

Integração de informação semântica e sintáctica no processamento de frases causais – atribuição de referência a sujeitos pronominais

Maria Armanda Costa

FLUL, Onset-CEL

A compreensão da língua escrita culmina na elaboração de uma representação semântica do conteúdo do texto lido integrada em estruturas conceptuais preexistentes. No entanto, para que haja sucesso na compreensão, têm de actuar previamente processos de tratamento da informação verbal que lidam com unidades linguísticas com estatuto categorial e funcional muito diverso, que possuem propriedades intrínsecas e contextuais, e que estão organizadas segundo regras típicas da componente da gramática a que estão afectadas. No estudo que desenvolvemos, interessa-nos particularmente a frase, enquanto unidade cujo processamento requer operações de tratamento que se iniciam com (i) a estruturação do material lexical, criando relações de ordenação, dependência e hierarquia (operações de análise que constituem a frase estrita do *parsing*), (ii) estabelecimento de relações de dependência entre constituintes descontínuos, criando relações dependenciais a longa distância, (iii) atribuição de uma estrutura semântica proposicional pela identificação dos predicadores e respectivos argumentos, (iv) fixação de cadeias de referência, ligando categorias identificadoras a categorias referencialmente dependentes. As duas últimas operações pertencem já ao domínio da interpretação (por oposição ao *parsing*) e é nelas que especificamente se centra o estudo agora apresentado.

O estudo do processamento da referência pronominal tem sido orientado por questões que têm dado lugar à emergência de linhas de investigação complementares. Considerando que a identificação de uma categoria referencialmente dependente impõe o recurso a várias fontes de informação linguística, pretende-se descrever a natureza da informação interveniente, saber como é que essa informação é integrada e, de entre os vários tipos de informação, qual é a mais relevante para uma atribuição eficaz da referência. Pressupondo que o processador opera sobre o *input* em função de pistas de processamento que criam restrições ao *parsing* e à interpretação, um objectivo da investigação é o de identificar as pistas de processamento relevantes para a tarefa de atribuição da referência e determinar o seu grau de validade na resolução de problemas criados, por exemplo, por condições de ambiguidade referencial. Sabendo que o mecanismo humano de processamento da informação funciona dentro de limites impostos pelo equipamento biocognitivo, procura-se identificar princípios e estratégias de processamento relacionados com propriedades gerais e particulares das línguas para a resolução da referência pronominal; pretende-se, ainda, compreender a dinâmica do

sistema, verificando quais os critérios de maximização de proveitos interpretativos e de redução de custos cognitivos que são usados em apoio da eficácia do processamento.

Para controlo do processamento da referência, são interessantes condições em que potencialmente há ambiguidade na selecção de um antecedente para um pronome. Vejam-se os exemplos de (1) a (3), em que há condições estruturais em que, no domínio interfrásico, o sujeito (SU) pronominal da segunda oração precisa de ser identificado por um SN antecedente, havendo uma dificuldade crescente à medida que aumenta o número de potenciais candidatos a essa função.

- (1) O João_i irritou-se porque [-]_i não arranhou bilhete para o concerto.
- (2) ?O João_i viu a Maria_j quando [-]_{ij} entrou para o concerto.
- (3) ?O João_i apresentou a Maria_j ao Luís_k quando [-]_{ij/k} estava no concerto.

Embora um pronome seja sempre livre na fixação da sua referência, o que poderá levar a que tome como seu antecedente uma qualquer categoria nominal existente no discurso, no processamento de frases em tempo real há automaticamente a busca de um antecedente no contexto próximo, no material existente em memória de trabalho. Assim, se em (1) está fechado o elenco de possíveis antecedentes para o pronome nulo, SU da frase subordinada, o que facilita a fixação da cadeia de correferência, em (2) e (3), há um conjunto de possíveis candidatos que se alinham legitimamente para a função de antecedentes do pronome. Há então um problema de processamento que será resolvido em função de informação estrutural, discursiva e, hipoteticamente, com a ajuda de estratégias de processamento que possam ser produtivas nestas condições.

Em termos de processamento, podemos dizer que há um problema de referência sempre que o processador encontra um pronome, visto que não há conteúdo referencial próprio mas dependente. As condições existentes para a identificação do seu conteúdo referencial são diversificadas, variando o seu grau de complexidade em função das pistas linguísticas de suporte. A gramática estipula as condições sintácticas e semânticas em que um pronome tem de ser livre no domínio frásico em que ocorre ou em que é obrigatoriamente ligado, distinguindo entre pronomes e anáforas em sentido sintáctico estrito (cf. *Binding Theory*, Chomsky, 1981). A análise das condições semânticas e discursivas determina o grau de acessibilidade de possíveis antecedentes para o pronome e estabelece escalas de acessibilidade do antecedente em função da natureza lexical da forma anafórica (Nicol e Swinney, 1989; Ariel, 1996; Fretheim e Gundel, 1996; Batoréo & Costa, 1997; Costa, 2003). A informação morfológica contribui comprovadamente na formação de cadeias referenciais através de traços flexionais de pessoa, género, número e caso (McDonald e MacWhinney, 1995; De Vincenzi 1999; Arnold et al., 2000). Veja-se que, nos exemplos de (6) a (8), apesar de o SU da subordinada ser sempre um pronominal nulo, não há qualquer ambiguidade na atribuição do seu antecedente.

