

# SEMÂNTICA FORMAL

## SEMÁNTICA FORMAL

## FORMAL SEMANTICS

**José Borges Neto\***

Universidade Federal do Paraná | Universidade Estadual do Oeste do Paraná

**RESUMO:** O propósito do texto é a apresentação de um mecanismo descritivo de estruturas semânticas, desenvolvido em parceria por lógicos, linguistas e especialistas em processamento computacional de línguas naturais, denominado por vezes *semântica de modelo teórico*, que se serve de uma metalinguagem de natureza lógica para descrever as estruturas linguísticas. Depois de uma introdução em que se discute a natureza dos modelos analíticos na Linguística e de uma seção em que se discutem as características de um sistema formal, em geral, busca-se a caracterização da semântica formal desenvolvida, em boa parte, com base nos trabalhos de Richard Montague. Para finalizar, apresentam-se alguns exemplos, claramente simplificados, dos procedimentos formais de que faz uso a semântica de modelo teórico.

**PALAVRAS-CHAVE:** Semântica formal. Semântica de modelo teórico. Semântica de tipos. Semântica de Montague.

**RESUMEN:** El propósito del texto es presentar un mecanismo descriptivo de estructuras semánticas, desarrollado en colaboración por lógicos, lingüistas y especialistas en procesamiento computacional de lenguas naturales, llamado algunas veces de *semántica de modelos teóricos*, que utiliza un metalenguaje de naturaleza lógica para describir las estructuras lingüísticas. Después de una introducción en la que se discute la naturaleza de los modelos analíticos en Lingüística y una sección en la que se discuten las características de un sistema formal, se busca la caracterización de la semántica formal desarrollada, en gran parte, con base en las obras de Richard Montague. Finalmente, se presentan algunos ejemplos, claramente simplificados, de los procedimientos formales utilizados por la semántica del modelo teórico

**PALABRAS CLAVE:** Semántica formal. Semántica de modelo teórico. Semántica de tipos. Semántica de Montague.

**ABSTRACT:** The purpose of this paper is to present the descriptive mechanism of semantic structures, developed in partnership by logicians, linguists and specialists in computational processing of natural languages, called sometimes *model-theoretic semantics*, which uses a logical metalanguage to describe linguistic structures. After an introduction in which the nature of analytical models in Linguistics is discussed and a section in which the characteristics of formal systems are discussed, the characterization of the formal semantics based on Richard Montague's works is presented. Finally, some clearly simplified examples of the formal procedures used by the model-theoretic semantics are presented.

**KEYWORDS:** Formal semantics. Model-theoretic semantics. Type semantics. Montague Semantics.

---

\* Professor titular aposentado da UFPR. Atua como professor sênior na Universidade Federal do Paraná (UFPR) e como professor visitante na Universidade Estadual do Oeste do Paraná (UNIOESTE). E-mail: [borgesnetojose@gmail.com](mailto:borgesnetojose@gmail.com).

## I INTRODUÇÃO

O propósito deste texto é apresentar, em linhas gerais, um tipo de semântica que se pode chamar de formal e que frequentemente é compreendida de modo enviesado pelas pessoas que não a praticam.

Começo dizendo que nenhuma teoria pode pretender dar conta de todos os aspectos de seu objeto: sempre trabalhamos com recortes de porções da realidade e com um subconjunto das propriedades do objeto, propriedades privilegiadas, que consideramos centrais (ou fundamentais).

Jorge Luís Borges, com imensa clareza, nos diz isso da seguinte maneira<sup>1</sup>:

DEL RIGOR EM LA CIENCIA

... En aquel Imperio, el Arte de la Cartografía logró tal Perfección que el mapa de una sola Provincia ocupaba toda una Ciudad, y el mapa del imperio, toda una Provincia. Con el tiempo, esos Mapas Desmesurados no satisficieron y los Colegios de Cartógrafos levantaron un Mapa del Imperio, que tenia el tamaño del Imperio y coincidía puntualmente con él. Menos Adictas al Estudio de la Cartografía, las Generaciones Sigüientes entendieron que ese dilatado Mapa era Inútil y no sin Impiedad lo entregaron a las Inclemencias del Sol y de los Inviernos. En los desiertos del Oeste perduran despedazadas Ruinas del Mapa, habitadas por Animales y por Mendigos; en todo el País no hay otra reliquia de las Disciplinas Geográficas.

SUÁREZ MIRANDA: VIAJES DE VARONES PRUDENTES, LIBRO CUARTO, CAP. XIV, LÉRIDA, 1658.

Ou seja, a tentativa de dar conta de todos os aspectos da realidade vai simplesmente nos devolver a própria realidade: tarefa inútil. Assim, as disciplinas científicas selecionam e recortam porções da realidade que tomam como seu objeto. Fazem mais. Não só dirigem sua atenção para (relativamente) pequenas porções da realidade, mas ainda selecionam aspectos desse seu objeto de observação, que privilegiam. Por exemplo. Os linguistas fazem recortes sucessivos na realidade: alguns “isolam” a linguagem do conjunto das atividades humanas, outros procuram tratá-la de forma associada com aspectos sociológicos ou neurológicos; alguns ocupam-se só da fonologia, enquanto outros tratam da sintaxe; alguns, que fazem morfologia, reconhecem palavras como suas unidades de análise, enquanto outros, alternativamente, reconhecem morfemas como suas unidades; ao estudar de forma privilegiada as combinações de palavras (ou morfemas) preocupam-se com as modificações de forma que as palavras sofrem ao se combinarem (e fazem apenas uma morfossintaxe) ou preocupam-se com a estrutura (abstrata) que resulta das combinações (e fazem uma sintaxe “lógica”) e assim por diante.

E existem alguns que resolvem se aventurar no estudo das significações associadas às expressões linguísticas. Região sombria. E, entre eles, alguns até ousam procurar estruturas. Em princípio, ninguém trata de todas as partes (nem mesmo reconhecem a existência das mesmas partes): de modo geral, os linguistas se especializam na abordagem de apenas alguns aspectos da linguagem e ignoram os aspectos que fogem de sua especialidade. E, acrescento, todos acham que estão certos.

