

# TYPES E TOKENS NA AQUISIÇÃO TÍPICA DE LINGUAGEM POR SUJEITOS DE 18 A 32 MESES FALANTES DO PORTUGUÊS BRASILEIRO

## *Types and tokens in typical language acquisition of Brazilian Portuguese subjects among 18 and 32 months*

Sabrina Scherer<sup>(1)</sup>, Ana Paula Ramos de Souza<sup>(2)</sup>

### RESUMO

**Objetivo:** analisar comparativamente a mudança em types e tokens e na taxa de type/token em crianças, de ambos os sexos, entre 18 e 36 meses, falantes nativos do português brasileiro, quanto à classe gramatical e à medida total e segmentar. **Método:** foram gravadas e transcritas as falas de 60 crianças com desenvolvimento típico de linguagem em atividades lúdicas com as mães ou professora. A seguir, efetuaram-se os cálculos da taxa de type/token (TTR) e de types e tokens (ty/to), na forma total e segmentar, com sua distribuição em classes gramaticais. Os resultados foram analisados estatisticamente, por meio dos testes t-student e análise de variância (ANOVA). **Resultados:** o número de types e tokens totais apresentou vantagens sobre o segmentar em termos de descrição de classes gramaticais e na diferenciação estatística das faixas etárias de 18, 24 e 32 meses. Evolutivamente substantivos surgem primeiro do que verbos, advérbios e adjetivos e demais classes gramaticais que se completam até 32 meses. **Conclusão:** não houve diferença estatística quanto ao gênero. A análise isolada dos types e tokens é mais efetiva do que a divisão dos mesmos (TTR), tendo validade estatística na diferenciação das faixas etárias de 18, 24 e 32 meses. O número de types e tokens segmentar é menos efetivo na descrição gramatical considerando faixa etária e gênero.

**DESCRITORES:** Avaliação; Testes de Linguagem; Linguagem Infantil; Vocabulário

### ■ INTRODUÇÃO

A taxa da relação type/token – número dos diferentes itens lexicais produzidos dividido pelo total de itens lexicais ou TTR – é uma medida da produção lingüística para estimar a proficiência lexical, ou seja, é uma forma de verificar a diversidade lexical ou a variedade de diferentes palavras faladas pela criança<sup>1</sup>.

Templin (1957), o criador da medida, examinou TTRs calculadas em amostras de 50 enunciados de 480 crianças entre 3 e 8 anos de idade. Verificou que, embora os números de ambos, tipos (types) e palavras faladas (tokens), aumentassem entre 3 e 8 anos, o número total de palavras, correspondia, na amostra, ao dobro do número de palavras diferentes (1:2)<sup>2</sup>.

A medida da diversidade lexical, por meio do número de types e de tokens, tem sido mundialmente utilizada em estudos de aquisição da linguagem<sup>3,4</sup>, na investigação de modelos de processamento da linguagem<sup>5-8</sup> e de lesões cerebrais<sup>9</sup>, demonstrando relevância enquanto forma de padronização de linguagem típica e atípica, como variável a ser controlada em medidas de fala ou escrita, tanto comparando a linguagem de distintos sujeitos, como a linguagem do mesmo sujeito em dois momentos distintos.

Os estudos analisados apontam a importância de padrões de referência para a medida de types e

<sup>(1)</sup> Fonoaudióloga; Supervisora de Atividades Práticas da Faculdade Fátima e Responsável pelo Setor de Fonoaudiologia do Hospital Virvi Ramos de Caxias do Sul; Especialização em Linguagem pelo CEFAC – Saúde e Educação; Mestranda em Envelhecimento Humano pelo Programa de Pós-Graduação da Universidade de Passo Fundo – UPF, Rio Grande do Sul.

<sup>(2)</sup> Fonoaudióloga; Docente do Curso de Graduação e Mestrado em Distúrbios da Comunicação Humana da Universidade Federal de Santa Maria – UFSM, Santa Maria, Rio Grande do Sul, Brasil; Doutora em Letras pela Pontifícia Universidade Católica do Rio Grande do Sul – PUCRS.

