

TRADUÇÃO INTERSEMIÓTICA DOS MECANISMOS DE BUSCA NO CIBERESPAÇO

TRADUCCIÓN INTERSEMIÓTICA DE LOS MOTORES DE BÚSQUEDA EN EL CIBERESPACIO

Silvana Drumond Monteiro - silvanadrumond@gmail.com
Doutora em Comunicação e Semiótica pela Pontifícia Universidade Católica de São Paulo (PUCSP). Docente do Programa de Pós-Graduação em Ciência da Informação da Universidade Estadual de Londrina (UEL).

Rogério Paulo Muller Fernandes - rogerio_muller@uol.com.br
Doutorando em Ciência da Computação pela Universidad de Alcalá (Alcalá de Henares, Madri, Espanha). Docente do Departamento de Ciência da Informação da Universidade Estadual de Londrina (UEL)

Maria Júlia Carneiro Giraldes - giraldes@uel.br
Mestre em Ciência da Informação pela Pontifícia Universidade Católica de Campinas. Docente do Departamento de Ciência da Informação da Universidade Estadual de Londrina (UEL)

Benjamim Luiz Franklin - belfra@gmail.com
Doutor em Gestão e Engenharia do Conhecimento pela Universidade Federal de Santa Catarina (UFSC). Docente do Departamento de Ciência da Informação da Universidade Estadual de Londrina (UEL)

Ilza Almeida Andrade - ilza.aandrade@gmail.com
Mestre em Gestão da Informação pela Universidade Estadual de Londrina (UEL). Bibliotecária da Universidade Estadual de Londrina (UEL).

Richele Grenghe Vignoli - richelevignoli@hotmail.com
Mestranda do Programa de Pós-Graduação em Ciência da Informação pela Universidade Estadual de Londrina (UEL). Docente do Departamento de Ciência da Informação da Universidade Estadual de Londrina (UEL)

Marcos Vinicius Fidencio - marcos.v.fidencio@gmail.com
Graduando em Biblioteconomia pela Universidade Estadual de Londrina (UEL). Bolsista de Iniciação Científica PIBIC.

RESUMO

Introdução: O processo de busca faz parte da sociedade atual e os mecanismos de busca são os responsáveis pela indexação e elaboração dos índices contemporâneos no ciberespaço.

Objetivo: Neste trabalho, a tradução intersemiótica verbo-visual do Google Images e oSKope, considerados híbridos por misturar linguagens, foi analisada visando avaliar o nível de equivalência dos signos nos resultados.

Metodologia: A partir da categorização dos mecanismos, o paradigma semiótico vem sendo perscrutado, visando entender as matrizes da linguagem-pensamento, de Santaella, no funcionamento dos buscadores. Entende-se a intersemiose como a mistura das linguagens, sendo a hibridização sua característica principal. O conceito de tradução, apropriado de Jakobson, engloba a transmutação do signo dentro de uma língua (intralingual), de uma língua para outra (interlingual), ou a tradução propriamente dita, e entre uma linguagem e outra (intersemiótica). Outro conceito incorporado na pesquisa foi a tradução intrasemiótica, dentro da matriz visual, para colaborar com a análise da linguagem visual, que tem a forma como eixo de inteligibilidade e a figura como representação do visível.

Resultados: O comportamento intersemiótico dos dois buscadores demonstrou-se, de maneira geral, satisfatório.

Conclusões: O aprendizado da linguagem pela máquina faz diferença entre os níveis de tradução, entre o igual e o parecido, sobre o princípio de similaridade da forma, seja de uma língua, seja de uma imagem.

Palavras-chave: Mecanismos de Busca. Semiótica. Tradução Intersemiótica.

1 INTRODUÇÃO

Viver é falar. Mas se assim é, resulta igualmente que viver é traduzir. (MOTTA, 2011, p. 49).

Os mecanismos de busca fazem parte da nova realidade pós-moderna e têm mudado o modo de circulação e acesso à informação e ao conhecimento na atualidade. Dessa forma, desde o seu início (final da década de 90), esses índices contemporâneos, operados por máquinas (robôs), têm se constituído objeto de nossa pesquisa, visando compreender o funcionamento da linguagem e da indexação no ciberespaço.

Esse objeto tem sido explorado por jornalistas especializados em tecnologia, além de constar em artigos científicos da área de Computação, porém é perceptível que não há um *corpus* estabilizado que sirva de lastro ao pesquisador. Aliás, esse objeto está longe de uma estabilização em seu *status quo*. A plasticidade do ambiente digital e a conseqüente matematização dos algoritmos colocam o signo em movimento ou, utilizando a expressão de Otávio Paz (1972), põem o signo em rotação.

Friedman (2007, p. 212), a propósito de um livro jornalístico, considera os mecanismos de busca, em especial o Google, uma das forças niveladoras do mundo. Em que medida isso ocorreria? “In-formar consiste na possibilidade de constituir e estruturar a sua cadeia de fornecimento pessoal, de informação, conhecimento e entretenimento.”

Os mecanismos estão cada vez mais inteligentes e o Google tem *corpora* constituídos de várias matrizes semióticas: fotos, vídeos (Youtube), acadêmico (*Scholars*), verbal, Geoespacial (*Maps*), Tradutor, entre outros.

Mas há outros mecanismos especializados em matrizes semióticas, como o Musipedia¹ (música) e o Retrievr² (imagem), assim o texto deixou de ser a única ou prevalente fonte de pesquisa no ciberespaço.

Em artigo anterior, abordamos essa nova realidade semiótica, que estuda as tecnologias e prospecta alguns mecanismos especializados em busca sonora, visual e verbal, o que categorizamos em Paradigma Semiótico dos mecanismos de busca (ABREU; MONTEIRO, 2010).

Após o estudo, percebemos que, além dessas realidades especializadas, o Google estava misturando essas matrizes em seu buscador geral, tornando-se um exemplo de Mecanismo Híbrido, fato que despertou o interesse em investigar a tradução intersemiótica de buscas verbo-visuais em um *corpus* específico de mecanismos híbridos.

Já a tradução, de acordo com Plaza (2010) e Jakobson (2007), é considerada a transmutação de signo em signo (pode ser a intralingual, mas também o próprio pensamento é um exemplo), de língua em língua (interlingual) e de linguagem em linguagem (intersemiótica). A tradução, analisada brevemente, neste trabalho, é a intersemiótica, especialmente a de linguagem em linguagem, embora os buscadores já sejam capazes de realizar todas elas, inclusive a intrassemiótica, conforme analisaremos.

¹ <http://www.musipedia.org>

² <http://labs.systemone.at/retrievr>

2 MECANISMOS DE BUSCA

Os mecanismos de busca são uma realidade concreta em que não há, para o senso comum, necessidade de definição. “Nunca antes, na história do mundo, tanta gente - *por conta própria* - teve a possibilidade de encontrar tantas informações sobre tantas coisas e sobre tantas outras pessoas.” (FRIEDMAN, 2007, p. 211, grifos do autor).

Só o Google, em 2012, fechou o ano com 1.2 trilhão de buscas³, em 146 línguas, e os números demonstram que a democratização do acesso à informação e ao conhecimento está impactando a sociedade contemporânea.

