AQUISIÇÃO DE PALAVRAS COMPLEXAS NO PORTUGUÊS BRASILEIRO: A EMERGÊNCIA DE MORFOLOGIA DERIVACIONAL NA FALA INFANTIL

Christina Abreu Gomes (UFRJ)
christina-gomes@uol.com.br
Maria Fernanda Moreira Barbosa (UFRJ)
fernanda136@gmail.com

RESUMO

Neste trabalho, investiga-se, com base nos pressupostos dos modelos baseados no uso, os aspectos morfológicos e lexicais do processo de aquisição de linguagem, incidindo na aquisição da morfologia derivacional e, mais especificamente, na aquisição de palavras complexas por derivação na fala de crianças de 2: 0 a 5: 0 anos de idade no português brasileiro. Adotamos a perspectiva teórica fornecida pelos modelos baseados no uso, que compartilham a ideia de que não há uma separação estrita entre léxico e gramática. Desse modo, a gramática emerge das representações do léxico, isto é, ela não tem existência independente. A amostra é constituída de dados de fala espontânea de crianças de 2;0 a 5;0 anos de idade, oriundos da Amostra de Fala Infantil e da amostra AQUIVAR, ambas pertencentes ao acervo do Programa de Estudos sobre o Uso da Língua (PEUL/UFRJ), de dados provenientes do *corpus* FLORIANOPOLIS/CHILDES e de dados do projeto Avaliação Sonora do Português Atual (ASPA/UFMG) a fim de categorizarmos os afixos de acordo com suas frequências do tipo.

Palavras-chave: Aquisição de linguagem. Derivação. Formação de palavras. Afixação. Português brasileiro.

1. Introdução

Este estudo visa fornecer uma descrição sobre a aquisição da morfologia derivacional no português brasileiro, permitindo-se conhecer como se processa a aquisição de afixos derivacionais e como se caracteriza a expansão do léxico infantil. É nesta perspectiva que se insere o presente estudo, tendo como foco principal analisar a emergência de afixos deri-

vacionais em crianças de 2:0 a 5:0 anos de idade que estão adquirindo o português brasileiro.

O texto apresenta-se estruturado da seguinte maneira: na segunda seção, tem-se as principais descrições sobre a aquisição de palavras complexas na literatura; Na terceira seção, apresenta-se um panorama dos modelos baseados no uso; Na quarta seção, especifica-se a metodologia empregada na coleta de dados; Na quinta seção, descreve-se os resultados preliminares obtidos a partir da coleta de dados, segundo o modelo teórico adotado neste estudo; Por fim, têm-se as considerações finais.

2. Revisão da literatura

Para Clark (2001), as crianças começam a produzir as primeiras palavras entre os 12 e 20 meses de idade. Nessa fase, elas produzem variações morfológicas sistemáticas de tais palavras. Logo, a partir do momento em que as crianças começam a articular suas primeiras palavras, seu vocabulário é constantemente expandido, tendo a habilidade de reconhecer cerca de 500 palavras ao completarem 2 anos de idade. Daí em diante, elas passam a adquirir cerca de 10 novas palavras a cada dia, ampliando consideravelmente seu vocabulário e alcançando uma média de 14.000 palavras ao completarem 6 anos de idade (CAREY, 1978, p.265) e eventualmente podem atingir de 20.000 a 50.000 palavras, mesma faixa de variação que um falante adulto tem prontamente à sua disposição (NATION, 1993; CLARK, 1993; AITCHISON, 1994). E antes da fase adulta, a expansão de seu vocabulário já está em torno de 50.000 a 100.000 palavras (CLARK, 2009).

Clark (2009) ressalta que as crianças percebem as palavras ao observarem o fluxo de conversação entre adultos e crianças (p. 283). Com isso, elas inferem os possíveis significados de palavras até então desconhecidas, a partir do modo como os adultos as empregam no discurso. Desse modo, o adulto serve de fonte-alvo para o *input* da criança e ela orienta-se na tentativa de comunicar-se seguindo as mesmas estratégias utilizadas pelos adultos. Sendo assim, o adulto emprega uma nova palavra e a criança observa o contexto no qual o vocábulo desconhecido foi utilizado. Posteriormente, ela formula a hipótese a respeito do conceito que possivelmente aquela palavra corresponde e vai confirmando, reforçando, modificando ou rejeitando a hipótese inicial à medida que observa outras instâncias de uso para o mesmo vocábulo (BLOOM, 1994, p. 11).

