



Universidade Federal de Juiz de Fora
Faculdade de Letras
Programa de Pós-Graduação em Linguística

MILENE CRISTINE DE CASTRO TEIXEIRA

**A contribuição das informações distribucionais de DNA nas etapas iniciais da
aquisição lexical**

**Juiz de Fora
2012**

Milene Cristine de Castro Teixeira

A contribuição das informações distribucionais de DNA nas etapas iniciais da aquisição lexical

Dissertação de Mestrado apresentada ao curso de Mestrado do Programa de Pós-Graduação em Linguística da Faculdade de Letras da Universidade Federal de Juiz de Fora, como parte dos requisitos para a obtenção do título de Mestre em Linguística.

Orientadora: Prof.^a Dr.^a Maria Cristina Name.

JUIZ DE FORA
2012

Teixeira, Milene Cristine de Castro.

A contribuição das informações distribucionais de DNA nas etapas iniciais da aquisição lexical/ Milene Cristine de Castro
Teixeira – 2012.

74 f. : il.

Dissertação (Mestrado em Linguística)- Universidade Federal de Juiz de Fora, Juiz de Fora, 2012.

1. Linguística. I. Título.

Milene Cristine de Castro Teixeira

A contribuição das informações distribucionais de DNA nas etapas iniciais da aquisição lexical

Dissertação de Mestrado submetida ao Programa de Pós-Graduação em Linguística da Faculdade de Letras da Universidade Federal de Juiz de Fora, como requisito parcial à obtenção do título de Mestre em Linguística.

Aprovada pela banca examinadora abaixo assinada.

Professora Doutora Maria Cristina Name (Orientadora)
Universidade Federal de Juiz de Fora

Professora Doutora Luciana Teixeira
Universidade Federal de Juiz de Fora

Professora Doutora Erika Maria Parlato de Oliveira
Universidade Federal de Minas Gerais

Juiz de Fora
2012

DEDICATÓRIA

*Aos meus pais, à minha irmã, ao meu noivo Rafael e a todos os bebês que
contribuíram para a realização desta dissertação.*

AGRADECIMENTOS

A Deus agradeço por mais esta vitória. Como tudo em minha vida, nada seria possível sem Ele, sem suas bênçãos, sem seu olhar sobre mim...

Aos meus pais, Mirian e Walter, e à minha irmã, Melissa, por constituírem essa família maravilhosa.

Ao meu (agora) esposo, Rafael, meu amor, por estar sempre ao meu lado e por me motivar sempre a continuar seguindo em frente. Obrigada por me fazer acreditar que tudo daria certo.

À minha Orientadora, Cristina Name, que, desde a Iniciação Científica, me guiou com muita responsabilidade, seriedade e competência nesta jornada acadêmica.

À Professora Luciana Teixeira, por ser presença indispensável nesta minha caminhada, pela sua dedicação, comprometimento e carinho.

À professora Erika Maria Parlato de Oliveira, por sua valiosa contribuição.

As “amiguitias” do mestrado Ana Paula, Vanessa e Carol por tornarem muito mais divertido esses dois anos e a todos os pesquisadores do NEALP – esse grupo que se torna cada vez mais sólido e competente. Em especial, à Flávia Faria pela super ajuda nos experimentos e pela companhia sempre agradável.

À Carol, bolsista do laboratório.

À Rosângela, secretária do Programa de Pós Graduação em Linguística, pela sua simpatia e presteza.

E em especial, aos 22 bebês lindos que participaram da atividade sem os quais este trabalho não teria sido possível. Que Deus abençoe sempre vocês e que vocês cresçam com muita saúde e paz.

Aos responsáveis dos bebês, por terem confiado na seriedade do meu trabalho e por terem depreendido um pouquinho do tempo de vocês contribuindo para minha pesquisa.

À CAPES, pelo financiamento da bolsa de Mestrado.

A todos que, de alguma forma, me ajudaram a concluir este trabalho, muito OBRIGADA!

RESUMO

Este estudo investiga as capacidades de abstração e generalização de padrões linguísticos nas etapas iniciais do processo de aquisição de língua materna, privilegiando relações de dependências não-adjacentes (DNA) e explora a relação entre itens funcionais e marcadores de palavras de Nome (*word markers*). Diferentemente de trabalhos anteriores, que usaram línguas totalmente artificiais – cujas propriedades fonológicas, morfológicas e prosódicas eram muito simples –, esta pesquisa usa uma língua pseudonatural, preservando tais propriedades. A conciliação entre um modelo de processamento voltado para a aquisição da linguagem (*Bootstrapping* Fonológico: Morgan & Demuth, 1996; Christophe *et al.*, 1997) e um modelo de língua que considere níveis de interface entre o sistema linguístico e outros sistemas – perceptuais e cognitivos (Programa Minimalista: Chomsky 1995; 1999 e obras posteriores) - permite compreender a passagem de um nível de representação fonético/fonológico para um nível de representação formal. Assumindo-se ainda que mecanismos estatísticos e de abstração e generalização são recursos precocemente disponíveis e explorados por bebês no processo de categorização de elementos lexicais (Marcus *et al.*, 1999; Gómez, 2002; Newport & Aslin, 2004), a conciliação entre o modelo de língua e o modelo de processamento pretende explicitar o modo como a criança extrai do *continuum* sonoro, além de pistas distribucionais e prosódicas, propriedades dos traços formais dos elementos funcionais nos quais o Sistema Computacional age resultando na derivação linguística. Bebês, com média de idade de 11 meses, foram expostos durante um curto período de tempo a uma língua diferente de sua língua materna e, posteriormente, expostos a novos DPs congruentes ou não aos apresentados anteriormente. Os resultados indicam que os bebês reagiram à diferença de padrões de DNA em função daqueles previamente apresentados, escutando por mais tempo padrões não familiarizados, sugerindo que a partir de uma curta exposição a uma língua pseudonatural – diferente do padrão do PB – bebês, aos 11 meses, são capazes de abstrair e generalizar o padrão dessa “língua”. Tais resultados sugerem que mecanismos de abstração e generalização podem ser recursos utilizados por crianças no processo inicial de aquisição lexical.

Palavras - Chave: Aquisição Lexical. Dependências não-adjacentes. Categorização. Determinantes.

ABSTRACT

This study investigates the capacities of abstraction and generalization of linguistic patterns in the early stages of the acquisition of mother tongue, emphasizing the relationship of nonadjacent dependencies (DNA) and exploring the relationship between functional items and word markers of nouns. Unlike previous studies, which used artificial languages - whose phonological, morphological and prosodical properties were very simple - this research uses a pseudonatural language, preserving such properties. The conciliation between a psycholinguistic model aimed at the acquisition of language (Phonological Bootstrapping: Morgan & Demuth, 1996; Christophe *et al.*, 1997) and a linguistic theory based on interface levels (Minimalist Program, Chomsky, 1995; 1999) allows us to understand the passage from one level of phonetic/phonological representation to a level of formal representation. Assuming also that statistical mechanisms of abstraction and generalization are resources available to infants and exploited early by them in the process of categorization of lexical (Marcus *et al.*, 1999; Gómez, 2002; Newport & Aslin, 2004), the conciliation between the two theories aims to clarify how the child extracts from the sonorous continuum as well as distributional and prosodic clues, properties of the formal traces of the functional elements in which the Computational System acts resulting linguistic derivations. Babies at 11 months were exposed for a short period of time to a language other than their mother tongue and later exposed to new DPs congruent or not to those previously presented. The results indicate that babies reacted to different patterns of DNA according to those previously presented, listening longer to unfamiliar patterns, suggesting that from a short exposure to a pseudonatural language, babies at 11 months are able to abstract and generalize the pattern of this language. These results suggest that mechanisms of abstraction and generalization can be used by children in the initial process of language acquisition.

Keywords: Lexical Acquisition. Nonadjacent dependencies. Categorization. Determiners.

SUMÁRIO

1 INTRODUÇÃO	11
2 OS ELEMENTOS DO LÉXICO E SUAS RELAÇÕES	16
2.1 ELEMENTOS LEXICAIS	17
2.1.1 <i>As categorias lexicais segundo Chomsky (1970)</i>	17
2.2 ELEMENTOS FUNCIONAIS	21
2.2.1 <i>Dependências não-adjacentes (DNA)</i>	23
2.3 CONSIDERAÇÕES SOBRE GÊNERO EM NOMES E DETERMINANTES	29
3 FUNDAMENTAÇÃO TEÓRICA	34
3.1 BOOTSTRAPPING FONOLÓGICO	34
3.2 PROGRAMA MINIMALISTA	39
3.3 A PROPOSTA DE INTEGRAÇÃO ENTRE O MODELO LINGÜÍSTICO (PROGRAMA MINIMALISTA) E O MODELO PSICOLINGÜÍSTICO (BOOTSTRAPPING FONOLÓGICO)	44
4 ATIVIDADE EXPERIMENTAL	46
4.1 TÉCNICA EXPERIMENTAL – OLHAR PREFERENCIAL	46
4.2 MÉTODO	48
4.2.1 <i>Participantes</i>	48
4.2.2 <i>Materiais utilizados</i>	49
4.2.3 <i>O ambiente</i>	49
4.2.4 <i>O procedimento:</i>	49
4.2.5 <i>Estímulos</i>	53
4.3 O EXPERIMENTO	54
4.4 RESULTADOS E DISCUSSÃO	59
5 CONSIDERAÇÕES FINAIS	62
REFERÊNCIAS	66
ANEXOS	69

1 INTRODUÇÃO

Esta dissertação¹ investiga as etapas iniciais do processo de aquisição lexical, enfocando capacidades de abstração e generalização de padrões linguísticos por bebês durante o primeiro ano de vida e focaliza, mais especificamente, dependências não-adjacentes (doravante, DNA) nas relações que se estabelecem entre determinantes (D) e nomes (N) no âmbito do Sintagma Determinante (doravante, DP, *Determiner Phrase*).

Assume-se, neste trabalho, a proposta inatista (Chomsky, 1965) que defende uma predisposição biológica para a aquisição da língua. Sabe-se que a pesquisa em aquisição de língua materna e, especificamente, em aquisição lexical, se depara com o problema de identificação de palavras no fluxo de fala exposto cotidianamente à criança. Dessa forma, mesmo assumindo tal postura inatista, faz-se necessário entender como a criança mapeia elementos lexicais a categorias sintáticas como N(ome), V(verbo) etc., a partir da exposição a um fluxo contínuo de fala. Em outras palavras, como o bebê segmenta os enunciados que ouve e ainda como relaciona esses elementos a entidades e/ou conceitos e os agrupa em classes (conceituais/gramaticais).

Toda e qualquer língua natural a ser adquirida possui padrões que as definem como determinada língua e, ao mesmo tempo, as diferenciam umas das outras. Assim, mesmo havendo alguma variação, as categorias sintáticas, por exemplo, se apresentam regularmente distribuídas nas sentenças. A alta frequência de determinados elementos (elementos de categorias funcionais), padrões de co-ocorrência entre os elementos da sentença e relações de concordância manifestas através de variações morfofonológicas são regularidades presentes no estímulo linguístico disponível para a criança, que podem facilitar a identificação de

¹Esta pesquisa se insere em um estudo mais amplo desenvolvido no NEALP (Núcleo de Estudos em Aquisição da Linguagem e Psicolinguística, UFJF) e se vincula ao projeto *Etapas Iniciais da Aquisição lexical*(Processo CNPQ n°401510/2010-7).

elementos lexicais de uma língua. A fim de explicar como tais pistas recorrentes no *continuum* sonoro podem facilitar a identificação de elementos funcionais, adota-se a possibilidade de uma habilidade computacional. Segundo tal abordagem, mecanismos estatísticos e de abstração seriam recursos disponibilizados precocemente no processo de aquisição da língua materna.

Neste trabalho, adota-se como perspectiva teórica a conciliação entre um modelo de processamento – *Bootstrapping* Fonológico (MORGAN & DEMUTH, 1996; CHRISTOPHE et al., 1997) – e um modelo de língua – Programa Minimalista (CHOMSKY 1995 e desenvolvimentos posteriores) – pressupondo assim que a análise fonológica do sinal de fala realizada pelo bebê forneceria informações sobre a estrutura sintática de sua língua e permitiria o início da aquisição lexical e sintática. Partindo dessa conciliação, assume-se que a distinção entre elementos funcionais e lexicais, a partir da sensibilidade a padrões prosódicos e fonotáticos no primeiro ano de vida, dá origem a um léxico mínimo, constituído de elementos com traços subespecificados, que, por sua vez, vão desencadear a aquisição da sintaxe (CORRÊA, 2006). Assume-se ainda que mecanismos estatísticos e de abstração e generalização são recursos explorados e disponibilizados precocemente pelos bebês permitindo, mesmo a partir de um conjunto restrito de enunciados, a (sub) categorização de elementos lexicais.

De maneira geral, como veremos durante este trabalho, diversos autores têm explorado as capacidades de abstração e generalização de padrões linguísticos tanto por adultos quanto por crianças. Entretanto, em sua grande maioria, exploram tais padrões em línguas totalmente artificiais (Marcus et al., 1999; Gómez, 2002; Newport & Aslin, 2004), sem levar em conta propriedades morfofonológicas e prosódicas, ou, quando trabalham com línguas naturais, investigam tais capacidades em crianças mais velhas (18 meses), considerando sobretudo o reconhecimento de padrões de sua língua materna, e não propriamente habilidades de abstração e generalização (SANTELMANN & JUSCZYK, 1998). Em língua pseudonatural e trabalhando com crianças mais jovens, NAME, SHI & KOULAGUINA (2011) sugerem que bebês canadenses de 11 meses, expostos durante um curto período de tempo (2 minutos) a uma língua pseudonatural semelhante ao português do Brasil (PB) – portanto, diferente de sua língua materna –, são capazes de reconhecer padrões de

relações entre D e N e, ainda, são capazes de abstraí-los de modo a identificá-los em novos enunciados.

A hipótese de trabalho que norteia este estudo é a de que bebês durante o primeiro ano de vida são providos da capacidade de tratamento do estímulo linguístico que lhe permite identificar, rapidamente, combinações recorrentes – particularmente DNA –, abstrair seus padrões e generalizá-los em novos enunciados.

A partir do trabalho usando língua pseudonatural com bebês canadenses, buscamos verificar robustez desses resultados, replicando a atividade experimental com bebês brasileiros. Mais ainda, avaliaremos se o contato prévio com o PB interfere na habilidade dos bebês brasileiros em reconhecer as DNA apresentadas durante um curto período de tempo na atividade experimental, dificultando a abstração e generalização dos padrões apresentados. Como veremos, usando a técnica do Olhar Preferencial, uma atividade experimental foi desenvolvida na qual foram utilizados estímulos baseados no PB, sendo construídos com determinantes do português e nomes inventados seguindo o padrão morfofonológico da língua. No entanto, foram criadas duas subclasses de nomes em função dos determinantes que os antecediam, diferenciando-as, dessa forma, daquelas do português. Sabe-se que a possibilidade de subgrupos dentro de uma categoria gramatical é recorrente em línguas naturais, e se faz presente no português na subclassificação de nomes em masculinos ou femininos, como veremos no item 2.3. Assim, usamos uma característica de língua natural, mas evitamos usar a subcategorização em dois gêneros, existente no português.

Nossas hipóteses específicas são: (1) bebês brasileiros aos 11 meses, tal como os bebês canadenses, são capazes de abstrair e generalizar os padrões das relações entre DNA a novas instâncias a partir da exposição em um curto período de tempo à uma língua pseudonatural, que preserva propriedades prosódicas, fonotáticas e distribucionais de uma língua natural, no caso, o PB; e (2) bebês brasileiros aos 11 meses são capazes de identificar uma propriedade comum às línguas naturais através da percepção de DNA nos estímulos linguísticos a que têm acesso durante um breve espaço de tempo, ainda que tais estímulos apresentem

elementos que lhes sejam familiares (devido à exposição ao PB), mas em uma nova configuração.

Assim, como objetivos específicos, temos:

- (1) replicar os resultados encontrados com bebês canadenses;
- (2) avaliar se a exposição prévia a um padrão semelhante ao de sua língua materna interfere no reconhecimento, abstração e generalização de DNA por bebês brasileiros.

