uma experiência colateral que associa a ideia ao seu símbolo na L1, isto é, o nome do objeto na língua materna do aluno. Por experiência colateral a semiótica entende o efeito de uma imagem (por exemplo o ícone de um gato) evocar no observador o símbolo correspondente (a palavra “gato”), e vice-versa.

²A teoria da representação é também estudada pela ciência cognitiva.

Cada cena do curso conterá uma imagem (foto ou um desenho), signo icônico, espacialmente relacionada a palavras (signos simbólicos) na L2. Quando o ILTS apresenta ao aluno um símbolo correspondente na L2, a relação temporal e espacial (vínculo indicial) começa a estabelecer uma associação entre a ideia evocada pela imagem e o vocábulo na L2. E as recorrências dessa situação tenderão a fixar essa experiência colateral. É verdade que a coocorrência de símbolo e ícone não é uma garantia de que o aluno estabelecerá uma associação entre os dois signos, e gradativamente memorizará (aprenderá) aquele símbolo na L2. Mas embora não haja uma certeza absoluta desse resultado, essa situação se parece, como discutido no Capítulo 3, ao aprendizado das palavras da língua materna pela criança. Por um lado, no caso da L1 a relação entre a ocorrência do símbolo (quando pronunciado na presença da criança) e a presença do objeto no campo de visão da criança é ainda mais tênue do que na interface do curso. Como salientado por Bloom (2000), para conseguir criar esses vínculos a criança precisa ter um modelo da mente dos adultos para assim poder saber a que eles se referem quando falam, e também ser capaz de enxergar (ou imaginar) o objeto referenciado. Ou, nos termos usados pela teoria ecológica, os *tokens* linguísticos não são automaticamente interpretados, simplesmente em razão de sua disponibilidade no ambiente: eles devem ser percebidos como relevantes ao foco de atenção atual. E não obstante, como é sabido, mesmo nessas condições adversas a grande maioria das crianças consegue assimilar a língua materna sem problemas, embora, como vimos, inicialmente em ritmo relativamente mais lento que o observado no final da infância.

No caso da interface aqui proposta, além da maior clareza da relação entre símbolo e ícone dois outros fatores contribuirão para facilitar o aprendizado. O primeiro deles tem a ver com o fato de que os alunos sabem de antemão, e com bastante certeza, que o objetivo daquele ambiente é justamente ensinar a L2. Não é então preciso ter um aguçado modelo mental das intenções dos tutores (humano e computacional) para descobrir que o rótulo junto ao objeto “rinoceronte” possivelmente refere-se ao nome do objeto ou a alguma qualidade ou parte dele. Ou, nos termos usados por Bloom, saber que a palavra usada para o coelho branco provavelmente se refere à cor, à espécie ou ao nome daquele indivíduo. Uma segunda vantagem dos aprendizes em relação à criança é que aqueles já têm um sistema cognitivo desenvolvido, já tiveram contato imediato ou mediado com inúmeros objetos e conceitos, e, sobretudo, já assimilaram todos os rótulos que esses objetos e conceitos têm na L1. Com isso, eles estão habilitados a reconhecer as imagens de apoio usadas no ILTS e, igualmente importante, em muitos casos estarão habilitados a reconhecer, por similaridade, também os símbolos (vocábulos) da L2, em razão da semelhança com o correspondente elemento no léxico da L1. Em Ávila e Gudwin (2009) foi proposta uma métrica para essa “similaridade intersimbólica”. No caso exemplificado abaixo, o símbolo na L2 “rinocero” guarda semelhança com os equivalentes em várias outras línguas, como “rinoceronte” em português e espanhol, “rhinoceros” em inglês, “rhinocéros” em francês, “rinoceront” em catalão, só para citar algumas. Isso não se aplica a todos os símbolos da L2, nem a todas as L1, mas, em comparação com a total inexistência dessas coincidências no caso da aquisição da L1, as condições em que ocorrerá o aprendizado pelo ILTS são até bastante favoráveis, sem serem todavia uma garantia absoluta de êxito, visto que isso dependerá de fatores como a persistência, a assiduidade, a atenção e a motivação de cada aluno.

Mas voltando à interface do ILTS, esses arranjos visuais têm o objetivo de tirar proveito da capacidade de reconhecimento icônico dos usuários para com isso evocar conceitos e ideias (imagens mentais) e então ligá-las aos vocábulos correspondentes na L2, os quais, por essa mediação visual, tenderão a se associar àqueles conceitos. Essa estratégia não é original nem sequer nova, pois já é utilizada em muitos dicionários visuais e em outros ambientes de ensino. A contribuição original deste estudo está, como discutido nos próximos capítulos, na estratégia de fazer com que os ciclos de apresentação de vocábulos e ícones na interface sejam autonomamente definidos por um mecanismo de seleção de ação que, por sua vez, é guiado pela lógica interna da L2, pelos objetivos didáticos definidos, e pelo histórico de interações de cada aluno individualmente.

Após os ciclos de exibição de um dado vocábulo e a interação do aluno com a cena e seus elementos considera-se que o aprendizado terá sido bem sucedido se isso habilitar no aluno a capacidade de ter a experiência colateral daquele vocábulo. Em outras palavras, se ao se deparar novamente com aquele objeto (ou uma representação dele) ele for capaz de se lembrar do vocábulo correspondente na L2, considerar-se-á que ocorreu o aprendizado. Mas cabe lembrar que, como postulado pela teoria ecológica, “os indivíduos nunca realmente 'adquirem uma língua' no sentido de serem capazes de reproduzir o sistema como um todo [...]. Os indivíduos aprendem a reproduzir **elementos** do sistema”.

Na Figura 4.5 é ilustrada uma possível cena do ambiente de ensino. Um conceito ou tema principal, neste caso “rinoceronte”, aparece no centro da imagem, e é anotado o elemento lexical que corresponde àquele conceito na L2. Há uma clara relação indicial de proximidade entre o ícone e o símbolo, mas essa não é absoluta no sentido de dar ao aluno a certeza de que o rótulo se refere à espécie do animal ou se a alguma característica desse. No entanto, como discutido no decorrer deste trabalho, a interface mostrará diversos rótulos (anotações), não só para o animal como para partes de seu corpo e qualidades dele. É o contraste entre esses diversos rótulos durante a interação e as sentenças de exemplo que ajudarão o aluno a, pouco a pouco, inferir o valor semântico de cada um deles.

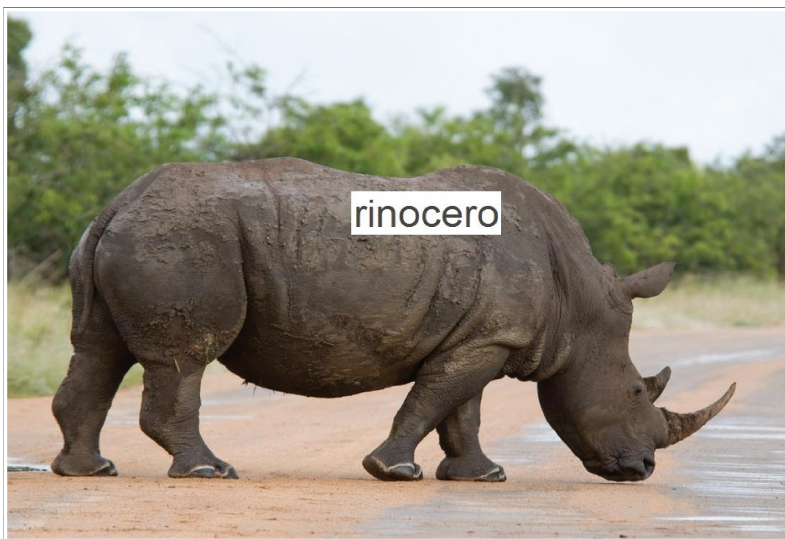


Figura 4.5: Exemplo de arranjo visual no curso

Visto que o especialista quer ensinar o símbolo da L2 para o objeto “rinoceronte”, o objetivo da cena é permitir uma sequência de ações semióticas:

- 1- O especialista usa uma imagem do objeto para exibição na interface do curso;
- 2- O especialista justapõe à imagem o símbolo correspondente na L2;
- 3- O aluno, ao fazer uso do ambiente de ensino, observa o arranjo visual;
- 4- O aluno reconhece o objeto iconicamente (pelas experiências visuais passadas);
- 5- Isso produz no aluno a imagem mental (conceito) daquele objeto;
- 6- O aluno reconhece os caracteres na cena e os interpreta como uma palavra;
- 7- O aluno, levando em conta as pistas indiciais (proximidade), associa a palavra à imagem mental (conceito) evocada pela imagem do objeto.

Esse é, em síntese, o percurso semiótico que leva o conhecimento do especialista até o aprendiz, pela mediação do ambiente de ensino, como ilustrado na Figura 4.6. Mas essa descrição sucinta não esgota a análise, visto que a inter-relação de textos e imagens tem sido motivo de debate ao longo da história da Semiótica, como será tratado na Seção 4.4.

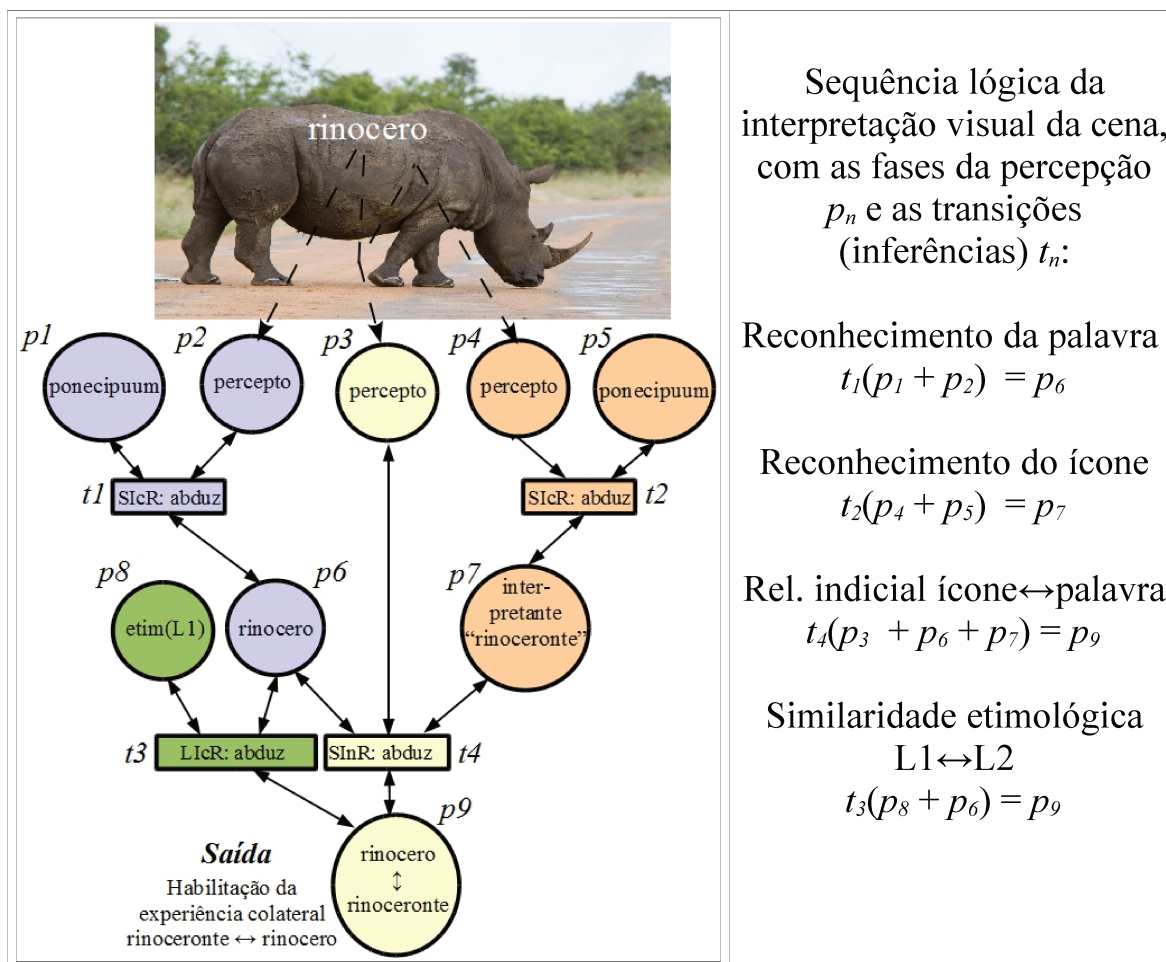


Figura 4.6: Sequência da interpretação visual da cena

4.4 A inter-relação de imagens e textos em uma cena

Dentro de uma visão logocêntrica sobre a dependência linguística da imagem, Barthes defendia que o entendimento de uma imagem é conduzido através da mediação da linguagem (Barthes, 1961). Já os partidários da tese da autonomia semiótica afirmam que os textos que acompanham uma imagem para comentá-la não provam a precedência da linguagem verbal sobre a visual (Santaella e Nöth, 2005: 45).

Segundo Santaella e Nöth (2005: 44), a diferenciação entre linguagem e imagens foi apresentada em pesquisas sobre suas eficiências respectivas, e cujos resultados sugerem que as imagens atuam mais intensamente no campo afetivo-relacional, enquanto na linguagem predominam efeitos cognitivos-conceituais (Janney e Arndt, 1993). Além disso, as imagens fomentam atenção e motivação, sendo então mais propícias à apresentação de informação espacial e a certos tipos de aprendizado (Weidenmann, 1988: 135-138), e sua eficácia emocional cresce com o seu grau de iconicidade (Reimund, 1993).

Santaella e Nöth comentam ainda que na visão do Gestalt, imagens são interpretadas como signos autônomos: as formas visuais são unidades de percepção independentes da linguagem, sendo percebidas em sua totalidade como formas, as quais são mais do que a soma de suas partes. A percepção é, assim, um processo construtivo (e não reprodutivo) determinado pelas leis da forma: fechamento, continuidade, proximidade (menor distância), igualdade e simetria (Metzger, 1975). Santaella e Nöth acrescentam ainda que, apesar de não possuírem um significado conceitual concreto, as formas podem, como invariantes visuais do campo visual, ser interpretadas como unidades semióticas autônomas, e nesse ponto haveria uma concordância entre a psicologia da forma e a interpretação estrutural-semiótica (Santaella e Nöth, 2005: 45). Esse valor signico das formas é defendido por Arnheim (1954: 65) nos seguintes termos: “Nenhum padrão visual existe somente em si mesmo. Ele sempre representa algo além da sua própria existência individual (...)”. Por fim, Santaella e Nöth salientam que a percepção da forma não é somente um processo de recepção, mas um processo de coordenação entre o percebido e formas já internalizadas, isto é, um processo semiótico tipo *token-type*.

Santaella e Nöth (2005: 55) comentam que, quando dispostos lado a lado, um texto e uma imagem permitem uma interpretação holística da mensagem (cf. Bardin 1975: 111), não se tratando nesse caso de uma mera adição de duas mensagens informativas diferentes. Ainda segundo aqueles autores, Barthes (1964c: 38-41) propôs duas possibilidades no que trata dos possíveis modos de inter-relação entre texto e imagem: *ancoragem* e *revezamento*. Na *ancoragem* o texto orienta a interpretação da imagem, ressaltando alguns aspectos e desprezando outros. No *revezamento*, texto e imagem estão numa relação complementar em que “a unidade da mensagem se realiza em um nível mais avançado”. Santaella e Nöth acrescentam que, embora em ambos os casos trate-se de referências indiciais entre palavra e imagem, no primeiro a referência se dá do texto para a imagem, enquanto no segundo a atenção do observador é dirigida do texto para a imagem e vice-versa.

Há ainda uma outra forma de referência indicial entre palavra e imagem, chamada de *denominação* ou *etiquetamento* (cf. Goodman 1968; Muckenhaupt 1986: 48), a qual corresponde àquelas situações em que a palavra designa a coisa ou pessoa mostrada na imagem (Santaella e Nöth, 2005).

Numa primeira análise a forma de referência indicial entre palavra e imagem que mais se aproxima da utilização feita no ambiente de ensino parece ser a de *etiquetamento*: imagens são justapostas a seus rótulos correspondentes na L2. Contudo, se aprofundarmos a análise um pouco mais, veremos que relações baseadas em *ancoragem* ou *revezamento* não estão totalmente ausentes. Por um lado, como discutido no Capítulo 3, o aprendizado de uma L2 em geral inclui o efeito das transferências positivas de conhecimentos da L1, por exemplo, no caso de símbolos cognatos nas duas línguas. Essa possibilidade é ainda mais efetiva no caso de uma língua de síntese como o esperanto, cujo léxico foi expressamente planejado no sentido da internacionalidade. Isso significa que, em muitas telas do curso, a interpretação dos elementos da cena vai fazer convergir duas linhas interpretativas: uma partindo do ícone (desenho), e outra partindo do símbolo (vocabulo), as quais se reforçarão mutuamente. Essas ocorrências sugerem relações do tipo *revezamento*. Também possíveis, embora menos prováveis, são os casos de *ancoragem*, sempre que em uma dada cena do curso as imagens não forem suficientemente claras para um aluno, mas ele puder recorrer à etimologia do símbolo adjacente para clarificar o conceito que a imagem tenta mediar.

Desse modo, fica claro que as interpretações das imagens e dos lexemas podem se apoiar reciprocamente na decomposição de uma cena. Mas cabe ainda discutir se o contexto cultural do aluno pode influir na sua capacidade ou sua maneira de interpretar imagens.

4.5 A influência de fatores culturais na interpretação de imagens

Na linguística, a hipótese Whorf-Sapir defende uma relatividade linguística extrema, segundo a qual aspectos da organização da linguagem de uma comunidade influenciam os hábitos perceptuais de seus membros, e até mesmo seus pensamentos. Por essa hipótese, não seria possível a existência de uma semântica de universais perceptuais. Mas, embora haja alguns dados para sustentar esse ponto de vista (Carroll & Casagrande, 1958), e não haja dúvida de que a linguagem modula a experiência do mundo, haveria ainda assim, segundo muitos autores, parca comprovação da extrema relatividade pretendida por essa visão, e algumas evidências claras em contrário (Hunt & Agnoli, 1991). Assim, segundo (Bloom, 2000), a afirmação Whorfiana de que as culturas têm números variáveis de nomes de cores para designar aspectos que têm relevância em seu contexto local, como, por exemplo, os diferentes termos para branco em culturas árticas, seria a conclusão de um raciocínio circular, em que se alega que os membros dessas culturas devem ter a percepção de distintos tipos dessa cor por terem eles diversos termos para branco, mas quando se questiona sobre como é possível saber se eles de fato têm uma percepção diferente no que diz respeito a tipos de branco, alega-se que sim, visto que eles têm vários termos para aquela cor... Nesse sentido, um dos estudos interculturais cujos resultados contradizem claramente a teoria Whorfiana trata justamente da terminologia das cores. No estudo de (Berlin & Kay, 1969), quando uma cartela de cores foi mostrada aos entrevistados, esses foram capazes de delimitar o mesmo repertório de cores básicas, independentemente dos diferentes rótulos linguísticos atribuídos por suas respectivas culturas. Os pesquisadores encontraram dessa maneira um total de onze categorias básicas de cor nas respostas dos respondentes, os quais eram provenientes das dezenas de culturas estudadas. Além disso, foi encontrada uma sequência evolutiva fixa que define a ordem em que as línguas acrescentam

novas categorias básicas de cores. A conclusão dos pesquisadores foi de que os referentes para os termos básicos de cor em qualquer língua parecem fundamentar-se em um conjunto de onze categorias perceptivas universais e, ao longo da história de uma dada língua, essas categorias são codificadas respeitando uma dada ordem fixa (Prince, 1993). Em outras palavras, diferentes línguas podem fazer mais ou menos distinções conforme necessidades ambientais ou culturais, mas não podem se sobrepor à base biológica da percepção de cores.

A existência de uma base não linguística, e até mesmo biológica, para a percepção de cores sugere que a percepção de imagens, essencial para este estudo, possa também estar ancorada em capacidades perceptivas inatas ou, ainda que adquiridas, resultantes de um processo de interação com o ambiente que não se subordina, em sua essência, ao contexto linguístico de cada indivíduo.

Outros estudos antropológicos e psicológicos mostraram a existência de uma inata capacidade de estabelecer rapidamente relações entre imagens e os objetos representados por elas (ou, em termos semióticos, entre ícones e objetos) como necessário à percepção (interpretação) pictórica. Em um estudo clássico (Hochberg e Brooks, 1962), um recém-nascido foi privado, até a idade de 19 meses, da exposição a todo e qualquer tipo de ilustração, tendo nesse período aprendido todo seu vocabulário por meio de objetos, sem receber qualquer treinamento em relação a significados ou conteúdos imagéticos. Apesar disso, quando lhe foi apresentada uma série de 21 fotografias e desenhos (de coisas e pessoas familiares ao seu ambiente), ele foi capaz de reconhecer os objetos representados. Essa demonstração da existência de habilidades inatas de percepção pictórica tem por consequência a invalidação de eventuais alegações de que sua ausência nessa ou naquela cultura ou em um dado indivíduo seria causada pelo “não aprendizado da linguagem imagética” (Prince, 1993).

Ainda em relação ao papel da cultura na percepção de imagens, (Prince, 1993) salienta que embora haja alguns indícios de variação das suscetibilidades culturais frente a ilusões visuais (como a de Muller-Lyer), talvez em razão dos variados ambientes físicos em que as distintas culturas estão imersas, isso não parece ter impacto nas habilidades básicas de percepção pictórica (Segall et al., 1966). Prince (1993) acrescenta que dados empíricos claramente sugerem que as habilidades de identificação pictórica não se desenvolvem em resposta a um longo período de exposição à significação (e o consequente aprendizado), como ocorre com as habilidades linguísticas, e isso se deve provavelmente ao fato de que as imagens mais realistas são isomórficas em relação às impressões visuais do mundo real, diferentemente do que se dá com os signos simbólicos, cuja relação com o representado é muito mais arbitrária. Isso faz com que as representações icônicas sejam processadas de modo muito mais imediato que as simbólicas, sobretudo por crianças, o que é comprovado em pesquisas (Calvert et al., 1982). E embora nenhum signo seja totalmente icônico, signos pictóricos apresentam similaridades estruturais em relação a seus referentes, o que traz para a percepção de imagens as implicações discutidas acima.

O efeito da iconicidade não se restringe todavia a desenhos e fotos. Como discutido na Seção 4.6, em muitas situações ele terá um papel considerável também nas situações de reconhecimento de palavras na L2 com base em sua similaridade com a palavra correlata na L1, tendo assim um efeito na facilidade de aprendizado e de memorização da palavra.

4.6 Iconicidade intersimbólica e iconicidade intersintática

Segundo Santaella e Nöth (2005: 142-143), a semiótica triádica de Peirce opõe-se radicalmente à tradição da semiologia estruturalista de propor oposições binárias, em que aquilo com qualidade sónica é oposto ao que não possui essa qualidade, em que o arbitrário é oposto ao icônico, o figurativo ao abstrato, etc., e propõe em lugar um pensamento triádico e no qual “as categorias descritivas devem ser entendidas como aspectos dos fenômenos, os quais podem estar presentes ao mesmo tempo em diferentes graus”.

Santaella e Nöth (2005: 145) comentam que para Peirce o ícone puro é somente uma possibilidade hipotética de um signo, não sendo um signo “genuíno” mas degenerado. Um signo genuíno que só se liga ao objeto por semelhança, por características comuns com ele, é denominado hipoícone (CP 2.276-277). Imagens figurativas, diagramas e metáforas são hipoicônicos em muitos aspectos, embora possam conter elementos indiciais e simbólicos.

Segundo Santaella e Nöth (2005: 146), a semiótica peirciana prevê, além do qualisigno icônico (signo puro, da primeiridade), duas outras possibilidades de iconicidade, nas quais o ícone está ligado às categorias da secundidade e da terceiridade: o sin-signo icônico, cujo veículo é singular, individual, e o legi-signo icônico, cujo veículo é uma regularidade.

A categoria de ícone é concebida por Peirce de forma mais geral e compreende também formas não visuais, por exemplo, acústicas ou mesmo táteis, olfativas ou também formas conceituais de semelhança sónica. (Santaella & Nöth, 2005: 38)

Segundo Santaella e Nöth a noção de ícone diagramático têm, na visão peirciana, importância fundamental no raciocínio (e linguagem) matemático e lógico: “O raciocínio deve estar especialmente relacionado com as formas que são os principais objetos do *insight* racional. Por isso mesmo, ícones são particularmente requisitados para o raciocínio” (CP 4.331). Os diagramas estão presentes em qualquer tipo de pensamento, isto é, com base em Peirce, todo pensamento é essencialmente diagramático. Aos ícones diagramáticos estão associados, segundo Santaella e Nöth, os padrões sintáticos, os quais assumem diferentes imagens diagramáticas conforme a língua de que se trate (Santaella e Nöth, 2005: 66). Essa noção é extremamente pertinente e útil para o entendimento dos aprendizados linguísticos, sobretudo se considerarmos, como já discutido, que o processo de aprendizado da L2 ocorre de maneira muito diferente daquele de aquisição da L1, em grande medida porque todos os padrões sintáticos e simbólicos (sintaxe e léxico) presentes na L1 desempenham um papel decisivo no aprendizado de novas línguas. Enquanto na aquisição da L1 esses padrões são assimilados de forma essencialmente simbólica, isto é, como padrões arbitrários regidos pelas convenções da L1, no aprendizado de uma L2 não é somente possível, mas provável que muitas dessas assimilações se deem de forma mais icônica, na medida em que os padrões sintáticos e lexicais da L2 guardarem semelhança com seus equivalentes na L1. Esse não é um aspecto secundário, mas central a todo o processo de aprendizado, e por essa razão neste estudo adotamos os conceitos de “iconicidade sintática” e “iconicidade simbólica”, como artefatos conceituais que permitirão compreender melhor o aprendizado nos contextos de ensino aqui tratados.

No Capítulo 5 os principais conceitos e ideias tratados até aqui convergirão para a proposição de uma arquitetura de tutor inteligente para ensino de línguas (ILTS). Serão destacados os principais aspectos dessa arquitetura e descrita sua implementação real.

Capítulo 5

Arquitetura ILTS e a gestão dos conteúdos

Diante do que foi discutido nos quatro capítulos precedentes, é possível aprofundar agora o estudo em direção a seu foco de contribuição, tendo-se em conta as duas principais complexidades envolvidas em um ILTS: (i) a diversidade dos aprendizes potencialmente atendidos por um curso de língua baseado na Web, com a variabilidade de *backgrounds* linguísticos, de estilos de aprendizado, de interesses e motivações e (ii) a riqueza e sutileza da linguagem humana, com sua considerável amplitude de léxico, as particularidades em seus recortes semânticos, em suas regras sintáticas e aspectos pragmáticos. Dados os limites de tempo de um estudo em nível de mestrado, optou-se, como discutido no Capítulo 1, por focalizar o trabalho no segundo ponto, aquele que trata da complexidade da automatização da gestão dos conteúdos do curso, isto é, da implementação de um modelo de domínio que permita recursos de seleção e de sequenciamento do itens lexicais para que sua exposição aos alunos seja pautada tanto por decisões de alto nível dos tutores humanos quanto pela própria semântica da L2, dando dessa maneira lugar à formação autônoma de sentenças-modelo, entre outros aspectos desejáveis. Esse foco foi ilustrado na Figura 1.4.

A definição desse foco de contribuição permite começar a esboçar as características necessárias à implementação do módulo de gestão de conteúdos do ambiente ILTS, tanto os seus aspectos mais fundamentais, tais como estratégia de modelagem do domínio e geração de sentenças, quanto os critérios de escalabilidade, generalizabilidade (para outras L2) e aderência a padrões e às práticas hoje disponíveis, como de intercomunicação de agentes, metadados para descrição de imagens e linguagem para representação de conhecimentos.

Portanto, é fundamental para o entendimento da arquitetura aqui proposta examinar mais a fundo a teoria e as técnicas de seleção de ação e de redes de comportamento, pois elas balizarão as decisões e intervenções do tutor inteligente. Isso será feito na Seção 5.1.

5.1 Modelagem de domínio proposta

O desenvolvimento de uma estratégia abrangente de modelagem do domínio em um ILTS é ainda um desafio. A fim de contribuir com essa área, um dos objetivos deste estudo foi propor uma nova modelagem do domínio para sistemas tutores inteligentes de línguas. A proposta consiste em modelar o léxico da L2 como um sistema multiagente em que cada agente representa um lexema da L2. Cada agente-lexema busca o objetivo individual de ser aprendido e que para isso deve competir por espaço na interface do curso, um recurso limitado, para que ali possa apresentar-se ao aluno com o apoio de imagens. Nesses termos, a disputa por espaço na interface do curso pode ser entendida como um processo de seleção de ação, cujas decisões também consideram uma ontologia da L2.

Um agente-lexema também pode aparecer formando sentenças. Isso é importante porque o aprendizado do valor semântico de um lexema pode depender, em muitos casos, de sua articulação com outros lexemas em sentenças-exemplo. Logo, a despeito de competir por espaço, o agente-lexema deve também cooperar com outros lexemas para favorecer a exibição daqueles com os quais ele geralmente forma sintagmas e sentenças comuns na L2.

E assim as sequências de exibição dos lexemas e de formação das sentenças emergirão dessas interações (competições ou cooperações), de modo autônomo, como nas coalizões propostas por Baars (1997) e implementadas por Dubois (2007) em outro cenário de ensino.

Para embasarem o aprendizado do léxico inicial as cenas do curso contêm imagens ilustrativas do significado dos lexemas ou das sentenças. Elas dão o contexto ou pano de fundo de cada cena. Essa informação, na forma de metadados das imagens, é levada em conta pelo mecanismo de competição e cooperação entre agentes para avaliar a pertinência de um lexema a uma dada cena. Se o lexema é um elemento presente na ou sugerido pela imagem, seu agente recebe ativação contextual e leva vantagem na disputa por espaço.

A formação de sentenças resulta de uma coalizão dos lexemas que as formam. Cabe ressaltar que elas só devem ser formadas quando ao menos a maior parte dos elementos dos quais elas são constituídas já tiverem sido exercitados, permitindo dessa forma que o(s) vocábulo(s) desconhecido(s) seja(m) inferido(s) com base naqueles já conhecidos.

A modelagem multiagente visa a dar ao sequenciamento dos conteúdos do curso uma lógica *bottom-up*, com a possibilidade de que sentenças sejam formadas de modo emergente, pela competição e coalizão de lexemas. Isso não significa, contudo, que não possam ser definidas metas didáticas *top-down* para o curso. O gestor do curso pode definir táticas que priorizem certos tópicos e temas, e assim favorecer os lexemas correspondentes na competição por espaço. Ele também pode priorizar palavras mais frequentes na L2, ou as mais parecidas com as correspondentes na L1, por exemplo.

Para estudarmos a modelagem de domínio proposta, a Seção 5.2 discute o uso de uma ontologia como base de conhecimento da L2, e das relações semânticas e sintáticas entre seus itens léxicos. Por sua vez, na Seção 5.3 se retomará a discussão da modelagem do léxico baseada em agente. Se discutirá como os agentes-lexemas interagem entre si e se comunicam com a ontologia da L2. Em seguida, na Seção 5.7 se tratará da semantização das imagens de apoio usadas no ILTS, examinando por que e como o uso de metadados (anotações) nessas imagens torna possível o funcionamento proposto para a arquitetura. Por fim, na Seção 5.8 serão apresentados diagramas esquemáticos da arquitetura proposta.

5.2 Uma ontologia do léxico da L2

A geração automática de sentenças-exemplo é uma funcionalidade útil a um ILTS, pois ela libera o especialista (tutor humano) da tarefa de elaborar essas sentenças sempre que necessárias e de forma individualizada para cada aprendiz, ou de ter que elaborar *a priori* um vasto repertório de sentenças para uso no curso. Entretanto, a geração automática não pode prescindir dos conhecimentos que o tutor humano tem da L2, isto é, de sua vivência com a língua, com os recortes semânticos desta, suas regras sintáticas e aspectos pragmáticos. Isso significa então que a automatização da geração de sentenças requer formas de sistematizar o conhecimento linguístico do tutor humano de maneira tratável pelo ILTS. Na IA, uma solução sistemática de representação do conhecimento se baseia no uso de ontologias. Assim, o uso de uma ontologia para definir as relações semânticas entre os lexemas da L2 é o segundo pilar da arquitetura aqui proposta. A ontologia define as relações de sinonímia, antonímia, hponímia e hiperonímia entre os lexemas da L2, além das regras que ligam os diversos conceitos segundo suas relações existenciais e associa a

cada substantivo os adjetivos com os quais ele pode se articular, ou os verbos dos quais ele pode ser sujeito ou objeto. Por exemplo, a ontologia pode informar, de modo tratável por um computador, que o ente *rinoceronte* é um animal mamífero, grande, que tem chifres, quatro patas, rabo e que ele se alimenta de vegetais. Essa ontologia constitui então uma base de conhecimento sobre a L2. Ela pode ser uma combinação dos conhecimentos que o especialista do domínio possui com outros dados provenientes de gramáticas e dicionários ou mesmo de registros de uso da L2 (obtidos, por exemplo, por motores de busca na Web).

O léxico da L2 é então a taxonomia na qual se baseia a ontologia que serve de base de conhecimento para o ILTS. Entretanto, além da taxonomia a ontologia deve conter as hierarquia de classes e subclasses de cada conceito da língua, como, por exemplo, tratar a classe *rinoceronte* como membro da classe *mamífero*, e a superclasse *mamífero* como membro da superclasse *animal*. Mais do que reproduzir a hierarquia conceitual dos itens lexicais, a ontologia deve ser um compêndio dos conhecimentos a respeito de cada palavra, pois essas informações fundamentarão semanticamente cada símbolo da L2.

Questões de escalabilidade foram consideradas na construção da ontologia da L2, pois, a despeito de o léxico desta, com cerca de três mil vocábulos básicos, ser econômico em comparação com outras línguas, uma ontologia com três mil classes cria problemas de portabilidade entre aplicativos. A opção foi construir várias ontologias menores, cada uma cobrindo parte do vocabulário. Essa decisão leva em conta o fato de que o aplicativo de criação de ontologias, apresentado a seguir, consegue tratar ontologias de ontologias.

Ferramentas para a criação e utilização de ontologias

Como mencionado no Capítulo 2, entre as ferramentas mais adequadas à construção de modelos de domínio estão os editores de ontologias. Em vista dos recursos que oferece e de suas APIs para interoperação com *reasoners* diversos, neste trabalho foi utilizado para a construção das ontologias o software Protégé, levando-se em conta a importância, para a arquitetura aqui proposta, de uma boa interoperabilidade do gerenciador de ontologia com outros aplicativos utilizados. O Protégé permite também integrar distintas subontologias no âmbito de uma ontologia principal. Esse recurso é particularmente necessário à proposta de se construir uma ontologia do léxico da L2, pois isso significa integrar todas as ontologias criadas para representar as classes correspondentes aos lexemas, o que no estudo de caso aqui tratado é algo da ordem de três mil classes, que se fossem criadas em uma única ontologia levariam a dificuldades de interoperabilidade com outros aplicativos. Isso é ainda mais crítico em futuras extensões para outras línguas pois essas poderão significar algumas dezenas de milhares de classes para cobrir só o léxico básico do inglês, por exemplo. Logo, o estudo de caso escolhido possibilitará ter uma real noção do esforço de implementação que essa estratégia pressupõe.

Na Figura 5.1 está ilustrada uma das subontologias, com a superclasse (*vivanto*) referente aos seres vivos e outros conceitos relacionados, como ação (*ago*), qualidade (*eco*), lugar (*loko*), etc. Na Figura 5.2 são ilustradas algumas das cerca de 300 classes (de adjetivos) da L2, agrupadas na superclasse qualidade (*eco*). Por sua vez, na Figura 5.3 são mostrados alguns dos cerca de 800 verbos selecionados para o ILTS, e que foram agrupados na superclasse ação (*ago*).

