

## Preservação Digital, Recuperação da Informação e Linguagem

**Ernesto Carlos Bodê**

[ernesto@unb.br](mailto:ernesto@unb.br)

**Renato Tarciso Barbosa de Sousa**

[renasou@unb.br](mailto:renasou@unb.br)

Universidade de Brasília, Faculdade de Ciência da Informação, Brasília, DF, Brasil

**Resumo:** O artigo refere-se aos primeiros resultados de uma pesquisa sobre consequências de problemas de funcionamento em sistemas de recuperação da informação em longo prazo, em relação ao uso da linguagem nesses sistemas. Parte-se do princípio de que tais problemas podem afetar negativamente a preservação de documentos digitais armazenados por longos períodos de tempo. Os resultados ora apresentados referem-se a um estudo do conceito de linguagem a partir de três áreas, com diferentes pontos de vista: filosofia da linguagem, semiótica e linguística. Com observações específicas a partir de cada uma dessas áreas são apresentadas observações que se consideram relevantes para o adequado funcionamento de sistemas de Recuperação da Informação (RI) em longo prazo. Ao final também são apresentados os futuros desdobramentos de estudos a partir do ponto atual das pesquisas.

**Palavras-chave** filosofia da linguagem; linguística; preservação digital; recuperação da informação; semiótica.

### Digital preservation, information retrieval and Language

**Abstract:** The article refers to the first results of a research on consequences of operating problems in information retrieval systems in the long term in relation to the use of language in these systems. Part it is assumed that such problems can negatively affect the preservation of digital documents stored for long periods. The results presented here refer to a study of the language concept from three areas with different points of view: philosophy of language, semiotics and linguistics. With specific from each of these areas observations considered relevant for the proper functioning of IR systems in the long run are presented. At the end of the future developments of studies from the current point of the research are also presented.

**Keywords:** digital preservation; information retrieval; Linguistics; Semiotics.

### Preservación digital, recuperación de la información y Lenguaje.

**Resumen:** El artículo se refiere a los primeros resultados de una investigación sobre consecuencias de problemas de funcionamiento en sistemas de recuperación de información a largo plazo, en relación al uso del lenguaje en esos sistemas. Se parte del principio de que tales problemas pueden afectar negativamente la preservación de documentos digitales almacenados por largos periodos de tiempo. Los resultados ahora presentados se refieren a un estudio del concepto de lenguaje a partir de tres áreas, con diferentes puntos de vista: filosofía del lenguaje, semiótica y lingüística. Con observaciones específicas a partir de cada una de esas áreas son presentadas observaciones que se consideran relevantes para el adecuado funcionamiento de los sistemas de Recuperación de la Información (RIE) a largo plazo. Al final también son presentados los futuros desdoblamientos de estudios a partir del punto actual de las investigaciones.

**Palabras clave:** Lingüística; preservación digital; recuperación de la información; Semiótica;

## 1 Introdução

A Preservação Digital é uma nova área que surge diante da realidade dos Documentos Digitais na vida das pessoas e das empresas públicas e privadas. Dentre os vários problemas enfrentados por essa área para “assegurar a integridade e o acesso aos documentos pelo maior prazo possível, idealmente para sempre” (BODÊ, 2008, p. 18) dos conteúdos registrados em documentos digitais, os mais desafiadores dizem respeito a equacionar hoje problemas que surgirão no futuro. E esse é um aspecto frequentemente mal compreendido na área de Preservação Digital.

Resolver problemas que surgirão no futuro não é algo muito comum de empreender-se. Ainda mais porque “futuro” aqui não se refere aos próximos anos, situação com a qual se preocupa qualquer atividade de planejamento. Quando nos referimos a “futuro” em preservação digital, estamos falando de várias décadas ou mesmo vários séculos à frente.

Talvez fosse melhor usar a expressão preservação para as futuras gerações de pessoas. Mas diferentemente do que é feito em termos práticos para a preservação de suportes materiais, ou seja, mantê-los intactos e/ou restaurá-los<sup>1</sup>, no caso da Preservação Digital, trata-se de agir muito mais ativamente no conteúdo atual e também em toda a tecnologia envolvida para criar ou transformar os bits que registram conteúdos sensíveis aos nossos sentidos. A atenção da Preservação Digital para aquilo que ainda está ativo hoje é fundamental, pois pequenas alterações nas tecnologias ou mesmo defeitos nas sequências de bits podem inviabilizar totalmente o acesso a esses conteúdos. Daí o paradoxo: atuar hoje pensando no futuro de longo prazo.

Ainda que inserido num projeto maior que tem como tema principal a Preservação Digital, esse artigo não tem como meta tratar desse assunto. Para uma visão geral do tema Preservação Digital em Arquivos, Museus e Bibliotecas ver (GORMAN; SHEP, 2006) e para princípios e práticas mais específicas ver (BORGHOFF; RODIG; SCHEFFCZYK; SCHMITZ, 2006).

Diante desse contexto, este artigo tem como motivação inicial abordar justamente um problema que em a longo prazo pode afetar o acesso a Documentos Digitais. Trata-se dos Sistemas de Recuperação da Informação (RI), que podem apresentar inconsistências e afetar parcial ou totalmente a *busca*, o *descobrimento* e a efetiva *recuperação* de considerado Documento Digital.

Partimos do princípio que todo Sistema de RI, pelo menos da forma como é estruturado hoje, baseia-se em três pilares: o *Documento*, a *Linguagem* e a *Informação*. Trata-

---

<sup>1</sup> Os verbos “manter” e “restaurar” se traduzem em uma série de procedimentos no âmbito da preservação de documentos tradicionais. Há várias normas e procedimentos para Armazenagem e Manuseio, Planejamento, Reformatação e Administração de Emergências, entre outras subáreas. Para maiores aprofundamentos, pode-se consultar: <http://www.argsp.org.br/cpba/>.

se de três aspectos conceituais diferentes que devem ser considerados para que a efetiva implementação tecnológica resulte em pleno funcionamento.

*Documento* é o aspecto mais óbvio nesses sistemas, principalmente porque se trata de algo físico e, portanto, muito concreto. Isso é válido, inclusive, para Documentos Digitais tanto quanto o é para Documentos em suportes tradicionais como o papel, por exemplo. Em outras palavras, todo Documento Digital sempre estará fixado em algum suporte material. O aspecto do Documento como uma coisa concreta acarreta problemas de espaço de armazenamento, velocidade de transmissão dos respectivos bits e plataformas específicas de *software* e *hardware* para efetivação do uso, entre outros problemas.

*Linguagem*, por sua vez, é um conceito bem mais abstrato se comparado ao conceito *Documento*. Tratando-se de Documentos com conteúdos passíveis de processos de significação humanos, a Linguagem estará presente necessariamente. E de fato, no presente trabalho, nosso objetivo é abordar especificamente esse aspecto na perspectiva da Preservação Digital de longo prazo, no que diz respeito à sua relação com Sistemas de RI. Mais especificamente, nossa meta aqui é explorar e detalhar ao máximo o conceito *Linguagem*, consoante a abordagem adotada.