- (6) a. O João viu a Maria quando entrou apressada no cinema.
b. O João viu a Maria quando entrou apressado no cinema.

- (7) a. O João viu os amigos mas ignorou-os ostensivamente.
 b. O João viu os amigos mas ignoraram-no ostensivamente.
- (8) a. O João apresentou a Ana aos amigos quando os encontrou no concerto.
 b. O João apresentou a Ana aos amigos quando a encontrou no concerto.

Em PE, parece funcionarem pistas específicas para a resolução da ambiguidade referencial em determinadas condições. Como exemplificado de (9) a (11), a forma lexical do pronome sujeito, nula ou realizada, parece funcionar como pista restritiva que designa o seu antecedente.

- (9) a. A Marta_i acha que [-]_i vai fazer uma grande viagem nas férias.
 b. A Marta_i acha que ela_{i/k} vai fazer uma grande viagem nas férias.
- (10) a. O João_i encontrou o Pedro_j quando [-]_{i/j} entrou no cinema
 b. O João_i encontrou o Pedro_j quando ele_{i/j} entrou no cinema
- (11) a. A Marta_i apresentou a Joana_j à Luísa_k mas [-]_{i/j/k} não a cumprimentou
 b. A Marta_i apresentou a Joana_j à Luísa_k mas ela_{i/j/k} não a cumprimentou.

A matriz fonética do SU pronominal de uma segunda oração na frase complexa parece funcionar como pista restritiva para a sua interpretação. Sempre que o SU pronominal assume uma forma nula, a interpretação preferencial é pela correferência entre SUs oracionais; a forma realizada parece ser pista assinaladora de referência disjunta. Este funcionamento tem sido interpretado como uma manifestação do designado *princípio evitar o pronome*, entendido mais como um princípio conversacional do que estrutural que interage com a gramática mas que não faz parte do seu núcleo, comportando-se como uma realização específica de princípios gerais envolvendo menos esforço e menos tempo (Chomsky, 1981). Em PE, o seu funcionamento tem sido estudado no âmbito da gramática (Brito, 1991; Barbosa, 1995; Barbosa et al., 2000; Mateus et al., 2003) e do processamento de frases (Costa et al., 1997). No caso em concreto, podemos também interpretar este funcionamento como uma consequência das predições da *teoria da acessibilidade*, que estabelece que formas anafóricas mínimas (no caso o pronome nulo) são as mais adequadas para designar antecedentes maximamente activos. Nas condições em análise, o SU da frase precedente, ao ocupar a posição estrutural mais alta, ser tópico e primeira referência, ganharia um estado de activação mais alto do que outras expressões nominais ocorrentes na frase. (cf. *The accessibility marking scale* in Ariel, 1996).

Em Costa et al. (1997), controlámos o funcionamento do *princípio evitar o pronome* no processamento de construções coordenadas, numa tarefa *off-line*. No presente estudo, verificamos se o mesmo princípio funciona como estratégia de processamento produtiva no processamento *on-line* de frases causais onde está disponível uma fonte semântica suplementar: as propriedades léxico-semânticas dos designados *verbos de causalidade implícita* (VCI).

Informação léxico-semântica verbal e atribuição da referência

No processamento de frases, a informação do predicador verbal tem sido controlada como pista válida para determinar a interpretação. O entendimento sobre a sua força de processamento distribui-se entre os que lhe atribuem um papel decisivo e competitivo com a informação sintáctica desde as fases iniciais da computação e os que consideram que o impacto das propriedades léxico-semânticas do verbo só se faz sentir numa fase ulterior à atribuição de uma representação estrutural (cf. Boland, 1997; Tanenhaus, Boland et al., 1993; Tanenhaus et al., 2000).

É certo que as propriedades de selecção semântica e sintáctica do verbo projectadas na respectiva construção verbal constituem *output* crucial para a interpretação. No caso da atribuição da referência a pronomes, a informação verbal pode determinar inequivocamente a cadeia referencial. É o que acontece com os chamados *verbos de controlo*, como em (12) e (13), onde a violação de condições estruturais estritas é causadora de agramaticalidade e bloqueadora da interpretação.

- (12) a. A mãe_i prometeu à Joana_j [-]_{v*}_j deixá-la ir à festa de finalistas.
 b. A mãe_i autorizou a Joana_j a [-]_v_j ir à festa de finalistas
- (13) a. Os professores_i prometeram aos alunos_j [-]_{v*}_j chegar antes do toque.
 b. Os professores_i impediram os alunos_j de [-]_v_j entrar depois do toque.