Mas, ao observar o fenômeno, percebe-se o quão complexa é essa nova realidade, seja do ponto de vista de Inteligência Artificial, seja do ponto de vista filosófico ou social.

Algumas definições são, para a pesquisa, importantes, pois tentam desvelar, ao menos linguisticamente, alguns aspectos significativos dos buscadores.

Batelle (2005, p. 241), em seu livro “A busca”, sugere que o futuro dos mecanismos de busca, portanto da Inteligência Artificial no ciberespaço, terá mais a ver com compreender do que simplesmente descobrir. Ademais, afirma que a busca “É nosso modo de navegar pelo único recurso infinito que move a cultura humana: o conhecimento.”

A busca (*searching*), para Nielsen/Netratings, é um fenômeno social e os mecanismos de busca são o ponto focal da experiência *online* (FRAGOSO, 2007). Com efeito, Soares (2008) utiliza a metáfora dos oráculos pós-modernos para os mecanismos.

Levy (2012, p. 260), em seu livro, define o Google como “[...] uma máquina; um colecionador onívoro de informações; um cofre hiperenciclopédico do conhecimento humano; um leiloeiro infalível; um estudante de línguas, comportamentos e desejos assustadoramente hábil.” Ou, ainda, o Google é “[...] uma gigantesca máquina de aprendizagem baseada em inteligência artificial.” (p. 15).

Pensando em estudar os mecanismos, e com isso compreender um pouco dessa nova realidade, uma categorização foi desenvolvida e está ainda em curso de

³ Medição realizada pela ferramenta “Zeitgeist”, que significa “espírito do tempo”, em alemão, e faz o monitoramento das buscas realizadas no Google (www.google.com/zeitgeist).

investigação, posto que alguns itens encontram-se inconclusos ou mostraram-se mais abrangentes, ou, ainda, estão em curso de criação de novas tecnologias, conforme o Diagrama 1.

A categorização, tomada em sentido filosófico, tem sido utilizada recorrentemente em nossas pesquisas como método de perscrutar e prospectar alguns objetos científicos. Para Plaza (2010, p. 46), “Perceber já é selecionar e categorizar o real, extrair informações que interessam num momento determinado para algum propósito.”

Considerando também que um esquema categórico geral baseia-se ou abrange uma ontologia, a categorização contida no Diagrama 1 procura a compreensão da base ontológica dos mecanismos de busca, ou seja, a sua natureza, ou, ao menos, intenta fazê-lo.

Diagrama 1 - Categorização dos Mecanismos de Busca

EIXO	CLASSIFICAÇÃO	EXEMPLOS	
ANATOMIA	Crawling (varrer)	Taxonomia da busca: Informacional, Navegacional, Transacional (BRODER, 2002) Anatomia da busca (MORVILLE; CALLENDER, 2010) Escaneamento, Browsing, Busca, Exploração, Passeio (CANTER; RIVERS; STORRS, 1985) Básica, Avançada (CUNHA; CAVACANTI, 2008)	
	Indexing (indexar)		
	Searching (buscar)		
FORMA GERAL DE ORGANIZAÇÃO OU INDEXAÇÃO (crawling/indexing)	Diretórios ou Catálogos	Yahoo!, Open Directory (dmoz)	
	Programas ou robôs de Busca	Google	
	Híbridos	Yahoo!	
	Metabuscaadores	Verbais	Metacrawler, Dogpile
		Visuais	Yometa
Federados		Science Search, MetaLib	
ORDENAÇÃO DOS RESULTADOS (indexing/searching)	Localização dos termos	Google e outros	
	Frequência dos termos	Google e outros	
	Análise de links	Google (autoridade e eixo)	
	Relevância	Google	
	Pagos, orgânicos e híbridos	Google	
APRESENTAÇÃO DOS RESULTADOS (searching)	Agrupamento ou clusterização	Verbais	Yippy
		Visuais	Carrot2, HighWire (Stanford), Mooter
	Especializados ou verticais	Technorati, Fizy, Google: Maps, Earth, Acadêmico, Books, Buscapé, e-Bay, Midomi, Buscatemática.net	
	Personalizados	A9	
	Federados	Science Research, MetaLib	
	Web Semântica	Semantização da Web por mec. gerais	Google (Knowledge Graph)
		Semânticos Gerais	Cluuz, Hakia, Lexxe
		Ontologias, RDF, XML	Watson, Swoogle
Web Invisível ou Profunda	Scirus, Pipl, Complete Planet, Hakia, Deep Dyve,		
PARADIGMA SEMIÓTICO (indexing/searching)	Sonoros	Musipedia, Midomi	
	Visuais e Georreferenciais	Retrievr (Systemone), Google Earth, Google Maps	
	Verbais	Google	
	Híbridos	Oskope, Lygo, Kmap	

Fonte: Adaptado de MONTEIRO, et al. (2013, p. 9).

A categoria, do ponto de vista ontológico, significa qualquer noção ou princípio que sirva de regra para a sua investigação (ABBAGNANO, 2003) e busca compreender as relações entre o todo e as partes.

A partir do ponto de vista lógico, a categoria significa os modos de predicar a substância e implica, ainda, estudar as relações hierárquicas, as quais tentamos evitar, já que os objetos no ciberespaço estão configurados rizomaticamente. Ainda no sentido lógico, Hessen (1999, p. 138) considera a categoria uma teoria especial do conhecimento, pois “[...] realizar determinações categoriais não é outra coisa senão descobrir as características e relações mais gerais presentes no objetos.”

Sendo o diagrama um ícone de secundidade que revela as relações inteligíveis entre o signo e seus objetos, iremos discutir e ilustrar, a partir do Diagrama 1, o Paradigma Semiótico.

E mais, para Deleuze, o diagrama é uma determinação do conjunto de relações de forças que jamais se esgotam, pode entrar em outras relações e em outras composições “[...] é um conjunto de linhas diversas funcionando ao mesmo tempo [...]. Por isso cada coisa tem sua geografia, sua cartografia, seu diagrama.” (DELEUZE, 1992, p. 47).

2.1 Paradigma Semiótico: sintaxe, forma, discurso e a hibridização

Desde o período inicial, os fundamentos dos buscadores já tinham sido estabelecidos: varredura, indexação e resultados (LEVY, 2012), ou seja, *crawling*, *indexing* e *searching*. São conceitos-acontecimentos que mostram o movimento do signo nesse processo de busca a partir de um mecanismo, indicados categoricamente, no Diagrama 1, como “Anatomia”.

Não obstante, é na categoria “Apresentação dos Resultados” que os índices contemporâneos tornam visíveis as múltiplas sintaxes e semióticas.

O “Paradigma Semiótico” é outra categoria que intenta compreender os mecanismos de busca a partir dos algoritmos de indexação e busca no ciberespaço, por meio das matrizes da linguagem-pensamento, conforme a teoria de Santaella (2005).

Para tanto, o aporte teórico empregado é a lógica formal dos signos, ou seja, a Semiótica peirciana, pois tem como função e objeto o estudo das representações,

daí que a representação é um pressuposto da qualidade sígnica, pois os objetos representados funcionam semioticamente (SANTAELLA, 2005).