Mas a região sombria da significação também comporta recortes e privilegiamento de aspectos diversos. Há quem se ocupe dos entrelaçamentos entre os significados das expressões e seus usos, entre os significados das expressões e as posições ideológicas dos falantes ou entre os significados das expressões e os contextos em que são proferidas. E há também quem apenas se ocupe do mero significado das expressões, sem levar em conta os contextos de enunciação, as intenções dos falantes etc.

O recorte continua. É possível passar a vida vendo apenas o significado das palavras – fazendo *semântica lexical*. É possível dedicar toda sua atenção apenas ao significado das proposições ou é possível estudar as proposições sempre relacionadas com as intenções dos falantes. E assim por diante.

<sup>1</sup> Ver Borges (1989, p. 847).

É no meio desses múltiplos recortes que se encontra a *semântica formal*. Sua peculiaridade se define tanto pela delimitação de seu objeto próprio quanto pelo modo de abordar esse objeto. Ainda, a própria semântica formal admite mais de uma maneira de definir o objeto e a metodologia de análise.

## 2 FORMAL?

Antes de seguir em frente, creio ser interessante explorar um pouco o adjetivo *formal*. Roberta Pires de Oliveira, num texto muito interessante chamado *Formalismos na Linguística: uma reflexão crítica* (OLIVEIRA, 2004), investiga os sentidos que o termo formal recebe na linguística contemporânea. Basicamente, os sentidos que Roberta destaca são: *formal equivalente a científico*; *formal equivalente a autônomo*; e *formal equivalente a cálculo*.

No primeiro sentido (formal equivalente a científico), está em jogo a consideração de que uma teoria científica qualquer deve ser construída de forma explícita e rigorosa, e apresentada numa linguagem matematicamente precisa. Como diz Apresjan (1980, p. 66)<sup>2</sup>: “Idealmente, todo modelo formal constitui um sistema MATEMÁTICO. Em um certo sentido, portanto, o conceito de formalidade é equivalente ao conceito de matematicidade, precisão ou univocidade”.

No segundo sentido (formal equivalente a autônomo), o termo remete à ideia de forma como oposta a substância ou conteúdo. Trata-se, no fundo, de uma delimitação específica de escopo: uma linguística formal toma como objeto apenas fenômenos relacionados à forma das expressões linguísticas. Roberta Oliveira toma a Gramática Gerativa como exemplo privilegiado de teoria formal nesse segundo sentido e por isso associa esse sentido à autonomia da sintaxe. Não precisaria ser assim. Creio que a questão aqui não é ser autônomo ou não, mas sim o privilegiamento que a gramática gerativa dá ao estudo das formas gramaticais enquanto objetos sintáticos, em detrimento do estudo de seus significados (semântica) ou de seus usos (pragmática).

Antes de passar ao terceiro sentido do termo formal, gostaria de fazer um passeio pela história e tentar descrever rapidamente a natureza do primeiro sistema formal proposto no mundo ocidental, que não a matemática. Trata-se da lógica de Aristóteles.

Os filósofos gregos da antiguidade conheciam a noção de *proposição* (sentença, oração). Sabiam que as proposições eram expressões que podiam ser ditas verdadeiras ou falsas. Sabiam que dada a verdade ou a falsidade de uma proposição a verdade ou falsidade de outras proposições podia ser inferida. Por exemplo, o filósofo grego sabia que a proposição “Todo homem é mortal” é verdadeira e que a proposição “Algum homem não é mortal” é falsa. Sabia que a falsidade da segunda proposição podia ser inferida da veracidade da primeira: se todo homem é mortal é uma proposição verdadeira, então a afirmação de que algum homem não é mortal só pode ser uma proposição falsa.

A grande inovação de Aristóteles foi a proposta de uma “análise sintático-semântica” para as proposições e a construção de um sistema de regras que permitisse que o tratamento das inferências sobre a verdade ou falsidade das proposições fosse feito levando-se em consideração apenas a estrutura das proposições.

Sua análise supunha que todas as proposições tinham uma estrutura tripartite: um sujeito, um atributo e uma relação. Uma proposição como “Todo homem é mortal”, por exemplo, apresentava “homem” como o sujeito da proposição (quantificado pelo *todo*, que dizia que a atribuição recaía em todos os indivíduos que pertenciam à extensão de *homem*); “mortal” como o atributo que a proposição dizia existir em todos os homens; e “é” como a afirmação de que o atributo existia no sujeito. A relação podia ser também “não é”, que negava a presença do atributo no sujeito.

Assim, podia ser dito de toda proposição que sua *estrutura* era SUJEITO + RELAÇÃO + ATRIBUTO, e no estabelecimento das inferências, deixavam de interessar os significados do sujeito e do atributo. As proposições – todas – resumiam-se a quatro *formas*:

<sup>2</sup> Citado também por Oliveira (2004, p. 224).

A: Todo X é Y

E: Todo X não é Y (= nenhum X é Y)

I: Algum X é Y

O: Algum X não é Y

Como, tanto nos casos em que o sujeito era um indivíduo (como “Sócrates”) quanto nos casos em que o sujeito vinha quantificado por *todo*, a distribuição era idêntica (o atributo era afirmado, ou negado, de todos os indivíduos denotados pelo sujeito), as proposições “Sócrates é mortal” e “Todo homem é mortal” podiam ser incluídas em uma mesma forma lógica (a forma A). Essa identificação das formas lógicas das proposições por meio das letras A, E, I e O surge na Idade Média e estão ligadas às palavras latinas *AffIrmo* e *nEgO*<sup>3</sup>.

Para Aristóteles, e para os aristotélicos, todas as proposições podiam ser reduzidas a estas quatro formas. Uma proposição como “Deus criou o mundo”, que é usada na Lógica de Port-Royal (1662), é uma expressão linguística que contém em si a proposição *Deus é criador do mundo*, que é exemplo da forma A.