Contrariamente aos verbos de controlo, os VCI representam uma classe semântica de predicadores interessante para o estudo da atribuição da referência pronominal, dado que possuem propriedades semânticas que restringem a interpretação, mas não criam condições de agramaticalidade estrita. Trata-se de verbos transitivos que possuem um significado inerente de causalidade; descrevem estados, acções ou realizações. Estes verbos possuem uma grelha temática com dois argumentos que têm entre si uma relação de causalidade: um possui o papel temático de *causador* (estímulo) e o outro o papel de *paciente*, ou de *experienciador*, dado que muitos destes verbos, por serem interpessoais, são também verbos psicológicos. Os papéis temáticos atribuídos pelos VCI relacionam-se com as funções sintácticas principais conforme o tipo de verbos: V1, se atribuem o papel de causador ao SU, V2 se atribuem esse papel ao OD (Garvey & Caramazza, 1974; Fiedler & Semin, 1988; Segui & Kail, 1991; McDonald & MacWhinney, 1990; Garnham et al., 1996; Arnold et al., 2000; Long & Ley, 2000; Stevenson et al., 2000). Em (14) temos o verbo *assustar* como paradigmático de VCI orientado para o SU (V1); na sua grelha argumental, o causador é SU, sendo o OD experienciador. Em (15), *admirar* representa o VCI orientado para o Objecto (V2); na sua grelha argumental o causador é agora o OD e o SU é experienciador. Em (16), *ver* apenas representa um verbo perceptivo que, contrariamente a *assustar* e a *admirar*, não possui informação sobre causalidade.

- (14) A Marta assustou a Joana porque empalideceu de repente.
 (15) A Marta admirou a Joana porque fez uma prova espectacular.
 (16) A Marta viu a Joana porque foi passear para o jardim.

Em termos da informação disponibilizada por cada tipo de verbo, parece haver uma certa assimetria entre V1 e V2, quanto ao papel temático de *causador*. Com verbos de causalidade implícita orientada para o SU, como em *assustar*, o causador parece possuir propriedades semânticas mais prototípicas do causador agentivo, do que quando ocorre em construções com V2 onde se associa à posição de OD, como em *admirar*. Se testarmos os adjectivos que podem ser derivados de V1 e de V2, verificamos que os primeiros são claramente mais agentivos no sentido de que possuem traços dinâmicos próprios de agente (*ser assustador, encantador, perturbante, irritante, surpreendente*), enquanto que os adjectivos derivados de V2 possuem traços mais estativos que o colocam mais próximo do paciente (*ser admirável, detestável, temível, odioso, censurável, desprezível*).

O efeito da informação semântica do verbo é o de focalizar o argumento *causador* colocando-o em condições privilegiadas de ser o antecedente do SU pronominal da subordinada onde é explicitada a relação de causalidade. Ora, da assimetria entre V1 e V2, há consequências para a dimensão aspectual do enunciado, importante na interpretação. Em V1, o causador mais agentivo é mais proeminente aspectualmente, ficando por isso mais acessível para o controlo da referência (Grimshaw 1990; Pesetsky, 1995).

Estudo experimental

De forma a avaliar a força das pistas de processamento em jogo, realizámos duas experiências, a primeira de produção, a segunda de compreensão. Partimos dos pressupostos teóricos do *Modelo de Competição* para o processamento de frases (Bates e MacWhinney, 1989, Kail, 1999), uma instância dos *constraint-based models* (Altmann e Steedman, 1988; MacDonald et al. 1994; Tanenhaus e Trueswell, 1995). Neste modelo, defende-se a existência de um processador que opera guiado por informação linguística que releva de todas as componentes da gramática, sob a forma de pistas de processamento. As pistas são formas cujas propriedades linguísticas e perceptivas lhes conferem validade suficiente para designar funções linguísticas relevantes para a interpretação. As pistas são integradas de forma interactiva e probabilística, impondo a sua integração custos cognitivos variáveis em função da sua detectabilidade e localidade. O grau de validade das pistas é calculado em termos da sua frequência, disponibilidade e fiabilidade. A convergência ou competição entre pistas cria condições favoráveis ou desfavoráveis à interpretação. O processador incorpora informação relativa a frequência e a restrições contextuais.

Na investigação realizada pretendemos captar o modo como interagem duas fontes de informação essenciais à compreensão da linguagem e ao processamento da frase: a informação léxico-semântica, relativa às propriedades temáticas dos verbos de causalidade implícita e a informação sintáctica que decorre de propriedades particulares do PE para codificação da referência pronominal. Fazendo interagir a informação semântica do verbo com a natureza lexical da forma pronominal, criam-se condições

interessantes para controlar o processo de atribuição da referência ao pronome. Nas frases (17) e (18), temos condições de convergência e de competição de pistas de processamento (condições a. e b. respectivamente).

- (17) a. O João_i desiludiu o Pedro_j porque [-]_i; não o deixou copiar no exame.
 b. O João_i desiludiu o Pedro_j porque ele_i não o deixou copiar no exame.
 (18) a. O João_i detestou o Pedro_j porque ele_j não o deixou copiar no exame.
 b. O João_i detestou o Pedro_j porque [-]_j; não o deixou copiar no exame.

Hipóteses

Em termos gerais, para o processamento de frases causais com VCI e ambiguidade de referência, prediz-se que:

(i) frases com verbos que focalizam semanticamente o SU atribuindo-lhe o papel de causador (*desiludir*) possuem condições mais favoráveis para a resolução da referência do SU pronominal da oração subordinada, porque se criam condições de convergência sobre SN1 de pistas sintácticas, semânticas e informacionais;

(ii) frases subordinadas com SUs pronominais nulos são mais fáceis de processar do que frases com SUs pronominais lexicalizados; a ocorrência de pronomes lexicalizados cria sobrecarga para o processador, quer em termos de processamento adicional quer porque induz mudança da estratégia mais frequente: coindexação de SUs oracionais.

