

## COMO COMPREENDEMOS SENTENÇAS?\*

Edson FRANÇOZO (*Universidade Estadual de Campinas*)

*ABSTRACT The principles underlying psycholinguistic theories of parsing are discussed as a basis for the computational simulation of linguistic processes.*

### 0. Introdução

Durante as duas últimas décadas houve um notável aumento na compreensão de como as pessoas compõem o significado das sentenças a partir das palavras que ouvem ou lêem. Tudo começou com a percepção de que o processo de composição tinha que ser restringido pelo conhecimento gramatical do ouvinte/leitor. Como a habilidade de usar a linguagem é uma das características mais marcantes e universais dentre as habilidades cognitivas humanas, muitos psicólogos passaram a supor que a compreensão de como essa habilidade funciona lhes daria *insights* sobre a natureza da cognição humana. Muitos lingüistas, por sua vez, passaram a confiar que esse estudo lhes daria informações sobre a razão pela qual a linguagem é como é.

O resultado dessa mudança no estudo da compreensão resultou no aparecimento de teorias – interessantes, específicas e testáveis – sobre como as pessoas usam seu conhecimento lingüístico para compreender sentenças. Algumas dessas teorias tomaram a forma de princípios de *parsing* (o processo de identificar partes das sentenças e as relações entre essas partes) que faziam afirmações explícitas sobre, por exemplo, a dificuldade de se compreender certas sentenças e os erros que são cometidos durante sua compreensão. Outras teorias tomaram a forma de programas de computador que atribuem uma análise estrutural explícita a cadeias de palavras, seguindo princípios que se aplicam à gramática de uma língua – uma vertente que veio a ser conhecida como psicolingüística computacional.

\* Agradeço à colaboração, inestimável, de Luiz Arthur Pagani (UEL). Este trabalho foi apresentado na palestra de encerramento "A Emergência de diferentes áreas do saber nas pesquisas sobre linguagem".

No entanto, a aproximação entre as teorias psicológicas e as computacionais não é trivial. Neste trabalho vamos examinar a forma geral das teorias de *parsing*, expondo seus princípios básicos de tal forma a permitir seu cotejo com as teorias computacionais<sup>1</sup>.

### 1. Parsing<sup>2</sup>

Vamos assumir que os ouvintes/leitores compreendem uma sentença identificando as palavras que ela contém, combinando essas palavras em constituintes de uma maneira que segue a gramática de sua língua, e compondo o significado dessa sentença a partir do significado das palavras e dos constituintes de forma a respeitar sua combinatória estrutural. A não ser por erros ocasionais, os ouvintes/leitores não violam seu conhecimento gramatical quando procuram pela interpretação de uma sentença: cada módulo consulta toda informação disponível em seu interior, e cada um deles é sensível a sutis distinções gramaticais<sup>3</sup>.

É claro que, a partir das restrições impostas pela gramática, os ouvintes/leitores geralmente chegam ao significado contextualmente mais plausível e apropriado. O desafio mais sério é o de construir uma teoria que respeite as restrições gramaticais impostas à sentença (globalmente, sobre toda a sentença, a despeito de limitações de memória imediata) e ao mesmo tempo seja sensível ao número ilimitado de fatores que influenciam a plausibilidade da interpretação. Uma teoria psicolinguística que se atenha a apenas um desses aspectos não poderá jamais ser bem sucedida.

Em geral, o desenvolvimento de uma teoria adequada assenta-se sobre o exame de dados experimentais como a resposta de mecanismos de *parsing* a ambigüidades estruturais locais ou temporárias, que são comuns nas sentenças. Vamos assumir aqui que o ouvinte/leitor atribui uma análise estrutural a uma sentença; nessa análise, cada palavra é atribuída a

<sup>1</sup> Uma tarefa que, por falta de espaço, não será realizada aqui

<sup>2</sup> Esta seção, bem como a seguinte, baseiam-se em Frazier e Clifton Jr. (1996).

