

# PERFIL EVOLUTIVO DA RELAÇÃO *TYPE/TOKEN* DE CRIANÇAS DE 3 A 5 ANOS DE IDADE\*

## EVOLUTIVE PROFILE OF THE *TYPE/TOKEN* RELATION IN THE CHILD BETWEEN THREE AND FIVE YEARS OLD

Sabrina Scherer<sup>1</sup>

Fabíola Schwengber Casarin<sup>2</sup>

Patrícia Zart<sup>3</sup>

Ana Paula Fernandes Ramos<sup>4</sup>

Vicente José Assencio-Ferreira<sup>5</sup>

### RESUMO

**Objetivos:** levantar um perfil evolutivo da relação *type/token* (TTR) nas faixas etárias de 3, 4 e 5 anos, verificando a *type/token* por categoria gramatical, por faixa etária e sexo e qual sua validade evolutiva. **Métodos:** foram analisadas 60 crianças em desenvolvimento típico de linguagem em atividades lúdicas com as pesquisadoras ou suas mães. Foram realizadas a gravação da fala espontânea e a sua transcrição, sendo efetuados os cálculos da *type/token* total e segmentar (50 enunciados). **Resultados:** tal medida não se apresenta produtiva nas idades estudadas, diferenciando agora nos quatro anos das demais idades. **Conclusão:** é necessário investigar novas medidas para estas faixas etárias.

**Descritores:** Desenvolvimento da linguagem; Linguagem infantil; Vocabulário; Comportamento verbal; Fonoterapia; Fala; Leitura

### ■ INTRODUÇÃO

A aquisição lexical inicial é marcada por alguns acontecimentos importantes: início da aquisição é basicamente a presença de substantivos na expressão, o crescimento do vocabulário das meninas é um pouco mais rápido do que os dos meninos, a existência de explosão do vocabulário ocorre ao redor dos 2 anos de idade e sobre esta idade é o momento de surgimento da sintaxe pelo aumento de verbos e o surgimento de outras categorias como conjunções, preposições.<sup>(1-3)</sup>

Tenta-se quantificar, de várias maneiras, os diversos aspectos da produção e compreensão da linguagem pelas crianças, através de diversos métodos de avaliação ou coleta do vocabulário inicial: o diário, a lista de verificação, o teste de evocação, as medidas quantitativas e qualitativas que são realizadas por meio de amostras da linguagem espontânea das crianças.<sup>(1)</sup>

As medidas quantitativas de linguagem avaliam componentes isolados do sistema de produção da linguagem da criança, como a extensão média de enunciado (MLU), número total de enunciados (TOT-UTT), número total de sentenças completas e inteligíveis (TOTCI), extensão média da sintaxe (MLS), número de diferentes palavras (NTW ou *types*), número total de palavras (NDW ou *tokens*) e a relação *type/token* (TTR). Algumas dessas são medidas evolutivas, pois estão associadas a

\*Trabalho realizado no CEFAC – Centro de Especialização em Fonoaudiologia Clínica.

<sup>1</sup>Especialização em Linguagem pelo Centro de Especialização em Fonoaudiologia Clínica (CEFAC). Fonoaudióloga graduada pela Universidade Luterana do Brasil (ULBRA).

<sup>2</sup>Especialização em Linguagem pelo Centro de Especialização em Fonoaudiologia Clínica (CEFAC). Fonoaudióloga graduada pela Universidade Luterana do Brasil (ULBRA).

<sup>3</sup>Especialização em Linguagem pelo Centro de Especialização em Fonoaudiologia Clínica (CEFAC). Fonoaudióloga graduada pela Universidade Luterana do Brasil (ULBRA).

<sup>4</sup>Doutora em Letras pela Pontifícia Universidade Católica do Rio Grande do Sul. Docente do Curso de Graduação em Fonoaudiologia pela ULBRA.