*Informação*, finalmente, é o último aspecto dessa tríade. E trata-se de algo bem mais abstrato e difícil de abordar, se comparado ao aspecto “Documento” ou “Linguagem”. Retomando a Teoria da Informação<sup>2</sup>, nos moldes como definida originalmente por Shannon, temos um transmissor A, um receptor B, um canal e o que é transmitido, isto é, a coisa transmitida. Ora, essa coisa não é o “Documento” em si, ainda que possa conter uma cópia do conteúdo de um documento ou referências a outro, em qualquer caso agregado com outros dados específicos naquela transmissão. Também não é a *Linguagem*, aspecto que, como veremos adiante, está associado ao conteúdo nesse mesmo *Documento*. Assim, temos de falar em outra instância num Sistema de RI, a qual utiliza necessariamente esse tipo de transmissão, para além do “Documento”, o que chamamos aqui, grosso modo, de “Informação”.

Para tanto, empregamos uma Metodologia que consiste na exploração do conceito Linguagem em três áreas distintas, mas bastante inter-relacionadas, a saber: *Filosofia da Linguagem*, *Semiótica* e *Linguística*. A partir da apresentação de brevíssimos resumos históricos do conceito Linguagem em cada área, derivamos e apresentamos nossas conclusões que julgamos relevantes para Sistemas de RI. Ao final, tecemos nossas considerações sobre os próximos desdobramentos em relação às conclusões apresentadas.

---

<sup>2</sup> Tomamos aqui Teoria da Informação como matematicamente tratada a transmissão de informação entre o elemento transmissor e o elemento receptor num meio com ruídos, vide por exemplo (MACKAY, 2003).

## **2 Aspecto linguagem**

A linguagem é de importância vital para qualquer área da ciência e da filosofia. Todo e qualquer processo de pesquisa, raciocínio ou comunicação não pode prescindir do uso de algum tipo de linguagem. De fato, é possível construir uma classificação hierárquica sobre esse conceito, considerando campos de estudo e outros (sub) conceitos inter-relacionados. Nosso problema consiste em explorar o conceito Linguagem à luz da *Filosofia da Linguagem*, da *Semiótica* e da *Linguística*.

Um caminho possível para iniciar uma discussão científica acerca de um conceito é apresentar definições prévias de outros autores sobre o mesmo conceito. No entanto, nos limites deste trabalho, isso traria mais problemas do que ajuda. A Linguagem instiga e instigou uma parcela importante de filósofos e cientistas, que originaram visões díspares, quando não conflitantes. Por isso, para dar cabo desse problema, utilizaremos outro método, que não passe somente ou essencialmente pelo (re) exame de conceitos autorais. O caminho alternativo escolhido foi apresentar uma análise do conceito de Linguagem a partir de áreas e não diretamente pelos autores dessas áreas.

Escolhemos as áreas que, a nosso ver, mais se destacam em relação ao estudo do problema da Linguagem. A vantagem nessa escolha reside no fato de que as três áreas elencadas tiveram sua grande evolução durante o século XX, a relativo pouco tempo, o que nos permite, portanto, considerar os últimos avanços do conceito estudado. Além disso, elas estão muito interligadas, inclusive com alguns expoentes, atuando em mais de uma área, ou nos limites delas. Finalmente, a problemática da Linguagem é o objeto principal das três áreas, e cada uma delas possui várias outras subáreas mais.

### **2.1 BREVE Resumo histórico: Filosofia da Linguagem**

Se considerarmos que a linguagem humana em suas manifestações específicas (línguas) são imprescindíveis para o desenvolvimento do conhecimento humano, então uma *Filosofia da Linguagem* pode, pelo menos em alguns aspectos, ser importante para qualquer outro ramo da Filosofia Geral, e, como consequência, potencialmente para qualquer outro ramo de ciência também, conforme essa observação: “os filósofos têm sido motivados por um desejo de dizer algo sistemático sobre essas noções de compreensão linguística, significado e conhecimento” (MILLER, 2010, p. 5).

Além disso, podem-se identificar problemas filosóficos que envolvem a linguagem desde os gregos, pelo menos<sup>3</sup>, de modo que é preciso delimitar o que iremos abordar nessa seção sobre *Filosofia da Linguagem*, ainda mais considerando tratar-se de um brevíssimo resumo histórico.

A figura 1 relaciona os principais autores que tratam da Filosofia da Linguagem nos séculos XIX e XX, tendo sido adaptada de uma tabela proposta para explicar o desenvolvimento da Filosofia Analítica. Na sequência, e a partir do resumo dessa figura, analisaremos o Empirismo Lógico, a Filosofia da Linguagem Comum e um pouco das ideias de Wittgenstein.

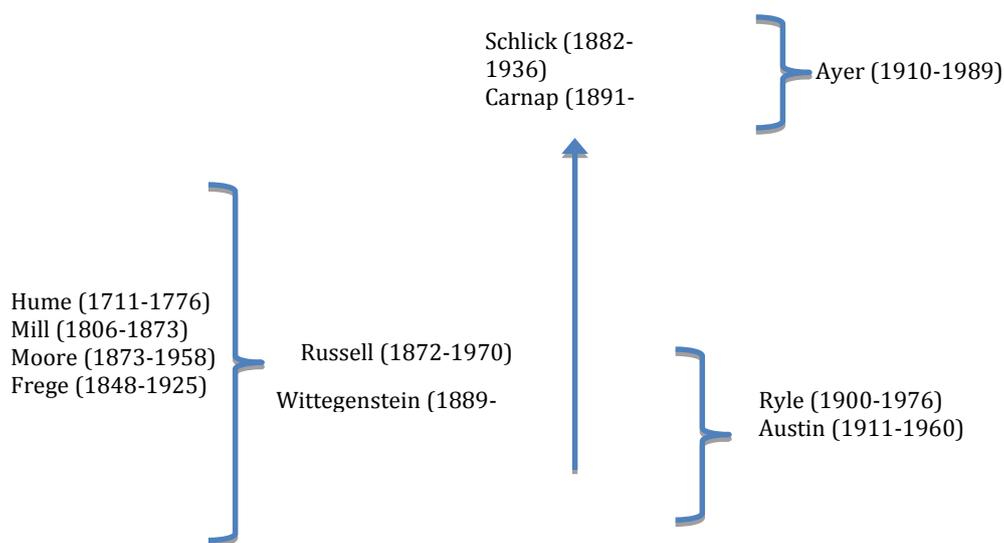


Figura 1 - Exponentes da Filosofia da Linguagem. Adaptado de Hunnex (2003, p. 130 e 131).

