

## XXV ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO - XXV ENANCIB

### GT 2 – Organização e Representação do Conhecimento

#### A LINGUÍSTICA SISTÊMICO FUNCIONAL NA CONSTRUÇÃO DE TESAUROS

##### *SYSTEMIC FUNCTIONAL LINGUISTICS IN THESAURUS CONSTRUCTION*

**José Mario de Oliveira Mendes** – Fundação Escola de Sociologia e Política de São Paulo (FESPSP)

**Vânia Mara Alves Lima** – Escola de Comunicações e Artes da Universidade de São Paulo (ECA-USP)

#### **Modalidade: Trabalho Completo**

**Resumo:** partindo do pressuposto que a linguagem é uma faculdade humana e a língua uma característica identitária, variável e mutável conforme experiências e conhecimentos vividos, esta pesquisa parte da hipótese de que há convergência teórico-metodológica entre a Linguística Sistêmico-Funcional (LSF) e os Sistemas de Organização do Conhecimento (SOCs), a partir da compreensão da língua como recurso de criação de significados. Trata-se de pesquisa bibliográfica e documental, de caráter exploratório que aborda as inter-relações entre a Linguística e a Organização do Conhecimento (OC) no âmbito da Ciência da Informação (CI). Identifica, na LSF, os fundamentos teóricos que contribuem para o aprimoramento da elaboração de SOCs, mais especificamente os Tesauros. Concluí que a integração entre a LSF e a OC possibilita não apenas abordagens descritivas, mas também um melhor entendimento do conceito e do termo, dos remas e dos temas, e da organização dos conceitos por meio de categorias, com o uso de definições na construção de SOCs como o Tesouro.

**Palavras-chave:** sistemas de organização do conhecimento; tesouro; linguística sistêmico - funcional.

**Abstract:** assuming that language is a human faculty and that linguistic expression is an identity-based, variable, and mutable feature shaped by lived experiences and acquired knowledge, this study is based on the hypothesis that there is a theoretical and methodological convergence between Systemic Functional Linguistics (SFL) and Knowledge Organization Systems (KOS), grounded in the understanding of language as a resource for meaning-making. This is a bibliographic and documentary research of an exploratory nature, addressing the interrelations between Linguistics and Knowledge Organization (KO) within the field of Information Science (IS). It identifies, within SFL, theoretical foundations that contribute to enhancing the development of Knowledge Organization Systems, particularly thesauri. The study concludes that the integration of SFL and KO enables not only descriptive approaches but also a deeper understanding of concepts and terms, rhemes and themes, and the organization of concepts through categories, with the use of definitions in the construction of KOS such as the Thesaurus.

**Keywords:** knowledge organization systems; thesaurus; systemic functional linguistics.

## 1 INTRODUÇÃO

A linguagem enquanto fenômeno social nos permite interagir com o mundo e com as pessoas. Aprendemos, construímos valores, identidades e representamos nossas experiências através da linguagem, construindo conhecimentos que podem ser transmitidos oralmente, por escritos, imagens e outros meios. A partir de invenção da prensa tipográfica por Johann Gutenberg, em 1450, e a Revolução Industrial, essas experiências, e principalmente, o conhecimento produzido, passam a ser registrados, compartilhados e disseminados em larga escala, o que exige a sua organização. Essa organização passa pela Enciclopédia de Diderot (2017, [1750]) pelos sistemas de classificação bibliográficos e chega aos modernos Sistemas de Organização do Conhecimento (SOCs), como os Tesouros. A linguagem é fundamental para organizar, representar e recuperar esse conhecimento, produzido por toda a sociedade ou comunidades específicas. A linguagem é uma faculdade humana universal, enquanto a língua é particular e variável, moldada pela visão de mundo, história, cultura e realidade social de seus falantes, permitindo a interação e apropriação do conhecimento ao longo do tempo. Estudada pela Linguística, a língua foi compreendida a partir do século XX, com Saussure (1972), como um sistema analisável sincronicamente, estruturado por seus elementos. Assim, língua e linguagem podem ser vistas como unidades básicas em sistemas de organização do conhecimento, pois os conceitos classificam experiências individuais que se relacionam formando redes de informação. Contudo, conceitos não são entidades físicas; para serem utilizados e compartilhados, precisam ser organizados, recuperados e expressos de alguma forma. Gnoli (2020, p. 45) diz que:

a forma mais comum de expressar conceitos é por palavras: outros símbolos, como gestos, ícones ou qualquer código compartilhado, também podem funcionar. Em certo sentido, então, as palavras na linguagem natural são uma forma básica de organização do conhecimento: quando dizemos 'aquele gato', estamos classificando o fenômeno que observamos como uma instância da classe dos gatos [e] porque muitos documentos estão na forma de textos escritos, suas palavras e frases podem ser um ponto de partida para a organização do conhecimento.

