



**JEL** UERJ  
Jornadas de Estudos da Linguagem  
02 a 04 de DEZEMBRO de 2010



### *Linguagem e Gramática nos Estudos Formais*

Livy Maria Real Coelho (UFPR/CAPES)

Ricardo Luiz de Andrade Abrantes (USP/CAPES)

livyreal@gmail.com

randrade@ime.usp.br

Considerando o crescimento da interface entre a linguística e a ciência da computação, entendemos ser necessário discutir a respeito dos objetos de estudo e da nomenclatura assumida por cada uma destas áreas. Para isso, propomos abordar de que forma são compatíveis as definições de *linguagem* e *gramática* utilizadas pela ciência da computação, a saber pela teoria de autômatos e linguagens formais, e pela linguística formal, particularmente de base montagoviana. Desconsideraremos aqui outras perspectivas da linguística, como a estruturalista e a funcionalista, por fugirem já do escopo deste trabalho que pretende discutir teorias exclusivamente formais. Para a ciência da computação, sob o rótulo *gramática* entende-se o conjunto composto por variáveis, terminais (também chamados de *constantes*), regras de produção e um símbolo de início. O símbolo de início pode ser entendido como o axioma primeiro da gramática; os terminais, como elementos do alfabeto; as regras de produção, como o modo a que se chega do símbolo inicial aos terminais; e as variáveis, como abstrações que tornam a aplicação destas regras possíveis. Desta forma, a produção gramatical se inicia com o símbolo de início e, através das regras de produção e de suas aplicações (utilizando variáveis, se necessário), termina com a produção de um subconjunto de *strings* da linguagem, que são compostos por terminais. Já *linguagem*, ainda na ciência da computação, é tida por um subconjunto do conjunto de todos os *strings* de um alfabeto específico. Um *alfabeto* é um conjunto de símbolos finito e não vazio; uma *string*, também chamada de *palavra* ou *cadeia*, é uma sequência finita de símbolos escolhidos a partir de um alfabeto. Desta forma, na ciência da computação, *linguagem* pode ser definida como os *strings* formados a partir dos elementos da gramática, que são, basicamente, regras de boa formação e léxico (*strings* formados por símbolos terminais). Já na linguística formal, *grosso modo*, *linguagem* é um termo mais abrangente que *gramática*, sendo a gramática parte da linguagem. Normalmente, gramática é o conjunto de regras capaz de originar uma parte da linguagem, em geral, a estrutura sintática. Há ainda na linguagem, elementos de fundamental importância para o linguista, como o léxico (comparável aos terminais da ciência da computação) e a ontologia da língua, isto é, elementos abstratos dos quais precisamos para que seja possível falar metalinguisticamente da língua (como as variáveis e o símbolo inicial da ciência da computação). Comparando, brevemente, o que entende-se por *gramática* e *linguagem* em ambas as áreas, percebe-se que enquanto na ciência da computação a linguagem é gerada pela aplicação da gramática, na linguística, a gramática é um elemento a mais na linguagem, sendo esta última o

rótulo pelo qual entendemos o objeto mais abrangente que temos. Considerando que para as duas áreas, o que entende-se por gramática e linguagem são objetos ligeiramente diferentes, o que pretende-se discutir aqui é **(i)** se, considerando, um conjunto composto por todos os elementos sob o rótulo *linguagem* acrescido de todos os elementos sob o rótulo *gramática*, tem-se os mesmos objetos teóricos tanto para a linguística formal quanto para a ciência da computação, **(ii)** em que medida o que define como cada área entende seus objetos teóricos é o fato de que a linguística formal preocupa-se com a língua natural e a ciência da computação com línguas artificiais, *i.e.*, de que forma a preocupação com linguagens artificiais permite que a gramática seja “priorizada” em relação à linguagem e como, a partir da preocupação com a linguagem natural, a gramática torna-se parte de um objeto maior, a linguagem.

**Palavras-chave:** Gramática, Linguagem, Linguística Formal, Ciência da Computação

### Referências

van BENTHEM, J.; ter MEULEN, A. (eds.) *Handbook of Logic and Language*. Amsterdã: Elsevier, 1997.

CHOMSKY, N. On Certain Formal Properties of Grammars. In: *Information and Control*, Vol 2, 1959.

CHOMSKY, N. 1986. *Knowledge of language*. NY: Praeger.

HOPCROFT, J.; ULLMAN, J.; MOTWANI, R. *Introdução à Teoria de Autômatos, Linguagens e Computação*. Rio de Janeiro: Campus, 2002.

WOOD, M. *Categorial Grammar*. Londres: Routledge, 1993.