Experiência 1

Visto que em PE não dispúnhamos de dados que confirmassem a existência de VCI, seguimos a metodologia usada para outras línguas em estudos prévios (Garvey e Caramazza, 1974; Fiedler e Semin, 1988; Kail, 1979; Segui e Kail, 1991; Garnham et al., 1996; Stevenson et al., 2000). Realizámos uma primeira experiência com o objectivo de verificar a existência de VCI em PE e avaliar a validade da sua informação semântica no processamento de frases. Pretendemos, ainda, avaliar de que forma a informação semântica deste tipo de verbos interage com a pista sintáctica criada pelo *princípio evitar o pronome*.

Tendo adoptado uma tarefa de produção (completamento de frases), formularam-se especificamente as seguintes hipóteses:

(i) em frases complexas causais que têm como oração principal uma construção transitiva com um VCI, o SU pronominal da subordinada correferente com o argumento da oração precedente que recebe o papel temático de *causador*;

(ii) o facto de o papel de *causador* poder ser atribuído ao SU ou ao OD da oração principal deverá ter consequências na forma fonética do pronome SU da subordinada: a forma nula é preferencialmente a escolhida quando há correferência entre SUs;

(ii) prevê-se que haja interacção entre o tipo de V e o tipo de pronome SU: oração

subordinante com V1, oração subordinada com SU pronominal nulo; oração subordinante com V2, oração subordinada com SU pronominalmente realizado.

Desenho experimental

Foram seleccionados 100 verbos transitivos, no pressuposto de possuírem significado implícito de causalidade. Desses 100 verbos, 50 eram supostos atribuidores de causalidade ao SU (V1), os outros 50 atribuidores de causalidade ao OD (V2). Construíram-se 100 frases incompletas que se suspendiam na conjunção causal *porque*, segundo o padrão exemplificado em (19)

- (19) SP ou SAdv, SN V SN porque...
 No tribunal, a Graça desiludiu o Paulo porque...

Os dois argumentos do verbo transitivo da oração subordinante possuem Nomes Próprios como núcleo de modo a eliminar possíveis efeitos semânticos que criassem enviesamentos na interpretação. De modo a controlar melhor a retoma do antecedente pelo pronome, foram usados pares de Nomes masculino/feminino, como em (20) e (21); a ordem de género foi balanceada.

- (20) No tribunal, a Graça desiludiu o Paulo porque...
 (21) No plenário, o Joaquim detestou a Sara porque...

Procedimento

As cem frases foram apresentadas em cadernos individuais, havendo uma frase por página de modo a evitar que o sujeito consultasse respostas anteriores. Fizeram-se 8 listas, correspondentes a oito ordens de apresentação dos 100 itens. A tarefa de processamento foi realizada em pequenos grupos. Aos sujeitos foi dito que lessem atentamente a parte inicial de cada frase e que a completassem de uma forma lógica, tão rapidamente quanto possível, evitando voltar atrás.

Participaram 60 adultos universitários, falantes nativos do PE. Cada sessão demorou cerca de 45 minutos.

Resultados

Os resultados obtidos confirmaram inequivocamente a existência destes dois grupos de verbos em PE e a validade da pista semântica na focalização do argumento *causador*. Os resultados mostram também um efeito forte do *princípio evitar o pronome* como estratégia para evitar quaisquer condições de ambiguidade referencial. Como se pode ver no Quadro 1, na totalidade das frases em que se empregou um V1, em 96% dos casos usou-se um pronome nulo como SU da causal. Pelo contrário, nas frases com

V2, em 81 % usou-se um pronome realizado como SU da subordinada para correferir com o OD da oração precedente.

Quadro 1

VCI	Tipo de pronominal SU	frequência	percentagens
V1	pro (pronome nulo)	2814	96%
	PRON (pronome realizado)	110	4%
V2	PRON (pronome realizado)	2225	81%
	pro (pronome nulo)	528	19%

Confirma-se a existência das duas classes de verbos possuidores de informação sobre causalidade implícita em PE: com V1, o SU é causador, com V2 o causador é o OD. O argumento focado semanticamente na primeira oração é o identificador privilegiado do conteúdo referencial do SU pronominal da subordinada. As cadeias referenciais interfrásicas foram marcadas de forma não ambígua: a forma nula do pronome é preferida sistematicamente à forma lexicalizada sempre que há correferência entre SUs, enquanto que a opção por formas lexicalizadas, sobretudo pronomes (há um emprego não significativo de SNs semanticamente relacionados), indica referência disjunta entre SUs.

Não se verifica a interação prevista entre informação sintáctica e semântica, como tinha sido predito nas hipóteses. Enquanto que com V1 o *princípio evitar o pronome* funciona sistematicamente para assinalar a correferência entre SUs oracionais, com V2, embora haja uma preferência significativa pela forma realizada, há contudo uma percentagem considerável de casos em que a forma nula é tomada como adequada para assinalar a correferência entre o SU da subordinada e o OD da subordinante (19%). Além disso, se tivermos em consideração a classificação de verbos que emerge dos resultados, e que é apresentada no quadro 2 (apenas os primeiros 20 verbos de cada categoria), também aí se verifica maior dispersão na classificação de V2.