No campo dos estudos linguísticos, a Linguística Sistêmico-Funcional (LSF), proposta por Michael Halliday no século XX, aborda a conexão entre sociedade e conhecimento mediado pela linguagem, inseridos em contextos culturais e situacionais, a partir de convenções e transmissão de significados definidos socialmente. No âmbito da Ciência da Informação, os Sistemas de Organização do Conhecimento (SOCs) têm por objetivo facilitar o

gerenciamento e a recuperação desse conhecimento em determinado domínio (Mazzochi, 2018). A partir das bases teóricas da Organização do Conhecimento (OC), como a Linguística Documentária (Lara; Tálamo, 2007), a Teoria do Conceito (Dahlberg, 1978) e a Terminologia (Cabré, 1999), esta pesquisa busca identificar na LSF (Halliday, 1985, 1994; Halliday; Matthiessen, 2004) conceitos e fundamentos que contribuam para o aprimoramento de SOCs do tipo Tesouro. Metodologicamente, trata-se de estudo exploratório bibliográfico e documental, voltado a oferecer visão geral sobre o tema (Gil, 2002). As fontes são registros disponíveis em livros, artigos, teses etc., já explorados por outros autores e que se tornam base para o trabalho do pesquisador (Severino, 2007). Quanto ao objetivo, é pesquisa exploratória, buscando levantar informações, delimitar o campo e mapear as condições de manifestação do objeto (Severino, 2007). A pesquisa bibliográfica foi realizada nas bases Dedalus (USP), Biblioteca Digital Brasileira de Teses e Dissertações (IBICT), Repositórios do ENANCIB, Anais da ISKO Internacional e Brasil, Catálogo da Biblioteca da Fundação Escola de Sociologia e Política de São Paulo, Repositório do Grupo Sistêmica, Ambientes e Linguagens, CAPES, SciELO, BRAPCI, LISA e ISTA. Foram usados termos de busca em português, inglês e espanhol relacionados a “informação e linguagem”, “informação”, “linguagem”, “linguística”, “linguística sistêmico-funcional”, “linguística hallidayana”, “linguística documentária”, “linguagem documentária”, “análise documentária”, “tesauros”, “organização da informação”, “organização do conhecimento”, “sistemas de organização da informação”, “sistemas de organização do conhecimento”. O levantamento resultou em 386 documentos, dos quais, após leitura objetiva (Cintra, 1987) e descarte dos que não atendiam aos objetivos, foram selecionados 105 com informações mais relevantes à pesquisa.

## **2 DESENVOLVIMENTO**

A Linguística Documentária (LD) é um campo da Ciência da Informação que estuda a linguagem como instrumento essencial para acessar, organizar e recuperar o conhecimento. Surgida na Espanha nos anos 1990 com García-Gutiérrez, parte do princípio de que os problemas informacionais são, sobretudo, problemas de linguagem. Difere das abordagens técnicas ao incorporar aspectos simbólicos, semióticos e pragmáticos nos processos de representação e mediação da informação. Inspirada na tradição da Documentação, especialmente nas ideias de Paul Otlet sobre organização e padronização do conhecimento, preocupa-se com a construção de linguagens documentárias, metalinguagens baseadas na

linguagem natural, voltadas à precisão, inteligibilidade e redução da ambiguidade. Nesse contexto, “a documentação ofereceu à linguística documentária uma estrutura cíclica que refletiu de forma direta na Análise Documentária” (Lima; Moreira; Moraes, 2016, p. 993). Para Santos, Maimone e Lima (2019), a LD investiga as estruturas simbólicas da documentação, os problemas linguísticos da mediação informacional, os meios e linguagens de representação e a produção da informação documentária. Visa minimizar ambiguidades e maximizar precisão e pertinência por meio da padronização e delimitação de sentidos, articulando o código da língua natural (baseado em saberes prévios), o subcódigo do domínio de conhecimento (que confere sentidos especializados) e o subcódigo da Ciência da Informação (que define recortes conceituais nos sistemas de organização). Assim, a LD atua como sistema simbólico e operativo essencial à circulação do conhecimento na cultura informacional contemporânea. Ao integrar pressupostos linguísticos, terminológicos e epistemológicos, conecta-se à Terminologia, à Teoria do Conceito e à Linguística Aplicada, consolidando-se como área estratégica para entender a mediação e estruturação do saber. Com o desenvolvimento da linguagem humana e a necessidade de comunicar e nomear o mundo, torna-se relevante refletir sobre a evolução da linguagem natural e das linguagens especializadas em contextos como química, física, engenharia, letras e sistemas de classificação (Dahlberg, 1978, p. 101).

O conhecimento fixou-se através dos elementos da linguagem. Novos conhecimentos apareceram com novos elementos linguísticos e também através destes tornaram-se mais claros e distintos. Podemos dizer que este processo de crescimento há de perdurar enquanto o homem existir sobre a terra e utilizar a linguagem como expressão de seus pensamentos.