Conflito de interesses: inexistente

tokens por faixa etária, tomando a aquisição típica de linguagem como referência para a diferenciação lingüística de sujeitos em aquisição atípica de linguagem, incluindo falantes tardios ou portadores de Distúrbio Específico de Linguagem (DEL).

Estudos de medidas de linguagem espontânea em crianças adquirindo o português brasileiro foram realizados para a extensão média de enunciados (MLU – Mean Length Utterance) nas faixas etárias de 1, 2, 3 e 5 anos<sup>10</sup> e para types e tokens em crianças de 3 a 5 anos<sup>11</sup>. No entanto, não havia estudos similares para a medida de types e tokens na faixa etária de 18, 24 e 32 meses, momento crucial da aquisição lexical, tanto por captar o ritmo inicial da evolução da diversidade lexical quanto por captar o momento de explosão do vocabulário<sup>12</sup>.

Assim, este trabalho tem como objetivos analisar comparativamente a mudança em types e tokens e a taxa de type/token em crianças, de ambos os sexos, entre 18 e 36 meses, falantes nativos do português brasileiro, quanto à classe gramatical e à medida total e segmentar.

## ■ MÉTODO

Foi realizada uma pesquisa de campo com 60 crianças, distribuídas em três faixas etárias (FE),

em desenvolvimento típico de linguagem (DTL), falantes do português brasileiro (PB), identificadas a partir da observação do professor e/ou da mãe como estando com linguagem esperada para sua FE. Tal aspecto foi confirmado pelas pesquisadoras por meio da observação do brincar, da interação comunicativa, da etapa cognitiva, da autonomia discursiva e da esfera afetiva, com base nos critérios de Golinkoff e Hirsh-Pasek<sup>12</sup> para referência do previsto em termos evolutivos para a faixa etária, sobretudo no que diz respeito aos períodos de maior proeminência de cada faixa etária escolhida.

As FE analisadas foram alternadas de seis em seis meses no desenvolvimento inicial, considerando a explosão de vocabulário que ocorre nesse momento. Foram três faixas etárias: FE1= 20 crianças de 18 meses (idade mínima 1 ano e 3 meses e máxima 1 ano e 8 meses), FE2= 20 crianças de 24 meses (idade mínima 1 ano e 9 meses e máxima 2 anos e 3 meses) e FE3= 20 crianças de 32 meses (idade mínima 2 anos e 5 meses e máxima 2 anos e 8 meses). A variação individual para cada faixa etária foi no máximo de 15 dias acima da idade limite, ou seja, cada criança poderia se afastar da sua faixa etária no máximo em 15 dias. Para cada faixa etária, havia 10 meninas e 10 meninos. O detalhamento da amostra considerando sexo e idade está nas Figuras 1, 2 e 3.

Sujeito-Sexo	Idade	Sujeito-Sexo	Idade
Sujeito 1 Feminino	19 meses	Sujeito 1 Masculino	20 meses
Sujeito 2 Feminino	17 meses	Sujeito 2 masculino	19 meses
Sujeito 3 Feminino	18 meses	Sujeito 3 Masculino	16 meses
Sujeito 4 Feminino	16 meses	Sujeito 4 Masculino	20 meses
Sujeito 5 Feminino	19 meses	Sujeito 5 Masculino	18 meses
Sujeito 6 Feminino	18 meses	Sujeito 6 Masculino	20 meses
Sujeito 7 Feminino	16 meses	Sujeito 7 Masculino	18 meses
Sujeito 8 Feminino	19 meses	Sujeito 8 Masculino	18 meses
Sujeito 9 Feminino	20 meses	Sujeito 9 Masculino	18 meses
Sujeito 10 Feminino	18 meses	Sujeito 10 Masculino	18 meses