As regras de inferência podiam ser estabelecidas exclusivamente a partir das quatro formas, independentemente de seus conteúdos. Por exemplo: dada a verdade de uma proposição com a forma A, a proposição I correspondente também é V, a proposição E é falsa e a proposição O é falsa (Todo homem é mortal = V, logo, Algum homem é mortal = V, Nenhum homem é mortal = F e Algum homem não é mortal = F). Por outro lado, supondo-se a verdade de uma proposição da forma I, podemos inferir que E é falso e que A e O podem ser verdadeiros ou falsos (Algum homem é mortal = V, logo, Nenhum homem é mortal = F, Todo homem é mortal = V ou F, Algum homem não é mortal = V ou F).

Esse é o tratamento formal inaugural. Pode-se notar que é *formal* no primeiro sentido (é explícito e preciso) e é também, aparentemente, *formal* no segundo sentido, na medida em que o tratamento das inferências é feito exclusivamente sobre as estruturas sintáticas das proposições. Digo aparentemente porque não podemos identificar a “formalidade” do sistema aristotélico com a ideia de autonomia da sintaxe.

Embora o sistema aristotélico funcione exclusivamente baseado em estruturas sintáticas, seu objeto são as inferências, isto é, as *relações semânticas* que podem ser feitas a partir de proposições. Quando um aristotélico diz que “se A = V, então E = F e vice-versa”, está dizendo que os significados das proposições A e E são contraditórios. Trata-se, então, de um sistema formal apenas enquanto metalinguagem.

Ainda, o sistema aristotélico também só é *formal* no terceiro sentido do termo (formal equivalente a cálculo) enquanto metalinguagem. Explico.

Vejamos inicialmente o que é um *cálculo*.

Creio que todos já tiveram contato com a linguagem dos algarismos romanos (hoje só utilizada na numeração de capítulos de livros e em alguns relógios analógicos). Vocês já tentaram fazer contas simples (somas, por exemplo) usando números escritos nesses algarismos? Tentem a seguinte soma: CMLIV + CLII. O resultado é MCVI. Não sei se alguém conseguiu chegar a esse resultado, mas, se conseguiu, foi provavelmente traduzindo as parcelas numa outra linguagem (a dos números arábicos, por exemplo), fazendo a soma e traduzindo o resultado, de volta, para números romanos. Qual é o problema com a linguagem dos números romanos? Ela não se presta ao cálculo; é muito difícil realizar operações matemáticas com ela. Notem que na primeira parcela da soma, o C está subtraindo cem do M e o I está subtraindo um do V; na segunda parcela, não há caso de subtração – todos os algarismos são somados. O resultado não pode ser obtido pela contagem dos símbolos (nas parcelas há um M, dois C, dois L, um V e três I). Não há posições fixas, como nos números arábicos: não há posição para unidades, dezenas, centenas e assim por diante. Não há algarismo para o zero. Os números IV e VI são compostos pelos mesmos dois algarismos, mas o I tem papéis radicalmente opostos em cada

<sup>3</sup> A = afirmo de todos; I = afirmo de alguns; E = nego de todos; O = nego de alguns.

um deles. A função de cada algarismo na composição do número depende do algarismo imediatamente seguinte – é contextual, portanto.

Sabemos, no entanto, que os romanos faziam a contabilidade de todo seu imenso império. E como agiam? Eles agiam como nós faríamos hoje para resolver a soma apresentada: traduziam as parcelas para uma outra linguagem – a linguagem do ábaco, em que os números eram representados por pedrinhas (*calculi*, em latim) em buracos ou contas em arames – faziam os *cálculos* ou as *contas* no ábaco e traduziam de volta o resultado para números romanos.

É fácil ver como os algarismos arábicos permitem que o cálculo seja feito sem a necessidade de traduzir as expressões em alguma outra linguagem. Usando as mesmas parcelas –  $954 + 152 = 1106$  – somam-se primeiro as unidades, depois as dezenas e assim por diante; se algumas dessas somas parciais apresentar dois algarismos, o primeiro algarismo vai ser somado no nível seguinte. A linguagem dos algarismos arábicos, portanto, é adequada para o cálculo.

Ou seja, um cálculo é um conjunto de operações que podem ser feitas na própria linguagem em que os objetos da operação estão apresentados. Dessa forma, a metalinguagem do sistema aristotélico pode ser dita um cálculo, mas as inferências entre as proposições do grego ou do português não são calculáveis, a menos que sejam traduzidas para a metalinguagem. É o mesmo que acontece com as contas dos romanos: é preciso traduzir os números romanos para outra linguagem (a do ábaco, que é adequada ao cálculo).

O sistema lógico dos aristotélicos, como a linguagem dos números romanos, não permitia que o cálculo fosse feito na própria linguagem em que as proposições eram apresentadas. Era preciso traduzir as proposições nas quatro formas, que constituem uma outra linguagem. Mais ainda. A todo momento, era necessário traduzir expressões em outras expressões para que o sistema funcionasse: o sistema não tratava expressões como “Deus criou o mundo”, que não possuía uma estrutura sintática que pudesse ser acomodada em alguma das quatro formas, o que implicava que tinha que ser traduzida em “Deus é criador do mundo” para que a tradução para a linguagem do sistema pudesse justificar as inferências. Nem mesmo a tradução da linguagem objeto para uma metalinguagem “formal” era sempre possível: era preciso “consertar” as proposições para que as estruturas resultantes pudessem ser traduzidas para a linguagem em que o cálculo seria feito. Talvez o exemplo mais claro desse procedimento esteja no tratamento dado pela Lógica de Port-Royal à proposição *Deus é o único ser necessário*.

Aparentemente, a estrutura da proposição poderia ser traduzida numa forma do tipo A, com sujeito, relação e atributo. Mas aí não temos como explicar a razão porque o silogismo (1) não é válido:

- (1) Deus é o único ser necessário.  
 Pedro não é Deus.  
 Logo, Pedro não é o único ser necessário.