É por meio da linguagem que o ser humano se relaciona com o mundo, elabora enunciados e experiencia o tempo e o espaço, formulando conceitos individuais e gerais. “É em base a tais enunciados que elaboramos os conceitos relativos aos diversos objetos. Cada enunciado verdadeiro representa um elemento do conceito” (Dahlberg, 1978, p. 102). Um conceito é constituído pela soma de características e enunciados verdadeiros sobre um objeto. Assim, ele é uma unidade estruturada composta por enunciados válidos sobre determinado objeto. A Teoria do Conceito considera os elementos conceituais (identificados com atributos dos objetos), a análise e síntese das características, a tipologia (simples ou complexas), a ordenação (essenciais ou acidentais, gerais ou individualizantes) e suas funções (classificação, definição e nomeação dos conceitos). Além disso, contempla as relações entre conceitos: lógicas (identidade, implicação, interseção, disjunção e negação), hierárquicas

(superior/inferior), partitivas (todo/parte), de oposição (contradição/contrariedade) e funcionais (relativas ao uso semântico da linguagem). Dahlberg (1978) também trata da intensão (conjunto de características) e da extensão do conceito (conjunto de conceitos mais específicos que compartilham essas características), bem como das espécies de conceitos, importantes para a construção e combinação de sistemas. As definições conceituais são essenciais à argumentação e comunicação, sendo fundamentais para a construção de sistemas científicos e de organização do conhecimento. Como afirma a autora, essas relações conceituais permitem comparações e organização em classificações e tesouros (Dahlberg, 1978, p. 104). Assim, o conceito precisa ser designado por um nome para integrar o léxico da linguagem natural. Do mesmo modo, na Análise Documentária, o conceito deve ser representado por um termo, o que nos conduz à discussão sobre Terminologia. Tudo o que existe é denominado. Seja por meio de desenho, fala ou escrita, os entes e coisas recebem nomes, materializados como “uma unidade da língua escrita, situada entre dois espaços em branco, ou entre um espaço em branco e um sinal de pontuação” (Houaiss, 2021). Segundo Barros (2004, p. 25), o avanço da ciência e as mudanças no modo de vida impulsionam transformações sociais, políticas e linguísticas, pois “cada descoberta ou invento recebe um nome [e] passa a ser designado por um termo”, gerando um “desenvolvimento terminológico tão importante quanto o econômico ou social”, com intensa criação neológica por diferentes mecanismos linguísticos. Desde a Revolução Industrial até a explosão informacional do pós-Segunda Guerra Mundial, a sociedade tem passado por profundas transformações, gerando novas relações, leis e tecnologias, mudanças que impactam o vocabulário; “a cada nova invenção, [...] surgiram [e continuam surgindo] novos termos correspondentes” (Barros, 2004, p. 26). O léxico e a terminologia se expandem continuamente. Para a autora, essas transformações geram novas demandas linguísticas e promovem o uso de uma língua padrão, com poder simbólico e social, reduzindo variações e reforçando noções de “valor, marca sociolinguística, certo e errado” (Barros, 2004, p. 27), tornando-se instrumento de discriminação e classificação social. Corbeil (1984, *apud* Barros, 2004, p. 27) observa que “os vocabulários de especialidade são produzidos pelos próprios especialistas” e difundidos por publicações, representando o conhecimento especializado em contraste com a língua natural dos leigos. Desde os tijolos de argila sumérios, de 2600 a.C., com termos sobre profissões, gado e divindades, passando por dicionários temáticos egípcios e glossários médicos antigos até os bancos de dados atuais, percebe-se a contínua necessidade de identificar, compilar e

representar vocabulários especializados, tema cada vez mais central para áreas como a Terminologia e a Linguística. Para Lara (2004, p. 234-235),

A Terminologia é uma área interdisciplinar que dá suporte a várias disciplinas no estudo dos conceitos e sua representação em linguagens de especialidade. Termo polissêmico, terminologia se refere tanto à teórica e metodológica, como à terminologia concreta. Na primeira acepção, a Terminologia fornece metodologia para a descrição, ordenamento e transferência do conhecimento, indicando princípios que regem a compilação, formação dos termos, estruturas de campos conceituais, usos e administração de terminologias. Na segunda acepção, refere-se a um conjunto de termos relacionados a uma língua de especialidades. (...) [Seus elementos centrais (...) são os objetos, conceitos, termos e definições

Conforme Lara e Tálamo (2006), a Linguística e a Semiótica, bases da Linguística Documentária, não resolvem a delimitação de subconjuntos linguísticos temáticos e funcionais por partirem de premissas especulativas. A Terminologia, por outro lado, contempla esses aspectos de forma teórico-metodológica e material, com duas correntes principais: a Teoria Geral da Terminologia (TGT) e a Teoria Comunicativa da Terminologia (TCT). A TGT, fundada por Eugen Wüster (1998), opõe-se à Linguística de Saussure e considera possível descrever conceitos de forma independente da língua. Para Wüster (1988), um conceito pode ser definido sem precisar identificar os termos que o nomeiam, havendo independência entre expressão e conteúdo (Barros, 2004). Ele propunha a univocidade entre termo e conceito, buscando eliminar ambiguidades e ruídos na comunicação entre especialistas. Defendia a normalização terminológica e elaborou um dicionário multilíngue de máquinas-ferramenta, organizado sistematicamente, publicado com apoio da ONU. Wüster (1998) rejeitava sinônimos, homônimos e termos polissêmicos e, quando não havia designação única aceitável, a Terminologia normativa poderia criá-la. A TCT, mais recente, proposta por Maria Teresa Cabré, critica o reducionismo da TGT e considera essencial a variação linguística, pois sem ela não há comunicação nem disseminação de conhecimento especializado. Os termos, segundo a TCT, são unidades linguísticas que expressam conceitos, mas permanecem signos da linguagem natural, com características linguísticas, cognitivas e sociais. A teoria considera ainda a variação conceitual e denominativa dos termos conforme o contexto sociocultural e situacional, bem como suas funções discursivas e textuais. Para Cabré (1999), o valor de um termo depende do lugar que ocupa no sistema conceitual e da forma como é utilizado nos diferentes domínios (Maimone, 2019). Como objetivos, a Teoria Comunicativa da Terminologia pretende, de acordo com Cabré (1999, p. 22)

descrever formal, semântica e funcionalmente as unidades que podem adquirir valor terminológico, dar conta de como são ativados e explicar suas relações com outros tipos de signos do mesmo ou distinto sistema, para fazer progredir o conhecimento sobre a comunicação especializada e as unidades que nela se usam.