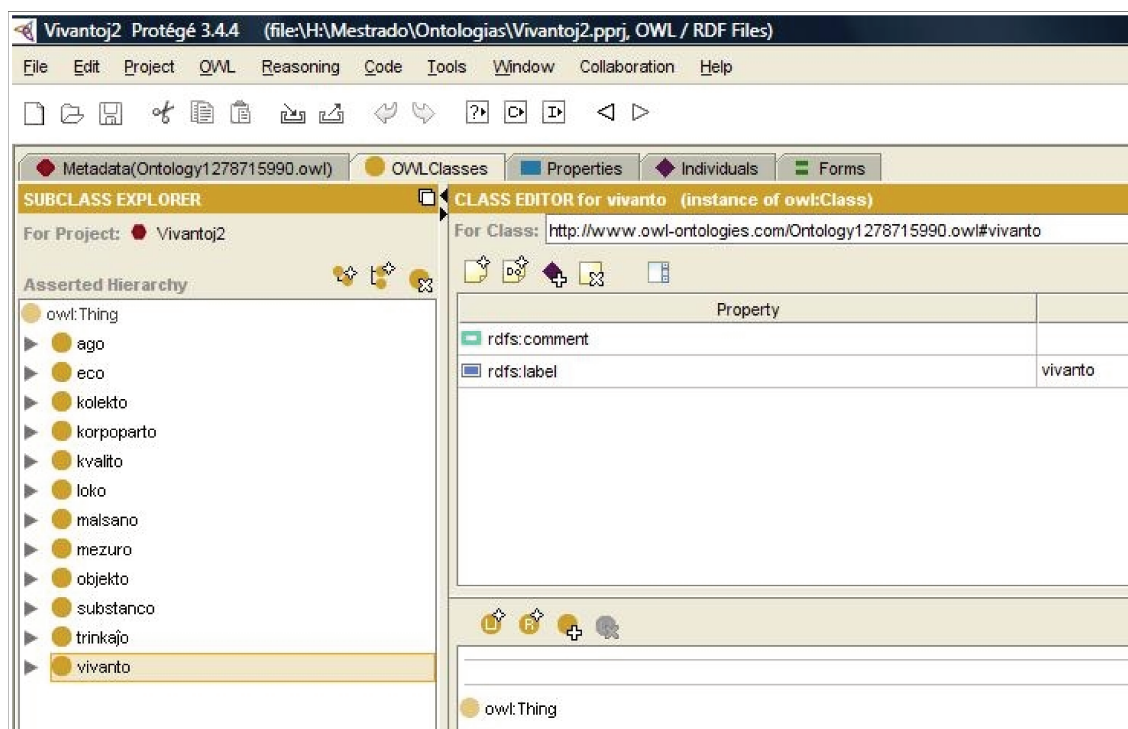


Figura 5.1: Superclasses da ontologia para seres vivos e conceitos relacionados

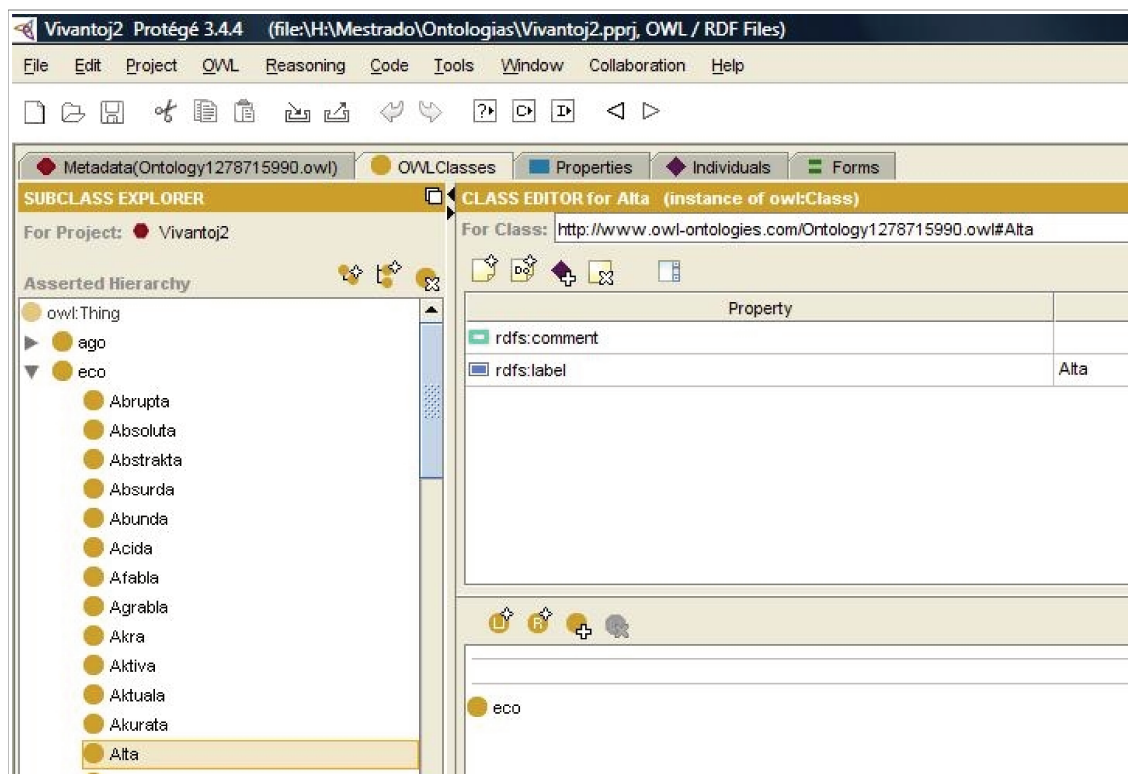


Figura 5.2: Classes de adjetivos dentro da superclasse “eco” (qualidade)

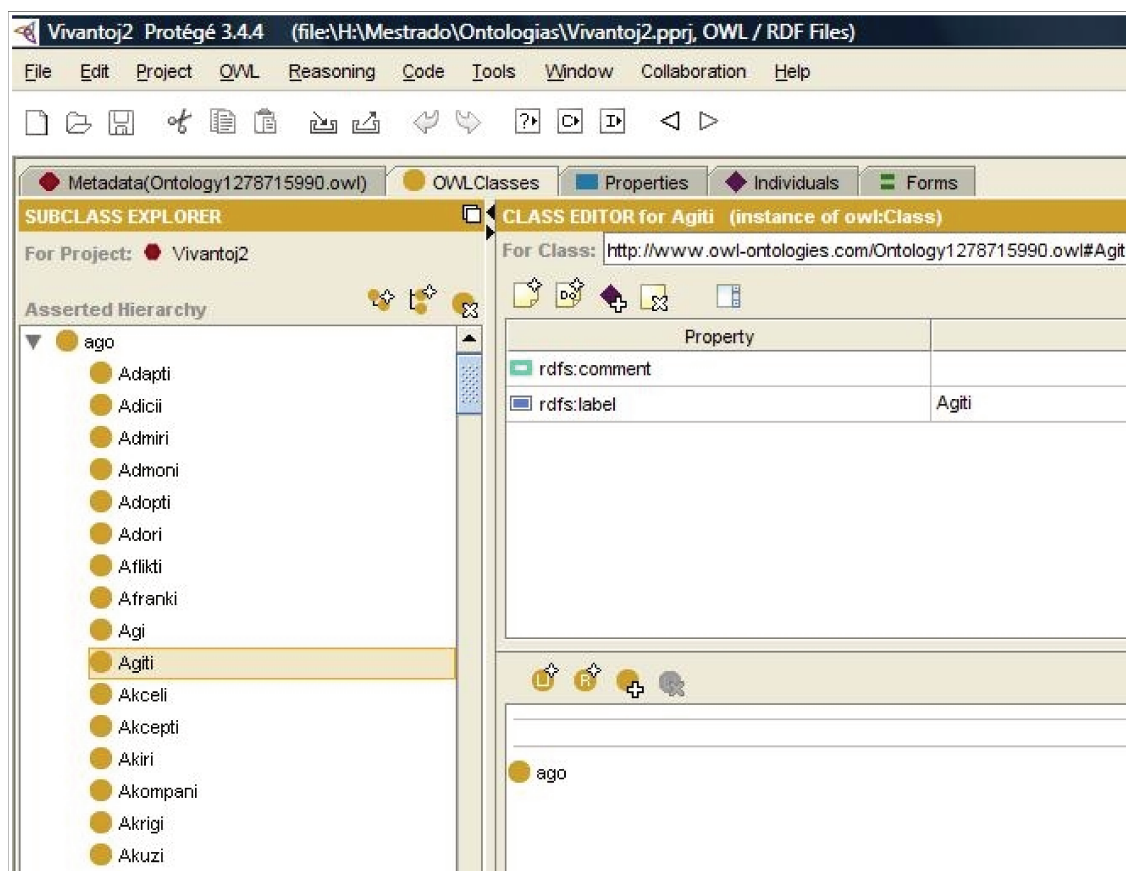


Figura 5.3: Classes de verbos dentro da superclasse “ago” (ação)

As classes são definidas como o conjunto dos indivíduos que representam os conceitos concretamente. As classes podem ser organizadas em taxonomias de superclasses e subclasses. Na Figura 5.4 está ilustrada a classe “rinocero”, na hierarquia que inclui ungulados, mamíferos, animais e seres vivos. A hierarquia de classes permite que atributos definidos para as superclasses sejam herdados por suas subclasses. Assim, quando definido nos dados da ontologia que a classe dos mamíferos é formada por animais de sangue quente, esse dado é herdado por todas suas subclasses, inclusive pela classe “rinocero”.

A definição de atributos semânticos para a classe pode também ser feita por meio da criação e atribuição de propriedades ou atributos que a correlacionam a outras classes e têm a função de interligar os indivíduos. As propriedades podem ser de objeto (conectam um indivíduo a outro indivíduo), de tipo de dados (conectam um indivíduo a um valor literal do tipo XML ou RDF) e de anotações (utilizadas para adicionar metadados as classes, indivíduos, propriedades do objeto e do tipo de dado). Por exemplo, entre as propriedades mostradas na Figura 5.5 algumas se aplicam ao ente “rinocero”, como, por exemplo, *vivasEn* (vive em), *manças* (come), *timas* (teme), *pezas* (pesa), etc. A primeira especifica o habitat do rinoceronte, e assim recebe o valor *savano* (savana), que é também ela própria uma classe da ontologia. Em destaque na figura está a regra *havasKorpoparton* (tem parte

do corpo), que aceita como valor a classe *Korpoparto* e se aplica às classes *Besto* (animal) e *Homo* (ser humano). No caso do ente “rinocero”, essa regra recebeu valores como *kapo*, *piedo*, *korno*, *vosto*, etc.

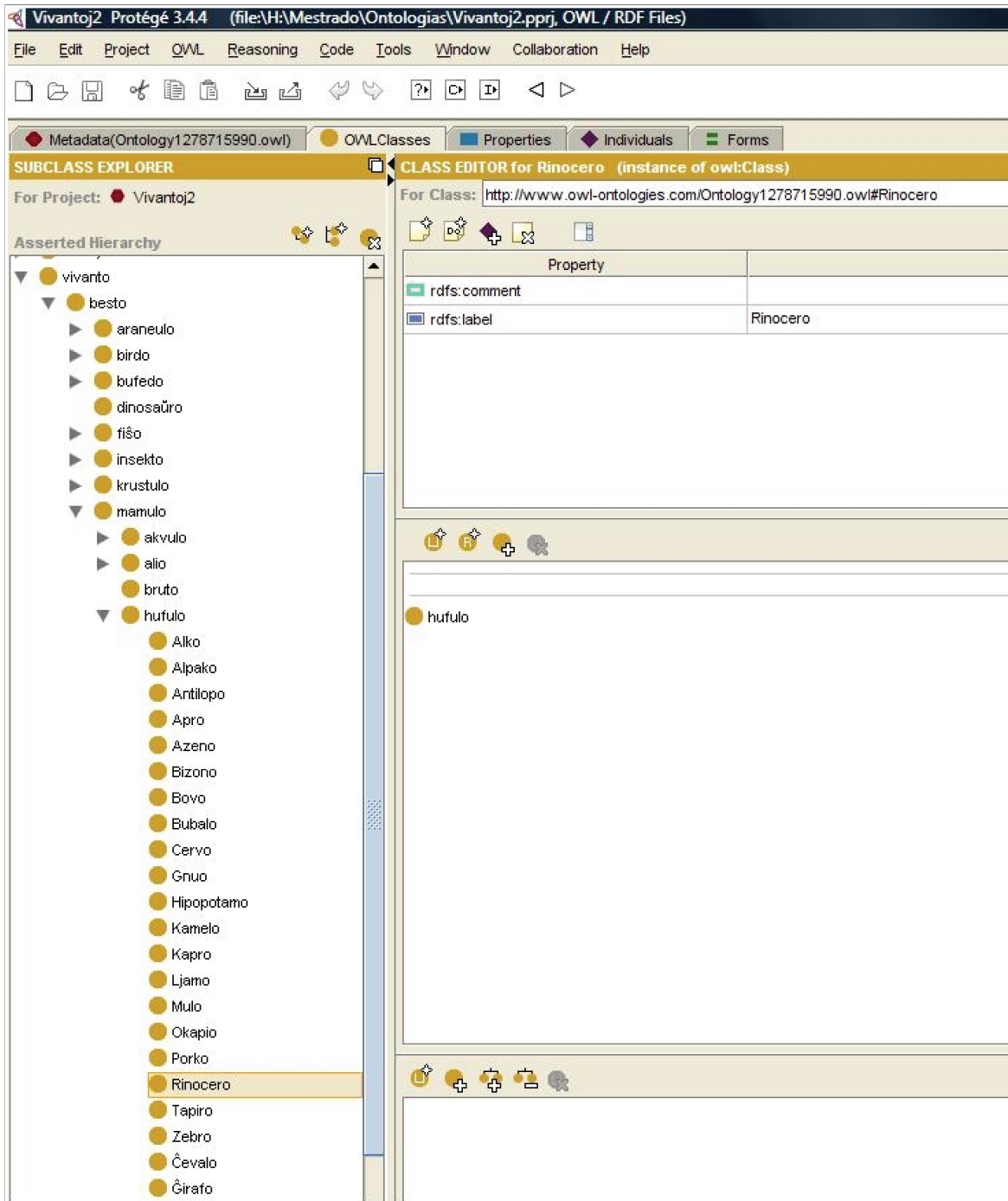


Figura 5.4: Classe “rinocero” dentro da superclasse “mamulo” (mamífero)

As ontologias podem ser construídas na linguagem OWL, recomendada pelo W3C (World Wide Web Consortium) e que oferece mais recursos para descrição de classes e propriedades, dando mais expressividade sobre o domínio definido. Um recurso importante é a possibilidade de especificar tipos para as propriedades, as quais podem ser: funcionais, funcionais inversas, transitivas, simétricas, assimétricas, reflexivas e irreflexivas.

Além das relações baseadas em propriedades, as informações sobre cada ente podem incluir dados factuais, como, no exemplo do “rinoceronte”, peso, altura, velocidade, etc. Na Figura 5.6 estão ilustradas algumas dessas informações, que, assim como as propriedades, podem ser consultadas pelo mecanismo de geração de sentenças para garantir que somente sentenças semanticamente válidas sejam geradas, como será discutido na Seção 5.7. Essa é uma das principais funções da ontologia na arquitetura aqui proposta: fundamentar com dados do mundo real cada um dos vocábulos (ítems léxicos) ensinados no curso. A construção dessa base de conhecimento não é, contudo, rápida nem fácil, e exigirá dos especialistas do domínio uma grande quantidade de esforço, sobretudo na criação das classes e na atribuição das propriedades que as interrelacionem. Entretanto, o volume de trabalho necessário à implementação do ambiente deve ser avaliado em termos dos ganhos que uma ferramenta de ensino baseada em um tutor inteligente, e portanto com alto grau de autonomia, pode trazer em um contexto de ensino baseado na Web. Em outras palavras, o esforço necessário para implementar o ILTS pode se tornar aceitável se considerada a economia de esforço e a qualidade do aprendizado que se espera ele seja capaz de propiciar.

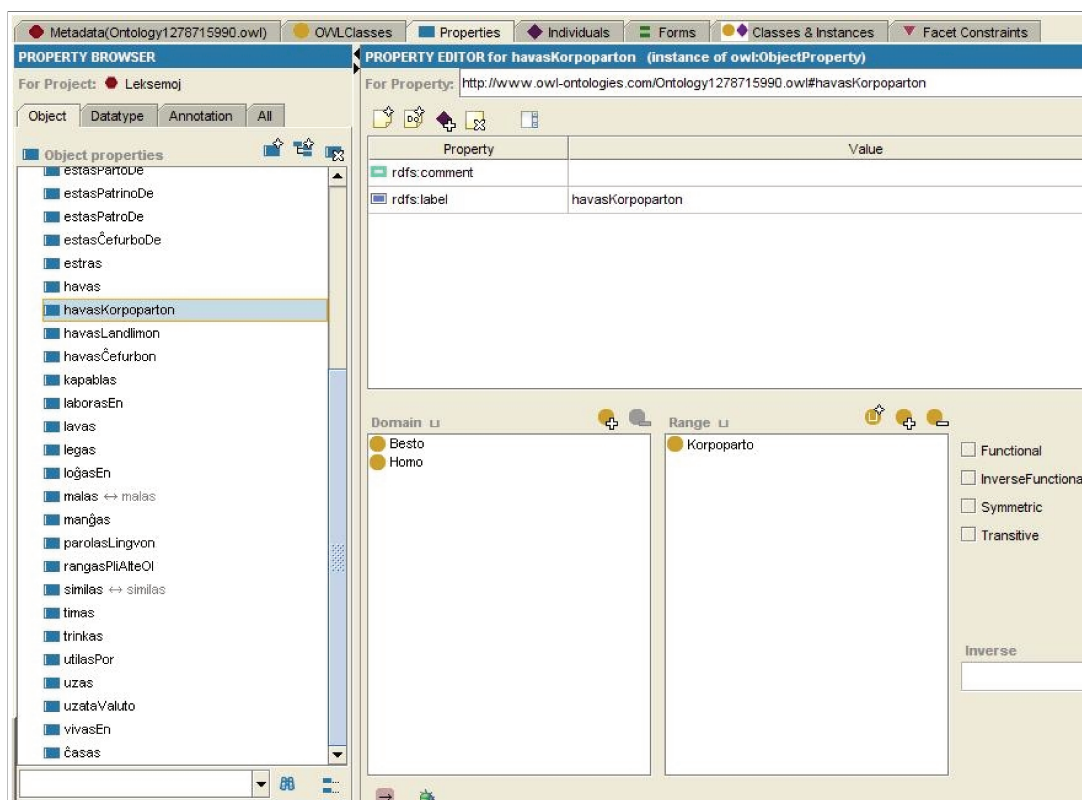


Figura 5.5: Regras de relacionamento entre classes

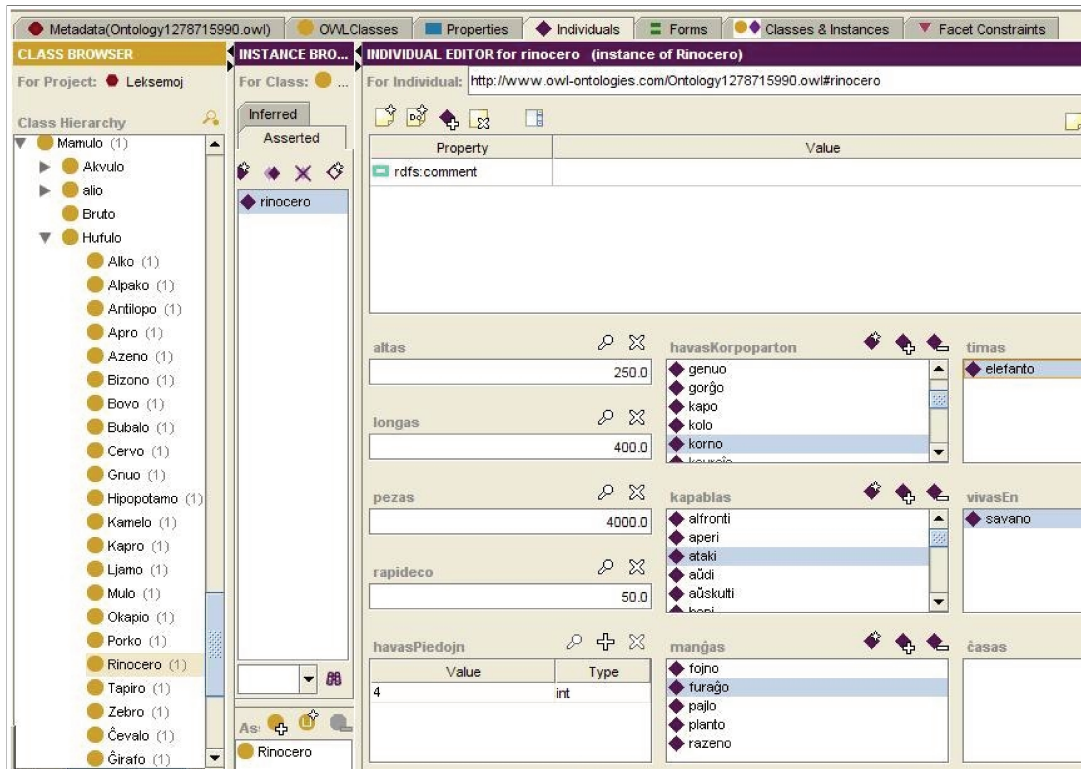


Figura 5.6: Exemplo de atributos para o conceito “rinocero”

5.3 Gestão do léxico baseada em agentes

Como discutido no Capítulo 2, a modelagem baseada em agentes é vantajosa em um grande número de situações, em razão de sua capacidade de modelar sistemas complexos em que se verifiquem fenômenos emergentes. Esse é o caso do ILTS aqui descrito, que deve ser capaz de modelar o domínio representado por uma L2. A estratégia de fazer uma modelagem multiagente da L2 significa que cada lexema será instanciado na forma de um agente e programado para demonstrar um comportamento autônomo, esforçando-se para ser aprendido, competindo por espaço na interface por um mecanismo de seleção de ação tal qual será discutido na Seção 5.5. Em outras palavras, cada agente tem a meta individual de ensinar seu próprio significado aos usuários do ambiente. A apresentação na interface do curso é uma condição necessária para que o aprendiz possa assimilar o sentido de um lexema, seja este exemplificado de modo isolado, seja em combinação com outros lexemas. Além disso, a sequência de apresentação dos lexemas na tela deve ser coerente com os objetivos pedagógicos e os critérios de individualização do ensino que, como visto no Capítulo 3, pode priorizar conteúdos mais frequentes na L2, ou os mais parecidos com os da L1, ou os mais simples, ou ainda os de um tema específico. Em posteriores versões do ILTS isso poderá levar em consideração informações detalhadas do modelo do aprendiz, mas em vista de sua amplitude isso está fora do escopo da primeira etapa do trabalho.

O aprendizado de um lexema em resposta à sua apresentação na tela pode ocorrer de diversas formas, dependendo dos elementos (signos) ali exibidos:

- 1) Com o apoio de uma imagem ilustrativa (desenho ou foto)
- 2) Pela semelhança que o lexema tem com o correspondente na L1
- 3) Devido à relação com outros lexemas no contexto de sentenças-exemplo
- 4) Por meio de uma definição formal de seu significado

Para cobrir as quatro possibilidades, cada agente-lexema tem associadas (i) uma imagem de apoio; (ii) uma métrica indicando quão similar ele é em relação ao lexema correspondente na L1; (iii) uma lista de lexemas com os quais ele pode formar sintagmas e sentenças válidos na L2; e (iv) uma definição formal de seu significado (escrita com o auxílio de outros lexemas da L2). Na Tabela 5.1 são exemplificados dados do agente-lexema *rinocero*.

Tabela 5.1: Exemplo dos campos de um agente-lexema

Campo	Tipo	Exemplo
Substantivo	texto	rinocero
Imagem (1)	imagem	G:\Desenhos_curso\rinocero.png
Posição do vocábulo	posição	[40.0 70.0]
Posição da imagem	posição	[45.0 70.0]
Nível de ativação	real	100
Repetições	inteiro	0
Número de sentenças	inteiro	20
Temas	lista	[ekologio; naturo; aventuro]
Semelhança com vocábulo na L1 (2)	real	0.85
Frequência no <i>corpus</i>	inteiro	10
Adjetivos (3)	lista	[granda, impona, sovaĝa, rapida, danĝera]
Sujeito (3)	lista	[ataki, kuri, manĝi, trinki, stari, minaci,]
Objeto direto (3)	lista	[ĉasi, rigardi, protekti, fotografii, filmi]
Hiperônimos (3)	lista	[hufulo, mamulo, besto, vivulo]
Hipônimos (3)	lista	[]
Sinônimos (3)	lista	[]
Antônimos (3)	lista	[]
Definições (4)	texto	Granda sovaĝa dikhaŭta mambesto, kun unu aŭ du pintaj kornoj elstarantaj sur la nazo.

Na arquitetura proposta os dados contidos nos agentes-lexemas dividem-se em duas categorias. A primeira reflete características do lexema representado pelo agente, tais como definição, tema, hiperônimos, hipônimos, antônimos, sinônimos, semelhança com o correspondente lexema na L1, frequência no *corpus* da L2, etc. A segunda categoria contém informações sobre o agente dentro do ambiente de simulação (nível de ativação, endereço de sua imagem de apoio, número de vezes em que apareceu em cenas, sua posição na tela, etc.). Esta última categoria de informações pode ser alterada durante a simulação pela própria dinâmica de interação dos agentes, enquanto a primeira é em princípio mais estável, podendo contudo ser alterada pelos especialistas do domínio em modo *offline*, para, por exemplo, enriquecer a lista de sinônimos, ou re-escrever a definição. É importante notar que a categoria de características do lexema pode ser preenchida com base na (a partir da) ontologia da L2, de onde se podem extrair as relações hierárquicas (sinonímia, antonímia, hiperonímia e hiponímia), as de tipo “sujeito” e “objeto”, listas de adjetivos, de verbos, etc.

Entretanto, para não onerar demasiadamente o processamento do ILTS, a maior parte dos dados da ontologia é consultada durante a inicialização dos agentes-lexema e não em tempo real. Após inicializados, os agentes conservam essas informações até nova atualização. Um mecanismo de consulta à ontologia em tempo real não é contudo descartado, e a arquitetura pode contar com um agente especializado para acessar o *reasoner* da ontologia para assim obter informações que não estejam explicitamente declaradas nela, mas que possam ser inferidas com base nas que estão explícitas.

5.4 Semantização das imagens de apoio do ILTS

Além do mecanismo multiagente e do uso de ontologias, a arquitetura aqui proposta baseia-se em um terceiro pilar, que é o da utilização, no ambiente de ensino, de metadados para seus conteúdos visuais. Essa utilização tem o objetivo de permitir aos agentes entender os elementos semânticos das imagens de apoio utilizadas no curso. Isso é particularmente necessário em vista do fato de o mecanismo de seleção de ação adotado estar baseado na metáfora do teatro, como será visto na Seção 5.5. Nela é o pano de fundo da cena que dá aos agentes o real contexto das difusões (*broadcasts*) feitas pelo agente-gestor. No ILTS proposto são as imagens de apoio que determinam o contexto de cada cena. Portanto, esses significados devem estar acessíveis aos agentes-lexemas diretamente nas imagens para estes saberem se e quando são pertinentes ao contexto da cena, e assim autonomamente se lançarem na disputa por espaço nela e pela participação em coalizões (sentenças-exemplo).

A associação entre imagem e significado é feita por meio de descritores semânticos, os quais se tornam os metadados da imagem. Os descritores podem ser associados inclusive a segmentos ou elementos constituintes de uma imagem. Por exemplo, na imagem do ente *rinocero* os metadados podem informar que segmentos da foto representam objetos e partes semanticamente ligados ao lexema-foco, tais como *chifre*, *patas*, *boca*.

Os metadados são definidos *a priori* e codificados em MPEG-7 pelo especialista do domínio, que assim pode, levando em conta cada imagem a ser utilizada no curso, agregar os descritores e mapas semânticos que identifiquem os principais elementos do ponto de vista didático, além das inter-relações e papéis desses elementos em um dado contexto. Podem também ser mapeadas relações tais como, por exemplo, *componentOf*, *agentOf*, *beneficiaryOf*, *dependsOn*, as quais são definidas no MPEG-7 e codificadas em XML, podendo com isso ser interpretadas pelos agentes.

As informações do contexto da cena somam-se àquelas provenientes da ontologia (inclusive as carregadas nos agentes na inicialização), complementando-as. Esses dois tipos de referencial permitem ao ILTS formar sentenças que sejam semanticamente válidas na L2 (apoiadas em dados da ontologia sobre o objeto ou conceito), ou que reflitam o contexto da cena e dos elemento(s) retratado(s) na imagem. No exemplo do *rinoceronte*, se na imagem este aparece parado, com seu dorso sujo de lama, e não está comendo, sentenças adequadas a essa cena serão: *O rinoceronte está parado* ou *o rinoceronte está sujo de lama*. Mas considerando, como é sabido (com base na ontologia), que animais dessa espécie se alimentam de vegetais, vivem em savanas africanas, são velozes e nervosos. Assim, além daquelas sentenças referentes à cena outras aplicáveis à espécie como um todo poderão ser criadas: *Rinocerontes vivem na África*, *Rinocerontes comem vegetais* ou *Rinocerontes são*

nervosos e velozes. Por outro lado, sentenças que não reflitam nem a cena de fundo, nem dados da ontologia, como, por exemplo, *O rinoceronte está voando* ou *O rinoceronte é carnívoro* não são formadas, pois são anômalas: contradizem informações contidas na base de conhecimento (ontologia) ou na descrição contextual (metadados), respectivamente.

Cabe ressaltar que embora gramaticalmente corretas, as sentenças anômalas podem ter no aprendizado de uma L2 o efeito de dificultar a compreensão do significado das palavras, uma vez que o sentido dos enunciados se choca com informações que o aluno tem do objeto em questão ou com aquelas apresentadas explicitamente na imagem na cena.

Um aspecto importante do uso de metadados no contexto do ILTS aqui proposto é que caso necessite incorporar novas imagens ao curso o especialista do domínio não precisa se preocupar em criar todas as sentenças correspondentes aos elementos e temas retratados nas novas imagens, pois a própria natureza autônoma do mecanismo de seleção de ação vai perceber os descritores semânticos nas novas imagens e, quando cada uma delas estiver em foco, produzirá uma geração automática de sentenças pertinentes àquelas imagens.

Uso de metadados em imagens

O uso de metadados em conteúdos multimídia tem o propósito de reduzir a chamada brecha semântica entre aquilo que um ser humano é capaz de interpretar de um conteúdo imagético (foto ou desenho) e o que uma máquina consegue extrair desse mesmo conteúdo. Como tratado no Capítulo 4, os seres humanos adquirem ao longo de seu desenvolvimento cognitivo uma refinada capacidade de interpretação de imagens. É essa capacidade que nos permite olhar para uma foto e prontamente reconhecer objetos, animais, pessoas, situações, emoções e, eventualmente, até a intenção de quem produziu aquela imagem. Somos assim capazes de nomear e classificar a imagem e seus elementos, isto é, de capturar a semântica da imagem a partir de suas características visuais. Esse não é absolutamente o caso da interpretação automática de imagens em vista da ainda limitada capacidade de interpretação semântica de imagens pelas máquinas (computadores). E ainda que um notável avanço tenha sido observado na última década no que concerne às implementações e técnicas de busca e recuperação de imagens, permanece enorme a distância entre o que um ser humano e um computador são capazes de inferir a partir de imagens.

Na abordagem mais tradicional um detector automático de conceitos em imagens baseia-se na extração de atributos de baixo nível (padrões de cor, gradientes, contrastes, etc.) e no uso de técnicas de aprendizado supervisionado de máquina. O classificador deve passar por uma fase de treinamento baseado em um conjunto de imagens conhecidas.

Contudo, no estado atual da técnica as implementações são capazes de classificar imagens para um número limitado de conceitos, e com probabilidades de acerto que variam de 30 a 70%, dependendo do conceito. Embora esses números cresçam ano após ano, as técnicas atuais de reconhecimento de imagens ainda são incapazes usar relações ontológicas entre elementos detectados nas imagens para assim inferir a presença de outros elementos, algo que o cérebro humano é capaz de fazer de forma totalmente automática e inconsciente.

De resto, em contraste com palavras e textos, imagens não podem ser interpretadas a partir de unidades discretas contáveis que, em combinação, possam sequenciá-las em ordem alfabética ou dar pistas sobre o significado dos elementos e temas ilustrados. Uma imagem

digital é, em essência, um conjunto de bits que representam um dado padrão de pixels. Mas isso se dá em um nível pré-semântico, não havendo relação direta entre a sequência de bits e os conceitos retratados. Em outras palavras, os descritores de uma imagem não são dados na própria imagem, é preciso extraí-los automaticamente com auxílio de uma das técnicas acima mencionadas ou acrescentá-los manualmente a partir da interpretação de um ou de vários anotadores humanos. Uma vez produzidos, esses descritores passarão a compor os metadados daquela imagem, na forma de um arquivo suplementar, a ela vinculado.

O Padrão MPEG-7 foi desenvolvido com o propósito de representar informações sobre um conteúdo e não o conteúdo propriamente dito. Isso atende ao objetivo de prover independência entre a descrição e o conteúdo. Outros objetivos do padrão são descrever os principais aspectos de um conteúdo (característica de baixo nível, estrutura, modelos, etc.), informar como os objetos são combinados em uma cena e possibilitar buscas, filtrações e identificações rápidas e eficientes de conteúdos. O MPEG-7 trata áudio, voz, vídeo, imagens, grafos e modelos 3D. O MPEG-7 armazena os metadados em formato XML.

Em razão dessas características, o padrão MPEG-7 foi adotado na codificação dos metadados do ambiente de ensino aqui descrito. E embora o MPEG-7 ofereça importantes funcionalidades para manipulação e transmissão de objetos e de seus metadados, a extração de descrições semânticas e a anotação do conteúdo com metadados está fora do escopo do padrão. Por essa razão, as imagens de apoio foram anotadas por aplicativos compatíveis com o MPEG-7, mas que oferecem recursos complementares, como a associação dos descritores a ontologias externas, como é, por exemplo, possível no aplicativo de anotação M-OntoMat-Annotizer. Isso é particularmente útil no nosso ILTS, visto que ele já utiliza uma ontologia do vocabulário da L2 como um dos referenciais para a formação de sentenças-exemplo. Desse modo, o uso das mesmas ontologias para fornecer os descritores a serem anotados nas imagens é não só conveniente, mas importante para criar um vínculo entre as ontologias e os metadados. Na Figura 5.7 é exemplificada uma tela do aplicativo de anotação, na qual é feita a correlação das áreas demarcadas com termos na ontologia.

No contexto do ILTS as imagens de apoio são anotadas a partir de um ponto de vista pedagógico, que privilegia aqueles termos que fazem parte do vocabulário a ser aprendido. As anotações descrevem não só o tema central de cada imagem (lexema-foco) mas também os lexemas acessórios, que se referem a partes daquele ser ou objeto, ou ainda a outros elementos presentes na cena e que possam ser usados nas sentenças-exemplo. Isso significa que as imagens de apoio devem ser selecionadas com base nos critérios de pertinência didática e de clareza de interpretação. Mas cabe notar que, diferentemente da maioria das implementações de metadados na Web, a função dos metadados neste caso é dar ao tutor inteligente as informações semânticas necessárias para que este saiba quais são os temas e contextos de cada imagem, e possa em função disso exemplificar as palavras e gerar as sentenças mais pertinentes a cada cena. Isso equivale a dizer que o objetivo dos metadados, no ILTS, é também reduzir a brecha semântica entre o que o sistema sabe sobre as imagens e aquilo que os alunos são capazes de interpretar quando as analisam. Mas ao fazer isso e ao apresentar ao aluno os lexemas da L2 para cada um dos elementos retratados busca-se, em última instância, criar as condições para que o ILTS reduza a brecha simbólica entre aquilo que o aluno reconhece na imagem e a maneira correta de exprimir tais e conceitos na L2.

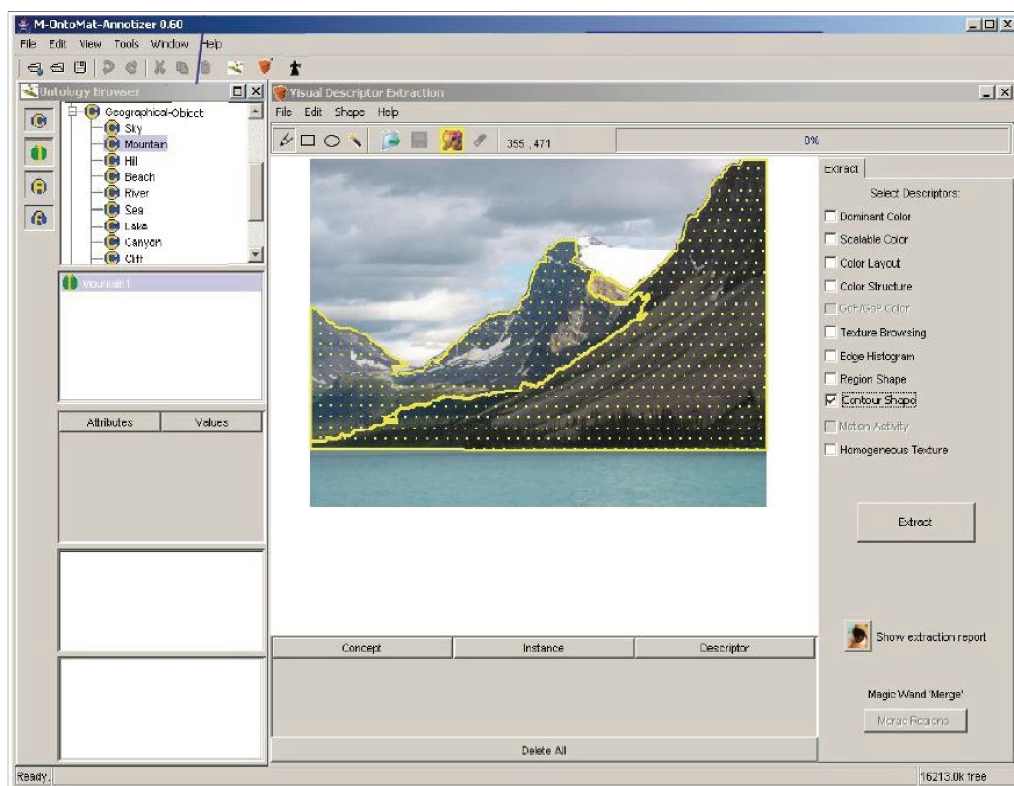


Figura 5.7: Exemplo de anotação de metadado em imagem

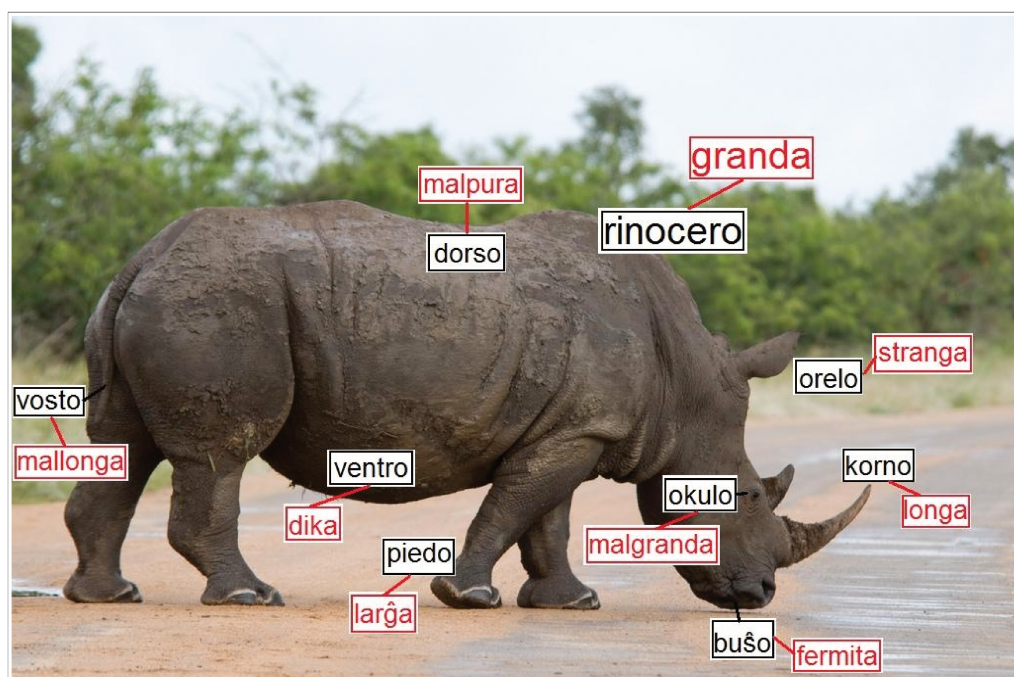


Figura 5.8: Exemplos de anotações do lexema principal e dos relacionados

Na Figura 5.8 é apresentado um exemplo de rotulagem didática de uma imagem, com descritores do lexema-foco *rinocero*, e de vários outros objetos/partes relacionados a ele, como *dorso*, *vosto*, *okulo*, *korno*, *orelo*, e os adjetivos correspondentes. Cada par de substantivo e adjetivo é associado a uma região da imagem. No modelo de interação previsto para a na interface do ILTS, quando uma região da imagem é apontada pelo aluno o foco do mecanismo de seleção de ação é dirigido para os lexemas ali anotados, com um acréscimo da energia de ativação dos agentes correspondentes e a consequente geração de sentenças relacionadas àqueles vocábulos. Isso não impede os demais lexemas de participar das sentenças, pois eles continuam pertinentes à cena (ativação contextual) e relacionados (na ontologia) aos lexemas selecionados.