Para os padres de Port-Royal, o problema estava na palavra *único*, que ocultava uma outra proposição: algo como *Nada além de Deus é ser necessário*<sup>4</sup>.

Outro silogismo “problemático” é o (2), que, ao contrário de (1), é evidentemente correto do ponto de vista da inferência, mas não é formalmente válido:

- (2) Os persas eram adoradores do sol.  
 O sol é uma estrela.  
 Logo, os persas eram adoradores de uma estrela.

<sup>4</sup> O argumento, corrigido, teria a seguinte forma:  
 Deus é ser necessário e nada além de Deus é ser necessário.  
 Pedro não é Deus.  
 Logo, Pedro não é ser necessário.

Para que ele fique formalmente válido, é preciso colocar na forma passiva a primeira premissa e a conclusão, obtendo-se (2'):

(2') O sol era adorado por todos os persas.

O sol é uma estrela.

Logo, uma estrela era adorada por todos os persas

Ou seja, antes de calcular as inferências, é preciso revelar a “verdadeira” estrutura das premissas e da conclusão (colocar na ordem direta, expor proposições ocultas, transformar ativas em passivas etc.).

Assim como os algarismos romanos foram abandonados em favor dos algarismos arábicos, a lógica aristotélica foi abandonada, por volta do final do século XIX, em favor de uma lógica “matemática”, desenvolvida numa linguagem que permitia o cálculo lógico, uma linguagem matematicamente precisa e absolutamente explícita. A linguagem em que lógica era desenvolvida era formal, no mesmo sentido em que era formal o sistema lógico aristotélico. Permitia dar conta de inferências que eram feitas sobre proposições que não apresentavam (e/ou não podiam apresentar) estruturas do tipo sujeito, relação e atributo, como as inferências matemáticas, por exemplo. Suas vantagens, enfim, eram evidentes. Mas, ao mesmo tempo em que as inferências matemáticas podiam ser tratadas, as inferências feitas sobre proposições das línguas naturais tornavam-se muito mais complexas e exigiam novos modos de tratar as estruturas sintáticas das expressões das línguas naturais.

Boa parte do desenvolvimento dessa lógica matemática, então, se reduziu à questão de formular uma linguagem que permitisse o tratamento das inferências em língua natural. A criação de uma metalinguagem que permitisse o cálculo lógico de inferências entre expressões matemáticas *e também* entre expressões das línguas naturais.

Ao contrário com o que acontece na matemática, todas as propostas sempre foram de criar uma linguagem, digamos, artificial, adequada ao cálculo, em que se traduzissem as expressões da língua natural. Não tenho nenhum conhecimento de propostas que tenham tentado fazer o cálculo inferencial diretamente nas expressões da língua natural. Mesmo Chomsky, que no modelo de *Syntactic Structures* (CHOMSKY, 1957) pretendia gerar diretamente as expressões da língua natural, no modelo de *Aspects* passou a gerar objetos abstratos que se interpretavam nas expressões das línguas (enquanto cadeias fonéticas estruturadas, de um lado, e enquanto “significados”, de outro. É interessante notar que os “significados” eram apresentados como *formas lógicas*).

Já em 1935, o lógico polonês Kazimierz Ajdukiewicz (1890-1963) propôs um sistema formal não-gerativo para estabelecer a boa-formação (a *gramaticalidade*) das expressões lógicas e de parcela das expressões linguísticas (AJDUKIEWICZ, 1935). Em 1954, o lógico israelense Yehoshua Bar-Hillel (1915-1975) propôs a Chomsky que formulasse o componente de base de sua gramática gerativa usando uma linguagem do cálculo lógico (a chamada *gramática categorial*). Ainda nos anos 1950, o lógico e matemático alemão Joachim Lambek (1922-2014) construiu um sistema lógico para dar conta da estrutura das expressões linguísticas e publicou um texto chamado *The mathematics of sentence structure* (LAMBEK, 1958)<sup>5</sup>. E no fim dos anos 1960 e início dos anos 1970, o lógico norte-americano Richard Montague (1930-1971) produziu um conjunto de textos em que tratava por meio de linguagens adequadas ao cálculo lógico diversos fenômenos da semântica das línguas naturais<sup>6</sup>.

Essas propostas associam a cada expressão da língua natural uma descrição sintática, formulada numa metalinguagem formal, e o cálculo, seja da gramaticalidade, seja do sistema inferencial (semântico) pode ser feito nessa metalinguagem.

As propostas que pretendem formular uma metalinguagem formal capaz de apreender a estrutura sintático-semântica das expressões das línguas naturais e tratar as relações semânticas entre elas, na forma de um cálculo, constituem um programa de investigação linguística conhecido como SEMÂNTICA FORMAL.

<sup>5</sup> E que inclui Noam Chomsky nos agradecimentos.

<sup>6</sup> Dentre os quais destaca-se Montague (1973).

### 3 A SEMÂNTICA FORMAL

De modo geral, as propostas de uma semântica formal (porque podemos dizer que há mais de uma, na medida em que temos várias propostas de metalinguagens formais) trabalham a partir de dois princípios gerais: (i) os significados das expressões são *composicionais* e (ii) os significados das expressões são tratados em termos de *condições de verdade*.

A composicionalidade é um princípio que diz que o significado de uma expressão resulta do significado de suas partes e das relações que as unem. Ou seja, o significado da sentença “Todo gato adora caixinha” resulta da composição dos significados de cada uma das quatro palavras que a compõem, numa estrutura hierárquica que os sintaticistas chamariam de estrutura de constituintes. O significado de “todo” se une ao significado de “gato” e resulta no significado de “todo gato”<sup>7</sup>; o significado de “adora” se une ao significado de “caixinha” e resulta no significado de “adora caixinha”; e, finalmente, o significado de “todo gato” se une ao significado de “adora caixinha” e resulta no significado da sentença completa.

Como se pode ver, a composicionalidade, no fundo, torna a semântica dependente de uma sintaxe. E a composicionalidade semântica aponta para uma composicionalidade sintática. Para saber que partes se unem a cada passo, dependo de uma análise sintática específica. E existem muitas análises sintáticas alternativas.