Assim, a teoria da terminologia migra de um paradigma prescritivo com a TGT para uma abordagem descritiva com a TCT, na qual “os problemas de identificação dos termos” não se limitam mais à frase ou a recortes paradigmáticos, mas se explicam pelos fenômenos da textualidade e discursividade, inseridos em um processo comunicativo (Lima, 2004, p. 109-110). Nesse contexto, insere-se a Linguística Sistêmico-Funcional (LSF), de Michael Alexander Kirkwood Halliday, que concebe a linguagem como um recurso semiótico voltado à criação de significados em contextos sociais e culturais. Diferente das abordagens tradicionais, a LSF entende a linguagem como funcional e contextualizada, resultante de escolhas feitas por usuários em situações específicas. Considerada *applied* por Halliday, a teoria tem aplicações em áreas como ensino, análise de discurso, tradução e construção de sistemas de organização de conhecimento. Autores como Fuzer e Cabral (2010), Eggins (2004), e Martin e Rose (2008) destacam sua relevância para entender a variação e funcionamento da linguagem em diferentes contextos e culturas. Para a LSF, a análise gramatical e discursiva mostra como o texto transmite significado e sua eficácia comunicativa (Halliday, 1994, p. 15). A linguagem é vista como prática social motivada por uma finalidade. A LSF estuda como os indivíduos produzem linguagem em contexto, considerando interação, comunicação e função social. Uma premissa central da LSF é a concepção da linguagem como sistema sócio-semiótico, em que os significados são compartilhados em contexto. Halliday (1994) aponta três fundamentos: a linguagem é funcional; suas funções criam significados; e o uso da linguagem envolve escolhas feitas pelos falantes. Para Halliday (1994, p. xiii-v)

Como uma teoria social que se preocupa com os usos da língua, a LSF reconhece que a linguagem é entidade viva, presente em situações, grupos, locais, eventos variados e, como tal, sofre a influência desses e outros fatores. Como propriedade de comunidades, culturas e indivíduos, a língua é variável, um potencial de significados à disposição dos falantes que dela fazem uso para estabelecer relações, representar o mundo e, com isso, satisfazer determinadas necessidades em contextos sociais específicos.

Ainda, de acordo com o autor, outro importante fundamento da LSF é denominado princípio das escolhas, as quais são realizadas dentro de uma gama de opções em cada contexto, sendo realizadas na construção dos textos falados ou escritos em qualquer situação

de uso da linguagem. Portanto, podemos enxergar e entender a linguagem como um recurso que é utilizado para se criar significados. Deste modo, quando um usuário ou falante de uma língua produz determinado texto, há a realização dos três tipos de significados ao mesmo tempo: significados que se relacionam à representação da experiência através da língua; significados que expressam as relações de poder, de solidariedade e de atitude em relação aos outros e; significados que se relacionam com o modo de organização do conteúdo da mensagem, com o modo que se diz o que foi dito. Nomeiam-se e classificam-se as coisas e, também, relatam-se as experiências. Por outro lado, pela linguagem, estabelecem-se as relações sociais entre os indivíduos e a manifestação da experiência, com a qual os usuários informam, conversam, indagam, ordenam, oferecem ou opinam sobre pessoas, eventos ou fatos. Ainda, através da linguagem, se organizam os ditos, criando-se formas em concordância com os significados a serem realizados por meio da organização e coesão das orações em termos de informações dadas e novas. Halliday propõe que toda linguagem realiza três *metafunções*: a **ideacional** (representar o mundo), a **interpessoal** (estabelecer relações sociais) e a **textual** (organizar o discurso), conforme o quadro 1.

Quadro 1 - Metafunções, Tipos de Significado e Papel na Oração

Metafunção	Tipos de Significado Expresso	O Papel na Oração
Ideacional	Constrói a experiência.	Oração como representação.
Interpessoal	Atua nas relações sociais.	Oração como troca.
Textual	Favorece a relevância no contexto.	Oração como mensagem.