Segundo Nagy & Anderson (1984), a capacidade de interpretar as palavras com base na análise morfológica pode ser explicada pela compreensão de leitura que acontece nos primeiros anos escolares. Nesse sentido, Wysocki & Jenkins (1987) realizaram um experimento com crianças falantes de inglês da 4ª, 6ª e 8ª série do ensino fundamental. Esses alunos receberam três lições de 15 a 20 minutos que eram compostas por palavras de baixa frequência, apresentadas isoladamente e contextualizadas em sentenças. Em seguida, foi solicitado ao professor que escrevesse tais palavras e suas definições no quadro negro e então os alunos liam em voz alta os respectivos vocábulos e suas definições. Quando as crianças já estavam familiarizadas com as palavras, o professor as apagava. Posteriormente, com o intuito de avaliar o conhecimento das crianças, foi-lhes pedido para produzirem a definição de determinada palavra ou para produzirem a palavra dada a sua definição. Em um segundo momento, foi-lhes ensinado que a partir de um vocábulo primitivo formavam-se novas palavras derivadas (p. ex.: gratuito para gratuidade), a fim de analisar se as generalizações morfológicas haviam ocorrido. Desse modo, os autores verificaram que, em todas as séries abarcadas pelo experimento, as generalizações morfológicas foram constatadas e os novos conhecimentos da consciência morfológica[34] foram utilizados para definir as palavras desconhecidas. Com isso, a pesquisa demonstrou que as crianças têm habilidade para decompor os vocábulos em morfemas como também que a consciência morfológica auxilia no processo de alfabetização nos primeiros anos escolares.

De acordo com Nunes e Bryant (2006), a compreensão de uma nova palavra é facilitada se a criança percebe que a mesma é formada por uma combinação de vários morfemas, cada um com o seu próprio significado, que se unem para formar esse novo vocábulo. Assim, se a criança conhece o significado do vocábulo "legal", poderá inferir também o significado de "legalmente" e "ilegal", por exemplo.

Os trabalhos de Nagy & Anderson (1984), de Wysocki & Jenkins (1987) entre outros, indicam que as crianças apresentam um conhecimento intuitivo e implícito da estrutura morfológica das palavras, ou seja, esse conhecimento é utilizado de forma inconsciente e não intencional. De um lado, observa-se que esses estudos foram realizados com crianças de

[34] A consciência morfológica corresponde à habilidade de compreender e manipular pequenas partes significativas que compõem as palavras, ou seja, a menor parte indivisível e significativa da palavra (morfema) como prefixos, radicais e sufixos, por exemplo.

ensino fundamental a partir dos 9 anos de idade, visto que esses autores argumentam que o processamento da morfologia derivacional começa tardiamente e está relacionado ao desenvolvimento da aprendizagem da leitura e da escrita.

Por outro lado, tal como é defendido por Clark (1993, 2001, 2009), à medida que a criança aprende mais palavras, as memoriza e as produz em seu discurso, torna-se capaz também de analisar a sua estrutura morfológica interna. Neste ponto, é pertinente destacar o fato de que as crianças começam a usar alguns sufixos derivacionais para construírem novos nomes por volta de 2 anos de idade. A partir desta idade, Clark (2009) ressalta que os primeiros sufixos tipicamente usados no Inglês são o agentivo *-er* e o sufixo diminutivo *-ie* e antes mesmo de completarem 3 anos já fazem uso do prefixo *un-*. Assim, as crianças começam a fazer uso de radicais e afixos em formas derivadas, desde cedo. Considerando, então, que se registra a aquisição de palavras complexas em crianças a partir de 2 anos de idade, esse trabalho se propõe a investigar a aquisição destes itens e, consequentemente, a aquisição da morfologia derivacional em crianças falantes do português brasileiro.

