

PROCESSAMENTO DA CORREFERÊNCIA E SUJEITOS ANAFÓRICOS – DADOS SOBRE O PORTUGUÊS EUROPEU E BRASILEIRO¹

Armanda Costa (Universidade de Lisboa)

Gabriela Matos (Universidade de Lisboa)

RESUMO

O presente estudo visa determinar as estratégias preferenciais, na escrita, de produção de sujeitos correferenciais em português europeu (PE) e brasileiro (PB), variedades de sujeito nulo consistente e parcial, respetivamente. Analisamos narrativas de 40 falantes de PE e PB, dos 4º e 6º níveis de escolaridade. Os resultados mostraram que o uso anafórico de sujeitos foneticamente realizados, nomeadamente expressões nominais repetidas, é uma estratégia comum a todos os grupos. Contudo, esta estratégia diminui do 4º para o 6º em PE, mas persiste no 6º do PB, o que interpretamos como consequência das propriedades gramaticais do PB. A produção de sujeitos nulos pronominais correferentes distingue os dois grupos: apenas os falantes do PE os usando recorrentemente em frases raiz e subordinadas finitas.

PALAVRAS-CHAVE: processamento de cadeias correferenciais; sujeitos anafóricos; línguas de sujeito nulo consistente e de sujeito nulo parcial

1. INTRODUÇÃO

Quando nos centramos na produção textual, algumas questões se colocam: O que é que determina a escolha, por parte de um locutor/escritor, de um elemento anafórico em posição de sujeito frásico que contenha propriedades suficientemente relevantes para recuperar o antecedente adequado? Na representação mental do locutor, que proeminência têm as entidades referidas no discurso e de que forma isso se reflete na forma das expressões nominais usadas para as retomar? Que propriedades da gramática da língua em uso determinam a retoma anafórica das formas sujeito? Tendo particularmente em consideração as diferenças entre o português europeu (PE) e o português brasileiro (PB), relativamente ao parâmetro do sujeito nulo, que diferenças poderão ser encontradas na produção de sujeitos frásicos lexicalizados ou foneticamente nulos, pronominais ou resultantes de movimento?

1. Este trabalho só foi possível com a colaboração de Eliane Nascimento, da Fundação Educacional de Paraíso do Tocantins – FEPAR, que fez a recolha dos dados dos sujeitos brasileiros em escolas municipais de Tocantins.

As questões enunciadas deverão ser respondidas, por um lado, considerando os sistemas cognitivos envolvidos no conhecimento da língua e no seu processamento, mais especificamente na sua produção, por outro, tendo em conta as propriedades gramaticais das variedades linguísticas em análise.

O presente artigo aborda as questões acima enunciadas do ponto de vista do desenvolvimento da linguagem. Assim, narrativas escritas por crianças entre os 10 e os 12 anos em ambiente escolar, falantes nativas do português europeu e do português brasileiro, são comparadas com vista a identificar as estratégias preferenciais de produção de sujeitos frásicos anafóricos (no sentido lato do termo) em cada uma das variedades em estudo.

Sendo o objetivo principal identificar as estratégias preferenciais no emprego de sujeitos anafóricos por falantes do PE e do PB, no estudo que apresentamos delimitamos a análise de dados aos contextos onde há condições para que o escritor possa usar uma de duas estratégias: numa primeira possibilidade, o escritor poderá usar um sujeito realizado lexicalmente, seja pela repetição literal, seja pelo emprego de uma expressão nominal semanticamente relacionada, ou pelo uso de qualquer tipo de pronome pleno (pessoal ou demonstrativo); numa segunda possibilidade, o escritor poderá optar por um sujeito omitido: por um pronominal (*pro*) em orações finitas, ou por um DP-gap em frases coordenadas, consequência de movimento de constituintes nominais por extração simultânea ('across the board'). Na análise de dados, também serão considerados contextos de frases adverbiais finais não finitas, em que, ao ocorrer o infinitivo flexionado, se admite a alternância entre pronome pleno e *pro*.

São excluídos da análise de sujeitos omitidos todos os casos em que não há opção de alternância entre sujeito lexicalizado ou omitido, tais como frases com sujeitos omitidos derivados de movimento de DP para posições argumentais ('*A-gaps*'), como em frases passivas, com verbos inacusativos ou em construções de elevação, ou derivados de movimento para uma posição não argumental, na periferia esquerda da frase (*A'-gaps*), como em frases relativas ou interrogativas.

O artigo está estruturado do seguinte modo: a secção 2 enquadra teoricamente o trabalho, considerando os aspetos cognitivos implicados na produção de narrativas escritas (secção 2.1.) e os aspetos relativos ao conhecimento linguístico, mobilizado na produção de sujeitos anafóricos (2.2); na secção (2.3), faz-se a caracterização das propriedades dos sujeitos em português europeu e em português brasileiro, tendo em vista a tipologia dos sujeitos realizados e nulos proposta em trabalhos recentes para estas variedades do português. Na secção 3, apresenta-se o trabalho experimental realizado, discutem-se aspetos metodológicos e analisam-se dados de escrita de narrativas, com foco na análise das cadeias anafóricas produzidas. Na secção 4, discutem-se os resultados.

2. PRODUÇÃO DE SUJEITOS ANAFÓRICOS NA PRODUÇÃO DE NARRATIVAS ESCRITAS EM PORTUGUÊS EUROPEU E BRASILEIRO - ENQUADRAMENTO TEÓRICO

Na presente secção procederemos à explicitação das propriedades centrais que, a nível do processamento e da gramática, determinam o funcionamento do objeto de estudo.

2.1. Aspectos cognitivos envolvidos na produção de relações referenciais em narrativas escritas

Assumindo o modelo de Levelt para a produção da linguagem (Levelt, 1989), a proposta arquitetura e funcionamento do sistema cognitivo para a produção de linguagem permitem pressupor que o locutor/escritor se envolve numa série de operações cognitivas e linguísticas, desde que há uma intenção comunicativa até que há produção de fala ou texto. Numa primeira fase, através de um conjunto de operações de conceptualização, há geração de uma mensagem pré-linguística, a que se segue a formulação de enunciados com apoio no léxico mental e na codificação gramatical (morfo-sintática e fonológica). Expressas as ideias em frases, há então lugar a processos de articulação fonética ou manuo-visual, no caso da escrita.

Na fase de conceptualização, distinguem-se operações de macro e de micro planeamento. As primeiras são realizadas tendo em vista a intenção comunicativa do locutor e o público a quem se dirige, assim como modelos de discurso. Para o estudo que agora se apresenta, são mais importantes as operações de micro planeamento, em que o locutor escolhe para cada ato de fala o formato que melhor responde à representação mental da informação selecionada e à perspectiva que quer passar ao seu interlocutor (*planning an informational perspective for an utterance will be called microplanning*, Levelt, 1989:5).

No caso da produção de narrativas, será importante a representação mental dos eventos, com base na sua sequencialização temporal e causal, bem como a introdução das entidades participantes, considerando quer a primeira referência, quer as retomas ao longo do discurso. Na construção do modelo mental da história, as entidades referidas ganham proeminência através da atribuição de funções discursivas às expressões linguísticas que as denotam, o que determinará a perspectiva informacional que o locutor quer fazer passar. O locutor, apoiando-se no léxico disponível, vai proceder a operações gramaticais que assinalem o *status* da informação que veicula. Distinções entre informação dada e nova vão repercutir-se no uso de sintagmas determinantes referenciais definidos ou indefinidos e na forma linguística das retomas anafóricas selecionadas.

Na produção de um discurso coerente, é obrigatório o recurso a mecanismos linguísticos que assegurem a coesão. No caso da produção de narrativas, a coesão referencial é assegurada pela criação de cadeias de correferência formadas por expressões associadas anaforicamente a antecedentes. Cognitivamente e linguisticamente, coloca-se a questão de saber quais as expressões mais adequadas para inequivocamente designar um dado antecedente. Numa outra perspectiva, será interessante saber até que ponto as propriedades discursivas e linguísticas, que cumulam uma dada entidade que ocupa um lugar informacional destacado ou posição estrutural elevado na frase, fazem prever a natureza da expressão que a vai retomar anaforicamente (veja-se a *Centering Theory*, proposta por Grosz, Joshi & Weinstein, 1995, em que é proposto o conceito de *forward looking center* para uma entidade que contém um potencial informativo e estrutural para ser retomadas posteriormente no discurso). Estão em questão aspectos de representação cognitiva da entidade em referência e a forma como ela é referida linguisticamente, assim como as propriedades linguísticas e de informatividade das expressões que se lhe referem anaforicamente no discurso.

O conceito de proeminência do antecedente e da sua acessibilidade aos processos de computação da frase e à sua interpretação é central para a *Teoria da Acessibilidade* proposta por Ariel (1996, 2001). O conceito de acessibilidade é cognitivamente motivado: tem a ver com foco atencional, com disponibilidade em memória de trabalho, com a natureza da informação codificada (dada/nova,

focalizada ou não). A Teoria da Acessibilidade estabelece uma correlação inversa entre a acessibilidade do antecedente no modelo mental em construção e o potencial referencial da expressão anafórica que o retoma: quanto maior for a acessibilidade do antecedente, menor poderá ser o valor referencial da anáfora. (*The more informative (the amount of lexical information), rigid (the ability to pick a unique referent) and unattenuated (phonological size) an anaphoric expression is, the lower the degree of accessibility it codes, and vice-versa...*”, 2001:32). Expressões anafóricas altamente informativas, como DPs modificados por orações relativas, correferem DPs com baixo nível de acessibilidade (como em (1), onde há grande distância entre antecedente e expressão anafórica), enquanto que pronominais nulos, em que não há sequer traços morfológicos e nenhum peso fonológico, assinalam antecedentes com acessibilidade alta (como em (2), onde a expressão anafórica nula retoma um antecedente na frase prévia).

(1) *O gato* observou o ninho. [...] *O gato, que tinha estado a observar o ninho*, decidiu subir à árvore sorrateiramente.

(2) *O gato* observou o ninho. [*pro*] Subiu à árvore sorrateiramente.