Enquanto que os verbos de tipo V1 são usados como tal entre 100% e 85%, os de tipo V2 apresentam uma distribuição mais dispersa entre 97% e 73%. A isto se acrescenta uma estratégia muito interessante que identificámos em alguns sujeitos e que não foi sistematicamente tratada. Acontece que, com alguns verbos de tipo V2, a frase causal é continuada assegurando-se a correferência entre SUs oracionais, através de uma estratégia sintáctica e discursiva de adiamento da explicitação da fonte de causalidade. O processo usado é o do encaixe de uma oração completiva com um verbo de tipo epistémico, como de (20) a (22), ou verbos predicativos de OD, como em (23) e (24).

Quadro 2

VCI sobre o SU		VCI sobre o OD	
1. desiludir	100%	1. adorar	97%
2. encantar	98%	2. detestar	97%
3. aborrecer	97%	3. temer	97%
4. desorientar	97%	4. aplaudir	95%
5. aterrorizar	95%	5. admirar	93%
6. chocar	95%	6. recriminar	93%
7. surpreender	95%	7. invejar	93%
8. comover	93%	8. apreciar	90%
9. escandalizar	93%	9. gozar	90%
10. irritar	93%	10. criticar	87%
11. perturbar	93%	11. reconhecer	87%
12. desesperar	92%	12. censurar	88%
13. emocionar	92%	13. lamentar	83%
14. enfurecer	92%	14. recompensar	88%
15. impressionar	92%	15. distinguir	88%
16. afligir	90%	16. elogiar	85%
17. frustrar	90%	17. odiar	85%
18. conquistar	90%	18. rezear	83%
19. cativar	90%	19. condenar	83%
20. envergonhar	90%	20. valorizar	82%
21. preocupar	87%	21. preferir	80%
22. assustar	85%	22. rejeitar	77%
23. prejudicar	85%	23. amaldiçoar	73%
24. divertir	85%	24. desprezar	73%

- (20) Face à análise de resultados, a Natália despromoveu o Roberto porque descobriu que ele mentiu. (S2, 2.66)
- (21) Na véspera do exame, o Gustavo sossegou a Rosário porque sabia que ela conseguia tudo o que queria. (S14, 2.92)
- (22) No átrio de entrada, a Daniela ameaçou o Simão porque suspeitava que ele lhe tivesse roubado o carro. (S16, 1.09)
- (23) Após as entrevistas, a Verónica preferiu o Torcato porque o considerou o mais eficiente. (S7, 2.94)
- (24) No fim-de-semana, a Antónia conquistou o Estêvão porque o achou muito charmoso e simpático.

Interpretamos tal estratégia como um reforço do que nos parece constituir uma estratégia por defeito no processamento de frases complexas: esperar que o SU da frase principal seja retomado como antecedente de qualquer categoria referencialmente dependente, isto porque ele ocupa uma posição dominante estrutural e cognitivamente. Em termos do Modelo de Competição, dir-se-á que com VI há convergência de pistas sintácticas e semânticas em SNI e que isso o promove em termos de activação, com consequências fortes para a interpretação.

Experiência 2

Na segunda experiência, utilizando como material experimental frases causais com as mesmas propriedades das usadas na experiência 1, recorreremos a uma tarefa de identificação de palavra subliminar (*visual probe task*) para captar o nível de activação do possível antecedente para o SU pronominal da subordinada. Trata-se de uma tarefa que capta o processamento *on-line*. Isto é, para o mesmo processo de atribuição da referência testado na produção de frases na experiência 1, tipicamente uma tarefa *off-line* em que o sujeito decide com base em toda a informação disponível e com tempo não limitado, agora pretende-se saber que pista de processamento é mais válida para a identificação do pronome antes de toda a informação ser computada.

Materiais

O material experimental foi constituído por 72 frases causais com a estrutura das usadas na experiência anterior (para uma descrição das frases causais usadas, veja-se Lobo (2000) e Mateus et al. (2003)). Seleccionaram-se 48 verbos com informação sobre causalidade confirmados no teste de produção: 24 V1 e 24 V2. Usamos ainda 24 verbos sem traços semânticos de causalidade (declarativos, perceptivos e outros); foram acrescentados 36 itens distractores. Cada lista experimental era constituída por 108 itens.

Para o desenho experimental, tomaram-se como variáveis linguísticas (i) a classe de VCI (3 níveis: V1, V2 e V3), (ii) a natureza lexical do pronome, SU da subordinada (2 níveis: nulo vs. realizado), (iii) a *probe word* que designa o primeiro Nome (W1) o segundo (W2) e, no caso das frases distractoras, um Nome não existente na frase (W3). Tratando-se do processamento *on-line* onde se quer controlar exclusivamente a contribuição da informação semântica e sintáctica, neutralizaram-se possíveis efeitos na percepção e reconhecimento dos Nomes Próprios, através de um emparelhamento rigoroso dos dois Nomes na mesma frase, considerando a sua extensão em número de caracteres, o seu formato fonológico em termos de número e estrutura das sílabas e do padrão acentual; controlou-se ainda a frequência dos Nomes na língua.