A figura apresentada sugere o papel importante e precursor de Russell no desenvolvimento da Filosofia da Linguagem, cujos trabalhos mais relevantes para nosso contexto estão em *Princípios da Matemática* (1903), *A filosofia do atomismo lógico* (1918) e *Investigação de significado e verdade* (1940). Influências importantes, conforme a figura, são Moore e Frege, e, antes desses, Stuart Mill e Hume.

A figura também evidencia o importantíssimo papel de Ludwig Wittgenstein, cujo trabalho fundamental, conhecido como o primeiro Wittgenstein, foi o *Tractatus logico-philosophicus* (1921). O que ficou conhecido como Círculo de Viena, por exemplo, foi explicitamente influenciado pelos estudos de Wittgenstein. Tanto as obras de Russell como também as de Wittgenstein são de grande importância para a Filosofia da Linguagem Comum

<sup>3</sup> Os gregos trataram de vários aspectos da Linguagem, em Platão com os textos *Crátilo*, *Sofista*, *Górgias* e *Fedro*, em Aristóteles com *Retórica*, *poética* e *Política* por exemplo (MARCONDES, 2010).

(AUDI, 2006, p. 358) e sua possível redução a formas precisas para comunicação e entendimento em filosofia e ciência.

A corrente dos *Positivistas Lógicos* ou *Empirismo Lógico* levou adiante essas ideias ao distinguir os critérios de *verificabilidade* e *falsidade*<sup>4</sup> para as proposições apresentadas na linguagem. Schlick e Carnap são dois importantes representantes nessa linha de pensamento. Desse último autor, são obras importantes: *A sintaxe lógica da linguagem* (1937), *Introdução à Semântica* (1942) e *Significado e necessidade* (1947).

Mais tarde, no segundo Wittgenstein (*Investigações Filosóficas*, 1953), algumas ideias importantes que serviram de base para o atomismo lógico e para o positivismo lógico são rejeitadas. Wittgenstein propõe a análise da linguagem como apresentada em sua forma e contexto originais,

“Assim, ao aceitar a linguagem essencialmente como ela é e ao buscar somente aclarar e esclarecer o que estava diante dele, ele procurou – como Moore – levar a linguagem do seu uso filosófico e frequentemente problemático de volta ao seu uso natural e ordinário” (HUNNEX, 2003).

O Positivismo Lógico<sup>5</sup> teve uma preocupação especial com enunciados ou proposições, parte essencial da linguagem utilizada em ciência e/ou filosofia. Um dos pontos mais característicos para aquele grupo de filósofos foi sua postura em relação aos enunciados metafísicos, “*a tese de que todos os enunciados metafísicos são desprovidos de sentido, porque não são verificáveis empiricamente.*” (ABBAGNANO, 2007, p. 381). O mesmo autor também afirma que há duas teses em comum entre esse grupo e as ideias do primeiro Wittgenstein: em primeiro lugar, “*os enunciados factuais, isto é, que se referem a coisas existentes, só têm significado se forem empiricamente verificáveis*” (ABBAGNANO, 2007, p. 382); em segundo, “*existem enunciados não verificáveis, mas verdadeiros com base nos próprios termos que os compõem [matemática e lógica]*” (ABBAGNANO, 2007, p. 382).

O segundo<sup>6</sup> Wittgenstein teve influência decisiva na corrente chamada de Filosofia da Linguagem Comum<sup>7</sup>, inclusive, bem antes de sua segunda obra. Já em 1929, ele defende:

---

<sup>4</sup> Vide mais sobre esses critérios em Popper (AUDI, 2006, p. 735).

<sup>5</sup> Essa corrente teve início no que é conhecido historicamente como Círculo de Viena, “um grupo de filósofos analíticos com inclinação científica e matemática, que se encontraram nos *seminários* privados de Moritz Schlick, nas manhãs de sábado, em Viena a partir de 1923” (MAUTNER, 2011)

<sup>6</sup> Apesar da predominância da visão de dois Wittgenstein “alguns intérpretes tem argumentado por uma unidade entre todos os estágios de seu pensamento, enquanto outros falam de uma divisão com mais nuances, adicionando” (BILETIZKI; MATAR, 2014, p. 1)

<sup>7</sup> A Filosofia da Linguagem Comum é uma linha filosófica que “*coloca a linguagem comum como objeto*”, além disso “*considerando-a como método julga que sua análise é o instrumento mais apropriado para resolver importantes problemas filosóficos*” (ABBAGNANO, 2007, p. 719).

“Eu costumava acreditar que havia a linguagem ordinária que todos normalmente falamos e uma linguagem primária que expressava o que realmente conhecíamos, a saber, os fenômenos. Eu também falava de um primeiro e um segundo sistema. Agora, desejo explicar por que eu não mais me ligo a essa concepção” (WCV: 45 apud CHILD, 2013, p. 92).

Das diferentes abordagens sobre Filosofia da Linguagem nos parágrafos anteriores, pode-se destacar alguns pontos. Primeiro, nota-se uma evolução na importância da linguagem comum em contrapartida às linguagens artificiais. Explicando melhor, ainda em Russell e alguns antecessores imediatos, a linguagem da Lógica era a solução para transformar os enunciados importantes em uma linguagem artificial (proposições), que pretendia ser perfeita para expressar ciência e filosofia. Tomando como referência o Círculo de Viena, houve uma evolução com as teses do empirismo, no sentido de avaliar enunciados ainda na linguagem comum. Finalmente, após meados do século XX a linguagem comum ou ordinária volta a ocupar o foco das atenções.

No entanto, como veremos mais adiante nas abordagens sobre Semiótica, e principalmente, sobre Linguística, “Linguagem Comum” não significa uma coisa simples ou menos complexa do que a própria Linguagem da Lógica, sendo, talvez, até mais complexa que a Linguagem Matemática.

## 2.2 BREVE resumo histórico: Semiótica

Várias áreas e linhas de pensamento na ciência tentam explicar o mundo ou, pelo menos, aspectos do mundo. Assim, a química procura explicar certos efeitos nos materiais, na água e nos seres vivos; a física procura dar conta do porquê e do como há movimento, propagação da luz ou produção da energia elétrica. Na filosofia podem-se detectar as mesmas metas, por exemplo, na *cosmologia*, na *filosofia da linguagem* ou na *teoria do conhecimento*. A Semiótica também tem a pretensão de explicar o mundo, e para alguns até o mundo como um todo<sup>8</sup>, de modo que tem sido difícil situá-la de maneira definitiva como parte das ciências ou da filosofia, pois não é totalmente homogênea em suas linhas de estudos<sup>9</sup>. Isso talvez seja um reflexo de sua curta existência<sup>10</sup>.

---

<sup>8</sup> Para Charles Sanders Peirce, a Semiótica possui um potencial imenso “*Trata-se, para ele, de um quadro de referência que engloba qualquer outro estudo*” (DUCROT; TODOROV, 2010, p. 89).

