

PONTIFÍCIA UNIVERSIDADE CATÓLICA DO RIO GRANDE DO SUL  
FACULDADE DE LETRAS  
CURSO DE PÓS-GRADUAÇÃO EM LETRAS

**PROJETO:**

**AQUISIÇÃO DA ANÁFORA PRONOMINAL NO PORTUGUÊS DO BRASIL**

por

Prof. Sergio Menuzzi (CPGL)

Porto Alegre, março de 1999.

## **Título do Projeto:**

Aquisição da Anáfora Pronominal no Português do Brasil

## **Sumário:**

0.Dados de Identificação

1.Introdução

2.Revisão bibliográfica preliminar

2.1. 'Efeitos tardios' do Princípio B

2.2. Análises recentes dos 'efeitos tardios do Princípio B'

3. Formas anafóricas em PB: hipóteses iniciais a serem investigadas

4. Objetivos e justificativa do projeto

5. Métodos de investigação

5.1 Investigação de corpora

5.2 Investigação experimental

6. Subprojetos vinculados

7. Cronograma

8. Referências bibliográficas

## **0.Dados de Identificação**

Responsável: Prof. Dr. Sergio de Moura Menuzzi

Instituição: Curso de Pós-Graduação em Letras, PUCRS

Áreas (CNPq): Teoria e Análise Lingüística; Psicolingüística.

Modalidade: Projeto Integrado

Número de bolsas solicitadas: seis (06) bolsas de iniciação científica

Período: Agosto de 1999 a Março de 2002 (duração total de 32 meses)

## **1.Introdução**

A pesquisa aqui proposta pretende investigar a anáfora pronominal intra-sentencial do ponto de vista da gramática gerativa, em particular, do ponto de vista da chamada 'teoria dos princípios e parâmetros' [TPP; ver nota 1 abaixo para referências]. O objetivo específico da pesquisa é testar empiricamente várias análises recentes da aquisição do uso anafórico de pronomes baseadas na TPP. Para isso, quer-se verificar se as predições que estas análises fazem para a aquisição do português falado no Brasil [PB] se confirmam ou não. O interesse desta investigação deriva do fato de que o PB possui um sistema anafórico para o qual as diversas análises a serem consideradas fazem diferentes predições, isto é, a aquisição do PB pode fornecer evidência crucial para se decidir qual das análises tem maiores chances de sucesso.

Este projeto está organizado do seguinte modo: na seção 2, faz-se uma revisão preliminar da literatura sobre os chamados 'efeitos tardios do Princípio B' em diversas línguas, com dois objetivos: (i) apresentar o tópico geral de investigação do projeto, e (ii) introduzir as análises cujas predições para a aquisição do PB este projeto quer verificar. Na seção 3, os aspectos relevantes do sistema anafórico do PB são brevemente descritos, e enumeram-se as diversas hipóteses específicas à aquisição de pronomes em PB a serem inicialmente investigadas. Na seção 3, os objetivos da pesquisa são definidos mais explicitamente e também justificados. Na seção 5, são descritos os métodos de investigação a serem adotados. Na seção 6, dois

subprojetos auxiliares são esboçados, indicando tarefas necessárias ao andamento do presente projeto que deverão ser executadas por bolsistas de iniciação científica. Na seção 7, apresenta-se um cronograma preliminar para a execução do projeto, e, na seção 8, as referências bibliográficas em que ele se baseia inicialmente.

## 2.Revisão Bibliográfica Preliminar

### 2.1 . ‘Efeitos tardios’ do Princípio B

No modelo clássico da teoria de princípios e parâmetros (Chomsky 1981), a subteoria responsável pelas restrições sintáticas à anáfora intra-sentencial é a ‘teoria da ligação’.<sup>1</sup> Os princípios da teoria da ligação que governam a distribuição de reflexivos e pronomes são os Princípios A e B, informalmente apresentados em (1) abaixo:

- (1) a. *Princípio A*: um reflexivo tem que ter um antecedente local.  
b.*Princípio B*: um pronome não pode ter um antecedente local.

De acordo com o Princípio A, uma sentença contendo um reflexivo é gramatical se o reflexivo possui um antecedente dentro de um certo domínio sintático, por exemplo, dentro da oração que o contém, como em (2a) abaixo (de acordo com a notação usual, dois NPs têm a mesma interpretação se possuem o mesmo ‘índice’ ou letra subscrita):

- (2a) [O Paulo<sub>j</sub> disse que [O João<sub>i</sub> já se<sub>i</sub> viu na TV ]]

Se o reflexivo não possui um antecedente dentro de um tal domínio, a sentença é agramatical, como em (2b):

- (2b) \* [O Paulo<sub>j</sub> disse que [O João<sub>i</sub> já se<sub>j</sub> viu na TV ]]

Analogamente, uma sentença é gramatical para o Princípio B se o pronome *não* possui um antecedente dentro, por exemplo, da mesma oração que o contém, como em (3a) abaixo; caso contrário, a sentença é agramatical, como em (3b):

- (3) a. [O Paulo<sub>j</sub> disse que [O João<sub>i</sub> já viu ele<sub>j</sub> na TV ]]  
b.\* [O Paulo<sub>j</sub> disse que [O João<sub>i</sub> já viu ele<sub>i</sub> na TV ]]

Como veremos brevemente, as versões subseqüentes da teoria da ligação a serem consideradas pelo presente projeto diferem da teoria clássica em vários aspectos, com conseqüências diretas para a investigação da aquisição do uso anafórico de pronomes.

No que diz respeito à teoria clássica, o incentivo inicial para a investigação de dados de aquisição veio da possibilidade de testar a hipótese de que os Princípios A e B fazem parte da gramática universal. Isto é, os Princípios A e B em (1) seriam parte constitutiva do conhecimento lingüístico inato que permite à criança adquirir a distribuição dos elementos anafóricos de sua língua.<sup>2</sup> Como Grimshaw & Rosen (1990:187) observaram, esta hipótese

---

<sup>1</sup> Para uma apresentação e história da teoria de princípios e parâmetros, ver Chomsky (1986), Chomsky & Lasnik (1993); para uma apresentação e história da teoria da ligação, ver Lasnik & Uriagereka (1988, capítulo 2), Lasnik (1989) e Menuzzi (1997, capítulo 1).

<sup>2</sup> Lembre-se aqui que o “programa gerativista” de pesquisa tem como objetivo determinar as propriedades da “faculdade de linguagem” ou “gramática universal” — i. é, a hipotética capacidade cognitiva específica aos seres humanos que lhes permite adquirir um complexo sistema de conhecimento, a gramática de sua língua materna, apesar da pobreza relativa de sua experiência lingüística. Ver Chomsky (1965, capítulo 1), (1986, capítulos 1 e 2), Lightfoot (1982), Hornstein & Lightfoot (1981), Crain (1991), entre outros, para discussão.

tem uma interpretação óbvia no que diz respeito à aquisição: as crianças deveriam demonstrar comando destes princípios desde o começo, isto é, desde a emergência do uso anafórico de pronomes e reflexivos; em particular, a emergência dos efeitos dos Princípios A e B deveria ser simultânea, já que ambos princípios são, por hipótese, inatos e pertencem a um mesmo módulo gramatical, a teoria da ligação.<sup>3</sup>

Entretanto, um dos fatos mais bem estabelecidos sobre o desenvolvimento da anáfora intra-sentencial contradiz esta ‘interpretação óbvia’ da teoria da ligação: resultados experimentais em diversas línguas demonstram que as crianças adquirem o comando de reflexivos antes do que o comando de pronomes. Isto é, as crianças atingem o desempenho adulto primeiro com sentenças envolvendo o Princípio A, e mais tarde com sentenças envolvendo o Princípio B (ver Lust 1986, Grimshaw & Rosen 1990, McKee 1992, Koster 1993, para revisões recentes da literatura). Considere, por exemplo, os resultados experimentais reportados na tabela 1 abaixo, obtidos por McKee (1992) em um teste de julgamento com crianças americanas:<sup>4</sup>

Tabela 1:

Porcentagem de respostas corretas em inglês (cf. McKee 1992)

Idade (média aproximada)	3 anos e meio	4 anos e meio
sentenças com reflexivos	88%	95%
sentenças com pronomes	65%	57%

Por volta dos três anos e meio, crianças em fase de aquisição do inglês já mostram um desempenho bastante satisfatório com sentenças envolvendo o Princípio A (em torno de 90% de respostas corretas), progredindo até atingir um desempenho próximo ao do adulto (95% de respostas corretas) aos 4 anos e meio. Com sentenças envolvendo o Princípio B, entretanto, o desempenho por volta dos 3 anos e meio é apenas regular (a literatura considera que as chances de uma resposta experimental correta ou incorreta são, em geral, as mesmas, isto é, em torno de 50% para cada) e parece não mostrar progresso no período que vai até os 4 anos e meio.

Também em holandês as crianças atingem um desempenho próximo ao dos adultos com reflexivos antes do que com pronomes, como ilustram na tabela 2 os resultados obtidos por Deutsch *et alii* com crianças holandesas em um teste de compreensão (cf. Koster 1993:37, tabela 3):

<sup>3</sup> As hipóteses implícitas por trás deste raciocínio são as de que: (i) se um princípio é inato, ele deve estar ativo desde o começo da emergência do domínio cognitivo que ele governa; (ii) se um conjunto de princípios se refere a um mesmo conjunto de conceitos (isto é, os princípios pertencem a um mesmo ‘módulo’ cognitivo), estes princípios devem emergir simultaneamente. Evidentemente, nenhuma das duas hipóteses é necessária: p. ex., a emergência de princípios e módulos pode depender de (diferentes etapas do processo de) maturação biológica do cérebro (cf. Borer & Wexler 1987).

<sup>4</sup> Note que, para atingir o desempenho de adultos com sentenças envolvendo o Princípio A e, portanto, responder corretamente testes com tais sentenças, a criança tem que: (i) aceitar aquelas em que o reflexivo possui um antecedente local (p. ex., (2a) acima), e (ii) recusar as sentenças em que o reflexivo *não* possui um tal antecedente (p. ex., (2b) acima). Analogamente, para responder corretamente testes com sentenças envolvendo o Princípio B, a criança tem que aceitar aquelas em que o pronome *não* possui um antecedente local (p. ex., (3a) acima), e tem que recusar as sentenças em que o pronome *possui* um tal antecedente (p. ex., (3b) acima). Quanto aos testes usados para verificar a competência das crianças, ver seção 5.2 abaixo e referências lá citadas.

Tabela 2:

Porcentagem de respostas corretas em holandês  
(cf. Deutsch *et alii*, *apud* Koster 1993)

Idade (média aproximada)	4 anos	6 anos	8 anos	10 anos
sentenças com reflexivos	51%	84%	95%	99%
sentenças com pronomes	56%	48%	70%	89%

O desempenho inicial das crianças holandesas é regular tanto com reflexivos como com pronomes, mas progride rapidamente com reflexivos, atingindo um desempenho próximo ao dos adultos por volta dos 8 anos de idade; com pronomes, progresso até os 8 anos é lento, e, mesmo aos 10 anos, o desempenho das crianças holandesas ainda está abaixo do dos adultos. Padrões semelhantes foram também encontrados na aquisição do francês (Jakubowicz 1990), do russo (Avrutin & Wexler 1992), e do islandês (Sigurjónsdóttir & Hyams 1992).