Para Pierce, a Fenomenologia tem por função responder algumas questões: como se dá a apreensão e compreensão do mundo pelo homem? Como a multiplicidade e diversidade infinitas do universo sensível são convertidas em realidade inteligível? (SANTAELLA, 1999).

Assim, depois de décadas de reflexão, Pierce concluiu que tudo o que aparece à consciência assim o faz por meio da gradação de três elementos: 1) qualidade ou sentimento (primeiridade); 2) reação (secundidade); 3) mediação (terceiridade). Resumidamente, todo fenômeno apresenta três elementos: a qualidade, o fato atual e a abstração.

Da mesma forma, o signo apresenta esses três momentos fenomenológicos, uma vez que ele representa seus objetos iconicamente, indicialmente e simbolicamente para o interpretante.

A partir das três categorias fenomenológicas propostas por Pierce, pode-se estudar ou mesmo classificar os eventos semióticos. Santaella (2005), em sua teoria, defende que a variedade e a multiplicidade de todas as formas de linguagem estão alicerçadas em apenas três, não mais que três, matrizes, a saber: **a sonora (ou virtual), a visual e a verbal.**

Para Santaella (2005), a Matriz Sonora está para a primeiridade e é uma questão do ícone, a Matriz Visual está para a secundidade e é uma questão do índice e a Matriz Verbal está para a terceiridade e é uma questão do símbolo.

Assim, para compor essa complexa malha diagramática das matrizes, foram fixados três eixos classificatórios, ou atributos internos em cada linguagem, de modo que o eixo da sintaxe está para a Matriz Sonora, o eixo da forma está para a Matriz Visual e o eixo do discurso está para a Matriz Verbal (SANTAELLA, 2005). Com efeito, a música traduz a iconicidade (qualidades de sentimento), a imagem, a referencialidade (contiguidade do real) e a palavra, a representatividade (pensamento em signos).

Igualmente, os mecanismos podem ser categorizados segundo o paradigma em que operam seus algoritmos, conforme se observa no Diagrama 1, e analisados de acordo com a Semiótica das três matrizes da linguagem-pensamento.

Considerando como partida o signo sonoro para a busca, pode-se deduzir que o item indexado é uma representação sonora e está no lugar lógico de um recurso sonoro. Essa cadeia lógica de representação está presente em todas as matrizes, sendo elas sonoras, visuais e verbais (ABREU; MONTEIRO, 2010).

A Semiótica no ciberespaço é um assunto complexo, uma vez que as linguagens que se apresentam na tela não são naturais ou puras, mas construções simbólicas advindas de um substrato lógico digital, portanto híbridas de partida. Entretanto, do ponto de vista do interpretante, em especial o intérprete, aquele que se encontra em frente à tela, o que se vê são linguagens, se não naturais, ao menos, aparentemente, analógicas.

Nessa direção,

O caráter tátil-sensorial, inclusivo e abrangente, das formas eletrônicas permite dialogar em ritmo 'intervisual', 'intertextual' e 'intersensorial' com os vários códigos da informação. [...] No movimento constante de superposição de tecnologias sobre tecnologias, temos vários efeitos, sendo um deles a hibridização de meios, códigos e linguagens que se justapõem e combinam, produzindo a Intermídia e a Multimídia. (PLAZA, 2010, p. 13).

Diante da hibridização das linguagens, o trabalho analisa o comportamento intersemiótico de dois buscadores híbridos, verbo-visuais, a saber: Google Images e o oSkope.

Já a noção de tradução, desenvolvida por Jakobson, em 1959⁴, pode ser dividida em interlingual, intralingual e intersemiótica:

- 1) a tradução intralingual ou reformulação (*rewording*) consiste na interpretação dos signos verbais por meio de outros signos da mesma língua;
- 2) a tradução interlingual ou tradução propriamente dita consiste na interpretação dos signos verbais por meio de uma outra língua;
- 3) a tradução intersemiótica ou transmutação consiste na interpretação dos signos verbais por meio de signos não verbais. (JAKOBSON, 2007, p. 64-65).

A citação supracitada permite inferir que esse tipo de tradução se dê com a linguagem verbal como partida ou original, mas o inverso acontece também, e pode

⁴ Obra publicada no Brasil, 1. edição, em 1971 pela Editora Cultrix. Utilizamos a edição de 2007, da mesma editora.

o conceito ser entendido e estendido como a passagem de uma linguagem à outra, em época de hibridização multimídia no ciberespaço, uma vez que:

A Tradução Intersemiótica ou 'transmutação' foi por ele definida como sendo aquele tipo de tradução que consiste na interpretação dos signos verbais por meio de sistemas de 'signos não verbais', ou 'de um sistema de signos para outro, por exemplo, da arte verbal para a música, a dança, o cinema ou a pintura', ou vice-versa, poderíamos acrescentar. (PLAZA, 2010, p. xi).

Assim, o objetivo da análise do *corpus* selecionado foi a tradução intersemiótica, conforme mencionado, mas, no caso do Google, outra noção foi introduzida neste escopo, a tradução "**intrasemiótica**". "Ela seria a tradução de uma linguagem em si própria, numa tentativa de melhor representar a mensagem a ser passada." (GONÇALVES, 2011, p. 7).

2.2.1 Mecanismos de busca híbridos

Santaella (2005) afirma que as modalidades existentes nas três matrizes da linguagem-pensamento objetivam apresentar focos de inteligibilidade à compreensão e análise da configuração e organização das linguagens e não devem ser empregadas como instrumentos estanques, mas como uma visão filigrânica das mesmas.

Na ambiência do ciberespaço, separar as linguagens é uma temeridade, posto que esse espaço, pela sua fluidez, transforma as linguagens em líquidas⁵, denominação de Santaella (2005, p. 320), que muito bem ilustra o fenômeno dos signos desterritorializados:

Longe de ser apenas uma nova técnica, um novo meio para a transmissão de conteúdos preexistentes, a hipermídia é, na realidade, uma nova linguagem que nasce da criação de hipersintaxes capazes de refuncionalizar linguagens (textuais, sonoras, visuais) que antes só muito canhestamente poderiam estar juntas, cambiando-as e retecendo-as em um mesma malha multidimensional.

⁵ Referência a Bauman (Modernidade líquida, 2001) sobre as análises de uma sociedade em constante mudança de forma e de sua conseqüente falta de permanência. Interessante destacar que Santaella (2007) reconhece que as obras de Deleuze e Guattari estão permeadas de conceitos, cujos campos semânticos remetem à liquidez (desterritorialização, nomadismos, corpos sem órgãos).

É nessa rede de tessitura emaranhada que surgem as linguagens híbridas:

[..] as quais, prefiro chamar de 'intersemióticas', pois ao fim e ao cabo, a hibridização não é outra coisa senão a justaposição, associação e interrelação dos mais variados sistemas de signos, verbais, visuais e sonoros, em hipersintaxes espaciais e temporais. [...] Texto, imagem e som já não são o que costumavam ser. Deslizam uns para os outros, sobrepõem-se completam-se, unem-se, separam-se e entrecruzam-se. (SANTAELLA, 2007, p. 391).