Fonte: Halliday (1994, p. 36, tradução nossa)

Neste trabalho, focamos na Metafunção Textual, que se realiza pela estrutura temática que dá à sentença seu caráter de mensagem (Halliday, 1994, p. 37), que usa a linguagem para organizar nossos significados experienciais e interpessoais de uma maneira coerente, coesa e linear. A realização léxico-gramatical dessa Metafunção se dá através do sistema Tema-Rema, cuja preocupação é a organização da oração como mensagem. As relações entre sentenças, a coesão entre o texto e seu propósito e a sequência lógica de pensamento são necessárias para que o usuário da língua possa expressar-se e ser entendido num momento de fala ou escrita. As diferentes escolhas do Tema na oração, ou seja, as diferentes escolhas do ponto de partida da mensagem, a maneira que o falante ou escritor inicia seu momento de fala ou escrita, é fator determinante na produção de sentido de fala. A escolha desses elementos iniciais de cada oração funcionará como sinal para o ouvinte, no caso de fala ou para o leitor, no caso de texto escrito, chamando sua atenção para o que virá a seguir, e para todo o significado textual

da sentença. De acordo com Halliday (1994, p. 37), presume-se que em todas as línguas a oração tenha a característica de uma mensagem, pois possui uma maneira de organização que lhe confere o status de evento comunicativo. Na Língua Portuguesa, bem como em muitas outras línguas neolatinas e germânicas, uma parte da oração recebe um status diferenciado do restante da sentença, e, ambos em conjunto, darão à sentença o status de mensagem, configurando-se por meio de elementos textuais, classificados do ponto de vista funcional como Tema e Rema. Tema é o elemento que funciona como ponto de partida, aquele que determina sobre qual será o assunto da oração (Eggins, 2004; Halliday, 1994; Thompson, 1996). É o ponto de partida da mensagem, a primeira parte da informação, realizada até o primeiro elemento da Transitividade (participante, processo ou circunstância). É classificado como Tema não-marcado quando é realizado pelo elemento na posição de Sujeito da oração, podendo ser um grupo nominal, um pronome, ou uma nominalização. É classificado como Tema Marcado quando realizado por um complemento ou um adjunto. Todo o restante da oração é classificado como Rema (Halliday; Matthiessen, 2004). Ainda conforme Halliday e Matthiessen (2004), o Tema pode trazer uma informação dada ou uma informação nova, dependendo do contexto e do que é dito anteriormente no texto. Essa unidade de informação permite que o escritor ou falante enfatize determinadas partes da oração e do texto, chamando a atenção do leitor para o que ele considera relevante. O Rema é a parte da oração na qual o tema será desenvolvido, ou seja, tudo o que não é tema, é rema (Eggins, 2004, p. 275) (quadros 02 e 03).

**Quadro 2** – Exemplo de relação Tema e Rema (1)

A teoria	discute a solução do problema via duas abordagens distintas.
Tema	Rema

**Fonte:** Mendes (2013, p. 17).

**Quadro 3** – Exemplo de relação Tema e Rema (2)

Diferentes autores	reportam evidências de uma relação negativa entre volumes negociados e retorno de ativos.
Tema	Rema

**Fonte:** Mendes (2013, p. 17).

Os termos “a teoria” e “diferentes autores” funcionam como os Temas das orações, apresentando a nova informação que é retomada em “discute a solução do problema via duas abordagens distintas” e em “reportam evidências de uma relação negativa entre volumes negociados e retorno de ativos”, os Remas, nos quais a informação nova é desenvolvida. Neste

ponto, nos parece realmente que a Metafunção Textual é o grande ponto de convergência e de interconexão entre a Linguística Sistêmica-Funcional e os Sistemas de Organização do Conhecimento. Para Martin e Rose (2008), a periodicidade diz respeito ao fluxo da informação e à forma como os significados são organizados para facilitar a compreensão do texto. Essa organização ocorre por meio da articulação de Temas e Remas nas orações, e de hiperTemas (sentenças-tópico) e hiperNovos no nível dos parágrafos. O hiperTema antecipa o conteúdo a ser desenvolvido no parágrafo, enquanto o hiperNovo sintetiza as informações mais relevantes, funcionando como um resumo. A análise de um texto via periodicidade permite compreender sua construção parágrafo a parágrafo, observando como o Tema introduz a ideia principal e como as orações subsequentes a desenvolvem. Os Temas, portanto, originam os hiperTemas e macroTemas, que operam em níveis mais amplos do texto. De forma complementar, os Novos ou Remas, que veiculam a informação nova, podem ser ampliados aos hiperNovos e macroNovos, expressando o conteúdo central a ser transmitido. Halliday e Matthiessen (2004) explicam que, em línguas como o inglês e o português, a oração é estruturada como uma mensagem dividida entre Tema e Rema. O Tema é o ponto de partida da oração, escolhido pelo falante ou escritor, enquanto o Rema representa a continuação da mensagem. Por exemplo: na oração “Os primeiros estudos sobre Sistemas de Organização do Conhecimento surgiram no século passado, nos Estados Unidos”, o Tema é “Os primeiros estudos sobre Sistemas de Organização do Conhecimento” e o Rema é “surgiram no século passado, nos Estados Unidos”. Além disso, há a distinção entre Informação Dada (conhecida pelo leitor ou ouvinte) e Informação Nova (elemento novo e relevante apresentado pelo emissor). A ordem Tema-Rema é fundamental: o Tema vem primeiro por ser o ponto de partida da mensagem; o Rema, por sua vez, desenvolve essa ideia. Assim, compreende-se como a linguagem organiza e transmite significado em um texto, conforme também apontado por Gouveia (2009, p. 38).

A definição [de Tema] é funcional, tal como acontece com todos os elementos na interpretação da estrutura gramatical. Ou seja, o tema é um elemento numa configuração particular, que, tomado como um todo, organiza a oração como mensagem, isto é, como a configuração Tema + Rema. [...] Em tal configuração, o Tema é o ponto de partida para a mensagem; é aquilo sobre o que a mensagem vai ser.