Este fenômeno generalizado da emergência tardia do comando de pronomes em relação ao comando de reflexivos poderia sugerir a seguinte hipótese: nos estágios iniciais de aquisição da anáfora intra-sentencial, a criança tem conhecimento do Princípio A (que governa reflexivos), mas não do Princípio B (que governa pronomes). Tal hipótese, entretanto, jamais chegou a ser proposta porque, mesmo nos trabalhos iniciais sobre o assunto, percebeu-se que há evidência de que algo como o Princípio B está em ação desde o começo do processo de aquisição.

Especificamente, há uma diferença sistemática entre o desempenho das crianças com as sentenças que o Princípio B define como gramatical, e seu desempenho com as que ele define como agramaticais: com sentenças que satisfazem o Princípio B, as crianças atingem altos níveis de acerto (em torno de 90% de respostas corretas); com sentenças que *não* satisfazem o Princípio B, por outro lado, seu desempenho é aleatório (apenas cerca 50% de respostas corretas inicialmente; ver Grimshaw & Rosen 1990:189-99).

Ora, crianças que, por hipótese, não conhecem o Princípio B não podem ser capazes de distinguir entre as sentenças que violam o princípio e as que não o violam. Tais crianças deveriam, portanto, aceitar igualmente ambos os tipos de sentenças. Como, na verdade, o desempenho das crianças mostra que elas distinguem sistematicamente entre os dois tipos de sentenças, conclui-se que elas dispõem do conhecimento responsável pelos efeitos do Princípio B. Disso segue que algum outro fator deve ser responsável por sua baixa performance com pronomes em relação a reflexivos.

## 2.2. Algumas análises recentes dos ‘efeitos tardios do Princípio B’

Que fatores poderiam estar afetando a performance das crianças com pronomes? Várias análises têm sido propostas para o problema, quase todas sustentando que os fatos de aquisição demonstram que a formulação clássica da teoria da ligação tem de ser modificada. Assim, a investigação da aquisição das restrições sintáticas sobre o uso anafórico de pronomes se tornou uma fonte crucial de evidência para diferentes versões da teoria da ligação e, com isso, o estudo dos dados de aquisição cresceu em importância para a investigação da teoria geral da anáfora intra-sentencial.

Chien & Wexler (1985, 1990) foram os primeiros a argumentar que os efeitos tardios do Princípio B oferecem evidência para uma versão alternativa da teoria da ligação. O argumento é baseado nos seguintes resultados que eles mesmos obtiveram: (i) as crianças americanas (entre os 5 e os 6 anos de idade) parecem de fato ter problemas em perceber violações do Princípio B quando o pronome é interpretado *co-referencialmente*, i. é, em sentenças como (4a) abaixo; (ii) entretanto, seu desempenho é tão bom quanto o com



de aquisição. Vejamos mais detalhadamente como isto explica o desempenho das crianças com (4a).

Em uma frase como (4a) o pronome pode, em princípio, ser interpretado ou como uma variável ligada, ou co-referencialmente.<sup>7</sup> Presumindo que as crianças não têm preferência por nenhuma destas interpretações, as chances são de que, em cerca de 50% do tempo, elas atribuirão a interpretação de variável ligada ao pronome, e, nos outros 50%, a interpretação co-referencial. Quando o pronome é interpretado como variável ligada (50% do tempo), frases como (4a), do mesmo modo que (4b), são excluídas pelo Princípio B (que, por hipótese, está disponível desde o começo da aquisição). Isto explicaria os cerca de 50% de respostas corretas com (4a).

Quando o pronome é interpretado co-referencialmente (os outros 50% do tempo), frases como (4a) deveriam ser excluídas pelo Princípio P. Entretanto, como vimos, Chien & Wexler propõem que este princípio não está disponível nos estágios iniciais de aquisição. Disso segue que, nestes estágios, a criança não é capaz de distinguir uma violação do Princípio P de uma não-violação. Logo, a criança deve aceitar, nestes estágios, todos os casos em que o pronome é interpretado co-referencialmente (50% do tempo). Isto daria conta, então, dos 50% de respostas incorretas para (4a).

Em suma, para Chien & Wexler (1990), o padrão de aquisição revelado pelos resultados em (4) pode ser interpretado como evidência a favor de uma teoria alternativa à teoria da ligação clássica. De acordo com esta teoria alternativa, somente relações anafóricas envolvendo variáveis ligadas são regidas pelos princípios da teoria da ligação; relações co-referenciais seriam regidas por um princípio pragmático, o Princípio P. Na análise de Chien & Wexler, o problema das crianças nos estágios iniciais de aquisição não se deve à ausência do Princípio B — que pertence à teoria da ligação —, mas à ausência do princípio pragmático P.<sup>8</sup>

Grimshaw & Rosen (1990), por outro lado, argumentam que os efeitos tardios do Princípio B são compatíveis com a formulação clássica da teoria da ligação, e que os aparentes problemas de desempenho mostrados pelas crianças são, na verdade, artefatos experimentais. Isto é, Grimshaw & Rosen sustentam que os resultados experimentais nos quais as conclusões da literatura precedente se baseiam não são confiáveis porque os experimentos não controlaram fatores que induzem sistematicamente ao baixo desempenho com pronomes.

Em particular, no que diz respeito ao contraste encontrado por Chien & Wexler (cf. (4a,b) acima), Grimshaw & Rosen sugerem que o aparente alto índice de respostas corretas com

---

<sup>7</sup> Como Lasnik (1976) e Reinhart (1983) demonstraram, antecedentes referenciais não requerem uma interpretação co-referencial para o pronome. Isto é ilustrado pela sentença (ia) abaixo, que é ambígua entre as interpretações em (ib) e (ic):

- (i) a. O João disse que a Maria confia nele, e o Paulo também  
b. (João disse que Maria confia em *João*) & (Paulo disse que Maria confia em *João*)  
c. João ( $\lambda x$  (x disse que Maria confia em x)) & Paulo ( $\lambda x$  (x disse que Maria confia em x))  $\rightarrow$   
(João disse que Maria confia em *João*) & (Paulo disse que Maria confia em *Paulo*)

<sup>8</sup> Grodzinsky & Reinhart (1993) propuseram recentemente uma variante desta análise: a baixa performance das crianças com sentenças envolvendo co-referência não se deveria à ausência do Princípio P nos estágios iniciais de aquisição, mas a limitações psico-cognitivas independentes. Especificamente, para ser aplicado, o Princípio P requer a comparação de representações semânticas alternativas para uma mesma sentença, uma tarefa que sobrecarregaria a capacidade de processamento da criança. Grodzinsky & Reinhart argumentam que esta análise é sustentada pelo fato de que: (i) adultos afásicos mostram o mesmo padrão de desempenho que as crianças com sentenças como (4a), e (ii) tanto afásicos quanto crianças têm problemas com tarefas que envolvam a comparação de duas representações linguísticas (1993:88-91).

sentenças envolvendo pronomes interpretados como variáveis ligadas (i. é, (4b)) pode ter uma explicação alternativa: as crianças podem ter dificuldades em processar estruturas quantitativas e/ou atribuir a interpretação de variável ligada a pronomes, já que uma tal interpretação é semanticamente mais complexa do que a interpretação co-referencial. Para distinguir esta possibilidade da possibilidade de que o bom desempenho com (4b) seja simplesmente uma consequência do Princípio B, Grimshaw & Rosen notam que seria necessário investigar o desempenho das crianças com sentenças em que os pronomes são interpretados como variáveis ligadas, mas *não* violam o Princípio B. Isto é, é preciso fazer uma comparação entre o desempenho das crianças com (4a,b) acima e seu desempenho com sentenças como (5a,b) abaixo:

- (5) a. *The boy* told Sam that Mary kissed *him*  
 b. *Every boy* tells Sam that Mary kissed *him*

Se (5a,b) repete o padrão de (4a,b), i. é, se as crianças têm maior dificuldade em (5b) do que em (5a) em interpretar o sujeito matriz como o antecedente de *him*, ter-se-ia razão para acreditar que os bons resultados obtidos com (4b) são, na verdade, influenciados por dificuldades com pronomes interpretados como variáveis ligadas. Tal possibilidade, entretanto, não foi investigada experimentalmente nem por Grimshaw & Rosen, nem, tanto quanto sei, pela literatura posterior.<sup>9</sup>

Além disso, a análise proposta por Chien & Wexler enfrenta problemas com um outro fato crucial observado pela literatura: os ‘efeitos tardios do Princípio B’ variam de língua para língua. Para dar alguns exemplos: (i) como vimos, experimentos reportados em Koster (1993) mostram que, em holandês, as crianças só atingem um desempenho próximo do dos adultos com pronomes (entre 80 e 90% de respostas experimentais corretas) por volta dos 10 anos; (ii) também vimos que, de acordo com os resultados obtidos por Chien & Wexler (1990), em inglês isto acontece por volta dos 7 anos de idade (85% de respostas experimentais corretas); (iii) em italiano, por outro lado, McKee (1992) descobriu que as crianças atingem cedo um desempenho próximo do de adultos (em torno de 90%), por volta dos 4 anos (os pronomes testados por McKee em italiano são os clíticos: ver discussão abaixo):

Tabela 3:

Variação entre línguas quanto à idade de emergência de desempenho próximo ao do adulto com pronomes

Língua	Emergência do Desempenho Adulto	Fonte
holandês	em torno dos 10 anos	Koster (1993)
inglês	em torno dos 7 anos	Chien & Wexler (1990)
italiano	em torno dos 4 anos	McKee (1992)

É improvável que o desenvolvimento de um princípio pragmático como o princípio P varie de acordo com a língua que as crianças estão adquirindo, já que este princípio não diz respeito a

<sup>9</sup> Koster (1993) reporta, entretanto, que experimentos com sentenças contendo elipse do VP em holandês parecem sustentar a idéia de que as crianças têm problemas com pronomes interpretados como variáveis ligadas, o que reforça a sugestão de Grimshaw & Rosen.

propriedades particulares das línguas (ver n. 6 acima). Portanto, a análise de Chien & Wexler nada tem a dizer sobre a variação ilustrada pela tabela 1.<sup>10</sup>

Grimshaw & Rosen sugerem que a diferença entre o padrão de aquisição do inglês e o do italiano é também um artefato experimental: os pronomes ingleses, sendo formas plenas, podem ser usados enfática ou contrastivamente, enquanto que os pronomes italianos, sendo formas clíticas, não podem (cf. Grimshaw & Rosen 1990:209-11). Como se sabe, pronomes enfáticos podem ter um antecedente local, em aparente violação do Princípio B. Isto significaria que os ‘efeitos tardios do Princípio B’ em línguas como o inglês poderiam simplesmente ser resultado do fato de que os experimentos relevantes não controlam a possibilidade de as crianças poderem estar interpretando os pronomes enfaticamente em inglês, uma possibilidade inexistente para os clíticos italianos.<sup>11</sup>

Há, entretanto, um problema para esta análise da variação entre o inglês e o italiano. Se o baixo desempenho inicial com pronomes em inglês se devesse pura e simplesmente à possibilidade inerente de, nas condições experimentais relevantes, pronomes serem interpretados enfaticamente em inglês, não se esperaria encontrar diferenças significativas no desempenho de crianças de diferentes faixa etárias, ou entre o desempenho de crianças e adultos. Mas é isso o que se encontra, i. é, há *evolução* no comando de pronomes em inglês ao longo do tempo.<sup>12</sup> Assim, a análise de Grimshaw & Rosen requer uma hipótese adicional que explique porque os aprendizes de inglês parecem interpretar pronomes enfaticamente com maior frequência nos estágios iniciais de aquisição.