Digamos que, na compreensão, a forma da expressão anafórica funciona como chave para o ouvinte/leitor encontrar o antecedente, enquanto que na produção, a expressão anafórica escolhida pelo falante deverá codificar o estatuto cognitivo e linguístico do antecedente, servindo princípios de economia recorrentes na produção da linguagem, mas também a função de dar pistas ao interlocutor que facilitem a interpretação das redes coesivas.

No modelo mental que vai sendo construído para a história, os participantes nos eventos relatados representam tópicos discursivos sobre os quais se vai acrescentando nova informação. Nos enunciados formados por frases complexas ou justapostas, em que os predicados são acerca do tópico mantido (propriedade captada pelo conceito de *topic continuity*, Givón, 1979, ou o conceito de *topic aboutness*, Cowles 2007), com grande probabilidade são encontrados sujeitos frásicos anafóricos, correferentes com o constituinte em tópico, seu antecedente. Esses sujeitos gramaticais, segundo a Teoria da Acessibilidade, deverão ser económicos, por isso, pronominais nulos serão então os mais adequados.

Ainda durante a fase de formulação, terá de ser equacionado o conhecimento linguístico acessível, a gramática que suporta a formatação das ideias em frases. Neste trabalho, colocamos a hipótese de que, perante uma mesma tarefa comunicativa – a produção escrita de uma história fortemente regulada por um estímulo visual - falantes de uma mesma língua, mas de variedades que se distinguem por propriedades morfológicas e sintáticas com impacto na formação de cadeias anafóricas, possam usar estratégias linguísticas distintas no estabelecimento da coesão referencial.

Acrescente-se ainda que, na situação particular de produção na modalidade escrita, o sistema cognitivo de monitorização dos processos em curso deve assegurar que as expressões nominais selecionadas sejam as mais adequadas para recuperar o antecedente apropriado, contribuindo para que a codificação gramatical assegure a transmissão da intenção comunicativa. Entendendo que o texto escrito usa mecanismos coesivos mais económicos do que a fala, evitando a repetição literal ou mesmo a paráfrase, teríamos condições ótimas para avaliar quais são as estratégias preferenciais em cada variedade do português dos grupos de participantes sob observação.

2.2. Sujeitos anafóricos e domínios de ocorrência

O conhecimento linguístico do falante restringe os tipos de expressões nominais que podem ocorrer em posição de sujeito, tendo em consideração a sua natureza e os domínios gramaticais em que ocorrem. No presente trabalho, consideraremos apenas aquelas que ocorrem em frases finitas e em frases infinitivas adverbiais em que alternam sujeitos realizados e omitidos.

No português, podem ocorrer na posição de sujeito frásico os seguintes tipos de expressões nominais:² sintagmas determinantes/nominais referenciais; pronomes realizados ou nulos localmente livres ou associados a expressões em posição-A' na periferia esquerda da frase; cópias-A (i.e., cópias de constituintes localmente movidos para a posição argumental de sujeito frásico); cópias-A' (i.e. cópias de constituintes localmente movidos para posições não-argumentais da periferia esquerda da frase).

Estes tipos de expressões nominais sujeito não participam aleatoriamente em relações anafóricas. No domínio das frases simples, apenas relações anafóricas locais são possíveis, i.e., relações em que o antecedente e o constituinte na posição argumental ocorrem nesse mesmo domínio frásico. É o caso das *cópias-A* dos sujeitos que se moveram do interior do sintagma verbal para a posição argumental de sujeito frásico, o especificador de Flex(ão) / T(empo), i.e os DP-gaps (o exemplo (3) ilustra esse movimento para uma frase com um verbo inergativo)³. É igualmente o caso das *cópias-A'*, decorrentes de movimento de constituintes para posições não-argumentais, na periferia esquerda da frase, como ilustrado em (4), ou de cópias resultantes do movimento de constituintes para posição de tópico marcado, como ilustrado em (5a) (topP-gaps'). É ainda o caso de pronomes associados a tópicos basicamente inseridos, como exemplificado em (5b):

- (3) *Os alunos_i [-]i* trabalharam afincadamente.
(4) a. *Quem_i é que [-]i* escreveu esse livro?
(5) a. *O Pedro_p, às crianças, [-]i* só oferece livros.
b. *As crianças, elas* veem sempre com agrado filmes de aventuras.

Em frases complexas subordinadas finitas, em que o sujeito é um pronome realizado ou nulo (*pro*), o sujeito pode estabelecer uma relação não local de correferência com uma expressão referencial que ocorra na frase subordinante, como em (6), ou ser associado a um constituinte que ocupe uma posição proeminente de tópico marcado de uma frase anterior, como ilustrado em (7), com exemplos representativos do português europeu.

- (6) a. *O autor_i* decidiu que {*ele/pro_i*} não publicaria esse artigo já.
b. O Pedro perguntou *aos amigos_i* se {*eles/pro_i*} queriam chá.
c. Quando {*ela/pro_i*} viu o João, a Maria cumprimentou-o.
(7) *A Maria_i,* o Pedro disse que {*ela/pro_i*} estava doente.

Em frases subordinadas infinitivas, diferentes tipos de sujeito podem ocorrer, dependendo da relação de dependência que a subordinada estabelece com a subordinante (argumento ou adjunto) e da natureza

2. As expressões nominais são designadas de acordo com a tipologia corrente da Teoria de Princípios e Parâmetros (Chomsky 1981, 1995, 2008).

3. Note-se que estamos a admitir a posição clássica de que no português em qualquer das variedades em análise, os sujeitos ocupam a posição de especificador de TP. Para argumentos a favor desta posição para o português europeu veja-se Costa & Duarte (2002); para a existência de pronomes fracos sujeito no português brasileiro, veja-se , por exemplo, Kato (1999). Para uma posição divergente relativamente ao português europeu veja-se Barbosa (1996, 2000); Kato (1999).

mais ou menos deficitária da oração subordinada. No presente trabalho, interessa-nos particularmente um subtipo de orações subordinadas adverbiais que em português europeu e brasileiro admite a presença de sujeitos realizados ou omitidos. No primeiro caso, ocorre tipicamente o infinitivo flexionado (cf. (8)) No segundo caso, o sujeito omitido é usualmente caracterizado como *pro* se ocorrer o infinitivo flexionado (cf. (9a)) e como PRO, quando o infinitivo é invariável (cf. (9b)):⁴

(8) O António chegou mais cedo para *as crianças* irem ao cinema.

- (9) a. Eles chegaram mais cedo para *pro* irem ao cinema.
b. Eles chegaram mais cedo para *PRO* ir ao cinema.

É de realçar que é impossível decidir qual a natureza do sujeito, *pro* ou *PRO*, quando as formas verbais do infinitivo flexionado e do invariável coincidem, como acontece na 3ª pessoa do singular:

(10) O António chegou mais cedo para *PRO/pro* ir ao cinema.

Em frases complexas formadas por duas orações coordenadas, a posição de sujeito do segundo termo pode ser ocupada por pronomes realizados correferentes com um constituinte do primeiro termo coordenado, por exemplo o objeto direto, como em (11).

(11) O João encontrou *o António*_i e *ele*_i abraçou-o cordialmente.

Em português europeu, a correferência do pronome sujeito realizado com o sujeito do primeiro termo coordenado é frequentemente sentida como marginal em contextos como o exemplificado em (12), em que o segundo termo coordenado não é separado por pausa do primeiro e em que o pronome realizado não é focalizado. Nestes casos a omissão do sujeito da oração coordenada é aparentemente requerida (Matos 1991, Costa et al. 1998):

- (12) a. ??*O João*_i encontrou o António e *ele*_i abraçou-o cordialmente.
b. *O João*_i encontrou o António e [-]_i abraçou-o cordialmente.

Essa marginalidade desvanece-se quando o pronome sujeito realizado surge em frases coordenadas associadas por justaposição (cf. (13a)), ou quando no segundo termo coordenado se interpõe material linguístico entre a conjunção e o pronome sujeito (cf. (13b)):

- (13) a. (?) *O João*_i encontrou o António. E *ele*_i abraçou-o efusivamente.
b. *O João*_i encontrou o António e, ao vê-lo sorrir, *ele*_i abraçou-o efusivamente.

Quanto à omissão do sujeito do segundo termo coordenado, em alguns casos é evidente que esse sujeito deve ser caracterizado como *pro*. É o que acontece quando a primeira oração exhibe um sujeito pós-verbal, como em (14), ou quando as frases coordenadas se encontram separadas por uma fronteira discursiva (cf. (15)):

(14) Chegaram as crianças e *pro* puseram a casa em alvoroço.

(15) a. As crianças chegaram a casa. E (pouco depois) *pro* brincavam na sala.

4. Chomsky e Lasnik (1993) caracterizam PRO como um pronominal nulo específico que ocorre em contextos de caso nulo. Porém, desde Hornstein (1999), vários autores admitem que PRO é uma categoria que não tem fundamento empírico e que, nos casos de Controlo Obrigatório, deve ser reanalisada como cópia de Movimento-A.

No entanto, em frases coordenadas em que as duas orações ocorrem em estrita contiguidade, com estreitas relações de coesão e partilha do mesmo sujeito pré-verbal, a estratégia de omissão do sujeito usada é a decorrente de movimento de *extração simultânea* do sujeito de ambos os termos coordenados (cf. (16)); neste caso o constituinte omitido resultante apresenta as propriedades de uma cópia-A (um DP-gap):

(16) *O João*_i [-]_i encontrou o António e [-]_i abraçou-o.

Note-se que, apesar de o português europeu ser uma língua de sujeito nulo consistente (cf. secção 2.2.), a estratégia preferencial quando os sujeitos das duas frases coordenadas são correferentes é a extração simultânea e não a inserção de *pro*. Admitimos que essa preferência está correlacionada com os requisitos de economia que guiam a gramática, uma vez que é usada uma estratégia geral, interlinguística, de supressão de um constituinte, sob condições estruturais de recuperabilidade local. Com efeito, o facto de, através das línguas, os sujeitos omitidos das frases coordenadas corresponderem sempre que possível a cópias de extração simultânea, e não a sujeitos pronominais, faz com que não constituam um fenómeno específico de línguas de sujeito nulo, e ocorram em línguas, como o inglês (cf. (17)), que excluem sujeitos omitidos em domínios de subordinação finita em que *pro* deveria ocorrer (cf. (18)), como mostram os seguintes exemplos retirados de Quirk et al. (1985:923, 924):

(17) a. I [may see you tomorrow or may phone you later in the day]
b. The Polish athletes [have succeeded today but may not repeat their success tomorrow].