<sup>3</sup> É fácil demonstrar o enorme efeito sobre o significado que pequenas diferenças podem ter; compare *O médico encaminhou o paciente na enfermaria* com *O médico encaminhou o paciente da enfermaria*. No primeiro caso, especifica-se o local de onde foi efetuado o encaminhamento médico e no segundo caso especifica-se o local em que o paciente estava.

um constituinte e cada constituinte relaciona-se a todos os outros numa árvore que especifica, entre outras coisas, as relações hierárquicas entre esses constituintes.

Os exemplos abaixo mostram como as sentenças podem ser temporária ou permanentemente ambíguas<sup>4</sup>:

- (1) O navio libanês [<sub>VP</sub> entrava no porto [<sub>NP</sub> **o navio brasileiro**]].
- (2) O navio libanês [<sub>VP</sub> entrava no porto [<sub>PP</sub> **ao lado do navio brasileiro**]].
- (3) João [<sub>VP</sub> disse que [<sub>S</sub> o padeiro chegou ontem]].
- (4) João [<sub>VP</sub> disse que [<sub>S</sub> o padeiro chegou] ontem].

As sentenças (1) e (2) ilustram o caso de ambigüidade temporária; o trecho em negrito indica onde a ambigüidade é resolvida. Nas sentenças (3) a (4) a ambigüidade não é resolvida gramaticalmente (embora um dos significados possa predominar sobre outro).

O mecanismo de *parsing* deve ser capaz de computar ambas as análises de cada seqüência de palavras e de chegar àquela que é determinada gramaticalmente (ou que é a mais plausível). Como isso é possível? A maioria das tentativas de resolver esse problema caracterizam o *parsing* em termos de dois parâmetros não totalmente distintos entre si:

- primeiro parâmetro envolve as relações lógicas e temporais entre as análises: o *parser* poderia computar todas as análises em paralelo, de modo que pudesse escolher dentre elas; ou o *parser* poderia computar uma única análise e mantê-las até que ela fosse contraditada; ou ainda o *parser* poderia adiar a computação quando detecta uma ambigüidade, retomando a análise apenas quando (e se) não houvesse mais ambigüidades.
- segundo parâmetro envolve a informação usada para resolver as ambigüidades na análise ou análises iniciais da sentença: para escolher entre análises alternativas, o *parser* deveria poder consultar imediatamente todas as informações disponíveis, incluindo conhecimentos discursivos e sobre o mundo, ou ele poderia consultar um domínio restrito de informações (por exemplo, informação gramatical).

<sup>4</sup> A parentetização é usada para mostrar os fatos estruturais relevantes.

Um simples exercício dedutivo nos mostra que há problemas para todas as três alternativas associadas ao primeiro parâmetro. Se o *parser* computasse todas as análises em paralelo, ele poderia muito bem exceder a capacidade da memória humana ao calcular todas as possíveis combinações das múltiplas ambigüidades que são comuns nas sentenças. Se ele computasse apenas uma análise de cada vez, haveria problemas a cada vez que a análise se mostrasse inadequada. Se ele adiasse a computação, também aí haveria problemas de memória; e, pior, se o *parser* não atribui de saída alguma análise estrutural às ambigüidades, como ele poderia reconhecer o trecho onde a ambigüidade pode ser resolvida? De fato, uma boa quantidade de evidências empíricas tem-se acumulado, evidenciando os problemas ou perturbações do *parsing* nos pontos de desambigüação. Em primeiro lugar, esses dados são evidência contra o paralelismo: se as análises são todas computadas, e o material lingüístico da região desambigüadora é usado para escolher entre elas, não haveria razão para se esperar problemas nesse ponto do processamento. Esses dados são também evidência contra a hipótese do adiamento: uma vez que a perturbação é assimétrica, aparecendo apenas quando a informação desambigüadora força uma interpretação, não se pode dizer que o processamento é perturbado sempre que uma análise adiada tiver que ser finalmente realizada. As evidências são portanto consistentes com uma concepção de *parsing* de análise única: se o *parser* está tipicamente comprometido com apenas uma análise, pode-se esperar perturbação sempre que o material que elimina a ambigüidade indicar que a análise inicial estava errada.