Em termos mais gerais, é evidente que a dicotomia clíticos/não clíticos não pode ser a única fonte de variação nos padrões de aquisição de pronomes: ela não explica, por exemplo, as diferenças entre o inglês e o holandês (cf. tabela 3), uma vez que pronomes holandeses podem ser usados enfaticamente, tal como os pronomes ingleses. Entretanto, a idéia de que a diferença entre clíticos e não-clíticos é relevante sugere uma estratégia plausível de aproximação ao problema da variação nos ‘efeitos tardios do Princípio B’: a de que a variação neste domínio, como a variação geral entre línguas, se reduz às diferenças que distinguem seus léxicos. Isto é, parece natural esperar que uma das fontes possíveis de variação nos padrões de

---

<sup>10</sup> O mesmo pode ser dito da variante proposta por Grodzinsky & Reinhart (1993), cf. n. 8: não há evidência, por exemplo, de que a capacidade de processamento de crianças holandesas de 7 anos de idade difira substancialmente da de crianças americanas da mesma idade.

<sup>11</sup> Segundo Grimshaw & Rosen, os resultados para o espanhol discutidos em Solan (1987), obtidos com crianças bilíngües de espanhol-inglês, sustentariam esta hipótese. Tais resultados, entretanto, não correspondem aos obtidos por McKee para o italiano; antes, demonstram que os aprendizes de espanhol têm dificuldades iniciais tanto com reflexivos quanto com pronomes, sendo que o que é notável em relação ao inglês é que mostram maior comando de *pronomes* (57% de respostas corretas) do que de reflexivos (41% de respostas corretas). Grimshaw & Rosen afirmam que tais resultados podem não ser diretamente interpretáveis nos mesmos termos que os resultados obtidos para o inglês (i. é, exclusivamente em termos da teoria da ligação), já que ‘as construções românicas [com clíticos] podem não ser comparáveis em complexidade às construções inglesas com não-clíticos’ (1990:211, n.17).

O problema com os clíticos românicos se torna ainda mais interessante quando se nota que Jakubowicz (1990) obteve para os clíticos do francês um padrão de aquisição que difere tanto do obtido por McKee para o italiano quanto do discutido por Solan com relação ao espanhol. Antes, o padrão obtido por Jakubowicz para o francês é similar ao encontrado em inglês: com pronomes, as crianças testadas por Jakubowicz só atingiram um desempenho próximo do de adultos (90% de respostas corretas no teste) entre os 6 e os 7 anos de idade; com reflexivos, elas já mostram um desempenho próximo ao dos adultos (90% de respostas corretas) em torno dos 4 anos de idade.

Evidentemente, estas diferenças entre o italiano, o espanhol e o francês podem ser consequência de diferenças nos experimentos utilizados. A fim de controlar tal possibilidade, seria interessante tentar replicar os resultados acima referidos sob condições experimentais uniformizadas.

<sup>12</sup> P. ex., Chien & Wexler (1990) reportam os seguintes estágios com antecedentes referenciais: 30% de respostas corretas por volta dos 3 anos de idade, em torno de 50% entre os 4 e os 6 anos de idade, e 75% por volta dos 7 anos.

aquisição de pronomes e reflexivos esteja justamente no processo de descoberta das propriedades lexicais destas formas, já que pronomes e reflexivos de línguas diferentes divergem quanto às suas propriedades lexicais (se são ou não formas clíticas, p. ex.).

Na verdade, esta idéia está em consonância com os esforços teóricos recentes que tentam reduzir a gama de variação lingüística possível no domínio da anáfora intra-sentencial às distinções lexicais que podem ser estabelecidas entre pronomes e outras formas anafóricas (ver Menuzzi 1997, especialmente o capítulo 1, para discussão extensiva). Tais esforços têm levado a modificações consideráveis no aparato teórico básico da teoria da ligação e, em conseqüência, a novas possibilidades de interpretação dos dados da aquisição. Um exemplo recente disso é a análise proposta por Coopmans e seus colaboradores para o contraste entre os processos de aquisição de pronomes em inglês e holandês.

Como vimos acima (cf. tabela 3) a diferença básica entre as duas línguas é a de que o comando de pronomes —isto é, a descoberta de que pronomes *não* podem ter um antecedente local —parece emergir muito mais tarde em holandês. Resultados obtidos por Sigurjónsdóttir & Coopmans (1996) e Philip & Coopmans (1996) sugerem, por outro lado, que a principal fonte de problemas para as crianças holandesas reside num tipo específico de sentenças: em torno dos seis anos, seu desempenho com sentenças como (6a) abaixo é comparável ao desempenho geral de crianças inglesas da mesma idade (em torno de 50% de respostas corretas), mas não com sentenças como (6b), com as quais seu desempenho é muito inferior (somente 20% de respostas corretas):

- |     |    |                                    |                               |
|-----|----|------------------------------------|-------------------------------|
| (6) | a. | Jan wilde dat Ad hem aaide         | 50% das respostas incorretas: |
|     |    | Jan queria que Ad ele(obj) tocasse | [hem] = [Ad]                  |
|     |    | ‘Jan queria que Ad tocasse (n)ele’ |                               |
|     | b. | Jan wilde dat Ad hem waste         | 80% das respostas incorretas: |
|     |    | Jan queria que Ad ele(obj) lavasse | [hem] = [Ad]                  |
|     |    | ‘Jan queria que Ad lavasse ele’    |                               |

Segundo Coopmans e colaboradores, tais resultados fornecem evidência direta para a revisão da teoria da ligação proposta por Reinhart & Reuland [R&R], uma das tentativas recentes de explicar a distribuição de formas anafóricas a partir de suas propriedades lexicais inerentes. Na teoria de R&R, os Princípios A e B da teoria da ligação clássica são reinterpretados como condições sobre a representação de predicados reflexivos. Tais condições distinguem anafóricos complexos como *himself* em inglês, de outros tipos NPs (incluindo pronomes e anafóricos como *se/si* em português). Além disso, relações anafóricas locais como as exemplificadas em (6a,b) também estão sujeitas a ação de uma condição sobre cadeias, a Condição da Cadeia, que distingue anafóricos subespecificados, como *se/si* em português, de outros tipos de NPs (incluindo, de novo, pronomes; cf. R&R 1992, 1993, 1995; ver Menuzzi 1997, seção 1.5 para um sumário detalhado, e as ns. 13 e 15 abaixo para uma breve apresentação).

Para Coopmans e seus colaboradores, o aspecto crucial da teoria da ligação de R&R é seguinte: o pronome *hem*, se coindexado com *Ad*, viola duas restrições em (6a) —o Princípio B revisado e a Condição da Cadeia; em (6b), por outro lado, *hem* viola somente a Condição da Cadeia:<sup>13</sup>

<sup>13</sup> Na teoria de R&R (1993), o Princípio B requer que predicados interpretados reflexivamente sejam marcados por um anafórico complexo (como *himself* em inglês e *zichzelf* em holandês) se o predicado *não* for inerentemente reflexivo. Como o predicado *wassen* ‘lavar(-se)’ em (6b) é um predicado inerentemente reflexivo, o anafórico complexo não é requerido e, portanto, o pronome *hem* não é excluído. Quanto à Condição da Cadeia, ver a n. 15 abaixo.

- (6\*) a. Jan<sub>j</sub> wilddat Ad<sub>i</sub> hem<sub>i</sub> aaide \*Princípio B, \*Condição da Cadeia  
 b. Jan<sub>j</sub> wilddat Ad<sub>i</sub> hem<sub>i</sub> waste ✓Princípio B, \*Condição da Cadeia

Para dar conta, então, do alto índice de respostas incorretas com sentenças como (6b), Coopmans e seus colaboradores propõem que as crianças holandesas têm dificuldades em determinar certas propriedades lexicais dos pronomes de sua língua materna que acionam Condição da Cadeia (propriedades que discutirei em seguida). Assim, pronomes *não* violariam esta condição nos estágios iniciais de aquisição do holandês e, por isso, seriam aceitos localmente ligados em sentenças como (6b) a maior parte do tempo. Em sentenças como (6a), por outro lado, a ação do Princípio B seria responsável pelo sucesso parcial das crianças em excluir pronomes localmente ligados (os 50% de respostas incorretas sendo resultado dos casos em que o pronome é interpretado co-referencialmente, com os quais as crianças também teriam problemas, cf. a análise de Chien & Wexler e de Grodzinsky & Reinhart).<sup>14</sup>

Retornando ao contraste entre o holandês e o inglês no desenvolvimento do comando de pronomes, a hipótese de Coopmans e colaboradores, então, é a de que as crianças holandesas adquirem mais tardiamente controle sobre os pronomes de sua língua materna porque demoram a identificar propriedades lexicais relevantes para a ação da Condição da Cadeia; o mesmo problema não se poria, por outro lado, à aquisição dos pronomes em inglês. Que propriedades lexicais seriam estas que acionariam a Condição da Cadeia em inglês, mas não em holandês, nos estágios iniciais da aquisição de pronomes?

Coopmans e seus colaboradores sugerem que a propriedade relevante é a especificação de pronomes para *Caso estrutural* (por oposição à especificação para Caso inerente, que *não* aciona a Condição da Cadeia; cf. Philip & Coopmans 1996:252-4).<sup>15</sup> Em ambas as línguas as formas pronominais para objeto direto e indireto são homofônicas (*him/her/it/them* em inglês, *hem/haar/het/hun* em holandês); em inglês, entretanto, elas recebem sempre Caso estrutural, enquanto que em holandês elas podem receber tanto Caso estrutural quanto Caso inerente.<sup>16</sup> Como esta distinção não recebe manifestação superficial nos pronomes holandeses, as crianças teriam dificuldades em determinar se estes estão ou não recebendo Caso estrutural e,

<sup>14</sup> Coopmans e colaboradores adotam a análise de Grodzinsky & Reinhart (1993, ver n. 8 acima) para (6a). Entretanto, pode-se demonstrar que a análise que propõem para o padrão em (6a,b) como um todo não é completamente consistente com o padrão descoberto por Chien & Wexler para o inglês (cf. (4a,b)), demonstração esta que deixarei de lado aqui por razões de espaço.

<sup>15</sup> A Condição da Cadeia de R&R (1993, 1995) requer que, em uma cadeia-A, a posição mais alta da cadeia (o antecedente) seja ocupada por um elemento completamente especificado para traços- $\phi$  e para *Caso estrutural*; as demais posições (i. é, as posições coindexadas com o antecedente) devem ser subespecificadas ou para traços- $\phi$ , ou para Caso estrutural. Pronomes violariam esta condição porque, por hipótese, seriam completamente especificados para os traços relevantes. Ver, entretanto, Menuzzi (1997), seções 4.5 a 4.7, para discussão extensiva dos inúmeros problemas levantados por esta formulação da Condição da Cadeia, bem como para uma alternativa.

