

Unidades de conhecimento especializado, caracterização e tipologia¹

M. Teresa Cabré Castellví e Rosa Estopà Bagot²

Tradução: Sue Anne Christello Coimbra³

Revisão: Cleci Regina Bevilacqua e Maria Lúcia
Machado de Lorenci⁴

Resumo: Neste texto propomo-nos a analisar a noção e tipologia das Unidades de Conhecimento Especializado (UCEs) para estabelecer os parâmetros pertinentes para sua classificação e determinar quais elementos são relevantes para caracterizar as Unidades Terminológicas (UTs) e diferenciá-las de outros tipos de UCEs que aparecem com elas no discurso especializado. Os materiais empíricos nos quais nos baseamos procedem do âmbito da Genômica.

1. Introdução

Muitos são os autores que se ocuparam em descrever as características presentes nas situações que convencionamos denominar *especializadas*. Entre os fatores relevantes, cabe destacar as condições do emissor, do receptor, da situação, da temática e da função comunicativa básica.

Cada um dos elementos anteriormente citados requer uma breve precisão. Inicialmente, temos que esclarecer que geralmente não se trata de que elementos com determinadas características imponham às suas produções a condição de especializadas - embora, em alguns casos, a correlação seja mais do que evidente - mas de constatar que nas situações comunicativas, que se reconhecem externamente como especializadas, os elementos que se dão em todo esquema de comunicação apresentam características peculiares constantes.

¹ Traduzido com a permissão das autoras para a publicação nos Cadernos de Tradução do IL, a partir do texto original que se insere no projeto de pesquisa TEXTERM2: *Fundamentos, estratégias y herramientas para el procesamiento y extracción automáticos de información especializada*. Foi publicado em espanhol em Cabré, M. T.; Bach, C.; (2005): *Coneixement, llengua i discurs especializat*. Barcelona: IULA.

² Institut Universitari de Lingüística Aplicada (IULA), Universitat Pompeu Fabra, Barcelona, Espanha.

³ Aluna do Bacharelado em Letras, UFRGS.

⁴ Professoras do Instituto de Letras, UFRGS.

O emissor ou emissores são sempre indivíduos com um conhecimento importante do assunto de comunicação, conhecimento que, em regra geral, aprenderam conscientemente em uma situação explícita de aprendizagem. Esse conhecimento é o que assegura que um assunto especializado seja tratado como tal, como veremos a seguir.

O receptor ou receptores da comunicação especializada são muito diversos e seria absurdo tentar classificá-los de maneira precisa, porém existe uma condição que se dá em todos eles: o fato de terem expectativa de receber informação, de serem informados, expectativa que não ocorre em outras situações comunicativas cujo propósito geral é de outra índole.

Em relação ao conhecimento transmitido, podemos dizer que somente será especializado caso se represente e se transmita respeitando escrupulosamente a estrutura conceitual do campo em questão. Isso significa que cada conceito de uma disciplina será transmitido sem ultrapassar os limites de suas características, além de salvaguardar as fronteiras nítidas existentes entre os conceitos no contexto de uma especialidade. A propriedade de precisão semântica dos termos só pode ser respeitada na comunicação se quem comunica controla bem a estrutura conceitual do âmbito (Cabré 2002b).

Esse controle conceitual, necessário para comunicar um conhecimento especializado, condiciona e explica que nesse tipo de transmissão a função comunicativa primordial seja a informativa, isso apesar de que em uma análise de texto possam ser encontradas pistas de outras funções.

Enfim, as situações nas quais se produz e se transmite conhecimento especializado costumam ser profissionais, essencialmente produzidas por especialistas e dirigidas a um público também especialista ou semi-especialista. Uma exceção parece ser a comunicação científica ou técnica com intenção divulgativa, mas não é assim se observamos que a transmissão de informação especializada sempre cumpre a condição de ser o especialista ou seu mediador quem transmite o conhecimento com a finalidade constante de aumentar a competência especializada em uma temática do destinatário (neste caso, o leigo). Nessas situações, que não seriam as mais significativas dentro do esquema da comunicação especializada, o contexto situacional adquire traços de "profissional", visto que quem produz o conhecimento atua como um especialista, mesmo que seja apenas um mediador atuando como um especialista na transmissão de formas adequadas de comunicação segundo as condições discursivas (Cabré 2002a).

Todo esse conjunto de características permite diferenciar a comunicação especializada em geral de outros tipos de comunicação (lúdica, literária, emotiva, etc.) e encontrar nos textos produzidos características estruturais específicas.

2. Estruturação dos textos especializados⁵

A observação dos textos especializados (orais e escritos), considerados especializados

⁵ Todos os exemplos aqui apresentados foram mantidos em castelhano.

pelas condições discursivas que os caracterizam, leva-nos a estabelecer, seguindo a proposta de Cabré (2002b), pelo menos a conjunção de três estruturas:

- uma estrutura formal, relacionada com o gênero textual de cada texto;
- uma estrutura informativa, melhor denominada cognitiva, que representa o conhecimento transmitido pelo texto;
- uma estrutura gramatical, conformada pela inter-relação das diferentes estruturas lingüísticas que formam o texto.

As três estruturas que constituem um texto estão intimamente imbricadas entre si, de forma que a aparição relevante de determinadas Unidades de Conhecimento Especializado (UCEs) ou sua maior ou menor concentração depende da estrutura formal e da estrutura cognitiva.

Cognitivamente, as Unidades Terminológicas (UTs) estão todas elas relacionadas, mais ou menos explicitamente, mais ou menos diretamente, e se agrupam em conjuntos que podem ser sequenciais e não sequenciais. O conjunto das UTs e suas relações pode ser representado topograficamente em forma de mapa cognitivo. O mapa de um texto corresponde à sua representação cognitiva. O mapa consensuado sobre uma matéria corresponde à estruturação geral de um âmbito especializado ou a uma escola de pensamento ou ação.

Finalmente, a inter-relação entre a forma e a gramática condiciona o tipo de estruturas sintáticas, os marcadores, a maior ou menor densidade de léxico especializado, etc.

Vejamos a seguir, brevemente, algumas características das três estruturas.

A estrutura formal do texto, associada ao gênero textual ao qual pertence, distribui o texto em partes relevantes que o caracterizam como pertencente a um tipo textual determinado. Assim, um *paper* científico, por exemplo, freqüentemente está estruturado segundo a Convenção de Vancouver, estabelecida em 1978 por um grupo de editores e que atualmente é apoiada pelo Comitê Internacional de Diretores de Revistas Médicas (CIDRM), nos seguintes itens (Junyent, 2003):

- Título
- Autor ou autores e organizações de referência
- Resumo
- Introdução ou justificativa
- Materiais e Métodos
- Resultados obtidos
- Discussão ou conclusões
- Bibliografia

A estrutura de conhecimento de uma matéria, bem como a que apresenta um texto especializado, poderia ser representada como um mapa de estruturas de núcleos e relações.

Os núcleos podem ter diferente alcance e, de acordo com esse critério, corresponder lingüisticamente a fragmentos textuais (por exemplo, as partes de um texto pertencente a um determinado gênero textual ou um parágrafo), a

unidades lingüísticas oracionais, sintagmáticas, léxicas ou morfológicas. As unidades mínimas de um núcleo são as léxicas, embora em determinadas matérias as unidades morfológicas marquem o caráter especializado dessas unidades léxicas e constituam, portanto, o foco de seu valor especializado. Em cada núcleo mínimo figura, pois, uma UCE de caráter léxico. As relações estão expressas lingüística ou topograficamente.⁶

A partir de um fragmento de um texto especializado extraído de uma revista médica, podemos observar representações das UCEs de diferentes níveis com suas relações. Na primeira ilustração, marcamos com um círculo os núcleos de conhecimento e sublinhamos as relações entre esses núcleos. Constatamos, também, que a relação entre fragmentos textuais estabelece-se normalmente (embora não exclusivamente) sobre a base dos denominados marcadores textuais. A ilustração 2 apresenta a estrutura textual mínima de conhecimento e, nela, podemos observar que os núcleos estão ocupados por unidades de conhecimento que correspondem a unidades do léxico. A última ilustração é uma abstração do conhecimento transmitido pelo texto.