**Figura 1 – Faixa Etária 1**

Sujeito-Sexo	Idade	Sujeito-Sexo	Idade
Sujeito 1 Feminino	24 meses	Sujeito 1 Masculino	27 meses
Sujeito 2 Feminino	21 meses	Sujeito 2 masculino	27 meses
Sujeito 3 Feminino	24 meses	Sujeito 3 Masculino	24 meses
Sujeito 4 Feminino	24 meses	Sujeito 4 Masculino	27 meses
Sujeito 5 Feminino	24 meses	Sujeito 5 Masculino	27 meses
Sujeito 6 Feminino	24 meses	Sujeito 6 Masculino	24 meses
Sujeito 7 Feminino	27 meses	Sujeito 7 Masculino	24 meses
Sujeito 8 Feminino	24 meses	Sujeito 8 Masculino	26 meses
Sujeito 9 Feminino	26 meses	Sujeito 9 Masculino	22 meses
Sujeito 10 Feminino	24 meses	Sujeito 10 Masculino	23 meses

**Figura 2 – Faixa Etária 2**

Sujeito-Sexo	Idade	Sujeito-Sexo	Idade
Sujeito 1 Feminino	28 meses	Sujeito 1 Masculino	30 meses
Sujeito 2 Feminino	32 meses	Sujeito 2 masculino	32 meses
Sujeito 3 Feminino	30 meses	Sujeito 3 Masculino	32 meses
Sujeito 4 Feminino	29 meses	Sujeito 4 Masculino	29 meses
Sujeito 5 Feminino	29 meses	Sujeito 5 Masculino	30 meses
Sujeito 6 Feminino	29 meses	Sujeito 6 Masculino	30 meses
Sujeito 7 Feminino	31 meses	Sujeito 7 Masculino	31 meses
Sujeito 8 Feminino	32 meses	Sujeito 8 Masculino	30 meses
Sujeito 9 Feminino	30 meses	Sujeito 9 Masculino	32 meses
Sujeito 10 Feminino	30 meses	Sujeito 10 Masculino	30 meses

**Figura 3 – Faixa Etária 3**

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.

A coleta de dados foi realizada pelas pesquisadoras em escolas infantis e/ou lar das crianças, por meio de gravação da fala espontânea em atividades lúdicas de interesse da criança, nas quais as mães ou professoras previam que haveria possibilidade de diálogo. A gravação teve duração de 45 minutos. Foram, portanto, descartados brinquedos em que a atividade psicomotora predominasse sobre a fala e a interação com pessoa íntima à criança (mãe ou professora). Buscou-se que a criança interagisse com a pessoa com quem ela passava a maior parte do dia e, por isso, em alguns casos, a professora foi privilegiada e em relação à mãe. Em alguns casos isso não foi possível no ambiente escolar, o que demandou uma visita domiciliar para a gravação.

Os dados foram registrados num gravador da marca Panasonic, modelo Rn 202, com fita da marca Sony Dat de 90 minutos. O próprio ruído da parada automática do gravador indicou o término da avaliação. Foram consideradas as sessões em que os adultos, a mãe ou a professora consideraram que a criança apresentou uma interação dialógica que lhe é comum, ou seja, em que a fala fluiu como usualmente ocorre um interlocutor privilegiado. Nos casos em que isso não ocorreu, a gravação foi refeita, procurando obter o melhor exemplo de possibilidades lingüísticas da criança. O pesquisador responsável, experiente na área de linguagem infantil, tomando os enunciados em seqüência, fez julgamentos acerca desse aspecto e também da ausência de sinais de distúrbios específicos da linguagem ou transtornos de desenvolvimento.