As seções desta dissertação obedecem à seguinte ordem: no segundo capítulo, apresenta-se a divisão do léxico entre categorias funcionais e lexicais, explorando com mais atenção as propriedades dos elementos constituintes de cada categoria, tendo em vista a importância atribuída a tais elementos no processo de aquisição da linguagem. Além de descrevermos sobre os elementos das categorias lexicais e funcionais, o primeiro capítulo trata também do foco deste trabalho – as dependências não-adjacentes (DNA) – e ainda apresenta as características da pseudolíngua utilizada na atividade experimental. No terceiro capítulo, apresentam-se, como fundamentação teórica, as caracterizações do modelo de processamento – *Bootstrapping* fonológico – e do modelo de língua – Programa Minimalista. A conciliação entre os dois modelos permite-nos supor que o *input* linguístico não é opaco, pois, em termos gerais, a informação prosódica disponível desde muito cedo para a criança em processo de aquisição desencadearia o funcionamento do sistema computacional inato. Assim, a partir da identificação de padrões fonológicos e prosódicos da língua pelo bebê, haveria a formação mínima do Léxico (sem necessariamente o Sistema Computacional funcionar). A formação desse léxico mínimo, contendo elementos de categorias lexicais e funcionais subespecificados, com um número mínimo de traços formais, permitiria o desencadeamento da aquisição, permitindo o *parsing* sintático inicial e viabilizando o funcionamento do Sistema Computacional, independentemente das distinções conceituais-intencionais que possam expressar (CORRÊA, 2006). O quarto capítulo trata da metodologia experimental utilizada, bem como descreve os

estímulos utilizados durante a atividade experimental e os procedimentos adotados na aplicação da mesma. Os resultados da atividade experimental analisados estatisticamente também são apresentados nesse capítulo. O quinto e último capítulo apresenta as considerações finais e a conclusão.

2 OS ELEMENTOS DO LÉXICO E SUAS RELAÇÕES

Conforme o modelo de *Bootstrapping* Fonológico (Morgan & Demuth, 1996; Christophe *et al.*, 1997), este trabalho parte da hipótese de que a sensibilidade a padrões prosódicos e fonotáticos da língua dá origem a um léxico mínimo, como veremos no próximo capítulo. Assim, o aprendiz aos poucos se torna sensível às orações, sintagmas e palavras. A criança, segundo esta concepção, não parte de uma sequência de elementos do léxico para a extração da gramática de sua língua, e sim de segmentos os quais podem ser decompostos em elementos de **categorias funcionais** e elementos de **categorias lexicais** que compõem o léxico de uma língua.

Em termos gerais, o léxico de uma língua natural pode ser, basicamente, dividido em dois grupos – categorias lexicais e categorias funcionais. Essa divisão, segundo a Teoria Linguística, se refere à capacidade ou não de certos elementos atribuírem papel temático e parece ser básica em todas as línguas. Os elementos pertencentes às categorias lexicais possuem traços semânticos e atuam na determinação do papel temático de seus argumentos. Pertencem à categoria lexical os elementos constituídos pela combinação dos traços categoriais (formais) N e V: Nome [+N, -V], Verbo [+N, +V] e preposição [-N, -V]². Às categorias funcionais pertencem os elementos D (Determinante), I (Flexão) e/ou T (tempo verbal) e C (complementizador), que não atribuem papel temático (não tomam argumentos) e codificam traços gramaticais, isto é, traços que determinam relações e processos gramaticais (Baker, 2003).

²Há vasta discussão sobre a classificação das preposições. Isto porque além de constituírem uma classe fechada, em determinados contextos podem ser classificadas como itens funcionais enquanto que em outros podem ser classificadas como itens lexicais. Em sentenças do tipo “Eu gosto *de* você”, a preposição serve apenas como um elo entre verbo e objeto indireto, c-selecionando seu complemento (item funcional) e em outros contextos como em “Eu vou *de* carro” comporta-se como um item lexical, s-selecionando seus argumentos.

Nas seções seguintes, exploraremos com mais atenção as propriedades dos elementos constituintes das categorias lexicais e funcionais, tendo em vista a importância atribuída a tais elementos no processo de aquisição da linguagem.

2.1 ELEMENTOS LEXICAIS

Tradicionalmente, associa-se a categoria Verbo a eventos, Nomes a entidades (objetos e seres) e Adjetivos a propriedades. Entretanto, o debate sobre a classificação das categorias lexicais é antigo e controverso, visto que tais definições não dão conta de explicar a complexidade do processo de categorização das línguas naturais. Não há a pretensão, neste trabalho, de expor a discussão teórica sobre a classificação dos elementos lexicais, mas há o anseio de explorar a importância desses elementos no processo de segmentação do *continuum* sonoro, contribuindo, conseqüentemente, para o entendimento dos processos que subjazem a aquisição de uma língua natural.

2.1.1 As categorias lexicais segundo Chomsky (1970)

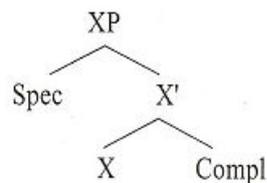
Para a Teoria dos Princípios e Parâmetros e, de modo geral, para a Teoria Gerativa (CHOMSKY, 1981), teoria base para o trabalho desenvolvido nesta dissertação, as categorias lexicais são distinguíveis pelo valor de seus traços binários. Dessa forma, os itens lexicais poderiam ser classificados pela combinação desses traços (Nome (N) e Verbo (V)) e seu valor (+) positivo ou (-) negativo para cada um desses traços, como podemos verificar no quadro abaixo, em que apenas dois traços seriam suficientes para descrever quatro categorias:

	+V	-V
+N	adjetivo	Nome
-N	verbo	preposição

Quadro 1- Diferenciação dos elementos lexicais N e V em relação a seus traços binários.

De acordo com tal classificação, os itens lexicais seriam ainda divididos em que contém elementos com ao menos um valor positivo para os traços – N, V, Adj³ – e possuem número indefinido, podendo novos elementos serem criados pelos falantes. O grupo de *classes fechadas* compreende somente Prep que apresenta o valor negativo para ambos os traços e são elementos fixos da língua, ou seja, não é possível acrescentar novos elementos a esse grupo.

Em suma, Chomsky chegou ao esquema de projeção de constituintes, abaixo, usando a notação X-barra, o que permitiu explicar que entre as categorias lexicais (N, V, Adj e Prep) e as categorias sintagmáticas (NP, VP, AP, PP) existiriam categorias intermediárias:



A Teoria X-barra, que permitiu captar as similaridades entre as categorias, pode ser descrita como uma tentativa de representação universal de um constituinte e busca apreender a relação sintática entre os elementos que compõem um constituinte:

A Teoria X-barra é módulo da gramática que permite representar um constituinte. Ela é necessária para explicitar a natureza do constituinte, as relações que se estabelecem dentro dele e o modo como os constituintes se hierarquizam para formar a sentença. Como acontece em qualquer módulo da gramática, a Teoria X-barra deve ser universal a ponto de configurar-se como um esquema geral, capaz de captar a estrutura interna dos sintagmas de qualquer língua; mas também deve prestar-se a dar conta da variação nas diferentes línguas. (MIOTO, 2007, p. 46)

Chomsky (1970), ao introduzir o componente central desta teoria – Teoria X-barra –, preocupava-se com as semelhanças, e não com as diferenças entre os elementos das categorias lexicais, o que influenciou outras propostas para a

³ Cabe ressaltar que os Adjetivos podem se constituir como classe com número finito de elementos.

classificação dos elementos lexicais e funcionais. A classificação a partir de traços das categorias lexicais, proposta por Chomsky, não era muito satisfatória até a formulação do Programa Minimalista, que será apresentado no capítulo 3 desta dissertação.

2.1.2 As categorias lexicais segundo Baker (2003)

Uma das propostas de classificação dos elementos do léxico, bastante relevante para este estudo é a proposta desenvolvida por Baker (2003). De acordo com este autor, as categorias lexicais diferem entre si muito mais do que pela simples classificação de traços binários. Sua abordagem se enquadra na perspectiva orientada pela sintaxe, pois o autor argumenta que é através da estrutura sintática, principalmente, que se define uma categoria (Baker, 2003, p.95). Em oposição à grande parte dos linguistas gerativistas que se preocuparam principalmente com as semelhanças entre as categorias, Baker (2003) busca investigar as diferenças entre as categorias lexicais, dentro de uma abordagem gerativa, sem desconsiderar totalmente os traços +/- N e +/- V propostos originalmente por Chomsky.

De acordo com a proposta de Baker (2003), as categorias lexicais se definem por seu comportamento dentro da estrutura sintática, ou seja, N, V e Adj são elementos sintáticos. Dessa forma, a categoria V é definida pela presença de um especificador (noção estritamente sintática), mas que se relaciona com propriedades morfológicas (como flexão de tempo) e semânticas (como atribuição de papéis temáticos), conforme Baker (2003, p.23) “X is a verb if and only if X is a lexical category and X has a specifier.”

Para definir a categoria N, Baker (2003) utiliza duas propostas: uma semântica e outra sintática:

“Semantic version: nouns and only nouns have *criteria of identity*, whereby they can serve as standards of sameness.”

“Syntactic version: X is a noun if and only if X is a lexical category and X bears a *referential index*, expressed as an

ordered pair of integers.” (BAKER, 2003, p. 95; grifos do autor)

Segundo tais classificações, N é definido a partir da relação de c-comando que estabelece dentro do sintagma. O índice referencial, em geral, corresponde a um referente no domínio semântico, o que não significa que seja um referente concreto no mundo, como proposto por concepções mais intuitivas, segundo as quais nomes representam coisas ou seres (Baker, 2003, p.96). Ainda de acordo com esse autor, a categoria Adj é definida a partir de traços negativos, diferentemente do que propõe Chomsky (1970), justamente por não apresentar as propriedades de Nomes e Verbos – não seleciona especificador nem possui um índice referencial. Torna-se possível estabelecer um quadro comparativo, contrastando-se a proposta de Chomsky (1970) e Baker (2003), como veremos abaixo. Cabe ressaltar que o trabalho de Baker teve como base a classificação já estabelecida por Chomsky; entretanto, buscou investigar as diferenças, e não as semelhanças, entre as categorias lexicais.

	CHOMSKY	BAKER
NOME	+N, -V	+N (tem um índice referencial)
VERBO	-N, +V	+V (tem um especificador)
ADJETIVO	+N, +V	-N, -V
PREPOSIÇÃO	-N, -V	As preposições fazem parte de um sistema diferente (funcional)

Tabela 1- Tabela comparativa das propostas de Chomsky (1970) e Baker (2003) sobre categorias lexicais.

Na próxima seção, trataremos de maneira mais específica sobre a importância dos elementos pertencentes às categorias funcionais no processo de segmentação do *continuum* sonoro.

2.2 ELEMENTOS FUNCIONAIS

Às categorias funcionais, classe fechada, pertencem os elementos como pronomes, preposições, morfemas e outros e, diferentemente dos itens lexicais, estes elementos não atribuem papel temático (não tomam argumentos) e codificam traços gramaticais, isto é, traços que determinam relações e processos gramaticais. Estes elementos, ao selecionarem seus complementos, têm em vista apenas as categorias às quais eles devam pertencer (c-selecionam).

As principais características dos elementos pertencentes a essa categoria são: não marcar papel temático (não tomam argumentos) e ter codificados traços gramaticais, isto é, traços que determinam relações e processos gramaticais. Outras características inerentes às categorias funcionais merecem ser destacadas, como veremos abaixo, para melhor compreensão de como tais elementos contribuem no processo de aquisição da linguagem:

- i) Categorias funcionais são classes fechadas, ou seja, apresentam número limitado de elementos que são altamente frequentes na fala;
- ii) Elementos de categorias funcionais tendem a ter durações vocálicas mais curtas, amplitudes mais fracas e estrutura silábica simplificada em comparação a itens das categorias lexicais (Morgan, Shi, Allopenna, 1996);
- iii) Elementos de categorias funcionais provêm a junção entre o léxico e a sintaxe de uma determinada língua;
- iv) Elementos de categorias funcionais facilitam a segmentação do contínuo da fala em unidades menores;
- v) Elementos de categorias funcionais tendem a ocorrer nas extremidades das unidades prosódicas; sendo assim, permitem, desde muito cedo, a identificação das palavras de conteúdo (i.e., itens lexicais) que as seguem (ou antecedem, de acordo com a língua).

Além disso, de acordo com a Teoria Gerativa, as categorias funcionais têm importância no processo de aquisição de uma língua, por considerar que a variação

entre as línguas pode ser entendida como uma consequência da força de traços das mesmas, e o processo de aquisição de uma língua como um processo de fixação de parâmetros ligados a essas categorias a partir da extração de padrões prosódicos e distribucionais da língua. Da mesma maneira, segundo o modelo de *Bootstrapping* fonológico, o desencadeamento da sintaxe parte da distinção entre elementos lexicais e funcionais, a partir da sensibilidade a padrões prosódicos e fonotáticos, o que daria origem a um léxico mínimo. De modo geral, considera-se que categorias funcionais são de fundamental importância no processo de aquisição de uma língua, visto que podem ser consideradas “âncoras” para predizer relações entre categorias em uma análise distribucional, e assim essenciais para a aquisição das categorias lexicais e a construção das relações sintáticas da língua.

Estudos revelam que a capacidade de distinção entre categorias funcionais e lexicais em uma língua artificial já estão presentes em bebês recém nascidos (Shi *et al*, 1999) e o reconhecimento de palavras funcionais em uma língua natural parece ocorrer entre 8 e 13 meses (Shi, Werker e Cutler, 2003). Além disso, uma série de estudos tem explorado a percepção de bebês e crianças a itens funcionais. Shi e colaboradores (1999), através da técnica experimental de sucção não nutritiva, em experimento realizado com bebês de 3 dias de vida, observaram que os bebês reagiram significativamente à mudança do tipo de estímulo a que foram familiarizados. Bebês familiarizados com itens funcionais reagiram à apresentação de itens lexicais e vice-versa. Bebês que escutaram o mesmo tipo de estímulo durante a familiarização e o teste (listas diferentes de itens lexicais nas duas fases ou listas diferentes de itens funcionais nas duas fases) não reagiram à mudança de lista. Esses resultados sugerem que bebês com poucos dias de vida são sensíveis a propriedades acústicas dos itens de sua língua que podem ser usadas, mais tarde, na identificação e distinção de itens funcionais e lexicais.

Shi e colaboradores (2006) investigaram, em tarefa de percepção de fala, que, em crianças aos 11 meses de idade, a alta frequência dos itens funcionais propiciaria a segmentação do vocabulário no *continuum* de fala, contribuindo para a aquisição de elementos do léxico de uma língua. Para a realização da atividade, as crianças eram familiarizadas com pseudonomes precedidos tanto de elementos funcionais reais (the – artigo definido em inglês) quanto por pseudoelementos

funcionais. A alta frequência de alguns itens funcionais (the) em relação a outros (her) na fala dirigida à criança também foi considerada. Assim, se a primeira sílaba do item funcional da sequência de familiarização fosse reconhecida como familiar pelos bebês, o pseudônimo com o qual é pareado poderia ser “aprendido” também em sua forma isolada e geraria um maior tempo de fixação do olhar quando aparecesse isoladamente na fase teste, sugerindo que a presença dos itens funcionais permite a segmentação do *continuum* sonoro.

Gómez e Lakusta (2004), com o objetivo de investigar como crianças começam adquirir categorias baseadas na forma e as relações entre elas, ou seja, se as crianças seriam capazes de identificar, abstrair e generalizar o padrão da língua às novas formas, propuseram duas atividades experimentais expondo crianças de 12 meses à estrutura sonora da forma aX e bY (correspondendo “a” e “b” a elementos de categorias funcionais e “X” e “Y” a lexicais, como em uma língua natural). Segundo as autoras, assimetrias categóricas desse tipo podem desempenhar um importante papel na aprendizagem, destacando pontos de ancoragem para a análise distribucional. De forma sucinta, o objetivo do teste era: identificar se crianças de 12 meses poderiam estabelecer as relações entre funções de palavras “a” e “b” e categorias X e Y baseadas em seus traços distintivos, a fim de generalizá-las a pares de palavras “a” e “b” com novos elementos “X” e “Y”. Os resultados sugerem que crianças, aos 12 meses, são capazes de discriminar pares de itens funcionais e lexicais congruentes de incongruentes a partir da distinção das pistas presentes.

Na próxima seção, trataremos ainda das categorias funcionais (marcadores morfossintáticos) com material interveniente – Dependências Não-adjacentes (DNA) – e sua contribuição para o processo de aquisição da linguagem.

2.2.1 Dependências não-adjacentes (DNA)

Como visto na seção anterior, as categorias funcionais são elementos essenciais no processo de aquisição de um língua materna, pois podem ser

consideradas “âncoras” para predizer relações entre categorias em uma análise distribucional, e assim essenciais para a aquisição das categorias lexicais – constituição do léxico – e construção das relações sintáticas da língua. O presente estudo investiga a sensibilidade de bebês brasileiros no fim do primeiro ano de vida e início do segundo ano a dependências não-adjacentes entre determinantes e a terminação fonológica de nomes.