La Esclerosis Múltiple (EM) es una enfermedad del Sistema Nervioso Central (SNC) que afecta de forma focal a la sustancia blanca. La lesión característica consiste en una desmielinización segmentaria de la sustancia blanca del SNC, asociada a un infiltrado inflamatorio perivascular en las fases activas (o agudas) de la enfermedad. Si bien la causa de la enfermedad sigue siendo desconocida, hay numerosos datos que indican que el sistema inmunológico juega un papel importante.

[A Esclerose Múltipla (EM) é uma enfermidade do Sistema Nervoso Central (SNC) que afeta de forma focal à substância branca. A lesão característica consiste em uma desmielinização segmentar da substância branca do SNC, associada a um infiltrado inflamatório perivascular nas fases ativas (ou agudas) da doença. Apesar de a causa da enfermidade continuar sendo desconhecida, há numerosos dados que indicam que o sistema imunológico desempenha um papel importante.]

Esse fragmento pode ser descomposto nas seguintes proposições:

La Esclerosis Múltiple (EM) es una enfermedad del Sistema Nervioso Central (SNC)

(La Esclerosis Múltiple (EM) afecta de forma focal a la sustancia blanca (del SNC)

La lesión característica consiste en una desmielinización segmentaria de la sustancia blanca del SNC

(La desmielinización segmentaria de la sustancia blanca del SNC) (está) asociada a

⁶ Para aprofundar nesse ponto, ver Cabré (2002b).

un infiltrado inflamatorio perivascular en las fases activas (o agudas) de la enfermedad

La causa de la enfermedad sigue siendo desconocida

Hay numerosos datos que indican que el sistema inmunológico juega un papel importante en (la causa de la enfermedad desconocida)

A representação conceitual geral desse fragmento, levando em consideração todas as informações do texto, poderia ser assim representada:

La (Esclerosis Múltiple (EM)) es una enfermedad del Sistema Nervioso Central (SNC) que afecta de forma focal a la sustancia blanca. La lesión característica consiste en una desmielinización segmentaria de la sustancia blanca del SNC, asociada a un infiltrado inflamatorio perivascular en las fases activas (o agudas) de la enfermedad. Si bien la causa de la enfermedad sigue siendo desconocida, hay numerosos datos que indican que el sistema inmunológico juega un papel importante.

figura 1

Esta representação, se prescindirmos dos dados não essenciais para a caracterização do objeto *esclerosis múltiple*, poderia ser reduzida a:

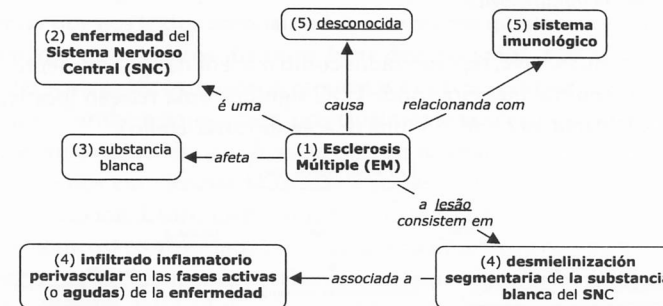


figura 1

Nessa representação reduzida pode ser observado que a informação se organiza em torno de diferentes classes de informação, classes que correspondem a diferentes aspectos (ou dimensões) do objeto conceitual *esclerosis múltiple*, como

é mostrado na ilustração 3:

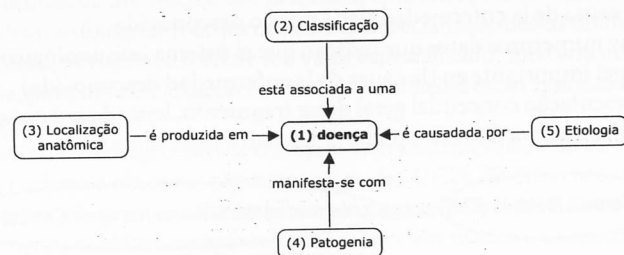


figura 3

Porém, se desse fragmento de texto extraíssemos não todos os conhecimentos que o texto transmite, mas somente os conceitos essenciais nos quais se baseia o conhecimento sobre o objeto *esclerose múltipla* (EM), percebemos que tais conceitos fazem parte da estrutura básica do objeto conceitual *esclerose múltipla*. Observa-se também que todos os conceitos coincidem com unidades terminológicas:

Conceitos-chave específicos:
Sistema nervioso central (SNC)
Sustancia branca
Desmielinização segmentaria
Infiltrado perivascular
 Conceitos genéricos
Lesión
Fase activa o aguda
Enfermedad
Sistema inmunológico

Os conceitos-chave, representados como núcleos de conhecimento, relacionam-se da seguinte maneira (onde LOC significa uma relação locativa, PART uma relação partitiva e CAUSA uma relação de causa-efeito):

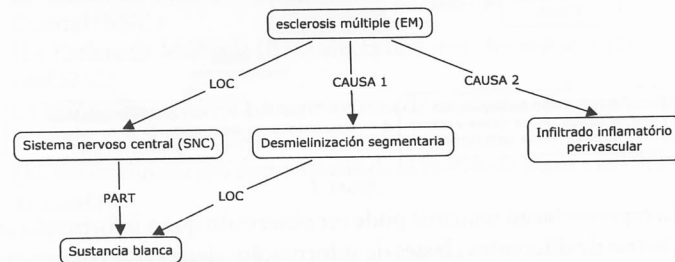


figura 4

Deixando momentaneamente de lado a estrutura conceitual ou cognitiva e observando sua estrutura gramatical, podemos ver que a estrutura gramatical desse fragmento organiza-se em diferentes níveis lingüísticos:

Um nível textual (que nesse caso corresponderia à totalidade do texto do qual procede esse fragmento), no qual aparecem distintas estruturas multioracionais e oracionais inter-relacionadas por marcadores discursivos⁷.

Um nível sintático, no qual as orações são compostas de constituintes estruturais ou sintagmas que ocupam um lugar e exercem uma função na oração.

No nível sintagmático infra-oracional, temos combinações estruturadas hierarquicamente de unidades do léxico que, em conjunto, podem corresponder a unidades do nível léxico inferior ou a outras unidades que são simples combinações sintagmáticas de discurso. No léxico temos uma série de unidades que estruturalmente podem proceder da sintaxe (sintagmas lexicalizados) ou ser originalmente lexicêmicas.

Todas essas unidades possuem uma estrutura morfológica, embora algumas se componham somente de um elemento morfológico. Portanto, no nível da estrutura léxica podem ser detectadas também unidades morfológicas de valor especializado.

À parte o critério compositivo na organização estrutural da gramática do texto especializado, dão-se muito sistematicamente algumas características gramaticais que, de modo orientado conforme os casos, permitem caracterizar especificamente o texto especializado em todos os níveis lingüísticos, ou, melhor ainda, que permitem contrastá-lo com as características dos textos não especializados.

No nível textual existe um controle bastante sistemático da distribuição da informação, que aparece mais ou menos concentrada e é de um ou outro tipo em função do gênero do texto. Também existe uma organização espacial precisa dos conteúdos mais rigorosa que nos textos não especializados. E, sobretudo, pode-se observar que nos textos especializados ocorre uma insistência em apresentar formalmente os conceitos distintos bem diferenciados. Por isso, aparecem freqüentemente gráficos, quadros, esquemas e outros recursos organizativos da informação. Também apresentam, principalmente nos textos de matérias técnicas e de algumas matérias científicas de ciências naturais, médicas, experimentais e exatas (menos em ciências humanas e sociais), símbolos e abreviaturas de caráter internacional, fórmulas, nomenclaturas científicas e outros signos (poucas vezes icônicos) representativos dos conceitos simples ou complexos do âmbito. Observa-se, também, nesse nível uma tendência à impessoalização ou ainda à assunção do EU individual por parte do EU coletivo (nós), à apresentação dos

⁷ É nesse nível textual que se pode detectar o tipo ou classe textual, através da organização da informação e da presença de marcadores textuais.

dados de forma que sejam percebidos como objetivos, às citações que corroborem a informação (polifonia discursiva), etc.