Reiterando que essa definição de Tema é funcional, e não morfossintática, pois se relaciona com a Progressão Temática, com o desenvolvimento do texto, com os constituintes

lexicais que se imbricam ao longo do texto, saindo do âmbito único e exclusivo da oração, engloba-se todo esse texto, uma vez que a cada oração temos um Tema, a cada parágrafo um conjunto de Temas e em, cada texto, uma coleção ou organização de Temas sobre aquele determinado assunto. E quando pensamos numa coleção de Temas sobre um determinado assunto, estamos mencionando também uma Organização de Temas que representam um assunto ou conhecimento específico no interior daquele texto. O conjunto desses Temas é que me dá a Informação Nova (ou Novo) do texto. É o que, inferimos, a partir daqui, o que Dahlberg (1978) chama de Conceito, e que na Terminologia vai ser designado por um Termo.

O significado destes padrões [Termos, grifos nossos] emerge quando nós viemos a considerar a importância do Tema da oração no desenvolvimento global de um texto [e na sua revocação, grifos nossos] (...) a escolha dos Temas das orações desempenha uma parte fundamental no modo como o discurso é organizado; isso, de fato, constitui o que tem sido chamado de “método de desenvolvimento” de um texto [...] (Halliday; Matthiessen, 2004, p. 99).

O mapeamento dos Temas nos parágrafos revela o conteúdo do texto e auxilia na indexação, classificação e organização do conhecimento para uma recuperação eficaz. Como afirma Fries (1983), “a informação contida nos Temas das orações de um parágrafo cria o método de desenvolvimento daquele parágrafo” e, assim, de todo o texto. Essa informação nos Temas é fundamental para a Progressão Temática, que orienta a organização do conhecimento, possibilitando a extração de conceitos e seus termos, essenciais para Sistemas de Organização do Conhecimento (SOCs) como Tesouros. Podemos analisar Tema e Rema em níveis variados: oração, parágrafo ou texto completo. A estrutura temática é função da oração, e mapear Temas e Remas em um parágrafo traz informações dadas e novas. Em textos maiores, como artigos, o Sistema de Periodicidade (Martin; Rose, 2008) identifica hiperTemas, sentenças-tópico que antecipam o assunto geral e guiam o texto, funcionando como Tema nas orações. No parágrafo, a Progressão Temática revela o desenvolvimento, e os Remas ou Novos apontam as informações que o texto transmite e que o leitor deve assimilar. Em textos simples, os Temas e Remas podem indicar “resumo”, “palavras-chave” ou termos específicos; em textos complexos, o Sistema de Periodicidade destaca hiperTemas e hiperNovos para os parágrafos, e macroTemas e macroNovos para o texto completo, sintetizando o conteúdo e as informações mais importantes, conforme destaca Gnoli (2020, p. 68):

[...] na maioria das vezes, assumimos que o recurso de conhecimento trata de apenas um assunto específico. No entanto, em muitos casos, isso não é verdade. De fato, vários 'temas' podem estar presentes em uma única e

mesma fonte. Isso é cada vez mais provável à medida que passamos de uma fonte simples, como uma mensagem curta, para fontes maiores e mais complexas. Um trabalho acadêmico geralmente se concentra em um tema específico e facetado, como 'digestão no estômago de insetos', com possíveis conexões com outros temas. Em uma monografia há espaço suficiente para tratar vários temas e para compará-los e conectá-los, embora geralmente se refira a um argumento principal que é dado na introdução e na conclusão.

Ainda, segundo o autor (Gnoli, 2020, p. 67), “a identificação e análise de temas em um documento [...] ainda é pouco desenvolvida, autores como Alberto Cheti introduziram a linguística textual na teoria [da organização do conhecimento] de indexação de assuntos” pensando em “um tema baseado que é o foco de seu argumento, e possivelmente uma série de temas particulares”. Entretanto, deve-se dar prioridade para o tema base, o que faz mais sentido para que as informações sejam agrupadas. Classificações tradicionais vêm sendo desenvolvidas com o objetivo principal de “organizar os livros em uma prateleira, de modo que um tema tenha que ser escolhido para selecionar a prateleira apropriada” (Gnoli, 2020, p. 67). Tal relação e importância de se pensar na periodicidade e identificação de Temas e Remas é descrita, inclusive, na Declaração dos Princípios Internacionais de Catalogação (Bertolini *et al.*, 2016, p. 18) como “qualquer entidade usada como assunto de uma obra. O *thema* pode variar substancialmente em complexidade. Um *thema* simples pode aparecer combinado ou agregado, resultando em *themas* mais complexos”. Para as autoras (*ibidem*, 2016, p. 7)

Thema [Tema, grifos nossos] [é] (qualquer entidade utilizada como assunto de uma obra) e Nomen [Rema, grifos nossos] [é] (qualquer signo ou sequência de signos pelos quais se conhece, identifica ou se faz referência a um Thema) são as entidades introduzidas e descritas no modelo FRASAD. Dentro do marco FRBR, Thema inclui as entidades do Grupo 1, Grupo 2 e, adicionalmente, todas os demais grupos que podem servir como assunto em uma obra (conceito, objeto, acontecimento e lugar do Grupo 3). No FRASAD Nomen é uma superclasse das entidades do modelo FRAD: nome, identificador e ponto de acesso controlado. Porém, nesta Declaração, Nomen não é considerada uma superclasse de nome, identificador e ponto de acesso controlado.