Dependências não-adjacentes são definidas como a co-ocorrência de marcadores morfossintáticos e/ou fonológicos, com material interveniente. Referem-se, portanto, a elementos de categoria funcional que estão “intimamente” relacionados, porém com um componente extra entre um elemento funcional e outro (funcional ou não funcional, como vogal temática de N). As dependências não-adjacentes são consideradas de fundamental importância para o processo de aquisição de uma língua, pois:

- i) São importantes para o processo de (sub)categorização como, por exemplo, na identificação do gênero de um nome (N masc. ou N fem.);
- ii) São determinantes para o estabelecimento das relações sintáticas e morfológicas, como em: “... **está correndo**,... **tem corrido**”;

Portanto, DNA podem ser consideradas ferramentas importantes no processo de aquisição de uma língua. Para tal, é necessário que a criança as reconheça no *continuum* sonoro disponível, as perceba como padrões abstratos e as generalize em novas ocorrências. Há evidências importantes da capacidade de abstração e generalização, como demonstram trabalhos experimentais tanto em língua artificial com adultos (Gómez, 2002; Newport & Aslin, 2004) e bebês de sete e dezoito meses, respectivamente (Marcus *et al.*, 1999; Gómez, 2002), quanto em língua natural em trabalhos com crianças de 18 meses (Santelmann & Jusczyk, 1998) em tarefa de reconhecimento de padrão da língua materna. Entretanto, como veremos, a maioria dos trabalhos utiliza línguas totalmente artificiais e quando utilizam língua natural trabalham com crianças mais velhas (aos 18 meses), destacando, por esse motivo, a importância do trabalho desenvolvido nesta dissertação.

Com o intuito de investigar os mecanismos de aprendizagem utilizados por crianças no início do aprendizado de uma língua, Marcus e colaboradores (1999) propõem ao menos 2 mecanismos de aprendizagem: mecanismo estatístico (Saffran *et al*, 1996) e regras algébricas. Assumindo a hipótese de que crianças são capazes de abstrair a estrutura adjacente de uma língua e generalizá-la a novas instâncias, os autores esperam que crianças, utilizando o paradigma de escuta preferencial, atentem por mais tempo, na fase teste, a itens inconsistentes em relação aos itens consistentes.

Nas três atividades experimentais, bebês de 7 meses de idade eram habituados, durante 2 minutos de fala, a sentenças de três palavras construídas a partir de uma língua artificial (no formato “ABA” ou “ABB”) e, posteriormente, testadas com novas sentenças de três palavras, as quais não haviam aparecido na fase de habituação. Devido ao fato de as sentenças testes serem novas, em relação às da fase de habituação, as crianças não poderiam distinguir as sentenças baseadas nas probabilidades transicionais, assim como por serem constituídas de mesma duração e contorno prosódico elaborado em computador, não era possível distingui-las baseando nas propriedades estatísticas como número de sílabas e prosódia. Os resultados da primeira atividade experimental evidenciaram que as crianças mostraram preferência por sentenças inconsistentes em relação às consistentes, como era esperado.

Entretanto, devido ao fato de a sequência dos traços fonéticos do teste ser próxima à da habituação, uma segunda atividade experimental foi proposta. Diferentemente da primeira, os itens da fase de habituação não proveriam informações sobre a relação entre consoantes vozeadas e desvozeadas, distinguindo os traços fonéticos das fases de habituação e teste. Mais uma vez, os resultados foram de acordo com a previsão, ficando as crianças atentas por mais tempo aos itens inconsistentes em relação aos consistentes. A partir da indagação de que crianças poderiam ter se habituado com a propriedade de reduplicação que distingue as duas “gramáticas” “ABA/ABB” (repetições em ABB contêm elementos imediatamente reduplicados (“ti ti”), o que não ocorre nas sequências de “ABA”), uma terceira atividade experimental foi proposta - comparação entre sentenças construídas pela formação “ABB” e sentenças construídas pela formação “AAB” com

os mesmos itens do experimento 2 –e os resultados foram também compatíveis com a previsão.

A partir dos resultados, os autores propõem um mecanismo no qual crianças extraem um tipo de regras abstratas que representem relações entre variáveis, como “o primeiro item de X é o mesmo do terceiro item de Y”. Assim, os resultados parecem mostrar que crianças possuem a capacidade de extrair rapidamente essas regras de um pequeno número de ensaios e de generalizar tais regras para novas instâncias. Segundo os autores, se essa posição estiver correta, o bebê, nos primeiros meses de vida, possui no mínimo duas ferramentas distintas de aprendizagem e de “investigação” do aprendizado da linguagem: um que trata de relações estatísticas, como probabilidades transicionais, e outro, que manipula variáveis, permitindo o aprendizado de regras por crianças.

Gómez (2002), com o intuito de investigar o processo de reconhecimento e generalização de dependências não-adjacentes (DNA) e, considerando que uma possível base para essa aprendizagem seriam as dependências adjacentes (DA) (condicionantes de ordem inferior numa hierarquia em que dependências adjacentes seriam aprendidas anteriormente e serviriam de base para o aprendizado de DNA), investigou a possibilidade de aprendizagem de DNA por adultos e bebês aos 18 meses. Para a realização das tarefas, duas línguas artificiais com 3 elementos distintos foram elaboradas de modo que os estímulos nas duas línguas começavam e terminavam com os mesmos elementos e continham os mesmos elementos de transição (**x**) que representa o número da variabilidade do elemento intermediário que poderia ser variável em 2, 6, 12 ou 24 elementos. Por exemplo, língua 1: “pel-**x**-rud”, “vot-**x**-jic”, “dak-**x**-tood”; língua 2: “pel-**x**-jic”, “vot-**x**-tood”, “dak-**x**-rud”. Dessa forma, os aprendizes poderiam distinguir as línguas somente pela aquisição das DNAs. Além disso, a maior variabilidade de **x** propiciaria uma maior focalização nas DNA e, conseqüentemente, um aprendizado mais eficaz das mesmas. Após 18 minutos de exposição e agora já na fase teste, os adultos eram informados que os estímulos ouvidos durante o treinamento pertenciam a um conjunto de regras de ordem de palavras e que ouviriam, nesta fase, mais 12 estímulos (6 com a mesma estrutura da fase de familiarização + 6 novos - pertencentes ao outro grupo da língua artificial). Eram também instruídos a selecionar determinada tecla em um

computador se os estímulos ouvidos fossem congruentes com os ouvidos anteriormente. Os resultados obtidos com adultos sugerem que o aumento de variabilidade do elemento intermediário propicia um aprendizado mais eficaz das DNAs. Assim, participantes expostos ao conjunto de 24 variações para “x” apresentaram resultados mais satisfatórios que os dos outros grupos, evidenciando melhor aprendizado de DNAs. A fim de se verificar a mesma habilidade em crianças e, ainda, se o aumento da variabilidade do elemento intermediário também facilitaria a identificação de padrões de DNAs, bebês, aos 18 meses, foram expostos aos mesmos estímulos, com período de familiarização de 3 minutos e testados com a técnica do Olhar Preferencial. Os resultados sugerem que o desempenho dos bebês, assim como o de adultos, aumentou radicalmente com o aumento da variabilidade do elemento intermediário, ou seja, bebês expostos ao conjunto de 24 variações para “x” também apresentaram resultados mais satisfatórios que os dos outros grupos. Ainda que seja um pouco forçado comparar os resultados das duas atividades experimentais, principalmente devido à diferença do tempo de exposição à língua artificial (18 min (adultos) e 3 min (bebês)) e pelo fato de que os adultos foram instruídos para a realização da tarefa, os resultados foram ao encontro da hipótese inicial de que a alta variabilidade do contexto poderia levá-los a concentrarem-se em estruturas menos variáveis, ou seja, DNAs. De modo mais amplo, os resultados sugerem que aprendizes são sensíveis a estrutura das línguas e, especificamente, aos elementos constituintes das DNAs.

Utilizando língua natural, Santelmann & Jusczyk (1998) focalizaram a sensibilidade de bebês americanos a DNAs entre o auxiliar *-is* e o morfema verbal *-ing* (na versão normal) e entre *can* e o morfema verbal *-ing* (versão modificada), utilizando a técnica de escuta preferencial. Crianças de 15 e de 18 meses escutaram pequenas histórias contendo verbos no presente contínuo (*is ...-ing*), na versão normal, e (*can ...-ing*) na versão modificada. Baseados em estudos anteriores (SHADY, 1996 *apud* SANTELMANN & JUSCZYK, 1998, p. 115) em que crianças entre 14 e 16 meses mostraram sensibilidade na distribuição e função da ordem dos morfemas em inglês, neste trabalho, bebês, aos 15 meses, não demonstraram sensibilidade a dependências entre o auxiliar “is” e o morfema “ing”, ou seja, não reconheceram a co-ocorrência entre esses morfemas. Entretanto, diferentemente do

resultado obtido com bebês aos 15 meses, bebês aos 18 meses se mostraram sensíveis para as relações entre esses morfemas, estabelecendo a relação entre os elementos não contíguos presentes na língua inglesa.

Os pesquisadores fizeram outros experimentos, variando o número de sílabas encaixadas entre os dois morfemas (colocando advérbios com duas ou mais sílabas), com crianças de 18 meses. As crianças continuaram sensíveis à relação entre os morfemas em uma janela de até 3 sílabas, i.e., quando havia entre *is* e *-ing* somente a sílaba do verbo ou, além dela, um advérbio de 2 sílabas. Porém, com uma “distância” maior (com advérbios de 3 e 4 sílabas), não houve diferença significativa no tempo médio de escuta das duas versões. Esses resultados apontam para a existência de uma “janela de processamento”, em que a criança é capaz de estabelecer relações entre morfemas.

Como vimos, os estudos existentes, em sua grande maioria, exploram tais padrões (DNA) em línguas totalmente artificiais (Marcus *et al.*, 1999; Gómez, 2002; Newport & Aslin, 2004⁴), sem levar em conta propriedades morfofonológicas e prosódicas, ou, quando trabalham com línguas naturais, investigam tais capacidades em crianças mais velhas (18 meses), considerando, sobretudo, o reconhecimento de padrões de sua língua materna, e não propriamente habilidades de abstração e generalização (Santelmann & Jusczyk, 1998). Diferentemente, o presente trabalho busca avaliar se a criança é capaz de identificar uma propriedade comum às línguas naturais através da percepção de DNA nos estímulos linguísticos a que tem acesso durante um breve espaço de tempo. Para tal, uma língua pseudonatural foi utilizada a partir de propriedades morfofonológicas e prosódicas presentes em língua natural.

Na próxima seção, exploraremos as características acerca do gênero por constituir um vasto campo para ilustração de DNA. Além disso, explicaremos a

⁴De modo sucinto, Newport & Aslin (2004) investigaram se adultos são capazes de “computar” regularidades de sons não-adjacentes e determinar, dessa forma, os tipos de regularidades não-adjacentes que são adquiridas facilmente. Para isso, propuseram quatro atividades experimentais a partir da elaboração de uma língua artificial contendo DNAs. Após o período de exposição aos estímulos, o aprendizado da coerência estatística de sílabas não-adjacentes entre palavras foi testado, a fim de verificar se os participantes – monolíngues em inglês – poderiam discriminar palavras de partes de palavras. Os resultados, de forma geral, contribuem para o entendimento dos mecanismos de aprendizagem estatística que subjazem o processo de aquisição de uma língua. Ressalta-se que, assim como em Gómez *et al.* (2002), não se pode afirmar que o que ocorreu foi a aquisição espontânea de uma língua, como o que ocorre com crianças, visto que se trata de adultos que foram orientados em uma tarefa.

subclassificação entre N e Det na pseudolíngua utilizada na atividade experimental e sua motivação.

2.3 CONSIDERAÇÕES SOBRE GÊNERO EM NOMES E DETERMINANTES

Nesta seção, transcorreremos acerca da importância do gênero como um campo produtivo para DNA.

Gênero é uma ótima ilustração de DNAs, pois se encontra manifesto morfofonologicamente em 75% das línguas naturais (Corbett, 1991; Van Berkum, 1997, *apud* Name, 2002). Portanto, trata-se de um fenômeno altamente produtivo, que nos ajuda a compreender e criar hipóteses que elucidam o processo de abstração e generalização dos padrões recorrentes na língua.

Classes de gênero podem refletir uma categoria semântica, gerando subcategorias na classe dos N nessas línguas e estabelecendo uma relação entre propriedades dos elementos da classe denotada pelo nome e o gênero gramatical. A principal característica marcada pelo gênero é o sexo (nomes masculinos são atribuídos a entidades do sexo masculino, e nomes femininos a entidades do sexo feminino), seguida da oposição animado/inanimado e [+ / - racional] (Corbett, 1991 *apud* Name, 2002). Entretanto, não há relação sistemática nos sistemas de gênero, visto que a correspondência entre gênero gramatical e classes semânticas nem sempre pode ser constatada, pois ainda que possa haver relações entre a distribuição de nomes em classes de gênero e um padrão fonológico da terminação ou declinação de nomes, esse critério não dá conta dessa distribuição (Name, 2002). Por isso, a noção semântica não se torna suficientemente forte para definir o que é gênero gramatical.

A noção de gênero é gramatical, isto é, faz parte do sistema gramatical de muitas línguas, e se manifesta nas relações de concordância entre o nome e diferentes elementos do sintagma/da oração. Entretanto, as línguas apresentam um padrão diferenciado em relação à subdivisão dos elementos do léxico. Assim, no

que tange a esta pesquisa, em francês, por exemplo, ainda que alguns autores chamem a atenção para certas regularidades, não existe um padrão fônico de terminação que facilite a identificação da distribuição dos nomes nas duas classes de gênero. Assim: *Table, maison, fleur, main* são nomes femininos, mas nomes com a mesma terminação - *cabale, ballon, coeur, pain* - são masculinos. Em Nomes, tradicionalmente, a terminação -e é marca flexional do gênero feminino *boulangier/boulangère, chien/chienne*. No entanto, o -e final é quase que exclusivamente uma marca da língua escrita, tendo uma pequena diferença em termos perceptuais/auditivos, visto que os nomes masculinos são pronunciados com “ê” fechado no final (“boulang[e]”) e o feminino, com “é+R” no final (“boulang[εR]”). Nos determinantes, o gênero só é manifesto no singular e também há acomodação fônica entre o determinante e o nome ou adjetivo começando por vogal(*cet escalier* – masc., *cette échelle* – fem.; *cet* e *cette* têm ambos a mesma pronúncia [sɛt]) (Name, 2002). No inglês, o gênero se manifesta somente em pronomes pessoais (e algumas profissões), em uma relação semântica direta com o elemento denotado pelo nome, não havendo manifestação de gênero (informação morfológica) entre determinantes e nomes.

No PB, há dois gêneros para nomes: masculino (não- marcado) e feminino (marcado) (Cunha & Cintra, 2008) e existe relação entre a terminação fonológica dos nomes e seu gênero. Assim, praticamente todos os nomes com a terminação -o são masculinos e a grande maioria com a terminação -a são femininos, enquanto que os nomes com terminação em -e podem ser tanto masculinos quanto femininos. Ex: bolo, corpo, papagaio (masc.); bota, cadeira, mobília (fem.); aste, ponte, (fem.) dente, pente (masc.). Alguns nomes fogem a essas regras e, em sua grande maioria, são os que contêm terminação em -a, mas são masculinos, como: planeta, cinema, drama, samba, panda, entre outros. O contrário também ocorre – nomes com terminação em -o e femininos – , porém ocorrem em número reduzido de palavras, exatamente em quatro palavras da língua portuguesa: libido, foto, moto, tribo⁵.

⁵ “Foto” e “moto” são reduções das palavras “fotografia” e “motocicleta”. Mas do ponto de vista da criança, são mais frequentes do que as palavras que lhes deram origem, por isso a consideramos como palavras.

De acordo com a Teoria Linguística Gerativa em sua vertente chomskyana (Chomsky, 1995), o gênero dos nomes, em português, pode ser intrínseco ou opcional, classificado assim quando o traço de gênero é parte integrante do nome ou quando o valor do traço varia morfofonologicamente, respectivamente. Em nomes [-animado], o traço é intrínseco (cama, bola, piscina, relógio, quarto, vestido). Em nomes [+animado], o traço *pode* ser intrínseco (homem, girafa) ou opcional (menino/a, macaco/a), sendo que estes últimos apresentam variação morfofonológica. Tanto o traço de gênero intrínseco quanto o opcional são interpretáveis nos nomes. Assim, os traços de gênero são semanticamente interpretados na interface da língua com os sistemas de desempenho. Entretanto, na grande maioria dos casos, o traço intrínseco é arbitrário, pois não possui valor semântico, o que faz com que ele seja subespecificado semanticamente. Por outro lado, os traços opcionais marcados morfofonologicamente possuem conteúdo semântico, remetendo às classes de gênero natural masculino e feminino.