<sup>16</sup> A evidência que Coopmans e seus colaboradores oferecem para isso é a seguinte: tanto objetos diretos quanto objetos indiretos podem ser passivizados em inglês, enquanto que somente objetos diretos o podem em holandês, cf.

- (i) a. John gave two books to him      Jan heeft twee boeken aan hem gegeven  
 b. Two books were given to him      Twee boeken worden aan hem gegeven  
 c. John gave him two books      Jan heeft hem twee boeken gegeven  
 d. He was given two books      \*Hij werd twee boeken gegeven

Se é verdade que somente NPs com Caso estrutural podem ser passivizados, então os exemplos em (i) demonstram que objetos indiretos recebem Caso estrutural em inglês, não em holandês.

portanto, se eles estão ou não violando a Condição da Cadeia. Em inglês, por outro lado, as crianças não teriam de enfrentar este problema, já que pronomes em posição de objeto recebem sempre Caso estrutural.

Philip & Coopmans (1996) oferecem um argumento adicional para esta análise. Em um experimento em que procuravam corroborar os resultados de Sigurjónsdóttir & Coopmans (1996), Philip & Coopmans obtiveram um resultado surpreendente: há um contraste na aquisição do comando das formas de 3a. pessoa do singular em holandês —a forma feminina (*haar*) só parece começar a sofrer os efeitos da Condição da Cadeia cerca de um ano após tais efeitos terem aparecido com a forma masculina (*hem*) (percentagens aproximadas, cf. Philip & Coopmans 1996:249):

Tabela 4:

Respostas incorretas com *hem* e *haar* em (6b) (cf. Philip & Coopmans 1996)

Idade	4-6 anos	7 anos	8 anos
<i>hem</i> ‘ele(obj)’	80%	<b>50%</b>	<b>50%</b>
<i>haar</i> ‘ela(obj)’	80%	80%	<b>50%</b>

Philip & Coopmans notam que a diferença básica entre *hem* e *haar* tem a ver com as distinções de Caso que estas formas podem expressar: neste sentido, *haar* parece ser ainda mais ‘ambíguo’ que *hem*, já que pode ocorrer num número maior de contextos, como mostram os paradigmas abaixo:

(7)	<i>Função Gramatical</i>	[3ps, masc.]	[3ps., fem.]
	Sujeito	<i>hij</i>	<i>zij</i>
	Possessivo	<i>zijn</i>	<b><i>haar</i></b>
	Objeto Direto	<b><i>hem</i></b>	<b><i>haar</i></b>
	Objeto Indireto	<b><i>hem</i></b>	<b><i>haar</i></b>
	Objeto de Preposição	<b><i>hem</i></b>	<b><i>haar</i></b>

O padrão de aquisição mostrado na tabela 4 acima sugere, então, segundo Philip & Coopmans, que quanto menos especificada uma forma pronominal é para contrastes de Caso gramatical, mas difícil será para a criança determinar se esta forma viola ou não a Condição da Cadeia.

Ou seja, a diferença no padrão de aquisição de *hem* e *haar* em holandês sustenta a hipótese de que diferenças lexicais quanto à especificação de Caso podem determinar diferentes padrões de aquisição do comando de pronomes. E isso, por sua vez, sustenta a análise de Coopmans e seus colaboradores dos contrastes entre o inglês e o holandês neste terreno. Mais geralmente, estes resultados corroborariam, por um lado, (i) a hipótese de R&R de que há uma condição atuando na anáfora intra-sentencial que é sensível à especificação para Caso estrutural das formas anafóricas, e, por outro, (ii) a idéia geral de que as condições que atuam sobre a anáfora intra-sentencial fazem referência às propriedades lexicais inerentes das formas anafóricas.

Também recentemente, vários trabalhos sobre a anáfora intra-sentencial têm sugerido um outro papel para o léxico na determinação da distribuição de formas anafóricas: a idéia básica, que tem sido formulada de diversas maneiras, é a de que as formas anafóricas de uma língua estão *em competição* entre si, e que pronomes são, em geral, o ‘elsewhere case’, isto é, eles são escolhidos quando os demais competidores —os reflexivos —são excluídos (ver Burzio 1989, 1992, e Menuzzi 1997 para implementações alternativas, bem como para a motivação desta idéia).

Em teorias como a teoria da ligação clássica ou as demais alternativas que consideramos até aqui, a distribuição de uma forma anafórica X depende exclusivamente das condições às quais X está sujeita. Numa teoria em há competição entre formas, por outro lado, a distribuição de X depende, também e crucialmente, da distribuição das demais formas anafóricas contidas no léxico da língua: para X ser escolhida em um contexto, as demais alternativas devem ser excluídas daquele contexto. Disso segue que a escolha da forma apropriada para um contexto depende de uma hierarquia de exclusão: é necessário primeiro que se determine a distribuição da forma que sofre maiores restrições para depois determinar a distribuição das formas que sofrem menos restrições. Daí porque pronomes são o ‘elsewhere case’ nestas teorias: pronomes estão sujeitos a menos restrições distribucionais que reflexivos.

Evidentemente, esta maneira de conceber a teoria da ligação tem conseqüências para a aquisição da anáfora intra-sentencial. P. ex., como Burzio (1992:10, n. 2) percebeu, o fato de que pronomes são o ‘elsewhere case’ nestas teorias explicaria porque se encontra somente ‘efeitos tardios do Princípio B em relação ao Princípio A’, e não vice-versa: a distribuição de pronomes emergerá sempre mais tarde, já que, para determiná-la, é preciso primeiro determinar a distribuição de formas reflexivas.

Uma outra predição feita pela hipótese de que as formas anafóricas estão em competição é a de que elas *não* serão adquiridas simultaneamente, mas sim de acordo com a hierarquia de exclusão que caracterizam: formas com distribuição mais limitada tem de ser adquiridas primeiro, formas com distribuição menos limitada depois. Embora esta predição não tenha sido investigada até aqui, há evidência de que ela pode estar correta. P. ex., em holandês há basicamente três competidores de 3a. pessoa: pronomes, o reflexivo *zichzelf* e o anafórico *zich*. Considere agora o desempenho das crianças com cada uma destas formas ao longo do tempo:<sup>17</sup>

Tabela 5:

Percentagem de respostas corretas com *zichzelf*, *zich* e pronomes  
(cf. Deutsch *et alii*, *apud* Koster 1993)

Idade (Média)	4 anos	6 anos	8 anos	10 anos
<i>zichzelf</i>	51%	<b>84%</b>	<b>95%</b>	<b>99%</b>
<i>zich</i>	22%	53%	<b>87%</b>	<b>90%</b>
pronomes	47%	51%	70%	<b>85%</b>

O padrão mostrado na tabela 5 é compatível com a seguinte ordem de aquisição: o comando de *zichzelf* emerge primeiro (por volta dos 6 anos), é seguido pela aquisição de *zich* (por volta dos 8), e finalmente de pronomes (por volta dos 10 anos). E essa ordem de aquisição corresponde justamente à hierarquia das formas no que diz respeito à sua distribuição (*zichzelf* só ocorre como co-argumento de seu antecedente; *zich* ocorre em todos os contextos locais em que *zichzelf* não é obrigatório, e pronomes, em todos os demais contextos; ver Menuzzi 1997, caps. 1 e 4).

Uma outra característica do padrão revelado pela tabela 5 reforça a idéia que a ordem de aquisição é determinada pela exclusão de competidores: o desempenho com uma forma só mostra progresso significativo quando a forma competidora de distribuição mais limitada está próxima de ser completamente adquirida, o que é justamente o que a teoria da competição

<sup>17</sup> Os resultados da tabela 5 compilam os reportados por Koster (1993) em suas tabelas 2 (p.35) e 3 (p.37). Deve-se, entretanto, tomar alguma cautela na comparação dos resultados para *zichzelf* e *zich*, já que foram obtidos em dois experimentos diferentes. Os resultados para pronomes são a média da porcentagem de respostas corretas nestes mesmos dois experimentos.

prediz. Assim, o desempenho com *zich* só ultrapassa a barreira dos 80% de respostas corretas depois que as crianças ultrapassaram os 90% de correção com *zichzelf* (o que se pode tomar como um sinal de que o comando de *zichzelf* está próximo da aquisição completa). Do mesmo modo, o desempenho com pronomes só ultrapassa a barreira dos 80% de respostas corretas quando as crianças já atingiram os 90% de correção com *zich*.

Finalmente, há uma outra predição feita pela teoria da competição que recebe algum suporte de resultados disponíveis na literatura. Se é verdade que a distribuição de pronomes só pode ser determinada uma vez que a distribuição de seus competidores foi previamente determinada, espera-se que quanto maior o número de competidores na língua, tanto mais tardiamente sejam os pronomes adquiridos. Isto poderia explicar porque pronomes em holandês são adquiridos mais tardiamente do que em inglês: como vimos, o holandês possui três formas alternativas de 3a. pessoa, pronomes, *zich* e *zichzelf*; o inglês possui apenas duas, pronomes e reflexivos (p. ex., *him* e *himself*).

Em suma, a hipótese de que a distribuição de formas anafóricas é determinada pela competição entre elas é uma inovação teórica cujas conseqüências para a aquisição merecem ser investigadas.

### 3. Formas anafóricas em PB: hipóteses iniciais a serem investigadas

O interesse em investigar a aquisição da anáfora intra-sentencial em PB deriva do fato desta língua ter um sistema de formas que difere de todos os demais já investigados pela literatura. Resumindo grosseiramente o quadro, o PB oferece, em princípio:<sup>18</sup> (i) quatro formas alternativas com verbos transitivos, como em (7a), das quais somente pronomes são excluídos em qualquer instanciação do contexto (ver (18)-(20) abaixo); (ii) quatro formas alternativas com verbos regendo preposição, como em (7b), todas possíveis em alguma instanciação do contexto:

- (7) a. O João<sub>i</sub> já {se}<sub>i</sub> viu {ele mesmo/a si mesmo/\*ele}<sub>i</sub> na TV  
b. O João<sub>i</sub> só fala d{ele/ele mesmo}<sub>i</sub> / de {si/si mesmo}<sub>i</sub>

Das formas anafóricas em (7), *si* e *si mesmo* são, em geral, sentidas como pertencentes ao registro formal e, por isso, possivelmente pouco frequentes no PB coloquial. Uma primeira hipótese que deveria ser investigada, então, é a de que *si* e *si mesmo* são aquisições tardias e resultantes da exposição à norma culta ensinada na escola:

- (8) HIPÓTESE 1 (verifica o status de *si/si mesmo* em PB coloquial):

As formas *si* e *si mesmo* aparecem tardiamente em relação as demais formas anafóricas de 3a. pessoa na produção de fala das crianças, possivelmente a partir da idade escolar (dos 7 anos em diante).

Se a hipótese 1 se confirma, a tarefa do aprendiz de PB se reduz basicamente a determinar a distribuição de três formas com verbos transitivos (*se/ele mesmo/\*ele*), e de duas formas com verbos regendo preposição (*ele/ele mesmo*). Este sistema de formas de 3a. pessoa difere:

- (i) do sistema do inglês, que dispõe de apenas duas formas (*himself/\*him*), sendo que pronomes são sempre excluídos, inclusive com verbos regendo preposição (compare (7b) com a sentença equivalente em inglês, *John speaks only about himself/\*him*);

---

<sup>18</sup> Evidentemente, entende-se por 'PB' aqui a variante a ser investigada, ou seja, a variante falada pelos habitantes da cidade de Porto Alegre, RS. Ver também a seção 5 abaixo.