Embora a anotação das imagens de apoio do curso possa demandar um considerável esforço na implementação do ILTS, uma vez constituído um banco de imagens anotadas o trabalho exigido dos especialistas do domínio e da aplicação tende a se reduzir. E, por outro lado, a disponibilidade de metadados nas imagens de apoio deverá reduzir muito o esforço necessário à produção de conteúdos textuais para o curso, além de tornar mais fácil a busca e recuperação dessas imagens, por exemplo em um contexto de uso compartilhado ou colaborativo desses conteúdos na Web.

5.5 Arquiteturas de referência para o ILTS aqui proposto

Nas seções precedentes foi mostrado como o ILTS aqui proposto se fundamenta em três tipos de artefatos tecnológicos, a saber, em ontologias, na modelagem multiagentes e no uso de metadados nas imagens de apoio do ambiente de ensino. É importante discutir agora o mecanismo proposto para conectar esses três artefatos em uma aplicação única que demonstre as desejáveis características de autonomia, responsividade, proatividade. Isso foi conseguido com o uso de mecanismos de seleção de ação (introduzidos no Capítulo 2).

Alguns autores, entre eles Stan Franklin, consideram que a perspectiva correta para se compreender o papel e a evolução da mente é aquela baseada em seleção de ação. Esse ponto de vista está, de resto, fortemente apoiado no modelo de consciência proposto por Baars (Baars, 1988, 1997) e que deu origem a algumas implementações de ambientes baseados em agentes, e cujas arquiteturas contêm elementos que são ao mesmo tempo mecanismos de planejamento, estrutura de decisão e memória procedural de longo prazo. Os planejamentos e as decisões que ele produz ocorrem de modo inconsciente mas se valem de comunicação com uma consciência central a fim de manter ciência da evolução das situações. A esse comportamento é dado o nome de seleção de ação conscientemente mediada. A primeira arquitetura desse tipo, chamada de IDA, foi proposta por Negatu e Franklin (2002) e é uma evolução da arquitetura de Maes. Na IDA cada comportamento é realizado por um grupo de agentes, denominados codelets, que permanecem sempre ativos, atentos às difusões (comunicações) que emanam do módulo central. Quando algum codelet se sente útil ou relevante ao contexto, ele se integra à rede de comportamento, alterando-a.

A arquitetura CTS (Conscious Tutoring System), proposta por Dubois (2007), baseou-se na arquitetura IDA para implementar um agente consciente com a função de um ITS para treinamento de astronautas na operação do braço robótico da Estação Espacial Internacional. Denominado Canadarm, esse braço robótico é um mecanismo dotado de sete

juntas e, portanto, possuindo sete graus de liberdade em seus movimentos, como ilustrado na Figura 5.9. Ele possui quatro câmeras com ângulos de visão complementares e, além disso, dependendo de sua posição em um dado momento, pode ser visto por algumas entre outras seis câmeras instaladas em pontos diferentes da Estação Espacial. A dificuldade da operação do braço está no fato de os astronautas não poderem vê-lo inteiro (e em três dimensões) através de uma janela, mas terem que inferir sua posição e seus movimentos pela combinação das imagens 2D dessas câmeras, enquanto comandam seus movimentos por meio de dois controles manuais. Isso implica um certo número de procedimentos obrigatórios e restrições no manejo do braço, o que faz desse aprendizado um desafio maior do que aparenta. O ILTS aqui descrito baseia-se em muitos conceitos aplicados no CTS.

O CTS assenta-se na ideia de que um agente deve apresentar quatro características: autonomia, responsividade, proatividade e sociabilidade. As três últimas são, além do mais, consideradas componentes da inteligência. Um tutor, por sua vez, deve ter a capacidade de perceber seu ambiente (incluindo o aprendiz) e possuir autonomia que lhe permita reagir ou agir da maneira mais apropriada, no momento certo. Isso implica uma capacidade de ver o contexto, reconhecer tendências, antecipar consequências, planejar e adaptar-se com base nisso e agir para tentar produzir o resultado mais apropriado ao contexto, interagindo às vezes com outros agentes para atingir sua meta. Como isso ultrapassa a capacidade de sistemas convencionais, justifica-se a opção pelo paradigma de agentes (Dubois, 2007).

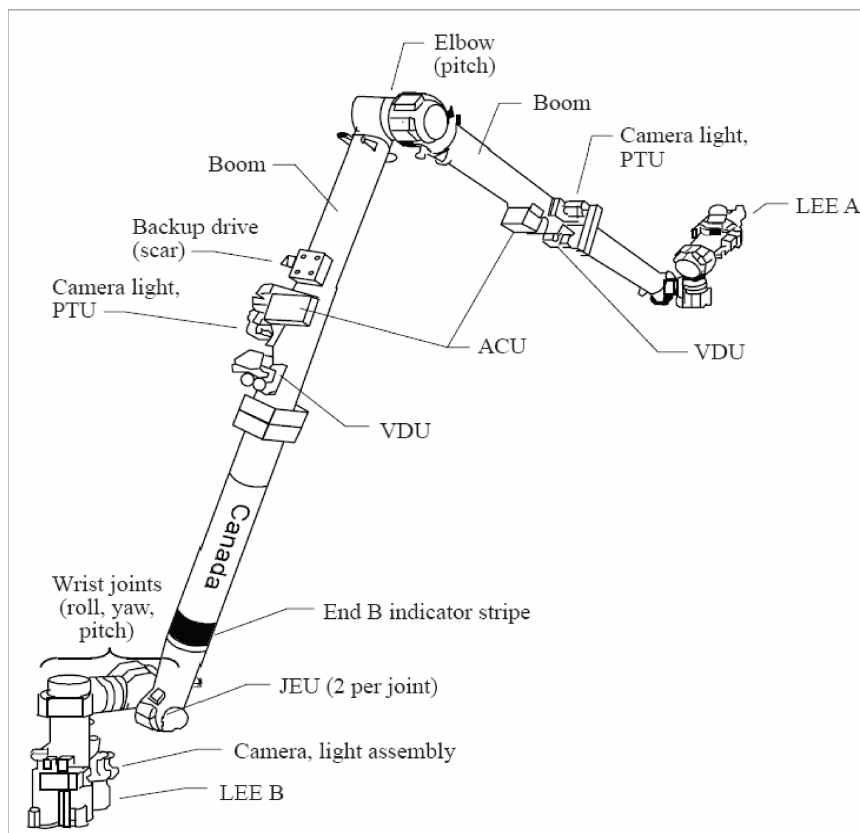


Figura 5.9: Braço robótico da EEI (Dubois, 2007)

O CTS é uma evolução da arquitetura IDA, proposta alguns anos antes por Franklin (1997 e 2001). Ambas as arquiteturas se caracterizam por adotar o modelo de consciência proposto por Baars (1997), e denominado “Global Workspace Theory”. Nele, a mente é modelada segundo a metáfora do teatro, a qual conta com um palco, uma vasta plateia de atores especializados e alguns outros elementos nos bastidores. Na metáfora, uma peça sem roteiro é encenada e seu desenrolar depende do talento dos atores presentes na plateia de intervir quando sentirem que podem contribuir para a trama que estão assistindo. Em geral, essa intervenção requer a participação simultânea de um certo número de atores no palco, dos quais só alguns ficam sob os holofotes. Esses poucos estão relacionados entre si de alguma forma e se apoiam mutuamente. Nos bastidores, a equipe de produção ouve o que é dito no palco, prepara os materiais de que os atores precisam e troca o pano de fundo que define o sentido do que é falado para a plateia. A metáfora está ilustrada na Figura 5.10.

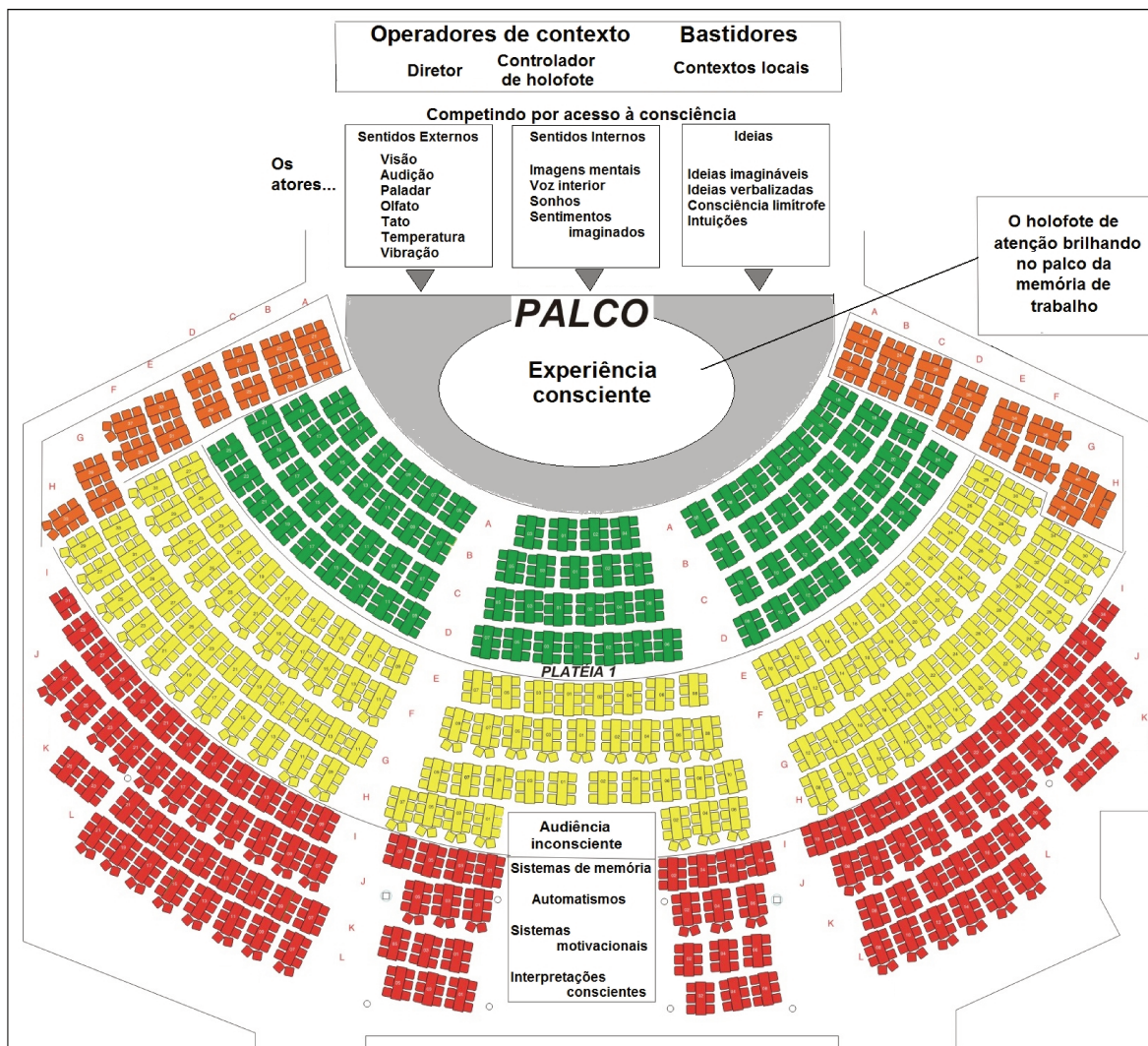


Figura 5.10: A metáfora do teatro da consciência

Na metáfora, o palco representa a memória de trabalho (MT) e a plateia corresponde aos processos cerebrais que respondem ao conteúdo da consciência (redes de neurônios distribuídas ao longo do cérebro e que desempenham funções inconscientes). Os holofotes simbolizam o mecanismo de atenção que destaca a informação a ser apresentada à plateia. Os contextos de interpretação são denotados pelo pano de fundo, que pode ser criado por redes semânticas que fornecem os possíveis referentes.

Assim como na IDA, no CTS os atores (representando processos cerebrais na forma de redes especializadas de neurônios) são denominados “codelets” e são “agentes simples e sem inteligência que simulam grupos neuronais”. E a exemplo das estruturas neuronais do cérebro, os codelets têm papéis especializados e são muito eficientes, a despeito de suas habilidades e alcance limitados. Isso significa que o CTS é um agente que contém uma estrutura multiagente (Dubois, 2007). Além disso, os codelets podem formar coalizões com outros codelets para atingir determinado fim. Os tipos de codelets implementados no CTS ou previstos para futuras ampliações são:

- **Codelets árbitros:** observam a MT e detectam quando uma informação clama por deliberação. Contam os ciclos desde o último enriquecimento da coalizão, selecionam a causa mais provável a ser acoplada a ela, declaram o fim da deliberação e tornam uma coalizão apta a entrar na competição para a seleção pelo mecanismo de atenção.
- **Codelets de atenção:** procuram por palavras específicas ou padrões de letras, prestam atenção a alguns tipos de informação, tentam localizar a presença de alguns padrões instantâneos de codelets ou constroem padrões temporais de eventos.
- **Codelets de expectativa:** de vida curta, são enviados após a execução de uma ação para verificar na MT o aparecimento de codelets de percepção/informação cujos conteúdos confirmem o efeito esperado. Se ocorreu, enviam energia de reforço ao nó que criou o efeito, se não, põem na MT um codelet de informação alertando sobre o problema.
- **Codelets de informação:** de vida curta, transportam uma ideia durante seu trânsito da e para a MT, representando-a lá até que ela seja publicada ou naturalmente abandonada. Formam associações que levam a novos conceitos, e tornam possível a deliberação por meio da qual uma ideia torna-se iterativamente mais rica ou inibida.
- **Codelets de comportamento:** em seu estado latente estão, junto com outros codelets de diferentes tipos, ligados ao nó de comportamento na rede de comportamento. Sabem como agir quando demandados (contatar a BD, apresentar material, etc.)

O nível de ativação de cada codelet de informação é um indicativo da importância da informação que este carrega e as transferências de energia entre agentes partem da hipótese fundamental de que a informação tem um valor. Codelets de informação quase sempre formam coalizões com outros codelets compatíveis. Essas coalizões são iniciadas com base em associações mútuas entre codelets, seja por serem vizinhos em um percepto, seja por associação espontânea na MT. Quando agrupados em coalizão, produzem um valor total de ativação que determinará se vão receber atenção. Dito de outra forma, na MT o nível de ativação dos codelets que se relacionam se soma de forma a favorecê-los na competição pela consciência. E assim como codelets criam coalizões, estas podem formar coalizões maiores, chamadas coalizões complexas (Dubois, 2007).

Assim como codelets podem formar coalizões, coalizões podem formar coalizões de coalizões, chamadas coalizões complexas, que são formadas pouco a pouco durante as deliberações, adicionando novos trechos de informação (codelets ou coalizões) e passos subsequentes até que a deliberação seja considerada completa pelo Codelet Árbitro.

O nível de ativação (energia interna) de várias entidades do CTS cresce com os estímulos vindos de fontes internas, de codelets que representam o ambiente ou com o passar do tempo. O fluxo de energia na rede de comportamento realiza uma importante parte do planejamento, sendo que a energia vinda de nós sensoriais indica objetivos/desejos do agente (planejamento *top down* e proativo) e a energia proveniente de estados ativados mantém a reatividade do CTS ao mundo externo (planejamento *bottom-up* e reativo).

Um Buffer Sensorial (BS) serve como interface de entrada para qualquer ator externo. O CTS usa uma gramática formal que descreve os elementos do ambiente e suas relações semânticas para traduzir as mensagens vindas do ambiente para um formato que possa ser processado pelos codelets. Um analisador sintático examina as mensagens recebidas e as posiciona em uma árvore hierárquica que é mais tarde inspecionada pelos codelets perceptuais em busca de informações reconhecíveis.

O CTS aplica o conceito de “espaço de trabalho global” (GW) na forma de um local único e ilimitado no qual todos os codelets se encontram quando precisam ser publicados. Isso corresponde ao palco na metáfora do teatro de Baars. É lá que as associações entre codelets são criadas e se fortalecem quando esses permanecem juntos. É para onde as coalizões são enviadas por todos os módulos, onde elas se combinam, se enriquecem e se opõem, e onde todos os codelets de atenção buscam informações. No CTS a MT permite a formação de vínculos entre codelets e coalizões que temporariamente a habitam.

No centro da arquitetura está a Consciência de Acesso, a qual publica, por meio de uma difusão (*broadcast*), a informação selecionada pelo mecanismo de atenção para torná-la acessível a todos os módulos (inconscientes). Esse mecanismo é crucial na colaboração das partes, por exemplo para chegar a um diagnóstico. A seleção da informação a ser publicada por sua importância para o contexto evita que muitos módulos se ocupem de informações menos importantes. Isso ajuda a direcionar o comportamento global do agente.

O aparato para produzir “consciência” consiste de um Gestor de Coalizão, um Controlador de Holofote e um Gestor de Difusão. O Gestor de Coalizão é responsável por formar e acompanhar coalizões de codelets. O Controlador de Holofote escolhe a coalizão com mais alta ativação total, e essa atenção conduz a coalizão à consciência. O Gestor de Difusão difunde o conteúdo da consciência a todos os codelets no sistema.

A Rede de Comportamento (BN) é ao mesmo tempo uma estrutura de decisão, um mecanismo de planejamento, e uma memória procedural de longo prazo. O planejamento que ela faz e as decisões que ela toma são inconscientes, mas dependem de difusões “conscientes” para seguir a evolução das situações. Isso é chamado de “seleção de ação conscientemente mediada”. A BN retém o repertório de conhecimentos do agente na forma de encadeamentos de nós de comportamento. Os encadeamentos são organizados sob nós motivadores e submetas. Cada nó de comportamento tem pré-condições necessárias e indica os efeitos que ele deve ter no ambiente, criando um vínculo natural com aqueles nós que têm esses efeitos como pré-condições. Efeitos e pré-condições existem como nós de estado

que interligam nós de comportamento, e para cada ligação de cima para baixo há outra de baixo para cima. Há também ligações inibidoras que partem de comportamentos em direção a outros comportamentos que poderiam desfazer as pré-condições daqueles.

O CTS adota uma versão modificada de BN, na qual cada comportamento é realizado por um grupo de codelets que, diferentemente dos nós de comportamento que não agem por si mesmos, são sempre ativos, atentos às difusões, e quando se consideram relevantes instanciam seu *template* de encadeamento como parte do nó de comportamento, e enviam ativação para seus nós respectivos. Quando escolhidos, os codelets são lançados com o nível de energia de seus nós e começam a fazer o que se espera deles: uns põem informação na MT, alguns solicitam informação, outros observam o que ocorre nos próximos ciclos para confirmar os efeitos esperados, etc.

A cada ciclo um comportamento é escolhido para ação. Se nenhum conseguir atingir o limiar necessário de energia, o limiar é reduzido pelo Gestor da BN no ciclo seguinte. Quanto mais alto for o limiar, mais longo será o planejamento da BN antes que um nó seja disparado, sendo assim explorados caminhos (planos) mais longos, com maior número de interações de influência mútua entre nós, criando um agente mais “prudente” e “analítico”.

O módulo chamado de Especialista no Domínio (DE) é um agente externo que é integrado à estrutura do agente consciente. Ele é uma entidade em três camadas contendo uma memória semântica (que indica qual câmara vê quais estruturas da ISS) e uma memória procedural sobre o domínio (que descreve os passos necessários antes de movimentar o braço robótico). Juntas elas codificam a *expertise* do CTS em relação ao domínio da aplicação e permitem ao CTS raciocinar sobre esse domínio.

A primeira camada do DE contém conhecimento descritivo, isto é, uma descrição física do mundo (simulado), seus elementos e inter-relações, e os conceitos que se referem àquele mundo, organizados na forma de uma ontologia. Isso permite o raciocínio lógico.

A segunda camada conecta a ontologia na primeira camada à descrição da memória semântica e descreve o conhecimento (correto e incorreto) que os aprendizes podem ter ao utilizarem o ambiente de ensino. O conhecimento é codificado a partir de uma perspectiva cognitiva na forma de memória semântica e procedural. Enquanto esta última descreve os conceitos segundo seu uso cognitivo (por exemplo, a intenção que o aprendiz pode ter com um conceito) e traz conhecimento didático específico do domínio, a segunda camada descreve os conceitos como parte de uma ontologia. O conhecimento procedural descreve como manipular o conhecimento semântico para realizar intenções.

A terceira camada define os objetivos de aprendizagem e abaixo desses organiza as duas outras camadas, empacotando-as na forma de unidades didáticas reusáveis.

O DE fornece informação em três situações:

- a) se ele ouve ou vê um pedido de informação que ele pode atender;
- b) se ele reconhece uma situação que requer um diagnóstico de primeiro-nível;
- c) se ele acredita que o aprendiz comete um erro.

Ele é informado das ações do aprendiz por publicações da consciência, e tenta reconhecer o plano e os passos seguidos pelo aprendiz (comparando com procedimentos armazenados no modelo de domínio). E quando o DE acredita que tem informação importante para comunicar, ele envia uma coalizão de codelets para a MT descrevendo o fato.

5.7 Transposição da arquitetura CTS para o ILTS

Como será discutido mais à frente, a transposição da arquitetura CTS para o caso do ILTS aqui descrito é em muitos aspectos bastante direta. Em primeiro lugar, por se tratar nos dois casos de tutores inteligentes, ambos preveem módulos de modelagem de aprendiz, de modelagem do domínio, entre outras similaridades. Além disso, como veremos abaixo, o ILTS adota a estratégia de usar codelets equivalentes aos utilizados no CTS, como Árbitro, de Atenção, de Expectativa, de Informação e de Comportamento. Mas tendo em vista o foco deste estudo, que é o modelo de domínio do ILTS, vai-se de início buscar uma implementação que tenha a mesma atribuição do módulo DE no CTS. Isso implica ter uma base de conhecimento na forma de uma ontologia da L2 que habilite o ILTS a raciocinar sobre esse domínio. Aqui há porém três diferenças significativas entre os dois:

- A amplitude muito maior de possíveis “ações” envolvidas no manejo de uma língua em relação ao manejo do braço robótico, cujo modelo de domínio deve conter dados sobre os sete graus de liberdade e sobre os arranjos recomendados de câmeras dependendo de qual será o movimento, enquanto que uma língua tem muito mais “graus de liberdade” no manejo de seus recursos, assim como muito mais restrições;
- O público-alvo do CTS é composto por um seletivo grupo de astronautas cujo treinamento é intenso, constante e acompanhado de perto por especialistas no domínio, ao passo que no ILTS o público-alvo é muito mais variado e numeroso e encontra-se disperso, sem contato direto com os especialistas do domínio;
- As implicações de possíveis erros no manejo do braço robótico são, em princípio, mais graves que eventuais erros dos alunos no uso da L2, pois os primeiros podem implicar grandes prejuízos materiais ou riscos à vida dos astronautas que porventura estejam realizando tarefas nas proximidades do braço robótico.

Assim, em razão dessas significativas diferenças em seu contexto e em sua área de aplicação, o ILTS aqui proposto teve que lidar com a considerável complexidade de se modelar o manejo da L2. A solução encontrada para atingir esse objetivo foi o uso de uma modelagem multiagentes atribuindo um agente específico a cada lexema da língua-alvo, pois, como já discutido, as aplicações ABMS são particularmente indicadas para problemas cuja complexidade não permite soluções baseadas em outras arquiteturas. E no contexto da arquitetura baseada em codelets esses agentes-lexemas podem ser vistos como codelets especializados (para o domínio das línguas) e por isso foram aqui chamados de codelexes.

O paralelismo entre as duas arquiteturas ocorre também no que diz respeito às coalizões, que no ILTS tem a função de, dadas certas condições de contexto e de estratégia didática, permitir que lexemas isolados se agrupem para gerar sintagmas e sentenças da L2. Além disso, assim como no CTS, no ILTS o primeiro nível do domínio do especialista contém conhecimento descritivo sobre o objeto do ensino, organizados na forma de uma ontologia com as relações semânticas entre os lexemas da L2.

No ILTS a segunda camada do DE codifica o conhecimento a partir da perspectiva semântica das imagens de fundo das cenas do curso. Enquanto a primeira camada descreve os conceitos dentro de uma ontologia, a segunda camada descreve os conceitos segundo seu contexto de uso e traz conhecimento didático específico do domínio.

O ILTS também tomou como referência a aplicação descrita em Silva (2009), uma criatura virtual (veículo autônomo) com consciência artificial baseada na arquitetura Baars-Franklin. Naquela implementação, o veículo deve se deslocar num espaço bidimensional buscando atingir um certo objetivo, mas tendo que constantemente monitorar o nível de energia em sua bateria e, no caso de nível baixo, buscar um ponto de recarga. Além disso, alguns obstáculos situados sobre o espaço bidimensional impedem a passagem do veículo, e devem ser evitados (contornados). O veículo tem um conjunto de sensores (de posição, visual, de contato e de carga da bateria) e de atuadores (velocidade, ângulo da roda dianteira e posição do sensor visual). A distribuição de tipos de codelet, assim como a atuação dos gerenciadores de coalizões, de foco de luz e de *broadcast*, estão ilustradas na Figura 5.11.

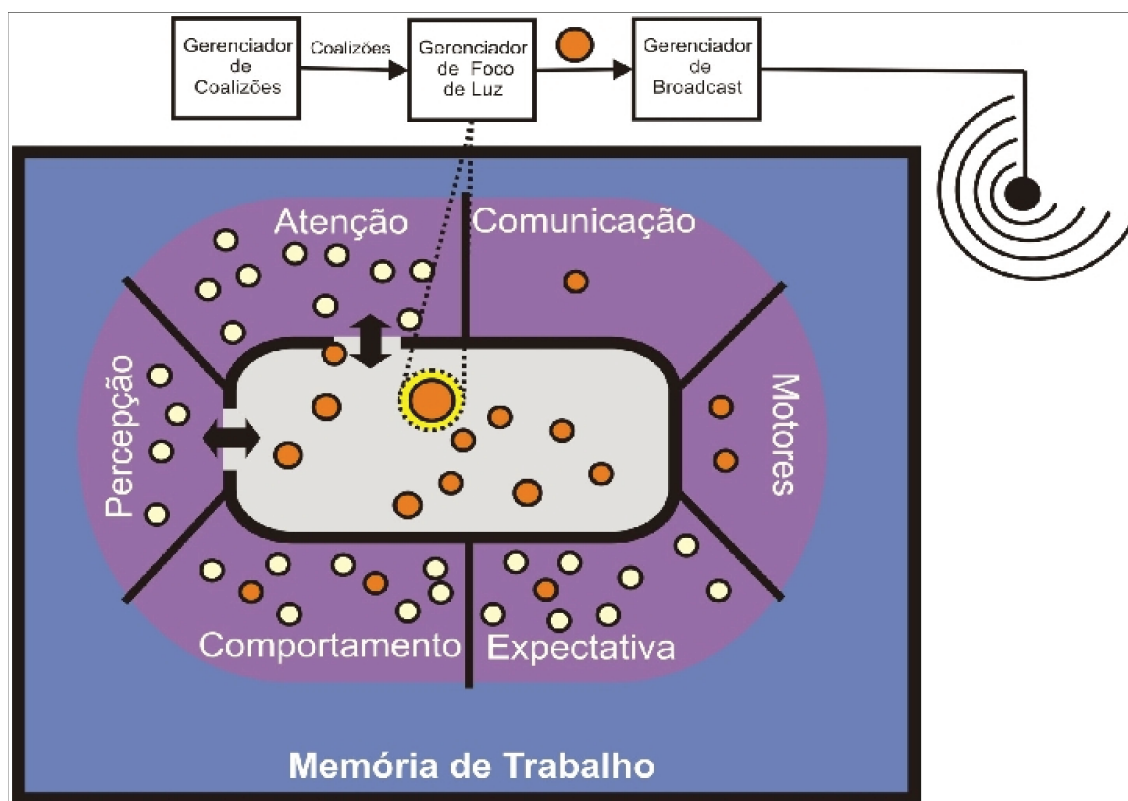


Figura 5.11: Segmentação dos codelets na implementação do veículo autônomo

A arquitetura do ILTS aqui proposta tem semelhanças e diferenças em relação à do veículo autônomo. Uma das diferenças é que, em vez de tentar conduzir o veículo a um local, a meta do gestor de conteúdos do ILTS é montar sentenças autonomamente a partir de um léxico pertinente a um dado contexto. Nesses termos, montar uma sentença equivale a traçar um percurso no espaço multidimensional da L2.

Na Figura 5.12 está ilustrada a rede de comportamento do módulo de gestão de conteúdos do ILTS. Ela se ocupa da montagem das sentenças na tela do curso e cuida para manter o nível de energia no sistema. As elipses representam nós de comportamento, os

triângulos representam as proposições adicionadas à BN, os retângulos são os codelets de atenção que interferem no estado da rede e os hexágonos são os objetivos da rede. As setas contínuas pretas representam os *links* predecessores e sucessores, enquanto que as setas tracejadas vermelhas mostram os *links* conflitantes (inibidores).

Diferentemente do veículo virtual autônomo implementado em Silva (2009), no módulo de gestão de conteúdos do ILTS o nível de energia não se refere à autonomia do veículo, mas à disponibilidade de conteúdos para a formação de novas sentenças. Isso é relevante porque à medida que várias sentenças vão sendo geradas para uma dada cena, o léxico inicialmente selecionado começa a se esgotar ou se repetir excessivamente. Isso deve motivar duas ações: a primeira é uma consulta à base de conhecimento (ontologia) para verificar se há alguma informação adicional sobre o lexema-foco da cena e, caso haja, trazê-la para fazer parte de novas sentenças; a segunda é, em último caso, se já não há mais informações relevantes na ontologia, é então o momento de substituir a imagem de fundo, dando início a uma nova cena e um novo ciclo de apresentação de léxico.

A necessidade de montar uma consulta à ontologia decorre então da verificação de que o estoque de lexemas está baixo e de que o plano didático prevê mais exemplos de conteúdos para o contexto em questão. Uma vez montada, a consulta é realizada e novos dados são trazidos ao módulo de gestão de conteúdos e passam a enriquecer a formação de sentenças como descrito abaixo.

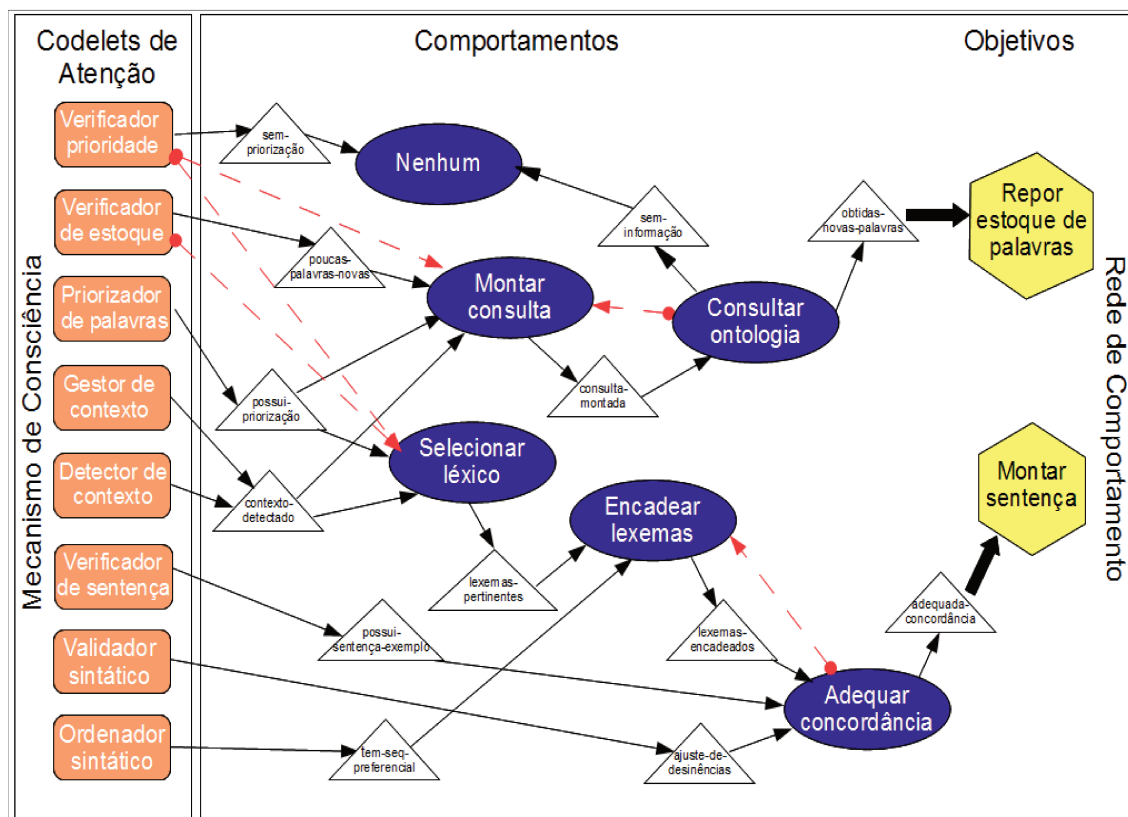


Figura 5.12: Rede de comportamento e mecanismo de consciência

Sequência de formação das sentenças

Uma vez escolhida uma imagem de fundo da cena e, conseqüentemente, definido o contexto para sua interpretação, o módulo de gestão de conteúdos mobiliza, com base nos metadados, uma lista inicial de lexemas pertinentes, substantivos e adjetivos, associados às regiões da imagem às quais esses se referem, como sugerido na Figura 5.8. Mas a geração de sentenças levará em conta três camadas (complementares) de informação lexical. A primeira é a associada aos metadados da imagem e contém, além do lexema-foco, vários lexemas acessórios: substantivos e adjetivos que refletem o contexto ilustrado na imagem. Esses metadados são a primeira referência na formação das sentenças da cena, recrutando-se com base nela os agentes-lexemas correspondentes (que aguardavam na plateia do teatro da consciência junto com todos os demais agentes). A segunda camada de informação lexical é transportada por cada agente-lexema, e corresponde às listas de possíveis verbos, substantivos, adjetivos com os quais eles em geral se articulam (esses dados vêm de registros de uso da L2). A terceira e última camada é a de informações lexicais provenientes da ontologia, e que provê dados complementares de natureza factual ou resultantes de inferências feitas a partir da base de conhecimento. As três camadas estão ilustradas na Tabela 5.2: as duas primeiras colunas referem-se aos substantivos e adjetivos que são anotados na imagem pelos especialistas do domínio, tendo em vista seu uso didático. Essas informações têm, em geral, uma natureza mais contextual, específica do ente mostrado ou de sua condição momentânea ou circunstancial. Na terceira coluna são exemplificados os dados transportados pelos próprios lexemas-agentes, e que dizem respeito às relações semanticamente mais usuais na L2 (por exemplo, “rinoceronte ataca”). Por fim, a quarta coluna traz informações adicionais inferidas pela ontologia para o lexema-foco.

Tabela 5.2: Fontes de léxico para a formação das sentenças-exemplo

Metadados da Imagem		Dados dos Agentes	Ontologia da L2
Substantivo	Adjetivos	Verbo	Relações/Fatos Inferidos
rinocero	[granda, sovağa, peza]	[estas, havas, manğas, atakas]	mambesto, besto, vivanto,
korno	[pinta]	[estas, havas, batas, boras]	—
piedo	[larğa]	[estas, havas, premas, stamfas]	—
bušo	[fermita]	[estas, havas, mordas, maças]	—
orelo	[stranga]	[estas, havas, aũdas, aũskultas]	—
vosto	[mallonga]	[estas, havas]	—
ventro	[dika]	[estas, havas]	—
dorso	[malpura]	[estas, havas]	—
okulo	[malgranda]	[estas, havas, rigardas]	—
nazo	[longa]	[estas, havas, flaras]	—
kapo	[granda]	[estas, havas, batas]	—
kruro	[forta]	[estas, havas]	—

A formação das sentenças-exemplo começa então com o lexema-foco, por exemplo, *rinocero* e com os verbos mais usuais *estas* (é) e *havas* (tem), contidos nos campos de dados do próprio lexema-foco, como ilustrado na terceira coluna da Tabela 5.2. O primeiro verbo é de ligação e deve ser seguido por adjetivos que qualifiquem o lexema-foco ou por substantivos que o definam:

rinocero estas granda (rinoceronte é grande),
rinocero estas sovaĝa (rinoceronte é selvagem),
rinocero estas peza (rinoceronte é pesado).