Após essa gravação, foi realizada a transcrição da fala das crianças e calculada a taxa entre types e tokens (types/tokens ratio – TTR), pelos critérios de Templin<sup>2</sup>:

$$\text{Type/token} = \frac{\text{número de palavras diferentes da amostra}}{\text{número total das palavras da amostra}}$$

Na TTR total foram analisados, a partir de todos enunciados infantis durante 45 minutos, os valores individuais do número de palavras diferentes (types) e o total de palavras enunciadas (tokens), para cada uma das seguintes categorias gramaticais: substantivos, verbos, preposições, artigos, pronomes, conjunções, adjetivos e advérbios. A seguir, verificou-se a soma total de types e tokens constituindo uma fração de type/tokens (X/Y), cuja comparação entre as faixas etárias incidiu sobre o número de types (X na faixa 1 e na faixa 2) e de tokens (Y na faixa 1 e na faixa 2) em separado e não sobre o resultado da operação de divisão. Por fim, o número resultado da divisão do número de types pelo número de tokens (taxa). A TTR segmentar foi realizada do mesmo modo, mas apenas sobre 50 enunciados contados a partir da segunda folha de transcrição.

A presente pesquisa foi um subprojeto do projeto de “Testes e medidas de linguagem: criação de um protocolo de avaliação”, aprovado pelo Comitê de Ética em Pesquisa da Universidade Luterana do Brasil sob protocolo CEP 106-2002.

Os resultados foram analisados estatisticamente, por meio dos testes t-student e análise de variância (ANOVA), para verificar diferenças entre grupos da mesma faixa etária e entre as distintas faixas etárias respectivamente.

## ■ RESULTADOS

Na Tabela 1 são expostos os resultados totais tanto de Types e Tokens como de TTR (type/token ratio) por faixa etária, por meio da análise de variância ANOVA. Verifica-se que há diferenças estatísticas significantes entre as faixas etárias. Observa-se que os três grupos estudados diferem entre si ( $p=0.001$ ), mas não para a TTR média que se manteve próxima para todas as faixas etárias. A análise da variância (ANOVA) mostrou que há diferenças estatisticamente significantes entre as faixas etárias de 18, 24 e 32 meses para as medidas de

Tabela 1 – Types, Tokens e TTR Totais Por Faixa Etária

Faixa Etária	Média Types	Desvio-Padrão	Média Tokens	Desvio-padrão	Média TTR	Desvio-Padrão
18 meses	60,89	34,60	191,16	124,31	0,37	0,14
24 meses	124,11	34,26	389,39	162,79	0,35	0,10
32 meses	158,26	30,20	541,11	166,77	0,30	0,06

Legenda: Faixa Etária – 18, 24 e 32 meses

Média Types, Média Tokens, Média TTR

Teste ANOVA ( $p=0.001$ )

types e tokens. Note-se que não há mudança no desvio-padrão apresentado em relação à TTR, visto que não difere nas faixas citadas, mas apenas o número de types e tokens isoladamente.

Na Tabela 2 são expostos os resultados de Types, Tokens e TTR segmentar (50 enunciados) por faixa etária. Por meio da análise de variância (ANOVA), foram encontradas diferenças estatísticas

significantes entre os grupos de faixa etária em relação ao valor médio de type. Observa-se que as crianças de 24 e 32 meses possuem um valor médio superior a de 18 meses ( $p=0,001$ ), mas não possuem diferença estatisticamente significativa entre si, demonstrando uma espécie de desaceleração na diversidade vocabular entre essas duas faixas etárias.

Tabela 2 – Types, Tokens e TTR Segmentar (50 Enunciados) Por Faixa Etária

Faixa Etária	Média Types	Desvio-Padrão	Média Tokens	Desvio-Padrão	Média TTR	Desvio-Padrão
18 meses	29,80	12,07	71,10	26,76	0,44	0,13
24 meses	58,37	17,13	125,32	33,00	0,46	0,06
32 meses	64,60	10,28	142,10	35,76	0,46	0,007

Legenda: Faixa Etária – 18, 24 e 32 meses

Média Types, Média Tokens, Média TTR

Teste ANOVA ( $p=0.001$ )

Não foi encontrada variação relevante nos coeficientes da divisão da TTR entre as faixas etárias. Pode-se verificar, no entanto, na comparação isolada do número de types e de tokens uma diferença estatisticamente relevante entre as faixas etárias estudadas. Observa-se que tanto no número de types como no número de tokens, por meio da análise da variância (ANOVA), as crianças de 24 e 32 meses possuem um valor médio superior às de 18 meses ( $p=0,001$ ).