3. Fundamentação teórica

Este trabalho fundamenta-se na base teórica fornecida pelos modelos baseados no uso – abordagens que compartilham a ideia de que não há uma separação nítida entre léxico e gramática, sendo essas duas esferas agrupadas em estruturas simbólicas (LANGACKER, 1987, 2008).

Diversos estudos têm trazido evidências para a postulação nos Modelos baseados no Uso (BYBEE, 2001, LANGACKER, 1987, 2008), da gradiência das estruturas linguísticas em oposição a um caráter estritamente discreto. Tem sido demonstrado que o comportamento dos indivíduos em situação de teste é gradiente em relação ao grau de complexidade dos itens lexicais, isto é, se a palavra é formada por mais de um morfema ou não. Em outras palavras, os indivíduos avaliam em escala se as palavras são ou não formadas por afixação ao invés de apresentarem um comportamento categórico, indicando, por exemplo, que *settlement* (do inglês, assentamento) é mais afixado do que *government* (do inglês, governo), conforme nos estudos de Hay (2001, 2003).

Esse tipo de evidência coloca uma questão em relação à representação e processamento das palavras complexas. Em que medida as palavras complexas são representadas e acessadas em função de suas subpar-

tes morfológicas? Segundo Bybee (1995, 1998), na proposição do modelo de redes (*Network Model*) tanto palavras simples quanto complexas estão representadas no léxico. Além disso, não há diferença de representação e processamento de palavras complexas regulares ou irregulares, sejam formadas por processo flexional ou derivacional, uma vez que ambos os tipos estão representados no léxico e seu grau de decomposicionalidade vai depender de sua frequência de uso. Segundo Bybee, o léxico não é uma lista de palavras, ao contrário, estas estão organizadas em redes de conexão lexical em função de sua semelhança sonora, semântica ou simultaneamente sonora e semântica. Dessa maneira, a estrutura interna da palavra é derivada do conjunto de conexões estabelecidas entre as palavras que apresentam as mesmas partes. Ademais, o modelo postula que a frequência de uso tem impacto na representação e determina a força lexical da palavra. Sendo assim, se a palavra é de alta frequência de uso, esta será representada de forma autônoma, não contribuindo para as redes de conexão que levam em conta as semelhanças entre os itens baseados na semelhança sonora e semântica, assim como também não irá contribuir para as redes de conexão estabelecidas em semelhança sonora e semântica. Desse modo, se um determinado padrão emergir de uma rede composta de um número alto de palavras que compartilham esse padrão, o resultado será uma forma de alta frequência de tipo, que tenderá a ser produtiva no sentido de ter seu uso estendido a novas palavras no léxico. Do ponto de vista do acesso lexical, a palavra de alta frequência tenderá a ser acessada e processada autonomamente ao passo que as de baixa frequência tenderão a depender da rede de conexões para o seu acesso, processamento e até mesmo aquisição.

Já Hay (2003, p. 5) apresenta como hipótese que a percepção de aspectos lexicais e sublexicais das palavras na fala afeta a representação de longo termo dos itens no léxico. O estudo apresenta evidências de que o processamento e a representação das palavras complexas no léxico dependem da frequência relativa entre a base e a forma derivada a ela relacionada, de maneira que se a forma derivada for mais ou menos frequente que sua base haverá impacto no tipo de processamento com consequência para a representação (HAY, 2003, p. 10). Por exemplo, os seguintes pares de palavras do inglês apresentam diferentes frequências relativas em relação às suas bases: *insane* (insano) é mais frequente que sua base *sane* (são), enquanto *insincere* (insincero) é menos frequente que sua base *sincere* (sincero). Sendo assim, *insane* tenderá a ser processado pela rota direta (*direct route*), isto é, como uma palavra inteira, reforçando uma representação com um grau baixo de decomposicionalidade, ao

passo que *insinsere* tenderá a ser processada pela rota de decomposição (*decomposed route*), na qual a palavra é acessada via suas subpartes, contribuindo para sua representação com grau alto de decomposicionalidade. As evidências do trabalho de Hay indicam que a frequência relativa parece ser mais importante que a frequência de ocorrência absoluta da proposta de Bybee (1995). Palavras de baixa frequência de ocorrência mas com frequência relativa maior do que a de suas bases tenderiam a serem processadas pela roda direta e representadas com baixo grau de decomposicionalidade (HAY, 2003, p. 73);

