

O ACESSO SEMÂNTICO NO PARSING SINTÁTICO: EVIDÊNCIAS EXPERIMENTAIS

Marcus MAIA¹

- RESUMO: Este estudo apresenta evidências em favor da hipótese de que o parser faz uso rápido e eficiente de alguns tipos de informação de natureza lexical associados ao verbo no processamento *on-line*.
- PALAVRAS-CHAVE: Parsing sintático; anáfora objeto; português brasileiro.

Introdução

Uma divisão tradicional no âmbito dos estudos sobre o parsing sintático separa, de um lado, as pesquisas cujos resultados parecem indicar que o mecanismo de compreensão de frases baseia-se inicialmente em procedimentos formais, puramente estruturais, não tendo acesso imediato a qualquer tipo de informação de natureza semântica. A estratégia da Ligação Mínima de Frazier (1979) é um exemplo clássico de procedimento de parsing inteiramente fundamentado na estrutura, sem acesso semântico inicial. De outro lado, alinham-se pesquisas que indicam a existência de acesso imediato a alguns tipos de informação de natureza semântica no processo de compreensão de frases. De fato, nos últimos anos, conforme apontado por autores como Pritchett (1992) e Gorrell (1995), tem havido um número crescente de propostas em favor

¹ Professor adjunto do Setor de Linguística do Museu Nacional e do Programa de Pós-Graduação em Linguística – Departamento de Linguística – Faculdade de Letras – UFRJ – Rio de Janeiro – RJ (maia@acd.ufrj.br).

da hipótese do acesso semântico rápido pelo parser. Segundo estes autores, um parser sensível ao conteúdo da frase seria intuitivamente mais adequado do que um parser estritamente formal, uma vez que a forma do marcador sintagmático seria derivada e, portanto, secundária, em comparação com as relações de licenciamento, por exemplo, de papel temático e de caso, que seriam mais centrais para a teoria linguística, no sentido de que o papel de uma configuração estrutural particular é o de permitir que as relações de licenciamento tenham lugar entre os elementos em uma representação sintática.

O presente estudo fornece evidências experimentais em favor da hipótese de que o sistema de processamento de frases faz uso rápido e eficiente de alguns tipos de informação de natureza lexical associados ao verbo no processamento *on-line*.

Apresentaremos um experimento baseado na teoria da reativação psicolinguística para investigar o processamento do objeto direto em português brasileiro, doravante PB. Diferentes autores demonstraram experimentalmente que diversos tipos de categorias vazias em inglês são psicologicamente reais, isto é, são relevantes perceptualmente no processamento *on-line* de forma análoga aos pronomes. Não há, todavia, registro de investigação experimental anterior a respeito da anáfora objeto no português do Brasil.

A teoria da reativação psicolinguística

A memória de reativação é uma forma não-consciente da memória humana implícita, que desempenha um papel fundamental na identificação perceptual de objetos linguísticos. Segundo Tulvin & Schacter (1990), um experimento psicolinguístico baseado na teoria da reativação desenvolve-se tipicamente em duas etapas. Em um primeiro momento, apresenta-se ao informante o objeto linguístico relevante, por exemplo, uma palavra. Em seguida, apresenta-se para reconhecimento uma pista reduzida ou um estímulo associado, por exemplo, uma palavra que seja semântica ou fonologicamente relacionada à primeira palavra. Diz-se que há reativação se a probabilidade de identificação do estímulo original aumentar significativamente ou se a latência da resposta de identificação for reduzida, em comparação com medidas similares para itens de controle. A diferença entre os desempenhos nas tarefas de reconhecimento do item experimental e do item de controle fornece a medida da magnitude do efeito de reativação.

A ocorrência de efeito de reativação está relacionada à propriedade que têm as unidades cognitivas na memória de apresentarem variação em seu nível de ativação ou excitação. A partir de um nível zero de ativação, uma unidade cognitiva pode ser ativada em razão do processamento perceptual ou lingüístico. Por exemplo, o conceito “mesa” torna-se ativo se alguém vê uma mesa ou ouve a palavra *mesa*. Dizer que um conceito tornou-se ativo é equivalente, em psicolingüística, a dizer que o conceito está sendo processado na memória de trabalho ou memória de curto prazo. Nota-se mesmo na literatura uma certa circularidade no uso desses termos, pois a memória de curto prazo é freqüentemente definida como o sistema de memória que é ocupado por informação ativa (Singer, 1990).

