

HPSG: arquitectura

António Branco e Francisco Costa

Introdução

HPSG é o acrónimo da expressão inglesa *Head-Driven Phrase Structure Grammar*, que designa um enquadramento epistemológico para o estudo da linguagem natural humana e em particular para o estudo científico daquilo que, na sequência de uma tradição milenar, se designa por gramática.

O presente capítulo é o primeiro de um conjunto de três capítulos cujo objectivo é apresentar uma introdução à HPSG.

Nos dois capítulos seguintes é progressivamente descrito o conteúdo de uma pequena gramática de um fragmento do Português, com o propósito de se ilustrar aspectos chave da HPSG. O Capítulo seguinte, sob o título "Representação Sintáctica em HPSG", ocupa-se essencialmente de aspectos relativos à sintaxe. O Capítulo subsequente, com o título "Representação do Significado em HPSG", por sua vez, ocupa-se da semântica. A gramática é aí ampliada com dispositivos que permitem representações do significado para expressões arbitrariamente complexas.

No presente capítulo são descritos os primitivos conceptuais, as assunções fundacionais, as bases empíricas e os dispositivos descritivos e computacionais básicos.

Na segunda Secção, sobre *Assunções Fundacionais*, descrevem-se os objectivos principais da HPSG: caracterização explícita e precisa do conhecimento linguístico, capacidade preditiva, neutralidade procedimental, processamento incremental, tratabilidade computacional, distinção entre princípios gramaticais universais e generalizações específicas a cada língua, e capacidade de integração da informação linguística com outras fontes de informação.

A terceira Secção, sobre *Bases Empíricas*, discorre sobre a obtenção de dados empíricos que alimentam a indução de generalizações gramaticais a ser incluídas na teoria.

Nas quarta e quinta Secções, sobre *Estruturas de Modelação* e *Meios Descritivos*, respectivamente, apresenta-se o formalismo utilizado para modelar os fenómenos linguísticos: estruturas de traços e hierarquias de tipos.

Na sexta Secção, sobre *Dispositivos Computacionais*, fala-se das operações computacionais de maior relevo assumidas pela gramática.

A terminar, a Secção 7 apresenta referências para algumas obras introdutórias, assim como outras fontes bibliográficas de interesse.

Assunções Fundacionais

Na HPSG é assumido que a faculdade da linguagem é uma capacidade cognitiva específica, com autonomia considerável em relação a outras capacidades cognitivas dos sujeitos humanos. É assumido também que os estudos gramaticais têm por objectivo a descoberta do conhecimento linguístico subjacente a essa capacidade cognitiva. A gramática é por conseguinte entendida como uma teoria explícita acerca desse conhecimento, que assegura uma caracterização precisa do mesmo.

Enquanto teoria, a gramática apresenta capacidade preditiva: através de uma caracterização finita, representa o conhecimento potencial que um sujeito detém acerca de um número eventualmente ilimitado de putativas entidades linguísticas. Esse conhecimento abrange quer entidades linguísticas pertencentes à linguagem específica de que o sujeito é um utilizador competente (e.g. Português, Chinês, Polaco, etc), quer para-entidades linguísticas não pertencentes a essa linguagem.

Não obstante a sua especificidade, a faculdade da linguagem é uma de entre as diversas faculdades cognitivas. Nesta medida, é concebida no quadro das assunções fundacionais da Ciência Cognitiva, em que os processos cognitivos são entendidos como tendo por base o processamento de informação cujo melhor modelo até à data é justificado pelo Resultado Fundamental da Teoria da Computação. Deste modo, a gramática de uma linguagem L é uma teoria compatível com um algoritmo de parsing que, para cada putativa entidade linguística, permite decidir se essa entidade pertence a L .

Tratando-se de uma teoria do conhecimento linguístico, a gramática oferece a melhor compatibilidade com a modelação do comportamento linguístico e em particular com o processamento de informação linguística. Nesta medida, a gramática é compatível com um algoritmo de parsing se não eficiente em termos gerais do ponto de vista da complexidade computacional, pelo menos tratável dentro das condições típicas da execução do parser mental natural.

Por outro lado, para ser compatível com diferentes regimes de processamento linguístico, e.g. compreensão, produção ou tradução, etc, a descrição do conhecimento gramatical é feita com neutralidade procedimental através da utilização de um dispositivo descritivo declarativo.

Adicionalmente, a gramática é compatível com modelos de processamento parcial de entidades linguísticas e com a articulação flexível entre os seus sub-módulos (i.e. fonologia, morfologia, sintaxe, semântica, pragmática). Nesta medida, o dispositivo descritivo que usa permite também a integração flexível das contribuições funcionais dos diferentes sub-módulos de forma incremental.