Mesmo não sendo a semântica a responsável pela identificação do sistema de gênero por parte das crianças, elas não apresentam dificuldades em adquirir o sistema de gênero de sua língua e não usam prioritariamente associações semânticas entre gênero e elemento denotado (por exemplo, sexo) (Pinker, 1995). Dessa forma, a criança teria que se valer de outra informação, que não seja a relação entre o padrão fonológico da terminação dos nomes e sua classe semântica, para identificar o sistema de gênero de sua língua.

A noção de concordância entre nome e outros itens é crucial para o conceito de gênero (Comrie, 1999). De acordo com NAME (2002, p. 34), a regularidade e transparência da marcação de gênero em uma classe fechada como a Categoria D permitiria à criança identificar os possíveis valores do traço de gênero de sua língua mais facilmente do que em elementos de uma classe aberta – Nome – com maior variação de terminações. Do ponto de vista do Programa Minimalista, concordância é vista como uma relação de checagem (Chomsky, 1995) ou valoração (Chomsky, 1999) de traços. Gênero é entendido como um traço formal, formando, junto com pessoa e número, o conjunto de *traços-phi*. Em português, a concordância de gênero é marcada em todos os determinantes. Dessa forma, sempre haverá concordância de gênero entre determinantes e nomes, ressaltando que a

subcategorização em dois gêneros é propriedade do português, assim como também o é em diversas línguas naturais (com duas ou mais subcategorias).

A percepção da relação entre DNA no DP, no primeiro ano de vida, poderia ser um precursor, uma etapa inicial desse processo de aquisição de gênero ou, de forma mais ampla, de identificação de subcategorias dentro da categoria N. O objeto dessa dissertação não é a aquisição de gênero gramatical, mas observamos que tal aquisição parece estar intimamente ligada ao processo de identificação e de generalização de DNA. Defende-se, portanto, que a sensibilidade da criança, sobretudo ainda no primeiro ano de vida, a regularidades do *continuum* sonoro e a capacidade de tratamento desse estímulo identificando, rapidamente, combinações recorrentes – DNA –, abstraindo seus padrões e generalizando-os em novos enunciados são habilidades cruciais para a (sub) categorização lexical.

Essa dissertação busca dar continuidade aos resultados de resultados encontrados por Name (Name, Shi & Koulaguina, 2011; Name & Shi, em prep.) em que bebês aos 11 meses, inseridos naturalmente na língua inglesa e/ou francesa, a partir de curta exposição (2 minutos) a uma língua pseudonatural diferente de sua língua materna – no caso, o PB, foram capazes de reconhecer padrões de relações entre D e N e, ainda, de abstrai-los de modo a identificá-los em novos enunciados. Assim, ao contrastar os resultados, considerando que o padrão da língua pseudonatural utilizada na atividade, apesar de não ser especificamente o padrão do PB, como veremos adiante nesta seção, é mais semelhante a este que a um padrão do inglês e do francês, buscamos verificar se crianças brasileiras apresentam resultados semelhantes e ainda avaliar se a exposição prévia a um padrão semelhante, porém distinto de sua língua materna, interfere no reconhecimento, abstração e generalização de DNA por bebês brasileiros.

Em nossa língua pseudonatural, utilizada na atividade experimental (tanto com crianças canadenses quanto com brasileiras), para se evitar a identidade fonológica entre determinantes e nomes, foram utilizados quatro determinantes masculinos (NOSSO, ESSE, MEU e SEU) seguidos de pseudopalavras que preservam terminações alternativas ao padrão masculino –o, ou seja, terminadas em –a ou –e, como, por exemplo: *guika*, *tova*, *mufe* e *pebe*. Como a possibilidade de subgrupos dentro de uma categoria gramatical é recorrente em línguas naturais, e

se faz presente no português na subclassificação de nomes em masculinos ou femininos, para esta língua pseudonatural, foram criadas duas subclasses de nomes em função dos determinantes que os antecediam. Assim, usamos uma característica de língua natural, mas *evitamos usar a subcategorização em dois gêneros*, existente no português, possibilitando, dessa forma, que a atividade experimental fosse aplicada tanto a crianças canadenses (que não tinham contato com o PB) quanto a crianças brasileiras (nativas), já que os padrões de DNA apresentados aos bebês brasileiros diferem do que já é conhecido, pois ainda que os determinantes lhes sejam familiares (devido à exposição ao PB), sua divisão em duas subclasses distintas foge ao padrão da língua (i.e., esses determinantes pertencem todos a uma única subclasse no PB, de masculinos, podendo, todos eles, anteceder os mesmos nomes). Ou seja, uma nova subclassificação foi criada permitindo que tanto bebês canadenses quanto bebês brasileiros estivessem de fato em exposição a um novo padrão, a uma nova língua. Como veremos, mais que uma replicação da atividade experimental, o trabalho desenvolvido nesta dissertação, se propõe a apresentar uma tarefa ainda mais complexa para as crianças brasileiras do que para as canadenses, isto porque enquanto estas estavam pela primeira vez diante de elementos de uma nova língua, aquelas estavam diante de elementos de sua língua (determinantes) e de um novo padrão de subcategorização diferente de sua língua materna. Dessa forma, a criança, exposta, por um curto período de tempo, aos estímulos, provavelmente, identificaria os elementos de sua língua, mas, tão rápido quanto e ao mesmo tempo, teria que abstrair e generalizar o novo padrão da pseudolíngua a novos estímulos nunca antes ouvidos.

3 FUNDAMENTAÇÃO TEÓRICA

Neste capítulo, apresentam-se os fundamentos teóricos que dão suporte à hipótese de trabalho que orienta esta pesquisa. Considera-se uma possível aproximação entre os desenvolvimentos recentes da teoria linguística gerativa, em sua versão chomskyana (Chomsky, 1995), que apresentam uma nova concepção da Faculdade da Linguagem, e o modelo psicolinguístico de processamento comprometido com a aquisição da linguagem – o *Bootstrapping* fonológico (Morgan & Demuth, 1996; Christophe *et al.*, 1997) de modo a compreender a maneira pela qual a criança extrai do *continuum* sonoro as propriedades fonéticas e fonológicas e como estas são transpostas para o nível formal.

3.1 BOOTSTRAPPING FONOLÓGICO

Uma das questões mais relevantes no estudo da aquisição da linguagem é explicar de que modo a criança, exposta à sua língua materna, dá início à identificação da gramática e as condições que tornam tal identificação possível. De forma geral, o estudo psicolinguístico da aquisição da linguagem pode ser distinguido, segundo Corrêa (2008), em duas fases: a primeira (décadas de 60 e 70), voltada para pesquisas sobre o tipo de informação usada pela criança na busca da interpretação dos enunciados linguísticos, bem como o tipo de informação gramatical a que ela estaria atenta (o que se evidenciaria em sua produção de fala); a segunda (a partir dos anos 80), caracterizada pela convergência dessas duas linhas de pesquisa, tendo em vista o problema do *bootstrapping* na aquisição da linguagem:

Como explicar o modo como a criança chega à sintaxe da língua partindo de informação de outra ordem – informação proveniente do possível significado dos enunciados, inferido pela criança do contexto em que estes são usados e/ou proveniente do próprio material que se apresenta à percepção (CORRÊA, 2008:173).

Essa indagação aponta para um problema de circularidade: assume-se que é necessário deter conhecimento de uma gramática para processar o estímulo linguístico, ao mesmo tempo em que é necessário processar o estímulo linguístico para identificar a gramática da língua na qual a criança se encontra inserida. Algumas soluções, que buscam explicar como o processo de identificação da gramática da língua é desencadeado, foram propostas, como o *bootstrapping* semântico (Pinker, 1987), sintático (Gleitman, 1990) e fonológico (Morgan & Demuth, 1996; Christophe et al., 1997).

A hipótese do *bootstrapping* semântico (Pinker, 1987) consiste na ideia de que as entidades gramaticais referem-se a classes semânticas no discurso. Assim sendo, objetos e pessoas seriam referenciados utilizando-se nomes, ações físicas e mudanças de estados, utilizando-se verbos, e definitude seria expressa por determinantes. Assim, papéis temáticos, como agente e tema, seriam tomados como correspondentes ao sujeito e ao complemento de orações, do mesmo modo em que o conhecimento conceptual de “objeto/pessoa” seria mapeado, por exemplo, na categoria “nome” e o conhecimento sobre ações na categoria “verbo”. Dessa forma, a partir do significado lexical, a análise semântica seria capaz de alavancar o domínio da sintaxe pela criança. As limitações dessa hipótese foram apontadas com o que veio a ser denominado de *bootstrapping* sintático.

A hipótese do *bootstrapping* sintático (Gleitman, 1990) considera que a criança interpreta um dado evento de fala guiada pela estrutura sintática. De acordo com essa hipótese, para que a criança adquira o significado dos verbos, ela teria de se basear em informações estruturais, deduzindo as estruturas argumentais, com as quais os significados dos verbos estão associados.

Saussure (2002), desde os primórdios da linguística moderna, atenta para a arbitrariedade entre o mapeamento de som e significado do léxico de uma língua. Dessa forma, o mapeamento entre forma e conteúdo, no curso da aquisição de língua materna, deve, ao menos, proceder a duas fases: Segmentação e Categorização (Christophe et al., 1997). A segmentação diz respeito à identificação da forma de palavra no contínuo da fala, enquanto que a categorização diz respeito ao mapeamento desta forma a seu significado. Chomsky (1965, p. 30) também chama atenção para o fato de que a criança, no curso de aprendizado de uma

língua, precisa ter uma capacidade para representar os sinais da fala e um modo de representar a informação estrutural destes sinais.

No entanto, a delimitação dessas unidades depende da percepção e análise do sinal acústico da fala, resultando na segmentação do mesmo em unidades sintáticas e lexicais. Cabe à criança, exposta aos dados linguísticos apresentados na fala natural do adulto, que ocorre de forma contínua, sem pausas entre as palavras, segmentar o fluxo da fala em unidades relevantes ao processamento linguístico, ainda que não tenha domínio do sistema fonológico e do léxico da língua com a qual está em contato. Essas unidades, uma vez combinadas, formam enunciados que remetem a objetos, animais, pessoas, estados mentais, eventos e transmitem propósitos e ideias. Mas, como se dá esse processo? Como a criança recorta essas unidades? De que forma elas se combinam? A hipótese do *bootstrapping*, seja ele semântico, sintático ou fonológico, refere-se à presença de pistas na superfície da fala capazes de revelar/ alavancar a estrutura sintática da língua e, assim, de favorecer o processo de aquisição da linguagem. Tanto na hipótese do *bootstrapping* semântico quanto na hipótese do *bootstrapping* sintático, há de se partir do pressuposto de que as unidades lexicais já estão previamente segmentadas (CORRÊA, 2006, 2008). Diferentemente desses, a hipótese do *bootstrapping* fonológico leva em conta a questão da segmentação do sinal acústico da fala pela criança, sendo o primeiro processo desencadeador da aquisição sintática e lexical.

De acordo com Morgan & Demuth (1996), o *bootstrapping* fonológico é a hipótese de que uma análise fonológica do sinal da fala pode permitir às crianças adquirirem o léxico e a sintaxe de sua língua. Esta proposta surge como questão relevante para o processo de aquisição da linguagem, visto que a tarefa inicial da criança deve ser a segmentação do fluxo da fala em unidades relevantes para o processamento linguístico, pressupondo que a criança, desde os primeiros dias de vida, esteja atenta a pistas recorrentes da língua. Christophe *et al.* (1997) apresentam quatro tipos de informações utilizadas pelas crianças para identificarem, no *continuum* da fala, as formas de palavras. São elas:

- i) regularidades distribucionais: posição frequente de determinados itens em sintagmas e frases;

- ii) regularidades fonotáticas: probabilidade de ocorrência contígua de determinados fonemas;
- iii) formas típicas de palavras: características inerentes e específicas às categorias funcionais e às categorias lexicais;
- iv) pistas de fronteiras prosódicas.

Essa hipótese sugere que, a partir da análise do sinal acústico da fala, a criança reconheceria padrões fonotáticos, distribucionais e prosódicos de sequências fonológicas, o que viabilizaria o acesso à estrutura sintática e aos itens lexicais de sua língua. A regularidade distribucional diz respeito à intuição por parte da criança de que sequências sonoras que ocorrem com frequência em uma variedade de contextos são melhores candidatas ao léxico do que as que ocorrem raramente em poucos contextos. As informações fonotáticas referem-se às limitações na ocorrência de sons entre palavras e sentenças, considerando que alguns sons ou combinações de sons ocorrem somente em fronteira de palavra, e outros ocorrem apenas no interior da palavra.

Gout & Christophe (2006) investigaram os processos iniciais da aquisição da língua materna, salientando a relevância da percepção de contornos prosódicos na delimitação de fronteiras lexicais e sintáticas. O objetivo deste estudo era investigar o papel do acento na frase fonológica para a fixação de parâmetros relativos ao posicionamento núcleo/complemento, na ordem linear em que estes se apresentam nas sentenças de uma dada língua. Através de atividades experimentais contrastando línguas como o francês e o turco, elas encontraram resultados compatíveis com a hipótese de que, a partir de oito meses de idade, a criança já é capaz de distinguir sequências de sílabas frequentes de sequências de sílabas raras na língua, sendo também sensível a fronteiras de unidades prosódicas.

A partir dos resultados apresentados e da visualização da figura abaixo, podemos supor que a prosódia forneceria uma segmentação inicial, cujo resultado seriam unidades menores, mais fáceis de serem segmentadas, a fim de se atingirem as unidades lexicais. Além disso, as propriedades prosódicas, fonológicas e distribucionais, agindo em conjunto, facilitariam a distinção dos elementos lexicais e seu mapeamento em categorias lexicais.

tempo, os bebês vão se tornando sensíveis a constituintes de orações cada vez menores. Aos 4 meses, eles são sensíveis à fronteira entre orações, aos 9 meses são sensíveis a fronteiras entre sintagmas e, aos 11 meses, sensíveis à fronteira de palavras. Por volta de 9 a 12 meses, a criança já se torna sensível à posição estrutural dos itens funcionais (ver Bagetti, 2009, para revisão da literatura).

A hipótese do *bootstrapping* fonológico torna-se, portanto, alicerce para a pesquisa desenvolvida nesta dissertação de mestrado, ao assumirmos que mecanismos de abstração e generalização de padrões do sinal acústico são recursos disponibilizados precocemente e explorados pelos bebês de modo a levar à categorização de elementos lexicais, mesmo expostos a um pequeno número de enunciados. Assim, diante de um *continuum* sonoro linguístico, em um primeiro momento, há a segmentação do fluxo da fala em unidades cada vez menores e, posteriormente, a distinção entre elementos lexicais e funcionais a partir da identificação e generalização de padrões – distribucionais, fonotáticos e prosódicos – viabilizando a aquisição da estrutura sintática e, conseqüentemente, do léxico de sua língua materna.

3.2 PROGRAMA MINIMALISTA

Como visto anteriormente, um modelo de processamento – *Bootstrapping* fonológico – ajuda, em parte, a explicar como as crianças, em processo de aquisição da linguagem, têm acesso à estrutura sintática e aos itens lexicais de sua língua a partir do reconhecimento de padrões fonotáticos, distribucionais e prosódicos de seqüências fonológicas. Entretanto, tomado isoladamente, não esclarece a passagem de um nível de representação fonético/fonológico para um nível de representação formal (CORRÊA, 2008), pois ainda que a criança identifique no *continuum* sonoro as propriedades fonéticas/fonológicas e os padrões recorrentes da língua e mais especificamente, os traços formais dos elementos funcionais da língua, é preciso assumir uma predisposição inata para a aquisição lexical e sintática de sua língua. Dessa forma, torna-se interessante a integração do modelo de

processamento assumido nesta dissertação – *Bootstrapping* fonológico – com um modelo de língua – Programa Minimalista (CHOMSKY 1995 e desenvolvimentos posteriores) – a fim de tentar explicar o problema lógico da aquisição de uma língua: como ocorre a representação formal da língua a partir dos dados prosódicos/fonológicos extraídos do *continuum* sonoro.