A decisão que se toma com relação ao primeiro parâmetro (serial/paralelo/adiamento) afeta o segundo, isto é, qual informação é usada para resolver a ambigüidade. Para que sentenças completamente ambíguas como (3) e (4) sejam compreendidas em seu sentido mais plausível, o ouvinte/leitor tem que levar em consideração toda a informação relevante – que potencialmente é toda a informação de que ela/ele dispõe. Isto é, todos os tipos de informação devem ser eventualmente usadas para resolver ambigüidades. Se um modelo da análise em paralelo é aceito, é fácil também aceitar um modelo de uso da informação no qual toda a informação relevante é usada para escolher entre as alternativas. O mesmo pode ser afirmado de um modelo de análise com adiamento.

No caso do modelo de análise única, uma análise é construída tão logo um certo tipo de informação se torna disponível. Temos então um problema tanto lógico quanto empírico: determinar qual tipo de informação está na base da construção de uma análise e da resolução da ambigüidade. Uma forte razão para se favorecer o modelo da análise única, para começar, é que esse modelo fornece alguma análise de uma sentença logo que possível. Como assumimos que o processador não tem capacidade para desenvolver simultaneamente múltiplas interpretações semânticas das sentenças, concluímos que uma análise sintática é necessária como suporte para a interpretação semântica, que presumivelmente é um precursor de uma avaliação de plausibilidade, de coerência discursiva, etc. Além disso, a estruturação sintática pode ajudar na identificação de outras análises sintáticas. Essas considerações sugerem que a informação relevante para a estruturação sintática será usada inicialmente para resolver uma ambigüidade sintática e para atribuir uma análise a uma cadeia de palavras.

A informação gramatical é o caso mais puro de informação que é relevante para a estruturação sintática. Essa informação pode ser identificada, previamente a qualquer sentença dada, como relevante à tarefa de compreender uma sentença. Isto é, o *parser* inicialmente consulta informações limitadas, que ‘sabe’ de antemão serem relevantes para sua tarefa. Como consequência, adotar um modelo de análise única encoraja a postulação de que é a informação gramatical que será usada inicialmente para resolver ambigüidades sintáticas. Mas, é claro, isso ainda é pouco. A presença de ambigüidade estrutural significa que a informação gramatical por si é incapaz de determinar uma análise única. Portanto, alguns princípios que regulamentem a aplicação da informação gramatical ao *parsing* devem ser postulados. Além disso, falar de ‘informação gramatical’ dá lugar a uma grande quantidade de questões fascinantes: qual a teoria sintática que se adota? que tipos de informação sintática ela disponibiliza, e de que maneira essa informação está representada? há distintas subclasses de informação sintática que podem ser usadas diferencialmente? Responder a essas perguntas, por sua vez, remete a perguntas ainda mais fundamentais sobre a mente humana: qual sua arquitetura funcional? quais unidades distintas de processamento ela contém (se é que distintas unidades de processamento, ou módulos, existem)? como os distintos módulos estão coordenados, e como eles se comunicam entre si? em que medida a linguagem é uma habilidade específica unicamente humana?

## 2. Princípios de Processamento Sintático

Como vimos, uma teoria do *parsing* atribui uma única e imediata análise a um fragmento ambíguo de uma sentença ambígua. A análise atribuída é determinada por princípios gerais que levam em conta a natureza da estrutura que está sendo construída, e não toma conhecimento de fatores não estruturais.

Os princípios do *parsing* são na verdade instanciações de um único princípio muito geral: “Escolha a primeira análise disponível”. Esse princípio está de acordo com a necessidade, de que falamos antes, de que é preciso determinar as relações sintáticas entre palavras e entre constituintes para que se possa avaliar suas relações semânticas. Está também de acordo com outros dados psicológicos – por exemplo, o fato de que é mais fácil reter na memória imediata material estruturado do que material não estruturado. Os princípios específicos mais estudados são:

- **Minimal Attachment** Não postule qualquer nó potencialmente desnecessário
- **Late Closure** Se for gramaticalmente possível, ligue novos itens a sentenças ou constituintes postulados mais recentemente
- **Minimal Chain Principle** Postule os membros de uma cadeia no primeiro ponto gramaticalmente possível, mas não postule desnecessariamente membros de cadeias.