No experimento psicolingüístico que reportamos neste artigo baseamo-nos na teoria da reativação para avaliar a compreensão da relação entre a anáfora objeto e seus antecedentes. A compreensão de relações anafóricas é uma tarefa fundamental do parser sintático, isto é, do mecanismo de processamento de frases. As questões básicas da área da atribuição da co-referência dizem respeito a se, quando e como as relações co-referenciais corretas são estabelecidas entre os componentes de uma frase. Tipicamente, o foco da pesquisa tem se centrado sobre a possibilidade que têm elementos referencialmente dependentes, tais como pronomes lexicais e categorias vazias, de facilitar a compreensão de um sintagma nominal mencionado previamente.

Como indicamos antes, um tema importante da pesquisa recente na área de Processamento de Frases diz respeito à natureza da informação a que o processador tem acesso durante o parseamento da frase. Assim, para alguns autores, tais como Clifton & Frazier (1989), apenas a informação de natureza estritamente sintática estaria imediatamente disponível na compreensão. Para pesquisadores como Tannenhaus & Carlson (1989) e Maryellen MacDonald (1989) o processador teria acesso rápido não apenas à informação sintática, mas também à informação de natureza semântica.

No experimento que ora relatamos estes são temas a serem investigados no que se refere à anáfora objeto em PB. Assim, nosso experimento tem o objetivo de comparar o efeito de reativação da categoria vazia em posição de objeto em PB com o efeito de reativação do pronome lexical na mesma posição de objeto, tomando como referência o processamento do sintagma nominal e de frases intransitivas. Passamos, em seguida, a relatar o experimento.

O experimento

No primeiro experimento, três variáveis foram comparadas: a natureza do objeto direto (categoria vazia, pronome lexical, sintagma nominal), a natureza da oração subordinada em que o objeto se encontrava (complemento/adjunto) e a natureza do papel temático atribuído ao objeto (afetado/não-afetado). Essas condições foram contrastadas ainda com estruturas de controle em que o verbo da oração subordinada era intransitivo. Como veremos, os resultados desse experimento atestam a interação entre informação semântica e estrutura sintática na compreensão do objeto nulo, permitindo concluir que o parser tem acesso imediato à informação sobre a grade temática dos predicados.

Nosso experimento buscou, portanto, encontrar evidências de processamento para duas hipóteses levantadas na literatura sobre o objeto nulo. Note-se que o objeto nulo em português brasileiro vem sendo objeto de pesquisa há mais de uma década do ponto de vista da Teoria Sintática, como atestam, por exemplo, os trabalhos de Teixeira (1985), Galves (1989), Maia (1990), Farrell (1990), Kato (1991), Cyrino (1994). Um tópico que provocou controvérsia na literatura diz respeito à aceitabilidade do objeto nulo vinculado por um antecedente intra-sentencial, como ilustrado pelos exemplos (1) e (2), adiante. Para alguns autores, tal como Teixeira (1985), tanto (1) quanto (2) constituiriam frases perfeitamente gramaticais em PB. Para outros pesquisadores, como Patrick Farrell em trabalho de (1990), apenas a frase (1), em que o objeto nulo encontra-se em uma oração adverbial (estrutura de adjunção) seria aceitável, contrastando com a frase (2) em que o objeto nulo está em uma oração substantiva (estrutura de complementação). A possibilidade de que evidências de processamento pudessem ser trazidas para o debate sobre a aceitabilidade do objeto nulo com um antecedente em posição argumental motivou-nos a incluir entre as variáveis a serem testadas o paradigma adjunto/complemento, exemplificado pelas frases (1) e (2).