Segundo Hay (2003), outros aspectos competem no processamento. O grau de transparência fonológica entre a base e a palavra derivada pode contribuir para sua decomposicionalidade. Por exemplo, a relação entre *insane* and *sane* é fonologicamente transparente, ao passo que em *sanity* (sanidade) não há a mesma transparência. Segundo Bybee, também palavras com baixa transparência fonológica tendem a ter representação autônoma. Além disso, a frequência fonotática das fronteiras morfológicas pode interferir no processamento. Se a frequência fonotática for típica de padrões internos à palavra, a tendência será a de um processamento sem fronteira morfológica, contribuindo para a baixa decomposicionalidade do item lexical, ao passo que, se for uma frequência alta, típica de fronteira de morfema, a consequência será a de processamento e representação das partes componentes da palavra.

Considerando, de acordo com os modelos baseados no uso, que a gramática é emergente e que alguns níveis de representação são emergentes de outros (BYBEE & HOOPER, 2001), em que medida os itens complexos produzidos pelas crianças expressam um armazenamento que contemple as subpartes componentes dos itens lexicais? A produtividade observada em algumas produções sufixais que resultam em não palavras da língua, como pintadeiro, formado a partir do acréscimo do verbo pintar ao sufixo *-eiro*, ou mesmo, grudador, constituído pela base verbal grudar anexada ao sufixo -dor, são indicativos de que aquelas formas estão sendo analisadas como componente de outros itens e estão sendo aplicadas para formar um item que, a princípio, não faz parte do léxico da criança ou não tem ainda uma representação forte que assegure a estabilidade da representação de sua forma como na língua alvo.

Retomando o objetivo inicialmente apresentado nesta seção, este trabalho se baseia em dados levantados de produção espontânea de crianças e 2;0 a 5;0 anos de idade com o objetivo de detectar que itens, que podem ser historicamente relacionados com formação derivacional, apa-

recem na fala dessas crianças. Esse levantamento irá nortear a elaboração de testes com vistas a avaliar o grau de decomposicionalidade de determinados itens lexicais por crianças de 2;0 e 5;0 com vistas a contribuir para os estudos sobre aquisição de morfologia derivacional.

4. Metodologia

A amostra é constituída de dados de fala espontânea de crianças de 2;0 a 5;0 anos de idade, oriundos da Amostra de Fala Infantil e da amostra AQUIVAR, ambas pertencentes ao acervo do Programa de Estudos sobre o Uso da Língua (PEUL/UFRJ), de dados provenientes do *corpus* FLORIANOPOLIS/CHILDES.

Para a constituição dos *corpora* desta pesquisa, selecionamos 09 crianças com idades entre 2 e 5 anos de idade da Amostra de Fala Infantil do PEUL, sendo 02 crianças do sexo masculino e 07 do sexo feminino. Da amostra AQUIVAR (PEUL/UFRJ), foram analisadas 34 crianças com idades entre 2 e 5 anos, sendo 15 meninos e 19 meninas. Além disso, do *corpus* FLORIANÓPOLIS/CHILDES, foram coletados dados da informante Ester, mãe de Paulo.

Na recolha de dados, utilizou-se as transcrições correspondentes às amostras de fala espontânea da Amostra de Fala Infantil e da amostra AQUIVAR (PEUL/UFRJ), visando listar os itens lexicais derivados que aparecem na fala infantil no período de 2 a 5 anos de idade. Além disso, dada a importância da fala do adulto em servir de *input* para a fala de crianças em estágio de aquisição de linguagem, coletou-se também as palavras complexas presentes no discurso da informante Ester, da amostra FLORIANOPOLIS/CHILDES[35], a fim de buscarmos os morfemas derivacionais presentes na fala dessa mãe em interação com seu filho, única criança falante do português brasileiro que faz parte da base CHILDES.

A partir dos itens derivados levantados anteriormente da Amostra de Fala Infantil do PEUL e do *corpus* do AQUIVAR (PEUL/UFRJ), extraiu-se a frequência de tipo dos morfemas derivacionais da base de dados do projeto Avaliação Sonora do Português Atual (ASPA/UFMG) a fim de realizarmos um levantamento de frequência em termos de *type* que nos permitiu definir e categorizar os afixos em dois grupos, alta e baixa frequência, em função de sua produtividade na língua.