Em face de aspectos comuns que se abstraem das gramáticas de diferentes línguas naturais, neste enquadramento epistemológico é assumido que existe um núcleo de conhecimento gramatical comum às diferentes linguagens. Esse conhecimento é designado por Gramática Universal. Cada gramática de um idioma individual integra a

descrição desse conhecimento universal comum e apresenta uma sua especificação e extensão particular, a qual constitui o conhecimento específico desse idioma particular. Deste modo, a gramática é compatível com uma leque de modelos de aquisição individual da linguagem, incluindo os que advogam alguma forma de inatismo.

Quando instanciado numa dada entidade linguística particular, o conhecimento gramatical associa a forma dessa entidade (viz. marcas simbólicas sonoras) e o seu respectivo significado — ou os seus respectivos significados, no caso de haver ambiguidade para a qual o contexto não é suficiente para possibilitar a sua resolução por retenção apenas do significado intencionado. Por conseguinte, a par da descrição precisa do conhecimento gramatical, a gramática garante a descrição precisa da informação linguística associada a cada entidade linguística em que esse conhecimento se instancia. Desta maneira, permite associar a representação da forma linguística e a do respectivo significado, no que se designa de representação gramatical.

As representações gramaticais são compatíveis com a integração da representação de informação especificamente linguística com a representação do conhecimento genérico, não-linguístico, do mundo. Permitem também a integração da representação do significado com modelos de processamento inferencial. Nesta medida, a gramática é compatível com o conhecimento não linguístico e com a dimensão regular do raciocínio.

Bases Empíricas

A base empírica para a construção de uma gramática de uma linguagem L é a base habitual do trabalho de investigação sobre gramática, a qual vem na linha de uma tradição milenar. Os dados empíricos são recolhidos através de acesso introspectivo a juízos acerca da pertença de putativas entidades linguísticas a L , também referidos como juízos de gramaticalidade ou de aceitabilidade.

A assunção aqui é a de que o investigador que executa esses juízos é um falante plenamente competente de L e a de que, em larga medida e na extensão relevante para a construção da gramática, a sua competência gramatical é idêntica à dos restantes falantes de L . A validade e produtividade desta metodologia é assegurada pela dimensão colectiva do trabalho científico, a qual neste caso se traduz na avaliação crítica dos dados empíricos recolhidos por parte de outros investigadores, também falantes de L , que confirmam se os dados utilizados coincidem com os juízos obtidos por si relativamente às mesmas entidades.

Para melhor aceder a esses juízos e aclarar as suas implicações, estes podem ser executados sobre pares de entidades linguísticas. Com o objectivo de ajudar a isolar determinadas propriedades linguísticas específicas, é comum procurar-se que as entidades assim colocadas em contraste difiram entre si de forma mínima e apenas em resultado de diferenças relativas a essas propriedades sob investigação.

Outra fonte de dados empíricos encontra-se nos repositórios de produções linguísticas conhecidos como corpora. As entidades linguísticas que se encontram nos corpora ajudam a sugerir dados empíricos que, sem o recurso à análise destes repositórios, poderiam escapar a um escrutínio baseado apenas na consideração de entidades linguísticas obtidas por manipulação mental com vista a se construir contrastes mínimos.

Estruturas de Modelação

Sendo uma teoria sobre um domínio empírico, a gramática e as entidades desse domínio são mediados por uma estrutura matemática que serve de modelo deste último. Esta estrutura e as suas partes estão em correspondência com as entidades observáveis por forma a que estas se encontrem representadas por entidades da estrutura simbólica que as modela. Neste arranjo ontológico tripartido — observáveis, modelo e teoria —, a gramática discorre acerca das entidades linguísticas discorrendo acerca das suas representações, sendo assim interpretada no modelo simbólico do seu domínio empírico.

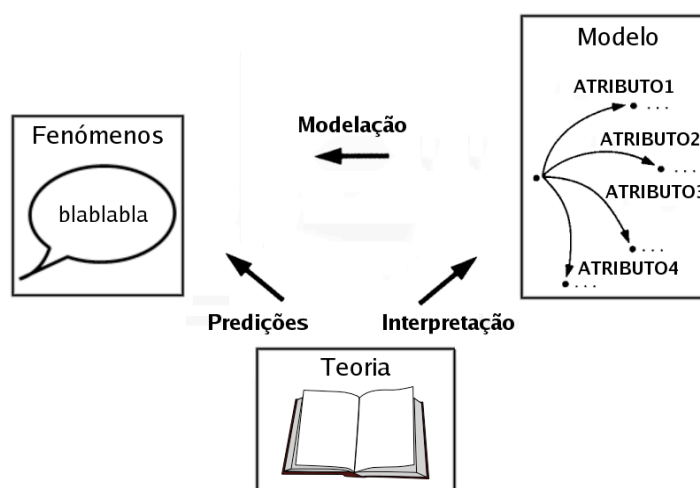


Fig. 1: Diagrama esquemático das relações entre observáveis, modelo e teoria. Exemplo adaptado de (Pollard e Sag 1994, p. 9).