A frase causal foi apresentada num ecrã de computador, com apoio no programa PsyScope para Macintosh (Cohen et al, 1993), em modo não cumulativo, segmentada em grupos funcionais. Em (22), apresenta-se o padrão de segmentação segundo as duas condições de P, e o ponto de inserção da *probe word* (W); em (23), exemplifica-se a variação de W. O tempo de exposição de cada segmento foi calculado a partir de dados recolhidos numa experiência complementar (leitura auto-segmentada ao ritmo do próprio sujeito, *self-paced reading task*); foi elaborado um algoritmo com base num tempo fixo acrescido de um tempo variável em função do número de palavras de cada bloco (cf. Costa, 2003).

- (22) a. SP / SN / V / SN / porque / Pron / V / **W** / SX
 b. SP / SN / V / SN / porque / V / **W** / SX

- (23) a. A Teresa desiludiu a Susana porque reprovou ^{Teresa} no exame de matemática.
 b. A Teresa desiludiu a Susana porque reprovou ^{Susana} no exame de matemática.

O ponto de inserção da *probe word* foi determinado pela condição de SU nulo. Optou-se pela sua inserção logo após o verbo da segunda oração, garantindo-se assim que a informação do verbo da oração principal já estivesse integrada e projectada na sua grelha argumental e que, no momento da percepção visual da *probe*, não houvesse ainda informação que pragmaticamente criasse restrições à interpretação. A *probe* permaneceu no ecrã durante 1800ms. A tarefa de processamento consistiu em dizer tão rapidamente quanto possível, premindo uma tecla, se a *probe* designava uma palavra já percebida na frase em leitura. Participaram 40 adultos universitários, falantes nativos do PE. Cada sessão demorou cerca de 30 minutos.

Resultados

Considerando os dados globais, e independentemente das variáveis V e P, há um efeito significativo de SN1 sobre SN2: $F_1(1,39)=9.3972$, $p<0.003$, $F_2(1,69)=18.664$, $p<0.0000$. Sempre que a *probe* designa SN1, os TR são significativamente mais baixos do que quando designa SN2. Tomando o TR como uma variável dependente que indica o nível de activação do SN, tal resultado indica que SN1 se encontra num patamar de acessibilidade mais alto que SN2, independentemente de qualquer das condições manipuladas.

Quadro 3

VCI	Tipo de pronominal SU	Probe Word 1	Probe Word 2
V1	pro	733 ms	767 ms
	PRON	734 ms	807 ms
V2	pro	765 ms	767 ms
	PRON	749 ms	785 ms
V3	pro	747 ms	759 ms
	PRON	791 ms	756 ms

Considerando a variável tipos de verbos (V) (cf. quadro 3), confirma-se a integração da pista semântica no ponto de inserção da *probe*, mas com efeito significativo apenas quando converge com a pista estrutural. Só no caso de frases com V1, é que a *probe* elicitava tempos de reacção significativamente mais curtos quando designa SN1 do que quando designa SN2: $F_1(1,39)=11.414$, $p<0.001$; $F_2(1,23)=16.705$, $p<0.0005$. No caso de V2, as diferenças não são significativas: $F_1(1,39)=2.067$, $p<0.1555ns$, $F_2(1,23)=3.436$, $p<0.07$. De acordo com indicadores que já haviam sido recolhidos na experiência 1, a pista semântica não tem as mesmas

consequências no processamento de frases com V1 ou com V2. As propriedades semânticas de V1 fazem coincidir a atribuição do papel temático mais alto na hierarquia aspectual (Grimshaw, 1990), que corresponde a um verdadeiro *causador* (Pesetsky, 1985), com a função gramatical de SU desempenhada pelo SN que ocupa o lugar estrutural mais alto, o de especificador do SV. Em V2, o *causador*, mais fraco nos seus traços agentivos, é atribuído ao SN/OD mais encaixado na estrutura, havendo por isso uma distribuição de pistas de processamento. Apesar disso, a força semântica de V2 em competição com a pista semântica activou o SN2. Disso é evidência a semelhança dos TR para W1 e W2, de outro modo, haveria uma réplica dos resultados de V1.

Quanto ao efeito da variável P, nas duas condições de Pronome, W1 dá lugar a TR mais baixos do que W2. Na condição de pronome realizado, as diferenças são significativas: $F_1(1,39)=8.049$, $p<0.007$, $F_2(1,69)=12.4603$, $p<0.0008$; na condição de pronome nulo, as diferenças são menos acentuadas mas ainda significativas: $F_1(1,39)=4.819$, $p<0.03$, $F_2(1,69)=5.634$, $p<0.01$. Verificando o que acontece no interior de cada categoria de V, apenas captámos efeitos do *princípio evitar o pronome* em V1 e em V3. No primeiro caso, a activação de SN/OD por W2 conjugada com pronome realizado, faz subir significativamente o TR: 807ms, o tempo obtido mais alto (v. quadro 3). Note-se que em V1, é SN1 que deverá estar maximamente activado pelas pistas semântica e sintáctica em convergência. Quando a *probe word* designa SN2, e mais ainda na condição de pronome realizado, que indica correferência disjunta do SU, o processador fica em sobrecarga, sendo disso prova a forte subida do TR.