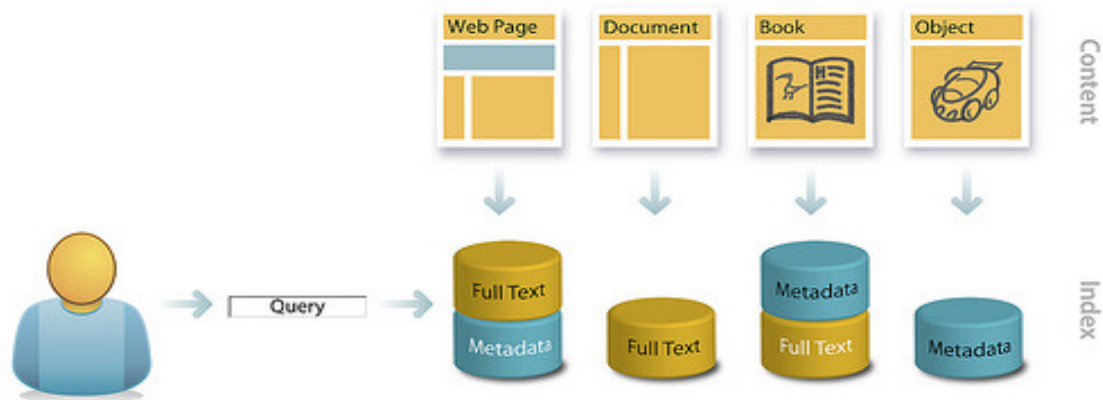
Da mesma forma, nos mecanismos híbridos, ocorrem processos de intersemiose, tal é o caso em que uma linguagem é utilizada para indexar e buscar produtos majoritariamente de outras linguagem, por exemplo, o texto é utilizado (linguagem verbal) para buscar música (linguagem sonora) e/ou fotos (linguagem visual). É comum encontrar mecanismos que oferecem caixas de busca em que se inserem palavras-chave para buscar fotos, músicas, vídeos, portanto mediados pela linguagem verbal escrita e efetuando a tradução ou a intersemiose. Observa-se essa hibridização, das linguagens verbo-visuais, em buscadores no ciberespaço, exemplificados a seguir, que compõem o *corpus* de investigação:

Google Images: serviço de busca por imagem, integrado ao buscador geral;
oSkope: é um assistente de busca visual que permite navegar e organizar os itens de Amazon, eBay, Flickr, Fotolia, Imagem Yahoo! e YouTube, de forma intuitiva.

Se a mediação é um conceito epistemológico que envolve o signo, a hibridização dos mecanismos de busca envolve hipersintaxes de tradução de objetos digitais de uma linguagem para a outra. E é justamente nessa tradução que problemas sintático-semânticos podem acontecer, uma vez que esses objetos, nos mecanismos híbridos, só podem ser buscados se rotulados, e mais, devem ser bem rotulados pela palavra, para efetuar uma mediação ou tradução satisfatória.

Morville e Callender (2010, p. 37) apresentam a indexação de conteúdos, a saber:

Figura 1 - Relative Value of Text and Metadata



Fonte: MORVILLE; CALLENDER, 2007, p. 37. Disponível em:
<<http://www.flickr.com/photos/morville/4273516677/>>.

Comparando a *Web page* com documentos existentes na intranet (*document*), livro disponível em *sites* de venda (*book*) e objeto (*object*) – especialmente este último, que compõe as imagens, só pode ser rastreado, indexado e buscado em mecanismos híbridos se houver algum metadado que possa fazê-lo –, segundo os autores, esse tipo de objeto apresenta o maior desafio para a busca.

Com o estudo da *Web Invisível*, percebeu-se que alguns objetos que eram invisíveis aos mecanismos, anos atrás, como imagens, fotos, vídeos e até mesmo documentos com extensão PDF, acabaram vindo à superfície, ganhando visibilidade, com o desenvolvimento de tecnologias específicas, mas trouxeram junto problemas de tradução intersemiótica. O trabalho aqui proposto permite, de certo modo, analisar o espectro de desenvolvimento dessas tecnologias a partir da simulação de *queries* e de seus respectivos resultados.

3 ANÁLISE DOS MECANISMOS HÍBRIDOS

A hibridização das linguagens não é novidade, tanto que McLuhan (1999, p. 75) já acreditava que o híbrido é o encontro de dois meios, que constitui um momento de verdade e revelação, do qual nasce a forma nova. “O momento de encontro dos meios é um momento de liberdade e liberação do entorpecimento e do transe que eles impõem aos nossos sentidos.”

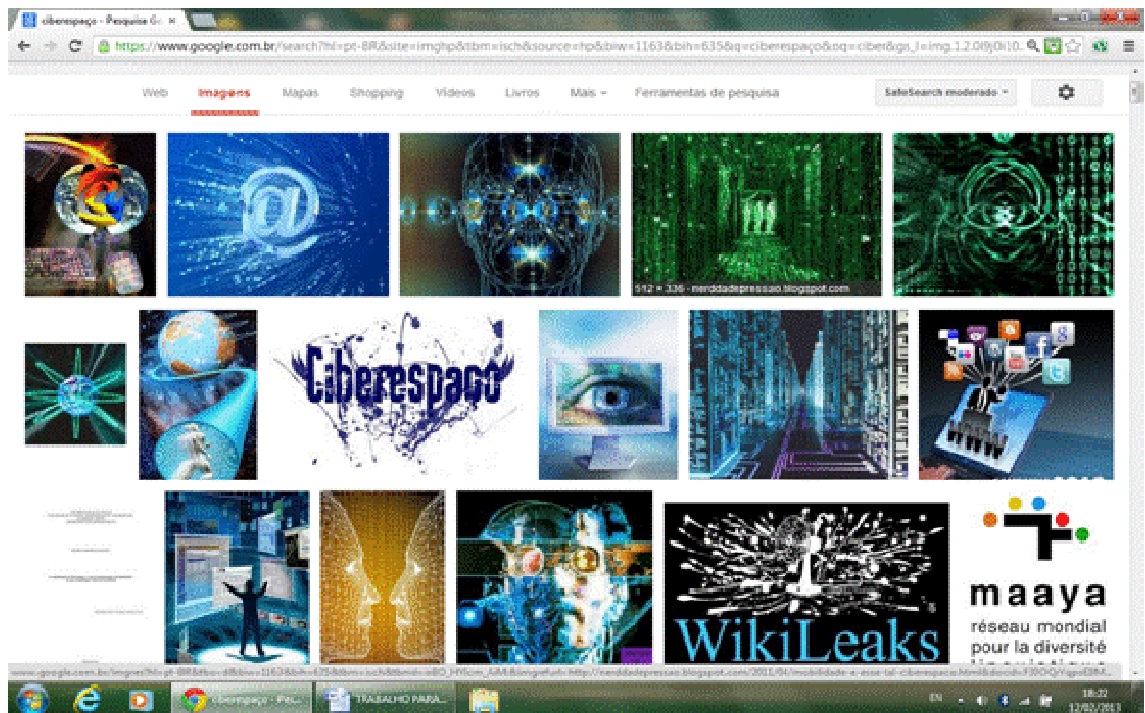
Nesse sentido, Plaza (2010, p. 12) cita alguns expoentes de linguagens híbridas: Mallarmé (no poema em forma de leque), Augusto de Campos (com a

Poesia Concreta) e Kandinski (no poema Klänge, que sugere a sinestesia entre sons, cores e formas). Apesar dessas características (fenômenos de interação semiótica de diversas linguagens), essas obras (e outras híbridas) não se confundem com a tradução intersemiótica “[...] dizem respeito às relações tradutoras mas não se confundem com elas.”