As crianças têm dificuldades iniciais em interpretar pronomes como variáveis ligadas independentemente da posição em relação ao antecedente (i. é, independentemente do domínio no qual o Princípio B opera).

Esta hipótese pode ser testada em PB comparando-se o desempenho das crianças em pares de frases equivalentes aos em (4a,b) e (5a,b) (ver (24)-(25) abaixo).

Uma outra análise dos ‘efeitos tardios do Princípio B’ cujas predições podem ser testadas em PB é a que Coopmans e seus colaboradores propuseram para as diferenças entre o holandês e o inglês. Como vimos, esta análise se baseia na teoria da ligação de R&R (1993), e sustenta que, além dos mesmos problemas que os aprendizes do inglês têm na aquisição do comando de pronomes, as crianças holandesas teriam um problema adicional: determinar a especificação de Caso dos pronomes holandeses, relevante para a operação da Condição da Cadeia de R&R. Como vimos, para Coopmans e seus colaboradores, os pronomes holandeses são especificados para a distinção entre Caso estrutural versus Caso inerente, mas esta especificação não pode ser determinada por suas formas superficiais (porque as formas que ocorrem em contextos de Caso estrutural ou inerente são homofônicas). Por isso, as crianças holandesas presumem, a princípio, que os pronomes holandeses *não* são especificados para Caso e, portanto, não estão sujeitos a Condição da Cadeia.

Considere agora as conseqüências dessa análise para a aquisição de pronomes em PB. Como em holandês, os pronomes em PB também têm de ser especificados para a distinção entre Caso estrutural e Caso inerente. Isso é particularmente claro em PB, em que formas diferentes de 1a. pessoa do singular (*me* versus *mim*) são exigidas para posições de Caso estrutural (regido por verbos) e Caso inerente oblíquo (regido por preposições). Além disso, embora tanto objetos diretos quanto objetos indiretos sejam regidos por verbos, somente objetos diretos podem ser passivizados:

- (12) a. O João *me*<sub>[+estr]</sub> viu no cinema  
b. *Eu* fui visto no cinema  
c. O João *me*<sub>[-estr]</sub> deu um livro  
d. \**Eu* fui dado um livro

O contraste em (12) sugere que a distinção Caso estrutural vs. Caso inerente é também relevante para os elementos regidos pelo verbo (cf. n. 16 acima). Em suma, há evidência de que o PB, como o holandês e diferentemente do inglês, distingue Caso estrutural de Caso inerente. E, de novo como em holandês, esta distinção não se manifesta superficialmente nas formas de 3a. pessoa, que ocorrem tanto em posições de Caso estrutural (sujeito e objeto verbal) quanto em posições de Caso inerente (objeto de preposição):

- (13) a. *Ele*<sub>[+estr]</sub> chegou tarde  
b. O João viu *ele*<sub>[+estr]</sub> no cinema  
c. O João deu um livro pra *ele*<sub>[-estr]</sub>

Assim, as formas de 3a. pessoa do PB possuem basicamente as mesmas propriedades que as do holandês. Por isso, a análise de Coopmans e seus colaboradores prediz que o padrão de aquisição de pronomes de 3a. pessoa em PB será o mesmo ou similar ao do holandês, hipótese que também pode ser investigada empiricamente:

- (14) HIPÓTESE 4 (verifica a análise de Coopmans *et alii* das diferenças entre a aquisição de pronomes em holandês e em inglês):  
O padrão de aquisição de pronomes em PB é o mesmo encontrado em holandês.

Se a hipótese 4 significa que os resultados específicos obtidos por Coopmans *et alii* e a literatura prévia sobre o holandês devem também ser encontrados em PB, então ela pode ser decomposta em duas sub-hipóteses:

(15) HIPÓTESE 4a:

Em PB as crianças atingem desempenho próximo ao dos adultos por volta dos 10 anos de idade (como em holandês, cf. tabela 3).

(16) HIPÓTESE 4b:

Em PB as crianças têm dificuldades particulares com pronomes em sentenças com predicados inerentemente reflexivos (como em holandês, cf. (6a,b), seção 2.2).

Evidentemente, a confirmação das hipóteses 4a e 4b não apenas corroboraria a análise de Coopmans *et alii* da aquisição de pronomes em inglês e holandês, mas também a versão da teoria da ligação que sustenta esta análise, i. é, a teoria da reflexividade e das cadeias de R&R (1992, 1993, 1995).

Finalmente, consideremos as predições feitas pela teoria da competição, proposta em diversas versões por Burzio (1989, 1992) e pelo proponente deste projeto (cf. Menuzzi 1997). De acordo com esta teoria, formas anafóricas alternativas de uma língua competem entre si no sentido de que a ocorrência de uma forma em um determinado contexto só pode ser determinada pela exclusão das demais competidoras daquele contexto. Como vimos, disso segue que a ordem de aquisição das formas de uma língua depende da hierarquia distribucional das formas: formas com distribuição mais limitada são adquiridas primeiro, formas com distribuição menos limitada depois. Em particular, para determinar a distribuição de pronomes —para determinar os contextos em que são excluídos, i. é, em que deveriam sofrer os ‘efeitos do Princípio B’—, é preciso determinar primeiro a distribuição das formas alternativas nestes contextos.

Uma breve consideração das possibilidades distribucionais das formas anafóricas de 3a. do PB revelaria que a forma *se* é a de distribuição mais restrita, seguida de *ele mesmo*, e finalmente de pronomes, que ocorrem em todos os contextos em que *se* e *ele mesmo* não são exigidos gramaticalmente.<sup>19</sup> Ou seja, a hierarquia distribucional e, portanto, a ordem de aquisição da distribuição destas formas, de acordo com a teoria da competição, deve ser a seguinte: *se* > *ele mesmo* > *ele*. Isso fornece-nos uma primeira hipótese a ser investigada:

(17) HIPÓTESE 5 (verifica a teoria da competição):

A ordem geral de aquisição das formas de 3a. pessoa do PB deve ser:  
*se* > *ele mesmo* > *ele*

Note que, se a distribuição de *ele* em PB é determinada pela competição com *se* e *ele mesmo* —i. é, os contextos em que *ele* é excluído são os contextos em que *se* e/ou *ele mesmo*

---

<sup>19</sup> Note que, embora a forma *se* do PB tenha propriedades similares a *zich* em holandês, e *ele mesmo* possa ter uma função análoga a *zichzelf*, o paralelo distribucional entre as formas brasileiras e holandesas não é completo. Por isso, a hipotética ordem de aquisição *se* > *ele mesmo* > *ele* em PB não se reflete na ordem de aquisição do holandês, que é *zichzelf* > *zich* > *hem*. *Zichzelf* tem a função única de ‘marcador de reflexividade’ e deve, portanto, ocorrer em todos e somente os contextos de reflexividade. *Zich* tem uma distribuição mais ampla e ocorre essencialmente nos contextos locais em *zichzelf* não pode ou não precisa ocorrer. Em PB, por outro lado, *se* tem distribuição limitadíssima —como forma anafórica, só ocorre em posições regidas verbalmente e ligadas pelo sujeito. E, embora *ele mesmo* possa ser usado funcionalmente como *zichzelf* (i. é, na função de ‘marcador de reflexividade’), este uso parece ser derivado do de marcador de ênfase, que é sua função primária. Sendo uma forma enfática, *ele mesmo* tem em PB uma distribuição muito mais ampla que *zichzelf* em holandês. Ver Menuzzi (1997), caps. 3 e 4, para discussão.

são exigidos —, então, a aquisição do comando da forma *ele* pode variar de contexto a contexto na medida em que estes difiram com respeito à disponibilidade das formas *se* e *ele mesmo*. De fato, os contextos variam em PB a este respeito, variação que parece depender de pelo menos dois fatores: a transitividade do verbo e o grau de ‘reflexividade inerente’ do predicado — i. é, o quanto as propriedades semânticas e/ou pragmáticas inerentes do predicado favorecem uma interpretação reflexiva (ver Menuzzi 1997, especialmente o capítulo 3, para uma discussão mais detalhada).

Quando a relação anafórica se dá entre o sujeito e o objeto de um *verbo transitivo* — i. é, um verbo com capacidade de atribuir Caso acusativo —, pronomes são sempre excluídos, e a disponibilidade de *se* ou *ele mesmo* parece depender do grau de reflexividade do predicado, como se vê em (18)-(20) abaixo.<sup>20,21</sup>

(18) Contexto 1 (verbo transitivo, reflexividade provável):

- a. O Paulo<sub>i</sub> {se}<sub>i</sub> vestiu {??ele mesmo/\*ele} rapidamente
- b. O Paulo<sub>i</sub> {se}<sub>i</sub> lavou {??ele mesmo/\*ele} com sabonete

(19) Contexto 2 (verbo transitivo, reflexividade possível):

- a. O João<sub>i</sub> {se}<sub>i</sub> viu {ele mesmo/\*ele} na TV
- b. O João<sub>i</sub> {se}<sub>i</sub> beliscou {ele mesmo/\*ele} (pra ver se ainda estava acordado)

(20) Contexto 3 (verbo transitivo, reflexividade improvável):

- a. O Paulo<sub>i</sub> {??se}<sub>i</sub> arranha {ele mesmo/\*ele} (quando sonha que é um tigre)
- b. O Paulo<sub>i</sub> {??se}<sub>i</sub> machuca {ele mesmo/\*ele} (pra não ter que ir a escola)

Padrões similares aos em (18)-(20) são também encontrados com verbos de complemento preposicionado, com a diferença de que, nestes últimos, pronomes se tornam possíveis.<sup>22</sup>

<sup>20</sup> Compare a aceitabilidade de *ele mesmo* em (18) com sua aceitabilidade em predicados de reflexividade obrigatória, nos quais esta forma é completamente excluído:

- (i) a. O João {se}<sub>i</sub> comportou {\*ele mesmo/\*ele}<sub>i</sub> bem na festa  
cf. \*O João comportou a Maria bem na festa
- b. Quando alguém critica ele, o João<sub>i</sub> {se}<sub>i</sub> fecha {\*ele mesmo/\*ele}<sub>i</sub> como um caramujo  
cf. \*Quando a Maria vai dizer besteira, o João fecha ela como um caramujo

O contraste entre (18) e (i) sugere que, embora desfavorecido em sentenças como (18a,b), *ele mesmo* pode ser ‘resgatado’ e se tornar apropriado sob certas situações discursivas nestas mesmas sentenças. Este e outros contrastes merecem uma investigação mais detalhada, que este projeto propõe como etapa inicial da pesquisa: ver seções 5 e 6 abaixo.

<sup>21</sup> Há também verbos que, em princípio, satisfazem a caracterização intuitiva dos predicados em (19) e, no entanto, desfavorecem *ele mesmo*, cf.

- (i) O João<sub>i</sub> {se}<sub>i</sub> apresentou {??ele mesmo/\*ele}<sub>i</sub> para a Maria

O contraste entre (i) e (19) também merece atenção, cf. n. 20 acima.

<sup>22</sup> Isto acontece, p. ex., com predicados cuja reflexividade é provável — com os quais *ele mesmo* é uma forma desfavorecida em relação aos pronomes, cf.