<sup>5</sup>Doutor em Medicina (Neurologia) pela Universidade de São Paulo (USP).

mudanças de idade, possibilitando, dessa forma, a diferenciação entre crianças com problemas de linguagem e crianças que apresentam um desenvolvimento normal. O autor ressalta a NTW (*types*) e a NDW (*tokens*) como bons índices para diferenciar faixas etárias e este último como bom parâmetro de diversidade lexical.<sup>(4)</sup>

A relação *type/token* (TTR) é uma medida da produção lingüística usada para estimar a proficiência lexical infantil, ou seja, é uma forma de verificar a diversidade lexical ou a variedade de diferentes palavras faladas pela criança.<sup>(5)</sup>

No estudo original, foram estudadas TTR calculadas em amostras de 50 enunciados de 480 crianças entre 3 e 8 anos de idade. Esta pesquisa revelou que, embora os números de ambos tipos (*types*) e palavras faladas (*tokens*) aumentassem entre 3 e 8 anos de idade, o número total de palavras faladas pelas crianças, correspondia, na amostra, ao dobro do número de palavras diferentes.<sup>(1,2)</sup> Os resultados foram obtidos opostamente à extensão de idade e não variaram por gênero e nível socioeconômico. Tanto a extensão média de enunciados (MLU) quanto o número de diferentes palavras produzidas (NDW = *types*) e o número total de palavras usadas (NTW = *tokens*) parecem se correlacionar com a idade o que as torna medidas evolutivas.<sup>(4,6)</sup>

Com o objetivo de avaliar medidas de diversidade lexical, em crianças com distúrbio específico de linguagem (DEL) e em crianças em desenvolvimento típico de linguagem com mesma idade lingüística (IL) ou cronológica (IC), coletou-se uma amostra de 100 enunciados de cada participante enquanto brincavam com um conjunto básico de brinquedos. Durante a amostragem, cada criança interagiu com um adulto. Medidas alternativas de diversidade lexical, como o NDW foram utilizadas para comparar os grupos (DEL, IL, IC), em amostras de 50 e 100 enunciados e, ainda, de 100 e 200 enunciados. Em cada análise, foi encontrado um significativo efeito de grupo. O padrão dos resultados foi o mesmo para ambas as amostras de 100 e 200 enunciados; o grupo IC (crianças em desenvolvimento típico, com mesma idade cronológica que as crianças com DEL) produziu um melhor NDW que os grupos IL (crianças em típico desenvolvimento igualadas a crianças com DEL em nível a capacidade lingüística) e DEL, e o grupo IL não diferiu do grupo DEL em NDW produzido. O autor concluiu que a TTR tradicional não diferenciou os grupos em ambas as amostras de 50 e 100 elocuições. Por isso, esses dados encontrados são duvidosos na aplicação recomendada de TTR como um indicador de diagnóstico de perda de linguagem para crianças com DEL, especificamente como uma medida de habilidades semânticas limitadas ou pobre capacidade de recuperação lexical. Sendo assim, esses resultados desencadeiam perguntas sobre as condições sob as quais TTR é uma medida informativa de diversidade lexical de crianças com DEL.<sup>(7)</sup>

Outra questão proposta por esse estudo foi verificar quais medidas alternativas de diversidade lexical foram sensíveis às diferenças nas capacidades lexicais de crianças com DEL e suas contrapartes em desenvolvimento típico. O NDW usado em amostras de 50 e 100 enunciados e nas amostras de 100 e 200 diferenciaram consistentemente as crianças com DEL dos seus pares IC. Acrescentando, os pares IL também foram distinguidos dos participantes IC. As crianças com DEL e suas contrapartes IL não se diferenciaram em nenhuma das medidas.<sup>(7)</sup>

Esses dados encontrados revelam que NDW distingue crianças com DEL dos pares de mesma idade com desenvolvimento típico de linguagem, ambos quando as amostras são igualadas em quantidade de enunciados ou por número de símbolos (NTW). Este fato implica muito para a medida de diversidade lexical em atividades clínicas ou de pesquisa.<sup>(7)</sup>

