

EXTENSÃO MÉDIA DO ENUNCIADO-PALAVRAS EM CRIANÇAS DE 4 E 5 ANOS COM DESENVOLVIMENTO TÍPICO DA LINGUAGEM

Mean Length of Utterance-words in children with typical language development aged 4 to 5 years

Maria Emília Santos ⁽¹⁾, Sofia Lynce ⁽¹⁾, Sara Carvalho ⁽¹⁾, Mariana Cacela ⁽¹⁾, Ana Mineiro ⁽¹⁾

RESUMO

Objetivo: analisar a Extensão Média do Enunciado-palavras (EME-p) em crianças entre os 4;00 e os 5;05. **Métodos:** foram observadas 92 crianças portuguesas com desenvolvimento típico: 49 meninas e 43 meninos, divididas em grupos etários com 6 meses de intervalo. Foi recolhida para cada criança uma amostra de 100 enunciados produzidos em discurso espontâneo. Os enunciados foram transcritos e analisados. **Resultados:** a EME-p variou de 4,5 a 5 palavras, aumentando com a idade. Esta progressão foi verificada anteriormente em crianças falantes de Inglês dos EUA e de Português do Brasil, embora no Português Europeu o número de palavras seja, no geral, um pouco superior. O desempenho de meninos e meninas foi idêntico. A escolaridade dos pais mostrou ter alguma influência, mas não em todos os grupos etários. Os resultados mostraram uma correlação positiva e significativa com um teste formal de linguagem, tanto na compreensão, como na expressão. **Conclusão:** a EME-p é uma boa medida de desenvolvimento da linguagem até aos 5 anos. Os valores encontrados podem servir como referência normativa relativamente às crianças portuguesas, mas também em estudos comparativos sobre o desenvolvimento da linguagem espontânea.

DESCRITORES: Linguagem; Desenvolvimento; Criança; Avaliação

■ INTRODUÇÃO

O uso de testes de avaliação de linguagem em crianças, apesar de interessante para o diagnóstico, não exclui a necessidade de compreender melhor o desempenho linguístico da criança num contexto natural. A recolha de amostras de discurso espontâneo torna-se importante, pois estas estão mais perto do ambiente quotidiano e da experiência da criança, ou seja, dos seus hábitos de comunicação, rotinas e parceiros. Vários autores apoiam este tipo de avaliação, uma vez que boa parte da pesquisa linguística envolve valores e dados qualitativos longitudinais, em que o processo de aquisição e desenvolvimento da linguagem é monitorizado considerando certos padrões que se desdobram

e mudam através do tempo¹. Por outro lado, muitas críticas têm sido levantadas ao processo formal de avaliação em favor de uma observação naturalística^{2,3}. Esta aproxima-se mais do comportamento linguístico da criança no seu dia-a-dia e permite observar o uso da linguagem em diferentes contextos⁴.

Apesar das vantagens na utilização de amostras de discurso espontâneo sobre os testes formais existem limitações, nomeadamente a forma como decorreu o processo de recolha e o número de enunciados recolhidos e analisados⁵. Devido a estas dificuldades, alguns autores têm-se dedicado à otimização deste processo, propondo diversas estratégias. O primeiro passo para a análise da produção linguística da criança é a obtenção fidedigna destas produções, devendo a interação comunicativa ser conduzida de modo a que a amostra resultante seja representativa da produção normal da criança, ou seja, descreva a sua

⁽¹⁾ Instituto de Ciências da Saúde da Universidade Católica Portuguesa, Lisboa, Portugal.

Conflito de interesses: inexistente

produção linguística usual, incluindo linguagem que possa ser um pouco inferior ou um pouco superior à sua performance comum⁶. Já nas décadas de 70 e 80, foi chamada a atenção para os materiais a usar na elicitación do discurso espontâneo, os diferentes contextos onde pode ser recolhido, a sua forma de registo ou a dimensão da amostra^{7,8}. Relativamente a este último aspecto Brown⁷ defendia que uma amostra de 100 enunciados seria suficiente. Outros autores têm sugerido amostras maiores e menores, sendo uma alternativa a determinação de um período de tempo 30 minutos, por exemplo, independentemente de quantas produções ocorram nesse período⁶. Geralmente, estes 30 minutos são suficientes para obter entre 100 a 200 enunciados para crianças com idades acima dos dois anos⁸.