No nível sintático, detecta-se nos textos especializados de caráter descritivo uma tendência a estruturas pouco complexas e a orações de menor extensão. Determinados tipos oracionais não aparecem, ou se o fazem não são representativos, como, por exemplo, frases exclamativas ou interrogativas. A extensão das orações, no entanto, costuma variar em função do tipo de texto e, por exemplo, nos de tipo argumentativo, a sintaxe é mais complexa e as orações de maior extensão. Cabe tratar também como caso excepcional o discurso estereotipado próprio de determinados âmbitos, como poderia ser o jurídico, de muito maior complexidade sintática.

No nível léxico, nível que juntamente com o textual mais diferencia o discurso especializado do não-especializado, observa-se uma seleção do léxico dependente de domínio, de léxico, seja formal e semanticamente ou apenas semanticamente. Formalmente e semanticamente, porque existem vocábulos que somente aparecem no discurso de uma determinada especialidade. Apenas semanticamente, porque, embora se trate de vocábulos utilizados também em outras matérias ou na comunicação geral, esses adquirem um sentido específico dentro de cada domínio especializado, sentido que se encontra no conjunto de sentidos que conformam o significado global de uma unidade léxica⁸. Do ponto de vista do léxico, observa-se no texto especializado uma maior proporção de siglas e acrônimos, e, às vezes, de estruturas pouco naturais morfológicamente (por exemplo, o nome de algumas proteínas como *1-amilasa*, *B-amilasa*, em que se usa o hífen como indicador de um falso prefixo). Enfim, é preciso insistir que em todos os textos especializados, ao lado das unidades léxicas específicas do âmbito em questão, aparece o vocabulário geral necessário para travar o discurso.

No nível semântico, detectamos, nos textos de maior nível de especialidade, uma tendência a evitar a redundância, tendência que desaparece gradualmente à medida que diminui o nível de especialização do texto. Dentro da redundância, no entanto, destaca-se, nos textos especializados, o uso de repetições explicativas ou parafrásticas do tipo UT seguida de definição, explicação ou paráfrase ou ao contrário explicação ou definição seguida de UT. Esses recursos, quer sejam UTs, quer sejam paráfrases ou definições, freqüentemente figuram entre parênteses, aparecem conectados mediante conjunções ou fazem parte de cláusulas explicativas, oracionais ou não. É de destacar também nos textos especializados, como consequência da busca da maior precisão, a introdução de marcadores denominativos explícitos do tipo Unidade parafrástica + Marcador metalinguístico + UT. A

⁸ Foi nessa acepção que em alguns trabalhos, como em Estopà (1999, 2002), utilizou-se a sigla USE (Unidades de Significação Especializada) para se referir a tais unidades.

redundância manifesta-se, também, no uso de ilustrações e gráficos que reforçam (situam, precisam ou complementam) determinados conceitos ou estruturas conceituais expressos no texto.

No nível morfológico, aparecem morfemas específicos de um âmbito, embora não em todos os âmbitos com o mesmo grau. Podemos observar também uma tendência à complexidade morfológica por parte dos termos, maior do que no vocabulário-chave dos textos não especializados, e uma tendência maior à nominalização dos verbos, dando ao texto um caráter mais designativo e aportando um ingrediente maior de abstração.

Finalmente, embora a ortografia dos textos especializados seja a comum da língua na que são elaborados, cabe notar, em determinadas matérias, algumas grafias não correntes, como a repetição ou alternância de maiúsculas e minúsculas nas siglas, contrariando a ortografia das palavras (*anti-VIH*, *ADN circular*), o uso de letras de outros alfabetos (*rayos γ*) ou a presença simultânea de letras e números em um mesmo vocábulo (*1-amilasa*, *ácido C terminal*, *acetil-CoA*).

3. Tipologia geral das Unidades de Conhecimento Especializado (UCEs)

A observação do texto, priorizando o aspecto de conhecimento transmitido, permite estabelecer elementos de diferente natureza na estrutura de conhecimento e seus elementos, os núcleos e as relações.

Quando o foco de nossa observação encontra-se na relação entre os elementos de conhecimento que conformam a estrutura cognitiva do texto e as unidades que veiculam esse conhecimento, entramos já no terreno da linguística. Nesse plano, forma e significado chegam a ser indissociáveis.

Desse ponto de vista, observamos que tanto as áreas de conhecimento do texto como os núcleos cognitivos correspondem-se a unidades linguísticas, sejam grupos de proposições relativas a um ponto ou núcleo de informação ou elementos que constituem essas proposições com o papel de predicados ou argumentos.

3.1. Questões de terminologia

Denominamos genericamente UCEs (*Unidades de Conhecimento Especializado*) as unidades de diferente nível descritivo que constituem os núcleos de conhecimento de um texto ou fazem parte deles. A condição essencial para considerar que uma unidade é uma UCE é o tipo de conhecimento que transmite (condições cognitivas e semânticas) e seu uso em discurso (condições pragmático-discursivas).

Apresentamos a seguir os termos e siglas mais relevantes deste trabalho, tanto os referidos à classificação como aos elementos de análise:

Unidade terminológica (UT): unidade léxica, cuja estrutura corresponde a

uma unidade léxica de origem ou produto da lexicalização de um sintagma. Possui um significado específico no âmbito ao qual está associada e é necessária na estrutura conceitual do domínio de que faz parte. As condições que uma unidade cumpre para ser uma UT são três:

- estrutura;
- especificidade semântica;
- necessidade na estrutura conceitual.

Unidade sintagmática ou *unidade poliléxica*: unidade complexa de estrutura sintática. Em função de seu grau de lexicalização, pode corresponder a uma unidade léxica ou a uma unidade fraseológica (David, 1993). Se essa unidade léxica cumpre os requisitos de especificidade semântica e necessidade, corresponde a uma *unidade terminológica*. Caso cumpra o requisito de especificidade, mas não o de necessidade, corresponde a uma *unidade fraseológica especializada*, que contém sempre uma UT.^{9,10}

Unidade monoléxica: unidade léxica formada por uma só unidade léxica. As unidades monoléxicas podem ser simples ou construídas, formadas por um único lexema ou por vários lexemas.

3.2. Critérios de classificação e tipologia das UCEs

As UCEs podem ser classificadas por diferentes critérios. Os seis que, em nossa opinião, são relevantes para distinguir os tipos de UCEs são:

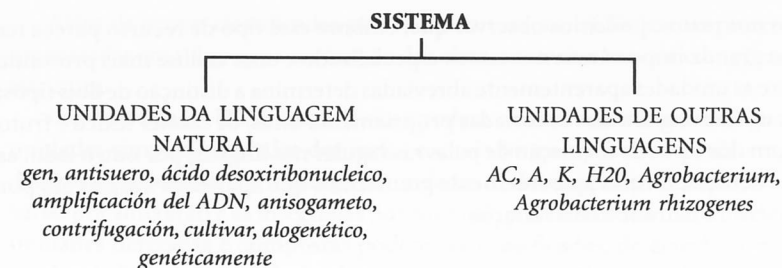
- o sistema ao qual pertencem;
- a estrutura;
- a categoria gramatical;
- a classe conceitual;
- a coesão interna;
- o papel lingüístico-conceitual que exercem no discurso especializado.

3.2.1. Classificação pelo sistema ao qual pertencem

Do ponto de vista do sistema ao qual pertencem, as UCEs podem ser unidades da linguagem natural ou pertencer a uma grande multiplicidade de sistemas artificiais.

⁹ Ver Bevilacqua 2004.

¹⁰ Corbin (1992, 1997) utiliza o termo *unidade polilexêmica* para denominar a unidade léxica formada por mais de um lexema, e inclui nesse conceito todos os tipos de unidades compostas, tanto as compostas cultas, como as patrimoniais e mistas, e as unidades sintagmáticas lexicalizadas. Neste trabalho não utilizamos essa denominação.