Os observáveis relevantes do domínio a modelar são os tipos de expressões da linguagem natural assim como os das suas subexpressões. Um tipo de expressão, entre muitos outros, será por exemplo aquele que se representa por "mesa", o qual se obtém por abstracção a partir das diferenças acidentais presentes nos enunciados particulares de "mesa".

Para representar os observáveis no modelo, são usadas hierarquias de tipos, as quais são ordens parciais, e estruturas de traços (*feature structures*), que são entidades definidas na teoria dos grafos.

Uma hierarquia de tipos é um grafo acíclico em que os nós são etiquetados com etiquetas de tipos e em que um nó s_2 onde chega um arco que sai de um nó s_1 é um subtipo de s_1 , e por isso mais específico que este último.

Uma estrutura de traços é um grafo em que os nós são etiquetados com etiquetas de tipos e os arcos com etiquetas de atributos. Por conseguinte, de uma dada estrutura de traços não atômica de tipo s — i.e. um grafo cujo nó de topo está etiquetado com s —, diz-se que tem atributos — aqueles que etiquetam os arcos que saem do nó s . Por sua vez, de um atributo a , diz-se que tem um valor, que é uma outra estrutura de traços (atômica ou não) para que o arco etiquetado com a aponta.

Dado que são modelos totais de entidades linguísticas, as estruturas de traços têm de ser (i) totalmente bem-tipadas e (ii) tipo-resolvidas. Em termos informais, isto impõe que: (i) para cada nó s no grafo, cada arco que devia sair de s por ser apropriado para caracterizar s está de facto presente; (ii) cada nó é etiquetado com o tipo s que é maximamente específico na hierarquia de tipos que s integra.

Uma propriedade crucial das estruturas de traços é que dois caminhos distintos no grafo podem levar a um mesmo nó. Nesta medida, a partilha de estrutura é possível ou seja, é possível dois caminhos no grafo terem uma mesma estrutura como seu valor comum.

Neste enquadramento, a teoria é uma especificação a ser interpretada no modelo do domínio constituído por estruturas de traços. Essa especificação é um conjunto de restrições que definem uma caracterização intensional finita de um conjunto eventualmente infinito de entidades desse modelo. Nesta medida, essas restrições estabelecem a predição de quais entidades de entre as que estão disponíveis no modelo (e através da sua mediação, que entidades de entre as que se encontram no domínio empírico) são entidades que pertencem à linguagem natural para a qual a teoria em questão é uma gramática.

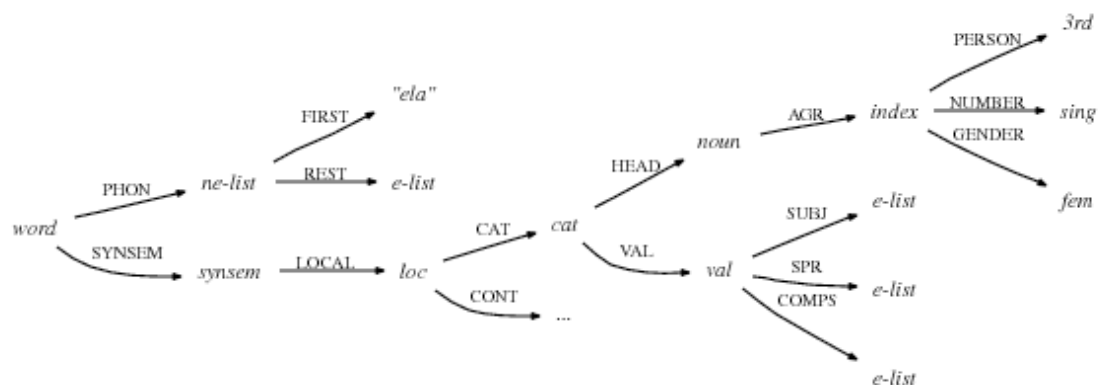


Fig. 2: Exemplo de um grafo que representa uma entidade linguística no modelo, no caso a palavra "ela" (o valor da maior parte das etiquetas é apresentado no Capítulo "Representação Sintáctica em HPSG"). Exemplo adaptado a partir de (Pollard e Sag 1994, p. 17).

Meios Descritivos

De um ponto de vista algébrico, uma gramática é uma especificação, que comporta uma assinatura e uma apresentação.

A assinatura define que tipos de dados estão disponíveis para representar entidades linguísticas. Inclui uma hierarquia de tipos e uma definição de adequação.

A hierarquia de tipos é uma ordem parcial de tipos em que os tipos possíveis para entidades linguísticas e seus subtipos são estabelecidos.

A definição de adequação, também designada por declaração de traços, estabelece quais são as características apropriadas para cada tipo na hierarquia. Consiste em associar a cada tipo restrições que definem quais as estruturas de traços apropriadas para esse tipo de entidade. A hierarquia de tipos é uma árvore taxonómica em que cada tipo herda as restrições de adequação dos seus supertipos.