### 3 CONSIDERAÇÕES FINAIS

Percebemos que o princípio das escolhas da LSF também é utilizado na OC no processo de indexação e recuperação, já que nestes processos se escolhe entre uma ou outra realização daquela linguagem para representá-la, indexá-la e recuperá-la. Pontos de convergência entre o Sistema Tema e Rema da LSF com a Organização do Conhecimento podem também ser percebidos nos

Princípios Internacionais de Catalogação, quando são tratadas Entidades, Atributos e Relações. As Entidades são os objetos chave de interesse para os usuários em um domínio particular, descritas por suas características primárias, chamadas de Atributos. Esses Atributos da Entidade servem também como um meio pelo qual os usuários elaboram consultas e interpretam as respostas quando buscam informações sobre uma entidade particular, com as Relações explicando as conexões entre as Entidades. Quando se diz que a catalogação deve considerar as Entidades, Atributos e Relações definidas nos modelos conceituais do universo bibliográfico diz-se que devem ser considerados (além da Obra, Expressão, Manifestação, Exemplar, Pessoa, Família, Entidade coletiva) o Thema e o Nomen, ou seja, é importante prestar atenção nas entidades utilizadas como assunto de uma obra e nos signos ou sequências de signos pelos quais se conhece, identifica ou se faz referência a um tema. Quando tratamos da linha evolutiva dos tesauros e de possíveis contribuições de outras teorias, podemos traçar mais um paralelo e mais um ponto de atuação da LSF. Quando Moreira (2003) menciona que “já a linha européia, principalmente a linha do *Thesaurfacet* lança mão do uso de categorias para organização dos conceitos em um domínio, o uso dessas categorias para o enquadramento de conceitos permite uma melhor organização das hierarquias e um posicionamento mais adequado dos termos associados aos conceitos. Ainda, reconhecendo-se a estrutura temática do texto, pode-se localizar o Rema por meio do Tema, mas também é possível localizar o Tema por meio do Rema. Organizando-se um SOC tipo Tesouro de modo que seja possível recuperar facilmente um Tema, por consequência seu Rema será localizado e, assim, vice-versa; localizando-se um pedaço, uma fatia do Rema, pode ser localizado seu Tema; localizando-se o Conceito, pode-se localizar seu Termo. E em outra linha, mais recente, a dos tesauros-com-base-em-conceitos, originada da junção do Conceito de Dahlberg com a Teoria da Classificação, podemos acrescentar que a Linguística Sistêmico-Funcional também pode contribuir com um melhor entendimento do conceito e do termo, dos remas e dos temas, com a organização dos conceitos por meio de categorias e com o uso de definições para o posicionamento do conceito no sistema de conceitos.

## REFERÊNCIAS

BARROS, Lídia Almeida. **Curso básico de terminologia**. São Paulo: EdUSP, 2004.

BERTOLINI, María Violeta; BOTHMANN, Robert L; RODRÍGUEZ, Elena Escolano; GALEFFI, Agnese; McGARRY, Dorothy. **Declaração dos Princípios Internacionais de Catalogação (PIC)**.

**XXV Encontro Nacional de Pesquisa em Ciência da Informação - XXV ENANCIB**  
**Rio de Janeiro, RJ - 03 a 07 de novembro de 2025**

2016. Disponível em [https://www.ifla.org/wp-content/uploads/2019/05/assets/cataloguing/icp/icp\\_2016-pt.pdf](https://www.ifla.org/wp-content/uploads/2019/05/assets/cataloguing/icp/icp_2016-pt.pdf). Acesso em: 17 maio 2025.

CABRÉ, Maria Teresa. **La terminología**: representación y comunicación: elementos para una teoría de base comunicativa y outros artículos. Barcelona: Institut Universitari de Lingüística Aplicada, 1999.

CINTRA, Anna Maria Marques. Estratégias de leitura em documentação. In: SMIT, J. W. (coord.) **Análise documentária: a análise da síntese**. 2. ed. Brasília, DF: IBICT, 1987.

DAHLBERG, Ingetraut. Teoria do conceito. **Ciência da Informação**, Rio de Janeiro, v. 7, n. 2, 1978. p. 101-107. Disponível em: <http://revista.ibict.br/ciinf/article/view/115>. Acesso: 15 maio 2025.

DIDEROT, Denis, **Enciclopédia, ou Dicionário razoado das ciências, das artes e dos ofícios** / Denis Diderot, Jean Le Rond D'Alembert; organizado por Pedro Paulo Pimenta, Maria das Graças de Souza; traduzido por Pedro Paulo Pimenta, Maria das Graças de Souza, Thomas Kawauche. - São Paulo: Editora Unesp Digital, 2017. Tradução de: Encyclopédie, ou Dictionnaire raisonné des Sciences, des arts et des métiers (Originalmente publicado em 1750).

EGGINS, Suzanne. **An introduction to systemic functional linguistics**: Grammar. 2nd ed. Londres: Continuum Intl. Publishing Group, 2004.