Extensão Média do Enunciado

A extensão média do enunciado (EME) é uma das medidas de linguagem que pode ser obtida através do discurso espontâneo e que tem como principal objetivo a obtenção de dados acerca dos aspetos morfológicos e sintáticos da linguagem das crianças com desenvolvimento típico ou com perturbação da linguagem^{2,7,9}. Este conceito derivou do conceito de extensão média de resposta, de Margaret Nice¹⁰, que já em 1925 considerava que o comprimento das frases deveria ser um dos critérios mais importantes para avaliar o progresso linguístico, sendo assim um marcador de maturação linguística. Atualmente a extensão média do enunciado em palavras (EME-p) tem vindo a ser considerada muito útil, mantendo-se a sua fórmula de cálculo igual à original: número total de palavras a dividir pelo número total de enunciados produzidos².

Após a publicação dos estudos de Brown⁷ as medidas de extensão de enunciados produzidos por crianças ganharam popularidade entre os investigadores. A extensão média do enunciado em morfemas (EME-m) proposta por Brown é menos utilizada do que a EME-p porque esta última é mais fácil de analisar^{4,11} e, embora as duas medidas foquem aspetos linguísticos diferentes, foram observadas correlações elevadas entre elas^{12,13}, o que levou muitos investigadores a preferir a EME-p por ser de mais fácil análise. Alguns autores sugerem que a EME-p seja utilizada, principalmente, quando se pretende realizar estudos comparativos entre línguas, uma vez que a contagem de palavras minimiza as diferenças morfológicas que podem interferir no seu cálculo¹⁴. Independentemente da variante utilizada, parece existir um patamar por volta dos 5 anos em que a EME estabiliza¹⁵ ou tem mesmo um declive¹⁶, desaparecendo a correlação

existente entre idade e EME¹⁷, pois as diferenças nos enunciados já não são ao nível da sua extensão, mas sim da sua complexidade¹⁸.

Apesar da maioria dos estudos não fazer distinção entre géneros, é uma questão importante de analisar visto que ao nível do desenvolvimento da linguagem, alguns resultados referem que os meninos começam a produzir as primeiras palavras e frases mais tarde do que as meninas e que, mesmo ao nível da utilização de pequenos e simples gestos, as meninas começam mais cedo¹⁹. Além disso, há estudos que mostram que as meninas têm um vocabulário mais completo e usam uma maior variedade de frases na sua comunicação mais inicial²⁰. Em idades escolares, as meninas também parecem ter mais sucesso em todas as competências verbais que os meninos²¹. Relativamente à EME não são normalmente observadas diferenças de género²², embora haja alguns resultados discordantes, indicando melhor performance das meninas, mas apenas até por volta dos 3 ou 4 anos de idade²³⁻²⁵.

Outro fator que interessa analisar é a possível interferência do nível sociocultural de origem das crianças, mas os resultados têm sido controversos, dependendo muito das metodologias usadas, incluindo as idades das crianças em estudo e o que se considera como indicador do nível sociocultural. Vários autores observaram uma forte influência desta variável mostrando que as crianças oriundas de nível alto produziam enunciados mais complexos²⁵⁻²⁷; o mesmo também foi constatado relativamente à escolaridade das mães³. Contudo, noutros estudos não foram encontradas relações significantes entre a EME-p e a escolaridade das mães, quer em crianças muito pequenas, cerca de 2 anos²⁸, quer em crianças mais velhas, até aos 9 anos^{9,29}.

Apesar da evidência da elevada qualidade da EME para a identificação e avaliação das produções efetuadas por crianças de língua inglesa e dos diversos estudos realizados noutras línguas, não existem análises relativas ao Português Europeu. No entanto, ao nível do Português do Brasil já foram realizados estudos³⁰ que descrevem um aumento significativo dos valores de EME-m e de EME-p com a idade, em crianças com desenvolvimento típico, entre os 2 e os 4 anos. A Tabela 1 mostra os valores médios de EME em dois estudos: um em Português do Brasil³⁰ e outro em Inglês americano²⁹ relativamente a crianças com desenvolvimento típico. Verifica-se alguma disparidade tanto para a EME-p como para a EME-m.