Quando se compara o TTR (relação *type/token*), NDW (*types*) e NTW (*tokens*) com os dados normalizados, recomenda-se o uso de uma TTR de 50 enunciados como uma norma geral para refletir a diversidade apropriada, assinalando que TTR abaixo de 50 enunciados indicam significativamente uma falta de diversidade, que pode ser associada com deficiência de linguagem. Embora não diretamente mencionada, uma consideração da extensão normal de variações para TTR seria muito importante, e a NDW e NTW poderiam ser realizados através de classes específicas (substantivos, verbos, adjetivos, entre outros), a fim de quantificar a diversidade em áreas particulares do léxico.<sup>(1)</sup>

Dois estudos realizados examinaram a TTR na classe gramatical dos verbos, comparando os tipos de verbos, enunciados dos verbos e TTR de verbos de crianças com desvios específicos de linguagem (DEL) e controles de linguagem (IL) e cronológicos (IC), para identificar áreas de dificuldade particulares para crianças e adolescentes com DEL. Este estudo envolveu as análises de amostras de 100 enunciados, coletadas de 14 crianças com DEL, 14 IC e 14 IL. Verificou-se que as crianças com DEL tiveram TTR verbais significativamente menores que as de seus pares IC e IL (0,42 para as crianças com DEL, 0,50 e 0,48 para os pares IC e IL respectivamente). Análises informais das crianças com DEL, revelaram que estas produziram menos tipos e enunciados de verbos que os seus controles de idade. Porém, as crianças com DEL produziram o mesmo número de tipos de verbos que os seus controles em idade de linguagem, entretanto mais enunciados verbais.<sup>(5,7)</sup>

Observaram uma limitação da TTR quando analisados os índices de desenvolvimento de linguagem. Isto porque a TTR é altamente influenciada pelo tamanho da amostra de enunciados. Com aumento do número de enunciados da amostra, as palavras tendem a se repetir e a TTR diminui. Mesmo quando o número de enunciados é mantido cons-

tante, o número de tipos utilizados pode diferir consideravelmente. A diferença da extensão da amostra em totais de palavras enunciadas, pode contribuir para a ausência de diferença entre as TTR de crianças em diferentes níveis de proficiência lingüística. Por exemplo, a TTR de uma criança com 25 tipos para 50 palavras enunciadas seria 0,5, enquanto a TTR de outra com 50 em 100 seria a mesma. Quando TTR são calculadas por números particulares de enunciados (por exemplo, 50 ou 100 enunciados), ou seja, TTR segmentar, e se comparam às crianças com DEL, seus controles de idade e linguagem, os controles de idade são aptos a produzir enunciados significativamente mais longos do que as das crianças com DEL. Mais especificamente, uma criança de 5 anos de idade com DEL pode produzir 60 palavras diferentes e 130 palavras em total em uma amostra de 50 enunciados. Já um controle de 5 anos de idade pode produzir 120 palavras diferentes e 260 palavras em total. Embora ambas as crianças tenham TTR de aproximadamente 0,46, existem diferenças óbvias em suas capacidades de produção lexical. Portanto, importantes diferenças lingüísticas podem ser mascaradas em cálculos de TTR. Dessa forma, os enunciados mais longos, incluem mais palavras e, por isso, podem reduzir a TTR e simular grupos de diferenças.<sup>(5)</sup>

Enfim, os resultados dessa investigação e das demais anteriormente citadas sugerem que TTR básica não é muito adequada para diversas atividades clínicas e de pesquisa, envolvendo crianças com DEL, assim como ela não difere crianças com DEL de seus controles de mesma idade cronológica e capacidades lingüísticas. Em contraste, o NDW surge para ser um índice de diversidade lexical mais informativo, porém uma investigação continuada do potencial de aplicação de NDW e suas propriedades de medição é solicitada, como no desenvolvimento de dados normativos mais intensivos em aspectos de aquisição lexical.<sup>(4,8-10)</sup>

O objetivo deste trabalho é levantar a relação *type/token* nas faixas etárias de 3, 4 e 5 anos, verificando a relação *type/token* por categoria gramatical e por faixa etária e analisar o desempenho na relação *type/token*, comparando meninos e meninas, os dados obtidos pela relação *type/token* total e segmentar (50 primeiros enunciados) e analisar se a forma mais efetiva de visualizar a evolução da diversidade do vocabulário é do coeficiente final *type/token* ou a comparação bruta do número total de palavras diversas (*types*) sobre o número total de palavras enunciadas (*tokens*).