Contudo, uma condição de natureza semântica, sugerida inicialmente por Rizzi (1986), a condição do objeto afetado, já havia sido analisada em Maia (1990), como adequada para caracterizar a distribuição do objeto nulo com antecedente argumental em PB. Segundo esta proposta, apenas o objeto nulo com papel temático afetado seria licenciado em PB. Rizzi (1986) define o objeto afetado como aquele que recebe de seu predicado papel temático capaz de implicar modificação em seu estado físico ou psicológico. Assim, por exemplo, verbos como

ver, conhecer, encontrar, saber etc. atribuiriam papel temático não-afetado ao seu objeto, enquanto verbos como *agredir, condenar, re-provar* etc., implicam, em maior ou menor grau, a existência de um processo de afetação sobre o estado físico ou psicológico de seu objeto. Este contraste levou-nos a propor, em Maia (1990), que uma das condições de identificação do objeto nulo em PB seria o requisito de que apenas os núcleos verbais capazes de atribuir papel temático afetado poderiam sancionar o objeto nulo com antecedente intra-sentencial nesta língua, aduzindo, portanto, evidência em favor da relevância sintática da distinção afetado/não-afetado da teoria theta. Restava, entretanto, cruzar o critério adjunto/complemento com o critério afetado/não-afetado para obter um quadro mais claro sobre a distribuição deste tipo de objeto nulo em PB. Foi isto exatamente o que fizemos no experimento que ora relatamos. Em resumo, o nosso experimento procurou responder às seguintes perguntas: (1) o objeto nulo em PB é psicologicamente real, isto é, apresenta um efeito de reativação de seu antecedente? (2) Pode-se identificar alguma diferença de processamento entre o objeto nulo em oração adverbial e o objeto nulo em oração substantiva? (3) Pode-se identificar alguma diferença de processamento entre o objeto nulo com papel temático afetado e o objeto nulo não-afetado?

O experimento utilizou uma tarefa de reconhecimento de sonda (o sujeito da oração principal, que era o antecedente intra-sentencial do objeto nulo e do pronome lexical) e uma técnica de julgamento intermodal (apresentação auditiva das frases com sonda visual). Os tempos de resposta necessários para o reconhecimento da sonda foram registrados para 98 períodos experimentais distribuídos em 14 condições, com sete frases cada, incluindo-se ainda 52 períodos distratores – construções com estruturas diferentes das frases experimentais destinadas a desviar a atenção consciente dos informantes das frases experimentais.

As frases de 1 a 14, abaixo, exemplificam cada uma das condições experimentais testadas. Note-se que as frases foram distribuídas de forma que cada informante testado fosse exposto a apenas uma versão de cada frase. A distribuição das frases em 14 diferentes conjuntos permitiu que todas as sentenças de um mesmo tipo fossem comparadas entre si.

1 Objeto nulo afetado em oração adjunto (ON/Adj)

O rapaz reclamou porque o cobrador agrediu **[nulo]-i**.

- 2 Objeto nulo afetado em oração complemento (ON/Comp)
O rapaz-**i** afirmou que o cobrador agrediu **[nulo]-i**.
- 3 Pronome afetado em oração adjunto (PR/Adj)
O rapaz-**i** reclamou porque o cobrador agrediu **ele-i**.
- 4 Pronome afetado em oração complemento (PR/Comp)
O rapaz-**i** afirmou que o cobrador agrediu **ele-i**.
- 5 Sintagma nominal afetado em oração adjunto (SN/Adj)
O rapaz reclamou porque o cobrador agrediu o estudante.
- 6 Sintagma nominal afetado em oração complemento (SN/Comp)
O rapaz afirmou que o cobrador agrediu o estudante.
- 7 Objeto nulo não-afetado em oração adjunto (ON-NA/Adj)
O rapaz reclamou porque o cobrador encontrou **[nulo]-i**.
- 8 Objeto nulo não-afetado em oração complemento (ON-NA/Comp)
O rapaz-**i** afirmou que o cobrador encontrou **[nulo]-i**.
- 9 Pronome não-afetado em oração adjunto (PR-NA/Adj)
O rapaz-**i** reclamou porque o cobrador encontrou **ele-i**.
- 10 Pronome não-afetado em oração complemento (PR-NA/Comp)
O rapaz-**i** afirmou que o cobrador encontrou **ele-i**.
- 11 Sintagma nominal não-afetado em oração adjunto (SN-NA/Adj)
O rapaz reclamou porque o cobrador encontrou o estudante.
- 12 Sintagma nominal não-afetado em oração complemento (SN-NA/Comp)
O rapaz afirmou que o cobrador encontrou o estudante.
- 13 Verbo intransitivo em oração adjunto (INTR/Adj)
O rapaz reclamou porque o cobrador desmaiou.
- 14 Verbo intransitivo em oração complemento (INTR/Comp)
O rapaz afirmou que o cobrador desmaiou