O verbo *havas* (tem), por sua vez, é transitivo e pede como complementos (objetos diretos) os substantivos referentes a coisas (como partes de seu corpo) que o lexema-foco possua:

rinocero havas kornon (rinoceronte tem chifre),
rinocero havas orelon (rinoceronte tem orelha),
rinocero havas ventron (rinoceronte tem barriga), etc.

Esses substantivos e adjetivos vêm da primeira camada de informação, aquela contida nos metadados, como ilustrado na Tabela 5.2. Em um segundo momento, após varrer todos os substantivos disponíveis, a sentença-exemplo passa a ser enriquecida pelos adjetivos que qualificam os lexemas acessórios:

rinocero havas pintan kornon (rinoceronte tem chifre pontudo),
rinocero havas strangan orelon (rinoceronte tem orelha estranha),
rinocero havas dikan ventron (rinoceronte tem barriga gorda), etc.

Por fim, quando o estoque de possibilidades das duas primeiras camadas de dados se esgota o gestor de conteúdos pode recorrer à terceira camada de dados, isto é, a dados contidos na ontologia, e referentes a fatos e relações de pertinência que podem ser inferidas dela, como ilustrado na quarta coluna da Tabela 5.2:

rinocero havas kvar piedojn (rinoceronte tem quatro pés),
rinocero havas du orelojn (rinoceronte tem duas orelhas),
rinocero estas mamulo (rinoceronte é mamífero),

É importante notar que após concatenada a sentença devem ser feitas as adequações dos lexemas às regras morfológicas de cada L2, como nas concordâncias de caso, gênero e número. Nos exemplos acima as sentenças mostradas já passaram pela etapa de adequação de concordância de forma que as desinências de plural *-j* e do caso acusativo *-n* foram acopladas aos lexemas originais, de acordo com as regras gramaticais da L2.

É importante ressaltar também que os codelexes são inicializados com um certo estoque de lexemas correlatos com os quais as primeiras sentenças válidas podem ser formadas. Contudo, uma vez usadas essas combinações, atinge-se um nível baixo no estoque de palavras disponíveis, e deve-se recorrer a novas consultas à ontologia.

O objetivo a ser atingido é então produzir sentenças-exemplo condizentes com o contexto, sempre de acordo com a ordem sintática predominante na L2 e, se necessário, enriquecidas por informações recém obtidas de consultas à ontologia. No Capítulo 6 será ilustrada uma simulação parcial desse mecanismo.

5.8 Configuração esquemática da arquitetura proposta

Como discutido nas três seções precedentes, a arquitetura proposta baseia-se em três elementos: uma ontologia do léxico da L2, a modelagem multiagente e o uso de imagens de

apoio com descritores semânticos baseados nesse léxico. A integração dos três elementos é vista na Figura 5.13. Nela estão o aplicativo para a ontologia, o de criação de metadados (descritores semânticos) para as imagens e o de organização dos dados internos dos agentes. Essas informações alimentam um simulador multiagentes, um *reasoner* da ontologia, e um aplicativo que interpreta os metadados. Os aplicativos são integrados na retaguarda da interface do curso, com a qual os alunos interagem.

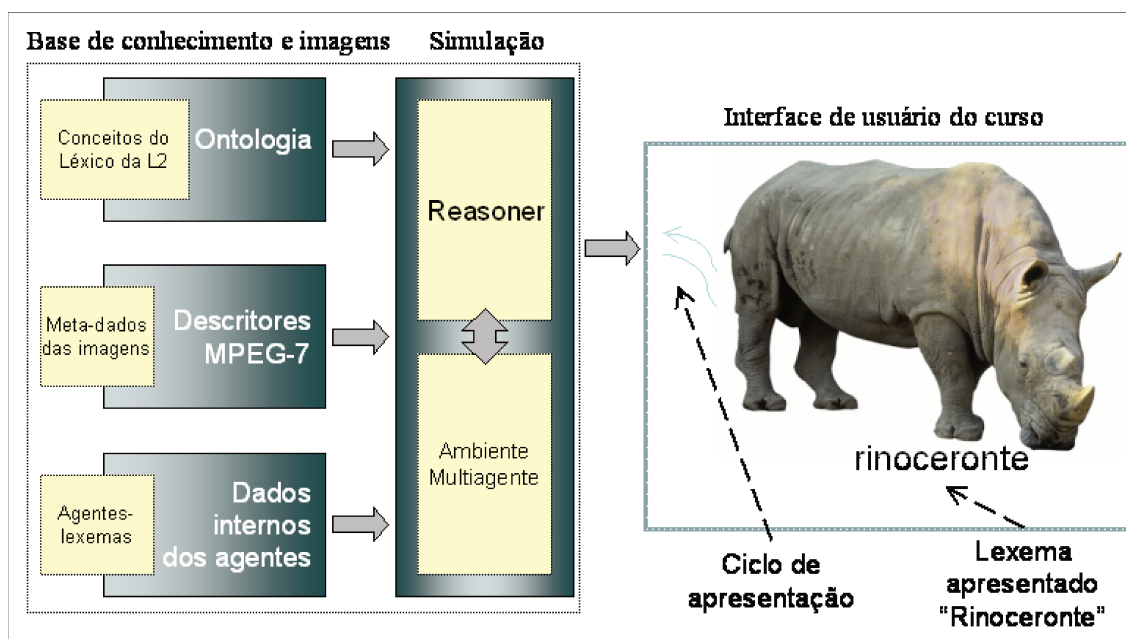


Figura 5.13: Síntese da arquitetura proposta para o ILTS

Os blocos funcionais da arquitetura proposta estão ilustrados na Figura 5.15. As etapas de preparação do ambiente de ensino são indicadas pela numeração nas setas que conectam os blocos. A primeira etapa consiste na delimitação do subconjunto do léxico da L2 que será objeto do curso e a transposição desse léxico-alvo para a planilha de dados em que se basearão os agentes-lexemas. A segunda etapa também parte do glossário da L2, mas seu objetivo é contrastar o léxico-alvo da L2 com os lexemas correspondentes na(s) L1(s) do público-alvo do curso (etapa 3), para, na etapa 4, usar os resultados dessas métricas de iconicidade intersimbólica e preencher o respectivo campo na planilha de agentes-lexemas.

Até esse ponto os especialistas da aplicação e do domínio utilizaram como dados de entrada glossários da L2 e das L1s. Mas tão importante quanto mapear o léxico de interesse em fontes normatizadoras da língua é pesquisar o uso real da língua e daí extrair as relações mais frequentes entre lexemas. Uma fonte rica de registros do uso real da L2 é a Web. Um motor de buscas como o Google permite direcionar a busca por páginas para uma língua específica. No caso particular do esperanto, L2 de nosso estudo de caso, o Google localizou mais de quatro milhões de páginas (no segundo semestre de 2010), das mais variadas naturezas, como blogs, listas de discussão, Wikipédia em esperanto, literatura original e traduzida, etc., e, como sugerido pela Figura 5.14, variado em termos de origem geográfica

e de época de produção. Esse vasto acervo de textos permite verificar como os lexemas se articulam no uso real da L2.

Contudo, tendo em vista que o léxico-alvo no estudo de caso totaliza três mil itens, torna-se inviável empreender essas buscas de forma manual. Em razão disso, na etapa 5 foi utilizado um aplicativo, denominado GugloKvek³, que é capaz de realizar essas buscas de forma mais automatizada (para o esperanto) e retornar os tipos mais comuns de relação (à esquerda, e à direita) do lexema em questão. Os resultados dessas pesquisas são então usados para preencher os campos pertinentes nas planilhas dos agentes-lexemas (etapa 6).

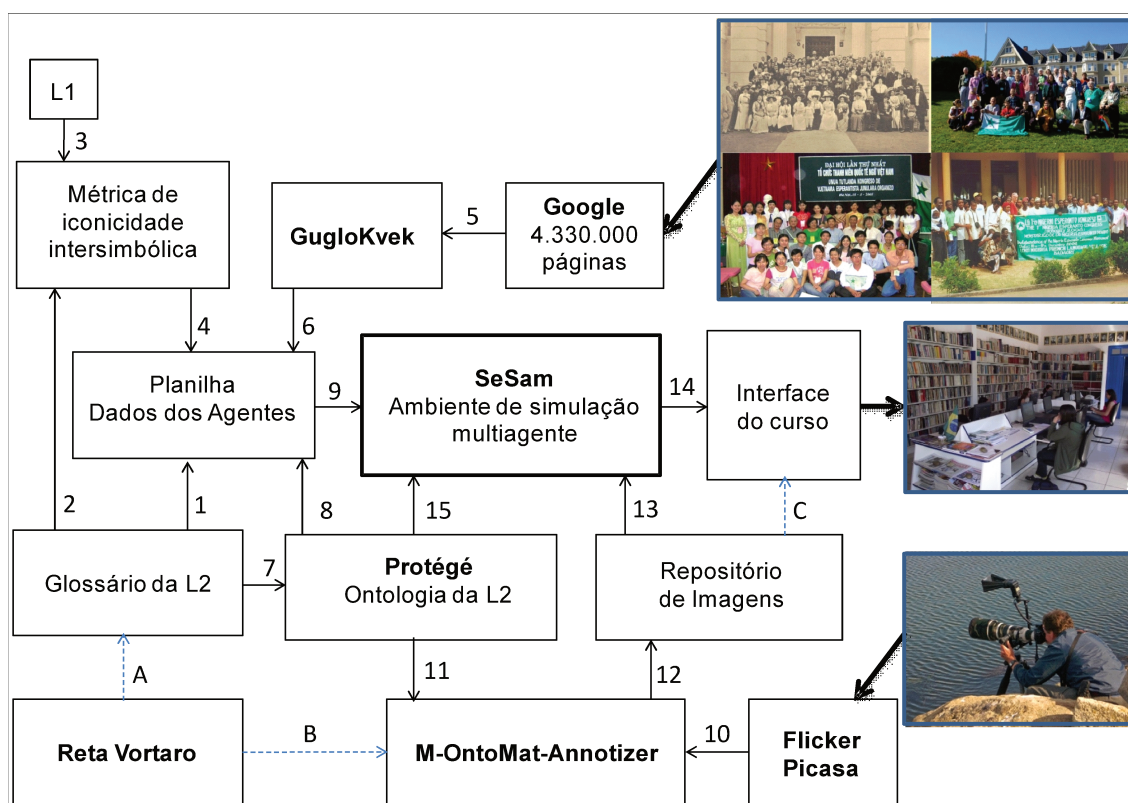


Figura 5.14: Diagrama de blocos do ILTS e sua relação com conteúdos da Web

A etapa 7 refere-se à criação da ontologia da L2 a partir de seu glossário e de outras fontes de informações (semânticas, sintáticas e pragmáticas/factuais) sobre cada um dos lexemas visados. Uma vez criada, a ontologia é usada durante a inicialização do ILTS para preencher nas planilhas dos agentes os campos sobre as relações entre lexemas.

Na nona etapa os agentes-lexemas das planilhas são importados para dentro do ambiente de simulação multiagente onde permanecem ativos mas ocultos, até que sintam que podem se manifestar em resposta a uma difusão feita pelo módulo central, engajando-se então em um processamento de seleção de ação.

³ Desenvolvido por Toon Witkam, pesquisador de linguística computacional e membro do Instituto Internacional de Esperanto (Haia, Holanda).

Na décima etapa as imagens de apoio do curso são selecionadas em repositórios da Web, como o Flickr e o Picasa, e nomeadas de forma que possam ser acessadas pelos agentes durante o funcionamento do curso.

A ontologia é acessada novamente durante a etapa 11 (que corresponde à fase de anotação semântica das imagens de apoio) pelo aplicativo M-OntoMat-Annotizer, conforme discutido na Seção 5.7. Na etapa 12 as imagens de apoio selecionadas e anotadas com os descritores semânticos são armazenadas no repositórios do ILTS, para, na etapa 13, serem requisitadas pelo ambiente de simulação multiagente sempre que forem necessárias à cena que se deseja apresentar ao aluno. A etapa 14 corresponde então à composição da interface por meio da qual o aluno interagirá com o curso e será apresentado ao léxico da L2.

Durante o funcionamento do ILTS, a ontologia pode novamente ser consultada (etapa 15) para obtenção de dados não explícitos, por meio de um mecanismo de inferência via aplicativo *reasoner*.

As setas indicadas pelas letras A e B indicam situações em que o especialista na L2 pode recorrer a um dicionário da L2 na Web, denominado Reta Vortaro, para assim enriquecer as definições e informações utilizadas tanto na planilha dos agentes-lexemas (A) quanto no processo de anotação semântica das imagens (B). Por fim, a seta indicada pela letra C indica a possibilidade de a interface do curso acessar as imagens diretamente no repositório de imagens, sem para isso depender dos recursos do ambiente ABMS.

Uma vez implementada a arquitetura, seu uso passa a ser mais simples. A função de especialista passa a ser a de manter as planilhas de dados dos lexemas, com a eventual inclusão de novos itens, manter a ontologia da L2, com eventual inserção de novas informações e acompanhar o desempenho dos alunos com base em seus modelos.

A proposição desta arquitetura foi a meta principal deste trabalho, e, espera-se, sua contribuição mais significativa. Os estudos subsequentes terão como objetivo implementá-la na sua totalidade e testar todos os cenários possíveis de uso, com usuários de diferentes origens culturais. Contudo, ainda que não tenha sido objetivo nesta etapa realizar toda a implementação da arquitetura, partes de seu princípio de funcionamento podem ser testadas por meio de simulações. Esse é então o objetivo do Capítulo 6.

Capítulo 6

Simulação ABMS da gestão de conteúdos

Neste capítulo será feita uma breve descrição de alguns cenários de simulação a que o mecanismo de gestão de conteúdos baseado em seleção de ação em ambiente multiagente vem sendo submetido. O objetivo dessas simulações é inicialmente fazer uma prova de conceito do elemento central da arquitetura, representado pelo ambiente para modelagem ABMS. Todavia, os cenários exemplificados aqui não esgotam as possibilidades que estão sendo vislumbradas para a arquitetura em sua configuração completa. Mas, como a total implementação do ambiente, tal como esquematizado na Figura 5.15 do capítulo anterior, demandará tempo e esforço maiores que os viáveis nesta primeira etapa (mestrado) da pesquisa, os cenários mais elaborados serão objeto dos estudos subsequentes. No trabalho atual foi dada ênfase a cenários factíveis dentro de uma simulação no ambiente ABMS.

6.1 Condições de realização das simulações

As simulações foram feitas em um aplicativo específico para modelagem ABMS, denominado SeSAM (Shell for Simulated Agent Systems), cuja interface está ilustrada na Figura 6.1. O SeSAM tem *plugins* para integração como o aplicativo de ontologias Protégé. Nas simulações as informações provenientes da ontologia (por exemplo, a lista de lexemas com os quais um lexema se articula) foram escritas nos campos da planilha dos agentes-lexemas e importados pelo SeSAM no início da simulação, como ilustrado na Figura 6.2.

Como discutido ao longo deste trabalho, a arquitetura de referência que serviu de base para o ILTS fundamenta-se, por sua vez, no conceito de seleção de ação envolvendo codelets, tipos de agentes simplificados mas capazes de interagir para buscar soluções para uma dada situação ou um determinado contexto. Assim, nas simulações aqui descritas usou-se uma imagem para estabelecer um contexto didático, uma cena na qual um elemento principal e vários elementos complementares prestam-se a embasar o aprendizado de vários itens léxicos da L2. E, como discutido na Seção 5.4 do capítulo anterior, a arquitetura do ILTS proposto baseia-se também no uso de metadados para a semantização das figuras de apoio do curso, a fim de permitir aos lexemas-agentes reconhecer os conteúdos dessas imagens. Isso torna possível reproduzir fielmente a ideia central da metáfora do teatro da consciência, na qual o pano de fundo da cena define o contexto com base no qual cada processo cerebral (representado no tutor por um codelet) avalia se sua participação na cena é ou não pertinente e oportuna. Nesse sentido, o primeiro cenário simulado partiu de uma imagem de fundo do ente “rinoceronte” e dos metadados que descrevem seus principais elementos. Esse contexto deveria então ser suficiente para os agentes pertinentes se darem conta de que são necessários e, após se deslocarem até as áreas da figura à qual cada um se refere (Figura 6.3), ali exibirem o rótulo correspondente àquele lexema na L2 (Figura 6.4). O lexema-foco da imagem, que é informado nos metadados, é destacado dos demais por meio de letras maiores. Uma vez atingido esse ponto o lexema-foco dá início à construção de uma sentença, deslocando-se para uma posição no alto da imagem. Tem aí início um processo de seleção de ação no qual os demais lexemas disputam a oportunidade de tomar

parte na formação da sentença, até que um deles tenha êxito (Figura 6.5). Isso transcorre com a troca de energia de ativação entre os agentes-lexemas, até que um deles, neste caso referente ao lexema *piedo* (pé), saiu vitorioso e teve acesso à sentença com o lexema-foco. O nível de ativação do agente vitorioso é visto na Figura 6.6. Após exibido, ele volta à sua posição anterior, junto ao elemento correspondente na imagem, e deixa o lugar na sentença para outro agente-lexema, e assim sucessivamente, até que o último, representando o lexema *dorso* (dorso) conseguisse um lugar na sentença exemplo (Figura 6.7).

É importante mencionar que este cenário de simulação somente buscou verificar se e como se daria esse processo de sequenciamento autônomo dos itens lexicais, sem ainda dar a esse mecanismo critérios efetivamente didáticos, como, por exemplo, favorecer os itens mais frequentes ou mais parecidos com a língua materna do usuário (hipotético). O cenário visou tão somente testar a interoperação do mecanismo de seleção de ação baseado em multiagentes com os dois outros pilares propostos para a arquitetura: os dados da ontologia e os metadados da imagem de apoio. Nesse ponto os resultados alcançados são promissores, mas a real validação da arquitetura proposta e a identificação de eventuais limitações vão demandar uma implementação mais ampla, como já discutido.

Deve-se mencionar que nesta primeira simulação não foram utilizados recursos de adequação (flexão) dos lexemas para refletir a sua função sintática na sentença, tendo em vista as normas gramaticais da L2. No exemplo da sentença simulada, *rinocero havas piedo* (rinoceronte tem pé), a gramática do esperanto prevê o uso da desinência “n” para marcar o caso acusativo (no objeto direto). Assim, a sentença correta deveria ser *rinocero havas piedon*. Isso significa que na implementação completa deverão ser previstos agentes com a função específica de fazer ajustes como desinências, artigos, maiúsculas, etc.

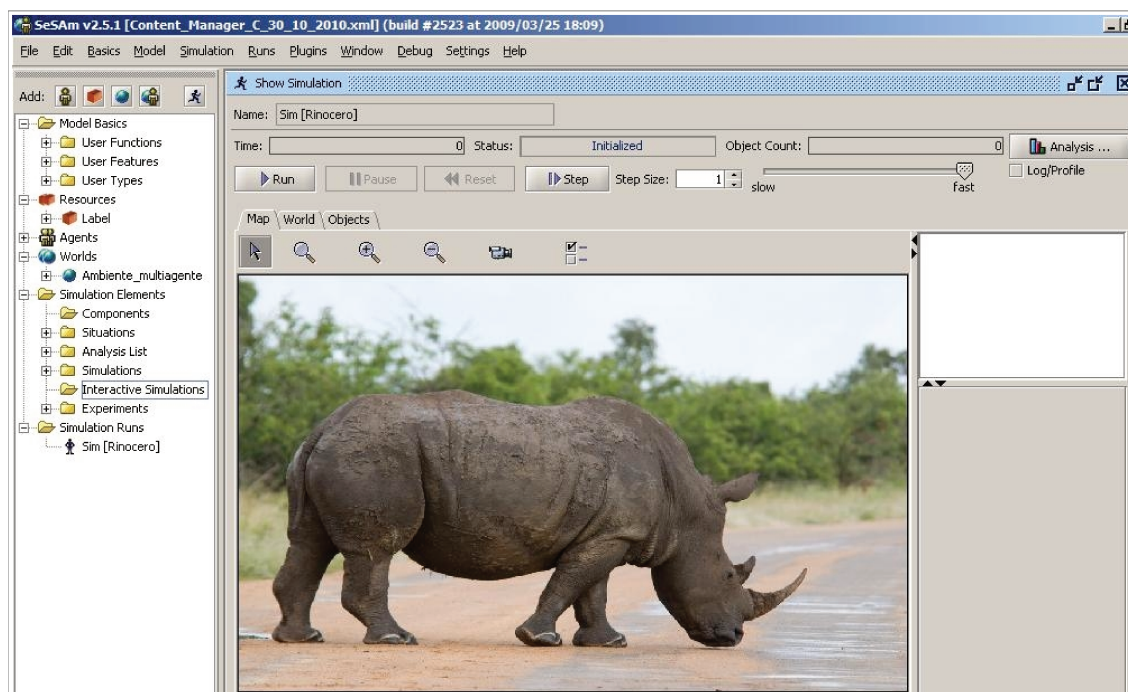


Figura 6.1: Condição inicial: pano de fundo define o contexto da cena

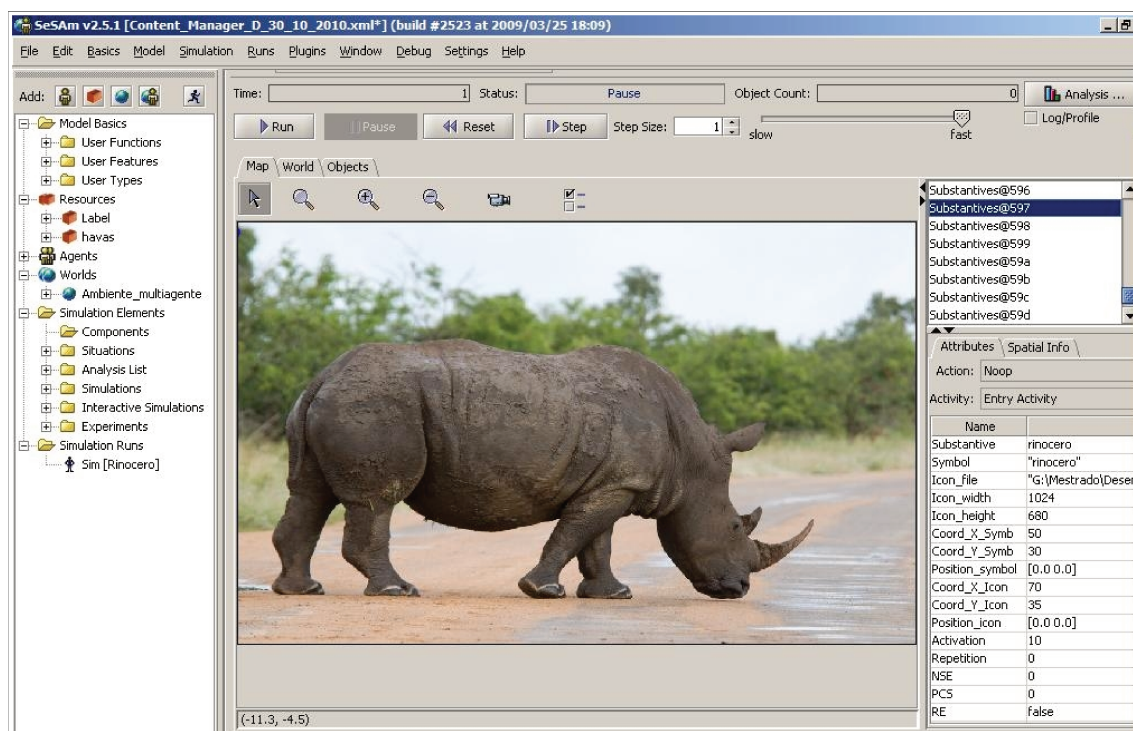


Figura 6.2: Agentes importados permanecem ativos mas ocultos

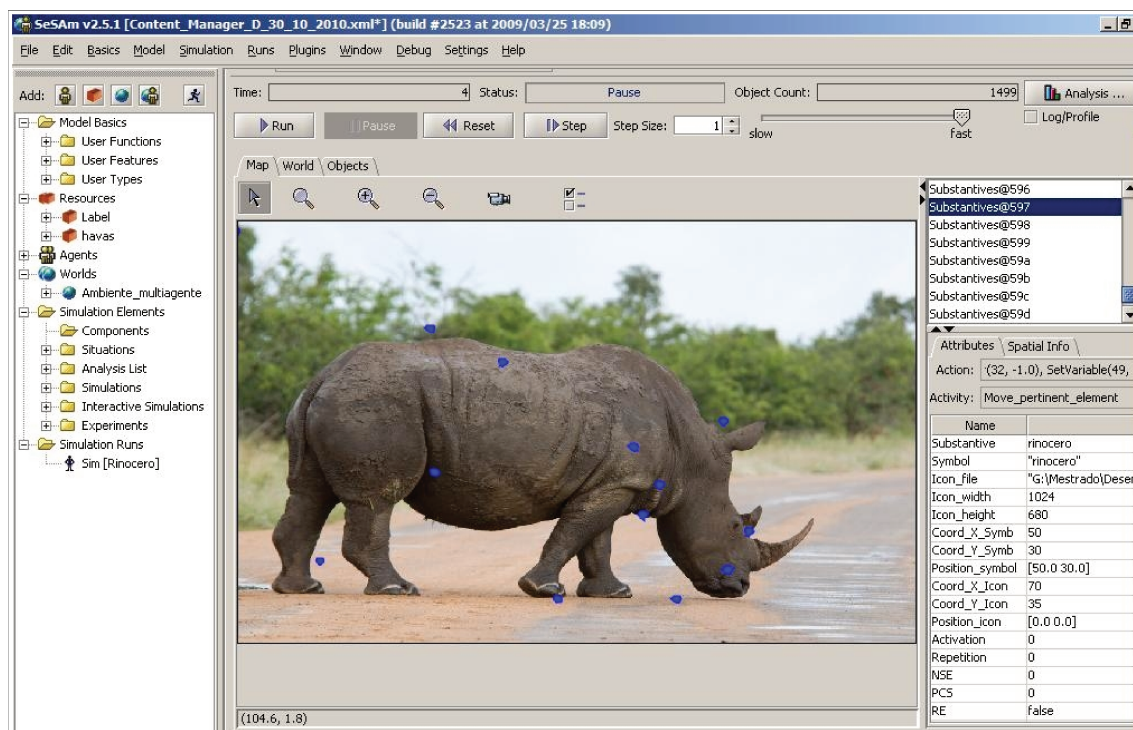


Figura 6.3: Agentes importados deslocam-se aos elementos indicados pelos metadados

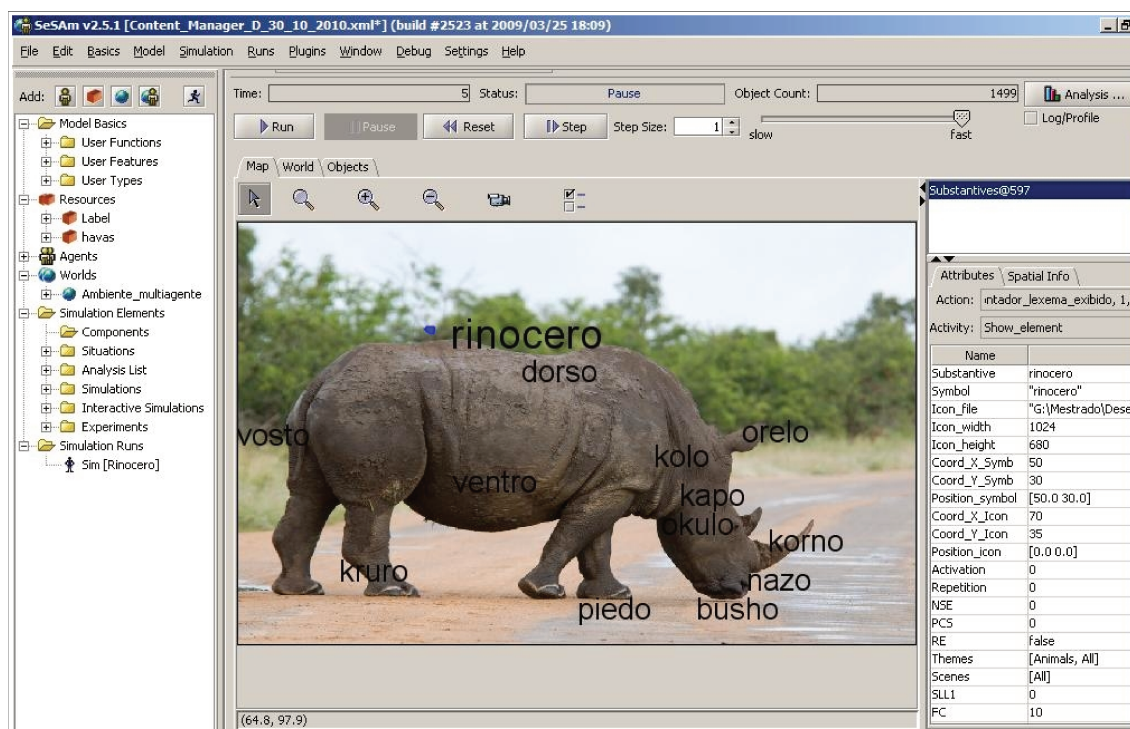


Figura 6.4: Processo de rotulagem dos elementos, destaque para o lexema-foco

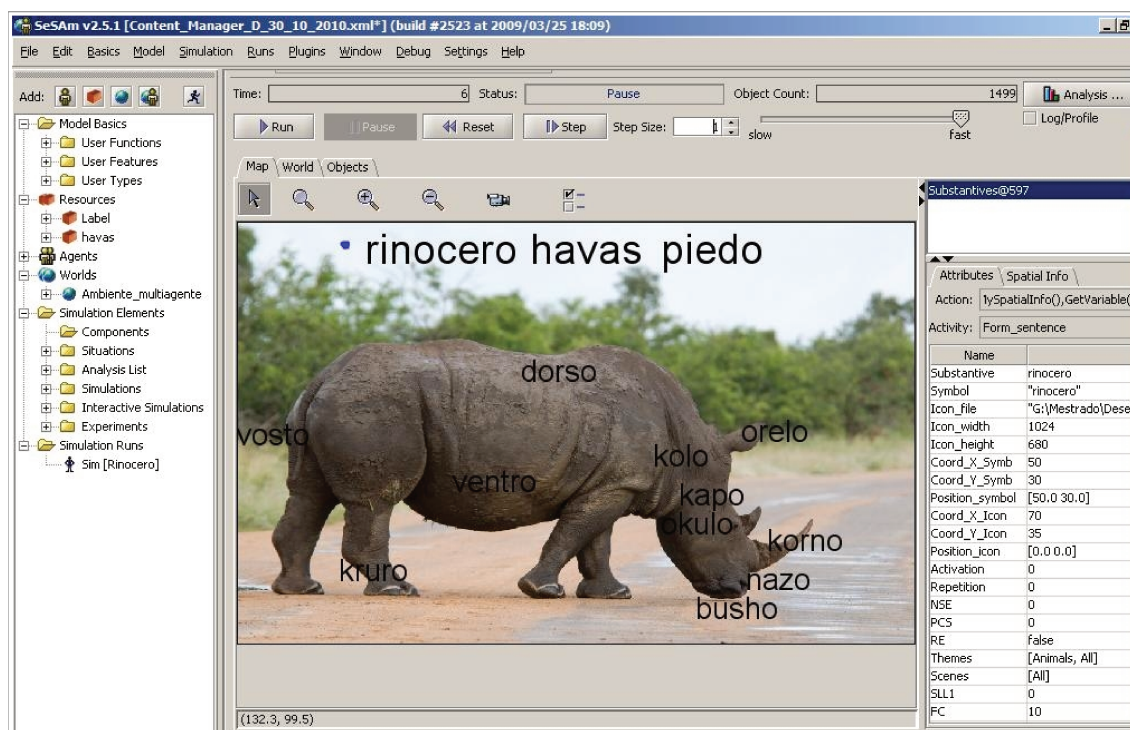


Figura 6.5: Lexema-foco inicia sentença e um dos lexemas correlatos junta-se a ela

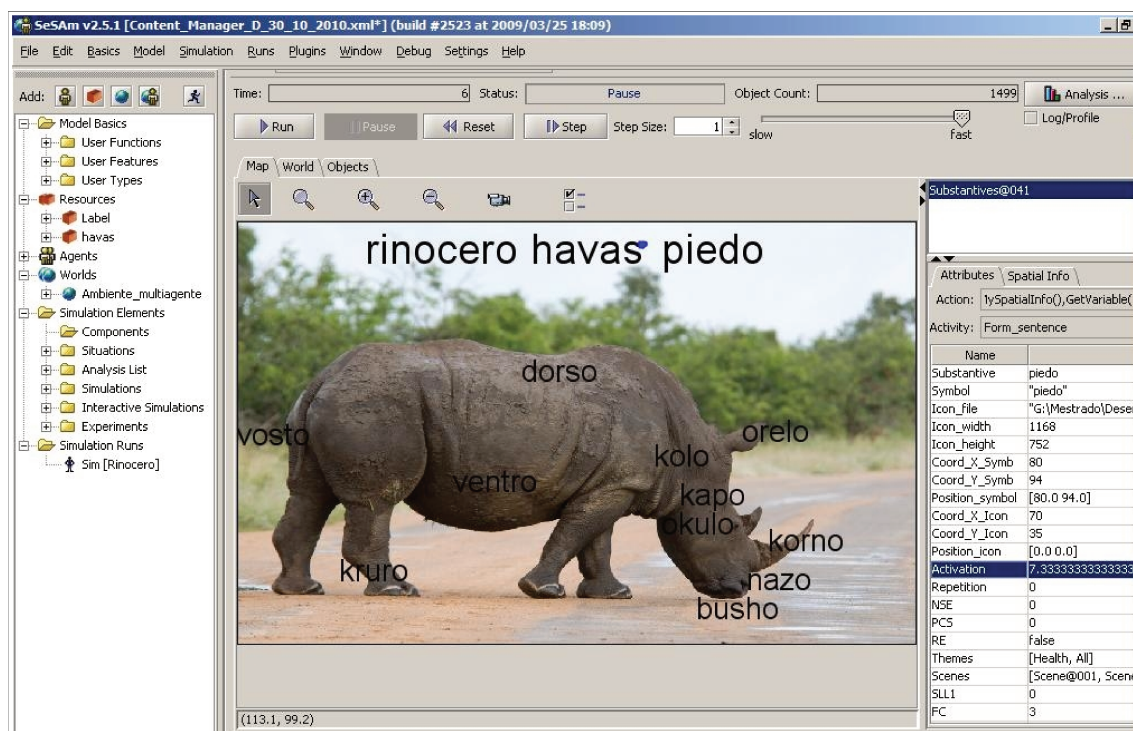


Figura 6.6: Nível de ativação do agente-lexema vencedor da primeira rodada

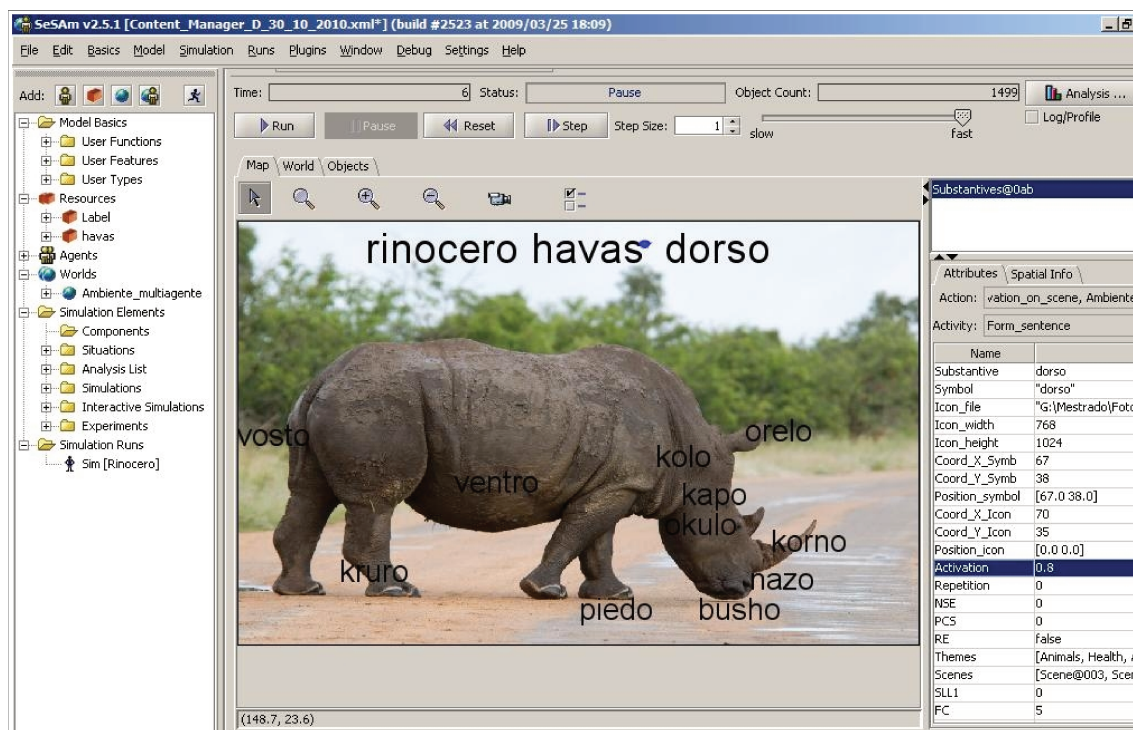


Figura 6.7: Nível de ativação do agente-lexema vencedor da última rodada

6.2 Uso de imagens de apoio baseadas em desenhos

Embora o uso de fotos como imagens de apoio tenha vantagens do ponto de vista da iconicidade em relação aos objetos no mundo real, isso tem duas desvantagens. A primeira refere-se à carga indicial de uma foto, que em geral está associada a um ponto no espaço e no tempo ou a um indivíduo mais que a uma classe de indivíduos. A outra é a dificuldade de manipulação dos elementos de uma foto, isto é, à pouca flexibilidade de uma foto para que suas partes sejam rearranjadas ou de algum modo alteradas (cor, tamanho, etc.).