## ■ MÉTODOS

Foi realizada uma pesquisa com 20 crianças por faixa etária, em desenvolvimento típico de linguagem, sendo dez do sexo masculino e dez do sexo feminino, totalizando

60 crianças. As faixas etárias (FE) analisadas foram as seguintes: FE 1 (3 anos), FE 2 (4 anos) e FE 3 (5 anos).

Foram excluídas as crianças com histórico de otites de repetição ou algum comprometimento auditivo, crianças com atraso de linguagem ou qualquer outro comprometimento de desenvolvimento.

Antes da coleta de dados, as crianças foram observadas pelas pesquisadoras. Foram considerados o modo como brincavam, a interação comunicativa, a etapa cognitiva, a autonomia discursiva e a esfera afetiva, tomando-se como referência Golinkoff e Hirsh-Pasek (1997).<sup>(6)</sup>

A coleta de dados foi realizada pelas pesquisadoras no primeiro semestre de 2001, em creches e/ou lar das crianças, através da gravação da fala espontânea em atividades lúdicas durante 45 minutos. Os dados foram registrados num gravador da marca Panasonic, com fita de 90 minutos. O próprio ruído da parada automática do gravador indicou o término da avaliação.

Após essa gravação, foi realizada a transcrição da fala das crianças e posteriormente os cálculos da TTR foram efetuados:

- *Type/token* = número de palavras diferentes da amostra/número total das palavras da amostra

Além da TTR total, foi realizado também o cálculo da TTR segmentar de 50 enunciados. Foram analisados os valores individuais do número de diferentes palavras (*types*) e total delas (*tokens*) pelas seguintes categorias gramaticais: substantivos, verbos, preposições, artigos, pronomes, conjunções, adjetivos, numerais e interjeições. Além disso, verificou-se também a divisão do número de *types* pelo número de *tokens*.

Ética: foi avaliado pela Comissão de Aconselhamento em Pesquisa do Curso de Fonoaudiologia da ULBRA resolução 196/96. Foi solicitada a autorização dos pais das crianças para o desenvolvimento da pesquisa.

## ■ RESULTADOS

Pelos resultados do teste observa-se que não existem diferenças significativas entre as variáveis estudadas entre os sexos (Tabela 1).

Pelo teste realizado observa-se que as variáveis que apresentaram diferenças significativas entre as faixas de idade foram:

- **TT 50:** Idade de 4 anos superior às idades de 3 e 5 anos.
- **TT TOTAL:** Idade de 4 anos superior às idades de 3 e 5 anos.
- **TYPE 50:** Idade de 3 anos inferior às idades de 4 e 5 anos.
- **TOKEN 50:** Todas as idades diferem entre si.

**Tabela 1.** Teste de comparação de médias *t-student* entre os sexos

Variável	Sexo	Média	Desvio-padrão	p-student
TTR 50	Masculino	0,48	0,08	0,59
	Feminino	0,50	0,12	
TTR Total	Masculino	0,44	0,15	0,72
	Feminino	0,42	0,14	
Type 50	Masculino	126,37	40,54	0,95
	Feminino	125,70	40,78	
Type total	Masculino	210,77	51,93	0,18
	Feminino	230,50	60,48	
Token 50	Masculino	272,07	99,67	0,35
	Feminino	250,93	72,66	
Token total	Masculino	529,93	227,36	0,48
	Feminino	575,37	266,66	

TTR = Relação *type/token*.