As evidências para sustentar esses princípios vêm de medidas da dificuldade de processar sentenças que têm uma ambigüidade temporária, mas que são finalmente resolvidas (interpretadas) no sentido previsto por esses princípios, e não no sentido oposto. A dificuldade de processamento é conhecida como ‘efeito labirinto’ (*garden-path effect*). Presumivelmente, o *parser* teria atribuído uma interpretação inicial como previsto pelo princípio relevante, mas então (tendo-se perdido no labirinto) teve que revisar a interpretação, o que causa um custo em termos de tempo ou dificuldade de processamento.

As primeiras evidências foram apresentadas por Frazier, em 1978<sup>5</sup>. Ela usou uma simples tarefa de apresentação visual em que as palavras de

<sup>5</sup> Citado em Frazier e Clifton Jr. (1996).

uma sentença eram apresentadas, uma de cada vez, por um breve período (300 ms. ou 350 ms.); ao fim da apresentação, pedia-se aos sujeitos que indicassem se a sentença era (ou não) gramatical. Para uma das listas de sentenças, que continha itens cuja interpretação poderia ser explicada pelo princípio do *'late closure'*, grandes diferenças de tempo foram encontradas.

?1? Though George kept on reading the story Sue bothered him.  
(1160 ms.)

?2? Though George kept on reading the story still bothered him.  
(1545 ms.)

?3? Without her contributions the funds are inadequate. (973 ms.)

?4? Without her contributions would be very inadequate. (1985 ms.)

Nas sentenças (6) e (8) acima a ambigüidade era resolvida (pela palavra ou constituinte em negrito) em favor de uma análise na qual o constituinte temporariamente ambíguo (por exemplo, *contributions* em (8)) deveria ser ligado ao constituinte que estava sendo processado (*her*) – isto é, uma interpretação compatível com o princípio do *late closure*. Nas sentenças (7) e (9), a ambigüidade era resolvida de maneira inconsistente com esse princípio – em (9), *contributions* deve constituir um novo nó, ao qual *would* será ligado, em vez de ligar-se a *her* como em (8). Resultados como esses – que tem sido obtidos por muitos pesquisadores desde então – são tomados como evidência de que os princípios acima corroboram a idéia de que do mecanismo de compreensão da linguagem faz parte um *parser*, isto é, um dispositivo destinado a atribuir uma análise sintática às sentenças.

### 3. Arquitetura e Processamento

Mas o estudo de como é que as pessoas processam a linguagem tão rapidamente, aparentemente sem fazer esforço, e de maneira automática exige mais do que princípios de *parsing*. Há questões relacionadas à arquitetura da mente e ao modo como os diversos dispositivos necessários à compreensão da linguagem interagem. Nesse contexto são duas as hipóteses fundamentais:

- **Modularidade** Um processador de sentenças constitui-se em um sistema distinto no interior da faculdade da linguagem, e tem uma

arquitetura modular, distribuída, em que a função de cada módulo é determinada por subsistemas da teoria linguística que são informacionalmente encapsulados (Crocker 1996: 1).

- **Incrementalidade** O funcionamento do processador de sentenças e seus módulos constituintes é determinado pelo princípio amplo da compreensão incremental. Esse princípio assegura que o uso máximo da informação linguística é possível, à medida em que cada palavra do enunciado é encontrada (Crocker 1996: 2).

Há tipos e graus diferentes de modularidade, mas a idéia subjacente é bastante simples: certas subtarefas da cognição, particularmente a percepção, são de responsabilidade de procedimentos distintos e especializados (Fodor 1983). De uma perspectiva formal, uma motivação para essa hipótese é simplesmente que as teorias de que dispomos atualmente sobre a fonologia, a sintaxe, a semântica e a pragmática, e sobre o raciocínio, fazem uso de mecanismos formais muito diferentes. Isto é, o sistema de regras e representações usado para caracterizar um determinado tipo de fenômeno é bastante distinto, em termos de poder expressivo e propriedades computacionais, dos usados para caracterizar outro tipo.