Por exemplo: posso supor que “todo” é uma função de três argumentos que quando se aplica a “gato” nos deixa com o significado “todo gato”; mas, “todo gato” ainda é uma função de dois argumentos (só um dos três argumentos de “todo” foi preenchido por “gato”) e, portanto, pode se aplicar a “adora”, nos deixando com “Todo gato adora”, que é uma função de um argumento e significa *o conjunto dos x tal que todo gato adora x*; essa função, agora, pode se aplicar a “caixinha”, resultando na sentença “Todo gato adora caixinha”. Ou, alternativamente, posso supor que “todo” pode ser representado, numa linguagem próxima da linguagem da teoria dos conjuntos, como “Todo  $X, Y = X \subseteq Y$ ”, que significa que “Todo gato adora caixinha” é equivalente a dizer que o conjunto dos gatos ( $X$ ) é subconjunto do conjunto dos que adoram caixinha ( $Y$ ).

Dentre as análises sintáticas alternativas, o semanticista deve selecionar a que melhor represente as expressões da língua natural e mais facilmente permita a composicionalidade e o cálculo semântico. Se é a análise sintática da gramática gerativa, como querem Angelika Kratzer e Irene Heim (HEIM; KRATZER, 1998), ou é a análise sintática da gramática categorial, como queriam Lambek, Montague e Carlos Franchi (em sua tese de doutoramento – FRANCHI, 1976), ou é alguma outra, só a tentativa de construir uma semântica formal adequada e consistente poderá nos dizer.

O segundo princípio são as condições de verdade. Falar em *condições de verdade* não é o mesmo que falar em *verdade*. Diante de uma sentença como “O número de folhas daquela árvore é ímpar”, o semanticista nunca se pergunta sobre a verdade ou a falsidade da proposição. Ele se pergunta sobre as condições em que a proposição poderia ser verdadeira. Explico.

A verdade da proposição não é indecidível, em princípio, já que basta ir até a árvore e contar suas folhas. Mas, ninguém faz isso para entender o que a proposição significa. Logo, conhecer o significado de uma expressão não equivale a saber se ela é verdadeira ou não. O que importa, para conhecer o significado de uma proposição, é saber como o mundo teria que ser para que ela fosse verdadeira. Se a proposição é, de fato, verdadeira ou falsa não interessa ao semanticista.

### 4 A SEMÂNTICA DE MODELO TEÓRICO

Vou me ocupar aqui, no que segue, apenas de uma das variedades de semântica formal, conhecida como *semântica de modelo teórico* (*model-theoretical semantics*)<sup>8</sup>.

<sup>7</sup> Por meio de alguma operação que vou ignorar aqui, mas que pode ser compreendida no que se segue.

<sup>8</sup> Ana Müller, Esmeralda Negrão e Maria José Foltran organizaram um livro – chamado *Semântica Formal* – que traz algumas de suas variedades (ver MÜLLER; NEGRÃO; FOLTRAN, 2003). Minha colaboração no livro é um capítulo intitulado justamente *Semântica de modelos* (ver BORGES NETO, 2003).

O filósofo e lógico finlandês Jaakko Hintikka e sua esposa Merrill, num livro sobre o pensamento de Wittgenstein (HINTIKKA; HINTIKKA, 1994), nos falam de duas possibilidades de entender a linguagem: *linguagem como meio universal* e *linguagem como cálculo*. O mesmo assunto é tratado por Martin Kusch (2001).

De acordo com a visão de linguagem como meio universal,

[Não] podemos observar a nossa linguagem como se estivéssemos fora dela e descrevê-la como fazemos com outros objetos [...] O motivo desse suposto impedimento consiste em que só é possível usar a linguagem para falar sobre algo quando podemos nos apoiar numa interpretação definida e estabelecida, numa rede estável de significados existentes entre a linguagem e o mundo. Consequentemente, não pode haver nem propósito nem sentido em dizer na linguagem o que essas relações de significado são, pois qualquer tentativa de fazê-lo implica de antemão sua existência. (HINTIKKA; HINTIKKA, 1994, p. 20)

Essa visão, portanto, implica a *inefabilidade da semântica*, ou seja, a impossibilidade de que os significados das expressões possam ser estudados sistematicamente, já que as relações de significado não podem ser expressadas. Esta é a conclusão a que chega Wittgenstein em seu *Tractatus*<sup>9</sup>, quando diz que “*O que não se pode falar, deve-se calar*”, e esta é a posição que Chomsky compartilha com ele.

Por outro lado, a visão da linguagem como cálculo considera que a linguagem é reinterpretável como um cálculo lógico. Considera que as interpretações mudam e que podemos, perfeitamente, falar sobre as relações de significado.

A semântica de modelo teórico entende a linguagem como um cálculo e, conseqüentemente, entende que a semântica não é inefável.

Creio que esta é a escolha metodológica central feita pela semântica formal em geral: é o que irmana as várias propostas: as línguas são cálculos (ou podem ser reduzidas a cálculos).

Passemos agora aos aspectos relativos à escolha dos objetos privilegiados pela semântica formal.

Normalmente, esse objeto é caracterizado como o *significado literal* das expressões. Quer dizer, a semântica de modelo teórico não está interessada em significados “figurados”, metafóricos, nem em significados que resultem dos contextos, ou mesmo em significados acrescentados pelas intenções dos falantes. Não que esses outros significados não mereçam atenção, apenas a semântica formal decidiu ignorá-los, porque não fazem parte de seu recorte.

Mas eu diria que essa caracterização não é exata. Ao menos se entendermos *significado literal* de um jeito convencional, como aquele significado que a expressão possui se retirarmos todo o contexto e todas as significações acrescentadas pelos falantes.

Para a semântica de modelo teórico, as expressões, na verdade, não possuem esse significado literal convencional. As expressões linguísticas não possuem significados inerentes. Elas são tratadas como objetos abstratos que adquirem algum significado quando associadas a um *modelo de interpretação*. Se pensarmos assim, uma dada expressão terá tantos *significados literais* quantos forem os modelos de interpretação a que ela se associar. Vou dar dois exemplos, um de uma linguagem lógica e outro da língua portuguesa.