A concepção de língua, conforme proposta por Chomsky e colaboradores no Programa Minimalista (PM), seria possibilitada por uma Faculdade de Linguagem, isto é, por uma disposição biológica para a língua, inata e comum a todos os humanos. A proposta do PM remete ao conceito de língua-I (interna), tomado como um procedimento gerativo que incorpora um Sistema Computacional Linguístico universal e um léxico. Tal proposta torna-se interessante pelo fato de que pela primeira vez se estabeleceria tanto o caráter inato (Sistema Computacional) quanto o adquirido (Léxico) no processo de aquisição de uma língua de maneira natural e espontânea, como veremos adiante. A Faculdade da Linguagem, proposta por este modelo, pode ser caracterizada em duas perspectivas: no sentido estrito – *FLN* – *Faculty of Language in the narrow sense* – e no sentido amplo – *FLB* – *Faculty of Language in the broad sense* (Hauser, Chomsky & Fitch, 2002). No primeiro sentido, constitui-se de um Sistema Computacional e um Léxico. No segundo sentido, compreende FLN e sistemas cognitivos perceptuais e conceptuais com os quais faz interface.

No que diz respeito ao primeiro sentido – FLN, o Sistema Computacional pode ser definido como um mecanismo estritamente linguístico, universal e comum a todas as línguas. O léxico de uma e qualquer língua, de acordo com o PM, é formado por elementos pertencentes a categorias lexicais (classes abertas) e funcionais (classes fechadas), como vimos anteriormente. Adquirido a partir da experiência linguística, o léxico reúne as especificidades de cada língua e seus elementos teriam em suas subdivisões três matrizes de traços:

- (i) traços semânticos e os traços formais semanticamente interpretáveis (categorias N e V), interpretados na interface semântica;
- (ii) traços fonológicos, interpretados na interface fonética;
- (iii) traços que não recebem interpretação em nenhuma das duas interfaces, os denominados traços formais, não interpretáveis, como veremos a seguir.

O segundo sentido – FLB – compreende FLN e os sistemas cognitivos de desempenho perceptuais e conceituais com os quais FLN faz interface.

De acordo com o PM, uma língua fornece informações aos sistemas cognitivos de desempenho – sistema perceptual-articulatório (A-P) e sistema intencional-conceitual (C-I) –, com os quais a Faculdade da Linguagem interage, os quais são mediados por dois níveis de interface – a forma fonética (PF, do inglês *Phonetic Form*) e a forma lógica (LF, do inglês *Logical Form*), respectivamente. Na PF, o resultado da computação sintática se apresenta em termos fonéticos e deve fornecer instruções para o sistema perceptual-articulatório. Na LF, o resultado da computação linguística se apresenta em termos proposicionais e deve fornecer instruções para o sistema intencional-conceitual. O sistema computacional opera sobre os traços formais (traços categoriais (N e V), de gênero, número e pessoa (também chamados de traços *phi*) e os traços de Caso, pois é somente pela presença destes traços que ocorre a computação sintática. Os traços formais são aqueles que se realizam em termos do que é regular/sistemático na língua, como, por exemplo, as distinções morfológicas de número e gênero presentes nos elementos funcionais, como afixos e determinantes em português (como os presentes nos estímulos da atividade experimental desta dissertação). O gênero, portanto, é um traço formal – um traço *phi*, bem como número e pessoa. Somente os traços formais são acessíveis ao Sistema Computacional e podem ser passados à Forma Lógica (LF) para serem interpretados. Embora os traços formais sejam compartilhados entre as línguas, eles variam entre si e é possível considerar que a variação entre as línguas está diretamente relacionada às propriedades de seus traços formais, mais especificamente nas propriedades dos traços formais das categorias funcionais.

Assim, operações do sistema computacional (*Select*, *Merge*, *Agree* e *Move*) formam estruturas sintáticas a partir de elementos do léxico com base em informação gramatical codificada nos traços formais. Por meio de *Select*, itens da Numeração (escolha do léxico) são selecionados, organizando-se os feixes de traços em itens lexicais. Agregando elementos de forma recursiva, a operação

Merge possibilita a construção de estruturas sintáticas hierárquicas⁶. A operação *Agree* permite a checagem dos traços formais não-interpretáveis, tomando como base seus correlatos interpretáveis e a valoração dos traços dos itens lexicais já concatenados. No momento em que a estrutura sintática é construída, os traços semânticos passam por operações semânticas que resultam em LF e os traços fonológicos dos elementos do léxico envolvidos na derivação passam por processos fonológicos que resultam em PF. A passagem da informação do sistema computacional para as interfaces com os sistemas de desempenho é denominada *Spell Out*.

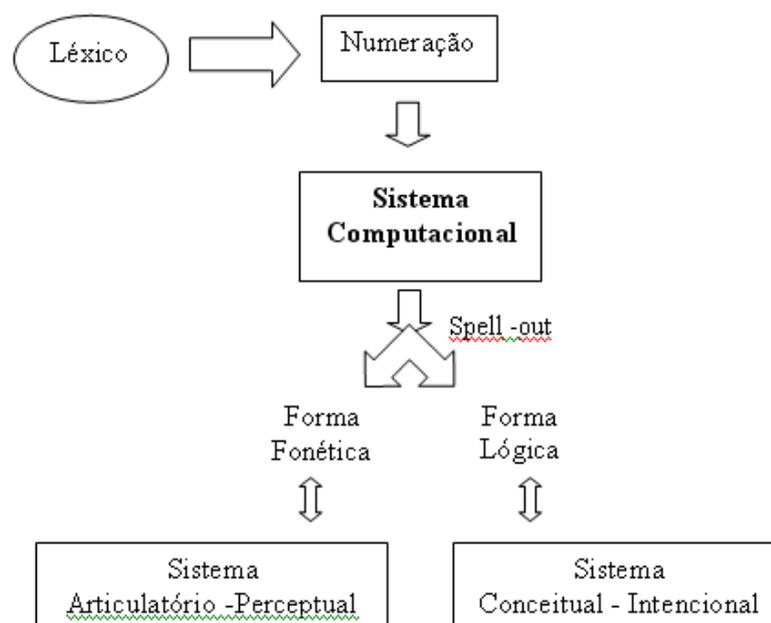


Figura 2 – O Programa Minimalista. Adaptada de CORRÊA (2008)

Sobre a vinculação da variação paramétrica das categorias funcionais cabe ressaltar que, sob a perspectiva do Programa Minimalista, os traços podem ser caracterizados em função de sua interpretabilidade e de sua natureza intrínseca ou opcional. Como vimos, o traço de gênero é opcional em alguns elementos, como

⁶ Chomsky (1995) considera que o Axioma da Correspondência Linear (Kayne, 1994) é uma condição que opera no componente fonológico, em decorrência da necessidade, imposta pela FORMA FONÉTICA, de linearização de uma estrutura, ou seja, de os sons da fala serem articulados num *continuum* temporalmente definido (linearmente, da esquerda para a direita).

Determinantes e Adjetivos, e sempre interpretável nos Nomes, porém nestes podem ainda ser intrínsecos, ou seja, especificados na entrada lexical, – por exemplo, em mesa, ônibus, homem, mulher, ou opcionais, especificados na *Numeração* – como em amigo(a), esposo(a). Assim, em termos gerais, categorias lexicais teriam traços de gênero [+ interpr] e categorias funcionais traços [- interpr]. Isso significa que um traço não-interpretável precisa ser checado (valorado) e apagado no curso da derivação sintática. Além disso, esse traço pode ser forte ou fraco. Se for forte, a checagem vai ocorrer na sintaxe sendo marcada morfofonologicamente no material enviado à Forma Fonética. Entretanto, se o traço for fraco, a checagem só ocorre na sintaxe encoberta, depois de *Spell Out* (sem repercussão morfofonológica). Utilizando como exemplo o Português do Brasil, o traço de gênero seria [+ interpr] no Nome e [- interpr] no Determinante. Dessa forma, o traço no Determinante precisaria ser valorado e apagado. Como o traço é forte, a checagem aconteceria na sintaxe aberta, ou seja, haveria o envio de material morfofonológico à Forma Fonética (PF), que seria lido pelo sistema Articulatorio-Perceptual (A-P), havendo assim a concordância de número e gênero entre Det e Nome (Name, 2002).

Dessa forma, a tarefa da criança, em fase de aquisição de uma língua, seria identificar, do conjunto de traços universais, os que são linguisticamente relevantes e selecionar os que pertencem à sua língua. Estes deverão ser associados a categorias do léxico, as quais deverão ser identificadas no *continuum* sonoro da fala do adulto em seu ambiente linguístico. A criança deverá, também, ser capaz de distinguir o que se constitui como traços formais, por terem implicações na sintaxe, considerando-se que, de acordo com essa proposta, o sistema computacional “atua” sobre eles, por meio de operações que não só combinam elementos de forma recursiva e estruturada (*Merge*), como também ajustam valores a eles conferidos por meio de concordância (*Agree*).

Considerando o tema focalizado neste trabalho, a percepção e distinção entre itens lexicais e funcionais no âmbito do DP, presente no *continuum* sonoro, favorecem a categorização dos elementos (Determinante e Nome) e, além disso, a identificação de itens funcionais por meio de informação de ordem prosódica e distribucional é condição necessária para que o processamento sintático seja iniciado (Corrêa, no prelo). Tais processos ainda favorecem a subcategorização dos

traços formais de gênero (masculino ou feminino), assumindo que o valor do traço de gênero no Nome é identificado, dentro do DP, com base em informação morfofonológica dos elementos da categoria D (NAME, 2002).

3.3 A PROPOSTA DE INTEGRAÇÃO ENTRE O MODELO LINGUÍSTICO (PROGRAMA MINIMALISTA) E O MODELO PSICOLINGUÍSTICO (*BOOTSTRAPPING* FONOLÓGICO)

Com a discussão do modelo de processamento psicolinguístico do *Bootstrapping fonológico* e do modelo da teoria linguística do *Programa Minimalista* nas seções anteriores, surgem-nos algumas questões que se tornam relevantes e intrigantes para nosso entendimento sobre o processo de aquisição da linguagem. O que desencadearia o funcionamento do sistema computacional? Partindo do pressuposto de que a Faculdade da Linguagem seja um aparato inato e comum a todo e qualquer ser humano, independentemente da língua no qual está inserido, como explicar a formação do Léxico, onde se encontrariam as especificidades de cada língua?

Como dito anteriormente, o modelo do *Bootstrapping* Fonológico, se tomado como única explicação para a aquisição e processamento linguístico, não consegue explicar a transposição do nível fonético-fonológico para o nível de representação formal no processamento da sintaxe (CORRÊA, 2008). De forma semelhante, o Programa Minimalista não dá conta, por si só, de explicar o desencadeamento do sistema computacional, ou seja, a formação inicial de um léxico durante a aquisição de linguagem, e a dinâmica de funcionamento do processamento linguístico. Assim, a Teoria Gerativa, ao postular a existência de uma predisposição biológica para a aquisição da linguagem e assumir a Faculdade da Linguagem como um sistema cognitivo interagindo com outros sistemas cognitivos e perceptuais, admite a associação entre este modelo de língua e um modelo de processamento, como o do *bootstrapping* fonológico, descrito acima.

De acordo com a proposta de integração entre o Programa Minimalista e o *Bootstrapping* Fonológico, o *input* linguístico não é opaco, pois, em termos gerais, a

informação prosódica disponível desde muito cedo para a criança em processo de aquisição desencadearia o funcionamento desse sistema computacional inato. Assim, a partir da identificação de padrões fonológicos e prosódicos da língua pelo bebê, haveria a formação mínima do Léxico (sem necessariamente o Sistema Computacional funcionar). A formação desse léxico mínimo, contendo elementos de categorias lexicais e funcionais subespecificados, com um número mínimo de traços formais, permitiria o desencadeamento da aquisição, desencadeando o *parsing* sintático inicial e viabilizando o funcionamento do Sistema Computacional, independentemente das distinções conceptuais-intencionais que possam expressar (CORRÊA, 2006). A gramática a ser identificada pela criança estaria inserida nos traços formais e, a partir dos padrões prosódicos e fonotáticos do fluxo da fala, a criança teria que identificar, na interface fônica, o que há de regular, sistemático na língua, baseando-se na frequência, regularidades, padrões prosódicos e fonéticos.

Em termos gerais, assumindo-se que mecanismos estatísticos e de abstração e generalização são recursos precocemente disponíveis e explorados por bebês no processo de categorização de elementos lexicais, a associação entre os dois modelos (modelo de língua e modelo de processamento) surge como possível investigação das questões suscitadas sobre o problema da aquisição da linguagem. Ressaltando a importância das DNA no processo de estabelecimento das relações sintáticas de uma língua e ainda na subcategorização dos elementos do léxico, considera-se que a informação acústica disponível desde muito cedo para a criança em processo de aquisição desencadearia o funcionamento desse sistema computacional inato, pois ainda que o bebê nasça com uma predisposição para a linguagem, é preciso que ele capte propriedades da língua de seu meio. Isto é, tal proposta busca explicitar o modo como a criança extrai do *continuum* sonoro, além de pistas distribucionais e prosódicas, propriedades dos traços formais dos elementos funcionais nos quais o Sistema Computacional age, resultando na derivação linguística.

No próximo capítulo, descreveremos a atividade experimental realizada que teve como objetivo verificar a sensibilidade a dependências não-adjacentes em crianças em fase inicial da aquisição lexical.

4 ATIVIDADE EXPERIMENTAL

Como dito anteriormente, na introdução, a proposta desta dissertação é um desdobramento dos resultados encontrados por Name (Name, Shi & Koulaguina, 2011; Name & Shi, em prep.). Bebês, com idade média de 11 meses, inseridos naturalmente desde o nascimento à língua inglesa ou francesa, a partir de curta exposição a uma língua pseudonatural diferente de sua língua materna – no caso, o PB, foram capazes de reconhecer padrões de relações entre D e N e, ainda, de abstrai-los de modo a identificá-los em novos enunciados (NAME, SHI & KOULAGUINA, 2011). Com o intuito de contrastar os resultados, considerando que o padrão da língua pseudonatural utilizada na atividade, apesar de não ser especificamente o padrão do PB (ver seção 2.3) é mais semelhante a este que a um padrão do inglês e do francês, buscamos verificar se crianças brasileiras apresentam resultados semelhantes. Dessa forma, foram utilizados a mesma técnica e os mesmos estímulos do experimento com crianças canadenses, que serão detalhados nas seções a seguir.

Para avaliar se aos 11 meses a criança é capaz de identificar uma propriedade comum às línguas naturais através da percepção de DNA nos estímulos linguísticos a que tem acesso durante um breve espaço de tempo, buscamos (1) replicar os resultados encontrados com bebês canadenses e (2) avaliar se a exposição prévia a um padrão semelhante ao de sua língua materna interfere no reconhecimento, abstração e generalização de DNA por bebês brasileiros.

4.1 TÉCNICA EXPERIMENTAL – OLHAR PREFERENCIAL

Em estudos com crianças mais jovens (fase pré-verbal) que apresentam como objetivo principal verificar a percepção da criança em relação ao estímulo linguístico, ou seja, a partir dos chamados “dados primários”, utilizam-se técnicas experimentais específicas, das quais se destacam a **Técnica de sucção não**

nutritiva, Técnica de escuta preferencial, Fixação preferencial do olhar⁷, entre outras.

Por ser um procedimento adequado para avaliação de habilidades perceptuais em crianças com até os 18 meses, a técnica escolhida para a aplicação do experimento é uma variação da **Técnica de Escuta Preferencial**.

Na Técnica de Escuta Preferencial, a criança é exposta a estímulos auditivos em eventos diferenciados em função de variáveis previamente selecionadas durante um período de tempo (tradicionalmente, 2min, podendo variar em função do objeto de pesquisa). É manipulada uma única variável em dois níveis. Na fase de teste, apresentam-se à criança dois tipos de estímulos distintos, coerentes ou não com o que foi anteriormente apresentado. Os estímulos são apresentados em áudio, por meio de dois alto-falantes situados à esquerda ou à direita da criança. O experimentador cronometra o tempo de escuta da criança - o tempo que ela se volta para o lado de onde vem o som do alto-falante (a variável dependente: o tempo de escuta, i.e., a medida comportamental do sujeito).

A variação dessa técnica, usada em diversos laboratórios psicolinguísticos e que foi usada nesta pesquisa, é a técnica do **Olhar Preferencial** (*Preferential Looking*). O procedimento é basicamente o mesmo. A principal mudança é o uso de apenas um alto-falante, centralizado junto a uma tela maior, o que permite a simplificação do *software* (*Habit*⁸) usado sem perder a confiabilidade dos resultados. Uma câmera escondida filma o olhar da criança e o experimentador mede o tempo do olhar que corresponde diretamente ao tempo de escuta do estímulo acústico. A medida pode ser definida como o olhar da criança ao estímulo visual apresentado.