FRIES, Charles C. On the status of theme in English. In: PETÖFI, Janous S.; SÖZER, Emel. (ed.). **Micro- and macro-connexity of texts**. Hamburg: Buske Verlag, 1983. p. 116-152.

FUZER, Cristiane; CABRAL, Sara Regina Scotta (org.). **Introdução à gramática sistêmico-funcional em língua portuguesa**. Santa Maria: Universidade Federal de Santa Maria, Centro de Artes e Letras Vernáculas, Núcleo de Estudos em Língua Portuguesa, 2010.

GIL, Antônio Carlos. **Como elaborar projetos de pesquisa**. 4. ed. São Paulo: Atlas, 2002.

GNOLI, Claudio. **Introduction to knowledge organization**. Londres: Facet Publishing, 2020. 148 p.

GOUVEIA, Carlos A. M. Texto e gramática: uma introdução à linguística sistêmico-funcional. **Matraga**, Rio de Janeiro, v. 16, n. 24, p. 13-47, jan./jun. 2009. Disponível em: <https://www.e-publicacoes.uerj.br/index.php/matraga/article/view/27795/19916>. Acesso em: 17 maio 2025.

HALLIDAY, Michael Alexander Kirkwood. **An Introduction to functional grammar**. Londres: Edward Arnold, 1985.

HALLIDAY, Michael Alexander Kirkwood. **An introduction to functional grammar**. 2nd. ed. Londres: Edward Arnold, 1994.

**XXV Encontro Nacional de Pesquisa em Ciência da Informação - XXV ENANCIB**  
**Rio de Janeiro, RJ - 03 a 07 de novembro de 2025**

HALLIDAY, Michael Alexander Kirkwood; MATTHIESSEN, Christian Matthias Ingemar Martin. **An introduction to functional grammar**. 3rd. ed. Londres: Edward Arnold, 2004.

HOUAISS, Antonio. **Dicionário Houaiss de Língua Portuguesa**. Elaborado pelo Instituto Antônio Houaiss de Lexicografia e banco de dados da língua portuguesa S/C Ltda. Rio de Janeiro: Objetiva, 2021.

LARA, Marilda Lopes Ginez de. Linguagem documentária e terminologia. **Transinformação**, Campinas, v.16, n.3, set./out. 2004. p. 231-240.

LARA, Marilda Lopes Ginez de; TÁLAMO, Maria de Fátima Gonçalves Moreira. O campo da lingüística documentária. **Transinformação**, Campinas, v. 18, n.3, p. 203-211, set./dez. 2006.

LARA, Marilda Lopes Ginez de; TÁLAMO, Maria de Fátima Gonçalves Moreira. Uma experiência na interface lingüística documentária e terminologia. **DataGramZero**, Rio de Janeiro, v. 8, n. 5, out. 2007. Disponível em: <https://brapci.inf.br/index.php/res/v/6681>. Acesso em: 10 maio 2025.

LIMA, Vânia Mara Alves. **Da classificação do conhecimento científico aos sistemas de recuperação de informação**: enunciação de codificação da informação documentária. 2004. Tese (Doutorado em Ciências da Comunicação) – Escola de Comunicações e Artes, Universidade de São Paulo, São Paulo, 2004.

LIMA, Larissa Mello; MOREIRA, Walter; MORAES, João Batista Ernesto. Linguística documentária e análise do discurso: um mapeamento em conceitos. *In*: SEMINÁRIO EM CIÊNCIA DA INFORMAÇÃO, 6., 2016, Londrina. **Anais [...]**. Londrina: Universidade Estadual de Londrina-UEL, 2016. p. 990-1003.

MAIMONE, Giovana D. **Terminologia**. 16 set. 2019. 50 slides. Material apresentado para a disciplina de Informação e Linguagem no PPGCI-ECA/USP.

MARTIN, James Robert; ROSE, David. **Genre relations**: mapping culture. Londres: Equinox, 2008.

MENDES, José Mário de Oliveira. **Processos verbais em artigos científicos da administração de empresas e engenharia**: perspectivas sistêmico-funcionais. 2013. Dissertação (Mestrado em Linguística) - Pontifícia Universidade Católica de São Paulo, São Paulo, 2013.

SANTOS, Cibele Araújo Camargo Marques dos; MAIMONE, Giovana Deliberali; LIMA, Vânia Mara A. **Linguística documentária**. 12 ago. 2019. 35 slides. Material apresentado para a disciplina de Informação e Linguagem no PPGCI-ECA/USP.

SAUSSURE, Ferdinand de. **Curso de linguística geral**. São Paulo: Cultrix, 1972.

SEVERINO, Antonio Joaquim. **Metodologia do trabalho científico**. 23. ed. São Paulo: Cortez, 2007.

**XXV Encontro Nacional de Pesquisa em Ciência da Informação - XXV ENANCIB**  
**Rio de Janeiro, RJ - 03 a 07 de novembro de 2025**

THOMPSON, Geoff. **Introducing functional grammar**. London: Arnold, 1996.

WÜSTER, Eugen. **Introducción a la teoría general de la terminología y a la lexicografía terminológica**. Tradução Anne-Cécile Nokerman. Barcelona: IULA, 1998.