O estudo usou uma técnica de reativação intermodal. Isto é, as frases foram pré-gravadas e apresentadas oralmente aos participantes. Logo após a apresentação oral de cada frase pelo computador, uma palavra aparecia na tela, devendo nesse momento o participante decidir se aquela palavra havia ocorrido na frase que ouvira apertando a tecla *sim* ou a tecla *não* em um dispositivo à sua frente. Mediam-se então os tempos de reação do participante em milésimos de segundos.

A lógica do experimento era a de que, no processamento da co-referência, os elementos anafóricos provocam a reativação dos seus sin-

tagmas nominais antecedentes na memória de curto prazo do ouvinte. Este efeito de reativação nos permitiria então avaliar se os pronomes lexicais e as categorias vazias em posição de objeto facilitarão o acesso aos seus antecedentes e se esse efeito de facilitação variaria com relação aos parâmetros adjunto/complemento e papel temático afetado/não-afetado. Tanto as orações intransitivas quanto as orações com SN pleno na posição de objeto poderiam servir de base de comparação para o estabelecimento de efeito de facilitação – as orações intransitivas, por não disporem de anáfora objeto para reativar o sujeito da oração principal, e os sintagmas nominais plenos, por ampliarem a carga de processamento, sem fornecer nenhuma pista sobre o SN sujeito cuja reativação se testava.

Os resultados estão expostos na Figura 1. Observe-se de imediato a existência de efeitos de facilitação significativos na compreensão do objeto nulo com papel temático afetado em oração adjunta (subordinada adverbial), não se registrando efeito de reativação significativo nem para o objeto nulo em oração complemento (subordinada substantiva), tanto com papel temático afetado quanto com papel temático não-afetado, e nem para o objeto nulo com papel temático não-afetado em oração adjunta.

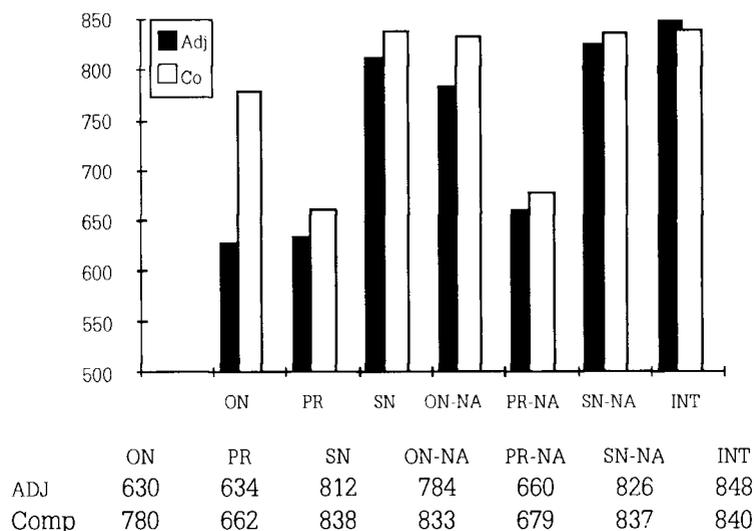


FIGURA 1 – Efeitos de facilitação da anáfora objeto em português brasileiro.