⁷**Sucção Não-Nutritiva:** adequada para bebês recém-nascidos e bebês de até 4 meses, tem como objetivo geral avaliar as habilidades discriminatórias disponíveis em face de estímulos linguísticos que se manipulam em propriedades acústicas. É dada ao bebê uma chupeta ligada a um computador. As taxas de sucção das fases de habituação e teste são comparadas e, se houver um aumento estatisticamente significativo da frequência de sucção a partir da mudança de estímulo ocorrida no grupo experimental, esse aumento sugere que há discriminação, por parte do bebê, das diferentes realizações da variável manipulada. **Fixação preferencial do olhar:** adequado para crianças de 4 meses a 4 anos de idade, explora o fato de o bebê, desde muito cedo, expressar a percepção de algo novo, fixando o olhar por mais tempo no objeto ou evento que difere em algum aspecto daqueles a que se encontra habituado. Há duas versões: a intramodal em que se lida com uma única modalidade perceptual, e a intermodal – usada nos estudos de aquisição da linguagem – que explora a capacidade de o bebê relacionar estímulos captados por diferentes modalidades perceptuais e tomá-los como vinculados a um único evento. Para mais detalhamento das técnicas aqui citadas ver Name & Corrêa (2006).

⁸(Cohen, Atkinson & Chaput, 2000). Trata-se de um software gratuito que exige a plataforma MAC.

Entretanto, torna-se bastante importante ressaltar que não há variação da imagem apresentada, diferentemente dos estímulos acústicos, que variam de acordo com as condições de teste. Além disso, do lado de fora da cabine, o experimentador não sabe que tipo de estímulo a criança está escutando, de modo que não há qualquer interferência de sua parte, ainda que involuntária, na medição do tempo de fixação do olhar. Adiante, serão apresentadas as descrições de todas as etapas e os procedimentos que envolveram a atividade experimental.

4.2 MÉTODO

4.2.1 Participantes

Participaram da atividade experimental 23 bebês com idade média de 11 meses⁹. Entretanto, 4 bebês foram eliminados por inquietação (1), choro (2) e problemas técnicos durante a sessão (1). Assim, os resultados encontrados são referentes a 19 bebês, dentre os quais 12 são do sexo feminino e 7 são do sexo masculino.. Para assegurar a confiabilidade dos resultados, os 19 sujeitos foram divididos em dois grupos, como será visto posteriormente. Dessa forma, 11 crianças constituem o Grupo 1 e 8 constituem o Grupo 2. Os participantes foram recrutados pelos pesquisadores do NEALP e não receberam nenhuma retribuição financeira e/ou de outra ordem pela participação na atividade. Os responsáveis foram informados que a atividade é aprovada pelo Comitê de Ética da UFJF¹⁰, não oferecendo, portanto, nenhum prejuízo físico e/ou mental à criança e foram orientados a assinar um Termo de Consentimento livre e esclarecido¹¹ e a preencher um formulário - Cadastro de Crianças¹² - com as informações sobre o bebê.

⁹ Todos os bebês foram recrutados a partir da iniciativa da autora desta dissertação e de membros do NEALP.

¹⁰ A aprovação do Comitê de Ética. Parecer no. 100/2011.

¹¹ Anexo 1

¹² Anexo 2

4.2.2 Materiais utilizados

Os materiais utilizados para a realização da atividade experimental foram: poltrona preta individual, fone de ouvido, aparelho reproduzidor de som MP3, monitor 42" MAC, mesa de apoio coberta por tecido preto, filmadora Sony, caixa preta ("esconderijo" da câmera), Computador MAC APPLE G5, televisor, monitor Samsung 15" e teclado.

4.2.3 O ambiente

Durante a realização da tarefa, dois experimentadores participam de todo o processo. Antes da chegada do bebê, o laboratório é todo preparado a fim de oferecer maior conforto e comodidade para o responsável e, principalmente, para o bebê. O laboratório é localizado no Centro de Pesquisas e Humanidades, na Universidade Federal de Juiz de Fora, e foi especialmente adaptado para a realização das atividades. É localizado em um ambiente silencioso da Universidade e possui seu interior dividido em três ambientes: A antessala, que funciona como um ambiente de adaptação para o bebê, e duas cabines acústicas: a cabine de atividade experimental e a cabine de controle do experimentador.

4.2.4 O procedimento:

O responsável, ao chegar com o bebê, em horário pré-agendado, é recebido por dois pesquisadores na antessala do laboratório. Trata-se de um ambiente aconchegante tanto para os responsáveis quanto para os bebês. Há, neste ambiente, poltronas, uma mesa, um tapete emborrachado colorido e alguns brinquedos, como é possível visualizar na imagem abaixo:



Figura 3- Sala de ambientação onde a pessoa responsável e o bebê são recebidos.

Durante alguns minutos, a criança se familiariza com o ambiente e nós fornecemos algumas orientações sobre a atividade para o responsável. Solicitamos que o mesmo não interfira no comportamento da criança que deve estar sentada no colo do responsável, mas não é revelado o objetivo da atividade. Dentre as orientações, o responsável é informado de que a atividade é aprovada pelo Comitê de Ética da UFJF e orientado a assinar o termo de Consentimento e a preencher o formulário de cadastro com as informações sobre o bebê. Assim que verificamos que a criança já se ambientou ao recinto e, nesta hora, estando presente apenas um pesquisador, enquanto o outro já se encontra posicionado na sala de controle, a mãe é convidada para adentrar a sala de experimento onde se encontram “visíveis” apenas uma poltrona e um monitor (distância de 1,5 m entre poltrona e monitor). Além desses dois equipamentos, encontram-se, abaixo do monitor, uma câmera que fornece as imagens ao experimentador e uma caixa de som amplificada, coberta embaixo da toalha da mesa, como podemos visualizar na imagem abaixo:



Figura 4- Cabine de atividades experimentais onde as atividades foram realizadas

Assim que responsável e bebê se acomodam no ambiente, um dos experimentadores coloca um fone de ouvido no responsável, com uma música clássica, a fim de evitar que ouça os estímulos auditivos da atividade e interfira na atividade, mesmo de forma inconsciente. O experimentador assistente se retira da sala e o experimentador que marcará o tempo de olhar da criança inicia a execução do programa. Vale ressaltar que o experimentador foi o mesmo em todas as atividades realizadas com todas as crianças, não existindo, portanto, diferenças em relação ao modo de aplicação da técnica. O experimentador, através da visualização do olhar da criança, o que é disponibilizado a partir de um monitor instalado na cabine de observação, “registra”, através do software Habit e do computador, o “tempo de olhar” de cada estímulo. Abaixo, encontra-se uma imagem da cabine onde o experimentador permanece durante toda a atividade:

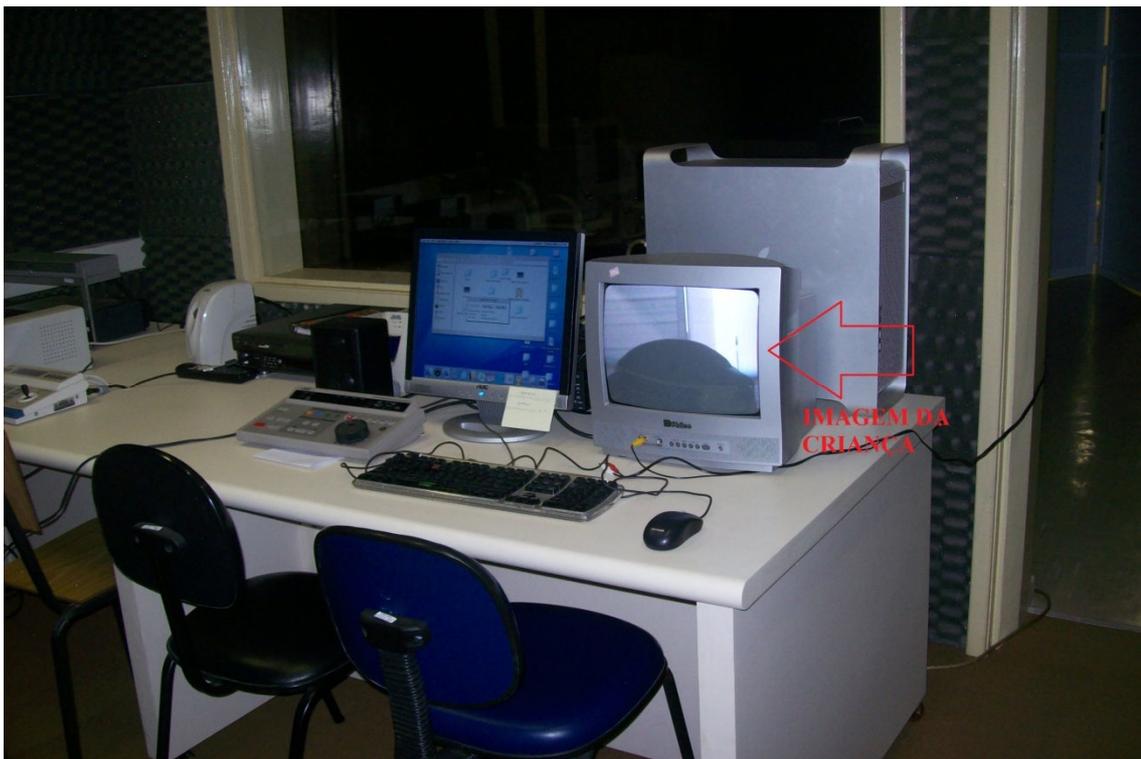


Figura 5- Foto da cabine de controle do experimentador.

Na realização da atividade experimental, buscou-se verificar o quanto a criança é perceptualmente sensível às DNA presentes nos enunciados linguísticos à sua volta e, ainda, se é capaz de abstrair os padrões de DNA identificados e reconhecê-los em novos enunciados. Dessa forma, a partir de um *design* intra-sujeitos, podemos definir as variáveis, as condições, a hipótese e a previsão:

Variável independente: a combinação entre Determinantes e pseudonomes nas DNAs.

Variável dependente: tempo de escuta dos estímulos ao longo do teste (congruentes ou incongruentes aos estímulos nos quais a criança foi familiarizada).

Condições Experimentais:

- GRAMATICAL: Os ensaios da fase teste, na condição gramatical, eram congruentes aos estímulos da fase de familiarização.

- AGRAMATICAL: Os ensaios da fase teste, na condição agramatical, eram incongruentes aos estímulos da fase de familiarização.

Hipótese: Bebês brasileiros, aos 11 meses, são capazes de abstrair e generalizar os padrões das relações entre DNA a novas instâncias a partir da exposição em um curto período de tempo a uma língua pseudonatural, diferente de sua língua materna.

Previsão: Baseada na hipótese que bebês, durante o primeiro ano de vida são providos de uma capacidade de tratamento do estímulo linguístico que lhes permite identificar, rapidamente, combinações recorrentes – particularmente DNA – abstrair seus padrões e generalizá-los em novos enunciados, a previsão é de que, se os bebês abstraírem o padrão da “nova” língua apresentada, eles “estranharão” os estímulos na condição agramatical, escutando “por mais tempo” os *ensaios* na condição agramatical dos que os *ensaios* na condição gramatical. Dessa forma, espera-se que haja diferença estatisticamente significativa da média do tempo de escuta entre as duas condições.

4.2.5 Estímulos

A atividade experimental explorou a sensibilidade de crianças aos 11 meses a DNA. O âmbito do DP foi o “lugar” escolhido para a construção dos estímulos, explorando as DNA entre Determinantes e as terminações dos Nomes. A fim de se evitar identidade fonológica entre Det e N foram escolhidos quatro variados determinantes masculinos do PB (*nosso, esse, meu e seu*) e criadas 16 pseudopalavras que preservam padrões alternativos para nomes masculinos.

Decidiu-se por usar pseudonomes com terminações diferentes da forma não-marcada (-o) a fim de evitar que frequência de palavras e terminação preferencial de nomes masculinos sejam variáveis não controladas no experimento. Assim, foram utilizadas pseudopalavras paroxítonas dissílabas terminadas em -a e -e: *bafe* [ˈbafi]; *mepe* [ˈmɛpi]; *guque* [ˈguki]; *pôfe* [ˈpovi]; *dabe* [ˈdabi]; *quive* [ˈkivi]; *zobe* [ˈzòbi]; *nêque* [ˈneki]; *ziba* [ˈziba]; *tafa* [ˈtafa]; *nôca* [ˈnoka]; *vêba* [ˈveba]; *bapa* [ˈbapa]; *goca* [ˈgòka]; *mefa* [ˈmɛfa]; *puva* [ˈpuva]. Dentre as pseudopalavras, 8 apresentavam terminação em -e e 8 apresentavam terminação em -a. Torna-se importante destacar que todos os estímulos auditivos, de todas as fases, foram gravados por uma falante nativa do PB e que determinantes e pseudonomes foram gravados separadamente, a fim de se evitarem diferenças entre eles devido às coarticulações. Com o objetivo de produzir uma articulação natural, foi solicitado que a brasileira pronunciasse o determinante imaginando, sem pronunciar, um nome na sequência. Entre determinantes e nomes, foram inseridos 200ms de pausa, e entre cada DP, 1 segundo de pausa silenciosa. Durante esta fase, a imagem da tela se mantinha constante (imagem de ondas verdes e azuis).

4.3 O EXPERIMENTO

O experimento é constituído de quatro fases consecutivas que serão, abaixo, detalhadas. Essas fases são necessárias para permitir um maior controle dos fatores que poderiam influenciar o resultado, como por exemplo, garantir a “atenção” do bebê.

1- Habituação

A fase de habituação consiste em um “aquecimento”. Nesta fase é apresentado auditivamente para a criança um estímulo contínuo de uma pequena história com prosódia dirigida à criança. O estímulo auditivo dura, em média, 7 segundos e é apresentado simultaneamente com um estímulo visual (o mesmo que será utilizado nas fases de familiarização e teste):

“Hoje é dia de passear no parque com todas as crianças. Quem quer vir comigo? Que festa maravilhosa!”

2- Familiarização

A fase de familiarização consiste em um áudio contínuo de 2 minutos constituído de sequências de DPs formados por determinantes e pseudonomes. Para tal, foram criadas 16 pseudopalavras dissílabas e paroxítonas. A fim de se evitar a preferência por determinada combinação, foram criadas duas “combinações” diferentes na fase de familiarização e as crianças que participaram da atividade foram divididas em dois grupos. Dessa forma, um grupo de crianças (11) foi familiarizado somente com o padrão “nosso, seu + -e e esse, meu + -a” (Grupo 1) e outro grupo de crianças (8) foi familiarizado somente com o padrão “nosso, seu + -a e esse, meu + -e” (Grupo 2). As combinações entre Det e N criadas para a fase de familiarização dos dois grupos encontram-se no seguinte quadro:

G R U P O	DPs: Det+X-e Nome		DPs: Det+X-a Nome	
	Determinante	Pseudo Palavra	Determinante	Pseudo Palavra
1	<i>nosso</i> [‘nòsu]	<i>bafe</i> [‘bafi] <i>mepe</i> [‘mèpi] <i>guque</i> [‘guki] <i>pôfe</i> [‘povi] <i>dabe</i> [‘dabi] <i>quive</i> [‘kivi] <i>zobe</i> [‘zòbi] <i>nêque</i> [‘neki]	<i>esse</i> [‘esi]	<i>ziba</i> [‘ziba] <i>tafa</i> [‘tafa] <i>nôca</i> [‘noka] <i>vêba</i> [‘veba] <i>bapa</i> [‘bapa] <i>goca</i> [‘gòka] <i>mefa</i> [‘mèfa] <i>puva</i> [‘puva]
G R U P O	DPs: Det+X-a Nome		DPs: Det+X-e Nomes	
	Determinante	Pseudo Palavra	Determinante	Pseudo Palavra
2	<i>nosso</i> [‘nòsu]	<i>ziba</i> [‘ziba] <i>tafa</i> [‘tafa] <i>nôca</i> [‘noka] <i>vêba</i> [‘veba] <i>bapa</i> [‘bapa] <i>goca</i> [‘gòka] <i>mefa</i> [‘mèfa] <i>puva</i> [‘puva]	<i>esse</i> [‘esi]	<i>bafe</i> [‘bafi] <i>mepe</i> [‘mèpi] <i>guque</i> [‘guki] <i>pôfe</i> [‘povi] <i>dabe</i> [‘dabi] <i>quive</i> [‘kivi] <i>zobe</i> [‘zòbi] <i>nêque</i> [‘neki]

Quadro 2- Combinações entre Det e N criadas para a fase de familiarização.