[35] A *homepage* do *Child Language Data Exchange System* (CHILDES) pode ser acessada pelo site http://childes.psy.cmu.edu.

5. Resultados

A frequência de uso será apresentada em três partes: um levantamento feito na amostra infantil do banco de dados do PEUL; um levantamento efetuado na amostra AQUIVAR (PEUL/UFRJ) e no *corpus* FLORIANOPOLIS/CHILDES; e um levantamento realizado na base de dados ASPA (UFMG) a fim de categorizar os afixos em função de sua produtividade na língua.

A seguir, na tabela 01, apresentamos o levantamento da frequência de *token* para as formas derivadas sufixadas encontradas na fala de crianças entre 2 e 5 anos. No entanto, cabe ressaltar que o tratamento aplicado aos dados será em termos da frequência de *token*, dado que nosso objetivo é a realização de um levantamento dos afixos presentes na fala de crianças entre 2 e 5 anos de idade.

TABELA 01

Frequência de ocorrência de sufixos na fala de crianças de 2 a 5 anos (continua)

Nº	**Sufixos**	**N**	**%**
1	-inho(a)	154	73,7
2	-eiro(a)	12	5,7

TABELA 01

Frequência de ocorrência de sufixos na fala de crianças de 2 a 5 anos (conclusão)

Nº	**Sufixos**	**N**	**%**
3	-dor	7	3,3
4	-aria /-eria	5	2,4
5	-ado(a)	4	1,9
6	-(i)dade	4	1,9
7	-oso(a)	6	2,9
8	-ão	3	1,4
9	-aço	2	1,0
10	-al	2	1,0
11	-eto(a)	2	1,0
12	-ino	2	1,0
13	-mento	2	1,0
14	-eco(a)	1	0,5
15	-edo	1	0,5
16	-ete	1	0,5
17	-ia	1	0,5
TOTAL		209	100

Fonte: AMOSTRA DE FALA INFANTIL - PEUL/UFRJ

A partir da tabela acima, nota-se que o diminutivo *-inho*(*a*) tem alta frequência de ocorrência nesta amostra, aparecendo em 73,7% das formas derivadas. Em seguida, temos o agentivo *-eiro* que ocorre em 5,7% das palavras complexas listadas no *corpus* em análise, seguido do sufixo -dor com 3,3% das ocorrências em formas derivadas. Dos 17 tipos listados na tabela 01, os sufixos com menor número de ocorrências são -*eco*(*a*), *-edo*, *-ete* e -ia, todos com 0,5%, ocupando as últimas posições da tabela.

Em relação às formas derivadas prefixadas, foram encontradas 11 ocorrências com o prefixo *des-*, único prefixo registrado neste *corpus*, como se vê na tabela 02.

TABELA 02

Frequência de ocorrência de prefixos na fala de crianças de 2 a 5 anos

Nº	**Prefixos**	**N**	**%**
1	des-	11	100
TOTAL		11	100

Fonte: AMOSTRA DE FALA INFANTIL - PEUL/UFRJ.

Os dados oriundos da amostra AQUIVAR (PEUL/UFRJ) encontram-se distribuídos em função da frequência de ocorrência para as formas derivadas sufixadas, conforme apresentado na tabela 03. Reiteramos que o tratamento aplicado aos dados do AQUIVAR (PEUL/UFRJ) será em termos da frequência de *token*, dado que objetivamos levantar as produções de fala manifestadas em crianças entre 2 e 5 anos de idade.