- (i) a. O João<sub>i</sub> comprou um pirulito pra {ele/??ele mesmo}<sub>i</sub>
- b. O João<sub>i</sub> carregou todos os brinquedos com {ele/??ele mesmo}<sub>i</sub>

Quando a reflexividade é simplesmente possível, *ele* e *ele mesmo* parecem igualmente aceitáveis:

- (ii) a. O Paulo<sub>i</sub> apontou pra {ele/ele mesmo}<sub>i</sub> na foto
- b. O Paulo<sub>i</sub> confia muito n{ele/ele mesmo}<sub>i</sub>

Não é claro quais são as predições que a teoria da competição, bem como as demais análises consideradas neste projeto, fazem com respeito a casos como os em (ii), em que não há distribuição complementar entre as formas anafóricas alternativas. Por esta razão, deixo a discussão destes casos para o desenrolar do projeto.

Assim, há um único destes contextos em que pronomes são excluídos ou fortemente desfavorecidos, e *ele mesmo* praticamente obrigatório: quando o predicado desfavorece a interpretação reflexiva, i. é, quando a reflexividade é improvável:

(21) Contexto 4 (verbo de complemento preposicionado, reflexividade improvável):

- a. O João<sub>i</sub> bateu n{ele mesmo/??ele}<sub>i</sub> com o porrete
- b. O João<sub>i</sub> jogou a bola pra {ele mesmo/??ele}<sub>i</sub>

Em suma, há pelo menos quatro padrões de competição entre formas anafóricas em que pronomes são excluídos em PB. Se é a competição com uma das formas alternativas o que os exclui, então basta que a criança adquira o competidor privilegiado pelo contexto para que o comando dos pronomes comece a melhorar neste mesmo contexto. Vejamos, agora mais detalhadamente, o que esta teoria prediz em relação ao comando dos pronomes nos contextos em (18)-(21).

No contexto 1 em (18), a forma competidora privilegiada é *se*, que, por hipótese, é a primeira cuja distribuição é adquirida (cf. hipótese 5). Isso significa que o comando de *ele* em (18) deve mostrar progresso assim que o comando de *se* for adquirido. O mesmo deve ser verdade para o contexto 2 em (19), se a competição com *se* for suficiente para a exclusão do pronome. No contexto 3 em (20), entretanto, a forma privilegiada é *ele mesmo*; portanto, o comando de pronomes neste contexto só deve mostrar progresso quando a distribuição de *ele mesmo* tiver sido adquirida. Mas, como a aquisição de *ele mesmo* é, por hipótese, mais tardia que a de *se* (cf. hipótese 5), segue que o comando dos pronomes no contexto 3 deverá emergir mais tardiamente em relação aos contextos 1 e 2. Por razões análogas, o comando dos pronomes no contexto 4 deve emergir simultaneamente com o comando dos pronomes no contexto 3 e, portanto, mais tardiamente do que nos contextos 1 e 2.

Ou seja, a teoria da competição, além de e juntamente com a hipótese 5, faz as seguintes predições com respeito a ordem cronológica de aquisição do comando dos pronomes de acordo com os contextos em (18)-(21):

(22) HIPÓTESES 6 (verificam a teoria da competição):

- a. o comando de pronomes deve mostrar progresso nos contextos 1 e 2 a partir da aquisição do comando de *se*;
- b. o comando de pronomes deve mostrar progresso nos contextos 3 e 4 a partir da aquisição do comando de *ele mesmo*;
- c. a ordem relativa de evolução por contexto deve ser: primeiro, emergência do comando de pronomes nos contextos 1 e 2; depois, nos contextos 3 e 4.

Do ponto de vista do presente projeto, há um aspecto crucial a ser observado acerca das hipóteses 1-6 acima listadas: elas não apenas verificam análises alternativas dos ‘efeitos tardios do Princípio B’ baseadas em formulações diferentes da teoria da anáfora intra-sentencial, mas principalmente elas expressam predições diferentes acerca de como a aquisição de pronomes se dará em PB. Isto significa que o PB pode fornecer evidência crucial para decidir entre as teorias em competição neste domínio.

#### 4. Objetivos e justificativa do projeto

O presente projeto tem três objetivos específicos (i)-(iii) e um objetivo geral (iv):

---

Note que as alternativas *si* e *si mesmo* foram ignoradas em (i), (ii) e (21) porque, por hipótese, são de aquisição tardia e não pertencem ao sistema do PB coloquial (cf. hipótese 1 em (8) acima).

- (i) determinar os estágios de aquisição uso anafórico intra-sentencial das formas de 3a. pessoa e, em particular, dos pronomes de 3a. pessoa do português falado no Brasil [PB] (ver a seção precedente para um inventário);
- (ii) verificar se as predições feitas pelas diferentes análises brevemente discutidas na seção 2.2 (i. é, as hipóteses 1-6 da seção precedente) recebem suporte ou são falsificadas pelos dados de aquisição a serem obtidos para o PB;
- (iii) postular e verificar hipóteses alternativas no caso da eventual falsificação de todas as hipóteses iniciais consideradas;
- (iv) contribuir, por meio dos objetivos em (i)-(iii), para a formulação apropriada da teoria geral da anáfora intra-sentencial —i. é, a teoria universal das restrições sobre a anáfora intra-sentencial e da aquisição destas restrições.

Um projeto com tais objetivos justifica-se por pelo menos três razões, duas de ordem puramente científica ((v) e (vi)), e uma terceira de ordem institucional ((vii)):

(v) Tanto quanto o proponente deste projeto sabe, não há literatura publicada sobre a aquisição das restrições à anáfora intra-sentencial no PB, em particular no que se refere à verificação das teorias mais recentes sobre o assunto; uma tal investigação, entretanto, pode contribuir decisivamente para a formulação da teoria geral da anáfora intra-sentencial (i. é, para o objetivo (iv); cf. a seção precedente).

(vi) Evidentemente, a elucidação das propriedades universais da anáfora intra-sentencial é uma das várias pré-condições necessárias à compreensão da faculdade da linguagem e de como ela possibilita a aquisição da linguagem pela espécie humana. I. é, a formulação apropriada da teoria geral da anáfora intra-sentencial é uma das pré-condições necessárias à execução do programa de pesquisa gerativista (cf. n. 2 acima). Por isso, qualquer investigação que contribua para esta teoria —o que inclui, portanto, a pesquisa proposta pelo presente projeto, cf. (v) acima —é do maior interesse para o programa de pesquisa gerativista.

(vii) O Curso de Pós-Graduação em Letras da PUCRS já dispõe de um forte grupo de trabalho na área de aquisição da linguagem, cuja atenção e o trabalho sistemático se desenvolve na investigação da aquisição da fonologia e dos desvios na aquisição da fonologia (cf. os vários projetos em andamento no CEAAL —Centro de Estudos sobre a Aquisição e a Aprendizagem da Linguagem). Embora uma das linhas de pesquisa deste grupo contemple o estudo de aquisição da sintaxe, os trabalhos nesta sub-área têm sido produzidos de modo menos sistemático, em particular no que concerne à abordagem gerativista. Do mesmo modo, o curso carece, no momento, de um programa de pesquisa com a produção sistematicamente dedicada à investigação em teoria sintática. Assim, sucesso na realização do projeto aqui descrito pode ter resultar na ampliação da área de *expertise* do curso para dois novos domínios, sintaxe e aquisição da sintaxe, ambas integradas ao trabalho atualmente desenvolvido em caráter sistemático na área da aquisição da linguagem no curso.

Além disso, pode-se antecipar aqui que os subprojetos vinculados ao presente projeto também terão resultados institucionais benéficos (ver seção 6 abaixo).

## **5. Métodos de Investigação**

As hipóteses iniciais apresentadas na seção 3 serão submetidas à verificação por dois métodos: (i) a análise da distribuição das formas anafóricas de 3a. pessoa em corpus de produção lingüística infantil, controlada por uma análise correspondente de um corpus de fala adulta, cf. seção 5.1 abaixo; (ii) condução e análise de um experimento de julgamento com

aprendizes do PB em várias faixas etárias, controlado por experimento análogo com falantes nativos adultos do PB, cf. seção 5.2 abaixo.

### 5.1. Investigação de Corpora

O corpus de controle a ser investigado é o que foi coletado pelo Projeto VARSUL com falantes adultos de Porto Alegre. Este corpus é constituído por entrevistas com 24 informantes, divididos em 12 perfis sócio-etários (portanto, o corpus contém 2 representantes de cada perfil). Estes perfis são definidos pelas seguintes variáveis: 2 sexos (masculino e feminino); 3 níveis de escolaridade (primário, ginásial e secundário); 4 faixas etárias (20 a 30 anos, 30 a 40, 40 a 50, e acima de 50 anos). Além destas características, os informantes satisfazem alguns critérios adicionais que asseguram a representatividade do corpus: todos falam apenas português, moraram em Porto Alegre por pelo menos 2/3 de sua vida, não moraram fora da cidade por mais de um ano no período de aquisição (dos 2 aos 12 anos), e sua fala não causa estranheza a outros moradores da cidade. Cada informante foi entrevistado 2 vezes, com a entrevista preliminar tendo durado cerca de 15 minutos e a entrevista principal, cerca de 1 hora. Portanto, o corpus é constituído de cerca de 30 horas de gravação, que estão transcritas e disponíveis em forma de texto computadorizado.

A análise deste corpus tem como objetivo determinar se a distribuição das formas anafóricas de 3a. pessoa corresponde à caracterização inicial apresentada na seção 3 acima. Para isso, investigar-se-á a distribuição quantitativa das formas *se*, *ele*, *ele mesmo*, *si* e *si mesmo* em contextos de reflexividade (i. é, em que estas formas tenham como antecedente um NP argumento que pertence ao mesmo predicado sintático e/ou semântico que as contém). Particular atenção será dada aos padrões descritos acima (contextos 1-4 em (18)-(20), respectivamente; ver também notas 19, 20 e 21).

Os corpora de fala infantil a serem investigados são os seguintes:

- (i) faixa etária entre 1;6 e 3;0: corpora longitudinais coletados por Luciene Simões e por Denise Issler (disponíveis no CEAAL/PUCRS);
- (ii) faixa etária entre os 3;0 e os 5;0: corpora transversais do projeto AQUIFONO correspondentes a esta faixa etária (também disponíveis no CEAAL/PUCRS);
- (iii) faixa etária dos 5;0 aos 8;0: corpora longitudinais do projeto DELICRI (UFRGS).

Os corpora acima especificados diferem não apenas na amostra (idade e tipo de amostra), mas principalmente no método de coleta e de transcrição. Isso introduz uma certa heterogeneidade indesejável e torna o trabalho de levantamento um pouco complicado. Entretanto, a investigação destes corpora se faz necessária já que não há presentemente um conjunto de dados coletados e transcritos uniformemente que cubra as faixas de idade de interesse deste projeto.