Chamamos a atenção sobre o fato de que as unidades especializadas representadas por signos diferentes dos do sistema da linguagem natural equivalem, na comunicação, às mesmas categorias gramaticais do sistema lingüístico natural. Conceitualmente, podem ser descritas através das mesmas classes semânticas, sua estrutura pode corresponder à tipologia das unidades do sistema lingüístico, podem também apresentar diferentes graus de coesão interna caso se trate de unidades sintagmáticas e cumpram os mesmos papéis no discurso que as unidades lingüísticas.

Freqüentemente encontramos nos textos especializados unidades híbridas em cuja formação intervêm elementos de diferentes sistemas de signos, combinando-se unidades da linguagem natural com unidades de linguagens artificiais diversas. Seria o caso dos elementos químicos: os sufixos do castelhano *-oso*, *-asa*, *-ico*, *-ina*, etc. que, por seu aspecto formal, parecem sufixos da linguagem natural, e, pelo contrário, são de criação totalmente artificial (*acetato*, *adenina*, *adenosina*, *azarosa*, *alanina*, *alcohol*, *amilasa*, *aminopeptidasa*, *anión*, *aplicón*, *ampilicina*). Outros exemplos de unidades híbridas são as combinações de símbolos (números e letras) com unidades léxicas (*ácido C terminal*, *acetil-CoA*, *ARNm*, *ADN A*, *actinomicina D*, *hepatitis*) ou as combinações de unidades léxicas de dois sistemas distintos (*1-amilasa*, *B-amilasa*, *Agrobacterium mediated*).

As siglas merecem, no discurso especializado, uma menção explícita, por três razões. Em primeiro lugar, por sua importância quantitativa, crescente na sociedade contemporânea. Em segundo lugar, pela variação que introduzem nos textos quando esses não são originais, mas traduzidos. A diferença estrutural das línguas, assumida totalmente na tradução quando se refere à morfologia e à sintaxe, resolve-se com variação no caso das siglas (*ADP*, *ARN*, *ADN* junto a *PDA*, *RNA* ou *DNA*). Em terceiro lugar, porque as siglas, enquanto unidades oscilantes entre o léxico de uma língua natural e o de um sistema artificial, combinam-se com unidades léxicas do sistema lingüístico, seguindo as regras combinatórias próprias da língua nas quais são usadas, tanto se correspondem a unidades próprias como se correspondem a empréstimos (*ADN circular*, *ADN copia ADN nativo*, *ADN pasajero*, *ADN polimerasa dependiente del ADN*, *amplificación del ADN*, *amortiguador TEB*). Enfim, em relação às unidades abreviadas que apare-

cem nos textos, podemos observar que, embora esse tipo de recurso pareça ter uma grande importância nos textos especializados, uma análise mais profunda sobre as unidades aparentemente abreviadas determina a distinção de dois tipos: por um lado, as formas abreviadas propriamente ditas, de caráter léxico e fruto de um dos tipos de formação de palavras regular nas línguas; por outro lado, as abreviaturas, formas aparentemente gramaticais que aparecem nas línguas por consenso e não admitem variação.

3.2.2. Classificação pela estrutura

Do ponto de vista de sua estrutura interna, as UCEs podem coincidir com morfemas (-itis, -ico, apo-, agluto-, -asa, -ina), com unidades léxicas, simples, derivadas e compostas (*virus, hepatitis, ocular, autonómico, anabólico, apógamo, cruzar, coagular, amplificar, genético, antisuero, Agrobacterium*), com sintagmas, sejam terminológicos ou fraseológicos (*esclerosis múltiple, nervio alveolar inferior, secuenciación del genoma, altamente calórico, amplificación del ADN, Agrobacterium tumefaciens*), ou com orações (que são muito raras).

No caso das UCEs lingüísticas (morfemas, palavras, sintagmas ou orações), as categorias ou classes estruturais que lhes são atribuídas devem ser interpretadas literalmente, de acordo com a terminologia gramatical. Cada unidade corresponde a um morfema, a um lexema, a uma combinação de lexemas ou a uma oração. No entanto, quando nos referimos às UCEs não lingüísticas, a classe estrutural deve ser interpretada como “correspondente” às próprias das unidades lingüísticas: cada unidade associa-se a uma classe que corresponde a um morfema, um lexema, um sintagma ou uma oração.

ESTRUTURA

UNIDADES MORFOLÓGICAS

-itis, -genia, apo-, agluto-, -oide, -soma, -asa, ina, -ol, -osa, -ato

UNIDADES MONOLÉXICAS

célula, caldo, clonaje, codificar, clonar, asepsia, anabólico, apógamo, autonómico

UNIDADES SINTAGMÁTICAS (unidades léxicas (poliléxicas) e unidades fraseológicas)

ácido desoxirribonucleico altamente repetitivo, ácido desoxirribonucleico de doble cadena, célula hija, ARN satélite, carácter de un gen individual, alteración cromosómica, adaptación del regulador del crecimiento, altamente calórico, célula ovárica de hámster chino, encapsular el ARN viral, purificar el ARNnh, acortar el ARN, responder a un antígeno

UNIDADES ORACIONAIS

capa obtenida por ingeniería genética, el ARNm se pliega, el ARNm incorpora las instrucciones genéticas, el ADN cromosómico permanece doblemente enhebrado

Além de serem classificadas por sua estrutura, as UCEs léxicas podem ser organizadas em subgrupos em função dos elementos internos que compõem sua estrutura.

Por exemplo, as UCEs monoléticas podem pertencer a três grandes grupos: unidades simples, unidades derivadas e unidades compostas; e ainda entre as derivadas podem ser distinguidas as unidades derivadas por prefixação, as derivadas por sufixação e as integradas por sufixos e prefixos conjuntamente¹¹. Já as unidades derivadas e compostas podem ser classificadas, de acordo com seus modos de formação, em derivadas e compostas cultas, formadas por morfemas ou bases léxicas de origem greco-latina, e derivadas e compostas patrimoniais, constituídas por morfemas ou bases léxicas de origem atual das línguas.

A classificação estrutural das unidades léxicas sintagmáticas realiza-se em função da categoria das unidades léxicas e gramaticais que as constituem e as descrevem como estruturas sintáticas. A maioria delas corresponde a sintagmas nominais, embora encontremos também sintagmas verbais e adjetivais (e, potencialmente, sintagmas adverbiais).

Apesar dessa formulação de possibilidades, é preciso esclarecer que a maioria dos sintagmas verbais cujos componentes co-aparecem com uma certa frequência nos textos de uma especialidade podem corresponder não a unidades terminológicas, mas a unidades fraseológicas próprias dessa especialidade. Ainda mais problemático é o caso dos sintagmas nominais cujo núcleo é um nome verbal relacionado com um verbo semanticamente “forte” no âmbito. Nesses casos, para diferenciar o que corresponderia a um fraseologismo ou a uma unidade terminológica, remetemo-nos a Lorente (2003), cujo trabalho de classificação dos verbos e nomes verbais em relação ao seu caráter fraseológico ou terminológico também faz parte dos resultados do Projeto Texterm2¹².

No entanto, há um dado que nos parece incontrovertível: na representação cognitiva da estrutura conceitual de um âmbito, os núcleos centrais estão “ocupados” por estruturas nominais (nomes ou sintagmas nominais), quer sejam de origem nominal, adjetiva ou verbal, até o ponto que quando uma noção que corresponde a um processo ou a uma ação possa ser representada verbalmente ou nominalmente, é a forma nominal a que “ocupa” o núcleo correspondente à estrutura conceitual de um âmbito. Esse dado permitiria formular duas questões: a primeira tem a ver com os modos de representação conceitual em terminologia e com os modos de conceitualização dos especialistas (não ousamos dizer, por falta de dados empíricos, dos modos de conceitualização dos falantes

¹¹ Essa composição não pressupõe que a formação se realize agregando simultaneamente um prefixo e um sufixo.

¹² Para uma descrição dos verbos que aparecem em um corpus textual de Genômica, pode-se consultar Lorente 2003.

em geral). A segunda levaria a questionarmos se é, efetivamente, através de unidades de categoria lingüística que o especialista organiza a estrutura de seu âmbito¹³; ou mais finamente, se é universal que seja através de unidades lingüísticas. A análise empírica sobre essas questões será objeto de trabalhos posteriores.