TABELA 03

Frequência de ocorrência de sufixos na fala de crianças de 2 a 5 anos (continua)

Nº	**Sufixos**	**N**	**%**
1	-inho(a)	478	70,8
2	-eiro(a)	55	8,1
3	-dor	39	5,8
4	-edo	22	3,3
5	-eto(a)	12	1,8
6	-ão	11	1,6
7	-ino	10	1,5
8	-oso(a)	10	1,5
9	-ado(a)	4	0,6

TABELA 03					
Frequência de ocorrência de sufixos na fala de crianças de 2 a 5 anos (conclusão)					
Nº	**Sufixos**	**N**	**%**		
10	-anca	4	0,6		
11	-ote	4	0,6		
12	-al	3	0,4		
13	-(i)ficar	3	0,4		
14	-mento	3	0,4		
15	-agem	2	0,3		
16	-eco(a)	2	0,3		
17	-(i)dade	2	0,3		
18	-nte	2	0,3		
19	-aço(a)	1	0,1		
20	-aria /-eria	1	0,1		
21	-ção	1	0,1		
22	-ento	1	0,1		
23	-ete	1	0,1		
24	-ico	1	0,1		
25	-ista	1	0,1		
26	-oto(a)	1	0,1		
27	-udo	1	0,1		
TOTAL		675	100		

Fonte: AMOSTRA AQUIVAR - PEUL/UFRJ.

Como se pode observar, nesta tabela, ocorre o predomínio do morfema *–inho*(*a*), apresentando uma frequência oito vezes maior (70,8%) que o sufixo seguinte, em ordem de ocorrência, o agentivo *–eiro*(*a*), com cerca de 8,1%. Em seguida, aparece o sufixo -dor que ocorre em 5,8% das palavras complexas listadas no *corpus* do AQUIVAR (PEUL/UFRJ). Nota-se ainda que, dos 27 tipos listados na tabela 03, os sufixos com menor número de ocorrências são os que apresentam um índice percentual de 0,1%: *-aço*(*a*), *-aria/-eria, -ção, -ento, -ete, -ico, -ista, -oto*(*a*) e *-udo*.

Em relação às formas derivadas prefixadas, foram encontradas 12 ocorrências que estão distribuídas em função da frequência de *token* dos prefixos extraídos do *corpus* AQUIVAR (PEUL/UFRJ), como se observa na tabela 04.

TABELA 04			
Frequência de ocorrência de prefixos na fala de crianças de 2 a 5 anos			
Nº	**Prefixos**	**N**	**%**
1	en-	6	50
2	des-	3	25
3	re-	2	16,7
4	es-	1	8,3
TOTAL		12	100

Fonte: AMOSTRA AQUIVAR - PEUL/UFRJ

Na tabela acima, observa-se que a distribuição dos prefixos apresenta um número de ocorrências bem reduzido em função do baixo número de formas derivadas encontradas, apenas 12 palavras complexas. Dos 4 tipos listados, o prefixo *en-* ocorre predominantemente em 50% dos dados analisados e apresenta uma frequência de ocorrência duas vezes maior que o prefixo *des-*, com exatamente 25%. Em seguida, temos o prefixo *des-* que ocorre em 16,7% das palavras complexas listadas no *corpus* em análise, seguido do prefixo es- com 8,3% das ocorrências em formas derivadas.

Os dados extraídos do *corpus* em análise encontram-se distribuídos em função da frequência de ocorrência dos sufixos detectados na fala de Ester, mãe de Paulo, conforme apresentado na tabela 05.

TABELA 05			
Frequência de ocorrência de sufixos presentes na fala da informante Ester			
Nº	**Sufixos**	**N**	**%**
1	-inho(a)	105	64,0
2	-eiro(a)	14	8,5
3	-ão	10	6,1
4	-edo	6	3,7
5	-ola	4	2,4
6	-eta	4	2,4
7	-oso(a)	3	1,8
8	-mente	3	1,8
9	-ura	2	1,2
10	-dor	2	1,2
11	-eco(a)	2	1,2
12	-izar	1	0,6
13	-ivo	1	0,6
14	-íssimo	1	0,6
15	-mento	1	0,6
16	-agem	1	0,6

17	-udo	1	0,6
18	-ia	1	0,6
19	-tório	1	0,6
20	-ista	1	0,6
TOTAL		164	100

Fonte: AMOSTRA FLORIANÓPOLIS - CHILDES.