Tabela 1 – Valores médios de EME-p e de EME-m para crianças brasileiras (PB) e crianças americanas (I) com desenvolvimento típico

Idade	Araújo e Befi-Lopes (2004)		Grupo etário	Rice <i>et al.</i> (2010)	
	EME-p PB	EME-m PB		EME-p I	EME-m I
2 anos	2.35	3.21	2;06 – 2;11	2.91	3.23
3 anos	2.83	3.72	3;00 – 3;05	3.43	3.81
			3;06 – 3;11	3.71	4.09
4 anos	3.52	4.55	4;00 – 4;05	4.10	4.57
			4;06 – 4;11	4.28	4.75
			5;00 – 5;05	4.38	4.88
			5;06 – 5;11	4.47	4.96
	--		6;00 – 5;05	4.57	5.07
	--		6;06 – 6;11	4.70	5.22

O principal objetivo do presente trabalho é estabelecer valores padrão de EME-p para crianças portuguesas dos 4;0 aos 5;05, divididas em faixas etárias de 6 meses, de modo a que esses valores possam servir de indicadores de desenvolvimento da linguagem. Pretende-se ainda verificar se existem diferenças dependentes do género e se a escolaridade dos pais está relacionada com a EME-p.

■ MÉTODOS

Participantes

A amostra incluiu 92 crianças dos 4 aos 5 anos e meio, 43 meninos e 49 meninas, divididas

em três grupos etários com 6 meses de intervalo (Tabela 2). Os dados foram recolhidos entre Janeiro e Maio 2012 em jardins de infância da região de Lisboa. A selecção dos jardins de infância foi feita de forma aleatória, tendo sido incluídas todas as instituições que responderam positivamente ao pedido de colaboração. Dentro de cada instituição foram integradas no estudo todas as crianças cujos pais deram a sua autorização por escrito e que correspondiam aos seguintes critérios de inclusão: terem Português Europeu como língua materna; terem obtido um resultado na avaliação formal da linguagem dentro dos valores esperados para a idade e não terem sido seguidos em Fonoaudiologia. Foram observadas 116 crianças e excluídas 24 por não cumprirem todos estes critérios.

Tabela 2 – Número e género dos participantes por grupo etário

Grupo etário	Género		Total
	Meninos	Meninas	
4;00 – 4;05	12	18	30
4;06 – 4;11	9	21	30
5;00 – 5;05	22	10	32
Total	43	49	92

Considerando a possível influência do nível de escolaridade dos pais, esta variável foi controlada. Utilizou-se como medida o número de anos de

escolaridade do pai ou da mãe, a que fosse mais elevada. A amostra mostrou uma variação entre os 4 anos de escolaridade e o ensino superior (Tabela 3).

Tabela 3 – Escolaridade dos pais por grupo etário

Grupo etário	Nº anos de educação formal		
	4 a 9 anos	10 a 12 anos	> 12 anos
4;00 – 4;05	10	8	12
4;06 – 4;11	10	10	10
5;00 – 5;05	9	5	18
Total	29	23	40

Procedimentos

Todas as crianças foram avaliadas através de um teste português de desenvolvimento da linguagem, dirigido a crianças dos 2 anos e meio aos 6 anos de idade³¹ (TALC). Este teste engloba tarefas relativas a Compreensão [Vocabulário (identificação de objetos e de imagens); Relações Semânticas (relações de duas e três palavras de conteúdo) e Frases complexas] e a Expressão [Vocabulário (nomeação de objetos e de imagens); Definição de Conceitos; Frases Absurdas – reconhecer e justificar; Constituintes Morfossintáticos e Pragmática (intenções comunicativas)]. Só foram incluídas no estudo as crianças cujos valores na avaliação se situassem dentro da média esperada para a respetiva idade (entre -1 e +1 DP).

De forma a elicitar o discurso espontâneo foi usada metodologia e materiais idênticos ao referido num estudo recente²⁹. As amostras de conversação foram recolhidas por três das autoras do presente trabalho, previamente treinadas na recolha dos dados, utilizando um conjunto de brinquedos apropriados à idade para elicitar diferentes formas gramaticais e tipos de frases, como objetos utilizados em casa e animais. Os observadores interagiam com as crianças num contexto lúdico, evitando o domínio das interações verbais, as respostas “sim/não” e as perguntas.

As amostras de discurso foram recolhidas em salas tranquilas, nas escolas que as crianças frequentavam, e foram registadas com um gravador Olympus WS-650S, escolhido pelas suas características de redução do ruído e boa captação de som a 50 cm de distância. Cada observadora ia à sala buscar a criança, interagindo com ela cerca de 5 minutos, durante o percurso para a sala onde era feita a avaliação. Este procedimento destinava-se a deixar a criança mais à vontade antes do início da recolha da amostra do seu discurso espontâneo. A interação com a criança demorava cerca de 20 a 30 minutos e tinha como objetivo obter um mínimo de 100 enunciados válidos. A recolha da amostra de discurso de cada criança foi feita numa única sessão.