Em relação à variável sexo, analisada por meio do teste t-student, não houve diferenças significativas entre as faixas etárias, tanto para a medida total quanto para a segmentar ( $p=0,55$ ).

A TTR total apresentou resultados mais importantes do que a segmentar em termos de classes gramaticais, demonstrando claramente que, na

faixa etária de 18 meses, substantivos constituem a classe dominante na fala espontânea da criança, enquanto verbos, advérbios, pronomes, artigos, adjetivos e conjunções são menos frequentes. Nas faixas etárias de 24 e 32 meses, há uma tendência de os verbos se igualarem ou superarem os substantivos. Essas duas classes são seguidas pelos advérbios. Praticamente todas as classes gramaticais estão presentes na faixa etária de 32 meses, conforme exposto na Tabela 3 na qual estão as TTR total e segmentar e as frações de types/tokens por faixa etária e categoria gramatical, de modo a serem visualizados os detalhes oferecidos por cada tipo de medida. Ressalta-se que a TTR não se apresentou como uma boa medida na diferenciação das faixas etárias.

Tabela 3 – Distribuição De Types E Tokens Por Categoria Gramatical

	18 meses		24 meses		32 meses	
	TTR 50	TTR Total	TTR 50	TTR Total	TTR 50	TTR Total
<b>Substantivos</b>	17/38	31/86	21/37	46/100	18/26	52/104
<b>Verbos</b>	6/15	15/43	18/37	37/105	20/47	55/191
<b>Artigos</b>	0,9/4	2/15	2/9	3/30	4/11	4/38
<b>Preposições</b>	0,7/0,9	¼	2/6	4/15	3/5	6/22
<b>Pronomes</b>	1/2	2/4	5/12	8/38	7/19	14/74
<b>Conjunções</b>	0,2/0,5	0,6/1	0,7/2	2/3	1/3	2/11
<b>Advérbios</b>	3/10	5/25	6/17	15/54	9/27	22/108
<b>Adjetivos</b>	0,5/1	1/3	0,8/1	3/5	3/4	8/14

Legenda: Faixa Etária – 18, 24 e 32 meses

Categorias Gramaticais: substantivos, verbos, artigos, preposições, pronomes, conjunções, advérbios, adjetivos  
TTR 50 e TTR Total

## ■ DISCUSSÃO

Nesta pesquisa pôde-se constatar um aumento significativo no número de types e tokens nas faixas etárias de 24 e 32 meses quando comparadas à faixa etária de 18 meses, conforme comprovado estatisticamente. Não houve, no entanto, distinção estatística em relação ao resultado da fração entre types e tokens nas três faixas etárias. Isso ocorre porque o resultado da divisão não explicita o domínio lexical. Assim, por exemplo, se uma criança que produz 100 types e 200 tokens ela teria a mesma TTR ou taxa de vocabulário, que é o resultado da divisão, igual a de outra que produz 300 types e 600 tokens, ou seja, de 0,5, mas a primeira criança teria uma diversidade vocabular 50% menor que a segunda. Esses dados confirmam aqueles de Klee<sup>13</sup>, que verificou que o número total de palavras (tokens) e o número de palavras diferentes (types) de uma amostra de linguagem em extensão fixa aumentam em função da idade, o que foi designado como “índice de facilidade lingüística”, o qual reflete inúmeros fatores, tais como: maturação motora da fala, habilidade para formular o enunciado e fazê-lo de modo extenso, ou seja, de produzir uma organização sintática mínima como sintagma nominal e verbal, até a possibilidade de produzir orações encaixadas que demandam um conhecimento sintático e lexical maior (maior número de conjunções, pronomes, artigos entre outros). Cabe ressaltar que os estudos nessa abordagem teórica, que descendem do gerativismo, não entram em questão sobre o quanto antes dos 24 meses há ou não análise lingüística e se os itens lexicais são evidência de domínio categorial, apenas demonstram em diversas pesquisas o fato de as medidas de types e tokens isoladas diferenciarem mais

as evoluções lingüísticas do que o resultado da fração<sup>3-8,11</sup>. É claro que tais resultados poderiam ser questionados à luz de abordagens como a interacionista brasileira, incompatível com medidas de linguagem como a aqui investigada.