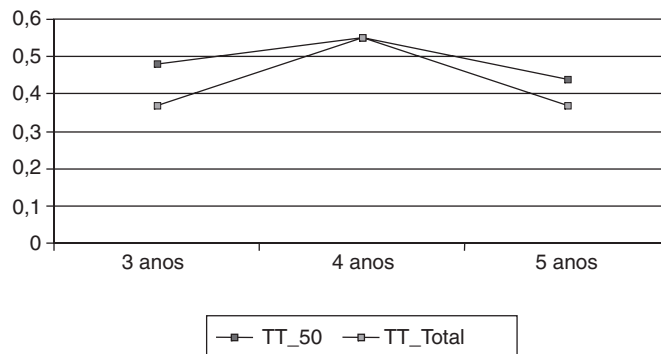
- Token Total:** Idade de quatro anos inferior à idade de cinco anos. A idade de três anos não difere de nenhuma das outras (Tabela 2).

A relação *type/token* da faixa etária de 4 anos foi superior às demais faixas etárias analisadas, possivelmente

**Tabela 2.** Teste de comparação de médias de análise de variância para a variável idade

Variável	Idade	Média	Desvio-padrão	P
TTR 50	3 anos	0,48	0,08	0,01*
	4 anos	0,55	0,11	
	5 anos	0,44	0,07	
TTR total	3 anos	0,37	0,07	0,01*
	4 anos	0,55	0,17	
	5 anos	0,37	0,07	
Type 50	3 anos	89,60	19,70	0,01*
	4 anos	148,55	43,54	
	5 anos	139,95	25,06	
Type total	3 anos	202,00	47,53	0,20
	4 anos	231,85	59,09	
	5 anos	228,05	60,82	
Token 50	3 anos	190,80	42,98	0,01*
	4 anos	266,15	47,46	
	5 anos	327,55	97,96	
Token total	3 anos	554,90	242,69	0,05*
	4 anos	455,75	222,95	
	5 anos	647,30	247,25	

TTR = Relação *type/token*.



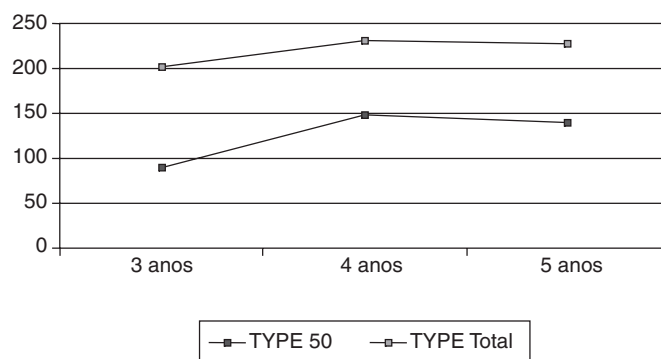
**Fig. 1.** Comparação da relação *type/token* entre as faixas etárias.

devido a uma variação individual e à diversidade vocabular (Fig. 1).

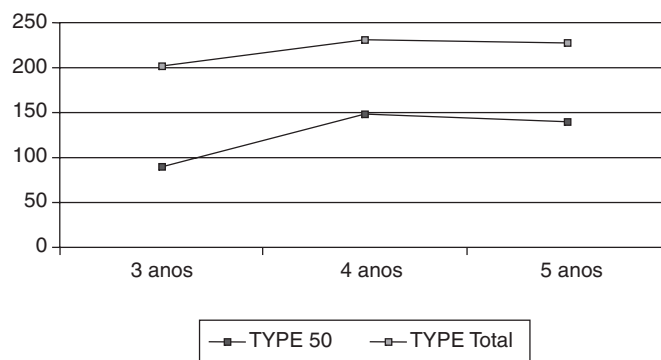
Observou-se que os *types* totais trazem mais dados que os *types* de 50 enunciados, embora não tenham diferido entre as faixas etárias (Fig. 2).