(18) *He did not spend very much, so that [could afford a trip abroad].

Em fragmentos discursivos distintos que apresentam frases independentes que são percebidas como uma só unidade de sentido, designadas frases *justapostas* (cf. Quirk et al. 1985, López García 1999), por vezes ocorrem no segundo fragmento discursivo sujeitos correferentes com um elemento do primeiro fragmento. Esses sujeitos podem ser pronomes, nulos ou realizados, mas também sintagmas determinantes/nominais referenciais repetidos (cf. (19)). Porém, no português europeu, no caso de ocorrerem pronomes nulos, o antecedente ocupa preferencialmente a posição de sujeito frásico, como mostra o contraste entre (20) e (21)⁵

(19) *Os estudantes*_i saíram da aula. Pouco depois {*pro/eles/os estudantes*}_i encontravam-se na biblioteca.

(20) a. O Pedro_i encontrou os amigos na biblioteca. Pouco depois, *pro*_i / *ele*_i convidou-os para jantar.

b. O Pedro, os amigos_i encontraram-no na biblioteca. Pouco depois, *pro*_i / *eles*_i convidaram-no para jantar.

(21) a. Os amigos encontraram o Pedro_i na biblioteca. Pouco depois *pro*_{?i} convidou-os para jantar.

b. O Pedro_i, os amigos encontraram-no na biblioteca. Pouco depois *pro*_{?i} convidou-os para jantar.

5. Refira-se o trabalho de Morgado (2011), que testou estas estruturas em PE numa tarefa de processamento off-line (leitura sequencial de duas frases em modo não cumulativo e resposta de interpretação da referência do sujeito pronominal da segunda frase (prônimo lexicalizado ou nulo). Os resultados mostram que a ocorrência de pronome lexicalizado é preferencial.

Em suma, as restrições gramaticais que regulam a distribuição das expressões nominais na frase presidem à sua boa formação e garantem que a codificação das ideias se faça de forma inteligível para os ouvintes/leitores. Também as condições discursivas deverão ser respeitadas para permitir a coesão entre fragmentos discursivos distintos e tornar o processamento possível e com custos cognitivos reduzidos.

2.3. Sujeitos nulos em português europeu e brasileiro

O português europeu e o português brasileiro exibem sujeitos nulos. Porém, o seu comportamento é diverso. O português europeu, porque admite generalizadamente sujeitos nulos pronominais, é considerado uma língua de sujeito nulo consistente. O português brasileiro, que admite restritamente a omissão dos sujeitos, é considerado uma língua de sujeito nulo parcial.

Desde o trabalho de Taraldsen (1978), tem sido prática corrente relacionar a possibilidade de omissão do sujeito pronominal nas línguas de sujeito nulo consistente com a riqueza da flexão verbal em pessoa e número⁶. A perda das marcas de flexão distintivas das diferentes pessoas gramaticais implicaria um progressivo desaparecimento do uso de pronomes nulos. Um caso que ilustra esta situação é o português brasileiro, como mostra a tabela 1, para o presente do Indicativo do verbo regular da 1ª conjugação, *cantar*, baseada na proposta de Duarte e Cavalcanti (2008) e adaptada no respeitante ao português europeu:

Pessoa	Português europeu		Português brasileiro	
	Número		Paradigma 1	Paradigma 2
1	S	<i>canto</i>	<i>canto</i>	<i>canto</i>
2	S (tu)	<i>cantas</i>	<i>cantas</i>	-----
2	S (você)	<i>canta</i>	<i>canta</i>	<i>canta</i>
3	S	<i>canta</i>	<i>canta</i>	<i>canta</i>
1	P (nós)	<i>cantamos</i>	<i>cantamos</i>	-----
1	P (a gente)	<i>canta</i>	<i>canta</i>	<i>canta</i>
2	P	<i>cantais</i>	<i>cantais</i>	-----
2	P (vocês)	<i>cantam</i>	<i>cantam</i>	<i>cantam</i>
3	P	<i>cantam</i>	<i>cantam</i>	<i>cantam</i>

Tabela 1

A presente tabela revela o enfraquecimento das marcas de pessoa e número da flexão verbal em português brasileiro dos falantes que usam (preferencial ou exclusivamente) o segundo paradigma, relativamente ao português europeu. No português europeu registam-se seis formas flexionais distintas para as pessoas gramaticais (*canto*, *cantas*, *canta*, *cantamos*, *cantais*, *cantam*). Dessas formas, a referente à 1ª pessoa do plural, *cantamos* é correntemente usada, embora alterne, em contextos informais, com a forma (*a gente*) *canta*; a forma *cantais*, da 2ª pessoa do plural, que é substituída por (*vocês*) *cantam* em muitas regiões de Portugal, subsiste em várias regiões do país. Em contraste, no português brasileiro, apenas quatro formas gramaticais ocorrem no uso quotidiano (*canto*, *canta*,

6. Note-se, porém, que há línguas, como o Chinês ou o Japonês, que apresentam sujeitos omitidos, apesar de não exibirem marcas de flexão verbal. Veja-se, por exemplo, Holmberg 2005, 2010, para uma distinção entre os três tipos de línguas.

cantamos, cantam), das quais a da primeira pessoa do plural (*cantamos*) tende a ser substituída na população mais jovem e na oralidade por (a gente) *canta* (cf. Duarte 1993, Kato 1999, Modesto 2011).

Como tem sido salientado por vários autores (Galves 1987, 2001, Duarte 1993, Kato 1999, Barbosa et al. 2005, entre outros), este enfraquecimento da Flexão em pessoa e número está correlacionado com o aumento de sujeitos realizados: a perda do valor referencial de Flexão levaria a uma menor ocorrência de sujeitos nulos no português brasileiro, nomeadamente em contextos em que são requeridos ou preferenciais em línguas de sujeito nulo consistente, como os casos de pronome ligado (cf. (22a)) ou de correferência com um antecedente na oração matriz (cf. (22b)), exemplos:

(22) [Mulher nenhuma]_i ela_i pode querer dominar o homem.⁷

(23) [A casa]_i virou um filme quando ela_i teve de ir abaixo. (Barbosa et al. 2005)

Vários autores salientam ainda que o enfraquecimento de Flexão no PB está associado à ocorrência de pronomes fracos em posição de sujeito frásico, que redobram um constituinte nominal realizado na periferia esquerda da frase (cf. Galves 1987, 2001, Figueiredo Silva 1996, Kato 1999):

(24) a. Essa competência ela é de natureza mental (Pontes 1981, *apud* Galves 1987, 2001)
b. Você, (vo)cê não me pega! (Barbosa *et al.* 2005)

Desde Rizzi (1982) tem sido proposto que a flexão verbal em línguas de sujeito nulo consistente, como o português europeu, apresenta um estatuto pronominal. Segundo alguns autores (Kato 1999, Manzini e Roussou 2000, entre outros), os traços de concordância nessas línguas são afixos pronominais inicialmente inseridos na posição argumental de sujeito de vP que se movem para o núcleo funcional da frase, Flex/T onde se associam ao verbo. Porém, em alternativa a esta posição radical, prevalece, em trabalhos recentes, a ideia clássica de que a Flexão/T(empo) pronominal nas línguas de sujeito nulo consistente está associada/o a uma categoria pronominal basicamente gerada como sujeito de vP, *pro* (cf. Barbosa 1996, 2009, 2011, Holmberg 2005, 2010, Roberts 2010).²

No que diz respeito às línguas de sujeito nulo parcial, a que o PB pertenceria, a presença de sujeitos omitidos é bastante restringida. Holmberg (2010), baseando-se nos dados de Ferreira (2004) e de Rodrigues (2004), considera, no português brasileiro, os casos de sujeito omitido com referência genérica, como (25), que no português europeu determinariam a ocorrência de *se* impessoal (cf. (26)), e os de referência definida sob c-comando local de um antecedente da frase subordinante, como (27).

(25) a. É assim *pro*_{arb} que faz o doce.
b. João me contou que na praia *pro*_{arb} vende cachorro quente.

(26) a. É assim que *se*_{arb} faz o doce.
b. O João contou-me que *se*_{arb} vende cachorros quentes na praia

(27) *João*_i me contou que *pro*_i vende cachorro quente na praia.

7. Repare-se que em português europeu, o sujeito da frase subordinante, dado o seu estatuto de palavra negativa/quantificacional (cf. 22a), só poderia ser retomado por um pronominal nulo interpretado como uma variável:

(i) Nenhuma mulher sabia que *pro*_i estava doente.

8. Estes autores diferem, no entanto, quanto às propriedades de Flexão/T pronominal e à relação que estabelece com *pro*. Para Barbosa, Flexão/T em línguas de sujeito nulo consistente funciona como afixo pronominal com propriedades referenciais plenas associado a uma categoria vazia basicamente inserida em vP; para Holmberg e Roberts, alguns dos traços de concordância de Tempo/Flexão não são especificados e têm de ser fixados por *pro*.

Holmberg (2010) assume que em (25), a posição de sujeito é ocupada por *pro* sem valor definido (arb(itrário)) e admite que em (27), o sujeito omitido corresponde a um pronome nulo localmente c-comandado por um antecedente (Holmberg 2010:118); deixa, porém, em aberto se estes casos são instâncias de correferência de *pro* ou de controlo obrigatório⁹, como proposto em Rodrigues (2004)¹⁰. De facto, como notado em Holmberg (2010), o português europeu, uma língua de sujeito nulo consistente, apresenta casos idênticos a (27), como ilustrado em (28):

(28) *O João_i disse que pro_i comprou um computador.*

No presente trabalho, consideraremos que o sujeito omitido que ocorre em frases como (27) e (28) deve ser caracterizado como *pro*, tanto em português brasileiro como em português europeu. Em português europeu estes sujeitos nulos podem alternar com sujeitos pronominais enfáticos (focalizados), localmente correferentes com o sujeito da subordinada (cf. (29a)). Todavia, quando o sujeito realizado não é enfático, é predominantemente interpretado como tendo uma referência disjunta do sujeito (cf. (29b)):

(29) a. *O João_i disse que ELE_i comprou um computador.*
b. *O João_i disse que ele_j comprou um computador.*

A preferência pelo pronominal omitido para dar conta da correferência entre sujeitos (cf. (28)), tem sido explicada em línguas de sujeito nulo consistente em termos do princípio *Evitar Pronome* (cf. Chomsky 1981 e Brito 1991, para o português europeu), um princípio de economia na fronteira entre a gramática e o processamento.