Por outro lado, os buscadores e seus índices contemporâneos realizam, explícita e intencionalmente, a tradução semiótica em seus *corpora* híbridos. Tal operação é de extrema complexidade da Engenharia da Computação voltada à indexação e à Recuperação da Informação.

Vejamos o Google Images, iniciado como experimento denominado Google Similar Images, podia ser categorizado, à época, como matriz eminentemente visual pela importação de imagens para a busca (*query*⁶ imagética). Entretanto, como o Google visava à integração de suas inovações em seu mecanismo, havia um campo para buscar também por palavras, como hoje aparece em seu buscador geral no domínio de língua portuguesa (nosso objeto de análise), conforme a Figura 2:

Figura 2 - Busca Intersemiótica no Google Images



Fonte: GOOGLE. Disponível em: <www.google.com.br>.

⁶ *Query* significa pergunta sem a estrutura de frase, mas entendida ou respondida como texto. (MORVILLE, Peter; CALLENDER, Jeffery. **Search patterns: design for discovery.** Canadá: O'Reilly, 2010).

Apesar de o ciberespaço ser de difícil definição, seu referente existe e está ligado às imagens do mundo virtual, conforme os resultados que puderam ser capturados no *print* da Figura 2.

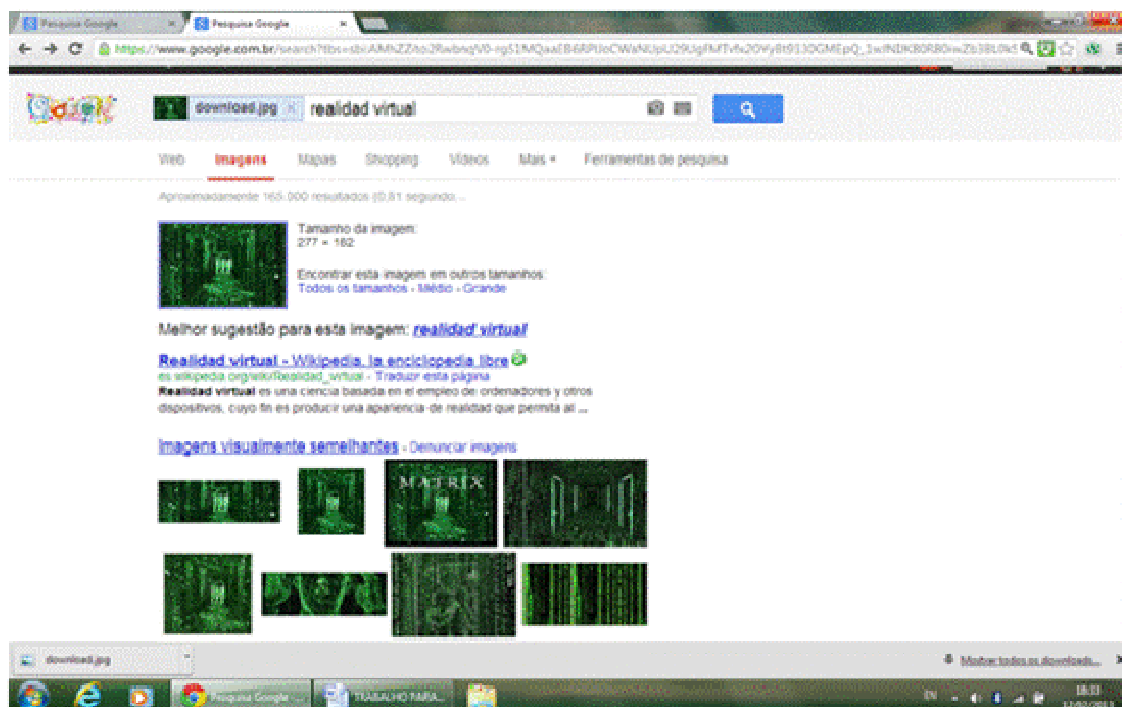
O Google lê os objetos a partir dos metadados, texto e *links*. Os resultados da Figura 2 ilustram, em certa medida, o *continuum* semiótico de autogeração de signos, base da teoria peirciana, só que realizada pela máquina (interpretante maquínico).

O resultado da tradução intersemiótica (Figura 2, entre linguagens) apresenta diversidade de sugestões para a imagem de “ciberespaço”, mas, via de regra, todas ou quase todas muito coerentes com a percepção que fazemos dela. A diversidade pode ser explicada, também, porque a tradução, especialmente interlingual/semiótica, significa signo equivalente e não estritamente igual, ou seja, sugestões atualizadas.

Levy, em entrevista com os engenheiros do Google, cita Singhal (especializado em Recuperação da Informação) que, a partir do contexto linguístico alimentado pelos sujeitos navegadores, desenvolve o trabalho com a linguagem no buscador: trabalha com sinônimos, com ortografia e estuda a teoria de Wittgenstein sobre como as palavras são definidas pelo contexto. “Tudo isso é aprendido. Tínhamos um computador aprendendo algoritmo, e nele construímos nosso classificador de nomes.” (*apud* LEVY, 2012, p.69).

Outra simulação também foi realizada, de forma intrasemiótica, isto é, importou-se uma imagem que foi anteriormente recuperada como “ciberespaço” para verificação dos resultados, conforme ilustrado na Figura 3.

Figura 3 - Busca Intrassemiótica no Google Images por Importação de Imagens



Fonte: GOOGLE. Disponível em: <www.google.com.br>.

Verifica-se, comparando os dois resultados, que a busca realizada pela *query* análoga (imagem), ou pela tradução intrassemiótica, mostra-se mais exata e delimitada.

No segundo caso, a delimitação dos resultados equivale à semelhança na forma, que é a representação (ou apresentação) do visível.