Os resultados dos *tokens* de 4 anos diferem dos demais, sendo, dessa forma, úteis para diferenciar as demais faixas etárias (Fig. 3).

Os substantivos apresentaram resultados superiores aos verbos, que apresentaram superioridade aos advérbios.



**Fig. 2.** Comparação de *types* entre as faixas etárias.



**Fig. 3.** Comparação de *tokens* entre as faixas etárias.

**Tabela 3.** Comparação de *types* e de *tokens* por categoria gramatical

	<b>3 anos</b> <b>TTR 50</b>	<b>TTR total</b>	<b>4 anos</b> <b>TTR 50</b>	<b>TTR total</b>	<b>5 anos</b> <b>TTR 50</b>	<b>TTR total</b>
Substantivos	23/37	61/110	36/52	60/88	41/59	75/115
Verbos	30/52	68/182	32/47	57/99	39/73	67/143
Artigos	2/12	3/18	5/26	6/46	5/27	6/52
Preposições	5/13	6/40	7/15	9/29	7/21	9/42
Pronomes	7/26	20/85	14/30	21/50	16/54	23/97
Conjunção	2/7	4/23	10/23	14/32	5/22	7/39
Advérbios	8/20	16/64	16/27	22/47	16/45	23/90
Adjetivos	4/6	13/20	14/20	18/27	7/8	13/16
Numerais	3/5	4/14	5/9	9/17	3/7	5/13
Interjeições	2/4	4/10	9/18	13/25	3/8	6/16

bios, que, por sua vez, foram superiores aos adjetivos, e estes se igualaram aos pronomes (Tabela 3).

Observou-se que a faixa etária dos quatro anos apresentou uma maior diversidade vocabular que a faixa etária de cinco anos ou uma coleta de fala mais rica.

## ■ DISCUSSÃO

Nesta pesquisa não foram observadas diferenças significativas entre meninos e meninas em termos de aquisição, o que contraria os achados da literatura estrangeira.<sup>(11-13)</sup> Existe uma variação individual nos padrões de crescimento do vocabulário inicial, existindo uma tendência das meninas adquirirem as palavras numa velocidade ligeiramente mais rápida que os meninos. Mas não se encontrou diferenças entre os sexos na TTR 50 e TTR total em crianças de 1 ano e 6 meses, 2 anos e 2 anos e 6 meses, semelhante ao constatado nesta pesquisa.<sup>(14)</sup>

A divisão de *types* e *tokens* tanto de 50 enunciados como do total foram superiores na idade de 4 anos, em razão de o grupo ter uma maior diversidade vocabular ou uma coleta de fala mais rica. Foi encontrado na literatura que, conforme aumenta o número de *types* e *tokens*, o coeficiente da TTR reduzia.<sup>(14)</sup> Outro estudo refere que com o aumento do número de enunciados da amostra as palavras tendem a se repetir, e a TTR conseqüentemente diminui.<sup>(9)</sup>

O número de *types* de 50 enunciados na idade de três anos foi inferior às idades de 4 e 5 anos.

Verificou-se que o número de *tokens* de 50 enunciados diferenciou em todas as idades analisadas. No número de *tokens* total a idade de 4 anos foi inferior à idade de 5 anos. A idade de 3 anos não difere de nenhuma das outras.

Em estudo anterior verificou-se que o número total de palavras (*tokens*) e o número de diferentes palavras (*types*) de uma amostra de linguagem em extensão fixa têm demonstrado aumentar junto com a idade e tem sido concebido como “índice de facilidade lingüística”, que reflete inúmeros fatores como: extensão de enunciado, maturação motora da fala, habilidade para formulação de enunciado e ênfase com palavras.<sup>(11)</sup>

Constatou-se um aumento significativo no número de *types* e *tokens* nas faixas etárias de dois anos e dois anos e meio quando comparadas à faixa etária de um ano e meio. Os testes estatísticos demonstraram uma diferença significativa entre um ano e meio, dois anos e dois anos e meio, diferenciando assim as faixas etárias analisadas.<sup>(14)</sup>

## ■ CONCLUSÃO

Observou-se que os dados de *type*, *token* e relação *type/token* de 3 a 5 anos foram menos produtivos do que aos dados de 1 ano e 6 meses, 2 anos e 2 anos e 6 meses. No entanto, esta pesquisa revelou um parâmetro para as faixas etárias de 3 a 5 anos, que servirão para a clínica fonoaudiológica, podendo identificar, dessa forma, crianças com desvio ou atraso de linguagem.