Suponhamos a seguinte expressão do cálculo de predicados de primeira ordem “ $P(a)$ ”. Essa expressão, a princípio, não significa nada<sup>10</sup>. Mas ela pode ser associada a um modelo de interpretação –  $M_1$  – em que encontramos (numa função  $F$  que dá a denotação das expressões atômicas) que  $P =$  o conjunto dos números pares e  $a = 3$ . Em  $M_1$ , portanto, nossa expressão denota o mesmo que a expressão portuguesa “3 é um número par” e é falsa. Mas, em outro modelo –  $M_2$  – a função  $F$  pode nos dizer que  $P =$  o conjunto

<sup>9</sup> Ver a afirmação final em Wittgenstein (1969).

<sup>10</sup> Essa afirmação não é totalmente correta: “ $P(a)$ ” significa, literalmente, que a entidade  $a$  possui a propriedade  $P$ .



dos números primos e  $a = 3$ . Em  $M_2$ , a expressão passa a denotar o equivalente à expressão portuguesa “3 é um número primo” e é verdadeira.

A expressão, em si, não significa nada, mas adquire significado ao ser associada a um modelo de interpretação. Justamente aí é que está a vantagem da linguagem do cálculo de predicados: porque suas expressões atômicas não possuem significado inerente, podem significar qualquer coisa, desde que mantidas as relações estruturais. A linguagem do cálculo de predicados é como a linguagem da matemática: por não significar nada, pode significar tudo (basta associarmos a um modelo de interpretação).

Pensemos agora na sentença da língua portuguesa “A massa está pronta” (o exemplo é de Sírio Possenti, comunicação pessoal). Eu diria que a sentença, a princípio, não significa nada<sup>11</sup>. Mas se eu a associar a um modelo de interpretação em que “massa”, por um equivalente da função  $F$  do cálculo de predicados, vai denotar *povo*, como faria um revolucionário, eu estarei dizendo que o povo está mobilizado (para a revolução, suponho); mas se nossa função  $F$  associar “massa” a uma pasta de farinha e água, como faria um cozinheiro, eu estarei dizendo que o macarrão já pode ser servido. E se eu for um pedreiro, poderia estar dizendo que já é possível começar a levantar a parede. Não por acaso, o dicionário Houaiss dá 28 acepções para o termo *massa*.

Entender que as expressões da língua natural, em si, não significam nada permite que digamos que os significados são *históricos* (ou seja, que os modelos de interpretação se alteram no tempo), são *ideológicos* (ou seja, posições ideológicas distintas selecionam modelos de interpretação distintos) e são *contextuais* (ou seja, contextos diferentes, falantes diferentes, lugares diferentes etc. selecionam modelos de interpretação distintos). Pensem em termos como *terrorista*, *democrata*, *golpe* e *ditadura*, por exemplo, e poderão perceber do que eu estou falando.

Se todos atribuíssemos significados semelhantes a boa parte das expressões, deve ser porque compartilhamos, ao menos parcialmente, de um mesmo modelo de interpretação. E talvez seja interessante pensar em como esse modelo de interpretação é construído.

Lembro de um caso em que estávamos – meus pais, irmãos e eu – na praia e havia uma charrete puxada por um cavalo, que fazia passeios pagos pela beira da praia. A charrete vinha cheia de guizos que faziam um barulho que se ouvia de longe. O termo infantil para cavalo em minha família era *tontom*. Meu irmão mais novo devia ter em torno dos dois ou três anos de idade e sempre que ouvia os guizos da charrete corria para a janela para ver a passagem. Alguém precisava levá-lo. A cena, que se repetia dezenas de vezes ao dia, era assistida e comentada: “olha o tontom”, “lá vem o tontom”. E meu irmão logo passou a dizer, desesperado, “tontom, tontom!” ao ouvir os guizos ao longe. Quando voltamos para casa, em Curitiba, o carro passou ao lado de um coletor de papéis que puxava um carrinho. E imediatamente meu irmão apontou e disse “tontom”. Claro que o riso foi geral.

O que aconteceu? O termo “tontom”, que para nós outros tinha denotação precisa, recebeu, tentativamente, por meu irmão alguma denotação em seu próprio modelo de interpretação e recebeu um significado que não correspondia ao significado dado pelo restante da família. O riso, e a zoeira geral, devem ter sido elementos suficientes de coerção que levaram a um ajuste em seu modelo, para que a denotação se aproximasse da denotação convencionalizada pela família. Ou seja, adquirir a linguagem é também adquirir um modelo de interpretação em que a linguagem se interpreta. E quanto mais interação entre as pessoas, mais próximos se tornarão seus modelos de interpretação. Quanto mais experiências em comum – leituras, filmes, escolarização etc. – mais próximos ficam seus “estoques” de expressões linguísticas e seus modelos de interpretação.

Antes que alguém diga que o número de modelos de interpretação caminha em direção ao infinito, o que é certamente verdadeiro, creio que agora já posso abordar a palavra *teórico* no nome “semântica de modelo teórico”.

Se a lógica ou a matemática fossem tratar de casos particulares, elas não teriam nenhuma utilidade. O mesmo acontece com a semântica. A lógica e a matemática estabelecem esquemas gerais de raciocínio (relações gerais, regras gerais, princípios gerais), dos quais se podem deduzir casos particulares, desde que estabelecidas as condições para a aplicação do esquema.

<sup>11</sup> A “ilusão” de que ela significa algo fora de qualquer modelo de interpretação ou de qualquer contexto se deve à compulsão que o ser humano apresenta de buscar uma significação para o que quer que seja dito. Isso, aliás, é a base do Princípio de Cooperação de Paul Grice.