Recuperando agora as unidades sintagmáticas nominais na classificação das unidades terminológicas, as unidades sintagmáticas nominais podem corresponder a três estruturas:

- as formadas por um núcleo nominal complementado por um adjetivo ou sintagma adjetivo;
- as formadas por um núcleo nominal complementado por um sintagma preposicional;
- as formadas por um núcleo nominal complementado por um sintagma nominal.

ESTRUTURA DAS UCEs SINTAGMÁTICAS NOMINAIS

NÚCLEO: NOME

COMPLEMENTO: S. ADJETIVAL

célula simple, alcohol graso, contaminación bacteriana, baja sensibilidad, ADN viral, ácido graso saturado, cultivo celular primario, ácido desoxirribonucleico circular cerrad, célula bacteriana donante componente

NÚCLEO: NOME

COMPLEMENTO: S. PREPOSICIONAL

célula de fusión, condición de crecimiento, coeficiente de transferencia de masa, alta frecuencia de transducción, compuesto de alto peso molecular, ADN de cloroplasto, ADN de célula animal, concentración por congelación, almacenamiento en frío, cromatografía en gel, ciclo sin sentido, aislamiento del virus, copia del ADN, cebador del ARN, aislamiento de las proteínas

NÚCLEO: NOME

COMPLEMENTO: S. NOMINAL

Cromosoma hijo, célula recipiente, célula madre, célula B, cromosoma W, cadena no ramificada, cromatidio no hermano, ADN A, ADN no codificador,

¹³ Nesse sentido, pode-se ver a tese de doutorado de Cleci Bevilacqua (2004) que mostra como em um corpus de tratamentos meio ambientais o campo em questão se estrutura conceitualmente ao redor de núcleos verbais.

Cada uma dessas estruturas pode receber expansões, de forma subordinada ou através da coordenação de algum de seus elementos, ou incluir entres seus componentes essenciais outras unidades, basicamente determinantes ou quantificadores (*alta frecuencia de transducción, actividad fijadora del nitrógeno, adaptación del regulador del crecimiento, ADN circular cerrado covalentemente, ADN moderadamente repetitivo, análisis de inmunosorbentes enlazados a enzimas, ARN nuclear pequeño, cromatografía en capa delgada, cultivo en agar blando*).

3.2.3. Classificação pela categoria gramatical

Do ponto de vista morfossintático, as UCEs léxicas ou UTs podem pertencer a quatro categorias gramaticais: nominal, verbal, adjetival e adverbial. Dentre todas elas, as de categoria nominal são mais representativas, diferentemente das de categoria verbal e adjetival. As UCEs adverbiais são relativamente poucas e costumam ter um correlato adjetival no âmbito.¹⁴

CATEGORIA GRAMATICAL

NOMINAL

caldo, código, célula, almidón, anión, adenina, ácido aspártico, ADN, ADN-ARN, citoplasma, antisuerio, amiosis, clonaje de tejidos, coagulante, condición de crecimiento, cultivo hidropónico, Agrobacterium

ADJETIVA

alelomófico, alcalino, avirulento, alogénético, coloidal, antiviral, citotóxico, contaminante, abortivo, medioambientalmente negativo

VERBAL

centrifugar, clonar, acetilar, cultivar, codificar, administrar penicilina, hacer un antibiótico letal, unirse mediante apareamiento de bases, transferir por apareamiento bacteriano, cortar el ARNnh

ADVERBIAL

biológicamente, por vía oral, por vía intravenosa, altamente molecular, en posición de decúbito, inmunológicamente

3.2.4. Classificação pela classe conceitual

Semanticamente, as UCEs podem ser reduzidas a quatro grandes classes conceituais, se seguimos a classificação de Vendler (1967): entidades, eventos, propriedades e relações.

¹⁴ É preciso especificar que, embora neste parágrafo tratemos de categoria intrínseca nominal, verbal, adjetival e adverbial, caberia acrescentar, por um lado, as estruturas sintáticas lexicalizadas e, por outro, as unidades que, apesar de não pertencer a uma classe gramatical, passam a funcionar como próprias dessa classe (*por vía oral* (adv), *analizador* (m), *antifúngico* (m), *coagulante* (m), *concentrado* (m)).

CLASSE SEMÂNTICA

ENTIDADES
<i>célula, agar, copia, alcohol, comatina, proteína, ADN, ADN donador, agente viral, código genético</i>
EVENTOS (ação ou processo)
<i>amortiguación, almacenamiento, asimilación, condensación, aislamiento del virus, amplificación génica, control genético, aislar, cruzar, clonar, codificar, amplificar</i>
PROPRIEDADES
<i>aAutooxidable, alcalinidad, androesterilidad, consanguíneo, alogénico, calor de fermentación, capacidad hidrófila, genético</i>
RELAÇÕES
<i>almacenamiento, administrar, cultivar, clonar</i>

Essas quatro classes não podem ser concebidas nem como uma lista absoluta (existem outras classificações que ampliam as classes conceituais básicas), nem como uma proposta discreta. Por um lado, essas classes podem ser subespecificadas; a proposta do próprio Vendler (1967) mostra isso, referindo-se aos verbos, estabelece quatro classes aspectuais baseadas nos critérios de dinamicidade e telicidade: estados, atividades, *accomplishments* e *achievements* ¹⁵.

Por outra parte, essas classes opõem-se a descrever muitas vezes aquelas unidades que estão oscilando entre duas classes e participam, portanto, das características de ambas, às vezes com o predomínio de uma sobre a outra, mas em alguns casos sem elementos que permitam decantar-se claramente para uma ou outra opção, já que esta pode depender do ponto de vista que o analista priorize. Assim, por exemplo, unidades como *câncer*, *aborto bacteriano*, *clonaje forzado*, *clonaje pasmídico*, *clonación in vitro*, *agroinfección* podem ser consideradas eventos (ações ou processos) ou entidades (resultados) em função do ponto de vista que se priorize; *actividad génica*, *cambio evolutivo*, *cambio genético* podem ser considerados eventos e propriedades, e *centrifuga refrigerada* poderia ser entidade e propriedade.

Cada classe conceitual pode desdobrar-se também em subclasses que *herdam* todas ou algumas das propriedades da classe e subclasses superiores hierarquicamente. Esse princípio constitui a base das classificações hierárquicas e das ontologias especializadas. Vejamos a seguir alguns exemplos:

a) Sager e Kageura (1994/1995) formularam, no início da década de noventa,

¹⁵ Ciercia & McConnell-Ginet (1990: 353) definem cada um desses subtipos: States are like snapshots of the world at a given instant. They lack a natural culmination or end point, and their subject is perceived not as an agent (as doing something) but as an experiencer (as experiencing something). (...) Activities share with states the property of lacking a natural culmination. Yet they are agentive in that they typically involve a subject doing something. (...) Telic eventualities have a natural end point or culmination.