Como se pode observar na tabela acima, o diminutivo *-inho*(*a*) também é muito frequente no discurso de Ester, mãe de Paulo, em exatos 64% dos dados. Em seguida, temos o agentivo *-eiro*, que ocorre em 8,5% das palavras complexas listadas no *corpus* em análise, seguido do sufixo *-ão* com 6,1% das ocorrências em formas derivadas. Dos 20 tipos listados na tabela 05, os sufixos com menor número de ocorrências são os que apresentam um índice percentual de 0,6%: *-izar, -ivo, -íssimo, -mento, -agem, -udo, -ia, -tório* e *-ista*.

De modo semelhante ao que ocorreu com os dados rastreados na Amostra de Fala Infantil e no *corpus* AQUIVAR, observa-se a ocorrência de um número bastante reduzido de formas prefixadas, sendo encontradas 12 ocorrências que estão distribuídas em função da frequência de *token*, como se vê na tabela 06.

TABELA 06			
Frequência de ocorrência de prefixos presentes na fala da informante Ester			
Nº	**Prefixos**	**N**	**%**
1	des-	10	83,3
2	es-	1	8,3
3	en-	1	8,3
TOTAL		12	100

Fonte: AMOSTRA FLORIANÓPOLIS - CHILDES.

Na tabela 06, nota-se que três prefixos foram rastreados na amostra FLORIANOPOLIS/CHILDES com base na fala da informante Ester, mãe de Paulo. Em função do baixo número de ocorrência das formas prefixadas encontradas, o prefixo *des-* ocorre predominantemente com maior frequência (83,3%) na fala de Ester em interação com seu filho do que os dois prefixos subsequentes. Em seguida, aparecem os prefixos es- e *en-* com aproximadamente 8,3% das ocorrências em formas derivadas.

Cabe ressaltar que os afixos listados na fala da mãe Ester no *corpus* FLORIANOPOLIS/CHILDES são relativamente os mesmos que aparecem na fala das crianças entre 2 e 5 anos listados nos dados da Amostra de Fala Infantil e da amostra AQUIVAR (PEUL/UFRJ), ou seja, não há diferenças substanciais entre os afixos encontrados na fala da mãe e aqueles coletados na fala das crianças, reiterando o fato de que a

fala do adulto serve de *input* para criança durante o período de aquisição da linguagem.

Realizou-se um levantamento de frequência em termos de *type* que nos permitiu definir e categorizar os afixos em dois grupos, alta e baixa frequência, em função de sua produtividade na língua. A partir dos dados extraídos da base ASPA (UFMG), os afixos foram distribuídos em função da frequência de *type*, como se observa na tabela 07.

TABELA 07

Distribuição de sufixos em função da frequência de tipo (continua)

Nº	**Sufixos**	**N**
1	-inho(a)	1223
2	-dor(a)	1088
3	-eiro(a)	740
4	-oso(a)	610
5	-mento	527
6	-ista	445
7	-(i)dade	406
8	-ção	392
9	-ão	391
10	-al	324
11	-nte	227
12	-agem	193
13	-ico	129
14	-ino(a)	124
15	-aria /-eria	106
16	-eto(a)	59
17	-(i)ficar	56
18	-ia	44

TABELA 07

Distribuição de sufixos em função da frequência de tipo (conclusão)

Nº	**Sufixos**	**N**
19	-aço(a)	37
20	-udo(a)	33
21	-ança	29
22	-ete	27
23	-ento	21
24	-ote	15
25	-eco(a)	14
26	-oto(a)	8
27	-edo	3
28	-ado(a)	3
TOTAL		7274

Fonte: Base de dados do projeto ASPA - UFMG.

Na tabela 07, nota-se que o tipo mais frequente é o diminutivo -*inho*(*a*) com 1223 registros na base ASPA (UFMG). No entanto, a frequência de tipo decai substancialmente a partir do sufixo -*eto*(*a*), com exatos 59 formas derivadas, chegando ao extremo de mais baixa frequência com os sufixos -*edo* e -*ado*(*a*), contando cada um com 3 formas derivadas presentes no *corpus* do projeto ASPA (UFMG).

A seguir, na tabela 08, apresentamos o levantamento da frequência de *type* para os prefixos.