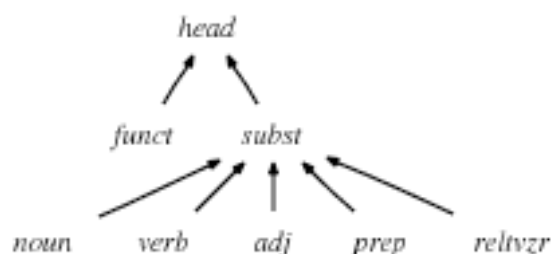


Fig. 3: Exemplo de uma hierarquia de tipos, no caso relativa à hierarquia dos subtipos de núcleo (*head*), que inclui no nível imediatamente inferior núcleo funcional (*funct*) e núcleo substantivo (*subst*), o qual tem como subtipos preposição (*prep*), relativizador (*reltvzr*), nome (*noun*), adjetivo (*adj*) e verbo (*verb*).

A apresentação é um conjunto de restrições implicacionais que são interpretadas (como sendo verdadeiras ou falsas) no modelo. São entendidas como descrições que circunscrevem o conjunto das estruturas de dados que, obedecendo ao formato dos tipos de dados estabelecido na assinatura, a teoria prediz serem entidades da linguagem natural em questão.

As restrições que integram a gramática são expressas numa linguagem de descrição específica cujas expressões são matrizes de atributos-valores (MAV). Numa apresentação abreviada da sua sintaxe, pode dizer-se que as MAVs não atómicas consistem em matrizes de duas colunas, em que na primeira coluna se encontram os atributos e na segunda coluna os valores correspondentes. Os valores dos atributos são MAVs. As matrizes recebem um subscripto ou superscrito à esquerda que indica o seu tipo. Uma MAV atómica consiste apenas num tipo atómico.

$$\begin{bmatrix} \textit{tipo} \\ \text{ATRIBUTO}_1 & \textit{mav}_1 \\ \dots & \dots \\ \text{ATRIBUTO}_n & \textit{mav}_n \end{bmatrix}$$

Fig. 4: Exemplo esquemático de uma matriz de atributos-valores (MAV) ou de traços.

Existe uma simbologia reservada para indicar partilha de estrutura. Quando cada um de dois atributos tem por valor uma etiqueta formada por uma cercadura em que se inscreve um número, e esses dois números são idênticos, isso indica que esses atributos têm o mesmo valor, ou seja que os seus valores são uma e a mesma instância. Essas etiquetas podem ser seguidas por MAVs, caso em que se está a expressar qual o valor a que a etiqueta corresponde.

As descrições de conjuntos são feitas da forma esperada, através do uso de chavetas em torno dos seus elementos, separados por vírgulas, e as descrições de listas são feitas através do uso similar de parênteses angulares.

Cabe salientar que as MAVs expressam restrições sobre as entidades (totais) do universo do modelo, ou seja são descrições (parciais) das entidades para cuja caracterização contribuem.

Dependendo da natureza da informação linguística que captam, é costume designar as restrições constantes da apresentação por princípios gramaticais ou entradas lexicais.

A gramática de uma dada linguagem natural inclui um conjunto de princípios que se assume, por generalização indutiva, serem parte da gramática de qualquer linguagem natural. Estes princípios são designados como princípios da Gramática Universal.

$$\begin{bmatrix} \textit{headed-phrase} \\ \text{SYNSEM|LOCAL|CAT|HEAD [1]} \\ \text{HEAD-DTR} \begin{bmatrix} \textit{sign} \\ \text{SYNSEM|LOCAL|CAT|HEAD [1]} \end{bmatrix} \end{bmatrix}$$

Fig. 5: Exemplo de um princípio gramatical, no caso o Princípio do Traço do Núcleo (*Head Feature Principle*), que impõe que se se trata de um sintagma com núcleo (*headed-phrase*), então os valores do atributo HEAD do constituinte que é o núcleo e os do sintagma são idênticos. Os valores do atributo HEAD são do tipo *head*, definidos na hierarquia da Fig. 3. Duas etiquetas

idênticas denotam identidade entre instâncias. Neste caso, a identidade entre o atributo HEAD do nó mãe e o atributo HEAD do núcleo (a HEAD-DTR) é representada pelas duas etiquetas $\boxed{1}$.

Juntamente com estes princípios, cada gramática para um dado idioma particular inclui princípios que valem apenas para esse idioma ou para um conjunto ou família de línguas a que esse idioma pertence. Todos estes princípios integram a gramática como uma conjunção de restrições, que qualquer entidade linguística tem de satisfazer.

A par do conjunto dos princípios gramaticais, as restrições que correspondem a entradas lexicais formam o conjunto dos princípios satisfazíveis disjuntivamente.