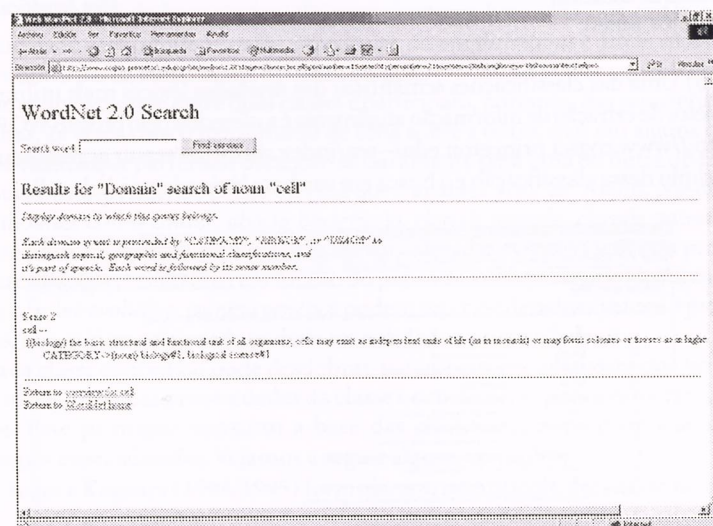
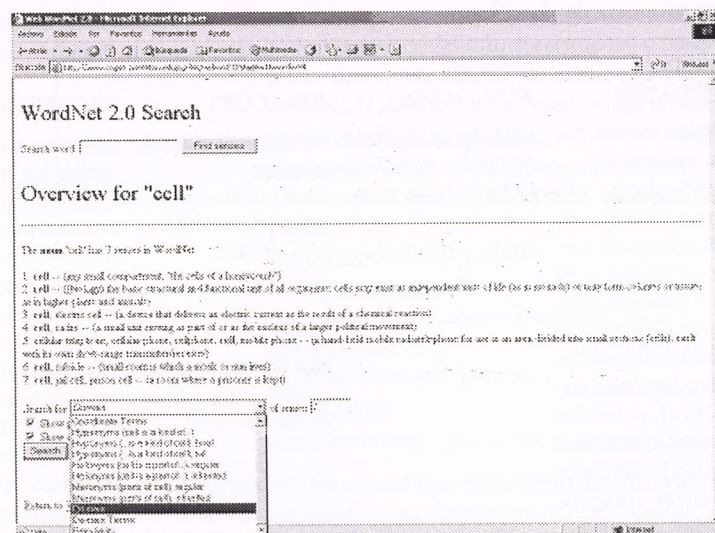
uma proposta de classificação hierárquica dos conceitos especializados a partir de quatro categorias muito básicas: entidades, atividades, qualidades e relações:

ENTIDADES	ATIVIDADES	RELAÇÕES	PROPRIEDADES
<u>Materiais:</u> animais (pessoas, organismos). <u>Classificadores:</u> elementos simples, elementos complexos, conceitos separados, parte de um termo complexo <u>Abstratos:</u> estrutura de conhecimento (geral, específico, linguístico, matemático); entidades abstratas (segundo uma temática); entidades subjetivas (papéis, avaliativas, psicológicas). <u>Representacionais:</u> Entidades de documentação (tipos, partes).	<u>Unidades de atividade:</u> Atividades específicas (ação/processo, transferência, mudança de estado, produção)	<u>Etiquetas de tipos de relações:</u> Gerais, conforme a área temática, medida na área temática <u>Valores de relações:</u> Físicos (espaciais, temporais, funcionais) Não físicos (modo de relação, comparação, situação relativa).	<u>Etiquetas do tipo de qualidades:</u> geral, segundo a área temática <u>Valores de qualidades:</u> Físicos (espaciais, temporais, funcionais, de situação, de forma, de cor, quantitativas, de escala, outras) Não físicos (gerais, especiais, naturais, ratificáveis, simples, complexos, negativos, de maneira, segundo a área temática).

b) Uma das classificações semânticas das unidades léxicas mais utilizadas em projetos de extração de informação atualmente é a oferecida pelo projeto WORDNET [<http://www.cogsci.princeton.edu/~wn/index.shtml>]. A seguir apresentamos um exemplo dessa classificação e a busca em sua base léxica da unidade *cell*:

The image shows a screenshot of a web browser displaying the WordNet interface. The title bar indicates 'The browser of the Synonym Graph and the corresponding word sets are listed below along with a brief description each file's contents'. The main content area is titled 'Linker Grapher Files' and contains a table with two columns: 'File Number' and 'Name'. The table lists 23 files, each representing a different sense of the word 'cell'. The descriptions are brief, such as 'cell as a container', 'cell as a unit of organization', 'cell as a unit of structure', etc.

File Number	Name	Description
01	cell	cell as a container
02	cell	cell as a unit of organization
03	cell	cell as a unit of structure
04	cell	cell as a unit of function
05	cell	cell as a unit of communication
06	cell	cell as a unit of information
07	cell	cell as a unit of energy
08	cell	cell as a unit of matter
09	cell	cell as a unit of time
10	cell	cell as a unit of space
11	cell	cell as a unit of temperature
12	cell	cell as a unit of pressure
13	cell	cell as a unit of volume
14	cell	cell as a unit of mass
15	cell	cell as a unit of length
16	cell	cell as a unit of area
17	cell	cell as a unit of weight
18	cell	cell as a unit of force
19	cell	cell as a unit of motion
20	cell	cell as a unit of sound
21	cell	cell as a unit of light
22	cell	cell as a unit of heat
23	cell	cell as a unit of electricity



c) O projeto SIMPLE propõe uma lista de quatorze papéis temáticos e traços semânticos para a descrição semântica do léxico das línguas (<http://www.ub.es/gilcub/SIMPLE/liege/spanish/Spanish.html>). Reproduzimos a lista dos papéis temáticos utilizados neste projeto e um exemplo dos traços que SIMPLE utiliza para descrever o conteúdo das unidades léxicas através da unidade:

SemanticRole list
 RoleKinship
 RoleIconic
 RoleHeadQuantified
 RoleHeadCollective
 RoleProtoAgent
 RoleProtoPatient
 Role2Participant
 RoleOblique
 RoleLocation
 RoleDirection
 RoleOrigin
 RoleSOAArg
 RoleAdjunct
 RoleUnderspecified

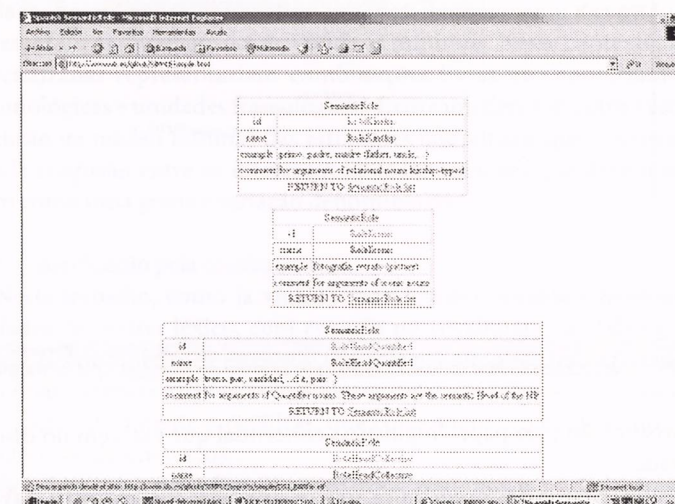
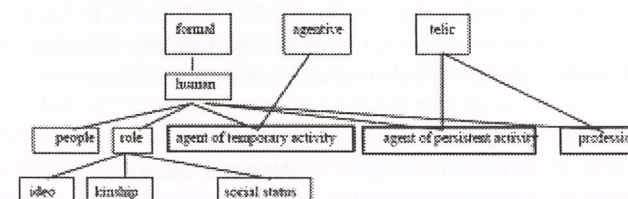
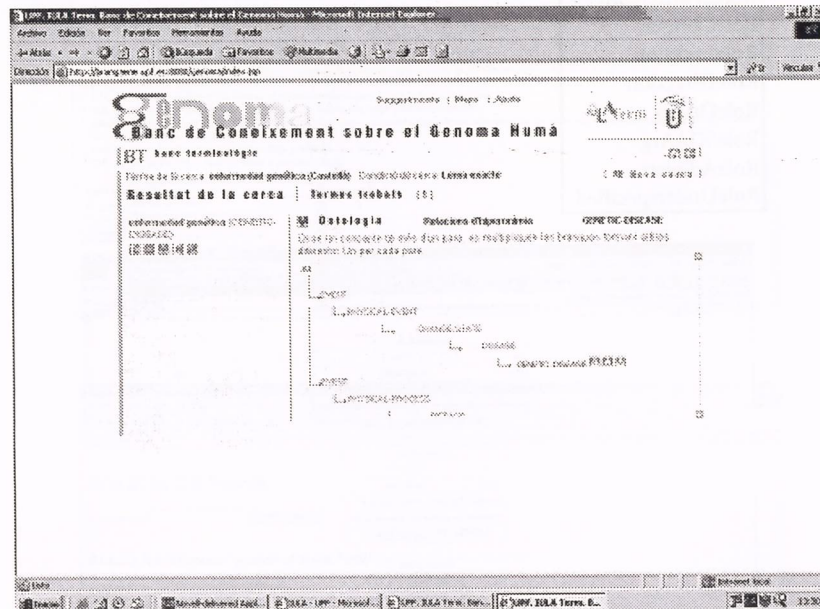


Figure 2: Subset of the SIMPLE Core Ontology representing human beings



Todas as classificações aqui expostas têm sido valorizadas para estabelecer a ontologia do protótipo de banco de conhecimento que se tem desenvolvido no presente projeto TEXTERM2. Esse banco, denominado GENOMA, pode ser consultado em <http://www.iula.upf.edu>. A seguir, como exemplo, reproduz-se o resultado da busca nesse banco do conceito *enfermedad genética* a partir do termo *enfermedad genética*:



3.2.5. Classificação pelo papel lingüístico-conceitual que exercem no discurso especializado

Diferentemente da classe semântica, mas relacionadas com ela, as UCEs representam, na estrutura conceitual de um âmbito, distintos papéis segundo constituam um núcleo ou se integrem em um, ou ainda representem relações entre núcleos, ou ambas as coisas simultaneamente, como já foi descrito em Cabré (2002b).