No eixo vertical da Figura 1 registram-se as latências em milésimos de segundos, ou seja, os tempos médios requeridos pelos participantes para decidir se o antecedente da anáfora estava ou não presente na frase. No eixo horizontal registram-se as anáforas: as três primeiras variáveis são o objeto nulo, o pronome lexical e o SN que recebem papel temático afetado; em seguida, registram-se o objeto nulo, o pronome lexical e o SN que recebem papel temático não-afetado. Finalmente, em última posição, estão os tempos médios para as frases com verbo intransitivo, que serviram de base de comparação para avaliar a ocorrência de efeito de facilitação nos demais casos. A altura das colunas indica os valores de cada condição: as colunas claras representam as orações substantivas enquanto as escuras representam as orações adverbiais. Como já dissemos, uma visão geral do gráfico indica de imediato que a divergência efetivamente significativa entre os dois tipos de oração subordinada está restrita à posição da extremidade esquerda, isto é, ao objeto nulo afetado. De fato, apenas o objeto nulo afetado em oração adjunto exibe um efeito de reativação significativo de seu antecedente, enquanto seu correspondente em oração complemento não apresenta latência significativamente distinta da oração intransitiva. Esse resultado oferece confirmação de natureza processual para a análise de Farrell (1990) de que o objeto nulo com antecedente argumental só é licenciado em português brasileiro nas orações subordinadas adverbiais, não ocorrendo nas subordinadas substantivas.

Já no que se refere ao processamento do pronome lexical em posição de objeto afetado – representado na segunda posição da esquerda para a direita na Figura – a diferença adjunto/complemento não parece ser relevante. Em ambos os casos, o pronome apresenta efeito facilitador do reconhecimento do seu antecedente argumental, quando comparado à condição intransitiva na extremidade direita da Figura. Note-se ainda que o efeito de reativação obtido pela anáfora pronominal não se perde quando o verbo atribui papel temático não-afetado, conforme se depreende da Figura. Da mesma maneira se comportam os SN plenos, que são processados nos mesmos tempos médios que os intransitivos, quer recebam papel temático afetado quer não. Em resumo, apenas no caso do objeto nulo podem ser identificadas diferenças no que se refere ao par adjunto/complemento e ao par afetado/não-afetado. Apenas o objeto nulo afetado em posição de adjunto apresenta efeito de reativação de seu antecedente, contrastando tanto com seu correspondente em posição de complemento quanto com seu correspondente não-afetado. Note-se que o objeto nulo não-afetado em oração adjunto não exibe

efeito de reativação de seu antecedente, permitindo concluir que um traço semântico do predicado é acessado no processamento *on-line*.

Conclusão

Os resultados do experimento indicam que o objeto nulo com antecedente em posição argumental só parece ser sancionado em PB em estruturas de adjunção com predicados capazes de atribuir papel temático afetado ao seu argumento interno. Com base em tais resultados, que fornecem subsídios para avançar a compreensão do fenômeno específico do objeto nulo em PB, caberia finalmente indagar o que se poderia ainda inferir a respeito da arquitetura do parser sintático no processamento *on-line*. A nossa proposta é a de que o Mecanismo Humano de Processamento de Frases lida com o fragmento de gramática que focalizamos da seguinte maneira: em primeiro lugar, o parser constrói a representação estrutural da frase, da esquerda para a direita. Ao encontrar um verbo transitivo o parser espera um objeto. Se não há um objeto pleno, isto é, com conteúdo fonológico, o parser verifica se a estrutura temática do predicado tem o traço [+ afetado]. Caso o resultado da busca seja positivo, o parser postula a existência de categoria vazia na posição de objeto. Em seguida, a co-indexação com um antecedente deverá ter lugar. Como sugerido por Nicol (1988), a atribuição da co-referência é levada a efeito pelo módulo de co-referência que constitui um estágio intermediário entre processos puramente estruturais e processos interpretativos, cabendo a este módulo determinar que referentes são antecedentes potenciais de um item dependente referencialmente a fim de acessá-los.