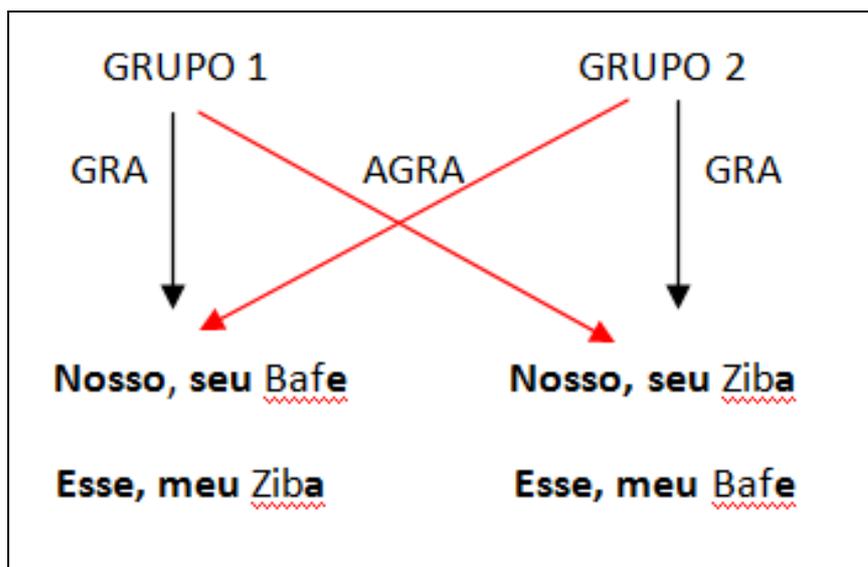
Durante esta fase, não nos era importante o rastreamento do comportamento visual da criança, pois o importante era a exposição ao novo padrão de língua. Assim, durante os dois minutos de exposição ao novo padrão de língua, a criança estava diante dos estímulos contendo as relações de DNA, os DPs eram apresentados consecutivamente e eram separados entre eles por uma pausa de 1 segundo. Cada uma das 16 palavras aparecia 2 vezes, cada vez precedida por um determinante diferente. Assim, “bafe”, no grupo 1, apareceria tanto antecedido pelos determinantes “nosso” quanto por “seu”, enquanto que, no grupo 2, apareceria tanto antecedido por “esse” quanto “meu”.

3- Fase teste

Logo após os 2 minutos de áudio da familiarização, iniciava-se a fase teste. Para essa fase, foram utilizados os mesmos determinantes da fase de familiarização – *seu*, *nosso*, *esse*, *meu* – com 4 novas pseudopalavras seguindo o padrão das palavras apresentadas na fase anterior. Foram, dessa forma, gerados 12 ensaios, cada um contendo 6 DPs. Em cada *ensaio*, os quatro determinantes eram combinados com duas novas palavras. Pausas de 200ms e 1 segundo foram inseridas entre os determinantes e nomes e entre os DPs, respectivamente. Cada *ensaio* apresentava duração média de 16 segundos.

Duas condições experimentais foram criadas: CONDIÇÃO GRAMATICAL e CONDIÇÃO AGRAMATICAL. Na condição gramatical, a combinação entre “Det + terminação do novo Nome” seguia o padrão apresentado anteriormente na fase de familiarização, enquanto que na condição agramatical o padrão não era congruente ao apresentado na fase de familiarização. Dessa forma, no grupo 1, na condição gramatical, *nosso* e *seu* eram combinados com pseudonomes com terminação –e e *esse* e *meu* eram combinados com novos pseudonomes com terminação –a, enquanto que, na condição agramatical, *nosso* e *seu* eram combinados com novos pseudonomes com terminação –a e *esse* e *meu* eram combinados com novos pseudonomes com terminação –e. De forma análoga, no grupo 2, na condição gramatical, *esse* e *meu* eram combinados com novos pseudonomes com terminação –e e *seu* e *nosso* eram combinados com novos pseudonomes com terminação –a,

enquanto que na fase agramatical, esse e *meu* eram combinados com novos pseudonomes com terminação *-a* e seu e nosso eram combinados com novos pseudonomes com terminações *-e*. Tal combinação pode ser mais facilmente identificada no quadro abaixo:



Quadro 3- Combinações entre Determinantes e Nomes nos estímulos nas fases de familiarização e teste.

As crianças foram divididas entre os dois grupos e, em cada grupo, a ordem de apresentação dos *ensaios* em gramatical e agramatical também foi contrabalanceada. Vale ressaltar que, como dito anteriormente, o estímulo visual apresentado na fase teste era o mesmo utilizado também na habituação e na familiarização. Afinal, a imagem apresentada não poderia ser mais atrativa que o estímulo auditivo apresentado. Entre um ensaio e outro e também entre uma fase e outra (por exemplo, familiarização e habituação, habituação e testes) aparecia outra tela, denominada *attention getter*, com um som não linguístico e constante.

A técnica de Olhar Preferencial através da utilização do *Software Habit* permite maior controle da “marcação” do tempo de olhar que equivale ao tempo de escuta. Isso porque o software é programado para que se a criança “olhe” no início do estímulo por menos de dois segundos contínuos para a tela e desvia o olhar, o estímulo “cai” para a tela *attention-getter* em que há a mudança do estímulo visual

e, assim que a criança retorne seu olhar, o estímulo é repetido, voltando a imagem de teste (mesma da habituação e familiarização). Caso a criança “olhe” por mais de 2 segundos consecutivos e desvie o olhar sem retornar a atenção, o estímulo “cai”, passando a tela para o *attention getter* e mudando para o próximo estímulo, sem repetir o anterior.

4- Pósteste

Após a apresentação dos 12 ensaios da fase teste, um estímulo auditivo, com duração de aproximadamente 16 segundos, foi apresentado sequencialmente para a criança. O tempo de escuta desse áudio também foi computado e posteriormente comparado com o tempo de escuta dos *ensaios* testes, a fim de garantir que a criança tenha se mantido atenta até o final da atividade. Assim como na fase de habituação, uma pequena história com prosódia dirigida à criança era apresentada:

“Você viu quantos presentes a menina ganhou? Que festa maravilhosa! Ela ficou tão feliz! Hoje é dia de passear no parque com todas as crianças! Quem quer vir comigo?”

Após o término da atividade, o *software Habit* fornece ao experimentador o relatório de dados contendo o tempo médio de escuta de cada *ensaio* que é arquivado para posteriores consultas. Um ou os dois experimentadores vão ao encontro do responsável e do bebê e a atividade se encerra. Ao final da atividade, presentamos o bebê com um simbólico certificado de participação na atividade¹³.

¹³Anexo 3.

4.4 RESULTADOS E DISCUSSÃO

Os resultados obtidos na atividade experimental, detalhada anteriormente, podem ser visualizados abaixo¹⁴. As médias gerais dos tempos de escuta de cada condição podem ser visualizadas no gráfico abaixo:

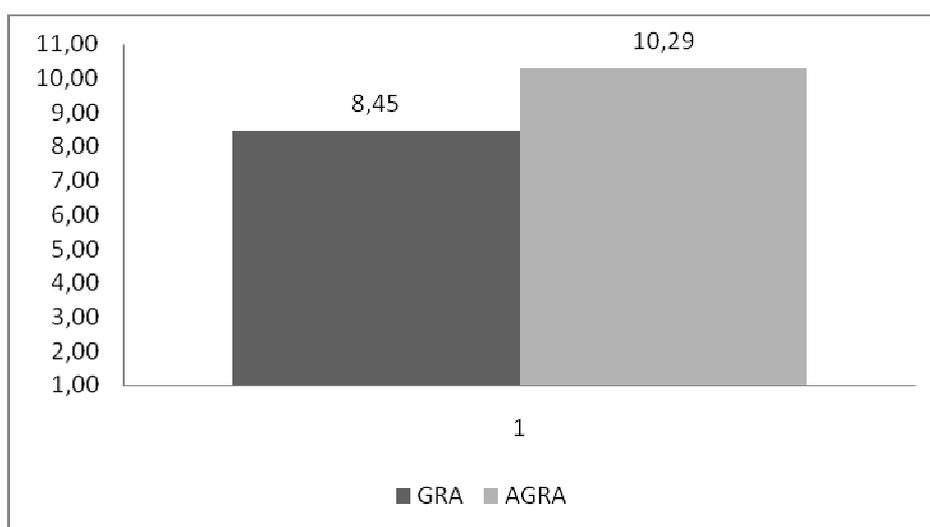


Gráfico 1- Comparação das médias totais de tempo de escuta, em segundos, entre as condições gramaticais e agramaticais.

Como podemos notar no gráfico acima, os resultados indicam que crianças se “atentaram” por mais tempo aos estímulos agramaticais do que aos estímulos gramaticais em relação ao padrão dos estímulos presentes na fase de familiarização. A média de tempo para a condição agramatical foi de 10,29 segundos, enquanto que, para a condição gramatical, foi de 8,45 segundos. Um teste estatístico foi realizado (*T-Student*) comparando-se os tempos médios de escuta das crianças nas duas condições e o resultado apresentou efeito significativo ($t(18)3.588$, $p=0.002$). Além do teste t, foi realizada uma análise da variância (ANOVA) mista, a qual apresentou o resultado significativo de $F(1,17) 10.08$ $p = .006$.

¹⁴ Para o tempo individual de cada criança, nas duas condições gramaticais, ver Anexo 4.

Análises estatísticas adicionais comparando-se o tempo médio de escuta nas duas condições em cada grupo, separadamente, revelaram diferença significativa no Grupo 1 ($t(10) 3.56, p=0.005$), e marginalmente significativa no Grupo 2 ($t(7)2.18, p=0.06$).O gráfico abaixo ilustra a diferença da média do tempo de escuta das duas condições gramaticais nos dois grupos.

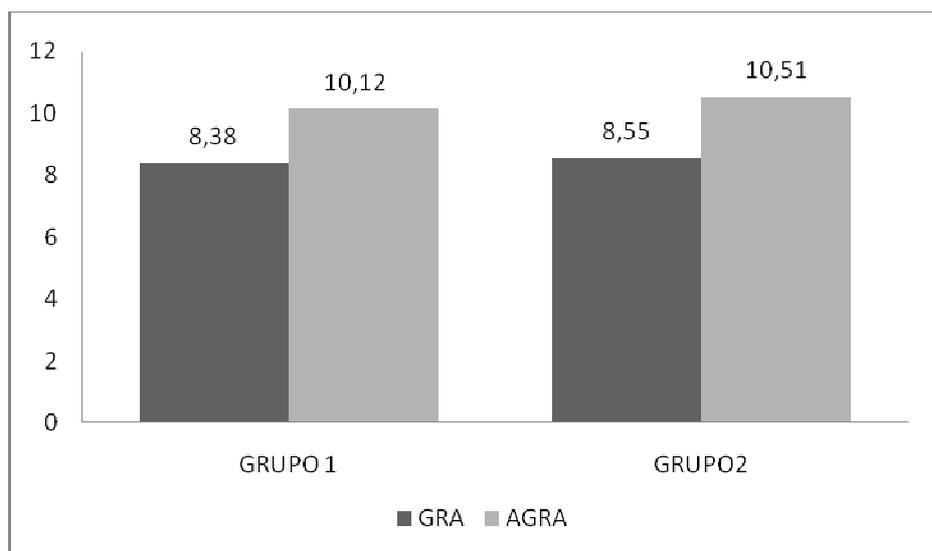


Gráfico 2- Comparação das condições gramaticais e agramaticais entre o Grupo 1 e o Grupo 2.

Acreditamos que esse resultado marginal no segundo grupo deve-se ao número ainda reduzido de crianças. De qualquer forma, o resultado, considerando-se os dois grupos, sinaliza uma reação ao tipo de DNA em função do que foi apresentado durante a familiarização. Ressalte-se que os dois grupos seguiram a mesma tendência, apresentando tempo maior de escuta para a condição incongruente em níveis semelhantes (8.38 e 8.55 sec na condição congruente; 10.13 e 10.51 sec na condição incongruente, Grupos 1 e 2, respectivamente).

Como vimos, os resultados vão ao encontro da hipótese inicial do experimento de que, diante de um curto período de exposição – 2 minutos – a uma língua pseudonatural diferente do padrão de sua língua materna, bebês, aos 11 meses, foram capazes de abstrair e generalizar um novo padrão de língua diferente do padrão de sua língua materna. Assim, DNA parecem favorecer o processo de categorização dos elementos, tal qual parece ter ocorrido no âmbito do DP entre Det

e Nomes. Além disso, notamos que a exposição prévia a um padrão semelhante ao de sua língua materna não interfere no reconhecimento, abstração e generalização de DNA por bebês brasileiros. Assim, bebês brasileiros também se mostraram eficientes na tarefa de abstrair um novo padrão – que mesmo contendo elementos de sua língua materna, apresenta um novo padrão de subcategorização – e ainda de generalizá-los a padrões nunca antes ouvidos. Os resultados parecem indicar que, assim como em curta exposição a uma língua pseudonatural (2 min.), os mecanismos de abstração e generalização parecem ser recursos/mecanismos utilizados por crianças no início do processo de aquisição lexical.

5 CONSIDERAÇÕES FINAIS

Esta pesquisa focalizou as etapas iniciais do processo de aquisição lexical, enfocando capacidades de abstração e generalização de padrões linguísticos por bebês durante o primeiro ano de vida e focalizou, mais especificamente, DNAs nas relações que se estabelecem entre determinantes (D) e nomes (N) no âmbito do DP.

A revisão bibliográfica acerca da divisão do léxico entre itens lexicais e funcionais destacou a importância dos itens funcionais e, mais detalhadamente, das DNAs no processo inicial de aquisição de uma língua materna. Em termos teóricos, buscou-se esclarecer como a conciliação entre um modelo de língua – Programa Minimalista –, que leva em conta a interface com sistemas conceituais e perceptuais, e um modelo de processamento – *Bootstrapping* Fonológico – propõe elucidar a passagem das propriedades extraídas do *continuum* sonoro para um nível de representação formal, ou seja, como a distinção entre elementos funcionais e lexicais, a partir da sensibilidade a padrões prosódicos e fonotáticos no primeiro ano de vida, dá origem a um léxico mínimo, constituído de elementos com traços subespecificados, que, por sua vez, vão desencadear a aquisição do léxico e da sintaxe.

A hipótese geral que norteou este estudo é a de que bebês durante o primeiro ano de vida são providos da capacidade de tratamento do estímulo linguístico que lhes permite identificar, rapidamente, combinações recorrentes – particularmente DNA – abstrair seus padrões e generalizá-los em novos enunciados. Nossas hipóteses específicas foram: (1) bebês brasileiros aos 11 meses, tal como os bebês canadenses, são capazes de abstrair e generalizar os padrões das relações entre DNA a novas instâncias a partir da exposição em um curto período de tempo à uma língua pseudonatural, que preserva propriedades prosódicas, fonotáticas e distribucionais de uma língua natural, no caso, o PB; e (2) bebês brasileiros aos 11 meses são capazes de identificar uma propriedade comum às línguas naturais através da percepção de DNA nos estímulos linguísticos a que tem acesso durante um breve espaço de tempo (2 minutos), ainda que tais estímulos apresentem

elementos que lhes sejam familiares (devido à exposição ao PB), mas em uma nova configuração.

Baseada nos resultados encontrados por Name (NAME, SHI & KOULAGUINA, 2011; NAME & SHI, em prep.) com crianças canadenses, uma atividade experimental foi realizada com o objetivo de verificar o quanto a criança é perceptualmente sensível à DNA presentes nos enunciados linguísticos à sua volta e ainda se é capaz de abstrair os padrões de DNA identificados e reconhecê-los em novos enunciados. Para a construção dos estímulos, foram construídos DPs com determinantes masculinos do PB e pseudopalavras dissílabas e paroxítonas que preservam padrões alternativos para nomes masculinos. Determinantes masculinos foram apresentados em dois conjuntos distintos (*nosso* e *seu*, de um lado, e *meu* e *esse* de outro) associando-se, cada um, a pseudonomes com uma determinada terminação (-a ou -e, de acordo com o grupo). Portanto, os padrões de DNA apresentados aos bebês brasileiros diferem do que já é conhecido, pois ainda que os determinantes lhes sejam familiares (devido à exposição ao PB), sua divisão em duas subclasses distintas foge ao padrão da língua (i.e., esses determinantes pertencem todos a uma única subclasse, de masculinos, podendo, todos eles, anteceder os mesmos nomes).

Como vimos, os resultados indicaram que crianças ouviram mais tempo os estímulos agramaticais do que os estímulos gramaticais em relação ao padrão dos estímulos presentes na fase de familiarização. Tais resultados vão ao encontro da previsão de que bebês aos 11 meses são sensíveis aos padrões recorrentes de DNA, sendo capazes de abstrair e ainda generalizá-los a novos DPs nunca antes ouvidos. Assim, diante de um curto período de exposição (2 minutos) a uma língua pseudonatural, bebês brasileiros, aos 11 meses, foram capazes de abstrair e generalizar um novo padrão de língua, tal como ocorreu com os bebês canadenses de mesma idade (NAME, SHI & KOULAGUINA, 2011).