<sup>9</sup> De acordo com Tzvetan Todorov “Apesar da existência desses trabalhos [após citar vários autores na América e Europa] e de quase um século de história (e vinte séculos de pré-história), a Semiótica continua sendo mais um projeto do que uma ciência constituída.” (DUCROT; TUDOROV, 2010, p. 93).

<sup>10</sup> A Semiótica atual tem dois pontos distintos, mas muito bem definidos entre o final do século XIX e início do século XX com as ideias de Peirce e Saussure, como veremos mais adiante. Portanto, como área de estudo, possui pouco mais de um século.

Certamente, o objetivo da *Semiótica Geral* é algo que se estrutura num campo de estudos que envolve ou está nos limites de conceitos como *informação, comunicação, sinal, código, ícone, símbolo, significação e representação do conhecimento*. Pelo menos esses são os termos evidenciados em trabalhos publicados da área<sup>11</sup>. Os primeiros autores, hoje clássicos, que trataram das bases da Semiótica e que contribuíram para o que ela é essencialmente hoje, foram Saussure (1857-1913) e Peirce (1839-1914). Ambos propuseram teorias sobre o signo, apesar das diferenças de pontos de vista de cada um. O primeiro tinha como área de trabalho a *Linguística*, e o último estudava a *Lógica*<sup>12</sup>.

Assim, um caminho para tratarmos da Semiótica é discutir sobre o que é um signo nessa área e sua relação com outros conceitos próximos, no mesmo campo de estudo. Para tanto, partimos do princípio de que as propostas e definições para signo estão relacionadas a vários fenômenos e situações do homem em sua relação com o mundo, nos processos de troca e obtenção de conhecimento.

Um esquema muito simplificado do mundo e útil aos nossos propósitos exige a presença de pelo menos dois indivíduos para que seja concretizada a comunicação das duas mentes, através de algum meio e tipo de linguagem. O que pode ser comunicado pode ser classificado de infinitas formas. Mas imaginemos, para efeitos de nosso esquema, uma separação entre coisas naturais e artificiais (produzidas pelo homem). A partir das coisas naturais temos as orgânicas e inorgânicas. Para os orgânicos temos os racionais (homem) e irracionais (demais criaturas vivas). As coisas artificiais podem ser subdivididas em coisas para registrar o conhecimento e demais artefatos com outras funções primárias (ferramentas de pesca, marcenaria, abrigos e casas, veículos e assim por diante).

Nesse esquema simplificado do mundo, o homem, ao mesmo tempo em que está inserido numa posição taxonômica bem definida, é também o criador desse mesmo esquema (no caso, do meu ponto de vista) e também um ser que se correlaciona com as demais coisas. Primeiro com outros seres semelhantes, outros indivíduos, e também com todas as demais coisas e processos diretamente relacionados a essas coisas: criação de artefatos artificiais, identificação e classificação de coisas naturais, ritos religiosos, processos de produção do

---

<sup>11</sup> Umberto Eco em seu *Tratado Geral de Semiótica* (ECO, 2012), ainda na introdução da obra, procura definir o que ele acredita serem os limites e fins de uma teoria semiótica. Sua visão é interessante, pois tendo publicado a obra no final da década de 70 (1976) permite uma visão histórica importante em relação às primeiras ideias de Peirce e Saussure.

<sup>12</sup> De fato, Peirce às vezes sugere que sua *Lógica* e o que ele chama de *Semiótica*, para ele, são a mesma coisa “*Em seu sentido geral, a lógica é, como acredito ter mostrado, apenas um outro nome para semiótica*” (PEIRCE, 2012, p. 45).

conhecimento e tantos outros processos, provavelmente impossíveis de serem descritos em sua totalidade, pelo menos aqui.

Esse esquema representativo do mundo condiz exatamente com as primeiras ideias sobre Semiótica, na visão peirciana, ou seja, a Semiótica como algo que poderia abarcar a totalidade do que o homem sabe ou pode saber:

*Segundo Peirce, não só as palavras da linguagem e os signos não verbais produzidos intencionalmente pelos seres humanos para comunicar-se, mas qualquer evento, estado, objeto do mundo externo e qualquer evento ou estado mental (representação, emoção, sensação etc.), podem entrar numa relação semiótica desde que interpretados por algum interpretante como signo de qualquer outra coisa. (ABBAGNANO, 2007, p. 1034).*

Em seu *Tratado Geral de Semiótica*, Umberto Eco declara que o objetivo de seu livro “é explorar as possibilidades e as funções sociais de um estudo unificado de todo e qualquer fenômeno de significação e/ou comunicação” (ECO, 1976, p. 1). A partir daí conclui-se que esse esquema de mundo interessa enquanto existam “fenômenos de significação e/ou comunicação” (ECO, 1976, p. 1). A análise desses fenômenos de significação depende totalmente da ideia de *Signo Semiótico*.

Uma das definições mais estudadas de signo é aquela proposta por Peirce, o qual afirma que “é tudo aquilo que está relacionado com uma segunda coisa” (PEIRCE, 2012, p. 28). Ela é conhecida como triádica, pois envolve necessariamente três elementos em sua composição: (1) um objeto (que no nosso esquema proposto pode ser qualquer coisa, inclusive os próprios indivíduos), (2) um interpretante (que não é necessariamente uma pessoa<sup>13</sup>) e (3) signo algo que só existe na relação (1) com (2).

Peirce desenvolve, então, em torno dessa proposta de Signo, toda uma complexa teoria. Por exemplo, deriva daí uma classificação de signos entre ícones, índices e símbolos. Se para Peirce, a Semiótica é uma disciplina através da qual o mundo poderia ser explicado, e não está claro como exatamente seria essa explicação, para muitos dos demais teóricos da Semiótica, ela é algo aplicável somente nos processos de comunicação e linguagem. De fato, há pesquisas até mesmo sobre a comunicação e linguagem de seres não racionais (abelhas, baleias e golfinhos são os mais citados), daí porque em nosso “esquema de mundo” essas coisas também foram destacadas.

É através dos processos de comunicação e uso da linguagem diretamente entre indivíduos racionais, ou até num mesmo indivíduo com si mesmo, utilizando ou não alguma

---

<sup>13</sup> Muitas das ideias de Peirce em relação ao signo são polêmicas e às vezes obscuras. Isso talvez seja o resultado da falta de uma obra única e definitiva sobre Semiótica. Apesar da tradução brasileira de sua obra (PEIRCE, 2012) ter o título “Semiótica”, o título original é “The collected papers of Charles Sanders Peirce”, um apanhado de vários de seus escritos, incluindo correspondência e resenhas.

forma de registro do conhecimento nesses processos, que o Signo Semiótico<sup>14</sup> aparece como um conceito útil para apoiar as explicações de fenômenos observados.