O corpus coletado por Luciene Simões (ver Simões 1997) acompanha um menino dos 2;4 aos 3;0 anos de idade, nascido em Porto Alegre e cujos pais são porto-alegrenses; o corpus é composto de 10 sessões de gravação com cerca de 45 minutos de duração média (portanto, um total de cerca de 8 horas de gravações). A amostra de fala é naturalística, e contém situações de interação espontânea entre o sujeito e a mãe e/ou a pesquisadora. O corpus está transcrito e digitalizado numa adaptação simplificada do sistema CHAT (do projeto CHILDES, cf. MacWhinney 1995) e, portanto, imediatamente acessível para o levantamento de dados. Nesta mesma situação se encontram os corpora coletados por Denise Issler, cuja amostra de fala é semelhante à coletada por Luciene Simões (i.é, composta de interações espontâneas entre o sujeito e membros de sua família e/ou a pesquisadora; cf. Issler 1998). Dois corpora de Denise estão disponíveis, um de um menino e outro de uma menina também porto-alegrenses de pais porto-alegrenses), ambos acompanhados entre 1;6 e 3;0 de idade, em cerca de 30 entrevistas de 30 minutos de duração aproximadamente (um total de cerca de 30 horas de gravações). Para a faixa dos 1;6 aos 3;0 dispõem-se, portanto, de cerca de 38 horas de gravações já transcritas e imediatamente acessíveis ao levantamento de dados.

Do mesmo modo, os dados para as faixas etárias de 5;0 a 8;0 de idade também está imediatamente acessíveis. Eles se encontram no banco de dados do projeto DELICRI (“Desenvolvimento da Linguagem da Criança na Fase de Letramento”), ligado ao sistema CHILDES (Child Language Data Exchange System). Este banco é composto de dados de crianças entre 5 e 9 anos de idade, todas aprendizes do dialeto de Porto Alegre, e é composto por diferentes tipos de amostras da sua produção lingüística: (a) diálogo entre o investigador e a criança; (b) narrativas livres orais e escritas; (c) narrativas a partir de quadros seqüenciais; (d) diálogo livre entre duas crianças; (e) o recontar de seqüências narrativas (os dois últimos tipos de amostras foram obtidos para somente 20% das crianças). Os dados foram coletados tanto transversal quanto longitudinalmente. Este projeto pretende investigar apenas os dados coletados longitudinalmente, porque permitem traçar melhor o quadro evolutivo individual. Estes dados foram obtidos com 10 crianças, acompanhadas dos 5 aos 9 anos, por meio de 5 coletas anuais. Todos os dados que compõem o corpus estão transcritos de acordo com o sistema CHILDES (cf. MacWhinney 1991), adaptado para o PB (cf. Guimarães & Lamprecht 1995); também estão, portanto, imediatamente acessíveis para levantamento.

Finalmente, deve-se dizer que, presentemente, não há nenhum corpus de fala de crianças nascidas em Porto Alegre na faixa entre 3;0 e 5;0 que esteja transcrito e digitalizado e, portanto, em formato apropriado para levantamento de dados. Há dados coletados que servem para os fins deste projeto; este é o caso particular das gravações feitas para a constituição do banco de dados AQUIFONO (“Aquisição da Fonologia do Português do Brasil, PUCRS). Estas contêm entrevistas de cerca de 30 minutos com 310 sujeitos entre os 2 e os 7 anos de idade (divididos por sexo e por faixas etárias de dois meses); no que diz respeito ao período entre 3;0 e 5;0, são 120 sujeitos (60 meninos e 60 meninas) gravados num total cerca de 60 horas. As entrevistas utilizam como método de elicitación a ‘homeação espontânea’ e a ‘narração e fala espontâneas’, em geral motivadas pela apresentação, por um experimentador, de objetos, figuras e bonecos à criança (cf. Yavas, Hernandorena e Lamprecht 1991). As entrevistas se revestem, em geral, do caráter de brincadeira e, na maioria dos casos, resultam em situação informal de interação. Por isso, oferecem uma amostra da fala espontânea da criança e podem servir para a investigação de seus aspectos não-fonológicos. Entretanto, até o momento estas entrevistas só receberam transcrição apropriada para a investigação da fonologia da criança (foram transcritas apenas palavras isoladas que atestam diferentes padrões fonológicos previamente determinados). Assim, para o levantamento dos dados correspondentes a faixa etária entre 3;0 e 5;0, será necessário executar uma transcrição adequada dos corpora correspondentes, tarefa que se pretende atribuir a bolsistas de iniciação científica (ver seção 6 abaixo, sobre os subprojetos vinculados ao presente).

Quanto ao levantamento e análise de dados propriamente ditos, procurar-se-á determinar nos corpora de produção lingüística infantil, tal como na análise do corpus adulto, a distribuição quantitativa das formas *se*, *ele*, *ele mesmo* nos contextos de reflexividade descritos na seção 3. Esta análise tem como objetivo principal determinar: (i) a idade em que o uso apropriado das formas *se* e *ele mesmo* (bem como, secundariamente, de *si* e *si mesmo*) emerge; (ii) a *evolução* no comando de pronomes de 3a. pessoa na produção de fala infantil (presumindo que, nos estágios iniciais de aquisição, pronomes ocorram em contextos de reflexividade dos quais são excluídos na fala adulta, cf. discussão da seção precedente).

## 5.2. Investigação Experimental

O tipo de experimento que se adotará para verificar o conhecimento lingüístico dos aprendizes de PB é chamado pela literatura em aquisição de ‘truth value judgment task’ (literalmente, ‘tarefa de julgamento do valor de verdade (de uma sentença-estímulo)’). Este tipo de experimento foi desenvolvido por Crain & McKee (1985) e tem sido adotado como padrão pela literatura por ter se revelado um método eficaz de obtenção de julgamentos lingüísticos infantis (ver Koster 1992, McDaniel, McKee & Kearns 1996 e Crain & Thornton 1998 para discussão e comparação deste com outros métodos experimentais de investigação do conhecimento sintático das crianças).

A idéia básica por trás de um ‘truth value judgment task’ é expor as crianças a um par sentença/significado e obter seu julgamento acerca da correspondência entre ambos (i. é, acerca da verdade ou falsidade da sentença) de uma maneira que as atraia. Por isso, estes experimentos são concebidos como jogos e tem o seguinte formato geral: (i) um experimentador usa brinquedos (ou mostra-se um filme, ou fotos, etc.) em que se encena um evento correspondente a uma interpretação da sentença-estímulo; (ii) outro experimentador manipula um boneco (ou finge ser um alienígena, ou um estrangeiro, etc.) que assiste ao evento com a criança e que tem como tarefa aprender a língua da criança (no caso deste projeto, o PB); (iii) o boneco (ou o ‘alienígena’, etc.) descreve, então, na língua que ‘está aprendendo’, o que aconteceu; (iv) a criança deve, então, julgar se o que o boneco disse é uma descrição correta ou incorreta do evento, premiando-o ou punindo-o (por exemplo, dando-lhe um beijo/um caramelo, etc., ou xingando-lhe/passando-lhe um castigo, etc.).

Para exemplificar como este tipo de tarefa permite avaliar o conhecimento que a criança tem das restrições sobre o uso de pronomes, considere a seguinte sentença:

(23) [s O Pimpão bateu nele] (com um porrete)

Como vimos na seção anterior, o pronome *ele* não pode se referir a *João* em um sentença como (23) em PB (porque predicados como *bater em* requerem um elemento que sirva de marcador de reflexividade, como *ele mesmo*: ver Menuzzi 1997, cap. 3). Ou seja, a sentença (23) só gramatical em PB quando *ele* se refere a um antecedente (ou indivíduo) que não está (representado lingüisticamente) dentro dela.

Para saber se a criança admite a interpretação gramatical (do ponto de vista da gramática adulta) de (23), poder-se-ia usar a seguinte cena: (i) Dois bonecos, Hermes e Pimpão, discutem para saber quem é o dono de um carrinho de brinquedo (um diz: ‘Ele é meu.’, o outro responde: ‘Não, ele é meu!’, etc.). (ii) Hermes pára de discutir e pega, simplesmente, o carrinho. (iii) Pimpão pega, então, um porrete e bate em Hermes. (iv) Hermes grita de dor: ‘Ai, ai, ai!’. (v) Jonas, o boneco que assiste à cena junto com a criança, então diz: ‘O Hermes tá gritando porque [s o Pimpão bateu nele]’. A resposta correta para este par sentença/significado é, evidentemente, ‘sim’, porque a sentença descreve apropriadamente o evento. A criança que faz este julgamento (corretamente) dá um caramelo para o Jonas; a criança que julga a sentença de Jonas inapropriada (incorretamente), dá-lhe um puxão de orelha.

Para saber se, por outro lado, a criança admite a interpretação *agramatical* (do ponto de vista da gramática adulta) de (23), poder-se-ia usar uma cena semelhante à acima descrita: (i) De novo, Hermes e Pimpão, discutem para saber quem é o dono do carrinho. (ii) Hermes pára de discutir e pega o carrinho. (iii) Pimpão pega o porrete e *tenta* bater em Hermes, mas *acerta* em si mesmo. (iv) Hermes dá gargalhadas: ‘Ha, ha, ha!’. (v) Jonas então diz: ‘O Hermes tá rindo porque [s o Pimpão bateu nele]’. A resposta correta para este par sentença/significado é ‘não’, porque a sentença descreve inapropriadamente o evento (ela sugere que Hermes está

rindo porque levou uma porretada!). A criança que faz este julgamento (corretamente) dá um puxão de orelha em Jonas, etc.

Pretende-se aplicar um experimento com o formato acima descrito a crianças nascidas e residentes desde o nascimento em Porto Alegre, e cujos pais também sejam falantes nativos do dialeto da cidade. Três faixas etárias serão inicialmente investigadas: entre 3 e 4 anos, entre 5 e 6, e entre 7 e 8 anos. Se a hipótese de que a aquisição do comando de pronomes em PB é tão tardia quanto em holandês se confirmar (cf. hipóteses 4-4a em (14)-(15) acima), então o experimento deverá ser estendido a crianças na faixa dos 9 aos 10 anos. Cada grupo deverá ter em torno de 20 indivíduos, a fim de se ter uma amostra apropriada para análise estatística. Pretende-se procurar indivíduos que se qualifiquem para o experimento em creches e escolas de bairros residenciais de classe média (portanto, pretende-se evitar maiores variações devidas, p. ex., à escolaridade dos pais). Dadas as dimensões da tarefa, pretende-se que ela seja executada por um grupo de bolsistas de iniciação científica especialmente treinados (cf. seção 6, sobre os subprojetos vinculados ao presente).

Quanto às formas e contextos a serem testados, em princípio serão alvo do experimento aqueles discutidos na seção anterior (ver contextos 1-4 em (18)-(20), respectivamente; ver também notas 19, 20 e 21). Além disso, a verificação experimental da hipótese 3 (em (11) acima) requer que estruturas como (24)-(25) sejam testadas (ver a discussão de (5) acima):

- (24) a. O menino já tinha batido nele (com um porrete)  
b. Tudo que era menino já tinha batido nele (com um porrete)
- (25) a. O menino disse pr'o Paulo que a Maria já tinha beijado ele  
b. Tudo que era menino tinha dito pr'o Paulo que a Maria já tinha beijado ele

A escolha do 'setting' experimental, das sentenças-estímulo, e demais decisões concernentes ao formato específico do experimento, seguirão procedimentos padronizados, quase que consensualmente adotados pela literatura sobre aquisição da anáfora (cf., p. ex., Crain & McKee 1985, Koster 1992, McDaniel, McKee & Kearns 1996 e Crain & Thornton 1998).