De acordo com esse critério, as UCEs que fazem parte das estruturas conceituais podem ser descritas em forma de conjuntos de dois traços: o traço *núcleo* e o traço *relação*. Assim teremos as seguintes combinações pertinentes para a representação cognitiva do conhecimento especializado:

[+núcleo –relação]	<i>gen, clon</i>
[+núcleo +relação]	<i>clonación, cruce</i>
[–núcleo+relação]	<i>clonar, cortar</i>

Os núcleos possuem “pesos” conceituais diferentes segundo sua maior proximidade ou afastamento do núcleo central. Quanto mais próximos do núcleo central, maior o peso conceitual; quanto mais distantes, menor o peso conceitual.

Outro dos elementos que condiciona o papel das UCEs em uma estrutura conceitual é sua condição terminológica ou fraseológica: ou são representações de núcleos ou fazem parte de predicções que especificam sua função, sua utilidade, suas características, etc. Desse ponto de vista semântico, são estruturas complexas normalmente descritivas de uma ação exercida pela unidade que ocupa o núcleo mínimo, ou uma expansão ou restrição, ou simplesmente especificação das propriedades desse núcleo. De acordo com esse critério, no qual intervêm ao mesmo tempo condições sobre a estrutura das unidades e condições sobre a relação e alcance dos núcleos, diferenciamos, por um lado, as unidades terminológicas, de estrutura léxica ou sintática, mas neste último caso lexicalizadas, fortemente coesas e representativas de núcleos mínimos, e, por outro lado, as unidades fraseológicas, de estrutura sintática, de grau menor de coesão interna, e representativas de expansões dos núcleos mínimos. Nesse contexto, as orações especializadas representariam combinações livres que conteriam unidades terminológicas e unidades fraseológicas. Existe um desnível entre a estrutura e a condição de núcleo mínimo das estruturas conceituais que provocaram uma grande confusão entre os diferentes tipos de unidades que devem ser retidas, assim como uma grande variação denominativa.

3.2.6. Classificação pela coesão interna

Neste trabalho, como já vimos, consideramos *unidades terminológicas* as unidades de caráter léxico, com entrada no vocabulário do falante enquanto especialista do âmbito, de estrutura morfológica simples, derivada, composta ou sintagmática (entre as quais se incluem as siglas e as formas lexicalizadas abreviadas), semanticamente específicas, que constituem um núcleo mínimo de um âmbito (não de um texto).

Consideramos *fraseologia*, e chamamos, portanto, a unidades fraseológicas, àquelas unidades pertinentes a um âmbito especializado, de estrutura sintática (nominal, verbal, adjetival ou adverbial), de combinação restrita, que representam expansões dos núcleos mínimos. Entre elas, as mais estudadas são as estruturas correspondentes a sintagmas verbais, nas quais as UTs exercem a função de complemento interno na maioria dos casos, embora possam exercer também, mais raramente, a de complemento externo.

Se essas unidades fraseológicas verbais são muito representativas de um âmbito, costumam alternar-se com estruturas sintagmáticas nominais formadas pelo verbo de base, realizado como nome deverbal seguido de um sintagma preposicional que introduz o complemento, geralmente determinado. Com menor frequência, encontramos também outras estruturas que se caracterizam por-

que o léxico é o mesmo, ainda que sob estruturas superficiais distintas (*aislar, aislamiento, aislamiento de proteínas, aislamiento de las proteínas, aislamiento del virus, AC aislado; agente alquilante, alquilar, alquilación*).

Como resumo do exposto, frisamos mais uma vez que somente consideramos UTs as UCEs que cumprem os requisitos de estarem lexicalizadas, ou seja, inseridas na gramática como unidades léxicas, serem usadas em um domínio temático com um sentido definido e serem necessárias na representação conceitual do referido domínio. Essas unidades lexicalizadas converteram-se, de fato, em unidades léxicas para todos os efeitos, mesmo que talvez não tenham todas as possibilidades gramaticais que têm as unidades originariamente léxicas.

As unidades de sistemas não lingüísticos que, à parte sua pertinência a outro tipo de sistema, cumprem todas as demais características das UTs e podem substituí-las em um texto, serão tratadas como *UTs subsidiárias*. Do ponto de vista funcional, atuam como unidades terminológicas (é o caso dos símbolos ou das nomenclaturas); do ponto de vista de suas possibilidades discursivas, apresentam diferenças em relação às unidades lingüísticas, já que, fora haverem sido cunhadas sempre *in vitro*, no exterior do discurso e para finalidades essencialmente normalizadoras, são imunes a propriedades da linguagem natural como a variação (*adenina* ou *agrobacterium*) e à mudança evolutiva. Essas unidades funcionam praticamente sempre como nomes, porque são os nomes as unidades mais prototípicas, tanto pela forma de expressar o conhecimento quanto pela evidência psicológica que delas têm os especialistas, como por seu número e frequência nos discursos especializados, são as unidades de caráter nominal.

4. Conclusões

A exploração de corpora textuais especializados abre o objeto de estudo da terminologia e permite estudar as unidades de conhecimento a partir de diferentes pontos de vista e de diferentes graus de profundidade. Na primeira parte deste trabalho caracterizamos os textos especializados, das Unidades de Conhecimento Especializado, para poder distingui-los dos textos que transmitem conhecimento geral. Descrevemos os textos especializados como unidades que possuem uma tríplice estrutura: formal, cognitiva e gramatical. A estrutura formal permite associar um texto a um gênero textual. A estrutura cognitiva, representada mediante mapas formados por núcleos de conceitos e relações, mostra diferentes níveis do conhecimento em função do grau de abstração do mapa e da dependência dos núcleos selecionados com o objeto de conhecimento principal de um texto. E finalmente, a estrutura gramatical reflete todos os componentes da gramática e como consequência pode-se estudar o texto a partir de qualquer um desses elementos.

Centramo-nos, na segunda parte do texto, nas Unidades de Conhecimento Especializado (UCEs) propondo uma definição e uma topologia de UCE a partir

de uma diversidade de critérios que servem para estabelecer uma classificação multivariante, na qual as UCEs se caracterizam por critérios de diversa natureza. Assim, ao lado de critérios de classificação mais clássicos como a estrutura, a categoria gramatical ou o sistema ao qual pertence a UCE, assinalamos a importância de levar em consideração critérios como a classe conceitual a que pertence a unidade, a coesão interna da unidade e o papel que exerce no discurso especializado, condicionado por uma série de parâmetros: estrutura semântica da unidade; papel de núcleo, de relação ou de núcleo e relação ao mesmo tempo; grau de proximidade do objeto de conhecimento focal do texto ou núcleo principal, etc.

Acreditamos que a multiplicidade de critérios propostos, por um lado, oferece uma visão mais complexa e enriquecedora das unidades que em um texto transmitem conhecimento especializado e, também, aporta uma definição de Unidade Terminológica mais completa; definição que se constrói mediante parâmetros condicionantes que nos permitem discriminar as UTs do resto de UCEs, especialmente das que estão mais próximas gramaticalmente e cognitivamente, como são as unidades fraseológicas especializadas.