Uma alternativa a essa limitação está no uso de cenas baseadas em arranjos de desenhos. Embora imagens baseadas em desenhos sejam menos icônicas que fotos, elas têm baixa carga indicial (referem-se muito mais à classe do que a um indivíduo em particular) e, sobretudo, dão mais flexibilidade à manipulação da cena quando essa é formada por vários desenhos organizados espacialmente. Um exemplo disso é a cena de um quarto de dormir, no qual podem estar presentes objetos como cama, poltrona, armário, janela, porta, lustre, entre outros. Assim, a cada um desses objetos corresponderá um agente-lexema do ILTS. Como exercício dessa alternativa foi também implementado um arranjo de cena baseado em agentes-lexemas ilustrados por desenhos de objetos. Na Figura 6.8 é mostrada a etapa da simulação em que o elemento *lito* (cama) foi selecionado.

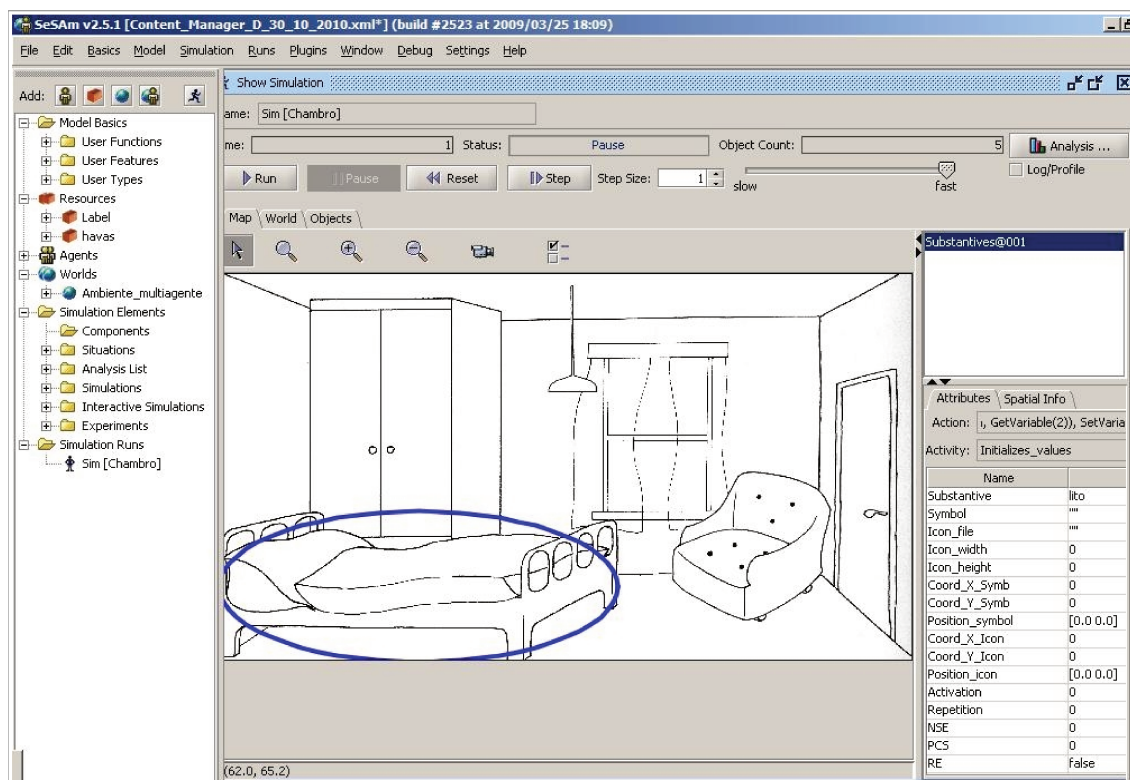


Figura 6.8: Exemplo de cena baseada em desenhos - destacado o agente “lito”

Na Figura 6.9 é ilustrada a etapa em que foi selecionado o agente *fenestro* (janela).

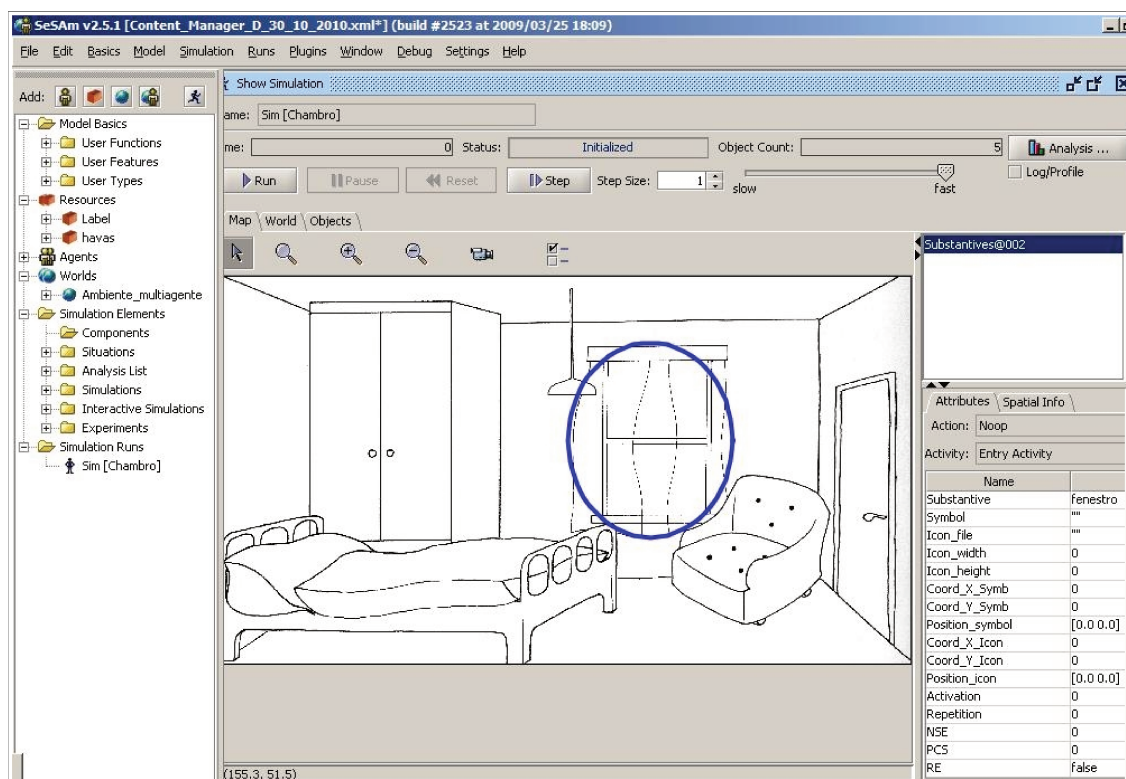


Figura 6.9: Exemplo de cena baseada em desenhos - destacado o agente “fenestro”

Assim, um modo de interação no qual os alunos podem mover os objetos na cena e com isso produzir uma alteração correspondente na sentença que a descreve é, em tese, possível na arquitetura proposta, visto que a partir das coordenadas posicionais de cada agente é possível inferir sua situação espacial em relação a outros agentes na mesma cena. Isso é particularmente útil no caso de preposições e expressões que indicam contexto ou lugar, como “sobre”, “sob”, “ao lado de”, “perto de”, “à direita de”, “entre”, etc. Essa possibilidade será estudada nas etapas subsequentes do trabalho, as quais avaliarão também a possível implementação das cenas em um ambiente 3D, para que essas relações espaciais façam mais sentido que em uma cena em 2D. Cabe lembrar que no caso de simulação da cena em uma ambiente 3D o uso das imagens (fotos e desenhos) 2D não se torna inviável, pois é possível definir, por exemplo, que uma das paredes do ambiente 3D terá um painel ou quadro no qual essas imagens serão projetadas quando for necessário definir um certo pano de fundo.

Assim, as imagens de apoio do ambiente de ensino podem ser fotos ou desenhos, e cada uma dessas modalidades tem suas especificidades seja no que se refere a níveis de iconicidade e de indicialidade, seja no quesito flexibilidade para modificações dinâmicas. Nesse particular, o uso de desenhos traz uma grande maneabilidade tanto na composição de cenas, que podem ser geradas e alteradas pela simples movimentação dos elementos que a compõem, quanto na alteração em tempo real de características da imagem, que podem ser mudanças de cor, de proporções, etc. Isso descortina algumas perspectivas muito atraentes,

como a de **geração emergente de cenas**. Como visto ao longo deste trabalho, uma das maiores vantagens da modelagem baseada em multiagentes é a possibilidade de atingir resultados não conhecidos ou esperados *a priori*. Essa é a natureza emergente da interação multiagente. No cenário aqui vislumbrado, uma das possíveis dinâmicas pedagógicas a ser explorada é a da formação de cenas a partir de um pano de fundo em branco (sem contexto predefinido), na qual o contexto final resultará justamente da interação dos agentes. Um exemplo disso seria o da formação emergente de um jardim a partir dos codelexes de flores, arbustos e árvores (e seus desenhos de apoio), só com base nas relações existenciais desses entes. Em outras palavras, sem que as características do jardim sejam predefinidas, se em um dado momento o primeiro elemento a ter acesso à tela (com fundo branco) é o codelex de um girassol, sua simples presença na tela já começa a criar um contexto mais provável para essa cena, e isso, em contrapartida, favorece outros entes relacionados a esse contexto (outros tipos de planta, por exemplo) na disputa por espaço na tela, e a cada nova presença de flores e plantas na imagem mais o contexto de “jardim” se fortalece e emerge. É claro, a disposição espacial desses agentes na tela deverá respeitar a disponibilidade de espaço e as proporções entre os vários entes ali presentes. Mas como os codelexes dispõem dessas informações, o arranjo da cena pode ser feito autonomamente. E o mecanismo de seleção de ação proposto nesta arquitetura possibilita que a formação do jardim seja acompanhada pela criação espontânea dos rótulos e das sentenças que descrevem cada elemento inserido e sua relação com os demais elementos. A flexibilidade de imagens baseadas em desenho permite também que as cores de alguns dos elementos sejam alteradas dinamicamente (produzindo uma alteração também em sentenças que descrevem a cena).

O exemplo mencionado para o jardim pode ser estendido para outros contextos, como o de uma cena subaquática, com peixes, plantas e outros entes. Essa possibilidade é não só atraente do ponto de vista do desafio de implementação mas, sobretudo, em vista dos efeitos positivos que ela pode ter do ponto de vista didático na retenção da atenção e do interesse dos alunos, que podem passar a encarar essas situações como um tipo de jogo.

As ideias principais da arquitetura aqui proposta estão descritas em Ávila e Gudwin (2010a) e em Ávila e Gudwin (2010b). Mas acreditamos que a arquitetura tem um potencial que vai além do que foi possível testar nas simulações acima. Assim, na discussão dos resultados atingidos até esse ponto, que será feita no Capítulo 7, serão consideradas não só as simulações, mas também as possíveis “avenidas de desenvolvimento” que a arquitetura pode abrir e discutidas sua validade e viabilidade já considerando alguns dos cenários de evolução previstos para as próximas etapas.

Capítulo 7

Discussões e considerações finais

A arquitetura ILTS aqui proposta para a gestão de conteúdos léxicos em um tutor inteligente tem características que podem ser úteis para o ensino de uma língua assistido por computador. Primeiramente, ela procura fazer a gestão dos conteúdos lexicais em seu nível elementar, em vez de segmentar o conteúdo por seções ou capítulos. A essa especificidade (granularidade) é também possível combinar recursos de gestão de nível mais alto, os quais permitem direcionar o processo de ensino para atender necessidades e temas específicos, como a priorização de determinados elementos lexicais em razão de sua conveniência para cada tipo específico de aprendiz. Isso vai ao encontro da ideia de oferecer um ensino mais adaptado às necessidades particulares de cada aluno. Em segundo lugar, ela torna possível o acompanhamento mais minucioso do aprendizado, ao garantir aos tutores humanos acesso direto ao ciclo de vida instrucional de cada lexema, uma vez que os dados do agente-lexema correspondente podem ser acompanhados e contrastados com as situações às quais o aluno foi exposto. Essa possibilidade, entretanto, pressupõe a existência de um mecanismo de modelagem do aprendizado, e portanto não foi possível testá-la nesta etapa do trabalho, sendo assim um dos pontos a serem investigados nos estudos subsequentes. Finalmente, a flexibilidade da abordagem *bottom-up* pode trazer à aplicação a capacidade de lidar com um dos aspectos mais complexos da área de ensino de línguas que é a da amplitude do léxico a ser aprendido, que, independentemente da L2 visada, será sempre da ordem de milhares de vocábulos.

A capacidade de auto-organização dos conteúdos em resposta à inclusão de novos lexemas nos objetivos curriculares do curso é uma característica que reduz a necessidade de os especialistas do domínio ou da aplicação proverem todo um cabedal de novas sentenças para as novas unidades léxicas. O mesmo se aplica às situações de inclusão de novas imagens ao repositório de apoios visuais do curso, pois uma vez que as novas imagens recebam os metadados que as descrevem (com vistas a uma estratégia didática) assim que incluídas elas se inserirão automaticamente na dinâmica do ILTS, levando à formação de sentenças baseadas nos itens semânticos que elas ilustram/apoiam. Essas imagens associadas a metadados têm além disso a característica de serem objetos de aprendizado que podem vir a ser compartilhados com outras instâncias do ambiente aqui descrito ou outras aplicações.

A característica da arquitetura proposta de buscar no uso real da L2 (tal como observável em conteúdos na Web) os parâmetros para estabelecer as relações mais usuais entre lexemas está em acordo com a teoria ecológica do aprendizado que, como discutido no Capítulo 3, caracteriza-se por valorizar comportamentos observáveis dos falantes reais, em contraste com a figura do “falante nativo ideal” e da visão dos dicionários e gramáticas como definidores mais válidos da língua (razão normativa). A arquitetura proposta permite inclusive que os conteúdos instrucionais instanciados nos agentes reflitam rapidamente mudanças no uso real da L2 tão logo essas se materializem em registros observáveis na Web, como, por exemplo, nos casos em que um novo termo se torna de uso corrente em razão de uma inovação social ou tecnológica, nos casos de novos conceitos, gírias, etc.

A estratégia de sequenciamento automático de lexemas e formação espontânea de sentenças significa, em última análise, que o ensino poderá se basear na exposição gradativa do aluno a *inputs* da L2, em contextos e sequências que façam sentido na lógica interna da L2, refletida na ontologia da língua, e que estejam fortemente ancorados em informações do ambiente, definidas pelas imagens e pelos metadados da cena. Tudo isso busca reproduzir aspectos do primeiro aprendizado linguístico, o da aquisição da língua materna que, como visto, é um processo gradual e contextualizado. Mas um aspecto-chave no caso do ambiente de ensino é que este tem consciência de que os conhecimentos linguísticos anteriores dos alunos (sobretudo a L1) são uma parte essencial do “contexto” em que se dá o aprendizado da L2, e isso significa que o ILTS vai oferecer apoios e sugerir sequenciamentos que sejam pertinentes a cada aprendiz. Assim, em vários aspectos o aprendizado da L2 diferirá da aquisição da L1, em particular pelo fato de o aprendizado não ser tão inconsciente quanto a aquisição, e sobretudo pelo fato de que os conhecimentos linguísticos de cada aluno têm grande influência no aprendizado da L2. Nesse sentido, o fato de a arquitetura proposta ser capaz de levar em conta, para cada lexema da L2, seu grau de similaridade com o lexema correspondente na L1 de cada aluno, dá aos tutores humanos maior visibilidade sobre os efeitos disso no processo de aprendizado da L2, sejam esses positivos (transferências) ou negativos (interferências).

Uma consideração importante é a de que o uso de imagens de apoio para facilitar o aprendizado da L2 por alunos de variadas origens culturais habilitando-os a traduzir para a L2 conceito e pensamentos de sua cultura não pressupõe um compartilhamento absoluto de significações dos diversos conceitos, nas diversas culturas. Para os propósitos de um ILTS basta que em seu uso intercultural os signos se refiram às mesmas coisas, sendo, como discutido no Capítulo 2, desnecessária qualquer sobreposição cognitiva adicional. Se para um jovem indiano o ícone de um “elefante” remete ao conceito um animal domesticável, e para um jovem africano o mesmo ícone leva ao conceito um animal selvagem e arisco, o que interessa é que ambos consigam aprender como se exprime o conceito “elefante” na L2.

O estudo de caso centrado no ensino do esperanto tem, além das justificativas listadas na Introdução e no Apêndice 3, a vantagem de permitir testar várias necessidades próprias dos diversos grupos linguísticos, uma vez que a gramática do esperanto apresenta características tanto de línguas inflexionais, quanto de línguas isolantes e aglutinativas (Golden, 1992; Jordan, 1987; Piron, 1981). O caso citado no Capítulo 6, da necessidade de flexão do lexema para refletir o modo acusativo, é ilustrativo do tipo de exigência que muitas outras línguas poderão ter, por exemplo nas desinências de plural, no uso de maiúsculas, etc. Além disso, como o esperanto permite grande liberdade na combinação dos lexemas, como discutido em Piron (2001), o estudo de caso exercita uma capacidade que o ambiente deve ter para refletir as construções que devem predominar em cada língua, como as sentenças do tipo SVO (sujeito-verbo-objeto) no português e outras línguas ocidentais, do tipo SOV (sujeito-objeto-verbo), como no japonês e hindi, ou VSO, como no gaélico. A abordagem multiagente permite incluir elementos (agentes) especializados para arbitrar o processo de formação de sentenças, a fim de que esse reflita as especificidades das línguas-alvo, sem que isso exija significativas reformulações do ambiente. E no caso das línguas que admitam as várias ordens, permite alternar exemplos com ordens diferentes.

7.1 Papéis e atores para a viabilização da arquitetura

O papel central na implementação da arquitetura cabe ao especialista da aplicação, a quem compete criar e manter a ontologia da L2, popular os dados internos dos agentes-lexemas, buscar e selecionar imagens com potencial didático para uso no ILTS e criar os metadados para essas imagens. Esse papel exige portanto conhecimentos tanto sobre a L2 quanto sobre a criação de ontologias, além de alguma familiaridade com o uso do aplicativo de anotação semântica. Conhecimentos de programação multiagente só serão necessários caso o especialista decida alterar aspectos funcionais do mecanismo de seleção de ação. A inserção de dados dos agentes-lexemas pode ser feita diretamente em uma planilha externa, sem que o especialista tenha que se ocupar do código do modelo ABMS. O mais provável é que a implantação inicial do ILTS para uma dada L2 demande o trabalho de uma equipe que combine as capacidades e conhecimentos mencionados acima. Cabe salientar que a arquitetura demanda um esforço considerável na sua implementação, sobretudo se o léxico visado for muito extenso, mas, em contrapartida, uma vez implementada ela tende a ser auto-organizável, restando aos especialistas do domínio funções menos técnicas e mais didáticas. Desse ponto em diante o ILTS passa a ter como principais atores os alunos e os tutores humanos. Caberá aos tutores humanos, que idealmente serão tanto especialistas do domínio, quanto da pedagogia, as tarefas de supervisionar o funcionamento do ILTS, avaliar que estratégias didáticas se mostram mais adequadas a determinados perfis de aluno, verificar se algum conteúdo necessita de novas imagens de apoio, se as imagens precisam de mais descritores, ou se a ontologia pode ser aperfeiçoada ou corrigida.

7.2 Desenvolvimentos futuros

Os três pilares da arquitetura, agentes, metadados e ontologias, são também os conceitos mais fundamentais da Web Semântica. Isso significa que, em certa medida, a arquitetura aqui proposta incorpora, em um âmbito mais restrito, muitos dos elementos que norteiam a evolução da Internet, sendo assim uma aplicação que testa em pequena escala ideias que podem vir a ser expandidas para a Web no futuro. Cabe então aos estudos subsequentes buscar possíveis transposições da arquitetura para um contexto mais aberto e totalmente baseado na Web, ou na nuvem, para usar um conceito mais atual. Nesse sentido, o estudo tem sido pautado pela aderência aos padrões propostos para a Web Semântica, como o uso de metadados baseados no padrão MPEG-7, entre outros.

Como tratado no Capítulo 1, nos estudos subsequentes vai ser também necessário investigar a questão da modelagem dos aprendizes para, entre outras razões, investigar de que forma essa pode tirar proveito da modelagem de domínio proposta nesta arquitetura. Espera-se que a harmonização dos dois modelos seja mais rica por conta das características de flexibilidade e de acompanhamento do ciclo de vida instrucional de cada lexema, como já descrito para a modelagem de domínio aqui proposta. Embora não tenha sido possível testar algumas dessas hipóteses nas simulações feitas até agora, será possível testá-las de forma aprofundada após a implementação completa do ambiente. Em particular, o recurso aos aplicativos de inferência da ontologia deverá ser exercitado para gerar sentenças com dados não explícitos nas imagens, mas sobre elementos presentes nelas.

7.3 Perspectivas da arquitetura

Como exposto na Introdução, existem justificativas e motivações de cunho prático para o desenvolvimento de novas arquiteturas para EaD de línguas, em especial baseadas na Web. E em vista dos desafios associados a essa área de estudo não foram discutidos outros possíveis usos para os mecanismos propostos e especificados. Mas um mecanismo para geração automática de sentenças com base em um contexto ou uma cena tem um potencial que extrapola o escopo do EaD de línguas. Esse seria aplicável, por exemplo, ao caso das aplicações georreferenciadas sensíveis a contexto, as quais, se associadas a uma ideia como a arquitetura multiagente para geração automática de sentenças, podem viabilizar aplicações de navegação ou de auxílio ao deslocamento de pedestres em locais desconhecidos, com a formação automática de sentenças de orientação espacial em função da pertinência maior ou menor dos obstáculos, objetos ou endereços do entorno, o qual varia com a coordenada espacial, e pode também levar em conta outros usuários acompanhados pelo sistema.

Em outro cenário de uso, se acoplada a softwares de reconhecimento e recuperação de imagens digitais dispersas na Web, a arquitetura proposta pode tornar-se uma ajuda técnica para usuários cegos se, com base no reconhecimento de elementos retratados em uma imagem e em dados de uma ontologia de apoio, for possível para o computador criar sentenças descritivas daquela imagem. Isso pode ser muito útil como recurso assistivo visto que nem todas as imagens disponíveis na Web têm os descritores textuais (metadados) recomendados pelas normas de acessibilidade.

Em um horizonte ainda mais distante podem-se vislumbrar aplicações em que câmeras capturam imagens do ambiente, recorrem a algoritmos de reconhecimento de padrões e, com base nos resultados obtidos, permitem a geração de sentenças descritivas do ambiente em que a pessoa com deficiência visual se encontra.

7.4 Considerações finais

Antes de concluir, é importante fazer algumas ressalvas com relação aos objetivos e o escopo deste trabalho. Por se tratar de um estudo da área de engenharia, seu objetivo não foi o de propor uma nova teoria para o aprendizado linguístico, nem questionar as várias teorias já existentes dentro da linguística para explicar tal processo. O objetivo, como comentado na Introdução, foi contribuir para a área de tutores inteligentes para ensino de línguas (ILTS) com a especificação de uma nova arquitetura de modelagem do domínio e, em particular, com uma solução que usa consciência artificial para tratar a complexa gestão dos conteúdos didáticos nesse tipo de aplicação. Com esse intuito, o escopo do trabalho incluiu uma revisão de várias disciplinas a fim de compreender as áreas de conhecimento pertinentes à aplicação. Essa extensa pesquisa resultou nos quatro primeiros capítulos da dissertação. A partir dessa base interdisciplinar foi então possível vislumbrar o potencial e as vantagens da arquitetura tal como especificada no Capítulo 5. Embora as simulações iniciais descritas não tenham esgotado todo o potencial da arquitetura nem exercitado todos os aspectos discutidos nos capítulos iniciais, a etapa subsequente do estudo avançará nessa direção, partindo da arquitetura já especificada e tentando avaliar todo o seu potencial.

Apêndice 1

Resumo das principais teorias sobre aprendizado linguístico

Muitas teorias já foram propostas para explicar o desenvolvimento linguístico nos seres humanos, das quais várias se fundamentam na psicologia e na linguística (aplicada). De modo geral, as teorias podem ser divididas em duas vertentes principais, contrastadas por sua ênfase na “natureza” (*nature*) ou na “educação” (*nurture*) como fator essencial para o aprendizado ou a aquisição de uma língua. Enquanto as primeiras adotam um paradigma nativista segundo o qual o conhecimento não é totalmente adquirido pela interação com o meio, sendo a aquisição linguística em parte determinada congenitamente, geneticamente transmitida, de maneira inata, as últimas têm um viés empiricista em que o conhecimento é considerado um resultado da experiência e da interação com o ambiente e de que a criança não precisa ter um mecanismo mental dedicado para resolver o problema do aprendizado de palavras; ela se vale de habilidades que existem para outros propósitos, como a capacidade de inferir a intenção dos outros, a de adquirir conceitos, a de avaliar a estrutura sintática, e certas habilidades gerais de aprendizado e de memorização. Embora não discordem por completo da vertente concorrente, os adeptos de cada corrente de pensamento tendem a enfatizar a preponderância de um ou outro fator.

Algumas das mais influentes teorias surgidas no século XX foram as da Aculturação (Schumann), do Comportamento Verbal (Skinner), da Polifonia ou Dialógica (Bakhtin), do Ato da Fala (Searle), da Acomodação (Giles), da Competência Variável (Ellis), da Gramática Universal (Chomsky), da Abordagem Modular (Fodor) e do Modelo Monitor (Krashen). Como o ambiente aqui proposto se baseia em uma visão ecológica e semiótica do aprendizado da L2, a discussão do mérito e das consequências dessa linha de raciocínio, central para nossos objetivos, encontra-se no capítulo 4. Mas como diversas outras teorias tratam do aprendizado linguístico, neste apêndice são sintetizadas as ideias e visões em que se apoiam as principais teorias e suas implicações na proposição de métodos de ensino de uma L2. Algumas das teorias aqui mencionadas, como as de Piaget e de Vigotsky, são na verdade construtos muito mais amplos e abrangentes, que tratam não somente do aprendizado linguístico, mas da cognição humana em sentido amplo. Nesses casos nos limitaremos a alinhar as visões dessas teorias para os propósitos deste estudo.

1) Teorias com ênfase nos efeitos do ambiente

Skinner e o comportamento verbal

O psicólogo Skinner propôs uma visão behaviorista da aquisição da linguagem, na qual se afirma que o desenvolvimento linguístico resulta de um conjunto de hábitos, sendo o conhecimento o produto da interação com o ambiente através de um condicionamento estímulo-resposta, com reforço positivo das respostas que, após um certo número de recorrências, faz com que o organismo aprenda como associar sua resposta ao estímulo com o reforço (Kiymazarslan, 2000). No âmbito linguístico, essa teoria defende que aqueles que adquirem uma L1 ou uma L2 recebem entradas linguísticas dos falantes no seu entorno, e reforço positivo de suas repetições e imitações, posição questionada por Chomsky (1959).

Ainda segundo os behavioristas, erros são interferências dos hábitos da L1 na aquisição dos hábitos da L2, e em razão disso, se as línguas se assemelham, a aquisição será fácil, e se são muito diferentes, será difícil. Essa é a chamada hipótese da análise contrastiva (HAC). Mas segundo (Lightbown e Spada, 1993), essa influência não seria meramente uma questão de hábitos, mas muito mais uma tentativa sistemática por parte do aprendiz de usar conhecimento já adquirido no aprendizado de uma nova língua. Além disso, há pesquisas (Ellis, 1985) que demonstram que nem todos os erros previstos pela HAC são de fato cometidos. A seção 4.4, abaixo, propõe uma explicação para os erros cometidos pelos alunos durante o processo de aquisição da L2.

Ainda segundo (Kiymazarslan, 2000), a visão behaviorista entende o aprendiz como uma *tabula rasa* desprovida de qualquer conhecimento inato. Entre os métodos de ensino influenciados por essa visão destaca-se o método Audiolingual. Entretanto, os métodos sustentados na teoria behaviorista falham ao suporem que as conexões estímulo-resposta pudessem explicar a aquisição e oferecer uma base sólida para uma metodologia de ensino.

Aprendizado linguístico segundo Piaget

As habilidades cognitivas nas crianças, segundo a teoria piagetiana, muito mais do que diferenças quantitativas, apresentam diferenças qualitativas em relação às dos adultos. Por essa razão, a criança é incapaz de realizar certas tarefas cognitivas caso ainda não tenha atingido um determinado estágio de desenvolvimento, e os quatro estágios identificados por Piaget estão associados a diferentes capacidades linguísticas. Assim, no estágio sensório-motor (0 a 2 anos) a maturação das regiões cerebrais frontais e pré-frontais corresponde às primeiras balbúncias, enquanto que o estágio pré-operacional (2 a 6 anos) testemunha a especialização dos hemisférios cerebrais, o desenvolvimento da percepção de formas e dos circuitos neurais da linguagem, que passa a organizar as ações, que se tornam intencionais. Por sua vez, o estágio operacional concreto (6 a 12 anos) dá à criança uma percepção da relação lógica entre acontecimentos e, em geral, corresponde à fase de aprendizado da língua escrita. Por fim, no estágio operacional formal (12 anos em diante) as habilidades cognitivas e linguísticas se tornam plenas.

Piaget entendia o processo de aprendizado como sendo iterativo, visto que as novas informações são adaptadas para encaixarem-se nos conhecimentos que o sujeito já possui, ao passo que o conhecimento que ele possui é também adaptado para acomodar a nova informação. Esse processo pressupõe uma etapa de *assimilação*, que ocorre quando a criança percebe novos objetos ou eventos em termos de esquemas ou operações existentes e tende a aplicar qualquer estrutura disponível a fim de assimilar o novo evento, e ativamente busca utilizar uma estrutura recém-adquirida, em um processo de adaptação às estruturas cognitivas existentes. Por sua vez, a etapa de *acomodação* ocorre quando os esquemas ou operações (estruturas cognitivas) devem ser modificados para levar em conta uma nova experiência (nova informação). E das sucessivas etapas de assimilação e acomodação resulta o processo de desenvolvimento, denominado *equilibração*.

Aprendizado linguístico segundo Vigotsky

Vigotsky (1962) partilhava a visão piagetiana de que o desenvolvimento infantil se dá em estágios distintos e mensuráveis, mas rejeitava alguns aspectos da teoria de Piaget, como a suposição de um desenvolvimento unidirecional em que todas as crianças atingem os mesmos estágios com aproximadamente a mesma idade. Vigotsky considerava possível, por meio do ensino, retardar ou acelerar esse desenvolvimento, que está sujeito à influência dos contextos cultural e social. Além disso, os desenvolvimentos biológico e cultural não ocorreriam de forma isolada, mas inter-relacionada, e o desenvolvimento intelectual seria uma evolução constante, sem ponto final. Por fim, enquanto Piaget defendia a noção de que alguns conceitos não deveriam ser ensinados até que as crianças estivessem no estágio de desenvolvimento apropriado, Vigotsky propôs a existência das zonas de desenvolvimento próximo (ZDP), que representam a distância entre o nível de desenvolvimento real na solução de problema por conta própria e aquele alcançado sob supervisão adulta ou em cooperação com colegas mais capazes.

Vigotsky considerava que a língua é a chave para todo desenvolvimento e que as palavras desempenham um papel central não apenas no desenvolvimento de pensamento, mas no crescimento da cognição como um todo, podendo o desenvolvimento linguístico ser visto como resultado da interação social. No que toca à definição da ZPD, ele ressaltou o papel da *língua compartilhada* no desenvolvimento da língua e do pensamento, como efeito de dois níveis de desenvolvimento que determinam o processo de aprendizado: interação e egocentricidade. As crianças precisam primeiro ser expostas à interação social, para então se capacitarem, por meio da fala egocêntrica, a criar seus recursos internos. Vigotsky não deixa claro, contudo, que recursos internos seriam esses, e não há consenso sobre o papel da fala egocêntrica no desenvolvimento do pensamento e da linguagem, como questionado, entre outros, por Piaget (Kiymazarslan, 2000).

Bakhtin e a teoria de polifonia ou dialógica

Essa teoria, influenciada tanto por Piaget quanto por Vigotsky, postula que o estudo da linguagem só é possível em conexão com os fatores sociais e políticos. Ela enfatiza o valor da diversidade linguística, e afirma que a língua evolui dinamicamente e é afetada pela cultura que a produz e em contrapartida ajuda a modelar aquela cultura, e a impor uma visão de mundo particular. É uma das bases da teoria ecológica, tratada na seção 4.4.

Teoria do ato da fala

Afirma que toda elocução possui dois tipos de significado: um preposicional, que se refere ao sentido literal da elocução, contido nas palavras e estruturas, e outro ilocucionário, que diz respeito ao efeito que o texto pronunciado ou escrito tem sobre os ouvintes ou leitores. Algumas elocuições podem ter efeitos subentendidos que o receptor deve tentar compreender. No que toca a questão do aprendizado da língua, é então necessário saber como esses efeitos ilocucionários são aprendidos, sobretudo no caso do aprendizado da L2. Alguns autores se perguntam se aprendizado em sala de aula é capaz de ajudar os alunos a adquiri-los ou apenas ter consciência de que eles existem.

Teoria do discurso

A teoria do discurso funda suas raízes na teoria do uso da língua. Ela enfatiza que o desenvolvimento da linguagem se deve ao modo como o aprendiz, ao tomar parte em um processo comunicativo, descobre a capacidade de significação da linguagem. A aquisição ocorrerá quando os aprendizes descobrirem como e quando usar a língua em situações variadas, e quando tiverem a cognição de várias formas de competências, seja gramatical (lexical, morfológica, sintática e fonológica), seja pragmática (atos de fala, etc.). O aprendiz precisa “conhecer” estratégias conversacionais para adquirir a língua. No caso da aquisição de uma L2, os alunos são encorajados a “aprender fazendo”, isto é, empreender ações que supostamente os ajudarão a adquirir a língua.

Segundo a visão nativista (Kiyamazarslan, 2000), uma das fraquezas da teoria é que ela superestima o papel dos fatores externos e subestima a importância de processos inatos do aprendiz. Embora mais sofisticada que a visão behaviorista, a teoria do discurso assemelhar-se-ia a ela na medida em que atribui uma posição central aos fatores e entradas (ou estímulos) externos no processo de aquisição da linguagem.

Teoria cognitiva da aquisição linguística

As origens da psicologia cognitiva são em geral atribuídas aos trabalhos de Edward Tolman, com sua Psicologia Gestalt (1932), e de Jean Piaget (1923), com sua Teoria do Desenvolvimento Cognitivo. Segundo a Psicologia Gestalt, o conhecimento vem de algo além da experiência, e pressupõe que o “conhecedor” imponha ativamente uma organização aos dados sensoriais. Para Kohler (1925), o comportamento não poderia ser explicado somente pelos princípios de associação. Tolman propôs a existência de um processo interno em que os aprendizes reconheceriam a interconexão com base nas propriedades das próprias coisas. Assim, o aprendizado não ocorreria de modo regular e contínuo a partir de um padrão de tentativas e erros, mas a partir da compreensão de um novo relacionamento: “a experiência do *insight*”.

Os psicólogos cognitivos veem a aquisição de uma L2 como sendo a construção de sistemas de conhecimento que num dado momento poderão ser automaticamente invocados para a fala e a compreensão. Nesse sentido, o aprendizado linguístico teria algumas características Gestalt, no qual esse é processo holístico que não pode ser analisado com base em associações estímulo-resposta, mas pode ser obtido de forma automática. Contudo, segundo a visão nativista (Kiyamazarslan, 2000), não está claro como isso se dá, nem como e quais estruturas da L1 podem ser transferidas para a L2.

Teoria da aculturação

Também conhecido como modelo sócio-educacional, advoga que aprendizado bem sucedido equivale à aculturação, isto é, tornar-se parte da cultura-alvo. Para melhor adquirir uma L2 os aprendizes devem se colocar em pé de igualdade com a comunidade linguística visada, sem se sentirem inferiores ou superiores.

Os críticos da teoria, embora admitindo a importância dos fatores socioculturais, questionam a necessidade da total aculturação para que o aprendizado possa ocorrer.

Teoria da acomodação

Afirma que a aquisição de uma língua depende somente da educação, e que os aprendizes adaptam sua linguagem àquela da pessoa com a qual falam, sempre que desejam reduzir a distância social, demonstrar solidariedade ou obter algo daquela pessoa, ou fazem o oposto quando querem criar uma distância social. As implicações dessa teoria para o entendimento da aquisição de uma L2 não são muito claras, mas é possível que o desejo de aproximar-se de uma determinada cultura torne o aprendiz mais propenso e motivado a acomodar-se à linguagem própria daquela sociedade, facilitando a aquisição da L2. No caso particular do esperanto, a propensão à acomodação pode se dar para aqueles sujeitos que se identifiquem com os valores que a língua carrega, de abertura para a diversidade cultural e a proteção das culturas minoritárias.

Modelo da competência variável

Essa teoria enfoca a aquisição de uma L2, concebendo-a como um fenômeno bidimensional, que envolve o processo e o produto da língua. O primeiro refere-se à distinção entre as regras linguísticas e a habilidade de fazer uso dessas regras. O segundo refere-se aos tipos de discurso gerados: planejados e não planejados.