Uma motivação adicional, e que é importante aqui, é que a modularidade permite um processamento mais eficiente: quando se examinam casos em que o processamento é perturbado, pode-se conseguir pistas sobre detalhes das propriedades do sistema, especialmente sobre a modularidade sintática. Há certos tipos de ambigüidade estrutural em que as pessoas parecem lançar mão, pelo menos inicialmente, de uma heurística puramente sintática para decidir entre alternativas. Por exemplo, porque é mais fácil ler (9) do que (10)?

?1? Pedro disse a Paulo que [ele viu um grande filme ontem].

?2? Pedro disse a Paulo que [ele vai ver um grande filme] ontem.

A sentença (9), cuja leitura não é problemática, conforma-se a uma estratégia que consiste em ligar o advérbio final à encaixada mais recente ou mais baixa - é o que se chama de *minimal attachment* (Frazier e Clifton 1996). Na sentença (10), cuja leitura é mais difícil, parece que as pessoas tentam um *attachment* similar ao da sentença (9), e só mais tarde percebem que isso é impossível devido a razões semânticas. Isso sugere uma preferência puramente estrutural para *attachments* mais imediatos, mesmo face a informações semânticas adversas. Uma inferência que se



faz desse tipo de evidência é que a análise do adjunto é feita por um processador modular, que obedece à estratégia do *attachment* contíguo. É somente na fase posterior da interpretação semântica que se percebe que isso não está certo, e uma reanálise faz-se necessária. De um ponto de vista computacional, é o privilégio de um tipo de representação, a sintática neste caso, que garante eficácia e rapidez ao sistema – mesmo ao custo de, em casos raros, tornar a reanálise necessária.

Quanto à incrementalidade, ela decorre em parte da idéia intuitiva de que as pessoas parecem compreender o que ouvem ou lêem à medida em que lêem ou ouvem, aparentemente num processo que avança palavra por palavra (em vez de esperar por fronteiras de sentença ou de sintagma). Isto é, parece que além de recuperar o significado das palavras que lemos ou ouvimos, rapidamente atribuímos estrutura ao *input*, de modo a tornar possível uma interpretação semântica composicional à medida que lemos ou ouvimos. Novamente, é quando algum tipo de perturbação atua nesse processo que podemos melhor compreender suas características. Considere as seguintes sentenças (adaptadas de Stowe 1989):

?3? Depois que a polícia parou o caminhão estacionou.

?4? Depois que a bicicleta parou o caminhão estacionou.

Ambas as sentenças acima são difíceis de compreender, talvez porque *caminhão* é inicialmente ligado a *parou* como seu objeto. Mas há evidências de que (11) é mais difícil de compreender do que (12). A explicação mais imediata é que em (12) as pessoas percebem que é pouco provável que uma bicicleta pare um caminhão, de tal forma que a leitura ergativa é privilegiada. Em (11), *polícia* é mais consistente com uma leitura causativa do verbo, tornando o *attachment* do objeto mais difícil de desconsiderar. Para que esta explicação se mantenha, é necessário que uma interpretação incremental seja aceita, isto é, que se tenha atribuído leitura ergativa ou causativa ao verbo, antes que se inicie a análise de seu (possível) complemento.

#### 4. Psicolinguística Computacional

O estudo da linguagem é um programa de pesquisas fundamental no desenvolvimento de nossos conhecimentos sobre a mente humana. E esse programa de pesquisa é essencialmente interdisciplinar. Nesse sentido, é natural que o modelamento computacional dos processos lingüísticos seja

devidamente informado por princípios como os acima discutidos. Resta discutir como isso pode ser feito.

#### REFERÊNCIAS BIBLIOGRÁFICAS

- Crocker, M.W.(1996) *Computational Psycholinguistics*. Dordrecht: Kluwer.
- Fodor, J.(1983) *Modularity of Mind*. Cambridge: The MIT Press.
- Frazier, L.(1978) On Comprehending Sentences: Syntactic parsing strategies. Dissertação de Doutorado inédita, Universidade de Connecticut, EUA.
- \_\_\_\_\_ e Clifton, Jr., C.(1996) *Construal*, Cambridge: The MIT Press.
- Stowe, L.(1989) Thematic Structure and Sentence Comprehension. In G.N. Carlson e M.K. Tannehaus (eds) *Linguistic Structure in Language Processing*, Dordrecht: Kluwer.