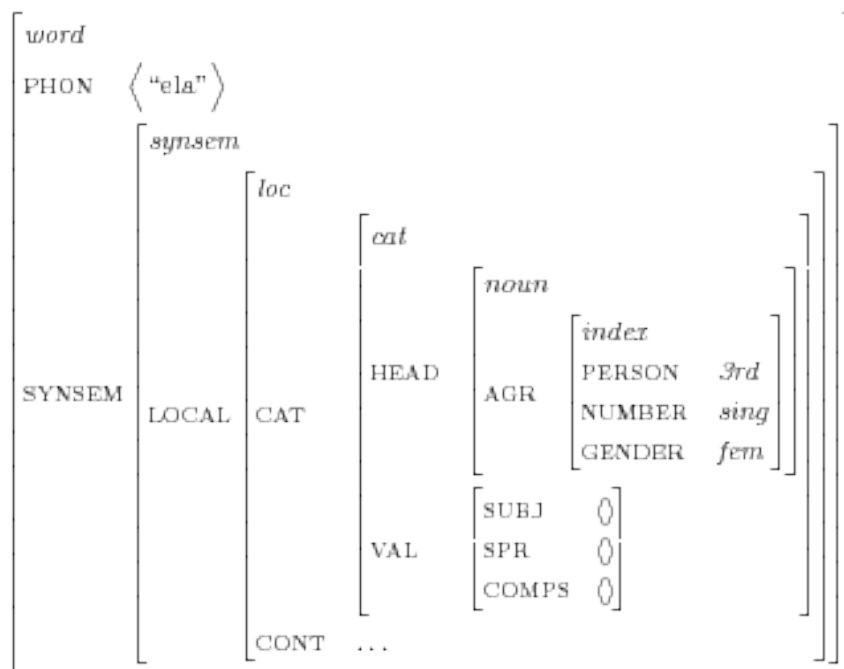


Fig. 6: Exemplo da descrição de uma entrada lexical, no caso da palavra "ela".

Outras restrições disjuntivas são também incluídas na gramática e são conhecidas como esquemas (para relações de dominância e de precedência).

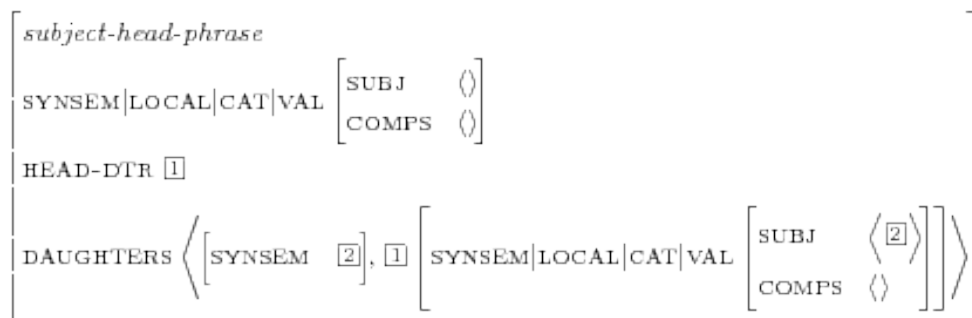


Fig. 7: Exemplo de um Esquema, no caso o Esquema Núcleo-Sujeito (*subject-head-phrase*), que entre outras coisas impõe que os atributos SUBJ e COMPS se encontram vazios, o que corresponde à circunstância de que um sintagma com o sujeito realizado não dispõe de capacidade de subcategorizar qualquer outro argumento.

Finalmente, um outro conjunto ainda de restrições, conhecidas como regras lexicais, inclui restrições com propriedades formais distintas, usadas para captar generalizações acerca de entradas lexicais. Uma regra lexical é um par de "meta-descrições" visto relacionar entidades da especificação (descrições) e não entidades do modelo (estruturas de traços). Ao se aplicar uma regra lexical, a restrição que é mencionada no seu lado esquerdo é usada para circunscrever as estruturas de traços que a satisfazem, às quais supletivamente se aplicam as restrições mencionadas no lado direito da regra.

$$\left[\begin{array}{l} \text{PHON } \boxed{1} \\ \text{SYNSEM } \boxed{2} \left[\text{LOCAL} | \text{CAT} | \text{HEAD} | \text{AGR} | \text{NUMBER } \textit{sing} \right] \\ \text{DAUGHTERS } \left\langle \begin{array}{l} \textit{noun-lexeme} \\ \text{PHON } \boxed{1} \\ \text{SYNSEM } \boxed{2} \end{array} \right\rangle \end{array} \right]$$

Fig. 8: Exemplo de uma Regra Lexical, no caso a que produz a forma singular dos nomes. A unificação dos atributos PHON modela o facto de que a forma singular de um nome é idêntica ao seu lema.

Na Fig. 9 encontra-se uma sinopse das diferentes classes de restrições que integram a gramática de uma linguagem L .