3. GRAMÁTICA E PROCESSAMENTO

No processamento de cadeias anafóricas, o ouvinte/leitor tem, primeiramente, de compor o material lexical em unidades sintáticas estruturadas e hierarquizadas, criando relações de ordem e de dependência. A esta fase correspondem as operações de *parsing* sintático. Na ocorrência específica de uma expressão referencialmente dependente, tem de fazer-se a interpretação desse objeto sintático e associá-lo a um antecedente que o identifique. Nesse caso, duas condições alternativas se colocam: existe um pronome lexicalizado, possuindo traços morfológicos de número, género e caso, suficientemente informativo para ser associado a um antecedente, ou há uma expressão nominal nula sem qualquer informação intrínseca que sustente a formação da cadeia. Em ambas as condições, há contextualmente pistas lexicais, estruturais e discursivas que podem contribuir para que essa associação se faça com mais ou menos custos (v. a *Hipótese da Causalidade Implícita*, Caramazza *et al* 1977, Costa 2003/2005, para o PE; a *Hipótese Estrutural do Antecedente*, Carminatti 2002, Luegi 2012, para o PE; a *Hipótese do Paralelismo Estrutural*, Chambers e Smyth 1998; Smyth 1994, Corrêa 1998, para o PB; a *Teoria da Centralidade*, Grosz *et al* 1995, Morgado 2011, para o PE; a *Hipótese da Carga Informacional*, Almor 2000). Uma terceira situação, é aquela em que o processador lida com lugares vazios na estrutura (*gaps*, em sentido lato), criados por movimento de constituintes: as cópias A e A', descritas na secção anterior. Neste caso, por se tratar de uma cadeia envolvendo dois

9. Adotando a posição de Hornstein (1999), a autora assume que o sujeito omitido resulta de movimento, i.e., é uma cópia-A, legitimada pelo facto de o português brasileiro ter uma flexão finita pobre.

10. Modesto (2010, 2011) argumenta contra a existência de controlo em frases finitas no português brasileiro, e propõe que as ocorrências de sujeito omitido em português brasileiro sejam consideradas casos de omissão de tópico.

sujeitos correferentes, decorrente de uma operação de movimento, a cópia (*gap*) e o antecedente que a c-comanda estabelecem uma relação de ligação, em sentido estrito, essa cadeia é governada por princípios puramente estruturais e não discursivos. Tal facto, faz prever que esta cadeia anafórica seja resolvida numa fase prévia à interpretação, a ser assumido um modelo de processamento modular e sequencial, em que operações de *parsing* precedem a interpretação (para uma discussão sobre o decurso temporal da atribuição da correferência e as fontes de informação usadas, considerando tipos de expressões anafóricas, veja-se Nicol e Swinney 1989, entre outros).

Sobre o processamento de vários tipos de relações anafóricas temos alguns resultados para o Português. Em diferentes condições sintáticas e discursivas, com diferentes desenhos experimentais, os resultados apontam para que em português europeu, pronomes realizados e pronomes nulos sejam resolvidos em função de pistas estruturais e discursivas. No domínio da frase subordinada complexa - adverbial causal, temporal, concessiva - há uma clara preferência para pronominais nulos tomarem antecedentes na posição estrutural mais alta (Sujeito ou consituente deslocado à esquerda) e pronomes plenos assinalarem referência disjunta com o sujeito prévio, ou correferência com constituintes argumentais mais encaixados na frase precedente (Costa *et al* 1998; Costa 2003/2005; Luegi 2012). Estes resultados confirmam a Hipótese da Posição Estrutural do Antecedente (como proposto em Carminatti, 2002). No domínio da justaposição de frases, o pronome pleno manifesta-se como preferencial sobre o pronome nulo para designar correferência com o Sujeito da frase precedente (Morgado 2011, para o PE).

No presente estudo, analisaremos as estratégias de produção escrita de relações anafóricas envolvendo sujeitos frásicos, e contrastaremos os dados obtidos com os recolhidos nos estudos de compreensão.

3.1. Hipóteses

As expressões usadas anafóricamente na produção de texto narrativo para assegurar a coesão referencial são função (i) da gramática da língua que suporta as operações de micro planeamento das frases que compõem o discurso, (ii) de princípios organizadores do discurso no que reporta à estrutura informacional e à atribuição de saliência às entidades em referência, (iii) das operações de processamento da linguagem que procedem sobre o conhecimento linguístico dentro dos constrangimentos do sistema cognitivo geral.

O facto de PE e PB se distinguirem por propriedades gramaticais que afetam a ocorrência de sujeitos nulos e as propriedades sintáticas dos sujeitos omitidos em geral, conduz à hipótese de que,

- (i) para referir uma entidade previamente introduzida no texto, os sujeitos falantes da variedade brasileira do português se distinguirão dos falantes da variedade europeia na escolha dos elementos anafóricos usados: em PB espera-se que haja preferência por sujeitos lexicalizados sobre sujeitos omitidos.

Tendo em conta a caracterização gramatical dos sujeitos omitidos e os resultados disponíveis do processamento de cadeias anafóricas e correferenciais, que apontam para que pronominais sejam mais custosos do que cópias resultantes de movimento local no âmbito da mesma frase complexa, espera-se que,

(ii) independentemente dos fatores escolaridade e variedade linguística do Português, haja uma elevada frequência de sujeitos omitidos em frases coordenadas quando há condições estruturais favoráveis a um processo de extração simultânea de sujeitos,

Considerando aspetos do discurso em articulação com teorias do processamento de cadeias anafóricas, como a *Teoria da Acessibilidade* (Ariel 1996) ou a *Teoria da Sobrecarga Cognitiva* (Almor 2000), prevê-se que,

(iii) independentemente da idade e da gramática da língua em uso, na retoma anafórica de entidades em posição de tópicos discursivo, sejam usados elementos anafóricos não redundantes, com evitamento de formas nominais repetidas, sempre que haja proximidade entre expressão anafórica e antecedente (aqui entendida como ocorrendo em frases contíguas), sempre que haja paralelismo estrutural (em que expressão anafórica e antecedente têm a mesma função sintática).

(iv) prevê-se, ainda, que haja um efeito de desenvolvimento (considerando o ano de escolaridade no uso de cadeias mais económicas, servindo princípios de parcimónia e de coesão).

4. ESTUDO EMPÍRICO

4.1. Questões preliminares à apresentação e análise dos dados.

Sendo o objetivo final fazer uma análise comparativa sistemática de estratégias de escrita relativamente à produção de cadeias anafóricas, duas questões preliminares se colocaram. Primeiro, tivemos de verificar se a condição experimental de escrita provocada levou à produção de textos comparáveis entre si. Não é possível tentar estabelecer comparações sobre estratégias preferenciais para a expressão da correferência por escrito, se os dois grupos de escolaridade e os dois grupos de falantes das duas variedades do português não as expressarem num texto com propriedades tipológicas e informacionais idênticas. Segundo, considerando que em ambos os grupos de escolaridade (4º e 6º anos do ensino básico) não estão ainda adquiridas todas as capacidades de escrita que presidem à competência textual narrativa, tivemos de tomar algumas decisões relativas à anotação das cadeias anafóricas, que explicitaremos.

4.2. Tarefa de escrita provocada e materiais produzidos

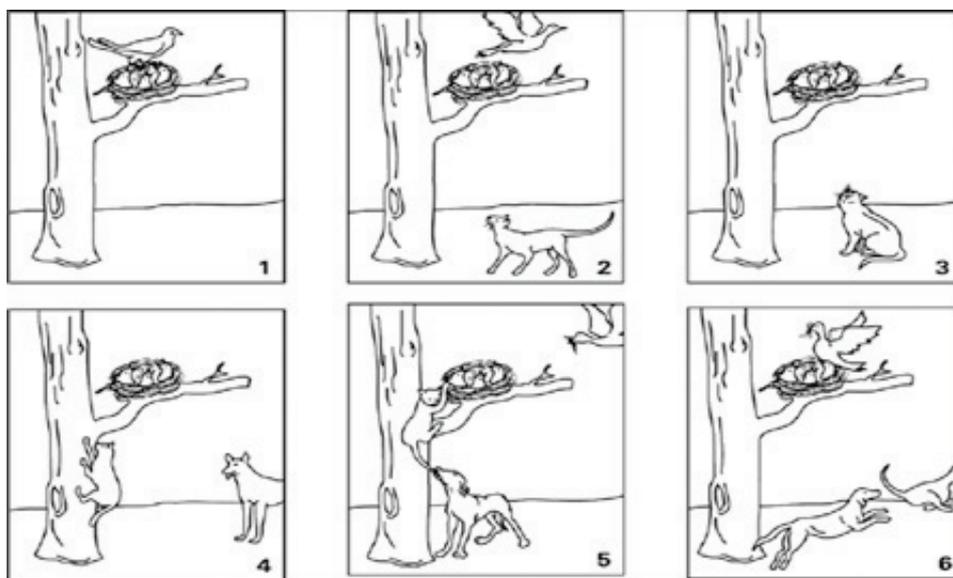
Os sujeitos e a tarefa de escrita

A amostra do nosso estudo é constituída por 40 sujeitos, distribuídos por dois grupos de escolaridade (4º e 6º ano de escolaridade) e dois grupos de língua (variedades europeia e brasileira do português, PE e PB), o que levou à formação de 4 grupos: 4º PE e 4º PB; 6º PE e 6º PB. Os dados foram recolhidos em escolas públicas portuguesas da área de Lisboa, em Portugal, e em escolas municipais de Palmas Tocantins, no Brasil¹¹. Embora seguramente haja diferenças pedagógicas nos sistemas

11. Os dados dos sujeitos brasileiros foram recolhidos por Eliane Gonçalves, no Brasil. Os dados dos sujeitos portugueses pertencem à base de dados do projeto *Diversidade Linguística na Escola Portuguesa*, ILTEC. 2008, <http://www.iltec.pt/divling/index.html>

de ensino em Portugal e no Brasil, valorizamos sobretudo o facto de nos dois casos haver o mesmo tempo de exposição à escrita e aos processos de aprendizagem.