Para o controle do experimento a ser conduzido com crianças, pretende-se aplicar um teste de julgamento a (cerca de 20) falantes nativos adultos do dialeto de Porto Alegre. As estruturas a serem testadas serão as mesmas, e o formato do teste também seguirá procedimentos padronizados (cf. Schütze 1996).

## **6. Subprojetos Vinculados**

Como foi apontado na seção precedente, o andamento do presente projeto depende de pelo menos três tarefas preliminares de envergadura: uma delas é executar uma transcrição apropriada dos corpora do projeto AQUIFONO correspondentes à faixa etária dos 3;0 aos 5;0 (entre 40 e 60 horas de gravação); outra é a de levantar e analisar quantitativamente os dados dos vários outros corpora transcritos na seção 5.1 (possivelmente envolvendo transcrições de mais de 120 horas de gravação); finalmente, é preciso preparar e aplicar o experimento descrito na seção precedente a cerca de 100 indivíduos e, depois, codificar e analisar quantitativamente os resultados. Tais tarefas estão, evidentemente, aquém daquela executável por um indivíduo no espaço de tempo relativamente curto previsto para a duração do presente projeto (32 meses). Por isso, pretende-se que sejam executadas por bolsistas de iniciação científica treinados para esse fim.

Assim, acompanhando o presente projeto, o pesquisador responsável pretende submeter, ao longo do período de duração deste, três subprojetos auxiliares:

(i) ‘Transcrição de Corpora do AQUIFONO’ [abreviadamente, AQUIFONO]: solicitação de dois bolsistas de iniciação científica para a execução de transcrição em formato CHAT (do projeto CHILDES, cf. MacWhinney 1995) dos corpora do projeto AQUIFONO selecionados, e para levantamento e análise dos dados nestes corpora (período: março de 2000 a março de 2001);

(ii) ‘Levantamento de dados no DELICRI’ [abreviadamente, DELICRI]: solicitação de dois bolsistas de iniciação científica para levantamento e análise dos dados dos corpora do projeto DELICRI selecionados; os bolsistas deste subprojeto podem também colaborar com a preparação do experimento descrito na seção 5.2 acima (ver (iii) abaixo), ou com a transcrição em formato CHAT de outros corpora do projeto AQUIFONO (período: agosto de 2001 a agosto de 2002);

(iii) ‘Realização do Experimento de Julgamento’ [abreviadamente, EXPERIMENTO]: solicitação de dois bolsistas de iniciação científica para preparação e aplicação do experimento descrito na seção 5.2, e para levantamento e análise dos dados coletados (período: março de 2001 a março de 2002);

Tais subprojetos atenderão não somente a necessidades o presente projeto, mas também a objetivos institucionais. De um lado, os subprojetos AQUIFONO e DELICRI colaborarão com o esforço de adaptação de corpora disponíveis no CEAAL ao sistema de banco de dados do projeto CHILDES, o que permitirá a eventual integração dos primeiros a este (um plano de longo prazo do CEAAL/PUCRS). Além disso, os três subprojetos estarão introduzindo alunos de graduação a problemas teóricos enfrentados pela pesquisa em aquisição da linguagem, bem como estarão treinando-os em técnicas de transcrição de dados fala e uso de recursos computacionais apropriados à análise destes dados. Com isso, espera-se estar colaborando com a constituição e a qualificação das próximas gerações de pós-graduandos em lingüística da PUCRS.

## 7.Cronograma

Atividades	Ago	Dez	Abr	Ago	Dez	Abr	Ago	Dez
	99 a	00 a	00 a	00 a	00 a	01 a	01 a	01 a
	Nov	Mar	Jul	Nov	Mar	Jul	Nov	Mar
	99	00	00	00	01	01	01	02
1. revisão da literatura	x	x	x	x	x	x	x	x
2. levantamento de dados corpora Simões e Issler	x	x						
3. AQUIFONO (treinamento bolsistas, transcrição, levantamento de dados, etc.)			x	x	x	x		

4. DELICRI (idem)	x	x	x	x
5. EXPERIMENTO (trei-namento bolsistas, prepa- ração experimento, aplica-ção, etc.)			x	x
6. Análise qualitativa e teórica dos resultados; pre- paração de relatório final.			x	x

## 8.Referências bibliográficas

- Avrutin, Sergei & Wexler, Kenneth (1992) 'Development of Principle B in Russian: Coindexation at LF & Coreference'. *Language Acquisition* 2:4, 259-306.
- Borer, Hagit & Wexler, Kenneth (1987) 'The Maturation of Syntax'. Em Roeper & Williams, eds., (1987), pp. 123-172.
- Burzio, Luigi (1989) 'On the Non-Existence of Disjoint Reference Principles'. *Rivista di Grammatica Generativa* 14.3-27.
- Burzio, Luigi (1992) 'The Role of the Antecedent in Anaphoric Relations'. *Paper* não-publicado, Universidade Johns Hopkins, Baltimore (USA).
- Chien, Y.-C & Wexler, K. (1990) 'Children' s Knowledge of Locality Conditions in Binding as Evidence for the Modularity of Syntax & Pragmatics'. *Language Acquisition* 1, 225-295.
- Crain, Stephen (1991) 'Language Acquisition in the Absence of Experience'. *Behavioral and Brain Sciences* 14, 597-650.
- Crain, Stephen & McKee, Cecile (1985) 'Acquisition of Structural Restrictions on Anaphora'. *NELS* 16, 94-110.
- Crain, Stephen & Thornton, Rosalind (1998) *Investigations in Universal Grammar: a Guide to Experiments on the Acquisition of Syntax and Semantics*. A sair, MIT Press, Cambridge (USA).
- Chomsky, Noam (1981) *Lectures on Government & Binding*. Foris, Dordrecht (Holanda).
- Chomsky, Noam. (1986) *Knowledge of Language*. Praeger, Nova Iorque.
- Grimshaw, Jane & Rosen, Sara Thomas (1990) 'Knowledge and Obedience: the Developmental Status of the Binding Theory'. *Linguistic Inquiry* 21:2, 187-222.
- Grodzinsky, Y. & Reinhart, Tanya (1993) 'The Innateness of Binding & Coreference'. *Linguistic Inquiry* 24:1, 69-102.
- Grodzinsky, Y.; Wexler, K.; Chien, Y.-C. & Marakovitz, S. (1989) 'The Breakdown of Binding'. *Paper* apresentado à Academia de Afasia, Santa Fé (USA).
- Guimarães, Ana Maria de Mattos & Lamprecht, Regina Ritter (1995) 'The Use of the CHILDES Database for Brazilian Portuguese'. Em I. H. Faria & M. J. Freitas, eds., *Studies on the Acquisition of Portuguese*, pp. 207-214, Colibri/Associação Portuguesa de Linguística, Lisboa.

- Hornstein, Norbert & Lightfoot, David (1981) 'Introduction'. In Hornstein & Lightfoot, eds., *Explanation in Linguistics: the Logical Problem of Language Acquisition*, Longman, Londres.
- Jakubowicz, C. (1991) 'Binding Principles and Acquisition Facts Revisited'. NELS 21.
- Kaufman, D. (1988) 'Grammatical & Cognitive Interactions in the Study of Children's Knowledge of Binding Theory and Reference Relations'. Dissertação de doutorado, Universidade de Temple, Filadélfia.
- Koster, Charlotte (1993) 'Errors in Anaphora Acquisition'. Dissertação de doutorado, OTS/Universidade de Utrecht, Utrecht (Holanda).
- Lasnik, Howard (1976) 'Remarks on Coreference' *Linguistic Analysis* 2.1-22.
- Lasnik, Howard (1989) 'A Selective History of Modern Binding Theory'. Em Lasnik, *Essays on Anaphora*, pp.1-36. Kluwer, Dordrecht (Holanda). (Série: Studies in Natural Language & Linguistic Theory.)
- Lasnik, Howard & Uriagereka, Juan (1988) *A Course in GB Syntax*. MIT Press, Cambridge (USA).
- Lightfoot, David (1982) *The Language Lottery: toward a Biology of Grammars*. MIT Press, Cambridge (USA).
- Lombardi, L. & Sarma, J. (1989) 'Against the Bound Variable Hypothesis of the Acquisition of Principle B'. *Paper não-publicado*, Universidade de Massachusetts (Amherst, USA) & Universidade de Connecticut (Storrs, USA).
- Lust, B., ed. (1986) *Studies in the Acquisition of Anaphora*, vol. I. Reidel, Dordrecht (Holanda).
- MacWhinney, B. (1991) *The CHILDES Project: Tools for Analysing Talk*. Erlbaum, Hillsdale (USA).
- McDaniel, Dana; McKee, Cecile & Cairns, Helen Smith, eds. (1996) *Methods for Assessing Children's Syntax*. MIT Press, Cambridge (USA).
- McKee, Cecile (1992) 'A Comparison of Pronouns and Anaphors in Italian and English Acquisition'. *Language Acquisition* 2:1, 21-54.
- Menuzzi, Sergio (1997) 'Topics in Binding Theory: Constraint Interaction, Chains, Indexing and Reflexivity, with Particular Reference to Brazilian Portuguese'. Dissertação de doutorado, Curso de Pós-Graduação em Letras da PUCRS, Porto Alegre.
- Philip, William & Coopmans, Peter (1996) 'The Role of Referentiality in the Acquisition of Pronominal Anaphora'. NELS 26, 241-255.
- Reinhart, Tanya (1983) *Anaphora & Semantic Interpretation*. Croom Helm, Kent (Inglaterra). (Série: Croom Helm Linguistic series.)
- Reinhart, Tanya (1986) 'Center & Periphery in the Grammar of Anaphora'. Em B. Lust, ed., *Studies in the Acquisition of Anaphora*, Vol. I, pp.123-150. Foris, Dordrecht (Holanda).
- Reinhart, Tanya & Reuland, Eric (1993) 'Reflexivity'. *Linguistic Inquiry* 24:4.657-720.
- Reuland, Eric & Reinhart, Tanya (1992) 'Binding Conditions & Chains'. *Proceedings of the 10th WCCFL*, pp.399-415.
- Reuland, Eric & Reinhart, Tanya (1995) 'Pronouns, Anaphors & Case'. *Paper não-publicado*, Universidade de Utrecht (Holanda) e Universidade de Tel Aviv (Israel).
- Roeper, Thomas & Williams, Edwin, eds. (1987) *Parameter Setting*. Reidel, Dordrecht (Holanda).
- Schütze, Carson (1996) *The Empirical Base of Linguistics: Grammaticality, Judgments, and Linguistic Methodology*. Chicago University Press, Chicago.
- Sigurjónsdóttir, Sigrídur & Hyams, Nina (1992) 'Reflexivization and Logophoricity: Evidence from the Acquisition of Icelandic'. *Language Acquisition* 2:4.

- Sigurjónsdóttir, S. & Coopmans, P. (1996) 'The Acquisition of Anaphoric Relations in Dutch'. *Amsterdam Series on Child Language Development*, vol. 5, Universidade de Amsterdam.
- Solan, L. (1987) 'Parameter Setting and the Development of Pronouns and Reflexives'. In Roeper & Williams, eds., (1987), pp. 189-210.
- Thornton, R. (1990) 'Adventures in Long-Distance Movement: the Acquisition of Complex WH-Questions'. Dissertação de doutorado, Universidade de Connecticut, Storrs (USA).