TABELA 08

Distribuição de prefixos em função da frequência de tipo

Nº	**Prefixos**	**N**
1	des-	553
2	re-	417
3	en-	107
4	es-	20
TOTAL		1097

Fonte: Base de dados do projeto ASPA – UFMG

Na tabela acima, observa-se que o tipo mais frequente é o prefixo *des*- com exatos 553 registros na base ASPA (UFMG), seguido pelo prefixo *re*- que ocorreu em 417 formas derivadas. No entanto, há um decréscimo substancial da frequência de tipo a partir do prefixo *en*-, apresentando 107 registros, e culmina com o tipo menos frequente, o prefixo es-, ocorrendo em 20 formas derivadas.

6. Considerações finais

Este estudo preliminar teve como objetivo realizar um levantamento de produção de fala espontânea de crianças de 2;0 a 5;0 anos, a fim de identificarmos os afixos que aparecem na fala infantil bem como na interação entre mãe e filho, registrando os morfemas derivacionais presentes nesta interação, dada a importância da fala do adulto em servir de *input* para a fala de crianças em estágio de aquisição de linguagem.

A partir dos itens derivados levantados anteriormente da Amostra de Fala Infantil do PEUL e do *corpus* do AQUIVAR (PEUL/UFRJ), extraiu-se a frequência de tipo dos morfemas derivacionais da base de dados do projeto Avaliação Sonora do Português Atual (ASPA/UFMG) com o propósito de realizar um levantamento de frequência em termos de *type* que nos permitiu definir e categorizar os afixos em dois grupos, alta

e baixa frequência, em função de sua produtividade na língua. Esse levantamento fornece indicações importantes para a elaboração de testes para avaliação da aquisição da morfologia derivacional que fazem parte da continuidade da presente pesquisa.

REFERÊNCIAS BIBLIOGRÁFICAS

AITCHISON, J. *Words in the mind:* an introduction to the mental lexicon. 2. ed. London: Basil Blackwell, 1994.

BLOOM, P. *Language acquisition*: Core readings. Cambridge: MIT Press, 1994.

BYBEE, J. Regular morphology and the lexicon. *Language and Cognitive Process*, vol. 10, n. 5, p. 425-455, 1995.

_____. *The emergent lexicon*. CLS 34: The panels. Chicago Linguistics Society, p. 421-435, 1998.

_____. *Phonology and language use*. Cambridge: Cambridge University Press, 2001.

_____; HOPPER, P. Introduction to frequency and the emergence of linguistic structure. In: BYBEE, J.; HOPPER, P. (Eds.). *Frequency and the emergence of linguistic structure.* Amsterdam: John Benjamins, p. 1-24, 2001.

CAREY, S. The child as word learner. In: HALLE, M.; BRESNAN, J.; Miller, G. A. (Eds.). *Linguistic theory and psychological reality*. Cambridge: MIT, 1978.

CLARK, E. *The lexicon in acquisition*. Cambridge: Cambridge University Press, 1993.

_____. Morphology in Language Acquisition. In: Andrew Spencer e Arnold M. Zwicky (Eds.), *The Handbook of Morphology*. Oxford: Blackwell, p. 374-389, 2001.

_____. *First language acquisition*. 2. ed. Cambridge: Cambridge University Press, 2009.

HAY, J. Lexical frequency in morphology: is everything relative? *Linguistics*, 39, p. 1041-1070, 2001.

_____. *Causes and Consequences of Word Structure*. Routledge: New York and London, 2003.

LANGACKER, R. *Foundations of Cognitive Grammar,* vol. 2. Stanford: Stanford University Press, 1991.

_____. *Cognitive grammar*: a basic introduction. Oxford: Oxford University Press, 2008.

NAGY, W. E.; ANDERSON, R. C. How many words are there in in printed school English? *Reading Research Quarterly*, vol. 19, p. 304-330, 1984.

NATION, P. Vocabulary size, growth, and use. In: SCHREUDER, R.; WELTENS, B. (Eds.). *The bilingual lexicon*. Amsterdam/Philadelphia: Benjamins, 1993.

NUNES, T.; BRYANT, P. *Improving literacy by teaching morphemes*. London: Routledge, 2006.

WYSOCKI, K.; JENKINS, J. R. Deriving word meanings through morphological generalization. *Reading Research Quarterly,* vol. 22, p. 66-81, 1987.