Saussure, por sua vez, vê a Semiótica como uma disciplina anterior à Linguística:

*Pode-se, então, conceber uma ciência que estude a vida dos signos no seio da vida social; ela constituiria uma parte da Psicologia Social e, por conseguinte, da Psicologia Geral; chamá-la-emos de Semiologia (do grego semeion, "signo"). Ela nos ensinará em que consistem os signos, que leis os regem. Como tal ciência não existe ainda [escrito começo séc. XX], não se pode dizer o que será. (SAUSSURE, 2012, p. 46).*

Ele não tentou, no entanto, desenvolver estudos semióticos, mas apenas procurou posicionar a Linguística num contexto maior em relação ao que entendia por Semiótica. O conceito de Signo em Saussure, que é praticamente contemporâneo ao de Peirce (como confirma as datas de seus trabalhos publicados), apesar de independente dele, aparece aplicado à linguagem humana concreta envolvendo a palavra e a língua. É relevante também ressaltar que o Signo de Saussure não é triádico, mas diádico.

Para Saussure, o signo linguístico:

*Não une uma coisa e uma palavra, mas um conceito e uma imagem acústica. Esta não é o som material, coisa puramente física, mas a impressão (empreinte) psíquica desse som, a representação que dele nos dá o testemunho de nossos sentidos; tal imagem é sensorial e, se chegamos a chama-la "material", é somente nesse sentido, e por oposição ao outro termo da associação, o conceito, geralmente mais abstrato. (SAUSSURE, 2012, p. 106).*

Se nosso esquema apresentado tem o objetivo de associar a ideia de signo no mundo como um todo, um outro esquema representaria a ideia de signo aplicada ao universo da língua e nos estudos de linguística. Isso na esteira das ideias de Saussure com relação a signos, ou seja, sua aplicação em relação à língua.

Para Hjelmslev a ideia de signo, desenvolvida a partir dos trabalhos de Saussure, está associada a planos que estruturam e permitem compreender fenômenos linguísticos. Em sua teoria Glossemática, ele propõe o plano da expressão e o plano do contexto. E para cada um desses planos o signo é composto por sentido (pensamento), forma (limites em cada língua) e a substância que associa sentido e forma (DUCROT; TODOROV, 2011, p. 33).

### **2.3 BREVE Resumo histórico: Linguística**

Assim como a Filosofia da Linguagem e a Semiótica, a Linguística também tem como objeto maior de estudos a Linguagem. Em relação à Filosofia da Linguagem por exemplo, "os interesses que justificam a atenção da Filosofia da Linguagem são no mais das vezes diferentes

---

<sup>14</sup> Importante observar aqui que o termo Signo também é utilizado em outras áreas como a Linguística: signo linguístico, permitindo em outros contextos outros significados.

e distintos dos interesses que estimulam e orientam a Linguística (mas que entre eles não há conflito nem antítese)” (ABBAGNANO, 2007, p. 721).

A Linguística, como é conhecida hoje, pode ser compreendida em seu processo de desenvolvimento histórico em, pelo menos, oito marcos temporais, apresentados na tabela 1.

<b>Marco</b>	<b>Ano</b>	<b>Expoentes</b>
Port-Royal	1660	Claude Lancelot
Linguística Histórica	1816	Bopp
Neogramáticos	1850-1900	G. Curtius / H. Paul
Distribucionismo	1914	L. Bloomfield
Saussurianismo	1916	F. de Saussure
Escola Praga	Década 30	Trubetzkói, Martinet, Jakobson
Glossemática	1943	L. Hjelmslev
Gerativismo	1950	N. Chomsky

Tabela 1 - Com dados reunidos a partir de (DUCROT; TODOROV, 2010)

A coluna “Ano” indica o ano de publicação das principais obras em cada fase, contudo, apesar da sequência cronológica, não se trata de uma simples “evolução” entre fases. De fato, alguns marcos na tabela ocorrem simultaneamente e de maneira bastante independente, como o Saussurianismo na Europa e o Distribucionismo na América do Norte. Trata-se muito mais de ideias e teorias que sobreviveram, e em alguns casos ainda sobrevivem, paralelamente.

É também importante notar que as influências das obras em cada marco, em diferentes graus, até hoje são percebidas, como o Saussurianismo, cuja obra “Curso de Linguística Geral” (SAUSSURE, 2012), ainda hoje é impressa e estudada em universidades ao redor do mundo. Além disso, o Gerativismo também é uma importante fase ainda presente em nossas universidades, notadamente na América do Norte.

Apesar de possuírem a Linguagem como objeto maior comum, qual é o interesse específico em Linguística, que a distingue de Filosofia da Linguagem e Semiótica? A resposta para essa pergunta é a *Língua Humana*<sup>15</sup> produzida naturalmente no meio social<sup>16</sup>. E esse é um ponto da mais alta importância para este estudo. Temos utilizado até agora, intencionalmente, somente o termo “Linguagem” ou “tipos de Linguagem”. E, de fato, parece ser o mais adequado quando se trata de Filosofia da Linguagem e Semiótica, visto que essas áreas não se

<sup>15</sup> Eu prefiro distinguir entre língua humana e linguagem humana, pois acredito que assim os termos ficam mais claros. Mas encontra-se com facilidade o termo Linguagem Humana (FIORIN, 2013) para o que é afinal Língua Humana. Reservo o termo linguagem humana para a capacidade de linguagem, que não se limita somente à língua nacional específica utilizada.

<sup>16</sup> Ainda que, excepcionalmente, casos isolados de línguas humanas artificiais como o Esperanto também possam ser parte da Linguística.

ocupam somente da Língua Humana natural, mas também de Linguagens ou Línguas artificiais como as proposições lógicas ou as linguagens de sinais. Mas a Linguística focaliza sua atenção na Língua Humana especificamente, falada ou registrada em suportes. Os estudos em Linguística podem ser divididos em pelo menos seis grandes grupos, conforme tabela 2.

<b>Grupos</b>	<b>Estudo de...</b>
Fonética	sons da língua fisicamente.
Fonologia	sons da língua nos aspectos cognitivos.
Morfologia	formação das palavras.
Sintaxe	formação das sentenças.
Semântica	significado na língua.
Pragmática	uso da língua.

Tabela 2 - Extraída e resumida a partir de (GREIMAS; COURTÉS, 2012)

Retomaremos mais adiante algumas outras ideias na área de Linguística, por ora basta elencar algumas características das línguas humanas que oferecem uma bela imagem da complexidade das línguas:

*[...] são duplamente articuladas, suas unidades definem-se umas por oposição às outras, seus signos são arbitrários; a redundância está sempre presente [...]; apresentam ambiguidades, dissimetrias e irregularidades; permitem a recursividade; estão em perpétua mudança; permitem a inventividade, a criatividade, o deslocamento de sentido (como nas metáforas por exemplo), os jogos com os sons e os sentidos; são estruturadas em três níveis (o dos sons, o da gramática e o dos sentidos); seus significantes são lineares nas manifestações orais das línguas (não podemos esquecer-nos de que temos também as línguas de sinais, como Libras, a língua brasileira de sinais); suas unidades são discretas, o que significa que elas são isoláveis, diferentemente do continuum das cores, por exemplo. (FIORIN, 2013, p. 72).*

### **3 Conclusões sobre o aspecto linguagem**

A partir do que pesquisamos e apresentamos de maneira resumida na seção anterior, sobre *Filosofia da Linguagem, Semiótica e Linguística*, é possível derivar algumas conclusões úteis para nosso objetivo específico nesse trabalho: *Recuperação da Informação*.