Em resumo, a estratégia de *default* que propomos é a de que o parser só busca acessar a grade temática dos predicados como último recurso, caso não encontre um SN ou um elemento pronominal com corpo fonológico na posição de argumento interno do verbo. Como demonstrado no experimento, o parser não parece se importar com o critério afetado/não-afetado nos casos dos SN's e dos pronomes lexicais, restringindo-se, nestes casos, à atividade de construção da estrutura e só ativando o subcomponente interpretativo caso não haja um candidato aberto para receber o papel temático que o predicado deve descarregar. Postulando-se a existência do objeto nulo, estabelece-se a possibilidade da cadeia de reativação que parece ter sido captada no experimento: a categoria vazia elicitava seu antecedente permitindo o seu

reconhecimento mais imediatamente do que ocorre nos casos dos verbos intransitivos, em que não há elemento elicitador do antecedente.

Agradecimentos

A pesquisa para este artigo foi realizada no Laboratório de Psicolinguística da University of Southern California. Agradeço a Maryellen MacDonald e Mark Seidenberg, por sugestões e comentários. Agradeço ainda comentários de Letícia Sicuro Correia e de outros participantes do VI Congresso da ASSEL-Rio, 1996, onde este trabalho foi apresentado.

MAIA, M. The semantic access in syntactic parsing: experimental evidences. *Alfa (São Paulo)*, v.42, p.101-111, 1998.

- *ABSTRACT: This study presents experimental evidence in favor of the hypothesis that the human parser makes rapid and efficient use of some types of lexical information associated to the verb in on-line processing.*
- *KEYWORDS: Syntactic parsing; object anaphora; Brazilian Portuguese.*

Referências bibliográficas

- CLIFTON, C., FRAZIER, L. Comprehending sentences with long distance dependencies. In: CARLSON, G. N., TANENHAUS, M. K. (Ed.) *Linguistic Structure in Language Processing*. Kluwer: Academic Press, 1989.
- CYRINO, S. *O objeto nulo no português do Brasil*. Campinas, 1994. 227p. Tese (Doutoramento) – Instituto de Estudos da Linguagem, Universidade Estadual de Campinas.
- FARRELL, P. Null Objects in Brazilian Portuguese. *Natural Language and Linguistic Theory*, v.8, p.325-46, 1990.
- FRAZIER, L. *On comprehending sentences: syntactic parsing strategies*. Bloomington: Indiana University Linguistics Club, 1979.
- GALVES, C. *O objeto nulo no português brasileiro: percurso de uma pesquisa*. Campinas, 1989. 227p. Dissertação (Mestrado) – Instituto de Estudos da Linguagem, Universidade Estadual de Campinas.

- GORRELL, P. *Syntax and Parsing*. Cambridge: Cambridge University Press, 1995.
- KATO, M. A. The distribution of pronouns and null elements in object position in Brazilian Portuguese. In: LINGUISTIC SYMPOSIUM ON ROMANCE LANGUAGES, 21, 1991, Santa Barbara.
- MACDONALD, M. C. Priming effects from gaps to antecedents. *Language and Cognitive Processes*, v.4, p.35-56, 1989.
- MAIA, M. *The null object in Brazilian Portuguese*. Los Angeles, 1990. M. A. Dissertation – University of Southern California.
- . *The Comprehension of Object Anaphora in Brazilian Portuguese*. Los Angeles, 1994. Dissertation (Doctoral) – University of Southern California.
- NICOL, J. *Coreference Processing during Sentence Comprehension*. Cambridge, 1988. Dissertation (Doctoral) – Massachusetts Institute of Technology.
- PRITCHETT, B. *Grammatical Competence and Parsing Performance*. Chicago: Chicago University Press, 1992.
- RIZZI, L. Null objects in Italian and the theory of pro. *Linguistic Inquiry*, v.17, p.501-57, 1986.
- SINGER, M. *Psychology of Language: an introduction to sentence and discourse processes*. New Jersey: Lawrence Erlbaum, 1990.
- TANENHAUS, M. K., CARLSON, G. N. Lexical structure and language comprehension. In MARSLEN-WILSON, W. D. (Ed.) *Lexical representation and process*. Cambridge: MIT Press, 1989.
- TEIXEIRA, R. F. *Zero Anaphora in Brazilian Portuguese Subjects and Objects*. UC Berkeley, 1985. Dissertation (Doctoral) – University of California.
- TULVIN, E., SCHACTER, D. L. Priming and Human Memory Systems. *Science*, v.247, p.301-6, 1990.