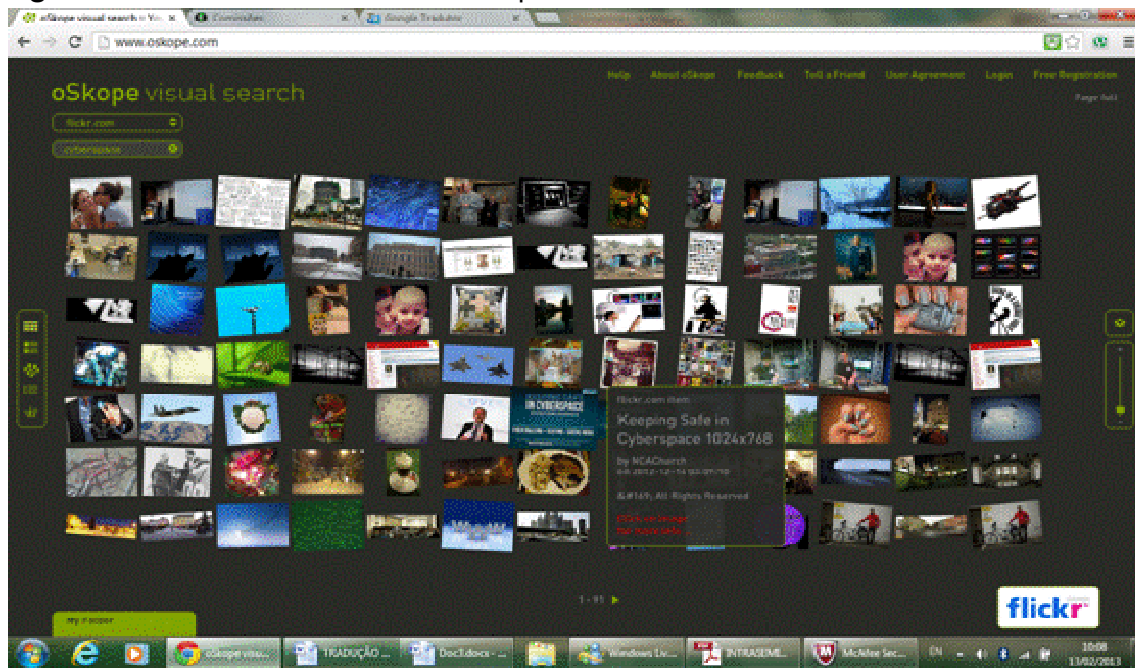
Como o eixo da matriz visual é a forma, a figura é, portanto, a principal classificação desse eixo, pois ela tem contorno, forma e organização. É o mais perto que uma linguagem pode chegar do mundo real. Nas Formas Figurativas (classificação da linguagem visual de SANTAELLA, 2005), a conexão dinâmica se impõe e coloca a figura como registro, seja imitativo, físico ou convencional, e transita entre as formas não-representativas e as representativas.

A foto, em especial, é o registro físico, uma emanção do referente, a mais fiel representação da uma forma. Por isso entendemos que os resultados foram mais restritos no resultado intrassemiótico, pois o índice foi buscar justamente a mesma forma (*query*) que foi indexada em outros endereços e *sites*.

O próximo mecanismo híbrido analisado é o oSkope Visual Search, especializado na busca das bases de dados da Amazon.com, e-Bay, Flickr, Fotolia,

Image Yahoo! e Youtube. Morville e Callender (2010) categorizam o oSkope como mecanismo de descoberta e o interessante dele é que apresenta sugestões (resultados) em vários formatos (em grade ou rede, ilustrado no exemplo infracitado, empilhado, embaralhado, em lista e gráfico).

Figura 4: Busca Intersemiótica no oskope Visual Search na Base do Flickr



Fonte: oSKope. Disponível em: <<http://www.oskope.com/>>.

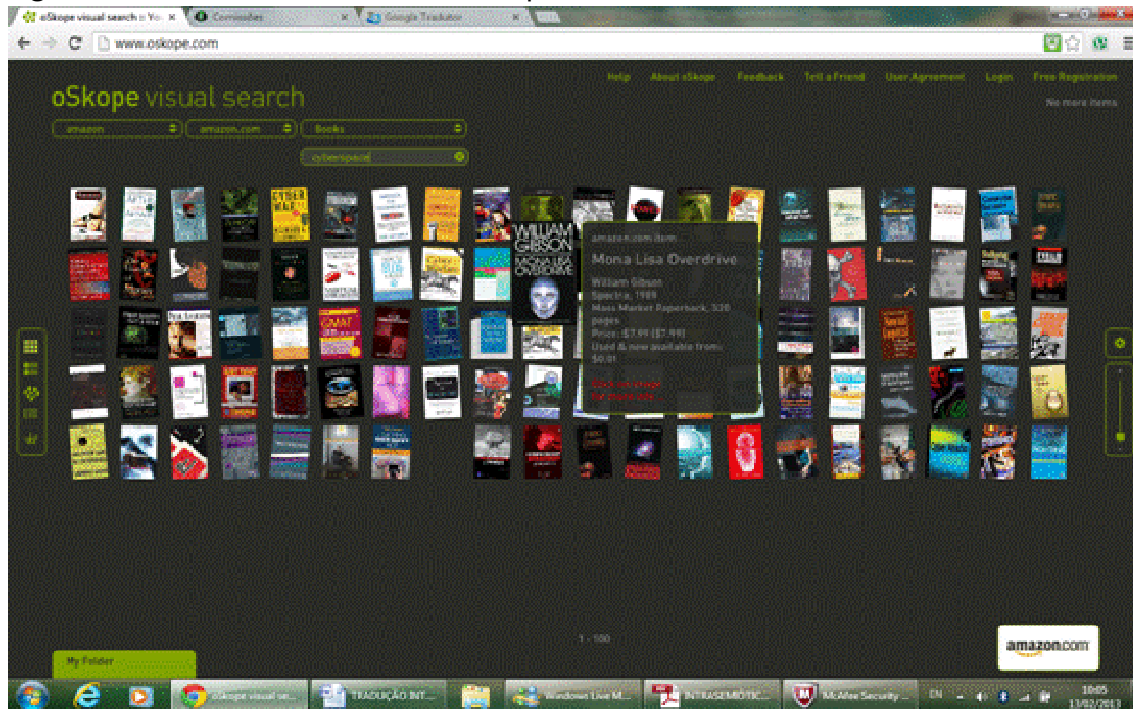
Percebe-se, a partir da Figura 4, que os resultados sobre *Cyberspace* não são similares e nem equivalentes, em sua grande maioria. Foi realmente insatisfatória a busca na base de fotografias do Flickr. Pode-se inferir que ou a tecnologia não está devidamente alinhada ou os sujeitos navegadores não estão rotulando devidamente suas fotos no Flickr, já que se trata de um serviço de mídia social.

Vale salientar que a imagem (objeto digital) é rotulada pelo seu produtor e, nesse sentido, vemos a indexação coletiva em funcionamento, a propósito da Web pragmática e sua equivalência cognitiva boa ou ruim dos sujeitos navegadores: “*Traduttore, traditore*, ou seja, traduzir, trair, pois se a linguagem liberta, também trai. “A linguagem é polissêmica, incompleta, ambígua.” (MOTTA, 2011, p. 49). E ainda, de acordo com o autor, toda tradução levanta uma série de apreensões porque a linguagem nunca obtém significações definitivas e universais.

Nesse quesito, por sua vez, a tecnologia do Google Images está muito mais aprimorada para compreender a linguagem e proceder as traduções com um nível de equivalência e similaridade satisfatório.

No exemplo, a seguir, fez-se uma busca na base de dados da Amazon por livros sobre o *Cyberspace*.

Figura 5 - Busca Intersemiótica no oskope Visual Search na Base da Amazon



Fonte: oSKope. Disponível em: <<http://www.oskope.com/>>.