*Linguagem ou linguagens?* A Linguagem, tomada inicialmente em sentido amplo, é algo que pode ser utilizada por seres tão peculiares como abelhas para comunicar a localização de alimentos ou entre vários outros animais para complementar rituais de acasalamento. No âmbito humano, que é nosso interesse aqui, pode ser utilizada desde as idades mais tenras, e algumas teorias defendem que, pelo menos, parte da capacidade de Linguagem é inata, ou seja, que nos acompanha desde o nascimento. Podemos também criar linguagens: a Filosofia da Linguagem, com a pretensão de maior clareza, objetividade e certeza, estabelece proposições lógicas; a matemática é uma linguagem artificial largamente utilizada; ainda no

campo das linguagens artificiais, desenvolvemos e utilizamos metalinguagens, dentre elas as linguagens documentárias são de particular interesse, pois são parte da tecnologia atual para Recuperação da Informação. Finalmente, temos a linguagem humana natural: as línguas e seus grupos linguísticos, ramos, dialetos e variantes regionais, sobre os quais a Linguística dedica considerável esforço de pesquisa e estudo.

Para mim, no que diz respeito às línguas humanas, e apesar de passado mais de um século, Saussure foi o grande observador dos principais aspectos da língua: o lugar da Linguística em relação à Semiótica, a distinção de uma linguística da língua e da fala, a importância da língua escrita (que veremos com mais detalhes adiante) e a mutabilidade da língua (também dedicamos atenção especial mais adiante), são alguns dos pontos fundamentais abordados de maneira brilhante em seu *Cours*. Saussure atuou nas fronteiras da Filosofia e da Semiótica.

A distinção de Saussure entre língua e fala é particularmente importante como exemplificação do que é Linguagem para ele. Aliás é curioso observar que nem todos os idiomas distinguem de maneira objetiva, tal como nos conceitos biológicos de macho e fêmea, os conceitos de linguagem e língua. Saussure vai além, distinguindo língua e fala (*Langue e parole*) (SAUSSURE, 2012).

Para Saussure, se não fossem feitas distinções adequadas em Linguagem, essa poderia ser estudada por várias áreas tão distintas como a Filosofia, a Antropologia ou a Psicologia (SAUSSURE, 2012, p. 40). Para ele, a língua é *“um produto social da faculdade de linguagem e um conjunto de convenções necessárias, adotadas pelo corpo social para permitir o exercício dessa faculdade nos indivíduos”* (SAUSSURE, 2012, p. 41). A Linguagem, por sua vez, é *“multiforme e heteróclita; o cavaleiro de diferentes domínios, ao mesmo tempo física, fisiológica e psíquica, ela pertence além disso ao domínio individual e ao domínio social”* (SAUSSURE, 2012, p. 41).

E vai além quando define de que *“o exercício da linguagem repousa numa faculdade que nos é dada pela Natureza, ao passo que a língua constitui algo adquirido e convencional”,* ou como também defende, *“não é a linguagem que é natural ao homem, mas a faculdade de constituir uma língua”* (SAUSSURE, 2012, p. 42). Saussure considera o ponto de vista individual em relação aos conceitos de linguagem e língua, naquilo em que o *“indivíduo é sempre senhor”* (SAUSSURE, 2012, p. 45) ele denomina de fala (*parole*). Para ele, essa distinção da fala individual é fundamental, pois separa-se *“o que é social do que é individual”* ou *“o que é essencial do que é acessório e mais ou menos acidental”* (SAUSSURE, 2012, p. 45).

*A Linguagem pode ir além das línguas?* Pelo que já expomos antes, a capacidade humana da Linguagem não se esgota nas línguas naturais, os idiomas falados, ou nas línguas artificiais como a língua de libras (surdos/mudos). Extrapolando os limites da Linguística, que está essencialmente preocupada com diferentes aspectos da língua humana natural, a Semiótica procura compreender o restante do potencial da Linguagem, desde códigos para comunicação (Morse, sinais de fumaça) até mesmo os processos de significação de coisas e seu papel na Linguagem: a significação de uma cruz para os cristãos ou o texto literário como um ícone de racismo ou intolerância são todos exemplos desses processos de significação mais sofisticados.

Não nos interessa aqui o aprofundamento em todos os aspectos da Linguagem, para além das línguas. No entanto, é importante destacar o que talvez já seja óbvio para muitos, a capacidade humana da Linguagem extrapola, em muito, todos os aspectos do uso linguístico propriamente dito. Mais adiante, quando apresentamos nosso Modelo para RI, retornaremos nesse ponto, a importância de entendimento da Linguagem nessa correta dimensão.

*Qual o papel da Língua Escrita na Linguagem?* A Linguagem existe internamente em nós, pois podemos pensar sem materializar externamente o que é pensado. Mas ela também pode materializar-se e de várias formas. Desde sinais gestuais ou faciais em nosso corpo, por meio do uso de instrumentos como tambores, fumaça ou através de uma língua. Mas é necessário reconhecer que o uso de algum tipo de escrita associada a uma língua falada ou não – pois existem línguas ideográficas – é um meio de materialização da Linguagem que se destaca, pelo menos hoje, com seu amplo uso e disseminação em nossa sociedade<sup>17</sup>.

O estudo de qualquer matéria, inclusive o estudo das próprias línguas, não pode prescindir do registro material através da escrita. Mesmo agora, produzindo este texto, necessito do ato de escrever para materializar minhas ideias e muitas vezes até amadurecê-las e corrigi-las.

A importância de registrar através da escrita parece particularmente importante na produção da Literatura. Até porque sem esse registro não chegariam até nós os textos clássicos, desde a antiguidade ou mesmo os textos falados do teatro. A importância da Literatura nas sociedades deve ser a grande responsável pelo status de relevância que a escrita possui entre nós, a ponto de os alfabetizados serem considerados como um grupo social de certa forma “superior”, o que é tanto mais acentuado quanto maior o grau de formação e titulação acadêmica. É fácil confirmar isso, pois um dos indicadores de país desenvolvido são os

---

<sup>17</sup> Os primeiros sistemas de registro escrito datam de 3000 AC (GELB, 1952, p. 61). Portanto, de uma perspectiva de história da humanidade, são bastante recentes.