Nos resultados desta pesquisa a distinção entre types e tokens se mantém tanto para a relação de types/tokens total (transcrição de 45 minutos), quanto para a segmentar (transcrição de 50 enunciados), demonstrando que para medidas de triagem, a relação segmentar permite estimar uma idade lingüística para o sujeito. Esse fato é debatido por Hess et al.<sup>14</sup>, que afirmam que, com o aumento do número de enunciados da amostra, as palavras tendem a se repetir, e a TTR conseqüentemente diminui. O número de types de 50 enunciados na idade de 3 anos foi inferior às idades de 4 e 5 anos. Verificou-se que o número de tokens de 50 enunciados foi diferente 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. Na pesquisa Scherer et al.<sup>11</sup>, verificou-se que a medida tem sensibilidade relativa para diferenciar essas faixas etárias, não devendo ser utilizada exclusivamente como avaliação de linguagem nas mesmas. Pode ser complementada por testes de vocabulário padronizados para o PB<sup>15</sup>, caso se busque analisar o conhecimento lexical de itens freqüentes no vocabulário infantil inicial. No entanto, ressaltamos, que tais avaliações são uma medida estática do vocabulário, que não captam o funcionamento lingüístico, conforme já abordado em um estudo no qual se discutiram as técnicas de avaliação por meio da elicitación visual e a análise da linguagem espontânea<sup>16</sup>. Neste estudo fica claro que crianças muito pequenas (entre 12 e 18 meses) podem

não entender a tarefa de nomeação de uma figura mas falam os itens em diálogo. Por isso, ressalta-se que a medida aqui apresentada é descritiva e não explanatória, podendo servir de referência para uma triagem de linguagem, com poder diagnóstico limitado para a terapêutica. Eventualmente, tal medida pode refletir alguns sinais patológicos, como, por exemplo, crianças com dificuldades do tipo anômica que apresentam *types* de substantivos muito distintos de crianças sem esta limitação processual<sup>17</sup>.

Outro aspecto importante desta medida é a possibilidade de dar valores específicos por categoria gramatical<sup>1,2</sup>. Os resultados apontaram um incremento importante nos verbos na faixa etária de 24 meses, aproximando esta categoria dos substantivos e sugerindo a entrada no período sintático. Parece haver um conseqüente incremento em categorias que estão mais diretamente ligadas à função sintática (pronomes, conjunções e preposições) e na atribuição de qualidade e especificações de modo, por meio do incremento gradativo dos adjetivos e advérbios. Embora algumas produções das crianças pudessem ainda não ter o status gramatical que possuem para o adulto, ou seja, fossem produções que abrangem processos como a especularidade, considerou-se o fato de a partir de 24 meses já haver uma forma sintática similar à forma adulta. Por isso, afirma-se o poder descritivo e não explanatório da medida, perceptível em vários estudos internacionais<sup>19,20</sup>.

Em relação à variável sexo, não foram observadas distinções estatisticamente significantes entre meninos e meninas, demonstrando um equilíbrio evolutivo entre os grupos. Os mesmos resultados foram encontrados na análise do perfil evolutivo de *types*, *tokens* e da relação *type/token* de crianças de 3, 4 e 5 anos<sup>1</sup> e também em trabalhos como o de Klee et al.<sup>13</sup>. Em contrapartida, Le Normand Parisse e Cohen<sup>18</sup> examinaram a influência do sexo e fatores socioculturais da produção de linguagem de 316 amostras de fala de crianças francesas, onde foram efetuadas medidas de desenvolvimento gramatical e lexical incluindo, MLU e relação *type token*. O resultado estatístico do estudo revelou forte influência de fatores socioculturais, demonstrando que crianças de famílias de alto nível socio-cultural têm produção lexical mais complexa e uma maior taxa de desenvolvimento. A análise também

revelou um efeito geral de gênero, mostrando uma pequena vantagem na produção de linguagem para as meninas sobre os meninos até 36 meses de idade. Tais resultados não foram verificados neste estudo porque buscaram-se crianças de um mesmo nível sócio-cultural (classe média ou classe média-baixa).