$\text{GU} = P_1 \wedge \dots \wedge P_n$	
$\text{Linguagem } L = \text{GU} \wedge P_{n+1} \wedge \dots \wedge P_{n+m} \wedge (S_1 \vee \dots \vee S_p \vee L_1 \vee \dots \vee L_q)$	
<i>Em que:</i>	
P_1, \dots, P_n	são princípios da Gramática Universal (GU)
P_{n+1}, \dots, P_{n+m}	são princípios específicos de L
S_1, \dots, S_p	são esquemas agrupados em princípios específicos de L
L_1, \dots, L_q	são signos lexicais de L (básicos ou o resultado de regras lexicais)

Fig. 9: Sinopse das diferentes classes de restrições implicacionais que integram uma gramática de uma linguagem natural L .

Um exemplo acessível de uma gramática HPSG para um fragmento de uma língua natural, no caso o Inglês, encontra-se nas 9 páginas do Apêndice de (Pollard e Sag 1994).

Dispositivos Computacionais

Tal como é descrito na gramática, o conhecimento linguístico constitui uma teoria com capacidade preditiva. Essa capacidade é tornada operacional no desempenho de duas tarefas computacionais chave: (i) síntese (ou geração): dada uma representação semântica, obter o conjunto de frases cuja representação gramatical inclui essa representação semântica, ou seja, a lista de frases com significado idêntico ao que é expresso pela representação semântica de partida; (ii) análise: dada uma expressão, verificar se essa expressão pertence à linguagem em questão, caracterizada pela gramática, e em caso afirmativo obter o conjunto das suas representações gramaticais.

Diversas linguagens de programação, também conhecidas por formalismos de implementação, têm sido desenvolvidas de tal forma que uma gramática pode ser escrita como um programa numa dessas linguagens, o qual é executável num computador. Estas linguagens de programação oferecem meios descritivos que se aproximam dos dispositivos descritivos definidos na HPSG, ao mesmo tempo que asseguram que a execução dos seus programas permitem o desempenho das tarefas acima referidas.

Dois procedimentos computacionais desempenham um papel nuclear na execução desses formalismos de implementação, o mecanismo de unificação e o de *parsing*.

O mecanismo de parsing permite decidir se uma expressão pertence ao conjunto das expressões definidas intensionalmente por uma gramática como formando uma linguagem.

Em termos muito genéricos, e no quadro dos formalismos de implementação, o mecanismo de unificação permite a construção de representações através da combinação de representações parciais compatíveis entre si e da eventual instanciação de variáveis nessas representações.

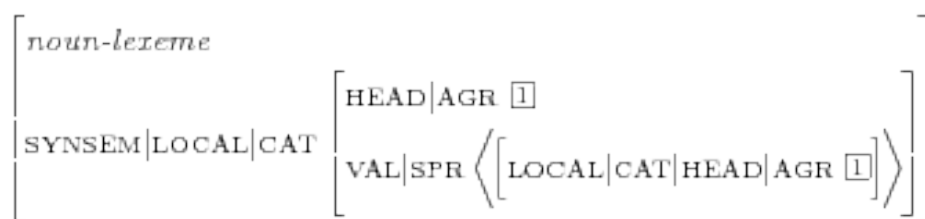


Fig. 10: Exemplo de uma operação de unificação, no caso a que impõe que os traços de género e número de um nome (representados sob o atributo AGR) devem ser idênticos aos traços de género e número do seu especificador (no traço SPR).

No seguimento da tarefa de análise, alguns formalismos de implementação permitem a execução automática de uma terceira tarefa, dita de resolução de ambiguidade. Essa tarefa consiste em indicar, de entre as representações gramaticais possíveis para uma dada expressão de partida, aquela que mais provavelmente é veiculada por essa expressão. Para esta finalidade, são usados procedimentos

automáticos de classificação baseados em parâmetros estocásticos obtidos a partir de corpora, conhecidos como bancos de árvores (*treebanks*), em que cada frase foi previamente anotada com a representação gramatical mais favorecida pelo seu contexto.