índices de analfabetismo da população, mesmo a Linguagem estando presente em populações analfabetas da língua oficial, com rica cultura, inclusive com relação à língua falada.

Por tudo isso, a escrita é um aspecto da Linguagem que se destaca socialmente. E quando temos acesso a uma língua estrangeira por meio da escrita, às vezes parece que ela é até tomada pela própria língua falada. Saussure já destacou que a escrita não apenas é independente da língua, mas evolui de maneira independente dessa (SAUSSURE, 2012, p. 59).

*A Linguagem evolui?* Se a Linguagem, considerada aqui como uma *capacidade humana*, muda ou evolui no transcorrer da vida de um indivíduo ou, para a humanidade, ao longo das gerações, não é a pergunta que farei aqui. Muito menos, tentarei respondê-la. Mas está claro que as línguas evoluem, tanto as faladas como as escritas. E é sobre isso que tratamos nessa parte do artigo: evolução das línguas.

Antes de tudo, é preciso esclarecer que há uma distinção entre língua falada e escrita no processo de mudança linguística, e essa distinção e suas características próprias são de suma importância para nossos objetivos, como entenderemos melhor mais adiante.

A percepção científica de que as línguas evoluem está nos primórdios da Linguística moderna. Foi essa área que mais se preocupou e ainda tem se preocupado com a fenômeno da mudança linguística. O termo técnico para esse ramo de estudos é *Linguística Diacrônica*, ou o estudo das línguas considerando etapas diferentes em sua evolução. Quando se consideram os estudos num estágio específico de uma língua, trata-se de estudos *Sincrônicos*. Essa distinção em linguística foi proposta por Saussure no começo do século XX e mantém-se até hoje.

Os estudos sobre esse fenômeno que nos interessam, a *Diacronia*, vão desde a segunda metade do séc. XIX com os neogramáticos, passando pelo estruturalismo e incluindo o gerativismo (SILVA, 2008, p. 28). Desse fenômeno podemos destacar alguns pontos bem significativos: primeiro, a mudança ocorre continuamente e afeta partes e não o todo da língua, mas mantém uma estrutura essencial (FARACO, 2005, p. 14). O falante, em geral, não percebe esse processo de mudança, exceto em casos em que se tem contato com pessoas de gerações muito antigas ou com a língua escrita utilizada no passado (FARACO, 2005, p. 14).

Aliás, o papel da língua escrita no fenômeno diacrônico é da maior importância para nossos objetivos neste trabalho.

A **Língua Falada** não apenas é distinta da língua escrita, como pode-se falar numa estruturação própria para a última, baseada no eixo da Mitografia e da Logografia (DUCROT; TODOROV, 2013, p. 188). A tabela 3 resume os aspectos importantes dessa estruturação.

<b>Mitografia (notação gráfica sem referencia à linguagem verbal)</b>
Aplica-se ao tato (objetos). Ex.: marcas entalhadas em madeira
Aplica-se à visão (figuras). Ex.: desenhos figurativos
<b>Logografia (notação gráfica para registro da linguagem verbal)</b>
Morfemografia (unidades linguísticas significantes). Ex.: morfemas
Fonografia <sup>18</sup> (unidades linguísticas não significantes). Ex.: consoantes

Tabela 3 - Categorias da Língua Escrita (DUCROT; TODOROV, 2013)

O papel da língua escrita pode ser problemático, sob vários pontos de vista, em relação à língua falada. Nas sociedades letradas, a língua escrita tem uma tendência a ser considerada como sinônimo da “língua culta”<sup>19</sup>. Mas a língua escrita não é senão uma “imagem”, na terminologia de Saussure, da língua falada que, obviamente, precede-a em existência.

Os dois tipos de línguas, a falada e a escrita, evoluem. Uma exerce força sobre a outra<sup>20</sup>, às vezes adiando e às vezes motivando mudanças. Mas é a língua falada que realmente muda nos contextos sociais, sendo que a língua escrita pode inserir-se nesse processo registrando tais mudanças. É possível até considerar a mudança de maneira sistemática, “elas costumam se desencadear na fala informal de grupos socioeconômicos intermediários; avançam pela fala informal de grupos mais altos na estrutura socioeconômica; chegam a situações formais de fala e só então começam a ocorrer na escrita.” (FARACO, 2005, p. 26).

*A Linguagem possui significados absolutos?* Essa pergunta origina-se das contribuições na área da Filosofia da Linguagem. A preocupação com os enunciados – principalmente em função do discurso científico e mesmo o próprio discurso filosófico – levou os filósofos dessa área a preocupar-se com a clareza, objetividade e verdade, isto é, os aspectos do significado. Wittgenstein tem um papel importantíssimo em relação ao problema da Linguagem. Parte das contribuições dele foi assim sintetizada por Huisman:

*Na linguagem corrente, acontece com muita frequência que uma mesma palavra designe de maneiras diferentes – pertença, pois, a símbolos diferentes – ou que duas palavras que designam de maneiras diferentes sejam empregadas, na proposição, superficialmente do mesmo modo. Assim, a palavra “é” aparece como cópula, como sinal de igualdade e como expressão da existência; “existir”, como verbo intransitivo, tanto quanto “ir”; “idêntico”, como adjetivo; falamos de algo, mas também de acontecer algo. (HUISMAN, 2000, p. 322).*

<sup>18</sup> A fonografia é de longe o sistema de registro mais utilizado. É esse sistema que empregamos hoje em nossos livros, documentos e mesmo neste artigo.

<sup>19</sup> Faraco vê um controle social mais intenso sobre a língua escrita (FARACO, 2005, p. 25).

<sup>20</sup> Veja-se um exemplo com a orações relativas com preposição (FARACO, 2005, p. 25)

Já no segundo Wittgenstein “*a significação de uma palavra é seu uso na linguagem*”. Daí a ideia de “jogo de linguagem” associado a uma “forma de vida” e a cultura (HUISMAN, 2000, p. 322).

É claro que a obra de Wittgenstein é complexa e ampla. Mas para nossos interesses específicos queremos destacar, muito simplificadamente, sua visão sobre os enunciados nas línguas: nada é absoluto. É preciso considerar especificamente quem diz, quando diz, onde diz e como diz cada enunciado.

#### **4 Implicações em RI e futuros desenvolvimentos**

De tudo que vimos, podemos agora falar sobre o conceito Linguagem com muito mais precisão. Podemos distinguir algumas características importantes para nossos objetivos, quais sejam: o aperfeiçoamento do funcionamento de Sistemas de RI.