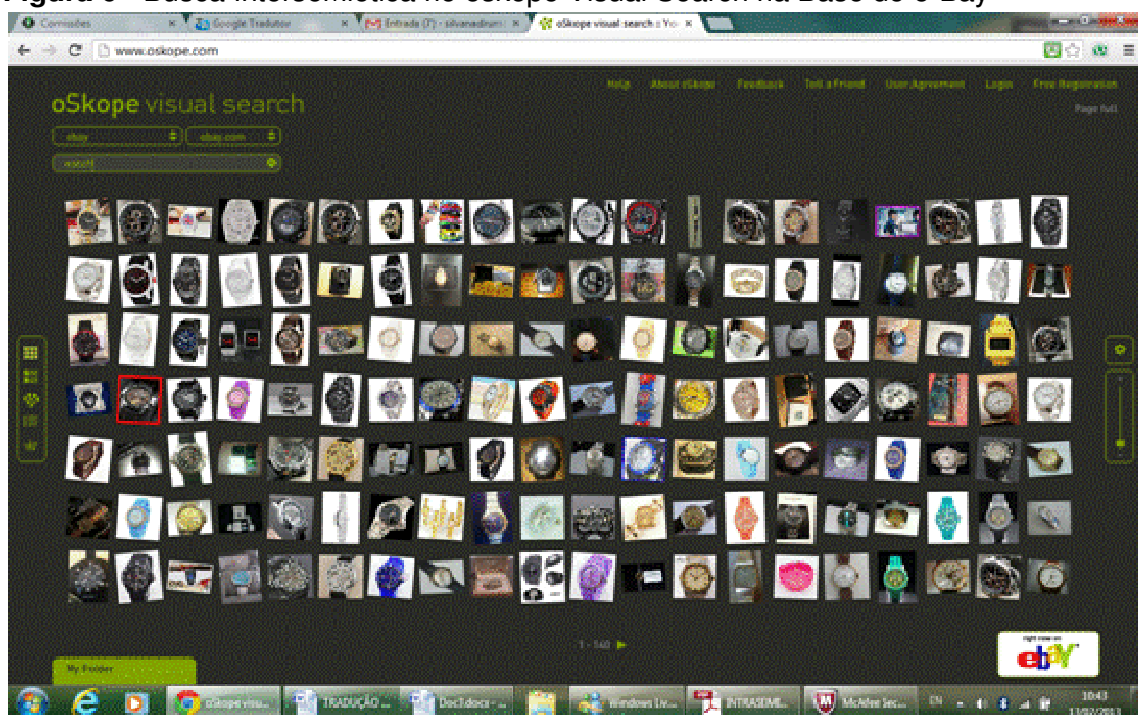
Diferentemente do Flickr, os metadados dos catálogos da Amazon fazem toda a diferença, pois o resultado é extremamente interessante e pertinente à *query* formulada. Com efeito, Morville e Callender (2010) ressaltam que a Amazon é campeã quando se trata em fazer uma pesquisa em sua base de dados e que navegar e pesquisar no *site* e em seus produtos têm a mesma importância, em termos de localização e descrição de seus produtos

Santaella (2007) afirma que é necessário compreender com profundidade a linguagem em mecanismos de intermediação e nas complexas mediações entre tecnologia e sociedade. O aprendizado da linguagem pela máquina, traduzido em um conjunto de normas e algoritmos, faz a diferença entre os níveis de tradução,

entre o igual e o parecido, enfim, sobre o princípio de similaridade da forma, seja de uma língua, seja de uma imagem.

Em outras palavras, os mecanismos de busca são construídos pelo homem e podem ser considerados como tecnologias intelectuais, tais mecanismos são constantemente aperfeiçoados e modificados. Nesse ambiente, temos as estruturas lógicas, ou seja, as linguagens formais, que são linguagens construídas que permitem que os sistemas de processamento da informação efetuem a mediação prática de nossas interações com o mundo simbólico.

Figura 6 - Busca Intersemiótica no oskope Visual Search na Base do e-Bay



Fonte: oSkope. Disponível em: <<http://www.oskope.com/>>.

Assim como no exemplo da Amazon, a consulta no e-Bay demonstrou-se muito satisfatória. A busca por *Watch*, conforme a Figura 6, mostra a qualidade dos resultados, inclusive ela pode ser reformulada por marca e modelo, conferindo muita especificidade à pesquisa.

De acordo com Feldman (1995), as tecnologias e sua qualidade semiótica que povoam o ciberespaço oferecem instrumentos de “[...] navegação, busca e atualização em um campo de referência multidirecionada.” (apud SANTAELLA, 2007, p. 395), e os mecanismos semióticos têm comprovado isso.

4 CONSIDERAÇÕES FINAIS

Os signos presentes no ciberespaço, tais como textos, imagens, áudios, vídeos e programações, têm gerado uma lógica a ser explorada. Não são poucas as consequências culturais, cognitivas e comunicacionais que o ciberespaço traz para a produção do conhecimento, da Arte e da informação em geral. Além disso, em consonância com Battelle (2006), consideramos que a busca está moldando nossa gramática cultural de forma que ainda não a compreendemos completamente.

Nesse processo intersemiótico dos mecanismos de busca, a tradução tem como objetivo buscar uma equivalência fundamental entre a *query* (construção inicial) e o traduzido (resultado ou aquilo que vai ser expresso na construção final), de forma que os valores sejam próximos, mas não idênticos (GONÇALVES, 2011).

Em grande medida, os resultados de todas as análises confirmam o enunciado anterior, à exceção da consulta da base do Flickr. Isso demonstra que as tecnologias digitais estão fazendo um bom trabalho de tradução intersemiótica e intrasemiótica. Pode-se incluir, embora não tenha sido objeto deste estudo, que o Google tem realizado, também satisfatoriamente, a tradução interlingual, abrangendo sua aprendizagem de máquina no escopo linguístico e semiótico.

Embora, aparentemente, os buscadores façam a tradução parecer um processo simples, ela é constituída de procedimentos complexos contidos em várias camadas de códigos, programas, linguagens e tecnologias.

Nessa malha cognitiva, em que traduzir é criar signos interpretantes, o Google tem encapsulado as *queries*, estudado a linguagem humana e aperfeiçoado seus algoritmos, fornecendo resultados mais equivalentes, do ponto de vista sintático e semântico.

As formas figurativas, dentro da classificação da linguagem visual, têm sido as mais recorrentes de ilustração em nossos estudos sobre os mecanismos de busca, pois as mesmas são mais próximas do mundo real e, como representação do visível, possuem uma forte proeminência indicial. Para o Google, o mundo (com seus textos) é um índice!

Já a palavra é, por si só, tradutora, pois o símbolo está ligado aos seus objetos simbolicamente. Mas a tradução verbo-visual, esse movimento entre o abstrato e o indicial, tem sido realizada com consistência, se comparada há alguns

anos. No entanto não é regra, por vários motivos: o padrão pragmático do ciberespaço, em que muitos rotulam seus objetos, e o investimento em Inteligência Artificial voltado para os mecanismos de busca.

Da metonímia⁷ (parte do discurso), que é a *query*, à forma da figura, os mecanismos híbridos operam o processo de tradução intersemiótico verbo-visual, e é nesse *continuum* semiótico que a geração de interpretantes se dá, homem-máquina, atualizando os sentidos e significando o ciberespaço.