O modelo conexionista

Enxerga o aprendizado linguístico como sendo o estabelecimento de reforços entre as inúmeras conexões existentes no cérebro. Segundo os críticos do modelo, ele supõe que a aquisição de uma língua não ocorreria de forma gradual, mas simultânea, e por só estudar como o cérebro estabelece as conexões quando a língua é processada, esse modelo não explica como a língua é adquirida.

A visão interacionista

Vê a aquisição de uma língua como sendo o resultado de uma interação entre as habilidades mentais do aprendiz (cognição) e as entradas linguísticas. Este modelo é, segundo (Kiyamazarslan, 2000), um dos melhores porque busca combinar a corrente nativista, que enfatiza os fatores inatos, com a corrente educacional, que enfatiza a influência do ambiente.

2) Teorias com ênfase nos fatores inatos**Teoria da Gramática Universal**

A teoria da Gramática Universal é uma proposta para explicar a relativa rapidez da aquisição da língua materna, em vista da “mínima exposição” a entradas externas, ou ao fato de que essas entradas seriam pobres e deficientes, degeneradas em razão de problemas em sua produção, e também desprovidas de correção gramatical. Isso parece sugerir a existência de um conhecimento universal específico para a linguagem.

Conquanto a maioria dos pesquisadores reconheça a importância tanto da biologia quanto do ambiente, um tema ainda polêmico é a hipótese de que o fator biológico inclui capacidades específicas para a linguagem, a chamada “Gramática Universal”, na forma de um dispositivo interno que nos predisporia a adquirir uma língua, a sistematicamente

perceber a linguagem ao nosso redor, facilitando a aquisição linguística. Críticos da GU alegam que a aquisição resulta somente de habilidades cognitivas genéricas e da interação dos sujeitos nas comunidades em que estão inseridos.

No que diz respeito à aquisição de uma L2, entretanto, o paradoxo sugerido acima parece não se aplicar, dados os frequentes fracassos dos alunos, que em geral são maduros do ponto de vista cognitivo, em atingir uma proficiência equivalente ao de usuários nativos. Como explicar que eles não consigam recorrer à GU? Em que medida a GU é acessível no caso do aprendizado de uma L2? Essas perguntas vêm dividindo os pesquisadores. Alguns sustentam que a aquisição da L2 é exatamente igual à da L1, isto é, as duas fazem uso do mesmo dispositivo de aprendizado. Outros acreditam que no caso da L2 o aprendizado é totalmente diferente, sem recurso à GU. Outros ainda postulam que somente uma parte da GU é usada na aquisição de uma L2. O foco da UG, contudo, está no funcionamento da linguagem, e a aquisição de uma L2 é algo secundário para esta teoria.

A abordagem modular

Essa teoria propõe que o cérebro é organizado de forma modular, e cada habilidade específica está relacionada a um agrupamento de neurônios, que funciona independente de outros módulos não diretamente conectados, em um arranjo denominado “encapsulamento informacional”. Como esses módulos tratam de domínios específicos, a linguagem tem um módulo próprio. Segundo a visão nativista, isso significa que conhecimento consciente não pode penetrar o módulo visual, o módulo de linguagem ou qualquer outro módulo subconsciente (Kiyamazarslan, 2000). Essa visão tem similaridades com a GU, pois ambas consideram que o ambiente e as entradas que este propicia não são suficientes *per se* para explicar a aquisição da linguagem, a qual em razão disso seria determinada geneticamente.

No que concerne ao ensino de uma L2, embora difundida, a ideia de que o aprendizado indutivo, no qual a língua é ensinada sem descrição de sua gramática, leva à aquisição, não teria sustentação nas teorias nativistas em geral, e na abordagem modular em particular, um vez que, para essas, aprender algo via aprendizado por descoberta não implica aquisição da língua, mas tão somente significa uma melhora da capacidade de resolução de problemas.

O modelo monitor

Fundamentado nas teorias de Chomsky, Piaget e Vigotsky, esse modelo estabelece uma clara distinção entre estudo formal (acumulação de conhecimento) e assimilação natural (desenvolvimento de habilidade) de idiomas. O modelo tem característica nativista e alicerça-se em cinco hipóteses principais: i) distinção entre aquisição e aprendizado no desenvolvimento de competência em uma L2; ii) a existência de uma ordem natural na aquisição de estruturas gramaticais; iii) a ênfase na importância, para a aquisição de uma L2, de que os alunos recebam entradas um pouco além de sua competência atual; iv) atuação de um “monitor” no processo de aprendizagem, em que o aluno conscientemente descobre regras; v) um filtro afetivo pode ajudar ou prejudicar as entradas necessárias à aquisição da L2.

A primeira hipótese é um dos pontos de discussão entre os teóricos. Segundo (Krashen, 1981; Krashen e Terrel, 1983), os processos de aprendizagem e aquisição são intrinsecamente diferentes: enquanto a aquisição requer foco consciente no significado, o aprendizado requer foco consciente na forma, e os dois processos não podem ocorrer de forma simultânea e consciente. Além disso, o aprendizado não se transforma em aquisição por meio da prática, mas pela exposição a entradas consistentes. Adquirir, em outras palavras, significa reaprender, a partir de experiências concretas, a estruturar o pensamento conforme uma nova língua. Cada pessoa adquire segundo seu próprio ritmo, num processo que produz habilidade prática e não necessariamente conhecimento, isto é, o sujeito torna-se capaz de utilizar corretamente a língua sem ser capaz verbalizar suas regras subjacentes. Contudo, vários autores contestam esse ponto de vista, afirmando que o aprendizado pode efetivamente se transformar em aquisição.

A hipótese da Ordem Natural baseia-se em estudos cujos resultados sugerem que, ao assimilarem a L1, as crianças apresentam desenvolvimento sequencial semelhante. Ou seja, tendem a assimilar determinadas estruturas antes de outras. Do ponto de vista de ensino, todavia, essa hipótese não tem implicações práticas, pois o próprio autor admite não ser necessário reorganizar os conteúdos de forma a refletir a ordem natural, mas sim enfatizar mais a aquisição e menos os métodos gramaticais. A Tabela A.1, relaciona algumas das teorias existentes aos métodos de ensino a que deram origem.

Tabela A.1.1: Teorias de aquisição linguística (adaptada de Kiymazarslan, 2000)

Teorias de aquisição linguística (L1 e L2)		Alguns métodos de ensino resultantes
Baseadas na educação (fatores ambientais)	Comportamento verbal (Skinner)	Método audiolingual Aprendizado linguístico em comunidade Abordagem comunicativa
	Aquisição da linguagem (Piaget)	
	Zona de desenvolvimento próximo (Vigotsky)	
	Teoria de polifonia ou dialógica (Bakhtin)	
	O modelo de aculturação (Schumann)	
	Teoria do discurso	
	Teoria cognitiva de aquisição da linguagem	
	Teoria do ato da fala (Searle)	
	Teoria da competição	
	Teoria da acomodação (Giles)	
	Teoria da competência variável (Ellis)	
	Modelo conexionista	
	Visão interacionista	
Baseadas na natureza (fatores inatos)	Teoria neurofuncional da aquisição linguística	Compreensão de Winitz Abordagem natural
	Teoria da gramática universal (Chomsky)	
	Abordagem modular (Fodor)	
	Modelo monitor (Krashen)	

Visão psicológica da formação de esquemas no aprendizado de uma língua

Para nossos propósitos, tanto a teoria piagetiana quanto a visão vigotskiana da língua são muito relevantes, pois elas ajudam a entender alguns aspectos fundamentais da aquisição da L1 e do posterior aprendizado de uma L2. A noção da formação de esquemas de ação é particularmente útil no entendimento desses dois processos cognitivos, e, mais do que isso, para a compreensão de como o primeiro interfere no segundo. Isso é importante nos casos em que os esquemas adquiridos para a L1 se mostram válidos também para a L2, facilitando assim seu aprendizado, ou quando esses são diferentes, dificultando o processo. A teoria dos esquemas sugere que o esforço cognitivo será maior se nenhum dos esquemas que a criança já possui (inatos ou relativos à L1) puder ser generalizado para a L2. Isso ocorre porque o aprendizado passa a ser mais dependente de acomodações das estruturas cognitivas aos novos esquemas da L2, sobretudo se estes últimos não forem, além do mais, generalizáveis no âmbito da própria L2 (ver Piron, 2002, pp.116-129).

Cabe salientar que mesmo certos argumentos utilizados por teorias nativistas para postular a importância de fatores inatos na aquisição linguística podem ser interpretados e explicados à luz da teoria dos esquemas e do mecanismo da assimilação generalizadora. Assim, um argumento utilizado por defensores da teoria da Gramática Universal baseia-se no fato de que, a exemplo das crianças que adquirem a L1, adultos que adquirem uma L2 apresentam dificuldades com alguns tópicos gramaticais específicos, como o emprego do “-s” da terceira pessoa do singular do presente do indicativo do inglês, em contraste com a relativa facilidade que eles demonstram na aquisição do “-s” indicativo de plural. Mas a validade da GU não necessariamente é a única explicação para tais casos. Uma explicação mais simples é a da aplicação, tanto pela criança quanto pelo adulto, do mecanismo da assimilação generalizadora, tal qual postulada por Piaget. Isso se dá porque, enquanto o “-s” indicativo de terceira pessoa é uma exceção à regra geral de que os verbos em inglês não recebem “-s”, no caso do plural ocorre exatamente o inverso, pois a desinência “-s” do plural é a regra geral, enquanto que os plurais irregulares representam exceções. Logo, é natural que o “-s” plural seja assimilado primeiro, visto que por respeitarem o princípio da assimilação generalizadora as regras mais gerais são retidas mais precocemente, enquanto a inserção dos casos irregulares e exceções da língua nos esquemas mentais/cerebrais é um processo muito mais trabalhoso e que exige muito mais prática e/ou imersão mais longa no universo da língua.

Como observou Piaget, esse fenômeno é universal, estando presente em todas as culturas, inclusive entre adultos que aprendem uma L2. No português, por exemplo, quando a criança aprende as formas verbais *comi*, *entendi* e *desci*, ela vai naturalmente extrapolar essa regra para criar formas verbais como *fazi* e *trazi*, consideradas incorretas pela norma culta da língua. Entretanto, como chama a atenção Piron (2002), tais “erros” não denotam imaturidade psicológica, mas tão somente a livre aplicação do mecanismo da assimilação generalizadora em um estágio de inserção incompleta das regras (e exceções) da L1 nos esquemas mentais da criança. É por esse motivo que nos primeiros anos de imersão em seu ambiente linguístico a criança tem que assimilar, com considerável esforço, irregularidades inúmeras da L1 e transformá-las em esquemas. É justamente a magnitude desse esforço que explica a dificuldade maior ou menor de aprendizado dessas mesmas línguas como L2,

sobretudo pelo fato de que, via de regra, no momento em que adquire a L2 o aprendiz não pode se valer dos esquemas adquiridos para tratar as exceções na sua L1, pois ainda que as regras sejam parecidas, as exceções podem ser diferentes na L1 e na L2. Seja, por exemplo, a criança de língua materna portuguesa que, por volta dos cinco anos, já possui um esquema para formar o tempo futuro com o auxílio do verbo *ir*. É observado que, em razão da assimilação generalizadora, ela tenderá a atribuir uma validade absoluta ao esquema, o que a levará ao erro quando o verbo principal é também *ir*, o que leva a uma exceção (um erro do ponto de vista da norma culta da língua) como visto na Tabela A.1.2. Após algum tempo, dependendo da influência do ambiente, ela se conformará à norma, incorporando a exceção a seu esquema mental. Se, posteriormente, ela tiver que aprender a língua inglesa como L2, ela encontrará um mecanismo muito semelhante na formação do tempo futuro por meio do verbo auxiliar *will*. Assim como em português, também no inglês o mecanismo não é absoluto e tem exceções, só que essas não correspondem àsquelas do português, e é por isso que a aquisição do esquema correto exigirá tanto o bloqueio da verbalização espontânea (assimilação generalizadora), quanto a inibição dos esquemas válidos para tratar as exceções da L1, que na L2 conduziriam a um erro:

Tabela A.1.2: Incompatibilidade de esquemas mentais de L1-L2

Português	Inglês
Eu vou comer	I will eat
Eu vou brincar	I will play
Eu vou poder	I will can (incorreta)
Eu vou ir (incorreta)	I will go

O mesmo fenômeno ocorrerá com as regras de flexão de grau de adjetivos. No português, os comparativos de superioridade e inferioridade seguem a forma analítica: “mais alto”, “mais velho”, “menos feliz”, etc. Mas a norma padrão estipula que a forma sintética deve ser usada para alguns adjetivos: “maior”, “menor”, “pior” e “melhor”, como ilustrado na Tabela A.1.3.

Tabela A.1.3: Incompatibilidade de esquemas mentais de L1-L2

Português	Inglês	Espanhol
Mais alto	Taller	Más alto
Mais inteligente	More intelligent	Más inteligente
Maior	Bigger	Más grande / Mayor
Melhor	Better	Más bueno / Mejor

O tratamento dessas exceções deve então ser incluído ao esquema mental para os comparativos em português. Contudo, quando a criança, cuja L1 é o português, tentar assimilar uma L2 como o inglês ou o espanhol ela infelizmente não poderá tirar total proveito do esquema adquirido para a L1, pois a morfologia daquelas línguas usa esquemas

específicos. No inglês, por exemplo, os comparativos de superioridade seguem duas formas alternativas, dependendo do adjetivo do qual se trate. Adjetivos de origem germânica tendem a utilizar uma forma sintética com a desinência *-er*, enquanto que os de origem latina tendem a exigir a estrutura analítica “more + adjetivo”. Já o espanhol segue um esquema mais parecido com o do português, só que com diferenças e nuances na lista de exceções.

Assim, o aprendizado das L2 estará sujeito, nesses casos, a dois tipos de erro: *interlinguísticos*, decorrentes das interferências (transferências negativas) dos hábitos da L1 para as L2, e *intralinguísticos*, causados por generalizações, analogias e simplificações baseadas na tendência natural de se generalizar as regras mais frequentes na L2.

Assim, tanto na aquisição da L1, quanto na da L2, é totalmente natural que as regras gerais sejam assimiladas mais cedo do que as exceções, sem que isso implique a existência de um mecanismo inato para a aquisição de cada regra específica, mas tão somente uma tendência genérica de reforço dos esquemas mentais/neuronais relacionados a casos mais frequentes. Esse aspecto é pertinente ao nosso problema, pois entre os fatores mais determinantes do êxito no aprendizado da L2 estão a aplicabilidade dos esquemas da L1 à L2 (quantidade de transferências positivas), o que depende essencialmente da proximidade linguística dos dois códigos, e a regularidade desses padrões na L2 (quantidade de extrapolações válidas), o que depende de quão regulares são os esquemas da L2. Em outras palavras, isso significa que o êxito do aprendizado linguístico dependerá em grande medida de dois fatores: da proximidade entre os esquemas usados na L1 e os necessários ao domínio da L2, e de quão generalizáveis estes últimos são no âmbito da L2.

Os esquemas, tais como entendidos pela psicologia, equivalem, na ciência cognitiva, ao conceito de *processos automáticos*. Esses automatismos não devem ser confundidos com reflexos inatos, visto que devem ser aprendidos. Uma vez assimilados, esses automatismos demandam pouco ou nenhum esforço cognitivo, são predominantemente inconscientes, e não estão, além do mais, submetidos a um comando voluntário detalhado. Esse é o caso das habilidades de digitação, de descer ou subir escadas, ou de falar a língua materna. Essas habilidades se formam de início sob controle do córtex, mas, após uma quantidade suficiente de treino, tendem a passar ao controle subcortical (Baars, 2000: p. 66). Com os avanços do escaneamento do cérebro é possível observar que a realização de tarefas baseadas em processos automáticos produz muito menos ativação cerebral do que tarefas equivalentes para as quais o sujeito não tenha desenvolvido os automatismos. Isso explica a diferença nos gastos de energia e na fadiga produzida quando falamos uma língua na qual temos fluência (em geral a L1) e quando falamos uma L2 não totalmente internalizada.

Teoria ecológica do aprendizado linguístico

A teoria ecológica descreve os efeitos mentais/cerebrais da aquisição com base no *paradigma dos esquemas*, admitindo com isso a existência de atributos inatos que habilitam o ser humano a construir significados e desenvolvê-los na forma de sistemas de linguagem. Mas a constatação de uma predisposição congênita para a linguagem simbólica é entendida como uma instância de capacidades cognitivas mais genéricas, e portanto não significa a

concessão de um status de gramática universal aos fatores inatos, que, embora não estejam explicitados em várias das teorias baseadas na educação, não são incompatíveis com elas.

Segundo Fettes (1999), o conceito de “mente” é frequentemente equiparado ao de “cérebro”, e a ideia da mente como um órgão físico para o pensamento, em contraposição ao corpo, não pensante e mecanicamente reativo, tem fundamentado muitos dos trabalhos da filosofia e da psicologia modernas. Por esse prisma, a mente estaria contida no cérebro, e este seria um *container* do conhecimento. O próprio conceito de *aquisição* linguística sugere um processo no qual um objeto externo, “língua”, é capturado e guardado pela mente do aprendiz. Para Fettes, há três inconvenientes nesse paradigma: (i) não há como mostrar que os “significados dentro da cabeça” correspondem aos “significados no mundo” (Millikan, 1984); (ii) a metáfora do *container* despreza o conhecimento individual, tratando toda pessoa como *tabula rasa* ou como se trouxesse faculdades inscritas geneticamente para a construção de significados; (iii) essa concepção de mente dá sustentação a sistemas educacionais altamente ineficientes e alienantes, que põem em desvantagem minorias e pessoas economicamente desfavorecidas (conforme Freire, 1972).

A alternativa defendida por Fettes abandona as dicotomias da filosofia Cartesiana, tais como mente/corpo, humano/animal, pessoa/mundo e adota uma perspectiva ecológica ou sistêmica, na qual a realidade é vista como um relacionamento entre coisas. Essa visão se apoia no trabalho de Reed (1996), que propõe que a evolução da vida e a seleção natural asseguraram o desenvolvimento de mecanismos de percepção e de resposta a **informações persistentemente disponíveis** no ambiente denominados *esforço em busca do significado*. Os seres vivos possuiriam essa consciência sobre o entorno, consciência não no sentido de um estado interno privado, mas como um cambiante relacionamento ativo entre organismo e entorno, como descreve (Fettes, 1999):

Os animais mais complexos desenvolveram formas sofisticadas de consciência retrospectiva e prospectiva, as quais os habilitam a recordar relacionamentos passados com o mundo, compará-los com o presente e desse modo projetar os desdobramentos mais prováveis no futuro imediato. O conhecimento ecológico não é dado, mas adquirido pelo indivíduo no ambiente ao empregar e modificar estratégias adaptativas desenvolvidas pelas comunidades e espécies durante longos períodos de tempo.

Para acomodar essa definição de mente, Fettes recorre ao *paradigma dos esquemas*, em que os esquemas são redes neurais hierárquicas complexas associadas a estados particulares de consciência e ação. Assim, à medida que um organismo cresce e acumula experiência, ele desenvolve a capacidade de “chegar a estados similares aos seus estados de consciência prévios ao mesmo tempo em que preserva o conhecimento de sua origem passada (memória ou consciência retrospectiva), e usa essa capacidade para aprofundar sua consciência da situação presente e suas consequências imediatas (consciência prospectiva)”. A partir desse ponto de vista, o conhecimento não é “armazenado”, isto é, guardado de forma permanente e segura, mas provisoriamente retido, como qualquer outra capacidade que enfraquece quando não usada regularmente (Ben-Ze'ev, 1995). Ainda segundo Fettes, o paradigma dos esquemas enfatiza a natureza holística da percepção e do aprendizado. Novas capacidades

adquiridas não são armazenadas em “prateleiras” ou “depósitos” separados, mas na forma de modificações no sistema cognitivo como um todo: estrutura e conteúdo são a mesma e única coisa. Duas consequências desse modelo são o fato de que o aprendizado é algo mais profundo do que em geral se supõe, pois ele implica mudar o que o organismo é, e envolve alterações em um sistema cognitivo que já é altamente interconectado. Nesse sentido, aprendizes, seja de uma língua, seja de outras habilidades, nada têm de *tabula rasa*: eles são sistemas complexos altamente adaptados à sua maneira, e cuja negociação com o processo de aprendizado será fortemente influenciada por seus **estilos individuais** frente ao mundo.

Nossa capacidade de consciência prospectiva é construída e refinada gradualmente, na medida em que aprendemos a obter com mais eficiência significados e valores no nosso ambiente cotidiano. Ao postular que somos adaptados, no âmbito filogenético, assim como no ontogenético, a descobrir significados reais no mundo real, essa teoria se mostra assim, segundo Fettes, em conformidade com o naturalismo epistemológico (Millikan, 1984), com o materialismo interno (Smith, 1990) e com o realismo crítico (Bhaskae, 1989).

Ainda segundo Fettes (1999), o paradigma ecológico, em contraste com o idealismo Cartesiano, implica que o conhecimento é **ativo**, resultado de um esforço que dura toda a vida na busca pelo significado, que ele é **realista**, derivado do contato direto com o mundo, e **não proposicional**, pois não depende de linguagem e existe em outro meio, que como a própria vida (Capra, 1996; Maturana e Varela, 1992), pode ser visualizado por meio da teoria dos esquemas. A teoria ecológica da cognição entende os esquemas como sendo dispositivos heurísticos abertos e nebulosos que cada indivíduo desenvolve ao ajustar suas percepções e ações para gerar encontros produtivos com o mundo (Fettes, 1999).

Segundo Fettes, a visão ecológica, em contraste com a visão neo-Cartesiana presente na corrente principal da linguística aplicada, vê a linguagem como um sistema emergente e evolutivo de atos públicos coordenados, preexistente e externo a todo agente individual:

Isso significa que os indivíduos nunca realmente “adquirem uma língua” no sentido de serem capazes de reproduzir o sistema como um todo em toda sua complexidade dinâmica. Os indivíduos aprendem a reproduzir **elementos** do sistema, que compreende uma miríade de agrupamentos naturais e subsistemas de dispositivos públicos de linguagem, para coordenar suas próprias ações e consciências, assim como a de outros de maneiras que sejam úteis a eles. É verdade que o estoque de capacidades de um indivíduo para interpretar e produzir dispositivos de linguagem constitui ele próprio um sistema, mas a ontologia e a epistemologia desse sistema privado difere nitidamente do **fenômeno emergente da linguagem pública** (realce nosso).

Ainda segundo Fettes, os *tokens* linguísticos não são automaticamente interpretados, simplesmente em razão de sua disponibilidade no ambiente: eles devem ser percebidos como relevantes ao foco de atenção atual. Por isso, a comunicação linguística se baseia no **atingimento de uma consciência compartilhada** entre os participantes. Essa consciência compartilhadas, como discutido na seção anterior, significa termos um modelo da mente de nossos interlocutores. Assim, no âmbito da psicologia ecológica, aprender quando e como atingir a consciência compartilhada com outra pessoa é uma capacidade vital que a criança

começa a adquirir logo após o nascimento, para então, por volta de nove a doze meses, se tornar apta a tomar parte em quadros de interação triádica que envolvem, além dela própria, uma pessoa adulta e algum objeto ou evento (Fettes, 1999). À medida que a criança cresce, ela aprende a participar de quadros de interação mais amplos, inclusive quando o interlocutor não está presente fisicamente, e na comunicação impressa e de gêneros textuais diversos. Esses quadros usam uma variedade de signos linguísticos, miméticos, etc., e novos quadros específicos podem derivar de outros mais genéricos. Fettes comenta que essa estruturação social da comunicação remete aos gêneros de fala propostos por Bakhtin (1986), segundo os quais nosso conhecimento da língua materna não resulta de dicionários e gramáticas, mas de elocuições concretas que ouvimos e reproduzimos na comunicação com pessoas ao nosso redor. Assim, aprender uma língua, seja L1 ou L2, significa aprender gêneros. Uma das maiores dificuldades da aquisição de uma L2 é que os gêneros de interação em sala de aula simplesmente não se parecem com os gêneros na L1, mesmo quando se dá ênfase à comunicação, pois os **relacionamentos** estão no âmago do gênero, e esses não podem ser “construídos à vontade”, pois nascem em determinados e complexos ambientes. Em razão disso, a pesquisa sobre aquisição tem que levar em conta os relacionamentos que são construídos no decorrer do aprendizado, e avaliar como os gêneros resultantes restringem a produção e a interpretação da língua (Fettes, 1999).

No que concerne à aquisição de uma L2, as linguísticas moderna e aplicada sofrem, segundo Fettes, forte influência da razão normativa, que trata dicionários e gramáticas como definidores mais válidos da língua do que os comportamentos reais observáveis dos falantes, o que dá lugar à noção de “falante nativo ideal” em contraste com os falantes reais.

Tão numerosos quanto as teorias propostas até hoje para explicar o processo de aquisição linguística do ser humano têm sido os métodos e as estratégias desenvolvidos para facilitar o ensino de línguas em sala de aula. Embora nossa proposta de estudo refira-se a ensino a distância mediado por um computador, é importante entender os desafios e as dificuldades que essa atividade implica e impõe mesmo às situações de ensino presencial. As seções seguintes trazem então uma síntese histórica da evolução dos métodos de ensino das L2 e discutem as possíveis estratégias para esse ensino.

História dos métodos de ensino de uma segunda língua

No livro “Semiotics in Language Education” (2000), o semioticista e professor de idiomas Marcel Danesi faz uma retrospectiva do ensino de línguas nos últimos séculos, em especial no século XX. Segundo aquele autor, na Idade Média e no início da Renascença prevalecia a chamada abordagem de *tradução da gramática* no ensino das línguas clássicas, em que o grego e o latim eram ensinados de forma *dedutiva*: primeiro com a apresentação de uma regra gramatical, seguida por tarefas de tradução oral e escrita que usassem a regra. Para as línguas nativas, ditas vernáculos, o aprendizado ligava-se às necessidades práticas, e era considerado mais eficaz quando ocorria no contato direto com falantes nativos, num tipo de *imersão social*. Esta pode então ser considerada a mais antiga teoria de aprendizado, remontando a sumérios e babilônicos (Titone, 1968, *apud* Danesi, 2000, p. 3).

Com o aumento da importância das línguas vernáculos a partir do século XVI, passou-se a crer que essas deveriam ser ensinadas segundo o mesmo método empregado

para as línguas clássicas, a *tradução da gramática*. Mas ao perceberem o quão ineficaz essa estratégia se mostrava no ensino de habilidades comunicativas, alguns educadores, entre os quais Guarino Guarini (1374-1460), Santo Inácio de Loyola (1491-1556) e Wolfgang Ratke (1571-1635), propuseram introduzir um pouco de tempero de *imersão social* na sala de aula, ressaltando a necessidade de que regras de gramática fossem aprendidas por indução. Eles argumentavam que esse é o modo como as crianças aprendem a falar e como são adquiridas as habilidades comunicativas nos casos de imersão. No século XVII, Comenius (1592-1670) também defendia uma pedagogia orientada à indução, e chegou a desenvolver a técnica dos *diálogos situacionais* para que os alunos pudessem induzir as formas corretas da língua antes de aprendê-las explicitamente (Danesi, 2000: 3).

Apesar desses esforços, a estratégia de tradução gramatical prevaleceu até o final do século XIX, quando uma visão radicalmente nova do ensino de línguas surgiu e se difundiu, motivada pela crença de que o ensino efetivo em sala de aula deveria ocorrer de acordo com princípios do aprendizado respaldados pela psicologia, em cujas teorias sobre aprendizado linguístico deveriam se basear as práticas pedagógicas que o método objetivava, antes de tudo, a sistematizar. Isso levou ao desenvolvimento da *noção de método*, uma metodologia padronizada de como instruir, associada a recursos pedagógicos criados especificamente para esse fim (Danesi, 2000: 2). Szanser (1992: p.5) define o método como um conjunto de procedimentos sistemáticos, premeditados e adequados, uma união de meios e técnicas escolhidos para melhor atingir determinado objetivo. O método depende, entre outros, dos objetivos do ensino, daquilo que se quer ensinar, do perfil dos alunos, inclusive de sua língua materna e da situação concreta em que o aprendizado ocorrerá.

A exemplo dos educadores renascentistas, os reformistas do século XIX também viam o aprendizado indutivo como o mais natural para o ensino de uma L2. Esse movimento reformista se fixou nas mentes dos professores durante o século XX, e ainda hoje é influente (Danesi, 2000: 3).

A rejeição das práticas de tradução da gramática no ensino das L2 e a elaboração do primeiro método de ensino indutivo, chamado de *método direto*, trouxe um compêndio de estruturas sequenciadas que o professor deveria ensinar numa ordem preestabelecida. Com essas práticas esperava-se ativar o modo natural de aprendizado dos alunos e assim levá-los a resultados proveitosos (Danesi, 2000: 4-5). Tendo como objetivo simular condições e processos que se supunham fundamentar o aprendizado da L1, os *métodos naturais*, dos quais o método direto era o mais conhecido, combinavam o uso de associações concretas, demonstrações, prática de padrões, imitação e outros procedimentos e rotinas num formato a ser seguido por professores e autores de livros-texto (Danesi, 2000: 5).

Segundo Danesi (2000: 5), a validade da noção de método só veio a ser questionada cientificamente em 1917, quando Harold Palmer propôs dar um tratamento teórico para a interface entre o ensino e o aprendizado de uma L2. Palmer salientou, acima de tudo, que o ensino de uma L2 deveria ser mais adaptado às necessidades particulares dos alunos. E em lugar de desenvolver um formato único e padronizado para ensinar todos os alunos, como preconizava o método direto, os professores deveriam fundamentar suas práticas em sala de aula na real capacidade de absorção de seus alunos. No início dos anos 20 já havia crescido bastante o desencantamento com o método direto e, como suspeitava Palmer, isso se deveu

ao fato de que a teoria psicológica na qual ele se baseava, de que o aprendizado de uma L2 é igual ao de uma L1, não refletia todos os aspectos do aprendizado da L2, nem tampouco o faziam os princípios pedagógicos que ele propunha ser aplicáveis a todas as situações em sala de aula (Danesi, 2000: 5-6).

Os métodos da leitura e oral

Ainda segundo Danesi (2000: 6), o descrédito do método direto reacendeu o debate entre indutivistas e dedutivistas, e originou dois novos e influentes métodos: o *da leitura* e o *oral*. O primeiro surgiu da visão, muito popular nos anos 20, de que o único objetivo realisticamente atingível em ambientes de aprendizado não imersivos é a compreensão da leitura. A exemplo da abordagem da *tradução da gramática* que prevaleceu nos séculos precedentes, o método da leitura enfatizava o papel do ensino da gramática na língua nativa do aluno, mas aliava a isso uma maior atenção à pronúncia correta da L2, por considerá-la essencial à compreensão. O método também frisava a importância de formular explicações gramaticais em termos de uma análise científica da L2, e trouxe às práticas comuns da área de ensino de L2 noções muito úteis, como a de aprendizado controlado de vocabulário e a de gradação de textos. Esse método ainda está em uso hoje em dia por ser pouco exigente em relação aos professores e por atender a situações em que a meta principal é o acesso à literatura na L2, tendo entretanto poucos defensores e sendo pouco usado como alternativa preferencial para situações reais de ensino de uma L2 (Danesi, 2000: 6).

O método oral, comenta Danesi, foi desenvolvido por linguistas britânicos nos anos 30, tendo sido então uma proposta verdadeiramente inovadora. Muitas de suas técnicas ainda estão em uso hoje. Tornou-se muito popular entre os professores por enfatizar um controle estrito do vocabulário, uma gradação dos itens gramaticais dos mais simples para os mais complexos (posteriormente conhecida como currículo estrutural), a apresentação oral dos novos conteúdos (de onde se originou seu nome), a introdução da leitura e da escrita somente após ter sido determinado que uma competência léxica e gramatical tenha sido atingida pelos alunos, o uso exclusivo da L2 na sala de aula, e a prática de se apresentar novos conteúdos de forma a refletir situações da vida real (Halliday, McIntosh e Stevens, 1964: 38, *apud* Danesi, 2000: 7).

O gradual desencantamento a partir dos anos 40 e 50, tanto com o método da leitura, quanto com o oral, acrescenta Danesi, se deveu ao fato de que nenhum deles era aplicável a toda e qualquer situação de ensino, ao passo que os educadores da época ainda buscavam desenvolver um método “universal” de ensino das L2 que pudesse ser aplicado em qualquer situação. Tal crença foi reforçada com o advento do *movimento behaviorista* na psicologia, produzindo uma nova teoria do aprendizado de línguas baseada na formação de hábitos, que poderia ser facilmente traduzida em práticas instrucionais (Bloomfield, 1942 *apud* Danesi, 2000: 7). Essa crença foi reforçada também pelo êxito sem precedentes do assim chamado *Programa de treinamento especializado do exército*, concebido durante a Segunda Guerra Mundial para atender membros das forças armadas dos EUA, e que seguia princípios behavioristas cuja concretização pedagógica se dava através do uso de imitação, repetição e estágios de prática de diálogos, e assim parecia ser a panaceia pela qual os professores esperavam desde o início do movimento reformista (Danesi, 2000: 7).

Os métodos audiolingual e audiovisual

Esses dois adventos logo inspiraram a criação de um novo método, o audiolingual, anunciado como sendo o definitivo. Além de enfatizar a formação de hábitos, a prática de padrões e procedimentos de treinamento indutivos, o método rejeitava a hipótese de o aprendizado da L2 ser igual ao da L1, e sustentava que os conhecimentos da L1 determinam fortemente a maneira como o aluno percebe e assimila a L2 (Danesi, 2000: 7). Tal *teoria da transferência* nasceu da observação frequente em sala de aula de que hábitos de pronúncia e categorias léxicas e gramaticais da L1 são inconscientemente transferidas para a L2 durante o aprendizado, sobretudo nos estágios iniciais (Danesi, 2000: 7). Esse entendimento levou ao desenvolvimento de uma técnica chamada análise contrastiva, na qual se acreditava ser possível determinar, por meio da comparação das estruturas da L1 e da L2, quais aspectos da pronúncia e da gramática demandariam mais esforço e quais não: os que coincidissem na L1 e na L2 seriam menos enfatizados, admitindo que o processo de transferência (positiva) permitiria aos alunos adquiri-los inconscientemente. Aqueles que fossem diferentes teriam, em contrapartida, mais ênfase pois nesses casos as transferências seriam negativas, isto é, elas interfeririam negativamente no processo de aprendizado (Danesi, 2000: 7).

Esse método deu origem a uma versão europeia denominada método audiovisual, desenvolvido nos anos 50 na França, e que adicionava ao original a característica inovadora de apresentar visualmente os novos conteúdos por meio de diapositivos. Ainda segundo Danesi (2000: 8), as expectativas entusiásticas despertadas por esses dois métodos foram aumentadas pela fé ingênua na tecnologia, de modo que a introdução dos laboratórios de línguas ao *modus operandi* do método audiolingual, e dos recursos visuais ao método audiovisual foram saudadas como as peças que faltavam no quebra-cabeça das práticas em sala de aula capazes de garantir o domínio da L2 (Danesi, 2000: 8). Mas a euforia durou só até meados dos anos 60, quando uma série de eventos e experiências ocasionou o quase total abandono de ambos os métodos.

Método do código cognitivo

Para o descrédito dos métodos audiolingual e audiovisual contribuiu, por um lado, o fato de as expectativas por eles criadas não terem sido atendidas na prática. Por outro lado, como comenta Danesi, o arcabouço psicológico sobre o qual os dois métodos tinham sido apoiados começou a ruir sob o peso do *cognitivismo* na psicologia (Ausubel, 1967) e do *generativismo* na linguística (Chomsky, 1957, 1965). Por fim, diversos experimentos psicolinguísticos demonstraram que nenhum resultado significativo era obtido por esses métodos quando comparados com as abordagens de tradução da gramática.

Se, por um lado, as teorias generativistas concorreram para desacreditar os métodos behavioristas, por outro lado elas suscitaram o aparecimento de um novo método, chamado *método do código cognitivo*, concebido no início dos anos 70 por estudiosos da área da linguística aplicada. Sua proposta se baseava na noção de *competência linguística*, isto é, na ideia de que saber uma língua significa saber as regras básicas de seu desenho gramatical (Danesi, 2000: 8-9). A estratégia consistia então em prover material didático e exercícios projetados para conferir aos alunos um conhecimento de regras abstratas da organização da

L2. Entretanto, esse método nunca deslanchou de verdade entre os professores em geral, e foi abandonado na metade dos anos 70. Mas uma contribuição da linguística aplicada que continua sendo muito útil ainda hoje é a da *análise de erro*, a qual se baseou na observação de que o tipo de competência que os alunos manifestam ao falar e escrever a L2, chamada *interlinguagem*, caracteriza-se, sobretudo, por erros recorrentes e previsíveis. Erros (*errors*) se distinguem de deslizos (*mistakes*), que são as escorregadelas que os alunos cometem, mas que eles são capazes de corrigir por si mesmos, enquanto que os erros revelam lacunas na competência linguística (Danesi, 2000: 9).

As análises de erros das interlinguagens produzidas por alunos deixavam claro que havia dois tipos principais de erros: (1) *interlinguísticos*, cujas causas eram as interferências (transferências negativas) de algum hábito da L1 para a L2; e (2) *intralinguísticos*, que eram causados pelos mesmos mecanismos de aprendizado que caracterizam o aprendizado da L1, isto é, por generalizações, analogias, simplificações, etc. Alguns teóricos sugeriram a ideia de que os alunos constroem uma “teoria de funcionamento” da gramática da L2, baseada em certos princípios gerais de desenho da língua, extrapolados a partir das entradas (*inputs*) às quais eles tenham sido expostos (Danesi, 2000: 9).