Em vários países há estudos utilizando softwares para a descrição de amostras de linguagem espontânea<sup>21,22</sup>, mas não podem ser esquecidas as limitações diagnósticas de tais descrições. Podem servir apenas de referência evolutiva para pesquisas que adotem a ótica quantitativa com o intuito de demonstrar estatisticamente mudanças na linguagem infantil.

## ■ CONCLUSÃO

Este trabalho permitiu visualizar uma evolução de *types* e *tokens* de 18 a 32 meses, demonstrando que tais medidas diferenciam estatisticamente as faixas etárias analisadas tanto com base na transcrição integral de 45 minutos de gravação como nos 50 enunciados para o cálculo das mesmas. A variável sexo não apresentou interferência nos resultados obtidos.

Observou-se que os dados de *type*, *tokens* e TTR segmentar são inferiores quando comparados à total, sendo esta mais rica, em termos de análise do que a segmentar. Os valores individuais dos *types* e dos *tokens* demonstraram ser mais significantes do que os valores dos coeficientes da divisão *type/token*, uma vez que podem mascarar diferenças lingüísticas importantes.

Os resultados demonstram, no entanto, que se trata de medida descritiva e não explanatória do funcionamento lingüístico, apresentando limitações conceituais quanto à presença de fala formulaica ou analisada e podendo variar de acordo com o contexto e a habilidade do interlocutor para interagir verbalmente com a criança. Por isso, os dados, com alguns sujeitos, tiveram sua coleta repetida de modo a permitir a melhor interlocução adulto-criança e este é um aspecto crucial a ser considerado em medidas de linguagem espontânea. Além disso, uma medida nunca capta a singularidade lingüística do sujeito, mas descreve padrões usuais em uma determinada faixa etária.

**ABSTRACT**

**Purpose:** to comparatively analyze the change in types and tokens and the rate of type / token in children of both genders, aged between 18 and 36 months, native speakers of Brazilian Portuguese, as for grammatical category and total measure and segmental aspect. **Method:** the speech of sixty children with typical language acquisition was recorded while playing with their mothers or teachers. The type/token ratio (TTR) and the total and segmental types and tokens were calculated, as for their grammatical distribution. The results were submitted to statistical analyses by Student Test and Variance Analysis (ANOVA). **Results:** the total number of types and the tokens is more effective to describe children grammatical evolution, among 18 and 32 months age, than the segmental one. Substantives emerge first than verbs, adverbs, adjectives and others grammatical categories. All of the grammatical categories are in production at the 32 months of age. **Conclusion:** no statistical differences were found among genders. The analysis of types and tokens is more effective than TTR and differentiated statically the ages: 18, 24 and 32 months. The segmental numbers of types and tokens is a less effective measure for the grammatical description in ages and genders.