As produções escritas foram induzidas a partir da visualização de uma história em banda desenhada, a história do gato (*the cat story*) usada inicialmente por Hickmann (1982; 1995), e corrente na investigação interlinguística para a avaliação de aspetos do desenvolvimento da linguagem e do uso particular que é feito nas várias línguas naturais para a expressão da referência, do espaço e do tempo.¹²



The cat story

Foi usado um protocolo que estabelecia que as crianças num primeiro momento observassem as imagens e que num segundo momento relatassem por escrito a história que as imagens ilustravam; foi-lhes pedido ainda que produzissem um texto redigido e não uma lista de frases que pudessem ser associadas a cada quadrinho. Na história ilustrada, participam três entidades, o pássaro, o gato, o cão, que partilham traços idênticos de género e número. Assim, em qualquer situação de correferência, por propósitos de clareza discursiva os sujeitos são condicionados a usar expressões anafóricas que designem um antecedente de forma não ambígua. Pela frequência com que participa nas cenas e por ser a entidade que desencadeia o conflito narrativo, o personagem *gato* institui-se como protagonista da história.

Ao pedir-se ao sujeitos a narração escrita da história, e por questões de economia discursiva e de coesão textual, configura-se uma situação de comunicação em que se espera que os sujeitos escolham as expressões anafóricas mais adequadas. A adequação será função da forma anafórica que, sem ser redundante, tem poder designativo forte relativamente ao respetivo Antecedente. Considerada a modelização do processo de escrita, os dados recolhidos são referentes à fase de *textualização*, em que o sujeito traduz as representações mentais ativadas, num formato tipificado de discurso – a narrativa – com apoio nos mecanismos morfossintáticos e semânticos que asseguram a coesão textual (Fayol 2003).

12. Em Português, com recurso ao mesmo estímulo e a condições de escrita provocada semelhantes, refiram-se os trabalhos de Batoréo 2000; Batoréo e Costa 1998; Arim 2005; o projeto do Instituto de Linguística Teórica e Computacional sobre *Diversidade Linguística na escola portuguesa* (Mateus et al 2008) e o mais recente, o de Nascimento 2009, que fez o primeiro estudo comparativo de escrita em PE e PB por sujeitos com perfil de escolaridade semelhante, falantes das duas variedades mas a residirem em Portugal

4.3. A extensão dos textos e a a estrutura das histórias

Numa primeira análise dos textos escritos, fomos verificar se todos tinham mantido a estrutura informacional esperada, considerando que *script* previsível, emergente das seis imagens/duas tiras, em seis passos, tendo o gato como protagonista:

- situação inicial: *um pássaro e um ninho numa árvore;*
- desenvolvimento: *o pássaro abandona o ninho e um gato chega; o gato observa o ninho; o gato sobe à árvore; (v) um cão ataca o gato (o gato é atacado por um cão);*
- resolução ou fechamento: *o pássaro chega e o cão persegue o gato.*

Calculou-se a extensão dos textos avaliando a média de palavras por grupo. Assinale-se que há uma enorme variação intragrupo 4º PB, como indica a medida de desvio padrão (dados apresentados na tabela1). Essa variação explica-se pelo uso de estratégias narrativas não usadas nos outros 3 grupos: inserção de diálogos ou excursos narrativos para tópicos laterais à história centrada no gato.

		Média de palavras por texto e grupo
PB	4º	95,8 (28,618)
PB	6º	85,9 (20,829)
PE	4º	73,7 (12,858)
PE	6º	83,4 (18,379)

Tabela 1: Número médio de palavras por texto; entre parênteses, o desvio padrão.

Independentemente da extensão do texto, registou-se a cadeia correferencial formada exclusivamente para *gato*, considerando a primeira referência que constitui o antecedente e todas as expressões anafóricas que a retomam, independentemente da sua forma ou função sintática. Neste estudo, apenas apresentamos dados para a cadeia referencial formada pelas expressões anafóricas em posição de sujeito frásico, por ser a mais produtiva.

A primeira referência a *gato* é canonicamente feita por sintagma nominal indefinido, sendo recorrente em todos os grupos o uso de verbos inacusativos de aparição (*chegar; aparecer; aproximar-se*). Interpretamos tais estratégias discursivas como manifestação de uma competência narrativa equilibrada entre grupos.

Os resultados obtidos na análise das cadeias correferenciais são apresentados no Gráfico 1. Feitas análises de variância, os fatores Grupo (PE e PB), Ano (4º e 6ª), não produziram quaisquer efeitos significativos na frequência de referências anafóricas-sujeito para *gato* ($p < 0.708$; $p < 0.501$, respetivamente). Quando comparados os quatro grupos entre si, o teste post-hoc Tukey HSD não revelou quaisquer diferenças significativas entre grupos. Interpretamos este resultado, por um lado, como a manifestação de uma competência narrativa no que respeita à referência do protagonista em posição de sujeito, idêntica entre grupos e anos de escolaridade, por outro, os resultados sustentam a hipótese que tínhamos colocado de que as características do estímulo criam restrições suficientemente fortes para que o número de referências ao protagonista da história seja semelhante. Consideramos, assim, que as cadeias usadas na produção escrita entre os vários grupos são comparáveis.

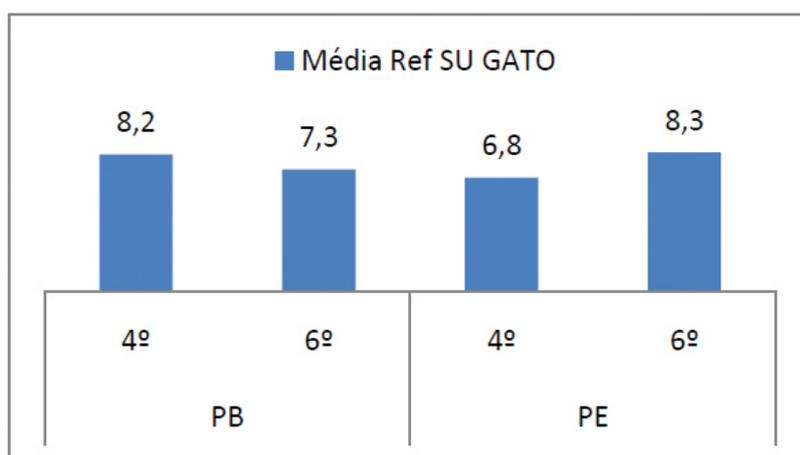


Gráfico 1: Média de referências a *gato*, em posição de sujeito frásico, por texto e grupo.

4.4. Anotação das cadeias correferenciais.

Nas tabelas de 2 a 4, dão-se exemplos de um texto por ano, com as respetivas anotações sintáticas. Como as crianças não dominam ainda completamente as convenções da escrita, no tocante à ortografia, corrigiram-se os erros ortográficos nos textos apresentados; quanto à marcação de fronteiras sintáticas para delimitação das unidades frásicas onde ocorrem as cadeias anafóricas, decidiu-se pela sua marcação sempre que há uso convencional de pontuação, ou sempre que há claramente transição de tópico (as fronteiras frásicas foram assinaladas com cardinal (#); dentro de cada frase, a vírgula assinala fronteiras oracionais). Recorde-se que apenas foram consideradas as ocorrências de sujeitos referentes a GATO, e, relativamente aos casos de sujeitos omitidos, para a análise restringimo-nos apenas àqueles em que poderia haver alternância entre sujeito omitido ou realizado. Assim, foram consideradas as categorias: *pro* em frases finitas; sujeitos omitidos associados a formas verbais homónimas para o infinitivo flexionado ou invariável em domínios de subordinação que admitem os dois casos, codificados como PRO/*pro*; sujeitos omitidos em frases coordenadas resultantes de extração simultânea do sujeito (ATB gaps), codificados no texto como [DP -]_{ATB}.

Na coluna da esquerda de cada tabela apresenta-se o texto exhaustivamente codificado, enquanto na coluna da direita apenas se apresentam as cadeias com os elementos que foram efetivamente analisados e acima já identificados. Não foram contabilizados os casos de controlo obrigatório, anotado nos textos como PRO_{CO}, nem os casos de sujeito com verbos (semi-)auxiliares (anotados no texto como [DP-]_R (Raising)).

1. Era uma vez um pássaro, que estava a tomar conta dos seus filhotes.	
2. Quando se foi embora, apareceu um gato .	2: # DP _[-def] pós verbal #
3. O gato queria PRO _{CO} apanhar os passarinhos.	3: # DP _[+def] pré-verbal #
4. <i>pro</i> Olhou para eles para PRO/ <i>pro</i> ver se <i>pro</i> conseguia [DP-] _R ter uma ideia para PRO/ <i>pro</i> os apanhar.	4: # <i>pro</i> , PRO/ <i>pro</i> , <i>pro</i> , PRO/ <i>pro</i> #
5. O gato lembrou-se de PRO _{CO} subir à árvore.	5: # DP _[+def] #
6. Quando <i>pro</i> já estava quase a [DP-] _R apanha-los, apareceu um cão.	6: # <i>pro</i> #
A mãe pássaro como estava quase a chegar, tirou o gato de cima da árvore, e foi atrás dele.	