Explico com um exemplo que ouvi mais de uma vez de Carlos Franchi. Se você perguntar para um físico qual o gasto de energia que alguém vai ter ao levantar uma cadeira do chão, o físico vai dar risada e dizer: me dê todas as variáveis – peso da cadeira, altura a que será levantada, pressão atmosférica etc. – e eu calculo para você. Nenhum físico está preocupado com casos particulares. Eles se ocupam do estabelecimento de regras gerais de cálculo que possam, por dedução, dar respostas aos casos particulares.

O mesmo acontece com a semântica de modelo teórico. Os semanticistas não estão interessados em estabelecer a verdade ou a falsidade de proposições particulares; eles estão interessados em estabelecer regras gerais de interpretação que possam ser aplicadas a qualquer proposição que se apresente. E entre as condições para a efetiva interpretação de uma proposição particular deverá estar o modelo em que ela deve ser interpretada.

A semântica de modelo teórico, então, não se preocupa em descrever modelos de interpretação concretos, mas frente a um modelo teórico, busca estabelecer as regras que devem reger as interpretações em qualquer modelo.

E qual a utilidade que uma semântica desse tipo pode ter? Eu diria que a mesma que aprender a regra geral para o cálculo do volume de um sólido geométrico teria. Se eu sei calcular o volume de um cubo teórico, eu sei calcular o volume de qualquer cubo, basta eu fixar o valor das arestas.

Em vez de associar expressões a coisas do mundo – como boa parte dos críticos da semântica formal pensam, erradamente, que fazemos – as expressões são associadas a *tipos lógicos*. Por exemplo, uma palavra como *Pedro* é associada ao tipo lógico das entidades ( $e$ ), uma palavra como *menino* é associada ao tipo lógico dos conjuntos de entidades ( $\langle e, t \rangle$ )<sup>12</sup>, o conjunto dos meninos, e uma palavra como *corre* também é considerada um conjunto de entidades ( $\langle e, t \rangle$ ): o conjunto das entidades que correm. Assim, numa sentença como *Pedro corre*, o que temos é a aplicação da função *corre* ao argumento *Pedro*, e a sentença como um todo significa que a entidade ‘pedro’ pertence ao conjunto das entidades que correm. E isso é verdadeiro ou falso. A decisão sobre a verdade ou falsidade da sentença vai depender de um modelo de interpretação que nos diga quem é Pedro e quais são as entidades que correm, mas o procedimento de obtenção das condições de verdade da sentença já está estabelecido nas relações entre os tipos lógicos<sup>13</sup>.

Creio que vale a pena ver exemplos um pouco mais complexos. Começemos com a sentença *Menino corre*.

Se *menino* denota um conjunto de entidades e *corre* denota um outro conjunto de entidades, a sentença deve significar uma relação entre conjuntos, em que todas as entidades do primeiro conjunto também pertencem ao segundo: algo como  $M \subseteq C$  (o conjunto  $M$  é subconjunto do conjunto  $C$ ).

Notem que se a sentença fosse *A menina corre*, não teríamos mais uma relação entre conjuntos, mas uma aplicação de *corre* a uma única entidade, a uma única menina. Portanto, o artigo definido deve funcionar como um operador que extrai uma única entidade de um conjunto (transforma uma expressão de tipo  $\langle e, t \rangle$  em uma expressão de tipo  $e$ ). E a sentença deve significar que a menina selecionada pertence ao conjunto dos que correm.

Vejamos agora a sentença *Pedro beija Maria*.

*Pedro* e *Maria* são expressões de tipo lógico  $e$ ; *beija* é uma função de dois argumentos (para chegar a um valor de verdade precisa preencher duas posições: a posição que nos diz *quem beija* e a posição que nos diz *quem é beijado*). Como sabem os lógicos, desde os anos 1920, a partir do trabalho do lógico alemão Schoenfinkel, uma função de dois argumentos pode ser dividida em duas funções de um argumento que se aplicam sucessivamente, sem perda de informação, e assim, *beija* pertence ao tipo lógico  $\langle e, \langle e, t \rangle \rangle$  – que quer dizer que *beija* é uma função que toma como argumento uma entidade e resulta numa nova função que toma uma entidade como argumento para resultar num valor de verdade. Em outras palavras, *beija* toma como argumento a entidade *Maria* e resulta no conjunto *beija Maria*, que é do tipo  $\langle e, t \rangle$  e que significa o conjunto dos  $x$  que beijam Maria. A função *beija Maria* se aplica ao

<sup>12</sup> O índice  $\langle e, t \rangle$  é um par ordenado que representa uma função que toma uma entidade ( $e$ ) como argumento e resulta num valor de verdade ( $t$ ).

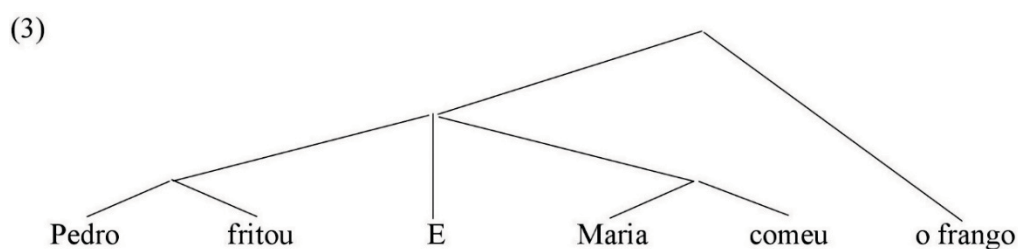
<sup>13</sup> Se  $\langle e, t \rangle$  é, semanticamente, um conjunto e, sintaticamente, uma função que se aplica a um  $e$  para resultar num  $t$ ; se *corre* é um  $\langle e, t \rangle$  e *Pedro* é um  $e$ ; a aplicação de *corre* a *Pedro* vai resultar em *Pedro corre*, que é um  $t$  e que significa que ‘pedro’ é membro do conjunto dos que correm.

argumento *Pedro* e resulta na sentença *Pedro beija Maria*, que significa que Pedro pertence ao conjunto dos que beijam Maria e que é do tipo lógico  $t$ , ou seja, que tem um valor de verdade (V ou F).