**KEYWORDS:** Evaluation; Language Tests; Child Language; Vocabulary

**■ REFERÊNCIAS**

- Williamson, G. Type-Token Ratio. *Speak Therapy Information and Resources –STIR*. 2009 [acesso em 05.09.2010]. Disponível em: <http://www.speech-therapy-information-and-resources.com/type-token-ratio.html>
- Templin MC. *Certain language skills in children: their development and interrelations*. Westport, CT: Greenwood; 1957.
- Dabrowska E, Szczerbinski M. Polish children's productivity with case marking: the role of regularity, type frequency, and phonological diversity. *J Child Lang*. 2006;33(3):559-97.
- Nicoladis E, Palmer A, Marentette P. The role of type and token frequency in using past tense morphemes correctly. *Dev Sci*. 2007;10(2):237-54.
- Heisler L, Goffman L, Younger B. Lexical and articulatory interactions in children's language production. *Dev Sci*. 2010; 13(5):722-30.
- Paradis J, Nicoladis E, Crago M, Genesee F. Bilingual children's acquisition of the past tense: a usage-based approach. *J Child Lang*. 2010; 26:1-25.
- Endress AD, Hauser MD. The influence of type and token frequency on the acquisition of affixation patterns: Implications for language processing. *J Exp Psychol Learn Mem Cogn*. 2010. [Epub ahead of print].
- Adelman JS, Brown GD, Quesada JF. Contextual diversity, not word frequency, determines word-naming and lexical decision times. *Psychol Sci*. 2006;17(9):814-23.
- Borovsky A, Saygin AP, Bates E, Dronkers N. Lesion correlates of conversational speech production deficits. *Neuropsychologia*. 2007 Jun 18;45(11):2525-33.
- Fensterseifer A, Ramos APF. Extensão média de enunciados em crianças de 1 a 5 anos. *Pro Fono Rev. At. Ci*. 2003;15(3):251-8.
- Scherer S, Casarim F, Zart P, Ramos A. Perfil evolutivo da relação type/token de crianças de 3, 4 e 5 anos de idade. Porto Alegre: Trabalho de especialização, CEFAC 2002.
- Golinkoff R; Hirsh-Pasek K. Reinterpretando a compreensão da frase pela criança em direção a uma nova estrutura. IN Fletcher P, MacWhinney B. *Compêndio de linguagem da criança*. Porto Alegre: Artmed; 1997. p. 355-78.
- Klee T, Stokes SF, Wong AM, Fletcher P, Gavin WJ. Utterance length and lexical diversity in Cantonese-speaking children with and without specific language impairment. *J Speech Lang Hear Res*. 2004;47(6):1396- 410.
- 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.
- Befi-Lopes D M. Verificação do Vocabulário in Andrade CMF, Fernandes FDM, Wertzner HF ABFW- Teste para avaliação de linguagem infantil. 1ª.ed. São Paulo: Pró-fono Editora, 2000, v.1, p.41-59.
- Bastos JC; Ramos APF; Marques J. Estudo do vocabulário infantil: limitações das metodologias tradicionais de coleta. *Revista da Soc. Bras. Fonoaudiologia* 2004; 9(1):1-9.
- Peets KF. Profiles of dysfluency and errors in classroom discourse among children with language impairment. *J Commun Disord*. 2009; 42(2):136-54.

- 18 Le Normand MT, Parisse C, Cohen H. Lexical diversity and productivity in French preschoolers: developmental, gender and sociocultural factors. *Clin Linguist Phon.* 2008;22(1):47-58.
19. Nicoladis E, Palmer A, Marentette P. The role of type and token frequency in using past tense morphemes correctly. *Dev Sci.* 2007; 10(2):237-54.
20. Vigário M, Freitas MJ, Frota, S. Grammar and frequency effects in the acquisition of the Prosodic Word in European Portuguese. *Language and Speech.* 2006; 49(2):175-203.
21. Boudelaa S, Marslen-Wilson WD. Aralex: a lexical database for Modern Standard Arabic. *Behav Res Methods.* 2010 May;42(2):481-7.
22. Hofmann MJ, Stenneken P, Conrad M, Jacobs AM. Sublexical frequency measures for orthographic and phonological units in German. *Behav Res Methods.* 2007; 39(3):620-9.

<http://dx.doi.org/10.1590/S1516-18462011005000058>

RECEBIDO EM: 07/05/2010

ACEITO EM: 06/12/2010

Endereço para correspondência  
Sabrina Scherer  
Rua Marques do Herval 386/601, Centro  
Caxias do Sul –RS  
E-mail:sabrischerer@yahoo.com.br