Tabela 2: texto 4, 4º ano PE

<p>1. Era uma vez um pardal que estava prestes a ter filhos e então foi buscar a comida para os alimentar.</p> <p>2. Mas mal sabia ele, que, um gato andava a rondar por ali.</p> <p>3. E os ovos chocaram naquele preciso momento.</p> <p>4. O gato mirou-os e [DP-]_{ATB} decidiu PRO_{CO} ir [DP-]_R comê-los.</p> <p>5. pro Trepou a árvore.</p> <p>6. Entretanto apareceu um cão.</p> <p>7. O cão viu o que o gato estava a [DP-]_R fazer e, então, agarram no gato pela cauda e puxou-o com toda a sua força. O pardal viu o que o cão fizera e agradeceu-lhe.</p>	<p>2: # DP_[-def]</p> <p>4: # DP_[+def] DP-gap_{ATB} #</p> <p>5: # pro #</p> <p>7: # DP_[+def] #</p>
---	--

Tabela 3: texto 6, 6º ano PE

<p>1. Era uma vez uma árvore e nessa árvore tinha um ninho.</p> <p>2. Aí o gato chamado Nino foi [DP-]_R chegando na arvore e [DP-]_{ATB} viu o ninho de passarinhos e o gato Nino ficou sentado se PRO_{CO} preparando para subir na arvore.</p> <p>3. E quando ele esta subindo, o cachorro aparece e vê o gato subindo e corre, corre.</p> <p>4. Pega no rabo do gato e a mãe dos passarinhos voltou e o gato saiu PRO_{CO} correndo.</p>	<p>2 # DP_[+def], DP-gap_{ATB}, DP_[+def] #</p> <p>3: # PRON, DP_[+def] #</p> <p>4: # DP_[+def] #</p>
---	---

Tabela 4: texto 6, 4º PB

<p>1. Num dia a mamãe pássaro estava com os seus filhotes quando um gato muito malvado queria PRO_{CO} comer os seus filhotes.</p> <p>2. A mamãe pássaro voou mas seus filhotes não podiam voar</p> <p>3. Quando o gato estava [DP-]_R subindo o cão pegou ele pelo rabo e puxou.</p> <p>4. O gato caiu no chão e o cão correu a atrás do gato, em seguida a mamãe pássaro viu seus filhotes são e salvos em seguida ela falou: [Aquele gato levou uma lição graças ao meu amigo cão por que se não fosse ele aquele gato malvado ia [DP-]_R comer os meus filhotes.]</p>	<p>1: # DP_[-def] #</p> <p>2: # DP_[+def] #</p> <p>4: # DP_[+def] #</p>
--	---

Tabela 5: texto 8, 6º PB

(em nenhum texto foram contabilizados os sujeitos referentes a GATO nos excertos de discurso direto, dada a reduzida representatividade desses fragmentos nos textos recolhidos).

4.5. Análise dos dados

Codificadas as cadeias correferenciais para GATO, procedeu-se à organização da base de dados, quantificando-se para cada texto o número de ocorrências dos elementos anafóricos previamente estipulados: DP-rep (sintagma determinante repetido); PRON, Pronome Pleno; pro, Pronome nulo; PRO/pro (sujeito omitido em oração adverbial final não-finita); DP-gap ATB. No gráfico 2, apresentam-se as ocorrências médias de sujeitos correferentes lexicalmente realizados (DP repetidos e pronomes plenos) contra sujeitos omitidos (todas as categorias).

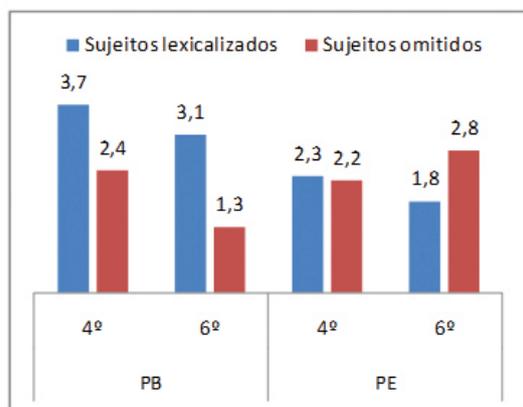


Gráfico 2

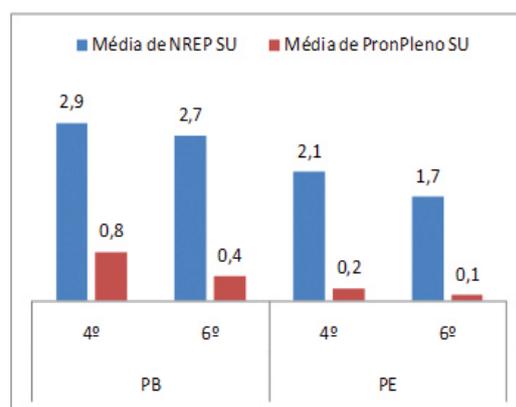


Gráfico 3

Os resultados mostram que em ambos os grupos de falantes do PB, há claramente um recurso mais frequente a sujeitos realizados do que a sujeitos omitidos. Nos grupos do PE, no 4º ano, o uso de sujeitos lexicalizados ou omitidos é equilibrado, mas no 6º ano é adotada a estratégia de sujeitos omitidos em desfavor de sujeitos lexicalizados.

A análise de variância revela que apenas a interação entre os fatores grupo/ano e o fator linguístico (SU omitidos/sujeitos plenos) produz diferenças significativas ($F(3,72) = 3.11$ $p < 0.03$). Feitas comparações individuais entre grupos com o teste post-hoc Tukey HSD, encontram-se diferenças significativas nas condições com sujeitos lexicalizados.

Quando considerada a frequência de sujeitos lexicalizados no 6º PE (1,8) contra o 6º PB (3,1) a diferença é significativa: $t(18) = 2.31$ $p < 0.032$; contra o 4º PB (3,7) também o é $t(18) = 2.73$ $p < 0.013$. Ainda quanto ao uso de sujeitos lexicalizados, o 4º PE apresenta resultados marginalmente significativos contra o 4º PB: $t(18) = 1.95$ $p < 0.06$.

Intragrupo, no 6º PB, encontram-se diferenças significativas entre o uso de SU lexicalizados (3,1) contra sujeitos omitidos (1,3): $t(18) = 2.03$ $p < 0.05$. A diferença no 4º PB entre tipos de sujeitos, embora aparente não é significativa: $t(18) = 1.77$ $p < 0.09$.

Em suma, o contraste sujeitos realizados contra sujeitos omitidos revela-se produtivo para distinguir grupos de língua, sobretudo se se tomar como referência o grupo do 6º ano PE comparativamente aos dois grupos do PB. Em Português Europeu, a estratégia de preenchimento lexical do sujeito frásico correferente é claramente abandonada por efeitos de desenvolvimento, a que estarão associados efeitos de aprendizagem da língua escrita. Em PB a preferência por sujeitos plenos diminui com a idade, mas continua a sobrepor-se ao uso de sujeitos não realizados.

Comparados os dois grupos de 6º ano, há uma assimetria evidente, que mostra que em PB se consolida o uso anafórico de sujeitos lexicais e se reduz o uso de omitidos, enquanto que em PE, os sujeitos omitidos são claramente preferenciais, embora as diferenças não sejam significativas ($t(18)=1.60$ $p < 0.12$).

Não há diferenças significativas entre grupos de ano da mesma variedade linguística.

Os valores apresentados no Gráfico 2 referem-se a uma análise interna que considera os dois tipos de sujeitos realizados: sintagmas determinantes repetidos e pronomes plenos. Como se pode verificar, o uso de pronomes plenos é residual em todos os grupos, especialmente em PE, em que a ocorrência de sujeitos pronominais nulos compete com a ocorrência de sujeitos realizados.

Nos gráficos 4 e 5, apresentamos as ocorrências dos sujeitos em duas condições. No gráfico 4, faz-se a comparação de ocorrências entre pronominais (pro e PRO/pro) contra DP-gaps ATB. No gráfico 5, discriminam-se os valores dos três tipos de sujeitos omitidos: sujeito nulo (pro), PRO/pro em frase subordinada final não finita e DP-gaps ATB.

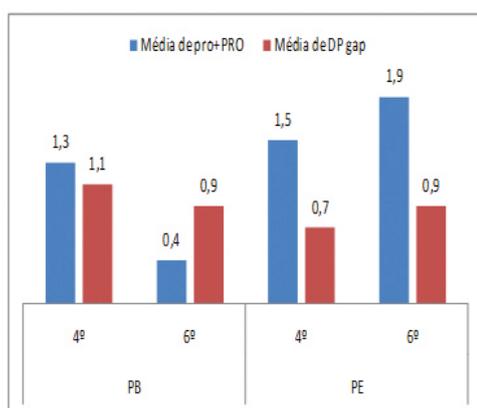


Gráfico 4

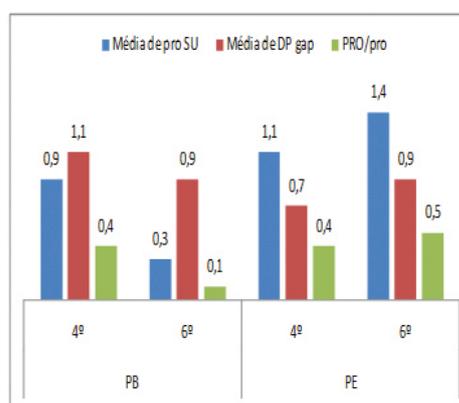


Gráfico 5

Considerando globalmente as categorias internas à categoria geral de sujeitos omitidos, sujeitos pronominais (pro ou PRO/pro) contra DP-gaps, não se encontram efeitos significativos considerando os fatores Grupo de língua ou nível de escolaridade.

Contudo, a análise com o teste post-hoc Tukey HSD evidencia algumas diferenças relevantes. Na categoria de pronominais (pro+ PRO/pro), quando comparado o 6º PE (1,9) com o 6º ano PB (0,4), a diferença é significativa: $t(18)=2.47$ $p < 0.02$; também 4º PE contra 6º PB mostra uma diferença perto da significância: $t(18)=1.94$ $p < 0.06$.

No referente à ocorrência de DP-gaps, não se encontram diferenças significativas entre quaisquer grupos (variedade linguística ou ano). Este fator não discrimina entre grupos.

No gráfico 5, desdobra-se a categoria dos pronominais (pro e PRO/pro) e apresentam-se os valores para os 3 tipos de sujeitos omitidos. Aí se encontram resultados significativos para o contraste entre sujeitos frásicos omitidos em frases coordenadas (DP-gaps) e adverbiais infinitivas (PRO/pro) entre os dois grupos de ano, intra grupo de língua: $t(18)=2.04$ $p < 0.05$, (4º vs. 6º PE), e menos relevante mas a assinalar, em PB (4º vs. 6º): $t(18)=1.98$ $p < 0.06$.

Note-se que as ocorrências de PRO/pro são exclusivas para o contexto de frase adverbial final infinitiva, que serve desígnios narrativos específicos (*o gato sobe para comer os passarinhos*), que se repetem de história para história.