**ABSTRACT**

**Purpose:** to establish an evolutionary profile of the type/token relation in the 3, 4 and 5 years old groups, and to verify type/token as regards of speech, considering age group and gender, verifying it's evolutionary validity. **Methods:** 60 children were analyzed in typical development of language in playful activities with the researchers or their mothers. The recording of the spontaneous speech, as well as it's transcription were accomplished. It was also done the calculations of the total and segment type/token (50 statements). **Results:** the types tokens analysis weren't more productive to show a significantly difference in children that are 4;0 and the other groups. **Conclusions:** types and tokens weren't an evaluative measure in language evaluation of children between 3;0 and 5;0 years old.

**Keywords:** Language development; Child language; Vocabulary; Verbal behavior; Speech therapy; Speech, reading

**REFERÊNCIAS**

1. Barret MD, Harris M, Chasin, J. Early lexical development and maternal speech: a comparison of children's initial and subsequent uses of words. *J Child Lang* 1991;18:21-4.
2. Benedict KS. Early lexical development: comprehension and production. *J Child Lang* 1979;6:183-200.
3. Burroughs E. Lexical diversity in listeners judgments of children. *Percep Motor Skill* 1991;73:19-22.
4. Chapman RS. Processos e distúrbios na aquisição da linguagem. Porto Alegre: Artes Médicas; 1996.
5. Reznick SH, Goldsmith PT. A multiple form word production checklist for assessing early language. *J Child Lang* 1989;16:91-100.
6. Golinkoff RM, Hirsh-Pasek K. Reinterpretando a compreensão da frase pela criança: em direção a uma nova estrutura. In: Fletcher P, Macwhinney B. *Compêndio da linguagem da criança*. Porto Alegre: Artes Médicas; 1997. p. 355-78.
7. Hess CW, Ritchie KP, Landry RG. The *type/token* ratio and vocabulary performace. *Psychol Rep* 1984;55:51-7.
8. Hess CW, Haug HT, Landry RG. The reliability of *type/token* ratios for oral language of school age children. *J Speech Hear Res* 1989;32:536-40.
9. Hess CW, Sefton KM, Landry RG. Sample size and *type/token* ratios for oral language of preschool children. *J Speech Hear Res* 1986;29:129-34.
10. Klee T, Fitzgerald MD. The relation between gramatical development and mean length of utterance in morphemes. *J Child Lang* 1985;12:251-69.
11. Klee T. Developmental and diagnostic characteristics of quantitative measures of children's language production. *Top Lang Dis* 1992;12:28-41.
12. Miller JF, Klee T. Abordagens computacionais à análise da deficiência de linguagem. In: Fletcher P, Macwhinney B. *Compêndio da linguagem da criança*. Porto Alegre: Artes Médicas; 1997. p. 447-66.
13. Miller JF, Freiberg C, Rolland MB, Reeves MA. Implementing computerized language sample analysis in the public school. *Top Lang Dis* 1992;12:69-82.
14. Scherer S. Perfil evolutivo da relação *type/token* em crianças de 1 ano e meio, 2 anos e 2 anos e meio. [monografia]. Canoas: Universidade Luterana do Brasil ULBRA; 1999.

Recebido para publicação em: 30/01/2002

Aceito em: 24/05/2002

Endereço para correspondência

Nome: Sabrina Scherer

Endereço: Rua Independência, 265 – CEP: 97970-000 – Roque Gonzales – RS

Fone/fax: (55) 3365-1141

sscherer@pro.via-rs.com.br