É apenas com os verbos sem traços de causalidade, V3, que a alternância pro/PRON surte efeito significativo. Quando a *probe* designa SN1, a condição de Pronome realizado provoca subida significativa de TR se comparada com os resultados obtidos para a condição de pronome nulo: 747ms contra 791ms, $F_1(1,39)=3.902$, $p<0.05$; $F_2(1,23)=3.288$, $p<0.07$. Poderá dizer-se que só em condições em que não há interferência da informação semântica verbal é que a codificação por via do *princípio evitar o pronome* encontra condições para se manifestar inequivocamente.

Conclusões

Os resultados obtidos permitem-nos extrair conclusões sobre alguns dos aspectos que tentámos esclarecer neste estudo. O primeiro tem a ver com o peso da informação semântica relativa à grelha argumental dos verbos de causalidade implícita. O que verificámos é que o seu efeito é diferente em função da tarefa de processamento. No processamento *off-line* é claramente usada levando à definição inequívoca das duas classes de verbos. Já no processamento *on-line*, o seu efeito só é evidente e significativo para os verbos que atribuem o papel de causador ao SU frásico. Estas são condições em que, segundo o Modelo de Competição, há convergência de pistas semânticas e sintácticas que activam o primeiro SN que linearmente ocorre na frase como controlador de processos sintácticos, morfológicos e semânticos que possam ocorrer em posições mais encaixadas na frase. Considerando as hipóteses teóricas que têm sido colocadas para o processamento de frases, subjacentes aos conceitos de estruturação

sintáctica e de interpretação, talvez possamos concluir que a pista semântica relativa a causalidade é integrada numa fase final posterior à informação sintáctica. Tais resultados vão no sentido da hipótese integrativa contra a de focalização (*integrative account vs. focusing account*) como discutido em Stewart et al., 2000. Com os resultados obtidos, poderemos dizer que a pista semântica só ganha força de processamento quando em convergência com a pista sintáctica.

Quanto ao funcionamento do que considerámos ser uma forma particular de codificar cadeias de correferência em PE, o *princípio evitar o pronome*, também revelou ter um comportamento diferente na produção *off-line* e na compreensão *on-line*. No domínio da frase complexa, a forma nula de pronome parece ser suficiente para designar os SNs, SU e OD que pode tomar como seus antecedentes. Contudo, no processamento *off-line*, a alternância nulo/realizado é sem dúvida uma pista produtiva para assinalar cadeias preferenciais de correferência. Já no processamento *on-line*, é uma pista muito pouco robusta. E isto poderá explicar-se por um princípio operativo no processamento do PE, que prediz que durante a computação de uma cadeia referencial, a forma nula é pista suficiente para indicar o antecedente se este está localizado na oração precedente. A ocorrência de pronome realizado funciona como instrução que contraria a estratégia por defeito: referência disjunta entre SUs oracionais. É disso prova a acentuada subida do TR sempre que ocorre pronome realizado como SU da segunda oração, sugerindo que nestas condições há informação suplementar a ser tratada que não é esperada pelo processador.

Em suma, os resultados sugerem que, apesar da confirmação da validade de cada pista, a sua integração é diferente em função das restrições impostas pela tarefa de processamento. Na compreensão *on-line* há aparente preferência pela integração de pistas em convergência que designam fortemente uma dada forma linguística para uma dada função (um SN para Antecedente de um pronome). Por outro lado, a estratégia mais previsível e com menos custos é a adoptada. Só isso explica que a pista semântica sobre CI só produza efeitos quando conjugada com a pista estrutural (*SU/Causador*). Por outro lado, o facto de a ocorrência do pronominal realizado elicitar tempos de reacção mais altos, parece ser indicador de que a forma nula é a esperada. A forma realizada é tratada como pista redundante ou que contraria a estratégia por defeito: correferência entre SUs e manutenção do tópico.

Referências Bibliográficas

- Altmann, G.T.M. e Steedman, M. J. (1988). Interaction with context during human sentence processing. *Cognition*, 30, 191-238.
- Ariel, M. (1996). Referring expressions and +/- coreference distinction. In T. Fretheim e J. Gundel (Eds.). *Reference and referent accessibility*. Amsterdam: John Benjamins Publishing Company.
- Arnold, J. E., Eisenband, J. G., Brown-Smith, S. & Trueswell, J. C. (2000). The rapid use of gender information: evidence of the time course of pronoun resolution from eyetracking. *Cognition*, 76, B13-B26.