Considerando apenas o fator grupo linguístico, verifica-se que em PE há diferença significativa no uso de [pro + PRO/pro] contra o uso de DP-gaps: $t(38)=2,20$ $p < 0,03$. Isso mostra uma clara preferência pelo uso de frases justapostas ou por frases subordinadas sobre frases coordenadas, e uma preferência por sujeitos pronominais nulos sobre sujeitos omitidos decorrentes de movimento (DP-gaps ATB).

Entre grupos linguísticos, se bem que não significativamente, PE e PB distinguem-se no uso de sujeitos pronominais nulos, com frequência mais baixa para o segundo: $t(38)=1,87$ $p < 0,06$.

Verificadas estatisticamente as diferenças entre grupos no tocante ao uso de sujeitos omitidos, e confirmada a hipótese de que falantes do PE usam mais sujeitos nulos do que os falantes do PE, fez-se uma análise qualitativa analisando os contextos de ocorrência de *pro*.

O comportamento dos falantes do PE e do PB é diferente consoante o nível de escolaridade em que se encontram, sendo particularmente distinto o uso dos falantes das duas variedades no nível mais avançado. Em frases raiz e em subordinadas, os alunos do PE do 6º nível usam mais sujeitos *pro* do que os alunos do 4º (5 contra 2); pelo contrário, neste mesmo contexto, os falantes de PB, que no 4º usavam, ainda que parcamente, sujeitos nulos (2), deixaram completamente de os usar no 6º.

No domínio das frases subordinadas finitas, os falantes do PE de ambos os níveis de escolaridade (4º e 6º) usam na mesma proporção e de forma expressiva (dado o número de ocorrências de subordinadas finitas) sujeitos nulos pronominais (*pro*) correferenciais (5 em ambos os grupos). O mesmo acontece, embora de forma um pouco mais restrita, com os alunos do PB (3/4). No âmbito da coordenação frásica, o número de sujeitos pronominais (*pro*) em PE aumenta substancialmente do 4º para o 6º nível de escolaridade, estando associado sobretudo à presença de sujeitos antecedentes ocupando uma posição pós-verbal, mas também à utilização como antecedente do sujeito do segundo termo coordenado de antecedentes repartidos (10, no 6º contra 2 no 4º). Por sua vez, no PB a utilização de *pro* em sujeitos de frases coordenadas só é encontrado garantidamente no 4º ano, num caso de coordenação entre fragmentos discursivos distintos.

6. CONCLUSÕES

As questões de investigação que desencadearam este estudo emergiram crucialmente de dois territórios: o do processamento e o da gramática teórica, tentando ver como um e outra se intercetam. Partimos da discussão das estratégias já identificadas na área do processamento da correferência, que apontam para que certos elementos anafóricos sejam preferenciais a outros, por imposição ou enviesamento de condições estruturais, discursivas ou até de informação léxico-semântica. Aceitando a universalidade de princípios de *parsing*, típicos dos modelos modulares e sequenciais, assumimos que cadeias formadas quer por elementos correferenciais com a mesma função sintática que o antecedente, quer por elementos que não oferecem redundância informativa, quer ainda por elementos cuja integração na estrutura e posterior interpretação depende mais de princípios gramaticais do que de informação discursiva, entram na formação de cadeias mais económicas. Assumimos que essas cadeias são mais facilmente interpretáveis na compreensão e que, por isso, serão mais usadas na

produção. Assim, criamos uma situação experimental tão controlada quanto possível e, a partir de uma tarefa de escrita provocada, recolhemos dados de escrita de falantes de PE e PB.

Na perspectiva da gramática das variedades do Português subjacentes às operações de microplaneamento das frases constitutivas dos textos, assumimos que PE e PB, ao distinguirem-se pelas propriedades relativas ao parâmetro do sujeito nulo (o PE sendo uma língua de sujeito nulo consistente e o PB uma língua de sujeito nulo parcial) determinam a produção de cadeias correferenciais distintas. Para a codificação dos dados, foram assumidas categorias sintáticas de sujeitos omitidos, de acordo com o quadro teórico da Teoria de Princípios e Parâmetros atual, no âmbito do Programa Minimalista.

Os textos foram produzidos por sujeitos que, para além da variedade do português que falam, se distinguem também por exposição a modelos pedagógicos de ensino e a cultura diferentes. Fez-se então uma prévia análise das narrativas para verificar a sua estrutura eventiva, a articulação cronológica e causal, a forma de introdução do protagonista (*o gato*) e a cadeia correferencial a partir daí construída. Os dados colhidos confirmaram a expectativa criada de que, apesar da variação existente entre sujeitos, os textos produzidos eram comparáveis entre si, mais ainda as cadeias de correferência sob observação. No entanto, é de notar que, sobretudo o grupo do 6º ano PB, manifesta comportamentos de escrita um pouco desviantes dos outros três grupos: apesar da instrução de redação de textos “ligados” (coesos), alguns dos sujeitos optam por listar frases em sequência. Esta constatação serve para a tomada de consciência de que este estudo tem limitações que deverão ser controladas pelo alargamento da amostra.

Em termos de resultados, o que se revela com maior força discriminatória entre grupos de língua é sem dúvida o contraste entre sujeitos lexicalmente realizados e sujeitos omitidos, para expressar relações anafóricas. Sobretudo se tivermos por referência os falantes mais velhos do PE (6º), verificamos que se afastam significativamente dos dois grupos do PB (4º e 6º): estes, para retomar sujeitos prévios, tópicos não marcados, preferem usar expressões nominais repetidas. Em PE, no 4º ano, essa é uma estratégia frequente, a par do uso de sujeitos omitidos, mas que é preterida com o desenvolvimento e a escolaridade.

Em termos das teorias de processamento da correferência, se aceitarmos que os princípios que regulam a compreensão são válidos para a compreensão, este é um resultado contraditório: em PB, prefere-se repetir em vez de omitir. Como interpretar tais resultados? As repetições sucessivas vão no sentido da violação do princípio da parcimónia, e deverão contribuir para a sobrecarga do sistema cognitivo, o que é inesperado. Se atentarmos no facto de que o 4º PE também se apoia na repetição de DPs, então poderemos pensar que essa é uma estratégia que, para além do mais, se poderá associar a uma competência de escrita em desenvolvimento, ainda a recorrer a formas fracamente coesivas. Porém, se considerarmos que essa estratégia distingue grupos de língua, então teremos de invocar efeitos de propriedades da gramática em uso. Esse aspeto é particularmente evidente nas produções dos falantes das faixas etárias mais altas, do 6º nível de escolaridade: no PE a produção de sujeitos correferenciais repetidos diminui substancialmente, enquanto no PB se mantém expressivamente a estratégia de repetição de nome, face a um uso muito restrito de pronomes realizados. Tais resultados contribuem para validar a hipótese de que, como afirmam vários autores, no PB há um enfraquecimento dos pronomes pessoais sujeito: esse enfraquecimento poderá estar na base da necessidade de o falante produzir, sempre que os domínios sintáticos o permitam, nomes repetidos cujo conteúdo garanta que as relações anafóricas pretendidas se estabelecem.

Quando entramos no domínio dos sujeitos omitidos dos diferentes tipos sintáticos considerados, uma primeira verificação é que em todos os grupos é homogênea a estratégia de uso de frases coordenadas com sujeito omitido no segundo termo coordenado. Sendo esta uma condição sintática comum a línguas de sujeito nulo e sem sujeito nulo, não se esperava que fosse distintiva entre PE e PB, como não é. Por outro lado, esta é uma estratégia que, na perspectiva do processamento e da gramática, serve princípios de economia: um DP-gap_{-ATB} é uma *cópia-A* deixada por movimento, que não cria requisitos de interpretação: é ligada localmente por um antecedente. Do ponto de vista do discurso aumenta o nível de coesão, enquanto que na perspectiva da gramática dá argumentos para que relações estruturais entre pronominais e antecedentes sejam mais custosas, porque menos locais e mais complexas, uma vez que ocorrem através de fronteiras frásicas e de fronteiras discursivas.

Esperava-se que a ocorrência de pronomes nulos fosse discriminativa entre grupos de língua. De facto, os falantes do PE (dos 4º e 6º anos) usam significativamente mais sujeitos nulos pronominais do que os do 6º ano do PB. Estamos a assumir por sujeitos nulos indistintamente os casos de sujeitos omitidos de frases finitas e de frases infinitivas adverbiais. Este é um resultado que vem confirmar a hipótese de que o PB é uma língua em que há um decréscimo de sujeitos omitidos por efeitos de mudança imputáveis ao enfraquecimento do sistema de flexão. Quando se faz uma análise qualitativa dos contextos de ocorrência de sujeitos nulos em PE e PB, mais esta hipótese se confirma. Em frases simples e frases raiz, o PB quase não usa sujeitos nulos (2 em PB, contra 7 em PE); em frases coordenadas finitas, em que no segundo termo há um sujeito pronominal nulo e não um DP-gap (casos em que não é possível a extração simultânea dos sujeitos), em PB há apenas uma ocorrência contra 10 em PE (das quais oito em contextos de sujeitos pós-verbais no primeiro termo coordenado, como antecedente do sujeito omitido do segundo termo coordenado).

Com as ressalvas devidas ao facto de que os dados de escrita em PE e em PB poderem não reflectir exactamente um mesmo nível de competência de escrita, consideradas as condições de ensino e de alfabetização em cada país, os resultados do nosso estudo mostram que há diferenças nas estratégias linguísticas para assinalar relações anafóricas em PE e PB escrito, imputáveis a diferenças das gramáticas das duas variedades. Em PE, língua consistente de sujeito nulo, prefere-se sujeitos correferenciais omitidos, enquanto que em PB, os sujeitos lexicalizados são preferenciais, ainda que isso acarrete redundância.

CO-REFERENCE PROCESSING AND ANAPHORIC SUBJECTS - DATA ABOUT EUROPEAN AND BRAZILIAN PORTUGUESE

ABSTRACT

The current study aims to identify the preferential strategies of coreferential subjects in written European and Brazilian Portuguese (EP and BP), a consistent and a partial null subject language, respectively. We analyzed written stories, produced by 40 EP and BP speakers attending the 4th and 6th grades. The results showed that the anaphoric use of overt subjects, namely repeated noun phrases, is a strategy adopted by all groups, but while it decreases with age in EP, it steadily remains in BP. The production of coreferential null pronominal subjects split apart the two linguistic varieties, only EP speakers making an expressive use of them in root and finite subordinate sentences.