A partir dos estudos acima, o recorte que fazemos sobre Linguagem resulta nas seguintes observações:

1. a Linguagem é antes de tudo uma *capacidade humana* que permite uma série de desdobramentos em relação ao pensamento e conhecimento humanos em sua interação com os demais indivíduos e a cultura social;
2. ela tem de ser compreendida em *níveis ou tipos de aplicação* em relação a uma pessoa e o meio social no qual essa mesma pessoa está inserida: linguagem no nível de signos, linguagem humana falada (que preferimos especificar como língua natural), linguagem humana registrada (que preferimos especificar como língua escrita), linguagem humana corporal (gestos, movimentos faciais). Esses níveis não são estanques, podemos, por exemplo, utilizar a linguagem registrada com imagens fixas ou em movimento: fotografias, filmes. Podemos inclusive registrar a linguagem falada. Um documento textual registrado (língua escrita) pode também ser um signo do tipo ícone de documento histórico ou sagrado;
3. pelo menos a língua falada e a língua escrita estão seguramente sujeitas a *processos diacrônicos*. Ou seja, esses níveis de linguagem evoluem e sofrem mudanças ao longo do tempo, e o uso da língua registrada precisa levar em consideração o momento temporal em que houve o registro ou o uso da língua falada;
4. se a Linguagem é uma capacidade humana, as línguas produzidas a partir dessa capacidade devem conter algo de idiossincrático. No caso da língua falada que é apreendida na sociedade, esse conflito entre o que é idiossincrático e social resulta em *individualidade*, o que precisa ser considerado também.

Os limites do formato ora apresentado não permitem o avanço da apresentação de nossos estudos. Mas podemos resumir os desdobramentos a partir do ponto atual. Considerando então a Linguagem compreendida através das quatro observações anteriores, isso resulta nas seguintes indagações para Sistemas de RI:

Primeiro, até que ponto os sistemas atuais de Recuperação da Informação levam em consideração a Linguagem no recorte que apresentamos aqui? Trata-se, portanto, de uma investigação que envolve definir como os sistemas de RI atuais são estruturados para levar em consideração o aspecto Linguagem, se é que o fazem, e quais aspectos são considerados ou não. Segundo, podemos propor e tentar responder perguntas mais específicas.

Sistemas de RI podem atender à necessidade de recuperação de signos? Por exemplo, a demanda de recuperação de documentos registrados textuais que sejam considerados sagrados, independentemente do tipo de conteúdo e informações específicas registradas.

Até que ponto sistemas de RI podem considerar a individualidade no uso da língua falada ou escrita? Uma aplicação direta seria tentar atender necessidades da língua em suas especificidades regionais, culturais e sociais. Frases como essa possuem um significado social muito além de suas palavras individualmente: *Queria dar um rolé nessa biblioteca aí na web e sacar umas coisas sobre hip hop!*

Ainda sobre individualidade, as pessoas assumem papéis em sistemas de RI e esses podem resultar em usos específicos da língua. É possível considerar que uma pessoa assuma pelo menos o papel de produtor de documentos e informações, pesquisador de documentos e informações e gestor (indexador, classificador, catalogador) de documentos e informações.

Até que ponto sistemas de RI podem considerar o efeito diacrônico da língua falada e escrita? Isso implica em tentar equacionar problemas como a língua escrita utilizada por alguém no papel de indexador no século passado ainda é aceitável para uma pessoa na atualidade no papel de pesquisador?

Retomando o que expusemos na introdução deste artigo, nossa preocupação inicial refere-se ao problema da Preservação Digital e ao uso de Sistemas de RI. Como se trata de um problema a ser tratado em longo prazo, a diacronia nos é particularmente cara. Mas atualmente, consideramos que não levar em conta todos os aspectos importantes da Linguagem, além da questão da diacronia nas línguas, também pode ter implicações nefastas para a recuperação de Documentos Digitais em longos períodos de tempo.

## Referencias

- ABBAGNANO, Nicola. **Dicionário de filosofia**. 5. ed. São Paulo: Martins Fontes, 2007.
- AUDI, Robert. **Dicionário de filosofia**. São Paulo: Paulus, 2006.
- BILETZKI, Anat; MATAR, Anat. Ludwig Wittgenstein. **Stanford Encyclopedia of Philosophy**. Disponível em: <http://plato.stanford.edu/archives/spr2014/entries/wittgenstein/> Acesso em 04 de novembro de 2014.
- BODÊ, Ernesto. **Preservação Digital: o papel dos formatos de arquivo**. Dissertação (Mestrado). Brasília: Faculdade de Ciência da Informação, Universidade de Brasília, 2008.
- BORGHOFF, Uwe M.; RODIG, Peter; SCHEFFCZYK, Jan; SCHMITZ, Lothar. **Long-Term preservation of digital documents: principles and practices**. Berlin: Springer, 2006.
- CHILD, William. **Wittgenstein**. Porto Alegre: Penso, 2013.
- DUCROT, Oswaldo; TODOROV, Tzvetan. **Dicionário enciclopédico das ciências da linguagem**. São Paulo: Perspectiva, 2010.
- ECO, Umberto. **Tratado Geral de Semiótica**. 4. ed. São Paulo: Perspectiva, 2012.
- FARACO, Carlos Alberto. **Linguística histórica: uma introdução ao estudo da história das línguas**. São Paulo: Parábola, 2005.
- FIORIN, José Luiz (Org.). **Linguística: o que é isso?** São Paulo: Contexto, 2013.
- GELB, Ignace Jay. **A study of writing**. 2. ed. Chicago, London: University of Chicago Press, 1952.
- GREIMAS, Algirdas Julien; COURTÉS, Joseph. 2. ed. **Dicionário de Semiótica**. São Paulo: Contexto, 2012.
- HUISMAN, Denis. **Dicionário de obras filosóficas**. São Paulo: Martins Fontes, 2000.
- HUNNEX, Milton. **Filósofos e correntes filosóficas em gráficos e diagramas: conheça melhor os filósofos e as correntes filosóficas por meio de gráficos e diagramas temáticos e cronológicos**. São Paulo: Vida, 2003.
- MACKAY, David J. C. **Information theory, inference, and learning algorithms**. Cambridge: Cambridge University Press, 2003. Disponível em: <http://www.inference.phy.cam.ac.uk/itprnn/book.pdf> . Acesso em: 04 de novembro de 2014.
- MARCONDES, Danilo. **Textos básicos de linguagem: de Platão a Foucault**. Rio de Janeiro: Zahar, 2010.
- MAUTNER, Thomas. **Dicionário de filosofia**. Lisboa: Edições 70, 2011.
- MILLER, Alexander. **Filosofia da linguagem**. 2. ed. São Paulo: Paulus, 2010.
- PEIRCE, Charles Sanders. **Semiótica**. 4. ed. São Paulo: Perspectiva, 2012.
- SAUSSURE, Ferdinand de. 28. ed. **Curso de linguística geral**. São Paulo: Cultrix, 2012.

SILVA, Rosa Virgínia Mattos e. **Caminhos da linguística histórica: ouvir o inaudível**. São Paulo: Parábola, 2008.

**Recebido/Recibido/Received:** 2014-11-08.

**Aceitado/Aceptado/Accepted:** 2015-07-15.