- Barbosa, M. P. (1995) *Null subjects*. MIT, Cambridge, Mass.
- Barbosa, M. P., Duarte, M. E. L. & Kato, M. (2000). A distribuição do sujeito nulo no português europeu e no português brasileiro. *Actas do XVI Encontro da Associação Portuguesa de Linguística*. Coimbra: APL
- Bates, E. e MacWhinney, B. (1989). Functionalism and the competition model. In B. MacWhinney e E. Bates (Eds.). *The cross-linguistic study of sentence processing*. New York: Cambridge University Press.
- Batoréo, H. e Costa, A. (1997). Referência nominal na narrativa oral e escrita aos dez anos de idade. *Actas do XIII Encontro Nacional da Associação Portuguesa de Linguística*, vol. I, 137-149. Lisboa
- Boland, J. (1997). The relationship between syntactic and semantic processes in sentence comprehension. *Language and Cognitive Processes*, 12 (4), 423-484.
- Brito, A. M. (1991) Ligação, co-referência e o princípio evitar o pronome. *Actas do Encontro de Homenagem a Óscar Lopes*, 101-121.
- Chomsky, N. (1981). *Lectures on government and binding*. Dordrecht: Foris.
- Cohen, J., MacWhinney, B., Flatt, M. & Provost, J. (1993). PsyScope: A new graphic interactive environment for designing and controlling experiments in the psychology laboratory using Macintosh computers. *Behavior Research Methods, Instruments & Computers*, 25 (2), 257-271.
- Costa, A. (2003). *Processamento de frases em Português Europeu – informação morfológica, sintáctica e semântica em competição*. Tese de doutoramento. FLUL.
- Costa, A., Faria, I. H. & Matos, G. (1997). Competitive information sources in referential ambiguity resolution. *Proceedings of 5th International Congress of the ISAPL*. Porto.
- De Vincenzi, M. (1999). Differences between the morphology of gender and number: evidence from establishing coreferences. *Journal of Psycholinguistic Research*, 28 (5), 537-553.
- Fiedler, K. & Semin, G. R. (1988). On the causal information conveyed by different interpersonal verbs: The role of implicit sentence context. *Social Cognition*, 6 (1), 21-39.
- Fretheim, T. & Gundel, J. (Eds.) (1996). *Reference and referent accessibility*. Amsterdam: John Benjamins Publishing Company.
- Garnham, A., Traxler, M., Oakhill, J. & Gernsbacher, M. A. (1996). The locus of implicit causality effects in comprehension. *Journal of Memory and Language*, 35, 517-543.
- Garvey, C. & Caramazza, A. (1974) Implicit causality in verbs. *Linguistic Inquiry*, 5, 459-464.
- Kail, M. (1999). Linguistic variation and cognitive constraints in the processing and the acquisition of language. In C. Fuchs e S. Robert (Eds.). *Language diversity and cognitive representations*. Amsterdam: John Benjamins Publishing Company.
- Lobo, M. (2000). Para uma sintaxe das orações causais do português. *Actas do XVI Encontro Nacional da Associação Portuguesa de Linguística*, 291-306.

- Long, D. L. & De Ley, L. (2000). Implicit causality and discourse focus: the interaction of text and reader characteristics in pronoun resolution. *Journal of Memory and Language*, 42, 545-570.
- MacDonald, M. C., Pearlmutter e Seidenberg, M. S. (1994). Syntactic ambiguity resolution as lexical ambiguity resolution. In C. Clifton, L. Frazier e K. Rayner (Eds.). *Perspectives on sentence processing*. Hillsdale, NJ: LEA, Publishers.
- McDonald, J. L. e MacWhinney, B. (1995). The time course of anaphor resolution: effects of implicit verb causality and gender. *Journal of Memory and Language*, 34, 543-566.
- MacDonald, M. C. & MacWhinney, B. (1990) Measuring inhibition and facilitation from pronouns. *Journal of Memory and Language*, 29, 469-492.
- Mateus, M. et al. (2003) *Gramática da Língua Portuguesa*. Lisboa: Caminho.
- Nicol, J. L. & Swinney, D. (1989). The role of structure in coreference assignment during sentence comprehension. *Journal of Psycholinguistic Research*, 18 (1), 5-20.
- Pesetsky, D. (1995) *Zero syntax. Experiencers and cascades*. Cambridge: The MIT Press.
- Segui, J. & Kail, M. (1991) Pronoun assignment in the processing of locally ambiguous sentences. In G. Piérault-Le Bonniec e M. Dolitsky (eds.). *Language bases ... discourse bases*. Amsterdam: John Benjamins Publishing Co.
- Stevenson, R., Knott, A., Oberlander, J. & McDonald, S. (2000). Interpreting pronouns and connectives: interactions among focusing, thematic roles and coherence relations. *Language and Cognitive Processes*, 15 (3), 225-262.
- Stewart, A. J., Pickering, M. J. & Stanford, A. J. (2000). The time course of the influence of implicit causality information: focusing versus integration accounts. *Journal of Memory and Language*, 42, 423-443.
- Tanenhaus, M. K., Boland, J., Mauener, G. e Carlson, G. N. (1993). More combinatory lexical information: thematic structure in parsing and interpretation. In G. T. M. Altmann e R. Shillcock (Eds.). *Cognitive models of speech processing*. Hillsdale: Lawrence Erlbaum Associates, Publishers, 297-319.
- Tanenhaus, M. K., Spivey-Knowlton, M. J. e Hanna, J. E. (2000). Modelling thematic and discourse context effects with a multiple constraints approach: implications for the architecture of the language comprehension system. In M. W. Crocker, M. Pickering e C. Clifton (Eds.). *Architectures and mechanisms for language processing*. Cambridge: Cambridge University Press.