KEY WORDS: coreferential chains processing; anaphoric subjects; consistent and partial null subject languages

REFERÊNCIAS

- Arim, E. (2005). Uso de expressões referenciais em narrativas escritas em Português Europeu por crianças chinesas. *Actas do XXI Encontro Nacional da Associação Portuguesa de Linguística*. Lisboa, 125-136.
- Almor, A. (2000). Constraints and mechanisms in theories of anaphor processing. In Crocker, M. W., M. Pickering, & C. Clifton (eds.) *Architectures and Mechanisms for Language Processing* (341-354). Cambridge, UK: Cambridge University Press.
- Ariel, M. (1990). *Accessing NP antecedents*. London: Routledge.
- Ariel, M. (2001). Accessibility theory: an overview. In Sanders, T., Schilperoord, J. & Spooren, W. (eds.). *Text representation: linguistic and psycholinguistic aspects*. Amsterdam: J. Benjamins Publishing Co.
- Barbosa, P. (1996). *Null Subjects*. PhD Dissertation. MIT, Cambridge, Massachusetts.
- Barbosa, P. (2000). “Clitics: a window into the Null Subject Property. In Costa, J. ed., *Portuguese Syntax: Comparative Studies*. New York/Oxford: Oxford University Press.
- Barbosa, P. (2011). Partial pro-drop as null NP-anaphora. *Proceedings of NELS 41*. Amherst, Mass.: GLSA Publications.
- Barbosa, P., M. E. Duarte e M. Kato (2005). Null subjects in European and Brazilian Portuguese. *Journal of Portuguese Linguistics*. Vol. 4, Nº2, 11-52.
- Brito, A. M. (1991). Ligação, correferência e o princípio evitar o pronome. In *Atas do Encontro de Homenagem a Óscar Lopes*, pp. 101-121. Maia: Gráfica Miradouro, APL.
- Batoréo, H. (2000). *Expressão do espaço no português europeu: contributo psicolinguístico para o estudo da linguagem e cognição*. Lisboa: FCG.
- Caramazza, A., Grober, E., Garvey, C., e Yates, J. (1977). Comprehension of anaphoric pronouns. *Journal of Verbal Learning and Verbal Behavior*, 16, 601–609.
- Carminati, M. (2002). *The processing of Italian subject pronouns*. Electronic Doctoral Dissertations for UMass Amherst.
- Chambers, C., e Smyth, R. (1998). Structural parallelism and discourse coherence: A test of centering theory. *Journal of Memory and Language*, 39, 593–608.
- Chomsky, N. (1981). *Lectures on government and binding*, Foris, Dordrecht.
- Chomsky, N. (2001). Derivation by phase. In Kenstowicz, M. and K Hale (eds.) *A life in language*, The MIT, Cambridge Massachusetts Press, pp. 1-52.

- Chomsky, N. (2008). On Phases. In *Foundational Issues in Linguistic Theory* (R Freidin, C Otero and M L Zubizarreta, The MIT Press, Cambridge, Massachusetts, pp. 133-166.
- Chomsky, N. e H. Lasnik (1993). The Theory of principles and parameters. In J. Jacobs, A. Stechow, W. Sternfeld e T. Vennemann, *Syntax: An international handbook of contemporary research*. Berlin: W. de Gruyter.
- Corrêa, L. (1998). Acessibilidade, paralelismo na interpretação do pronome sujeito e o contraste pro/pronome em português. *DELTA: Documentação de Estudos em Lingüística Teórica e Aplicada*, 24. p. 295-392
- Cowles, H. W. (2007). The influence of “aboutness” on pronominal coreference. *ZAS Papers in Linguistics*, 48, 2007: 23 – 38
- Costa, A. (2005). *Processamento de frases em Português Europeu. Aspectos cognitivos e lingüísticos implicados na compreensão da língua escrita*. Lisboa: Fundação Calouste Gulbenkian [versão publicada da tese de doutoramento de 2003, FLUL].
- Costa, A., Faria, I. F. e Matos, G. (1998). Competitive information sources in referential ambiguity resolution. *Proceedings of 5th International Congress of the ISAPL*. Porto, pp. 173-187.
- Costa, A. e Nascimento, E. (2011). Cadeias correferenciais em narrativas escritas em português europeu e português brasileiro. *Actas do III Simpósio Mundial de Estudos da Língua Portuguesa*. Macau, 30 agosto a 2 de setembro de 2011 (em publicação).
- Costa, J. e Duarte, I. (2002). Preverbal subjects in null subject languages are not necessarily dislocated. *Journal of Portuguese Linguistics*. Vol.2, Nº2, 159-176.
- Duarte, M. E. (1993). Do pronome nulo ao pronome pleno. In I. Roberts e M. Kato (orgs.) *Português Brasileiro: uma viagem diacrónica (homenagem a Fernando Tarallo)*. Campinas: Editora da UNICAMP, 107-128.
- Duarte, M. E. e Cavalcante, S. (2008). The Subject Position in Brazilian Portuguese: the Embedding of a Syntactic Change. *University of Pennsylvania Working Papers in Linguistics*, Volume 14, Issue 2 Selected Papers from NWA 36.
- Fayol, M. (2003). Comprendre et produire des textes écrits: l'exemple du récit. In M. Kail & M. Fayol (Eds.) *L'acquisition du langage. Le langage en développement. Au-delà de trois ans*. Paris: PUF. p. 183-214.
- Ferreira, M. (2004). Hyperraising and null subject on Brazilian Portuguese. *MIT Working Papers in Linguistics*. 46, 57-85.
- Figueiredo Silva, M. C. (1996). *A posição do português brasileiro*. Campinas: Editora da UNICAMP.
- Galves, C. (1987) A sintaxe do português brasileiro. *Ensaio Lingüísticos*. 13, 31-50.
- Galves, C. (2001). *Ensaio sobre as gramáticas do Português*. Campinas: Editora da UNICAMP, 43-58.

- Givón, T. (1979). From discourse to syntax: grammar as a processing strategy. In: Givón, T. (ed.) *Discourse and syntax*. New York: Academic Press.
- Grosz, B., Joshi, A., e Weinstein, S. (1995). Centering: A framework for modeling the local coherence of discourse. *Computational Linguistics*, 21, 203-225
- Hickmann, M. (1982). *The Development of Narrative Skills: Pragmatic and Metapragmatic Aspects of Discourse Cohesion*. Tese de Doutorado. University of Chicago.
- Hickmann, M. (1995). Discourse organization and the development of reference to person, space, and time. In: Fletcher, Paul e Brian MacWhinney (eds.) *Handbook of Child Language*. Oxford.
- Holmberg, A. (2005). Is there a little *pro*? Evidence from Finnish. *Linguistic Inquiry* 36 (4), 533-564
- Holmberg A. (2010). Null subject parameters. In Biberauer T., Holmberg A., Roberts I., Sheehan M., (Eds.) *Parametric variation : null subjects in minimalist theory*. Cambridge: Cambridge University Press, 2010, 88-124.
- Holmberg A, Nayudu A, Sheehan M. (2009). Three partial null-subject languages: a comparison of Brazilian Portuguese, Finnish, and Marathi. *Studia Linguistica* 2009,63(1), 59-97.
- Hornstein, H. (1999). Movement and Control. *Linguistic Inquiry*. Vol. 30, Nº1, 69-96.
- Kato, M. (1999). Strong and weak pronominals in the null subject parameter. *Probus*, vol. Nº11, 1-37.
- Levelt, W. (1989). *Speaking from Intention to Articulation*. Cambridge, Mass.: The MIT Press.
- Luegi, P. (2012). *Processamento de sujeitos pronominais em Português: efeito da posição estrutural dos antecedentes*. Tese de doutorado. FLUL.
- López Garcia, Á. (1999). Relaciones paratáticas y hipotáticas. In Bosque, I. e V. Demonte (dir.) *Gramáticas Descriptiva de la Lengua Española*. Madrid: Espasa, vol. 3, 3507-3547.
- Manzini M. R., Roussou A. (2000). A minimalist theory of A-movement and control. *Lingua* 110: 409-447
- Matos, G. (1991). Frases Coordenadas, Sujeito Nulo e Correferência. In *Atas do Encontro de Homenagem a Óscar Lopes*, 123-140. Maia: Gráfica Miradouro, APL.
- Modesto, M. (2010). What Brazilian Portuguese Says about Control: Remarks on Boeckx & Hornstein. *Syntax*, 13:78–96.
- Modesto, M. (2011). Finite control: Where movement goes wrong in Brazilian Portuguese. *Journal of Portuguese Linguistics*, vol. 10, Nº 2, 3-30, 2011.
- Morgado, S. (2011). *Processamento da co-referência pronominal. Informação sintática e semântica*. Tese de mestrado. FLUL.

- Nascimento, E. (2008). *Aprendizagem da escrita em Português Europeu por falantes do Português Brasileiro*. Tese de Mestrado. FLUL.
- Nicol, J. e Swinney, D. (1989). The Role of Structure in Coreference Assignment During Sentence Comprehension. *Journal of Psycholinguistic Research*, Vol. 18, No. 1, 5-19.
- Pontes, E. (1981). “Da importância do tópico em português” *Anais do 5º Encontro de Lingüística*. PUC-RJ. In Pontes, E., *O Tópico no português do Brasil*. Campinas: Pontes.
- Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech & Jan Svartvik (1985). *A Comprehensive Grammar of the English Language*. London: New York: Longman.
- Rizzi, L. (1982). *Issues in Italian Syntax*. Dordrecht: Foris Publications.
- Roberts, I. (2010) A deletion analysis of null subjects”. In Biberauer T., Holmberg A., Roberts I., Sheehan M., (Ed.) *Parametric variation : null subjects in minimalist theory*. Cambridge: Cambridge University Press, 2010, 58-87.
- Rodrigues, C. (2004). *Impoverish Morphology and A-Movement out of case domains*. PhD Dissertation: University of Maryland.
- Smyth, R. (1994). Grammatical determinants of ambiguous pronoun resolution. *Journal of Psycholinguistic Research*, 23, 197–229.
- Taraldsen, K. (1978). *On the NIC, Vacuous Application and the That-Trace Filter*. Bloomington, Indiana University Linguistics Club.