Danesi comenta ainda que o objetivo principal dos métodos direto, audiolingual, de leitura e audiovisual era prover um currículo apropriadamente estruturado e um plano de ensino que conferissem competência linguística de forma sequencial e ordenada. Esperava-se com isso que a habilidade de aplicar as estruturas da L2 em situações comunicativas da vida real emergisse espontaneamente daí em diante. Os criadores do método oral eram a única voz dissonante nesse cenário. Nos outros quatro métodos, os conteúdos de diálogos orais eram tão sistematizados e controlados que, embora tentassem simular diálogos do mundo real, quase nunca eles se pareciam com conversas verdadeiras (Danesi, 2000: 9-10).

Método comunicativo

No início dos anos 70, acrescenta Danesi, o linguista Dell Hymes (1971) propôs a noção de que o conhecimento da estrutura de uma língua fosse interconectado com o conhecimento de como utilizá-la apropriadamente em situações sociais específicas. A essa interconexão ele deu o nome de *competência comunicativa*. Esse conceito ganhou um apelo imediato entre professores, os quais naquela época começavam a suspeitar que a noção de método era fundamentalmente errônea, visto que ela implicava o estudo artificial da língua como um código, separado de seus usos. Isso conduziu ao *movimento comunicativo de ensino de línguas*, que durou até os anos 80 (Danesi, 2000: 10).

Os primeiros currículos comunicativos a substituir os currículos estruturais foram desenvolvidos pelo Conselho da Europa. Ainda na década de 70, tais currículos passaram a se organizar em torno do conceito de *atos da fala*, definidos como a estratégia comunicativa que falantes nativos empregam inconscientemente para manter certos tipos de interações sociais. Assim, por exemplo, um protocolo simples como dizer *olá* é um ato da fala que postula um detalhado conhecimento das palavras, frases, estruturas e pistas não verbais apropriadas, que se juntam de forma coesa como que seguindo um *script* que permite ao falante estabelecer um contato social bem sucedido com outro falante. Uma violação de

qualquer dos detalhes procedurais desse *script* pode causar uma quebra na comunicação, ou ainda pior, um desentendimento (Danesi, 2000: 10).

Métodos humanistas

Uma segunda mudança ocorrida na filosofia do ensino das L2 durante os anos 70 e 80, segundo Danesi, foi o surgimento do chamado movimento *humanista*, que buscava uma alternativa ao ensino comunicativo de línguas, retomando a ideia, defendida sessenta anos antes por Palmer, de que as características de personalidade e necessidades emocionais dos alunos são de importância capital para o êxito do ensino de uma L2. Entre os métodos humanísticos mais conhecidos estão o *community counseling* de Curran (1976), o *silent way* de Gattegno (1976), o *total physical response* de Asher (1977, 1988), a *suggestopedia* de Lozanov (1979), a *abordagem natural* de Krashen e Terrel (1983), e a *abordagem de cenário* de Di Pietro (1987). Apesar de diferenças consideráveis nas técnicas usadas, todos eles, ainda segundo Danesi, compartilham três crenças centrais acerca da necessidade de: (1) prover um ambiente de ensino agradável, amistoso e não-ameaçador; (2) envolver os alunos diretamente no processo de aprendizado; e (3) estabelecer um clima de amizade entre professor e alunos. Essas premissas parecem explicar o alegado êxito desses métodos, mas sua desvantagem é que eles requerem do professor um total compromisso com os *scripts* de ensino do método, deixando pouco espaço para improvisações e criatividade. Essa é, segundo Danesi, a razão para eles nunca terem conquistado parcela substancial dos profissionais de ensino de línguas. Mas eles foram importantes ao chamarem a atenção às necessidades emocionais dos alunos. Algumas das ideias dos métodos humanísticos que merecem destaque são (Danesi, 2000: 12):

- Como afirmou Curran, a ansiedade sentida pelo aluno nas salas de aula tradicionais é o maior obstáculo ao aprendizado de uma língua;
- Como defendeu Gattegno, quanto mais o professor fala e explica, tanto menos o aluno tende a aprender, sugerindo a necessidade de dar mais oportunidades para os alunos se expressarem;
- Como enfatizou Asher (e também os adeptos do método direto), os resultados do aprendizado melhoram quando a L2 é associada com ações físicas, e que novos conteúdos sejam apresentados por meio de comandos aos quais os alunos tenham que reagir fisicamente;
- Como propôs Lozanov, o potencial para aprendizado duradouro cresce notavelmente se são criadas em sala de aula as condições capazes de ativar processos subliminares com os quais os alunos aprendam sem se darem conta;
- Como propôs Krashen, se o processo é inconsciente (como no caso da L1), então se trata de aquisição, e se o processo é consciente, trata-se de aprendizado, e este se baseia no monitoramento pelo aluno das entradas (*inputs*) às quais é exposto na sala de aula.

Apesar de todas as teorias e modelos propostos ao longo de várias décadas, no final do século XX, paradoxalmente, os professores ainda se deparavam com o mesmo incômodo dilema que os reformistas do século XIX tentaram solucionar: por que após estudar uma língua por variados períodos de tempo os alunos ainda parecem não ter a capacidade de

falar com o mesmo tipo de naturalidade que se percebe imediatamente na fala de um falante nativo? Danesi sustenta que a fonte de falta de naturalidade nasce da inabilidade dos alunos em interconectar os níveis superficiais e conceituais da linguagem, que aquele autor chama de fluência conceitual (Danesi, 2000: 13). Na busca por uma maior e melhor compreensão da fluência conceitual, Danesi recorre à noção de semiose, “a habilidade inata do cérebro e dos sistemas sensoriais de produzir e compreender signos”.

Apêndice 2

Aspectos principais da teoria Semiótica

Como teoria geral dos signos, a etimologia da palavra semiótica remonta ao grego *semeion*, que significa ‘signo’, e *sêma*, que pode ser traduzido por ‘sinal’ ou ‘signo’.” Esse termo foi primeiramente utilizado na medicina pelo médico grego Galeno de Pérgamo (139-199) ao referir-se à diagnóstica como sendo ‘a parte semiótica’ (*semeiotikón méros*) da medicina.” Segundo Nöth (1995), a semiótica propriamente dita iniciou-se com filósofos como John Locke (1632-1704) que, no seu “Essay on human understanding”, de 1690, postulou uma doutrina dos signos denominada Semeiotiké, e com Johann Heinrich Lambert (1728-1777) que, em 1764, foi um dos primeiros filósofos a escrever um tratado específico intitulado Semiotik. Enquanto a semiótica medieval desenvolveu-se no âmbito da teologia, da gramática, da retórica e da dialética, a partir do século XIX surgem as modernas teorias de significado, sentido e referência da semântica linguística e tem início o questionamento científico da linguagem.

Embora a Semiótica finque suas raízes na Grécia antiga, e tenha também recebido, ao longo da idade média e do renascimento, os aportes de inúmeros pensadores tais como São Tomaz de Aquino, Leibniz, Kant, Locke, Descartes, entre outros, o período moderno da Semiótica (século XX) é inaugurado por Edmund Husserl (1859-1938) com a sua teoria fenomenológica dos signos e significados. Mas a sedimentação mais importante da teoria resultou dos trabalhos simultâneos (e independentes) de Charles Sanders Peirce, nos Estados Unidos, de Viessé-Iovski e Potiebníá, na Rússia e de Ferdinand de Saussure, na Suíça. A partir desses trabalhos, que todavia se distinguiam tanto em abordagem quanto em amplitude, a teoria passou a ser desenvolvida ou reinterpretada por outros semiólogos e semioticistas ao longo de todo o século XX, com particular destaque para os trabalhos de Hjelmslev, Jakobson, Morris, Greimas e Eco. Essas contribuições levaram à consolidação da Semiótica como uma meta-ciência, visto que contribuições mais recentes ampliaram sua abrangência, inicialmente restrita à linguística, à medicina e à filosofia, a áreas tão variadas quanto a da biologia, da física, e finalmente a da cibernética e dos sistemas de controle. Essas últimas iniciativas, por sua vez, descortinaram os horizontes de implementações de sistemas computacionais baseados em estruturas semióticas, como o objeto deste estudo.

Assim, segundo Nöth (1995), a “Semiótica é a ciência dos signos e dos processos significativos (semiose) na natureza e na cultura”, e suas principais linhas são:

- Peirceana (Peirce), cujo foco de atenção está na universalidade epistemológica e metafísica.
- Estruturalista/Semiologia (Saussure; Lévi-Strauss; Barthes; Greimas), cujo foco de atenção está nos signos verbais.
- Russa ou Semiótica da cultura (Jakobson; Hjelmslev; Lotman), cujo foco de atenção situa-se na linguagem, na literatura e em outros fenômenos culturais, tais como a comunicação não verbal e visual, mito, religião.

Segundo Santaella (2002), a Semiótica é um dos membros da tríade das ciências normativas – estética, ética e lógica ou semiótica –, cuja função é estudar ideais, valores e normas. Enquanto a estética se ocupa dos ideais que guiam nossos sentimentos, a ética trata dos ideais que orientam nossa conduta, e a lógica estuda os ideais e normas que conduzem o pensamento. Ao constatar que o pensamento não pode desenvolver-se somente por meio de símbolos, nem mesmo o raciocínio matemático puro, Peirce compreendeu a importância das demais modalidades de signos, e isso o levou a buscar a extensão da lógica para uma teoria mais ampla, chamada Semiótica, capaz de tratar não somente das leis do pensamento e das condições da verdade, mas também de sua evolução, o que pressupunha compreender as condições gerais dos signos, a transmissão de significado de uma mente para outra e de um estado mental para outro. Em razão dessa diversidade de tarefas, a lógica ou semiótica têm, segundo Santaella (2002), três ramos que se ocupam de diferentes questões:

- **Gramática especulativa:** estuda todos os tipos de signos e as formas de pensamento que eles possibilitam;
- **Lógica crítica:** estuda os tipos de inferências, raciocínios ou argumentos que se estruturam com base nos diversos signos;
- **Metodêutica ou retórica especulativa:** analisa os métodos a que cada tipo de raciocínio dá origem, com base na validade e na força própria de cada argumento.

A gramática especulativa fornece as definições e classificações para a análise de todos os tipos de linguagens, signos, sinais, códigos etc., de qualquer espécie, e de tudo que está neles implicado — a representação em seus três aspectos: *significação*, *objetivação* e *interpretação* (Santaella, 2002). Assim, Peirce ampliou a noção de signo, concebendo-o como uma relação triádica:

Defino um Signo como qualquer coisa que, de um lado, é assim determinada por um Objeto e, de outro, assim determina uma ideia na mente de uma pessoa, esta última determinação, que denomino o Interpretante do signo, é, desse modo, mediatemente determinada por aquele Objeto. Assim, um signo tem uma relação triádica com seu objeto e com seu interpretante. (8.343).

Em outras palavras, signo ou *representâmen* é “aquilo que, sob certo aspecto ou modo, representa algo para alguém”. Uma relação triádica pressupõe o signo ou *representâmen*, o objeto a ser representado e o interpretante (um ser vivo, um computador, etc.) (Santaella, 2000). Logo, as três dimensões da significação são:

- **Signo:** alguma coisa que está no lugar de outra, em relação à ideia que esta produz na mente de um intérprete. É o meio pelo qual um objeto ausente acaba por produzir uma ideia na mente do intérprete;
- **Objeto:** não necessariamente material ou com existência real no mundo, pode ser inclusive uma outra ideia;
- **Interpretante:** sempre uma ideia na mente de um intérprete; atuará como signo num futuro processo interpretativo; mediador na relação do signo com o objeto.

A Figura 4.1 ilustra a relação entre signo, objeto e interpretante.

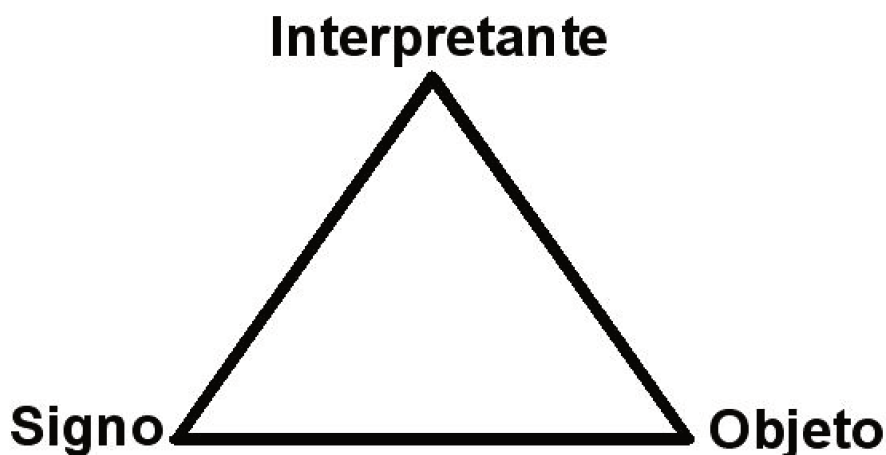


Figura A.2.1: As três dimensões do signo na Semiótica de Peirce

Santaella chama ainda a atenção para o fato de que, estando fundamentada na fenomenologia, a Semiótica de Peirce não se limita aos chamados signos genuínos, e inclui também os quase-signos, que sequer têm a natureza plena de uma linguagem (palavras, desenhos, etc.), e podem ser uma mera ação ou reação, emoção ou sentimento, por mais vago que seja. Assim, qualquer coisa que esteja presente à mente tem a natureza de um signo e dá corpo ao pensamento, às emoções, reações etc. Tanto quanto o próprio signo, o objeto do signo também pode ser qualquer coisa de qualquer espécie. O que faz dela um objeto é o fato de estar sendo representada por um signo. Em outras palavras, o que define signo, objeto e interpretante é a posição lógica que cada um deles ocupa no processo representativo (Santaella, 2002).

Peirce distinguiu dois tipos de objeto, o dinâmico e o imediato. O primeiro é aquilo que o signo sugere, indica ou representa, e o segundo é o modo como o signo sugere, indica ou representa o objeto dinâmico. Só o objeto imediato nos é acessível, pois na sua função mediadora é ele que nos coloca em contato com a realidade (Santaella, 2002).

Ainda de acordo com Peirce, os signos se diferenciam dependendo da relação entre os elementos que os compõem e a sua ação específica (ou semiose). Essa diferenciação determina três categorias, isto é, uma tríade ou tricotomia. As três categorias de signos, isto é, as três tricotomias, estão descritas a seguir.

Primeira tricotomia (signo com relação a si mesmo)

As três categorias fenomenológicas (qualidade, existência e lei) estão na base dos signos. No primeiro elemento da tríade, quando na relação do signo consigo mesmo (no que toca à natureza do seu fundamento ou daquilo que lhe capacita a atuar como signo) predominar sua qualidade intrínseca, ele se classificará como um *quali-signo*; quando se destacar o seu aspecto de existência concreta, ele pertencerá à categoria dos *sin-signos*; e quando preponderar o caráter de lei, ele será um *legi-signo*:

- **Quali-signos** (primeiridade): correspondem a uma qualidade e ao poder de sugestão que essa mera qualidade apresenta (exemplo: azul-claro sugere céu, etc.);

- **Sin-signos** (secundidade): o signo existe na natureza e no mundo real e a existência funciona como signo, numa síntese de múltiplas determinações frente a ou do ponto de vista de outros existentes;
- **Legi-signos** (terceiridade): o signo não existe na natureza, e é uma lei geral que subordina os existentes na natureza; por exemplo, as palavras são leis porque pertencem a um sistema fora do qual deixam de ter sentido.

Segunda tricotomia (signo com relação ao seu objeto)

Dependendo do fundamento do signo em questão, a maneira como ele pode determinar seu objeto será diferente, e também diferente será a relação do signo com seu objeto dinâmico, aquele que determina o signo e que é representado por ele. Assim, no segundo elemento da tríade em função da relação com o seu objeto dinâmico o signo pode ser classificado como ícone, índice e símbolo:

- **Ícone** (primeiridade): seu fundamento é um quali-signo, e por isso ele apresenta em si próprio qualidades do objeto por ele representado, seu objeto dinâmico. Este só pode ser evocado ou sugerido pelo objeto imediato (que tem natureza descritiva), seja por meio de uma imagem (similaridade na aparência), de um diagrama (similaridade nas relações internas) ou de uma metáfora (similaridade no significado do representante e do representado). A relação entre signo e objeto independe do interpretante e do próprio objeto;
- **Índice** (secundidade): seu fundamento é um sin-signo, e por isso seu objeto imediato (que tem natureza designativa) mantém uma conexão física ou relacional com seu objeto dinâmico, que é por ele indicado. A relação entre signo e objeto independe do interpretante, mas depende do objeto, cuja existência concreta fundamenta o índice. Todos os índices envolvem ícones;
- **Símbolo** (terceiridade): seu fundamento é legi-signo, e por isso seu objeto imediato (que tem natureza copulante) mantém uma relação totalmente arbitrária com seu objeto dinâmico, que é por ele representado. A relação depende intrinsecamente do interpretante e do objeto. Todo símbolo inclui dentro de si quali-signos icônicos e sin-signos-indiciais.

Terceira tricotomia (signo com relação ao seu interpretante)

No terceiro elemento da tríade, é possível compreender como uma mente real ou potencial pode interpretar os signos. Enquanto o objeto que determina o signo se divide em imediato e dinâmico, segundo uma relação dual, a relação do signo com o interpretante é completa, isto é, triádica, pois há pelo menos três passos para que se complete o percurso da interpretação: interpretante imediato, dinâmico e lógico. O primeiro é interno ao signo e refere-se ao seu potencial interpretativo ou interpretabilidade, é algo que diz respeito ao signo em sua objetividade. O segundo refere-se ao efeito do signo na mente de cada intérprete em particular (dimensão psicológica), e pode se dar em três níveis: emocional (qualidade ou sentimento, em geral produzidos por ícones), energético (ação física ou mental, em geral produzidas por índices) ou lógico (associação de ideias por meio de símbolos e suas regras interpretativas). Por fim, o terceiro nível é o interpretante final, que se refere à interpretação a que todo intérprete tenderia a chegar no infinito (Santaella, 2002).

Na relação do signo com o interpretante final há uma vez mais três níveis de interpretante: rema, dicente ou argumento, conforme descrito abaixo:

- **Rema** (termo): seu interpretante é uma mera possibilidade (condições possíveis de verdade), não indo além de uma hipótese interpretativa, uma possibilidade qualitativa. Assim, os interpretantes remáticos são produzidos por quali-signos icônicos que sugerem hipoteticamente uma qualidade do objeto (Santaella, 2002);
- **Dicente** (proposição): seu interpretante é um existente, podendo ser afirmado como verdade ou falsidade (condições reais de verdade) e, portanto, não pode ser um ícone, uma vez que este não fundamenta interpretação de que algo se refere a uma existência real. Por isso os dicentes são interpretantes de sin-signos indiciais (Santaella, 2002);
- **Argumento**: seu interpretante é uma lei, devendo ser afirmada como verdade ou negada como falsidade (condições necessárias de verdade). O signo é interpretado por meio de uma regra interpretativa interna ao intérprete. Em razão disso, a base dos interpretantes argumentativos está nas sequências lógicas de que depende o legi-signo simbólico. Sem essas regras o símbolo não poderia significar (Santaella, 2002).

Os três tipos de argumentos

Um argumento é um tipo de signo composto por duas proposições acopladas por uma relação de implicação, de tal forma que a verdade da primeira proposição (premissa) implicará necessariamente a verdade da segunda proposição (conclusão). Há três tipos de argumentos, conforme descrição a seguir:

- **Abdução** (primeiridade): conclusão potencialmente verdadeira (criativo);
- **Dedução** (secundidade): conclusão realmente verdadeira;
- **Indução** (terceiridade): conclusão geralmente verdadeira.

A Semiótica de Peirce contribuiu decisivamente para ampliar o entendimento da abdução, que, segundo Peirce, é a única forma de raciocínio que produz conhecimentos novos. Isso deixa claro portanto a importância indiscutível da abdução em qualquer teoria sobre o aprendizado, inclusive aprendizados linguísticos.

Como neste trabalho temos necessidade de compreender não somente o aprendizado dos signos verbais correspondentes à L2, mas também o reconhecimento de signos visuais que servirão de apoio para esse aprendizado, buscaremos na semiótica de Peirce as referências para compreender a interação desses diversos universos sógnicos, sobretudo em situações em que signos visuais permitem a inferência dos significados dos signos textuais.

Segundo Fann (1970), Peirce via todos os processos mentais como sendo inferenciais. Assim, “inferência” inclui não somente dedução e indução, mas também a abdução, uma hipótese, que é “uma operação sobre dados que resulta em cognição” ou “uma explicação de um fenômeno por meio de uma hipótese.” Quando percebemos que uma previsão após a outra é confirmada pela experiência, passamos a destacar a hipótese.

Cunningham (2001) comenta que esses três tipos de argumentos, ou de inferência, são as maneiras fundamentais de se preservar a crença ou de passar da dúvida para a crença. Segundo ele, a semiose, ou ação dos signos, é o processo de se aplicar signos para entender fenômenos (indução), raciocinar de signo para signo (dedução), e/ou usar signos para dar sentido a alguma experiência nova (abdução). Esses modos de inferência são cíclicos e

caracterizam o desenvolvimento dos *Umwelten* durante toda a vida: os signos são inventados ou descobertos para dar conta das experiências.

Todo processo de comunicação pressupõe um meio através do qual uma mensagem é conduzida de uma cognição para a seguinte. Essa função mediadora do signo levou Peirce a defender a ideia de que um signo é um tipo de “meio de comunicação” entre duas ideias ou entre um objeto e uma ideia (interpretante que o signo produz ou modifica).

Do ponto de vista da semiótica Peirceana, o aprendizado está relacionado em sua essência ao fluxo do tempo, sendo assim um processo contínuo. O aprendizado é raciocínio, interpretação e representação (De Tienne, 2002). Segundo Houser, a teoria sêmica de Peirce foi de uma “fundamental importância para uma correta teoria do aprendizado”, e seu poder explanatório se deve ao “foco na estrutura triádica da ação do signo” e à consequente capacidade de, como uma teoria consistente e até mesmo completa, explicar de maneira muito eficaz muitos fatos relativos ao aprendizado, como o papel desempenhado pelos conhecimentos anteriores, pelas metáforas e analogias (Houser 1987, p.270-271, *apud* De Tienne, 2002). Como os signos são a condição de possibilidade do fenômeno mental, o entendimento da mente pressupõe o entendimento dos signos, e a visão de Peirce era genérica o bastante para falar de uma “quase-mente”, não a instância mental particular que cada indivíduo possui, nem mesmo uma “mente social” que alguns indivíduos podem compartilhar, mas um fenômeno (ou uma entidade) mental mais geral, cujo funcionamento a semiótica permite analisar. E é essa a razão para a semiótica preceder a psicologia, e para a visão Peirceana postular que aquilo que experimentamos como “mente” é o que é não por valer-se de signos, mas por ser feita de signos (De Tienne, 2002).

Causas finais são possibilidades gerais que podem se realizar no futuro, leis gerais que ditam a direção geral que sequências particulares de eventos podem tomar para que o processo do qual são constituintes assumam uma identidade cada vez maior ao longo do tempo, tomando a forma de uma corporificação da ideia geral representada pela causa final. A ideia geral de fazer uma torta de maçã, por exemplo, guiará a sequência de ações que tendem a produzi-la, mas não ditará com precisão qual receita deve ser usada, nem a quantidade, que ingredientes e em qual ordem usar, nem tampouco o tempo de preparo. Todos esses fatores poderão variar dentro dos limites permitidos pela ideia geral. Assim, a causa final somente indica uma certa tendência, mas não tem poder de ditar a concatenação exata de ações e reações que devem ter lugar a fim de que o futuro venha como o esperado (De Tienne, 2002). Da mesma forma, aprender é crescer dentro dos limites de um plano condicional geral cuja natureza dá uma ordem especial e identidade ao que quer que seja que dela resulte, e cuja realização cria personalidade, como disse Peirce no artigo “The Law of Mind” (De Tienne, 2002).

Raciocinar é passar de uma crença para outra. Todo raciocínio, seja abutivo, dedutivo ou indutivo é composto de uma sequência de proposições (premissas) na qual alguma ideia que ou era desconhecida ou não totalmente revelada é trazida à luz em virtude de transitar através da sequência de premissas. Cada premissa representa uma crença de algum tipo, seja particular ou universal, (...) que traz outra [premissa] em virtude de sua própria associação com outras crenças já posicionadas dentro da sequência cuja identidade geral é ditada pelo que Peirce às vezes chama de “princípio guia”.

Apêndice 3

Resumo das principais características do esperanto

A proposta do esperanto foi publicada em 1887, quando o doutor Lejzer Ludwik Zamenhof editou em língua russa um livro com o título “Internacia Lingvo – Prefácio e Método Completo”. O nome esperanto foi adotado universalmente porque o autor assinou o livro com o pseudônimo *Doktoro Esperanto* (Eco, 2001). A gênese da língua teve início por volta de 1870, com as primeiras tentativas de Zamenhof, ainda adolescente, de criar uma língua que fosse mais acessível às pessoas comuns e das diversas origens culturais. Seus esforços se estenderam por vários anos, culminando em pelo menos três versões sucessivas, até que uma versão mais amadurecida e bem resolvida tomou forma em 1887.

Nuessel (1995) comenta que as línguas planejadas (LP) podem ser classificadas em duas categorias básicas: línguas *a posteriori* e línguas *a priori*. A distinção das duas categorias pode ser resumida nos seguintes termos (Large, 1985: 51):

Enquanto uma língua *a priori* é composta de elementos inteiramente inventados e não encontrados em qualquer língua existente, e é geralmente baseada em uma classificação lógica de ideias, uma língua *a posteriori* é baseada em elementos de gramática, vocabulário e sintaxe extraídos de uma ou mais línguas naturais. Embora na prática a maioria das línguas universais tenha incluído uma mistura das duas abordagens, é em geral possível e conveniente situar cada esquema específico em um ou outro desses tipos gerais.

Uma subcategoria das línguas planejadas é formada pelas línguas internacionais auxiliares (LIA), cuja função essencial é resolver o problema da comunicação internacional. Em contraste com as línguas filosóficas dos séculos XVII e XVIII, que eram eminentemente *a priori*, é razoável admitir que as LIAs sejam *a posteriori*, pois, segundo Eco (2001), elas nascem da comparação com as línguas naturais, das quais tentam ser versões racionalizadas (ainda que preservando o léxico), sendo assim uma síntese equilibrada das línguas naturais.

O esperanto foi concebido visando à neutralidade e à facilidade de aprendizado especificamente para facilitar as relações e a convivência entre nações e culturas, e assim fortalecer os laços entre os diferentes povos. Daí se depreendem dois de seus atributos essenciais: o de ser uma língua planejada (LP) e uma língua internacional auxiliar (LIA).

Segundo Nuessel (1995), o esperanto pode ser classificado como língua planejada *a posteriori*, visto que seus elementos léxicos, gramaticais e sintáticos foram emprestados de línguas étnicas. Porém, ele difere dessas línguas por ter nascido de uma proposição de sua gramática, e não como uma língua oral que depois se desenvolveu em um conjunto de símbolos gráficos com um subsequente compêndio de sua gramática.

A despeito de sua simplicidade superficial, o esperanto possuiria, segundo (Balbin, 1982; 1987; Golden, 1992; Schubert, 1989), vários atributos que se aproximam de vários grupos linguísticos e sua gramática teria características de línguas inflexionais, aglutinativas e isolantes (Golden, 1992; Jordan, 1987; Piron, 1981). Ainda segundo Nuessel (1995):

A língua esperanto é um signo criado que funde as propriedades linguísticas das principais famílias linguísticas (românica, germânica e eslava). Essa língua é uma invenção, no sentido etimológico da palavra, isto é, ela é a descoberta de Zamenhof de uma nova forma de comunicação linguística planejada visando à facilidade de aquisição por meio de um projeto gramatical simplificado, e derivada do vocabulário de um amplo espectro de línguas étnicas.

Como toda língua, o esperanto se baseia em signos arbitrários, isto é, em símbolos. E embora os símbolos iniciais tenham sido convencionados por Zamenhof (a partir de outras línguas), com a difusão da língua a partir de 1887 a adoção de novos símbolos passou a decorrer da dinâmica de uso mais amplo da língua, e resultar da necessidade natural de acompanhar a crescente complexidade da sociedade moderna. Nesse particular, não há grandes diferenças com relação ao processo de evolução das demais línguas, exceto, talvez, pela busca de se atender a critérios de internacionalidade e economia do léxico.

Assim, após mais de um século de uso no seio de sua comunidade internacional, a língua passou, segundo Nuessel (1995), do estágio de signo criado para o de signo vivo:

Em termos semióticos, isto é, no âmbito da clássica tríade semiótica, signo, objeto e interpretante, a língua esperanto é o signo ou representação, seu uso em contextos diversos é o objeto ou referente, e seus usuários ou os proponentes do esperanto como língua internacional auxiliar são seus interpretantes. Como signo, o esperanto “representa” uma variedade de significados acordados a essa língua por seus interpretantes. Esses significados podem incluir a significação normatizada da língua (linguagem comercial, turística, de conferências) e seu significado orientado a valor (paz mundial, justiça e unificação da humanidade).

Em termos gerais, as línguas planejadas podem ser classificadas segundo sua orientação mais naturalista ou mais esquemática (Mattos, 1987:1). No primeiro caso, o objetivo é que sejam tão parecidas quanto possível com um certo número de línguas étnicas; no segundo caso, procura-se garantir que sejam mais fáceis de se aprender, independentemente da língua do aprendiz. Embora toda língua planejada tenha algo de esquemático, visto que resultam invariavelmente de uma certa simplificação, algumas têm feições claramente naturalistas. Nesse particular, o esperanto é, segundo (Mattos, 1987:1), moderadamente esquemático, dentro de uma estratégia de aliar uma estrutura extremamente regular a um léxico formado por termos internacionais, compartilhados por vários idiomas e que possuem etimologia comum. Embora quase todas as línguas possuam tais interônimos, segundo Szerdahelyi (1987), a diferença essencial entre sua utilização no esperanto e nas línguas étnicas é que, enquanto nestas últimas eles somente complementam o léxico básico, no esperanto todo o léxico básico originou-se deles. Essa predominância implica que aquele que aprende o esperanto encontra uma abundância de termos cognatos aos de sua L1 ou aos de uma outra L2 já conhecida. Isso tem um nítido efeito na facilitação do aprendizado e na retenção do vocabulário do idioma. Mas muito embora toda LIA seja predominantemente *a posteriori*, a fim de tirar proveito dessas semelhanças lexicais e sintáticas com línguas étnicas, e assim facilitar seu aprendizado como L2, os benefícios dessa abordagem podem

variar conforme o caso, dependendo dos critérios adotados, isto é, do grau de esquematismo imposto aos vocábulos. Enquanto uma intervenção comedida tende a preservar o aspecto internacional do *corpus*, a interferência excessiva no léxico pode fazer com que sejam perdidas todas as referências etimológicas, anulando o efeito mnemônico dos termos de etimologia comum, conforme comenta Eco (2001) acerca do volapük, LIA proposta em 1879, na Alemanha, por Johann Martin Schleyer:

Ele não é de fato *a priori*, pois empresta as raízes de línguas naturais; mas tampouco é *a posteriori*, pois ele as deforma sistematicamente, com base em decisões *a priori*, eliminando seu aspecto cognato. Por sua tendência a não se parecer com qualquer língua conhecida, ele cria dificuldades para todos seus usuários.

Assim, embora o uso de interônimos também tenha ocorrido no volapük, neste as raízes sofreram deformações que resultaram na eliminação de toda e qualquer associação com os lexemas originais, como mostram os exemplos a seguir (Passini, 1993):

Tabela A.3.1: Exemplos de modificação de raízes

Português	Volapük	Esperanto
Europa	Yulop	Eŭropo
América	Melop	Ameriko
Brasil	Basil	Brazilo

A esse respeito, Mattos (1987) comenta que em 1880 Zamenhof já havia finalizado uma primeira versão de sua língua internacional, quando tomou conhecimento do volapük, recém lançado, e observou que entre os critérios adotados por Schleyer em seu projeto estavam a eliminação de vogais iniciais e finais e a redução do número de sílabas das palavras. Como essa estratégia de encurtamento dos vocábulos tem, em princípio, algumas vantagens em termos de economia, Zamenhof decidiu pôr à prova no seu próprio projeto os métodos utilizados por Schleyer na concepção do volapük. Ao fazê-lo, se deu conta dos inconvenientes que as excessivas deformações das raízes traziam para a inteligibilidade do idioma e, diante disso, retomou as premissas iniciais de seu projeto. Em 1887, ele publicou a versão final da *lingvo internacia*, mostrando um prudente compromisso entre naturalismo lexical e regularidade gramatical. Consciente ou intuitiva, a valorização por Zamenhof das referências etimológicas é hoje justificada pela noção de que para o sistema perceptivo deformações excessivas nas raízes cognatas implicam a perda de seu valor mnemônico.

Assim, diferentemente do que se dá com o volapük, grande parte dos conhecimentos lingüísticos dos aprendizes são referências válidas na aquisição e na memorização do léxico do esperanto, o que sustenta a referencialidade do vocabulário. Isso é sobretudo verdadeiro para sujeitos cuja L1 é de origem indo-européia, como comenta Eco (2001):

(...) analisando o vocabulário do esperanto, o falante de qualquer língua europeia encontrará: (i) vários termos cognatos, pois são iguais ou muito parecidos aos de sua língua; (ii) outros, estrangeiros, que ele de algum modo já conhece; (iii) alguns à primeira vista incompreensíveis, mas que, após o aprendizado de seu significado, tornam-se reconhecíveis; e, por fim (iv) um

número aceitavelmente pequeno de termos desconhecidos que devem ser aprendidos por completo.

Para os aprendizes cuja L1 não é indo-europeia, o aprendizado do esperanto como L2 não necessariamente tira proveito de um grande número de palavras cognatas, embora também outros grupos linguísticos tenham contribuído com vários interônimos adotados no esperanto: *kanguruo*, *kimono*, *algebro*, *ĝiraf*o, etc. Além disso, como muitos dos falantes de línguas pertencentes a outros grupos linguísticos na Ásia e na África estudam uma língua indo-europeia (inglês, francês, russo, português) como língua estrangeira, o acesso ao léxico indo-europeu não é incomum. Na Figura 6.1 é ilustrado um caso em que a adoção de um interônimo no esperanto facilita o reconhecimento por parte de aprendizes com diversas L1:

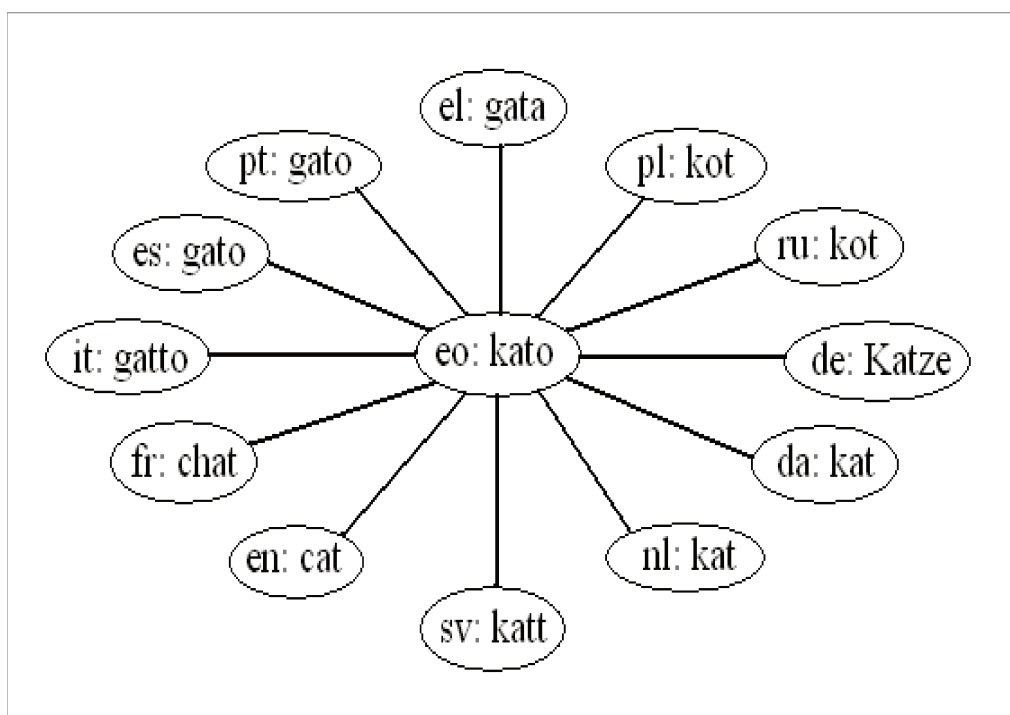


Figura A.3.1: Exemplo de interônimo - a palavra *kato*

Mas qual seria, em termos semióticos, o efeito dos interônimos e da consequente presença de termos cognatos no aprendizado de uma L2 em geral, e do esperanto em particular? Como toda linguagem humana é convencionada, arbitrada, ela se baseia em legi-signos simbólicos. Por outro lado, todo símbolo inclui dentro de si quali-signos icônicos e sin-signos indiciais (Santaella, 2002: p.23). Excetuando-se o caso do bilinguismo infantil, nos demais processos de aprendizado de uma L2 o sujeito já terá domínio de uma L1. Assim, se os símbolos da L2 guardam semelhanças com seus correspondentes na L1, em razão da cognação, a relação de similaridade entre essas palavras é icônica e não simbólica, visto que a L1 e a L2 são convenções linguísticas independentes. Por esse prisma, a iconicidade é a principal referência no reconhecimento dos termos cognatos, e isso se dá

tanto em razão da semelhança das imagens dos lexemas na L2 com os da L1, quanto em termos diagramáticos, visto que as relações internas (letras, fonemas, sílabas, etc.) dos dois vocábulos podem exibir um certo grau de correspondência funcional.