Destaque-se que a situação apresentada não era exatamente a mesma para os dois grupos. Para os bebês canadenses, tratava-se de dois conjuntos de determinantes (*nosso* e *seu*, de um lado, e *meu* e *esse* de outro) que se associavam, cada um, a nomes com uma determinada terminação (-a ou -e, de acordo com o grupo). Para os brasileiros, no entanto, a situação era mais complexa,

visto que os determinantes já eram familiares (devido à exposição ao PB), mas não em subclasses distintas. Como sabemos, esses determinantes pertencem todos à subclasse de masculinos. Assim, sua apresentação em dois conjuntos de determinantes distintos, diferentemente do padrão de sua língua materna, era um desafio maior para esses bebês, que teriam de reconhecer novos padrões a despeito de uma falsa semelhança com o PB.

Nossos resultados sugerem fortemente que as capacidades de reconhecer padrões de relações entre D e N, de abstraí-los e generalizá-los, identificando-os em novos enunciados, são recursos disponíveis aos 11 meses de idade. Assim, a identificação e generalização de DNA parecem favorecer o processo de categorização dos elementos, tal como parece ter ocorrido no âmbito do DP entre Dets e Nomes, evidenciando uma eficiente capacidade de tratamento de estímulos linguísticos para uma rápida e eficiente aquisição lexical.

O presente estudo, como já foi observado anteriormente, busca contribuir para o entendimento dos processos mentais que subjazem as etapas iniciais do processo de aquisição da linguagem e de modo mais específico, a aquisição lexical. Vimos que bebês de 11 meses fazem uso de as habilidades de identificação, abstração e generalização de DNA diante de novos estímulos linguísticos. Mais ainda, observamos que o contato prévio com o PB não interferiu na capacidade dos bebês brasileiros em reconhecer as DNA apresentadas, que se distinguiam parcialmente do padrão do PB.

Como vimos, o presente estudo pode ser considerado um avanço nos estudos psicolinguísticos acerca da aquisição lexical e de abstração e generalização de padrões linguísticos recorrentes na língua. Os trabalhos existentes, em sua maioria, exploram tais padrões em línguas totalmente artificiais (MARCUS et al., 1999; GOMEZ, 2002; NEWPORT & ASLIN, 2004), sem levar em conta propriedades morfofonológicas e prosódicas, ou, quando trabalham com línguas naturais, investigam tais capacidades em crianças mais velhas (18 meses), considerando sobretudo o reconhecimento de padrões de sua língua materna, e não propriamente habilidades de abstração e generalização (SANTELMANN & JUSCZYK, 1998).

Esta dissertação não pretendeu esgotar todas as discussões apresentadas, pois diante dos resultados descritos neste trabalho, uma série de questões foram

traçadas: verificar se tais habilidades estão disponíveis em idade anterior aos 11 meses para o tratamento de língua pseudonatural, já que os resultados encontrados na literatura com bebês mais novos (7 meses) usaram língua artificial; questões relativas à influência da maior exposição das crianças brasileiras aos Dets na pseudolíngua e suas consequências: seriam as crianças brasileiras sensíveis às propriedades das DNA antes dos 11 meses por já “conhecerem” os determinantes, ainda que com diferentes subclassificações para os nomes? Questões relativas à sensibilidade de adultos a padrões de DNA em uma língua que não seja artificial, dentre outras. Diante do exposto, considera-se que tal área investigativa encontra-se em pleno potencial de desenvolvimento e ainda cheia de indagações a serem respondidas. Espera-se, com esta dissertação, ter contribuído para abrir os campos e os desafios de estudo da talvez mais fascinante capacidade humana: a capacidade de adquirir naturalmente uma e qualquer língua natural do meio no qual se encontra inserido.

REFERÊNCIAS

BAGETTI, Tatiana. **Um estudo experimental do processamento na interface fônica e da análise sintática inicial**: o papel de elementos funcionais na aquisição da linguagem. Tese (Doutorado), PUC RJ, 2009.

BAKER, M. C. **Lexical categories** - verbs, nouns and adjectives. Cambridge: Cambridge University Press, 2003.

CHOMSKY, N. Derivation by Phase. **MIT Working Papers in Linguistics 18. Cambridge, MA: MIT Working Papers in Linguistics, 1999.**

_____. (1995) **The Minimalist Program**. Cambridge, MA: MIT Press, 1995.

_____. (1970). Remarks on Nominalization. In Jacobs, R.; Rosenbaum, P. (eds.). **Readings in English Transformational Grammar**. Massachusetts: Ginn, Waltham, 184-221.

_____. (1965). **Aspects of the theory of syntax**. Cambridge, Mass.: MIT.

CHRISTOPHE, A.; GUASTI, T.; NESPOR, M.; DUPOUX, E.; VAN OUYEN, B. **Reflections on phonological bootstrapping**: its role for lexical and syntactic acquisition. *Language and Cognitive Processes*, v. 12, nº 5/6, p. 585-612, 1997.

COMRIE, B. **Language universals & Linguistic typology**. 2. ed. Chicago: The University of Chicago Press.

CORRÊA, L. M. S. (2008). O desencadeamento (*bootstrapping*) da sintaxe numa abordagem psicolinguística. In: Quadros, R. M. de. **Teorias de aquisição da linguagem**. Florianópolis: Ed. da UFSC, p. 169-220.

_____. (2006) Conciliando processamento lingüístico e teoria de língua no estudo da aquisição da linguagem: habilidades discriminatórias de bebês, categorias funcionais e a disponibilidade de um sistema computacional lingüístico. In : Corrêa, L.M.S. (org.). **Aquisição da Linguagem e Problemas do Desenvolvimento Lingüístico**. Rio de Janeiro: Editora da PUC-RJ, p. 21-78.

CUNHA, Celso & CINTRA, Lindley. **Nova Gramática do Português Contemporâneo**. 5 ed. Rio de Janeiro: Lexicon, 2008.

DUPOUX, E. MEHLER, J. Monitoring the lexicon with normal and compressed speech: Frequency effects and the prelexical code. **Journal of Memory and Language**, 29, 316-335, 1990.

FARIA, Flávia C. **O fenômeno da Concordância no processo de Aquisição do Pronome Possessivo por Crianças Adquirindo o PB**. Dissertação de Mestrado. UFJF, 2005.

GLEITMAN, L. (1990). The structural sources of verb meanings. **Language Acquisition**, v.1, p.3-55.

GÓMEZ, R.L. Variability and detection of invariant structure. **Psychological Science**, 13, 431–436, 2002.

GÓMEZ, R. L. ;LAKUSTA, L. A first step in form-based category abstraction by 12-month-old infants. **Developmental Science**, 7, 567-580, 2004

GOUT, A.; CHRISTOPHE, A. O papel do *bootstrapping* prosódico na aquisição da sintaxe e do léxico. In: CORRÊA, L.M.S. (Org). **Aquisição da linguagem e problemas do desenvolvimento linguístico**. Rio de Janeiro: Ed. PUC-Rio; São Paulo: Loyola, 2006.

HAUSER, M.; CHOMSKY, N.; FITCH, W. T. The Faculty of language: what is it, who has it, and how did it evolve? **Science**, v. 298, p.1569- 1579, 2002.

MARCUS, G.F.; VIJAVAN, S.; BANDI RAO, S.; VISHTON, P.M. Rule learning by seven-month-old infants. **Science**, 283, 77–80, 1999.

MIOTO, C. S.; SILVA, M. C. F.; LOPES, R. E. V. **Manual de sintaxe**. 2 ed. Florianópolis: Insular, 2005.

MORGAN, J. L.; DEMUTH, K. Signal to Syntax: an overview. In: J. L. MORGAN & K. DEMUTH (Orgs.), **Signal to Syntax: Bootstrapping from speech to grammar in early acquisition**. Mahwah, New Jersey: Lawrence Erlbaum Associates, 1-22, 1996.

MORGAN, J. L.; SHI, R.; ALLOPENNA, P., Perceptual Bases of Rudimentary Grammatical Categories: Toward a Broader Conceptualization of Bootstrapping. In: J. L. MORGAN & K. DEMUTH (Orgs.), **Signal to Syntax: Bootstrapping from speech to grammar in early acquisition**. Mahwah, New Jersey: Lawrence Erlbaum Associates, 263-286, 1996.

NAME, M. C. Bootstrapping sintático: o papel da ordem estrutural na aquisição de nomes e adjetivos. **Letras de Hoje**, v. 42, p. 53-63, 2007.

NAME, M.C. **Habilidades perceptuais e lingüísticas no processo de aquisição do sistema de gênero no português**. Tese (Doutorado) PUC RJ, 2002.

NAME, M. C. L. & CORRÊA, L. M. S. Explorando a escuta, o olhar e o processamento sintático. In: In : CORRÊA, L.M.S. (org.). **Aquisição da Linguagem e Problemas do Desenvolvimento Lingüístico**. Rio de Janeiro: Editora da PUC-RJ, p. 79-100, 2006.

NAME, C.; SHI, R.; KOULAGUINA, E.; Non- adjacent dependency learning and grammatical categorization in 11- month-old infants. **12th International Congress for the Study of Child Language**.vMontreal, 2011.

NEWPORT, E.L.; ASLIN, R.N. Learning at a distance: I. Statistical learning of non-adjacent dependencies. **Cognitive Psychology**, 48, 127-162, 2004.

PINKER, S. Why the child holds the baby rabbits: A case study in language acquisition. In L. Gleitman, & M. Liberman (Eds.), **Invitation to Cognitive Science, 2nd Edition. Volume 1: Language**. Cambridge, MA: MIT Press, 1995.

_____. (1987) Productivity and Constraints in the Acquisition of the Passive. **Cognition**, 26(3): 195 – 267.

_____(1984) **Language learnability and language development**. Cambridge, MA: Harvard University Press.

SANTELMANN, L.; JUSCZYK, P. **Sensitivity to discontinuous dependencies in language learners**: evidence for limitations in processing space. *Cognition*, 69, 105-134, 1998.

SAUSSURE, Ferdinand de; Curso de Linguística Geral. São Paulo, Editora Cultrix, 2002

SHI, R.; WERKER, J.; MORGAN, J. Newborn infants' sensitivity to perceptual cues to lexical and grammatical words. **Cognition**, 72, B11-21, 1999.

SHI, R.; WERKER, J.; CUTLER, A. Function words in early speech perception. **15th ICPHS Barcelona**, 3009-3012, 2003.

SHI, R.; CUTLER, A.; WERKER, J.; CRUICKSHANK, M. Frequency and form as determinants of functor sensitivity in English-acquiring infants. **Journal of Acoustical Society of America**, 119(6), 2006.

ANEXOS

Anexo 1 – Termo de consentimento livre e esclarecido

TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO

Seu/sua filho(a) está sendo convidado(a) como voluntário(a) a participar da pesquisa “Etapas iniciais da aquisição lexical”. Neste estudo pretendemos observar o modo como os bebês adquirindo o português começam a segmentar a fala e a reconhecer palavras. O motivo que nos leva a estudar esse assunto é contribuir para o entendimento do processo de aquisição de vocabulário desde suas etapas iniciais.

Para este estudo adotaremos o(s) seguinte(s) procedimento(s): o bebê participará de uma atividade lúdica (uma “brincadeira”), durante a qual lhe apresentaremos imagens na tela do computador acompanhadas de frases curtas. Observaremos sua atenção e interesse aos estímulos apresentados. **A atividade não tem nenhum caráter de avaliação do desempenho e/ou de conhecimento da língua.** Seu único objetivo é observar o modo como a criança se relaciona com a língua em uma situação que simula uma atividade espontânea. A atividade dura cerca de 15 minutos e no total (desde a chegada do bebê, sua adaptação ao ambiente e saída) não ultrapassa 30 minutos.

Para seu/sua filho(a) participar deste estudo você não terá nenhum custo, nem receberá qualquer vantagem financeira. Você será esclarecido(a) sobre o estudo em qualquer aspecto que desejar e estará livre para participar ou recusar-se a participar. Poderá retirar seu consentimento ou interromper a participação a qualquer momento. A participação de seu/sua filho(a) é voluntária e a recusa em participar não acarretará qualquer penalidade ou modificação na forma em que é atendido(a) pelo pesquisador. O pesquisador irá tratar a sua identidade e a de seu/sua filho(a) com padrões profissionais de sigilo. Ele(a) não será identificado(a) em nenhuma publicação que possa resultar deste estudo. Este estudo apresenta risco mínimo, isto é, o mesmo risco existente em atividades rotineiras como conversar, tomar banho, ler, etc. Apesar disso, você tem assegurado o direito a ressarcimento ou indenização no caso de quaisquer danos eventualmente produzidos pela pesquisa.

Os resultados da pesquisa estarão à sua disposição quando finalizada. O nome de seu/sua filho(a) ou o material que indique a participação dele(a) não será liberado sem a sua permissão. Os dados e instrumentos utilizados na pesquisa ficarão arquivados com o pesquisador responsável por um período de 5 anos, e após esse tempo serão destruídos. Este termo de consentimento encontra-se impresso em duas vias, sendo que uma cópia será arquivada pelo pesquisador responsável, e a outra será fornecida a você.

Eu, _____, portador(a) do documento de Identidade _____, fui informado(a) dos objetivos do

presente estudo de maneira clara e detalhada e esclareci minhas dúvidas. Sei que a qualquer momento poderei solicitar novas informações e modificar minha decisão de autorizar meu/minha filho(a) participar se assim o desejar. Declaro que concordo em participar desse estudo. Recebi uma cópia deste termo de consentimento livre e esclarecido e me foi dada oportunidade de ler e esclarecer as minhas dúvidas.

Juiz de Fora, ____ de _____ de 20____ .

Assinatura do(a) participante

Assinatura do(a) pesquisador(a)

Em caso de dúvidas com respeito aos aspectos éticos deste estudo, você poderá consultar:

CEP- COMITÊ DE ÉTICA EM PESQUISA - UFJF
PRÓ-REITORIA DE PESQUISA / CAMPUS UNIVERSITÁRIO DA UFJF
JUIZ DE FORA (MG) - CEP: 36036-900
FONE: (32) 2102-3788 / E-MAIL: cep.propesq@ufjf.edu.br
PESQUISADOR(A) RESPONSÁVEL: MARIA CRISTINA LOBO NAME
ENDEREÇO: FACULDADE DE LETRAS – UFJF CAMPUS UNIVERSITARIO - MARTELOS
JUIZ DE FORA (MG) - CEP: 36036-300
FONE: (32) 2102.3150 / E-MAIL: CRISTINA.NAME@UFJF.EDU.BR

Anexo 2 – Cadastro de crianças NEALP

NEALP Universidade Federal de Juiz de Fora		
Cadastro de Crianças		
Projeto em Aquisição da Linguagem		
Nome da Criança: _____		
Data de nascimento da Criança: __/__/__	Sexo: () masculino () feminino	
Frequenta Creche ou escola? Qual? _____		
Nome do responsável: _____		
Profissão do pai/mãe: _____		
Telefone para contato: _____		Celular: _____
E-mail: _____		
Os pais da criança são falantes nativos do Português do Brasil? Sim () Não ()		
A criança viveu no exterior ou tem contato regular com falante de língua estrangeira (diferente do pai ou da mãe): Sim () Não ()		
A criança possui parentes próximos ou antepassados com qualquer problema de linguagem: Sim () Não ()		
Marque caso a criança tenha apresentado qualquer dos quadros abaixo:		
() otite de repetição		
() lesão cerebral		
() qualquer deficiência genética		
Este formulário é para facilitar o contato entre o NEALP e os pais interessados em nossa pesquisa. O preenchimento do mesmo NÃO IMPLICA qualquer engajamento da parte dos pais. Toda informação nele contida é confidencial e para uso exclusivo do NEAL		

Anexo 3 – Certificado simbólico de participação



O NEALP - Núcleo de Estudos
em Aquisição da Linguagem e
Psicolinguística da UFJF
certifica que

participou brilhantemente
das atividades de pesquisa.

Juiz de Fora,

_____ de _____ de 20_____.



Anexo 4 – Tempo individual da atividade experimental, nas duas condições, de cada criança

	GRAMATICAL	AGRAMATICAL
GRUPO 1		
bebê 1	12,3	13,8
bebê 2	9,8	12,3
bebê 3	5,9	4,61
bebê 4	12,2	13,7
bebê 5	3,6	7,9
bebê 6	6,3	8,7
bebê 7	9,8	7,8
bebê 8	7	9,6
bebê 9	7,7	10,4
bebê 10	10,6	13
bebê 11	7	9,6
GRUPO 2		
bebê 12	6,8	8,5
bebê 13	7,3	4,6
bebê 14	7,6	13,6
bebê 15	5,1	9,9
bebê 16	11,15	13,7
bebê 17	7,76	12,2
bebê 18	11,6	9
bebê 19	11,1	12,6