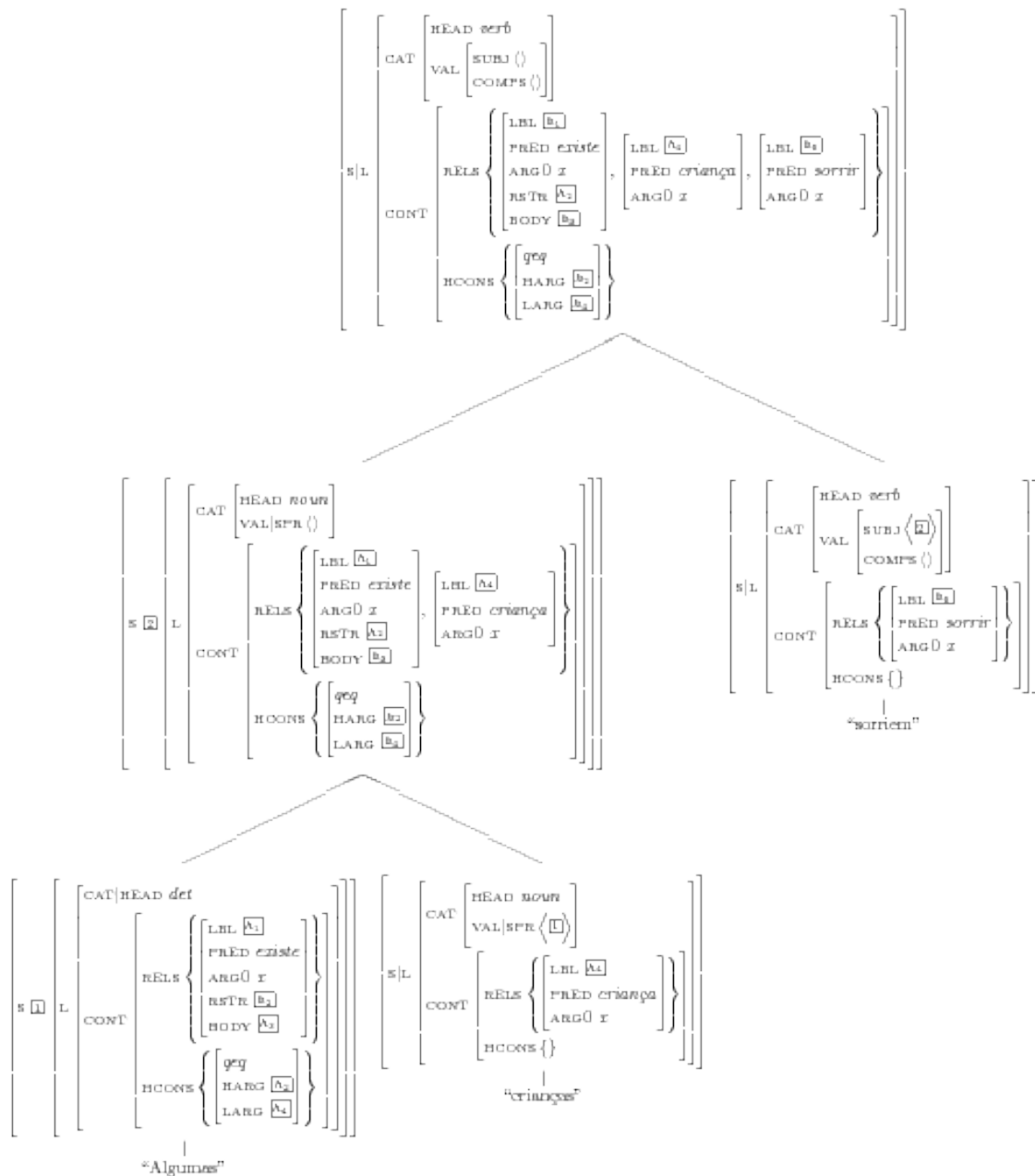


Fig. 11: Exemplo de uma representação gramatical abreviada, no caso a que é relativa à frase "Algumas crianças sorriem". O atributo SYNSEM está abreviado como S, e LOCAL como L. Este exemplo ilustra a forma como a subcategorização é codificada na HPSG, através dos atributos dentro de VAL: SUBJ e SPR (*vide* Capítulo "Representação Sintáctica em HPSG"). Também

ilustra a representação semântica atribuída aos vários constituintes, dentro do atributo CONT: RELS é uma colecção de relações e HCONS restringe as possíveis relações de âmbito entre elas (*vide* Capítulo "Representação do Significado em HPSG").

Leituras Subsequentes

Para uma introdução ao estudo científico da gramática, a partir de um nível básico e sem pré-requisitos, (Kim e Sells 2008) é um excelente manual pedagógico que progressivamente expõe conceitos linguísticos chave, exercita técnicas nucleares de investigação e argumentação e usa meios formais de modelação e descrição. Para o leitor que procura uma exposição mais alongada e que não se inicie a partir de um nível tão básico, (Sag *et al.* 2003) é um ponto de partida alternativo.

O enquadramento epistemológico da HPSG encontra-se definido nos dois volumes fundacionais de (Pollard e Sag 1987, 1994), os quais constituem leituras complementares e de aprofundamento em relação a estas duas obras de introdução referidas acima.

Os leitores de um nível avançado têm como fonte principal de documentação o repositório bibliográfico online da HPSG, com acesso livre à maior parte das publicações aí listadas, que se encontra em (HPSG Página Web).

As assunções fundacionais da HPSG têm a sua base nos pressupostos epistemológicos que na segunda metade do século XX estiveram na origem e amadureceram sob a forma de Ciência Cognitiva, sendo que de entre estes, têm especial relevo os pressupostos que emergiram no âmbito dos estudos sobre a chamada gramática formal. Os manuais de (Friedenberg e Silverman 2006) e de (Fromkin 2000) são introduções conceituadas a cada uma destas áreas e oferecem um amplo leque de referências a outras obras de maior profundidade e especialidade.