O mais interessante é que poderíamos começar a construção do significado da sentença aplicando a função de dois argumentos *beija* ao argumento *Pedro*, obtendo *Pedro beija*, que também é do tipo  $\langle e, t \rangle$  como *beija Maria*, e que significa o conjunto dos  $y$  que Pedro beija; e só no segundo passo aplicar a função *Pedro beija* ao argumento *Maria*, dizendo que Maria pertence ao conjunto das entidades que Pedro beija, o que será verdadeiro ou falso.

Essa flexibilização na ordem das aplicações funcionais pode ser útil para tratar casos como *Pedro fritou e Maria comeu o frango*. Notem que podemos construir *Pedro fritou* exatamente como construímos *Pedro beija*. E o mesmo pode ser feito com *Maria comeu*. Tanto *Pedro fritou* quanto *Maria comeu* são expressões do tipo lógico  $\langle e, t \rangle$  (o conjunto dos  $x$  que Pedro fritou; o conjunto dos  $y$  que Maria comeu). Se considerarmos que a conjunção “e” pode ser associada à operação de intersecção entre conjuntos ( $\cap$ ), podemos chegar ao significado de *Pedro fritou e Maria comeu*, que é a intersecção do conjunto das coisas que Pedro fritou e do conjunto das coisas que Maria comeu e que também é uma função que pertence ao tipo lógico  $\langle e, t \rangle$  (também é um conjunto de entidades). Ao aplicar essa nova função à entidade *o frango* (estou tratando “o frango” como uma entidade do mesmo modo como fiz para “a menina”), vou dizer que a denotação de *o frango* pertence ao conjunto das coisas que Pedro fritou e Maria comeu (está na intersecção dos dois conjuntos), o que é verdadeiro ou falso.

Toda análise semântica supõe uma análise sintática, mas essa análise sintática é essencialmente dirigida por razões semânticas. A análise sintática que fizemos de *Pedro fritou e Maria comeu o frango* teria a aparência da estrutura (3), que não corresponde à forma que teria na sintaxe gerativa, por exemplo:



Ou, numa representação parentetizada, a forma (3')

(3') (((Pedro<sub>e</sub> fritou<sub>\langle e, \langle e, t \rangle \rangle</sub>)<sub>\langle e, t \rangle</sub> E (Maria<sub>e</sub> comeu<sub>\langle e, \langle e, t \rangle \rangle</sub>)<sub>\langle e, t \rangle</sub> o frango<sub>e</sub>)<sub>t</sub>

Essa análise permite a construção do significado da sentença sem usar categorias vazias, índices referenciais ou relações anafóricas (catafóricas), e permite manipular como constituintes (sintático-semânticos) coisas que não são, normalmente, consideradas constituintes pela sintaxe gerativa ou pela sintaxe tradicional.

## 5 CONCLUSÃO

Enfim, as técnicas e procedimentos de uma semântica de modelo teórico são relativamente simples, rigorosamente formuladas numa linguagem formal adequada ao cálculo. A sintaxe que a acompanha (sintaxe categorial, basicamente, embora não seja necessariamente a única possível) também é bastante simples e rigorosa e permite a flexibilidade necessária para que os processos de construção dos significados das expressões possam ser descritos.

Seus propósitos são claramente identificados: destina-se a descrever o processo de construção do significado composicional das proposições. Desse modo, é sem sentido a crítica de que não dá conta das intenções dos falantes, de algumas das variações

contextuais ou dos significados figurados, já que esses componentes da significação geral das expressões linguísticas ficam fora de seu escopo.

Deixo como última observação – para que não se pense que ajo como um fã sem senso crítico – que a semântica de modelo teórico, como todos os modelos linguísticos formais, não tem respostas para todos os casos e enfrenta alguns problemas de difícil solução.

## REFERÊNCIAS

AJDUKIEWICZ, K. Die syntaktische Konnexität. *Studia Philosophica*, n.1, p. 1-27, 1935.

APRESJAN, JU.D. *Ideias e métodos da linguística estrutural contemporânea*. Trad. Lucy Seki. São Paulo: Cultrix; Campinas: Fundação de Desenvolvimento da Unicamp, 1980.

BORGES, J. L. *Obras completas*. v. 1. Buenos Aires: Emecé editores, 1989.

BORGES NETO, J. Semântica de modelos. In: MÜLLER, A. et al. (org.). *Semântica formal*. São Paulo: Contexto, 2003, p. 9-46.

CHOMSKY, N. *Syntactic Structures*. The Hague: Mouton, 1957.

FRANCHI, C. *Hipóteses para uma teoria funcional da linguagem*. 1976. Tese (Doutorado em Linguística) – Universidade Estadual de Campinas, Campinas, 1976.

HEIM, I.; KRATZER, A. *Semantics in generative grammar*. Trad. Enid Abreu Dobranszky. Oxford: Blackwell, 1998.

HINTIKKA, M; HINTIKKA, J. *Uma investigação sobre Wittgenstein*. Campinas: Papirus, 1994.

KUSCH, M. *Linguagem como cálculo vs. Linguagem como meio universal*. Trad. Dankwart Bernsmüller. São Leopoldo: Editora da Universidade do Vale do Rio dos Sinos, 2001.

LAMBEK, J. The mathematics of sentence structure. *American mathematical monthly* 65, p. 154-170, 1958.

MONTAGUE, R. The proper treatment of quantification in ordinary English. In: HINTIKKA, J. et al. (ed). *Approaches to natural language*. New York: Springer, 1973. p. 221-242.

OLIVEIRA, R. P de. Formalismos na linguística: uma reflexão crítica. In: MUSSALIM, F.; BENTES, A. C. (org.). *Introdução à linguística*, v. 3 – fundamentos epistemológicos. São Paulo: Cortez, 2004. p. 219-250.

WITTGENSTEIN, L. *Tractatus Logico-Philosophicus*. Trad. José Arthur Giannotti. São Paulo: EDUSP e Companhia Editora Nacional, 1969.



Recebido em 23/03/2020. Aceito em 27/03/2020.