Bibliografia

- BEVILACQUA, C. R. *Unidades fraseológicas especializadas: estado de la cuestión y perspectivas*. Barcelona: Institut Universitari de Lingüística Aplicada (IULA), Universitat Pompeu Fabra (UPF), 1999. (Trabalho de pesquisa, inédito)
- BEVILACQUA, C. R. *Unidades fraseológicas especializadas eventivas: descripción y reglas de formación en el ámbito de la energía solar*. Barcelona: Institut Universitari de Lingüística Aplicada (IULA), Universitat Pompeu Fabra (UPF), 2004. (Tese de doutorado inédita)
- CABRÉ, M. T. *Terminology. Theory, Methods and Applications*. SAGER, J. C. (Ed.). Trad. J. DeCesaris. Amsterdam/Philadelphia: John Benjamins Publishing Company, 1999a. 252 p. HB: 90-272-1633-9. PB: 90-272-1634-7.
- CABRÉ, M. T. *La terminología. Representación y comunicación*. Una teoría de base comunicativa y otros artículos. Barcelona: Institut Universitari de Lingüística Aplicada (IULA), Universitat Pompeu Fabra (UPF), 1999b. (Sèrie Monografies, 3).
- CABRÉ, M. T. Elements for a theory of terminology: Towards an alternative paradigm. *Terminology. International Journal of Theoretical and Applied Issues in Specialized Communication*, 6, 1, 2000a, p. 1-23.
- CABRÉ, M. T. Sur la représentation mentale des concepts: bases pour une tentative de modélisation. In: BÉJOINT, H.; THOIRON, P. (Eds.). *Le sens en terminologie*.

- Lyon: Presses Universitaires de Lyon, 2000b, p. 20-39. ISBN: 2-7297-0641-0.
- CABRÉ, M. T. Textos especializados y unidades de conocimiento: metodología y tipologización. In: GARCÍA PALACIOS, Joaquín; FUENTES, M. Teresa (Eds.). *Texto, terminología y traducción*. Salamanca: Ediciones Almar, 2002a, p. 15-36. ISBN: 84-7455-079-3.
- CABRÉ, M. T. Análisis textual y terminología, factores de activación de la competencia cognitiva en la traducción. In: ALCINA CAUDET, A.; GAMERO PÉREZ, S. (Eds.). *La traducción científico-técnica y la terminología en la sociedad de la información*. Castellón: Publicacions de la Universitat Jaume I, 2002b, p. 87-105.
- CABRÉ, M. T. Theories of terminology. Their description, prescription and explanation. In: *Terminology*, 9, 2, 2003a, p. 163-200.
- CABRÉ, M. T. El lenguaje científico desde la terminología. In: GUTIÉRREZ, Bertha. (2003). *Aproximaciones al lenguaje de la ciencia*. Burgos: Fundación Instituto Castellano y Leonés de la Lengua, 2003b, p. 19-52. ISBN: 84-607-7286-1.
- CABRÉ, M. T.; ESTOPÀ, R. On the units specialised meaning uses in professional communication. *Terminology Science and Research*, 2003, 1-2. (no prelo)
- CABRÉ, M. T.; ESTOPÀ, R. El conocimiento especializado y sus unidades de representación: diversidad cognitiva. *Sendébar*, 2002, 13, p. 141-153.
- CABRÉ, M. T.; FELIU, J. (Ed.). *La terminología científico-técnica: reconocimiento, análisis y extracción de información formal y semántica* (DGES PB96-0293). Barcelona: Institut Universitari de Lingüística Aplicada (IULA), Universitat Pompeu Fabra (UPF), 2001. ISBN: 84-477-0744-X.
- CABRÉ, M. T.; LORENTE, M. L'estructura argumental dels predicats derivats. In: *Estudi General*, Revista de la Facultat de Lletres de la Universitat de Girona, 1999, v. 17-18, p. 67-78.
- CHIERCHIA, G.; MCCONELL-GINET, S. *Meaning and Grammar*. An Introduction to Semantics. Cambridge: The MIT Press, 1990.
- CORBIN, D. Hypothèses sur les frontières de la composition nominale. *Cahiers de grammaire*, 17, 1992, p. 26-55.
- CORBIN, D. Locutions, composés, unités polylexématiques: lexicalisation et mode de construction. *Actes du colloque de 1994: La Locution, entre langue et usage*. Textos reunidos por Michel Martins-Baltar. Saint-Cloud :ENS Éditions Fontenay, 1997, p. 1-29.
- DAVID, S. *Les unités nominales polylexicales*. Éléments de description et reconnaissance automatique. Tese de doutorado. Paris: Université Denis Diderot, 1993.
- ESTOPÀ, R. Les unités de signification spécialisées: élargissant l'objet du travail en terminologie. *Terminology*, 7, 2, 2001, p. 217-237.
- ESTOPÀ, R. Los adjetivos en las unidades terminológicas poliléxicas: un análisis morfosemántico. *Organon*, 14, 28/29, 2001, p. 233-246.
- ESTOPÀ, R. *Extracció de la terminologia: elements per a la construcció d'un SEACUSE*. 2003 [1999]. CD-ROM. 84-89782-02-4.

- ESTOPÀ, R. *Les unitats terminològiques polilèxemàtiques en els lèxics especialitzats: dret i medicina*. Barcelona : Institut Universitari de Lingüística Aplicada (IULA), Universitat Pompeu Fabra (UPF), 1996. (Trabalho de pesquisa de doutorado, inédito)
- FELIU, J. *Relacions conceptuals i variació funcional: elements per a un sistema de detecció automàtica*. 2000. (Trabalho de pesquisa, inédito)
- FELIU, J.; SOLÉ, E.; TEBÉ, C. CABRÉ, M. T. Las relaciones conceptuales: un elemento esencial en la estructuración del conocimiento especializado. In: *Actas del VIII Simposio Iberoamericano de Terminología*. 28-31/10/2002, Cartagena de Indias. 2002. ISBN: 958-33-4022-7.
- FELIU, J.; VIVALDI, J.; CABRÉ, M. T. Ontologies: a review. *Working Paper*, 34. Barcelona: Institut Universitari de Lingüística Aplicada (IULA), Universitat Pompeu Fabra (UPF), 2002. DL: 23.735-2002.
- FELIU, J.; VIVALDI, J.; CABRÉ, M. T. Towards an Ontology for a Human Genome Knowledge Base. LREC2002. In: *Third International Conference On Language Resources And Evaluation*, 05/2002, Las Palmas de Gran Canaria. *Proceedings*. Las Palmas de Gran Canaria: 2002, p. 1885-1890. ISBN: 295-1740-808.
- FELLBAUM, C. (Ed.). *WordNet*. An Elechtronical Database. Cambridge: The MIT Press, 1998.
- JUNYENT, C. La transmisión del conocimiento especializado. Conocimiento y conceptos. In: CABRÉ, M. T.; FREIXA, J.; TEBÉ, C. *Terminología y conocimiento especializado*. Barcelona: Institut Universitari de Lingüística Aplicada (IULA), Universitat Pompeu Fabra (UPF), 2003.
- KAUFMANN, U.; BERGENHOLTZ, H. *Diccionario enciclopédico de Ingeniería genética*. Toronto: Lugus Libros LatinAmerica Inc, 1998, 2 vols.
- LORENTE, M. Verbos y discurso especializado. In: *Estudios de Lingüística Española* (ELIES), 16, 2002. Publicação eletrônica disponível em: < <http://elies.rediris.es> >.
- LORENTE, M. Procesos de creación del léxico con significado especializado: unidades verbales en el discurso del Genoma Humano. In: BISELTO, A.; IACOBINI, C.; THORTON, A.M. *Seriti di Morfologia*. In onore di Sergio Scalise. Roma: Caissa Italia Editore, 2003, p. 129-148.
- PUSTEJOVSKY, J. *The Generative Lexicon*. Cambridge/Massachussets: The MIT Press, 1995.
- SAGER, J. C.; KAGEURA, K. Concept Classes and Conceptual Structures: Their Role and Necessity. In: *Alfa*, 7/8, 1994/1995, p. 191-216.