A obra de referência para os meios descritivos explorados na HPSG é a de (Carpenter 1992). Para uma panorâmica sobre questões relacionadas com os fundamentos computacionais do processamento de linguagem natural, incluindo os que se aplicam no quadro da HPSG, veja-se (Branco 2005) e as referências para que remete. Uma introdução à utilização do mecanismo de unificação em gramáticas encontra-se em (Shieber 1986).

O desenvolvimento de dispositivos computacionais para gramáticas da família da HPSG foi alvo de diversas iniciativas durante a última década do século XX, documentadas em (Uszkoreit *et al.* 1996). Dessas contribuições, destacam-se os sistemas de implementação TRALE (TRALE Página Web) e LKB (LKB Página Web), por o seu desenvolvimento continuar a ser mantido.

De entre estes sistemas, cabe assinalar o LKB (*Linguistic Knowledge Builder*) por um conjunto de factores de que a seguir se destacam alguns dos mais importantes. Trata-se de software cujo acesso é gratuito, encontra-se apropriadamente documentado, oferece uma interface gráfica amigável em relação ao utilizador e apresenta níveis de eficiência de desempenho que permitem a sua utilização em aplicações computacionais comerciais.

Existe uma obra de introdução acessível ao desenvolvimento de gramáticas computacionais através da utilização do LKB (Copestake 2002), com pré-requisitos mínimos relativamente às competências dos leitores.

O LKB é utilizado pela maior comunidade de investigadores em engenharia de gramáticas para o processamento linguístico profundo, reunidos no consórcio Delph-In (DELPH-IN Página Web).

LXGram é uma gramática computacional para o Português desenvolvida pela Universidade de Lisboa sobre o LKB no âmbito deste consórcio. Encontra-se disponível em (LXGram Página Web).

Referências

BRANCO, António (2005). “Cognitive Science and its Computational Foundations: A Natural Language Perspective”, in BRANQUINHO, J. (org.), *Cognition and Content*. Lisboa, Centro de Filosofia da Universidade de Lisboa.

CARPENTER, Bob (1992). *The Logic of Typed Feature Structures*. Cambridge, Cambridge University Press.

COPESTAKE, Ann (2002). *Implementing Typed Feature Structure Grammars*. Stanford, CSLI Publications.

DELPH-IN (Página Web). Página com a apresentação da iniciativa Deep Linguistic Processing with HPSG. <http://www.delph-in.net>.

FRIEDENBERG, Jay e SILVERMAN, Gordon (2006). *Cognitive Science: An Introduction to the Study of Mind*. Thousand Oaks, Sage Publications.

FROMKIN, Victoria (2000). *Linguistics: An Introduction to Linguistic Theory*. Oxford, Blackwell.

HPSG (Página Web). The HPSG Bibliography. Repositório de publicações de trabalhos de investigação realizados no quadro da HPSG. <http://hpsg.fu-berlin.de/HPSG-Bib>.

KIM, Jong-Bok e SELLS, Peter (2008). *English Syntax: An Introduction*. Stanford, CSLI Publications.

LKB (Página Web). Página que resume o procedimento para obter e instalar o software LKB. <http://wiki.delph-in.net/moin/LkbInstallation>.

LXGram (Página Web). Sítio da Web em que se encontra disponível a gramática LXGram para o português, <http://nlxgroup.di.fc.ul.pt/lxgram>.

POLLARD, Carl e SAG, Ivan (1987). *Information-Based Syntax and Semantics*, Vol. 1. Stanford, CSLI Publications.

_____ (1994). *Head-Driven Phrase Structure Grammar*. Stanford, Chicago University Press and CSLI Publications.

SAG, Ivan A., WASOW, Thomas e BENDER, Emily M. (2003). *Syntactic Theory – A Formal Introduction*. Stanford, CSLI Publications.

SHIEBER, Stuart (1986). *An Introduction to Unification-Based Approaches to Grammars*. Stanford, CSLI Publications.

TRALE (Página Web). Página que resume o procedimento para obter e instalar o software TRALE. <http://www.sfs.uni-tuebingen.de/hpsg/archive/projects/trale>.

USZKOREIT, Hans, BACKOFEN, Rolf, CALDER, Jo, CAPSTICK, Joanne, DINI, Luca, DÖRRE, Jochen, ERBACH, Gregor, ESTIVAL, Dominique, MANANDHAR, Suresh, MINEUR, Anne-Marie e OEPEN, Stephan (1996). *The EAGLES Formalisms Working Group – Final Report Expert Advisory Group on Language Engineering Standards*. Saarbrücken, Relatório Técnico LRE 61-100 do DFKI – Deutsches Forschungszentrum für Künstliche Intelligenz. <http://www.dfki.de/lt/publications/show.php?id=563>.