REFERÊNCIAS

- ABBAGNANO, Nicola. **Dicionário de Filosofia**. São Paulo: Martins Fontes, 2003.
- ABREU, Joel Gomes de; MONTEIRO, Silvana Drumond. As matrizes da linguagem e a organização virtual do conhecimento. **Ci. Inf.**, v. 39 n. 2, p. 9-26, maio/ago., 2010. Disponível: <<http://revista.ibict.br/ciinf/index.php/ciinf/article/view/1788/1363>>. Acesso em: 08 fev. 2013.
- BATELLE, John. **A busca: como o Google e seus competidores reinventaram os negócios e estão transformando nossas vidas**. Rio de Janeiro: Elsevier, 2005.
- CANTER, David; RIVERS, Rod; STORRS, Graham. Characterizing user navigation through complex data structures. **Behaviour and Information Technology**, v. 4, n. 2, p. 93-102, 1985.
- CUNHA, Murilo Bastos da; CAVALCANTI, Cordélia Robalinho de Oliveira. **Dicionário de Biblioteconomia e Arquivologia**. Brasília: Briquet de Lemos, 2008.
- DELEUZE, Gilles. **Conversações**. São Paulo: Ed. 34, 1992.
- FRAGOSO, Suely. Quem procura acha? O impacto dos buscadores sobre o modelo distributivo da Word Wide Web. **Eptic: Revista de Economia Política de las Tecnologías da La Información y Comunicación**, v.9, n.3, n. p., sept/dic. 2007. Disponível em: <<http://ufrgs.academia.edu/SuelyFragoso>>. Acesso em: 15 mar. 2009.

⁷ Metonímia, aqui, significa parte do todo que é o discurso. Nesse sentido, a *query* é uma metonímia e há uma relação de contiguidade entre os mecanismos (índices) e os textos em suas múltiplas matrizes semióticas.

GONÇALVES, Paulo J. Veras. Tradução intra-semiótica. CONGRESSO BRASILEIRO DE CIÊNCIAS DA COMUNICAÇÃO, 34, 2-6 set. 2011, Recife. [Anais...] Recife: Intercom, 2011. Disponível em: <<http://www.intercom.org.br/papers/regionais/nordeste2011/resumos/R28-0328-1.pdf>>. Acesso em: 08 fev. 2013.

JAKOBSON, Roman. **Linguística e Comunicação**. São Paulo: Cultrix, 2007.

LEVY, Steven. **Google**: a biografia. Como o Google pensa, trabalha e molda nossas vidas. São Paulo: Universo dos Livros, 2012.

McLUHAN, Marshall. **Os meios de comunicação como extensões do homem**: (*understanding media*). Reimp. São Paulo: Cultrix, 1999.

MONTEIRO, Silvana Drumond; FERNANDES, Rogério Paulo Müller; GIRALDES, Maria Júlia C.; FRANKLIN, Benjamim Luiz; ANDRADE, Ilza Almeida de; VIGNOLI, Richele Grengue; FIDENCIO, Marcos Vinicius. A tradução intersemiótica dos mecanismos de busca no ciberespaço. In: SEMINÁRIO EM CIÊNCIA DA INFORMAÇÃO, 5., 22-24 abr.2013, Londrina. **Anais...** Londrina: Dep. Ciência da Informação, 2013. 1 CD- ROM, p. 1-20.

MORVILLE, Peter; CALLENDER, Jeffery. **Search patterns**: design for discovery. Canadá: O'Reilly, 2010.

MOTTA, Luiz Gonzaga. Um só mundo, múltiplas interpretações. **Revista de Jornalismo Científico e Cultural da Universidade de Brasília**, n.7, p. 48-49, ago./set. 2011.

PAZ, Otávio. **O signo em rotação**. São Paulo: Perspectivas, 1972.

PLAZA, Júlio. **Tradução intersemiótica**. 2.ed. São Paulo: Perspectiva, 2010.

SANTAELLA, Lúcia. **Linguagens líquidas**. São Paulo: Paulus, 2007.

_____. **Matrizes da linguagem e pensamento**: sonora visual verbal: aplicações na hipermídia. 3. ed. São Paulo: Iluminuras: FAPESP, 2005. 431 p.

_____. Três matrizes da linguagem-pensamento. **Cult**, Revista Brasileira de Literatura, n. 29, p. 1-4, 1999.

SOARES, Claudio. A anatomia da busca. **Pontolit**, 2008. Disponível em: <<http://blog.pontolit.com.br/autor/>>. Acesso em: 08 abr. 2009.

Title

Intersemiotic translation of search engine in cyberspace

Abstract

Introduction: The search process is part of today's society and the search engines are responsible for the preparation of indexes and indexing contemporaries in cyberspace.

Objective: In this work, the verb-visual translation intersemiotic Google Images and oSkope, considered to be hybrid by mixing languages, was analyzed to assess the level of equivalence of signs in the results.

Methodology: From the categorization of the mechanisms, the semiotic paradigm has been scrutinized in order to understand the matrix of language-thought, Santaella, the operation of search engines. Means intersemiosis the like mixture of languages, and hybridization its main feature. The concept of translation, appropriate Jakobson, encompasses the transmutation of the sign within a language (interlingual) from one language to another, or the translation itself, (interlingual) and between one language and another (intersemiotic). Another concept incorporated in the study was the intra-semiotic translation, within the matrix to collaborate with visual analysis of the visual language that is shaped like intelligibility and the figure axis as the visible representation.

Results: The behavior of the two searchers intersemiotic was demonstrated, in general, satisfactory

Conclusion: The language learning machine makes the difference between the levels of translation, between equal and similar, on the principle of similarity of form, is a language, it is an image.

Key words: Search Engine. Semiotics. Intersemiotic Translation.

Título

Traducción intersemiótica de los motores de búsqueda en el ciberespacio

Resumen

Introducción: El proceso de búsqueda es parte de la sociedad actual y los motores de búsqueda son los responsables de la preparación de índices y indexación contemporáneos de en el ciberespacio.

Objetivo: En este estudio, la traducción intersemióticas verbal-visual de lo Google y oSkope considerados híbridos de la mezcla de idiomas se evaluó para evaluar el nivel de equivalencia de los signos en los resultados.

Metodología: A partir de la clasificación de los mecanismos, el paradigma semiótico está siendo objeto de estudio con el fin de entender la matriz del lenguaje-pensamiento, Santaella, el funcionamiento de los motores de búsqueda. Se entiende intersemiose tal como una mezcla de idiomas, siendo la principal característica de la hibridación. El concepto de traducción, Jakobson caso, incluye la transmutación de la señal en un idioma (intra lingüística) de un idioma a otro traducción (interlingüística) o propiamente dicho, y entre un idioma y otro (intersemiótica). Otro concepto incorporado en el estudio fue la traducción intrasemiótica verbal-visual entro de la matriz visual, para colaborar en el análisis del

lenguaje visual, que es la forma y la figura como inteligibilidad de la representación del visible.

Resultados: El comportamiento de los dos motores intersemióticos se ha demostrado, en general, satisfactorio.

Conclusiones: La aprendizaje de idiomas por la máquina que hace la diferencia entre los niveles de la traducción entre los pares y similares, en el principio de similitud de la forma, sea un lenguaje, sea una imagen.

Palabras clave: Motores de Búsqueda. Semiótica. Traducción Intersemiótica.

Recebido em: 28.06.2013

Aceito em: 10.08.2013