Por fim, mesmo que a etimologia do esperanto não seja muito familiar a asiáticos e africanos, outros aspectos da língua, em particular o amplo uso de referências indiciais, como discutido a seguir, têm, segundo Piron (2002), muito em comum com várias línguas não indo-europeias, como o turco, o indonésio e o chinês, o que também facilita seu aprendizado por alunos provenientes daqueles contextos culturais.

Relações de indicialidade na formação do léxico do esperanto

No aprendizado ou na aquisição de uma língua, o pleno domínio do léxico implica a assimilação de uma grande quantidade de símbolos (palavras). Embora inevitável, esse processo pode ser notavelmente otimizado nos casos em que a estrutura da L2 permite um amplo uso de referências indiciais na derivação de símbolos correlatos. Assim, embora a iconicidade dos signos cognatos tenha um papel importante na facilitação do aprendizado de qualquer língua, no esperanto essas relações icônicas de primeiro nível são amplificadas pela validade absoluta das referências indiciais no nível seguinte. Isso contrasta fortemente com o que ocorre nas línguas ocidentais (flexionais) e em várias LIAs de índole claramente naturalista, nas quais, ainda que um símbolo possa ser assimilado iconicamente, a derivação de novos símbolos a partir dos já conhecidos raramente pode se basear em índices diretos, estando sempre sujeita a exceções e irregularidades. Na prática isso significa que quase toda a carga simbólica desses idiomas deve ser aprendida/adquirida incrementalmente. Para entender isso, é preciso lembrar que, como salienta Santaella (2002: p.32), todo sin-signo indicial é, em alguma medida, uma atualização de um legi-signo, sendo um caso particular de uma classe mais geral. Como discutido, formas mais complexas de referência são construídas sobre formas mais simples: as referências indiciais dependem das referências icônicas e as referências simbólicas dependem das indiciais.

Para ilustrar o efeito da indicialidade na assimilação lexical, em português (como nas demais línguas ocidentais), para aprender/adquirir os adjetivos referentes a substantivos já assimilados, o aluno deve memorizar um a um cada novo vocábulo, uma vez que não existe uma *regra geral* de formação e, em razão da inconstância dos mecanismos de indicialidade, um símbolo aprendido em geral não indica claramente o caminho para os símbolos derivados/correlatos. Isso eleva enormemente o esforço necessário para que os alunos dominem a “carga simbólica” da L2. Numa situação equivalente, o aprendizado ou a aquisição de adjetivos em esperanto exigem esforço muito menor, pois as raízes (símbolos) já assimiladas podem ser reutilizadas sem restrição para produzir os correspondentes adjetivos, bastando para isso substituir a desinência substantiva *-o* pela desinência adjetiva *-a* (que estabelece a relação indicial com o conceito substantivo denotado pela adjunção da raiz à desinência *-o*), como ilustrado na Tabela A.3.2:

Tabela A.3.2: Formação dos adjetivos em português e em esperanto

Português		Esperanto	
Substantivo	Adjetivo	Substantivo	Adjetivo
irmão	fraternal	frato	frata
amigo	amistoso	amiko	amika
tio	avuncular	onklo	onkla
cachorro	canino	hundo	hunda
gato	felino	kato	kata
carneiro	ovino	ŝafo	ŝafa
casa	doméstico	domo	doma
estrela	estelar	stelo	stela
fala	oral	parolo	parola
ajuda	auxiliar	helpo	helpa
mão	manual	mano	mana
orelha	auricular	orelo	orela
bispo	episcopal	episkopo	episkopa
ilha	insular	insulo	insula
cidade	urbano	urbo	urba
água	aquático, aquoso , hídrico	akvo	akva
fogo	ígeo, ardente, incandescente	fajro	fajra
exército	militar	armeo	armea
arte	artístico	arto	arta
pássaro	ornitológico	birdo	birda
sonho	onírico	sonĝo	sonĝa
dinheiro	pecuniário, monetário , financeiro	mono	mona
mar	marítimo	maro	mara
lua	lunar	luno	luna
porco	suíno	porko	porka
Carga simbólica: 56		Carga simbólica: 25	

A Tabela A.3.2 traz somente uma amostra limitada de casos, mas ela é útil para exemplificar dois mecanismos fundamentais no aprendizado ou na aquisição de uma L2: os efeitos da iconicidade e da indicialidade na assimilação do léxico da L2. Assim, para aquele cuja L1 é, por exemplo, o português, o aprendizado do esperanto como L2 tira proveito de um grande número de raízes iconicamente relacionadas às da L1. Essas raízes podem tanto se assemelhar a substantivos, indicados em negrito na primeira coluna da tabela, quanto a adjetivos, destacados na segunda coluna. Na amostra acima, de um total de 25 conceitos, 17 podem ser iconicamente associados a seus correspondentes na L1, e, caso o aluno já saiba uma outra L2, como o inglês, francês, italiano ou alemão, o reconhecimento de raízes tais como *onkl-*, *help-*, *parol-* e *arme-* amplia o efeito da iconicidade na aquisição do esperanto.

Assim, a relativa facilidade de assimilação do léxico do esperanto, ao contrário de outras línguas, não se deve somente ao valor icônico (cognato) que várias raízes apresentam em razão dos conhecimentos prévios do aluno (L1 ou outra L2) mas, sobretudo, ao fato de

que cada raiz aprendida/adquirida permite produzir um elevado número de outros conceitos, derivados por meio de mecanismos indiciais simples, como o ilustrado na Tabela A.3.2: a desinência *-o* denota conceitos substantivos, e da desinência *-a* produz os correspondentes adjetivos. Esse potencial não se esgota aí, visto que outras desinências e sufixos permitem estender o alcance do léxico: a desinência *-e* forma advérbios e o sufixo *-et* e *-ar*, só para citar alguns, formam respectivamente os diminutivos e coletivos.

Os atributos de iconicidade e indicialidade permitem uma clara redução da carga simbólica exigida para a assimilação do léxico do esperanto e, de resto, reforçam estruturas neuronais correspondentes a cada termo, visto que a derivação de conceitos sempre a partir de uma mesma raiz ou germe fortalece as conexões dos neurônios especializados em reconhecer aqueles agrupamentos. A possibilidade de se utilizar relações indiciais para, a partir dos símbolos já aprendidos (com ou sem recurso à iconicidade), produzir toda uma série de palavras derivadas, torna o manejo do léxico extremamente eficaz, como discutido na seção 6.2.

De resto, a presença de um estoque de raízes internacionais dá aos aprendizes do esperanto como L2 uma considerável ajuda na aquisição de uma L3, em particular se esta for de origem indo-europeia. Esses dois fatos explicam o notável valor propedêutico do esperanto descrito em (Piron, 2002). Se, por outro lado, o esperanto é visado como uma L3, aprendida após uma L2 indo-europeia, o efeito de correspondência léxica se mantém válido, e os aprendizes podem reconhecer as raízes cognatas e, diferentemente do que se deu no aprendizado da L2, podem somar a isso os recursos de indicialidade que o esperanto possibilita. Em razão disso, cada vez que um símbolo é assimilado, as formas derivadas podem ser automaticamente generalizadas pelo aprendiz, sem que os símbolos correlatos (representantes de noções semânticas derivadas) tenham que ser aprendidas elemento por elemento. Como nota Deacon (1997), aprender é função da probabilidade das correlações.

Em suma, o aspecto mais fundamental da estruturação do esperanto, confirmado em ensaios psicolinguísticos (Piron, 2002), reside no fato de que seu funcionamento respeita o mecanismo da assimilação generalizadora. Nesse ponto o esperanto é mais *natural* para o sistema nervoso (Piron, 2002), dada sua estrutura mais regular, o que reduz o esforço necessário para a formação de todas as relações simbólicas necessárias ao domínio da língua. Ele se estrutura com base no princípio da combinatória infinita, com lexemas invariáveis que podem ser justapostos sem restrição, com regularidade elevada. Como discutido, quando novos estímulos exibem características comuns ao conjunto de estímulos já conhecidos, sem inconsistências, esses estímulos potenciais são também aprendidos. Isso torna a assimilação e a utilização do esperanto mais fáceis do que a das línguas com estruturação irregular. A regularidade, em vez de significar pobreza de recursos, na verdade permite a flexibilidade e a adaptabilidade do idioma a partir de uma quantidade moderada de símbolos. Ou, nas palavras de Eco (2001), enquanto as línguas *a priori* e as gramáticas lacônicas visam tanto quanto possível à economia, o esperanto visa ao princípio da otimização.

A estruturação do esperanto

O esperanto é uma linguagem simbólica (formada por símbolos) na medida em que seu léxico e sua gramática, como os de qualquer outra língua, resultam de uma convenção arbitrária, conforme a própria definição de símbolo. Mas ao contrário do que ocorre na maior parte das línguas, no esperanto as convenções arbitrárias são balizadas pelo princípio do “necessário e suficiente” (Saussure⁴, 1910). Isso significa que sua carga simbólica é tão leve quanto possível, o que não implica necessariamente uma carga simbólica pequena, tendo-se em conta que ser capaz de veicular todas as nuances do pensamento humano pressupõe riqueza de repertório. Todavia, no esperanto essa variedade de possibilidades de expressão é obtida de forma muito eficaz. As principais razões para isso são:

- Léxico fundamental reduzido, extensível por meio de afixos;
- Gramática regida pela assimilação generalizadora (combinatória infinita);
- Alfabeto fonético (relação unívoca entre cada letra e cada fonema).

Quanto ao léxico, é importante lembrar que a carga simbólica é reduzida ainda mais em decorrência do uso, no esperanto, de raízes internacionais, isto é, raízes comuns a um grande número de línguas. Esse fato faz com que o processo de aprendizado do vocabulário possa se valer de referências diretas a encadeamentos já existentes na memória: ao nos depararmos com a raiz *elefant-* uma correspondência imediata é feita com o símbolo já existente na língua materna do aprendiz. Isso não é verdadeiro no caso de um aprendizado como primeira língua, nem nos casos de indivíduos em cuja língua materna o radical não seja cognato. Assim, para os propósitos do ambiente de aprendizado aqui proposto, voltado para aprendizes de qualquer origem cultural, todo o léxico será construído com base em referências icônicas e indiciais. Assume-se com isso que a validade da base icônica é muito mais abrangente que a dos radicais internacionais, cuja validade é mais restrita. Mas em qualquer dos casos a assimilação de um dado radical desencadeia toda uma série de palavras derivadas por mecanismos generalizáveis.

A estrutura básica da palavra em esperanto é descrita a seguir (Mattos, 1987:2):

Prefixo+raiz+sufixo+desinência

Exemplos:

Dis + don + ad + o (distribuição)

Re + kon + at + a (reconhecido/a)

Os prefixos e sufixos são opcionais e podem aparecer em número maior que 2. A raiz é obrigatória; uma palavra pode ser composta por mais de uma raiz. A desinência é accidental, porém obrigatória: é ela que vai determinar, de forma clara, a função sintática da palavra (substantivo *-o*, adjetivo *-a*, advérbio *-e*, verbo no infinitivo *-i*, etc.). Os tempos verbais seguem as desinências: *-as* (presente), *-is* (pretérito), *-os* (futuro), *-us* (condicional). O modo imperativo: *-u*. A ausência de conjugação e de formas verbais irregulares torna o sistema relativamente simples.

⁴ Trata-se aqui de René de Saussure, irmão do renomado semiólogo Ferdinand de Saussure. A pedido de Ferdinand, René dedicou-se ao estudo do esperanto no início do século XX, tendo posteriormente publicado o primeiro texto sobre os mecanismos de formação dos vocábulos dessa língua.

Cada um dos elementos que formam as palavras pode aparecer como uma palavra isolada. Como os morfemas são invariáveis, eles podem ser usados sem restrição na totalidade dos casos. Isso dá ao aprendiz segurança para construir as palavras ou para desconstruí-las, decodificando seu significado.

Segundo Deacon (1997), o relacionamento referencial entre palavras forma um sistema de relacionamentos de ordem mais alta que permite às palavras ter intrinsecamente relacionamentos indiciais, e não somente índices nelas próprias.

Em termos de aprendizado simbólico, tem-se que o processo se inicia por meio da formação de um repertório de lexemas que são assimilados de forma indicial. Mas uma vez aprendidas, essas relações permitem a busca por uma lógica de mais alta ordem, e uma vez que essa é descoberta, os sujeitos se tornam capazes de usá-la para generalizar novas associações. Isso permite uma notável aceleração na eficiência do aprendizado.

Embora essa lógica nunca tenha sido explicitamente aprendida, esse terá sido um tipo de conhecimento implícito, subproduto espontâneo da recodificação simbólica. Assim, associações indiciais formam a base para a referência simbólica, mas elas devem ser substituídas para que esta possa atuar (Deacon, 1997).

Embora a referência indicial seja mantida na transição simbólica, ela não mais é determinada pela ou dependente da correlação física entre sinal e objeto. Isso possibilita um novo tipo de generalização: a generalização de categoria (lógica), em contraponto à generalização de estímulos e de conjuntos de aprendizado (Deacon, 1997).

Formação de palavras em esperanto

Segundo Mattos (1987:2), a formação das palavras em esperanto tem aspectos que a distinguem do processo observado nas línguas flexionais. Para compreender isso, é necessário estudar os diversos níveis da linguagem como em um processo reducionista. Se analisados sob esse ângulo, palavra, locução, oração e período seguem as mesmas regras de construção. Isso em geral não é percebido pelo falante, pois as palavras são aprendidas desde a infância como unidades e suas partes constituintes passam despercebidas. Em razão da rigidez de formação das palavras, a liberdade dos usuários é quase nula nesse nível, mas vai aumentando à medida que passamos para as locuções, para as orações até atingir o máximo de liberdade na formação dos períodos. Vamos então tentar entender em que medida há semelhanças entre as formações nos diversos níveis da linguagem. Assim, tomando exemplos de (Mattos, 1987:2), observamos que:

O período é um conjunto de orações, das quais uma é **essencial** e as outras incidentais:

Chovendo, **o rio transborda**, alagando tudo.

Obs.: a oração essencial é a única que se liga a todas as demais e que pode aparecer isolada.

A oração é um conjunto de locuções, em que uma é **essencial** e as outras incidentais:

O menino **deu ontem um livro ao pai**.

Obs.: a locução essencial é a que se liga a todas as demais.

A locução é um conjunto de palavras, em que uma é **essencial** e as outras incidentais:

Aquelas tuas lindas **revistas encadernadas**.

Obs.: a palavra essencial é a que se liga a todas as demais.

A palavra é um conjunto de morfemas, em que um é **essencial** e os outros incidentais:

Enferm eir inh a s

Obs.: o morfema essencial é o que se liga a todas os demais.

Aqui aparece a diferença fundamental entre a estruturação das palavras em esperanto e a das línguas flexionais: nestas, as palavras permanecem inalteradas quando se juntam para formar os conjuntos superiores (locuções, orações e períodos), ao passo que, em contraste, os morfemas que constituem as palavras assumem formas especiais que só se aplicam quando acompanhados de determinados outros morfemas, o que obriga o usuário ou aprendiz daquela língua a assimilar (decorar) cada palavra do vocabulário, sem ter jamais a liberdade de extrapolar um morfema aprendido para usar na formação de outras palavras. No esperanto, ao contrário, os morfemas nunca têm mais de uma forma para o mesmo significado. Por exemplo, comparemos, no português e no esperanto, a formação de substantivos abstratos derivados de adjetivos:

Tabela A.3.3: Formação dos substantivos abstratos em português e em esperanto

Significado	Forma em português	Forma em esperanto
qualidade de ser alegre	alegr ia	gaj eco
qualidade de ser beato	beat itude	beat eco
qualidade de ser enorme	enorm idade	enorm eco
qualidade de ser amplo	ampl idão	larĝ eco
qualidade de ser fresco	fresc ura	freŝ eco
qualidade de ser verde	verd or	verd eco
qualidade de ser belo	bel eza	bel eco
qualidade de ser tolo	tol ice	stult eco
qualidade de ser azedo	azed ume	acid eco
qualidade de ser selvagem	selvag eria	sovaĝ eco

Para um mesmo significado (neste caso o conceito abstrato relativo a um adjetivo) o esperanto usa sempre e invariavelmente o mesmo monema *-ec* (pronúncia = ets), para todo e qualquer adjetivo, o que estabelece uma relação unívoca entre os elementos constituintes das palavras e o significado de cada um deles. Isso tem profundas repercussões no uso e na assimilação da língua.

No âmbito do alfabeto, a carga simbólica do esperanto é extremamente leve: cada letra deve ter seu valor aprendido, mas o alfabeto usa caracteres latinos (os mais difundidos no mundo). Além disso, sendo o alfabeto totalmente fonético, cada letra age como um índice único para um dado fonema (e vice-versa). Uma vez que essas relações grafema-fonema básicas estejam estabelecidas, elas se aplicam a 100% do vocabulário.

A título de comparação, para quantificar a carga simbólica associada ao aprendizado do sistema de escrita da língua inglesa, é preciso considerar que as 26 letras do alfabeto inglês podem ser combinadas de 1120 maneiras diferentes para representar os 40 fonemas da língua. Isso significa que é extremamente difícil estabelecer um conjunto de referências indiciais: a relação grafema-fonema é sempre sujeita a exceções, tornando necessário em

geral conhecer cada palavra para saber qual a pronúncia correta, o que significa que a assimilação do sistema de escrita deve ser concomitante com a assimilação do corpo léxico. Essa é uma carga tão pesada que uma pesquisa (Paulesu, 2001) demonstrou uma forte correlação entre o sistema de escrita e a incidência de distúrbios da linguagem: nos países anglo-saxões as taxas de dislexia são muito altas em comparação com aquelas observadas na Itália, isso porque a língua italiana possui somente 33 maneiras de grafar seus 25 fonemas.

No caso específico aqui tratado, ensino do esperanto por meio da Web, o uso de um referencial semiótico é promissor. Segundo os três modos de inferência definidos por Peirce: abdução, dedução e indução, a *semiosis* pode ser entendida como um processo de se aplicar signos para entender algum fenômeno (indução), raciocinar de signo para signo (dedução), e/ou usar signos para fazer sentido de alguma nova experiência (abdução) (Cunningham, 2001). No ensino do esperanto, as situações em que a abdução se manifesta são muito frequentes por conta da regularidade dos esquemas: ao ser confrontado com um texto, o estudante, por exemplo, se dá conta de que a cada vez que encontra uma palavra terminada em *-o*, ela representa um substantivo, e todas aquelas terminadas em *-a* representam adjetivos. Por força dessas repetidas experiências ele abduz a idéia (verdadeira) de que os substantivos se formam com a terminação *-o* e os adjetivos com *-a*, e essas estruturas mentais se tornam firmes, pois elas se baseiam no nível mais fundamental e espontâneo da verbalização.

O mecanismo da indução, por sua vez, se manifesta quando com base em ideias o estudante induz experiências válidas. Isso se dá, por exemplo, quando ele aprende que os substantivos coletivos são sempre formados com auxílio do sufixo *-ar*, e assim ele extrapola a ideia para todos os casos, sempre chegando a um resultado correto: se *stelaro* é o coletivo de *stelo* (estrela), e *bovaro* é o coletivo de *bovo* (boi), ele induz que o coletivo de *abelo* (abelha) é *abelaro*.

Finalmente, o processo de dedução ocorre quando a partir de uma ideia o estudante é levado a outra ideia: esse é o caso quando o estudante aprende que *kio*, *nenio* e *tio* significam o que, nada e isso, e que *kie*, *nenie* e *tie* significam onde, em nenhum lugar e lá, ao se deparar com o correlativo *kiam* significando quando, ele deduzirá que nunca e então são, respectivamente, *neniam* e *tiam*.

Da discussão que precede é possível concluir que enquanto o aprendizado de uma língua estruturada de forma irregular pressupõe inserções na memória de número elevado de esquemas cuja assimilação raramente autoriza extrapolações, o aprendizado do esperanto permite uma ampla extrapolação dos esquemas básicos da língua. Por essa razão, o desenvolvimento de um ambiente computacional voltado ao ensino do esperanto (de seu sistema simbólico) pode tirar proveito desse aspecto.

Referências bibliográficas

AITCHINSON, J. *Words in the Mind. An introduction to the mental lexicon* (2^a Ed.), Blackwell, Oxford, 1994.

ARNHEIM *Art and Visual Perception*, University of California Press, Berkeley, CA, 1954.

ASHER, J.J. *Learning Another Language Through Actions: The Complete Teacher's Guidebook*. Sky Oaks Productions, Los Gatos, Califórnia, 1977.

ASHER, J.J. *Brainswitchin: A Skill for the 21st Century*. Sky Oaks Productions, Los Gatos, Califórnia, 1988.

ÁVILA, I., GUDWIN, R. *Geração e sequenciamento automáticos de lexemas e sentenças em Sistema Tutor Inteligente para ensino de línguas*. In: Anais do Simpósio Brasileiro de Informática na Educação – SBIE 2010, João Pessoa, Brasil, 2010a.

ÁVILA, I., GUDWIN, R. *Image semantization with MPEG-7 descriptors in a Web language course*. Poster apresentado na Summer School on Multimedia Semantics – SSMS 2010, Amsterdã, Holanda, 2010b.

ÁVILA, I., GUDWIN, R. Lexical similarity metrics for vocabulary learning modeling in Computer-Assisted Language Learning (CALL). In: *Proc. of the AIED Workshop on Natural Language Processing in support of Learning*, v.10. pp. 9-17, Brighton, UK, 2009.

BALBIN, J. Is Esperanto an Artificial Language? In *Geolinguistics* 8, pp. 11-27, 1982.

BALBIN, J. Is Esperanto a Romance Language? In *Geolinguistic Perspectives*, Jesse Levitt, L.R.N. Ashley e Kenneth, H.R. (eds.), pp. 85-103, Lanham, MD: University Press of America, 1987.

BARDIN, L. Le texte et l'image. *Communication et langages*. v.26, N.26, pp. 98-112, 1975.

BAARS, B.J. *In the theater of consciousness: The Workspace of the Mind*. Oxford University Press, 1^a edição, 1997.

BARTHES, R. Le message photographique. *Communications* 1: 127-138, 1961.

BEN-ZE'EV, A. Is there a problem explaining cognitive progress? In R. F. Goodman & W. R. Fisher (Eds.), *Rethinking Knowledge: Reflections Across the Disciplines* (pp. 41-56). Albany: State University of New York Press, 1995.

BERLIN, B., KAY, P. *Basic color terms: their universality and evolution*. Berkeley: University of California Press, 1969.

BILLSUS, D., PAZZANI, M. A hybrid user model for news stories classification. In *Proc. of the 7th International Conf. on User Modeling*. Banff, Canada, pp. 99-108, 1999.

BLOOM, P. *How children learn the meanings of words*. The MIT Press: Cambridge, 2000.

BLOOMFIELD, L. *Outline Guide for the Practical Study of Foreign Languages*. Linguistic Society of America, Baltimore, 1942.

BRUSILOVSKY, P. Methods and techniques of adaptive hypermedia. *User Modeling and User Adapted Interaction*, 6, 2-3, pp. 87-129, 1996.

CALVERT, S.L., HUSTON, A.C., WATKINS, B.A., & WRIGHT, J.C. The relation between selective attention to television forms and children's comprehension of content. *Child Development*, 53, 601-610, 1982.

CAMERON, K. *CALL: Media, Design and Applications*. Taylor & Francis, 1999. ISBN 902651543X.

CAPRA, F. *The Web of Life: A New Scientific Understanding of Living Systems*. New York: Anchor/Doubleday, 1996.

CARROLL, J., CASAGRANDE, J. *The function of language classification in behavior*. Readings in Social Psychology. Holt, Rinehart & Winston, Nova Iorque, 1958.

CHOMSKY, N. Language and nature. *Mind*, 104, 1-61, 1995.

CHOMSKY, N. *Syntactic Structures*. Mouton, Haia, 1957.

CHOMSKY, N. *Aspects of the Theory of Syntax*. MIT Press, Cambridge, 1965.

CORE, M.G., MOORE, J.D., ZINN, C. Supporting Constructive Learning with a Feedback Planner. In *Building Dialogue Systems for Tutorial Applications, Papers of the 2000 AAAI Fall Symposium*, C.P. Rose e R. Freedman (Ed.), 1-9, 2000.

COSTA, E. B. *Um modelo de Ambiente Interativo de Aprendizagem Baseado numa Arquitetura Multi-Agentes*. Tese de Doutorado. Univ. Fed. de Campina Grande, PB, 1997.

CRISTEA, A., DE BRA, P. ODL Education Environment Based on Adaptivity and Adaptability. *Proc. of the AACE ELearn`2002 conference*, pp. 232-239, 2002.

CUNNINGHAM, D. J. Cognition as Semiosis: the role of inference. Indiana University, 2001. In: www.indiana.edu/~educp550/theory.html, acessado em abril de 2005.

CURRAN, C.A. *Counseling-Learning in Second Languages*. Apple River Press, Apple River, 1976.

DANESI, M. *Semiotics and Education*. ISAS'97, 1997.

DANESI, M. *Semiotics in Language Education*. Mouton de Gruyter, 2000. ISBN 3-11-016915-0.

DANESI, M. *Second Language Teaching: a view from the right side of the brain*. Topics in Language and Linguistics. Kluwer Academic Publishers, 2003.

DEACON, T. W. *Symbolic Species – the coevolution of language and the brain*, W. W. Norton, 1997.

DE TIENNE, A. *Learning qua semiosis*. II Workshop on Computer Intelligence and Semiotics. Orgs. Queiroz, J. e Gudwin, R., Brasil, 2002.

DI PIETRO, R.J. *Strategic Interaction*. Cambridge University Press, Cambridge, 1987.

DUBOIS, D. Constructing an Agent Equipped with an Artificial Consciousness: Application to an ITS. PhD Thesis at the University of Quebec, Montreal, 2007.

ECO, U. *A busca da língua perfeita*. Editora: Edusc, 2001. ISBN: 85-7460-109-8.

ELLIS, N. The epigenesis of language: Acquisition as a sequence learning phenomenon. In A. Ryan & A. Wray (Eds.), *Evolving Models of Language* (pp. 41-57). Clevedon: British Association for Applied Linguistics/Multilingual Matters, 1997.

EMBENSON, S. *Diagnostic testing by measuring learning processes psychometric considerations for dynamic testing*. Frederiksen N. (Ed.) Diagnostic Monitoring of Skill and Knowledge Acquisition. Lawrence Erlbaum Associates, Hillsdale, N.J., 1990.

FETTES, M. *Critical realism and ecological psychology: foundations for a naturalist theory of language acquisition*. Ecology of Language Acquisition Workshop, University of Amsterdam, 11-15 de janeiro de 1999.

FOX, B. *The human tutorial dialogue project*. Lawrence Erlbaum Associates, Inc. Hillsdale, N.J. 1993.

FREIRE, P. *Pedagogy of the Oppressed*. Harmondsworth: Penguin, 1972.

GASSER, F.J., HOFFMANN, D. A Multi-Agent Architecture for Operation and Maintenance of Telecommunications Networks. In: *International conference on AI, expert systems and natural language*, 12., Avignon, France, 1992. p. 427-436, 1992.

GATTEGNO, C. *The Common Sense of Teaching Foreign Languages*. Educational Solutions, Nova Iorque, 1976.

GOLDEN, B. When did the first conversation in Esperanto take place? *Geolinguistics* 16, pp. 85-106, 1990.

GOLDEN, B. The Agglutinative Character of Esperanto. *Geolinguistics* 18, pp.89-128, 1992.

GONÇALVES, J. P. *A integração de testes adaptativos informatizados e ambientes computacionais de tarefas para o aprendizado do inglês instrumental*. Monografia de Mestrado, USP, São Carlos, 2003.

GOODMAN *Languages of art*. Bobbs-Merrill, Nova Iorque, 1968.

GRAESSER, A.C., PERSON, N.K., MAGLIANO, J.P., Collaborative Dialogue Patterns in Naturalistic One-to-One Tutoring. In *Applied Cognitive Psychology*, v.9, 495-522, 1995.

HALLIDAY, M.A., McINTOSH, A., STEVENS, P. *The Linguistic Sciences and Language Teaching*. Longman, Londres, 1964.

HOCHBERG, J., BROOKS, V. Pictorial Recognition as an Unlearned Ability: a Study of one Child's Performance. *The American Journal of Psychology*, University of Illinois Press, 1962.

HOUSER, N. Toward a Peircean Semiotic Theory of Learning. *The American Journal of Semiotics* 5, n.2, 251-274, 1987.

HULSWIT, M. A Semeiotic Account of Causation. The 'Cement of the Universe' from a Peircean Perspective. Ph.D dissertation, Katholieke Universiteit Nijmegen, 1998.

HUNT, E., AGNOLI, F. The Worfian Hypothesis: A cognitive psychology perspective. *Psychological Review*, 98 (3), pp. 377-389, 1991.

HYMES, D. *On Communicative Competence*. University of Pennsylvania Press, Filadélfia, 1971.

JANNEY, R.W., ARNDT, H. Universality and relativity in cross-cultural politeness research: A historical perspective. *Multilingua*, 12-1 pp.13-50, de Gruyter, Berlim, 1993.

JORDAN, D.K. Esperanto and Esperantism: Symbols and Motivations in a Movement for Linguistic Equality. *Language Problems and Language Planning* 11 (1), pp.104-125, 1987.

KIYMAZARSLAN, V. A promising approach to second language acquisition. *Science Journal of Army Academy*, 2000.

KLÜEGL, F., OECHLEIN, C.; PUPPE, F. Developing Multi-Agent Simulations Based on Explicit Model Representation. Dept. Of Artificial Intelligence and Applied Computer Science, University of Würzburg, Alemanha, 2001.

KRASHEN, S.D., TERREL, T. *The Natural Approach: Language Acquisition in the Classroom*. Pergamon, Oxford, 1983.

LARGE, A. *The Artificial Language Movement*. Basil Blackwell, Oxford, 1985.

LEVY, M. *Computer-Assisted Language Learning: Context and Conceptualization*. Oxford University Press, ISBN 019823631X, 1997.

LOZANOV, G. *Suggestology and Outline of Suggestopedy*. Gordon and Breach, Nova Iorque, 1979.

MAGOULAS, G.D., PAPANIKOLAOU, K.A., GRIGORIADOU, M. Neuro-fuzzy Synergism for Planning the Content in a Web-based Course. *Informatica* 25, pp.39-48, 2001.

MATTOS, G. *La Deveno de Esperanto*. Fonto, Chapecó, 1987.

MATTOS, G. *Gramática completa do Esperanto*. Fonto, Chapecó, 1987.

MATURANA, H.R., VARELA, F.J. *The Tree of Knowledge: The Biological Roots of Human Understanding* (Robert Paolucci, Trans.). (Revised ed.). Boston and London: Shambhala, 1992.

MERRILL, D.C., REISER, B.J., RANNEY, M., TRAFTON, J.G. Effective tutoring techniques: a comparison of human tutors and intelligent tutoring systems. In *The Journal of the Learning Sciences*, 2, pp. 277-305, 1992.

METZGER *Die Gesetze des Sehens*. Frankfurt a. M., Waldemar Kramer, Alemanha, 1975.

MILLER, G.A. *The Science of Words*. Freeman, Nova Iorque, 1996.

MILLIKAN, R. G. *Language, Thought, and Other Biological Categories: New Foundations for Realism*. Cambridge, MA: MIT Press, 1984.

- MUCKENHAUPT, M. *Text und Bild*, Tübingen: Narr, 1986.
- MURPHY, G.L. *Meaning and Concepts*. In P.J. Schwanenflugel (Ed.), *The Psychology of Word Meanings*. Hillsdale, NJ: Erlbaum, 1991.
- NEISSER, U. *Cognitive Psychology*, Appleton Century Crofts, Nova Iorque, 1967.
- NKAMBOU, R. *Managing inference process in student modeling for intelligent tutoring systems*. Proc. of the 11th International Conf. on Tools with Artificial Intelligence, Chicago, Estados Unidos, pp. 19-23, 1999.
- NÖTH, W. *Panorama da Semiótica: de Platão a Peirce*. São Paulo: Annablume, 1995.
- NÖTH, W. *Handbook of Semiotics*. Bloomington e Indianapolis: Indiana University Press. Tradução de: *Handbuch der Semiotik*, 1985/1995.
- NUESSEL, F. *The Esperanto Language*. Nova Iorque, Ottawa e Toronto: Legas, 2000:1.
- NUESSEL, F. Second-language instruction as a semiotic activity. Review article. *Semiotica* 144-1/4 (2003), pp. 389-403, 2000:2.
- NUESSEL, F. *The Semiotics of Esperanto*. University of Lousville, Kentucky, 1995.
- PASSINI, J. *Bilinguismo – utopia ou antibabel*; Editora Pontes, 1993.
- PAULESU, E. Dyslexia: Interaction of Genes with Culture. *Science Magazine*, v.291, n.5511, 16 de março, 2001.
- PIAGET, J., INHELDER, B. *The Psychology of the Child*. Basic Books, New York, 1971.
- PINKER, S. *The Language Instinct*. HarperCollins, Nova Iorque, 1994.
- PIRON, C. *O Desafio das Línguas*. Ed. Pontes, Campinas, 2001.
- PIRON, C. *Esperanto: European or Asiatic Language?* Esperanto Documents 22A, Associação Universal de Esperanto, Roterdã, Holanda, 1981.
- PLUNKETT, K., SINHA, C., MOLLER, M.F., STRANDBY, O. Symbol grounding or the emergence of symbols? Vocabulary growth in children and a connectionist net. *Connection Science*, 4, 293-312, 1992.
- PRINCE, S. The Discourse of Pictures: Iconicity and Film Studies. *Film Quarterly* v. 47, n.1, pp. 16-28, University of California Press, 1993, <http://www.jstor.org/stable/1213106>.

REED, E. S. *Encountering the World: Toward an Ecological Psychology*. Oxford: Oxford University Press, 1996.

REIMUND, W. *Ikonizität und emotionale Bedeutung bildlicher Darstellung*. Lang, Frankfurt/Main, 1993.

ROSENTHAL, S. *Categories, pragmatism, and experimental method*, In Digital Encyclopedia of C.S. Peirce, www.digitalpeirce.org (Ed.) J. Queiroz, 2001.

RÖSLER, A. *Charles S. Peirce on the emergence of the code of perception*. Disponível em: <http://user.uni-frankfurt.de/~wirth/texte/roeslere.html> Acessado em julho de 2009.

SANTAELLA, L., NÖTH, W. *Imagem: cognição, semiótica e mídia*. 4ª Edição, Iluminuras, São Paulo, 2005.

SANTAELLA, L. *Semiótica Aplicada*. Pioneira Thomson Learning, São Paulo, 2002.

SANTAELLA, L. *A teoria geral dos signos: Como as linguagens significam as coisas*. 2ª ed. São Paulo: Pioneira, 2000.

SANTAELLA, L. *The muscular, the sensorial, and the cerebral machines*. Universidade Católica de São Paulo, 1998. In: www.pucsp.br/~lbraga/epap_cult4.htm Acessado em maio de 2009.

SANTAELLA, L. *O computador como um meio semiótico*. Cultura das mídias, 4ª. ed. São Paulo: Experimento, 2004.

SCHUBERT, K. An Unplanned Development in Planned Languages: A Study of Word Grammar. In *Interlinguistics: Aspects of the Science of Planned Languages*, Klaus Schubert (ed.), pp. 249-279, Mouton de Gruyter, Berlin, 1989.

SEGALL, M.H., CAMPBELL, D.T., HERSKOVITS, J.M. *Influence of culture on visual perception*. Bobs-Merill, 1966.

SILVA, R.C.M. *Análise da Arquitetura Baars-Franklin de Consciência Artificial Aplicada a uma Criatura Virtual*. Tese de Mestrado, FEEC, Unicamp, Campinas, 2009.

SISON, R., SHIMURA, M. Student Modeling and Machine Learning. *International Journal of Artificial Intelligence in Education*, 9, pp. 128-158, 1998.

SULLIVAN, T. *Intelligent User Interfaces*, ACM Press, Nova Iorque, 1991.

SZANSER, I. *Teorio kaj praktiko de lingvo-instruado*. ILEI, Budapeste, 1992.

SAUSSURE, R. *La logika bazo de vortfarado en Esperanto*, 1910.

SCHAR, S. G., KRUEGER, H. Using new learning technologies with multimedia. *IEEE Multimedia* 7(3), 40-58, 2000.

SELF, J.A. *Formal Approaches to Student Modelling*. In Greer, J. and McCalla, G. (Eds.), In: *Student Modelling: The Key to Individualized Knowledge-Based Instruction*. pp. 295-352. Springer-Verlag, 1994.

SMITH, D. E. *The Conceptual Practices of Power: A Feminist Sociology of Knowledge*. Toronto: University of Toronto Press, 1990.

SOWA, J. F. *Crystallizing Theories out of Knowledge Soup*. *Intelligent Systems: State of the Art and Future Directions*, ed. Zbigniew W. Ras and Maria Zemankova, Ellis Horwood, New York, pp. 456-487, 1990.

SZERDAHELYI, I. *Principoj de Esperanta Etimologio*. Studoj pri la Internacia Lingvo. Ed. Duc Goninaz. AIMAV Gent, pp. 109-138, 1987.

THOMPSON, J. B. *The media and modernity: a social theory of the media*. Stanford: Stanford University Press, 1995.

TITONE, R. *Teaching Foreign Languages: An Historical Sketch*. Georgetown University Press, Washington, D.C., 1968.

VASILAKOS, T., DEVEDZIC, V., KINSHUK, W.P. Computational Intelligence in Web-Based Education: a Tutorial. In *Journal of Interactive Learning Research*, v.15, 2004.

VIGOTSKY, L. S. *Thought and language*. Cambridge, Mass.: MIT Press, 1962.

VIGOTSKY, L. S. *Mind in Society*. Cambridge, Mass.: Harvard University Press, p.51, 1978.

WEIDENMANN, B. *When good pictures fail. An information processing approach to the effect of illustrations*. In H. Mandl, & J. R. Levin (Eds.), *Knowledge acquisition from text and pictures*. Amsterdam: Elsevier, 1988.