As amostras foram transcritas e codificadas pelas observadoras utilizando o software ELAN (EUDICO Annotator linguísticas), uma ferramenta de anotação criada no Instituto Max Planck de Psicolinguística (<http://www.lat-mpi.eu/tools/elan>).

Para efeitos da análise, foram consideradas os primeiros 100 enunciados de cada criança, seguindo o critério de Brown⁷. Os critérios estabelecidos para contagem dos enunciados e das palavras foram os seguintes:

- Segmentação dos enunciados: considerou-se que o enunciado é delimitado pela sua curva entoacional^{32,33}; a utilização da expressão “e depois” (“e *depois*” bengala linguística utilizada com regularidade pelas crianças portuguesas) foi considerada como dividindo enunciados, sendo um marcador para o início do enunciado seguinte; foram contabilizados todos os enunciados, mesmo que tivessem erros morfossintáticos, pois estes são comuns nas faixas etárias estudadas; os enunciados repetidos exatamente da mesma forma foram contabilizados apenas uma vez⁷; não foram contabilizadas músicas, enumerações, contagens e palavras isoladas³²; foram ainda excluídos os enunciados resultantes de imitações produzidas imediatamente após a sua utilização pela observadora, os que não se percebessem devido à existência de palavras ininteligíveis, os enunciados parciais, resultantes por exemplo de mudança do foco da atenção da criança, falsos começos e reformulações (contou apenas a última formulação), marcadores de discurso (oh, ah...) não integrados no significado do enunciado e palavras isoladas³².
- Contagem de palavras: foi definida “palavra” como qualquer sequência semanticamente interpretável e delimitada por espaços em branco ou sinais de pontuação³⁴; a palavra foi contabilizada apenas uma vez, na forma em que foi produzida mais completa, com exceção dos casos em que a repetição teve por objetivo enfatizar uma ideia⁷; as contrações foram contabilizadas como uma palavra¹⁷; os clíticos,

as palavras compostas, as combinatórias cristalizadas (ex. David Beckham; Central Park) e as palavras onomatopaicas foram contabilizadas como uma palavra; palavras ininteligíveis não foram contabilizadas¹⁷; os auxiliares de discurso como exclamações não foram contabilizados⁷.

Os observadores tiveram treino prévio para transcrição, contagem dos enunciados válidos e das palavras. Cada um transcreveu os enunciados e esta transcrição foi verificada pelos outros dois observadores e também por dois investigadores com experiência. Este procedimento decorreu até ao final do trabalho, com os desacordos resolvidos através de consenso.

■ RESULTADOS

Foi calculada a EME-p de cada criança, utilizando-se em cada grupo etário de 3 000 a 3 200 enunciados para análise. Verificou-se que entre os

4 e os 5 anos e meio a EME-p varia em média entre as 4,5 e as 5 palavras, verificando-se um aumento progressivo nos três grupos de idade (Tabela 4.). Uma vez que a EME-p tinha uma distribuição normal nos três grupos foi usado o teste estatístico ANOVA oneway que mostra diferenças significantes entre eles ($F(2) = 7.72$, $p = .001$), mas o post hoc test Scheffé indica que as diferenças são apenas entre o primeiro grupo e o segundo grupo dos 4 anos ($p = .037$) e entre o primeiro grupo e o das crianças mais velhas ($p = .001$). Não há diferenças entre o segundo e o terceiro grupos ($p = .49$).

Comparativamente com os valores obtidos por Rice e colaboradores²⁹, verifica-se exatamente a mesma progressão da EME-p nas três faixas etárias, com maior diferença entre a primeira e a segunda faixa dos 4 anos (Figura 1). Em Português Europeu os enunciados são um pouco mais extensos em todas as faixas etárias. Não foi possível comparar com os valores do Português do Brasil³⁰, pois os grupos etários são diferentes.

Tabela 4 – Resultados num teste de linguagem formal (compreensão e expressão), valores médios de EME-p e respetiva distribuição percentilica

Grupo etário	Teste de Linguagem (TALC)		EME-p			
	Compreensão Max: 69	Expressão Max: 54	M ± DP	P10	P50	P90
4;00 – 4;05	61.63±2.46	44,70±2.97	4.49 ± .38 (3.80 – 5.18)	3.98	4.45	5.03
4;06 – 4;11	63.97±2.00	49.20±2.29	4.90 ± .76 (3.60 – 6.32)	3.99	4.93	5.95
5;00 – 5;05	65.03±1.84	49.00±2.46	5.09 ± .65 (4.05 – 6.18)	4.22	4.95	6.00

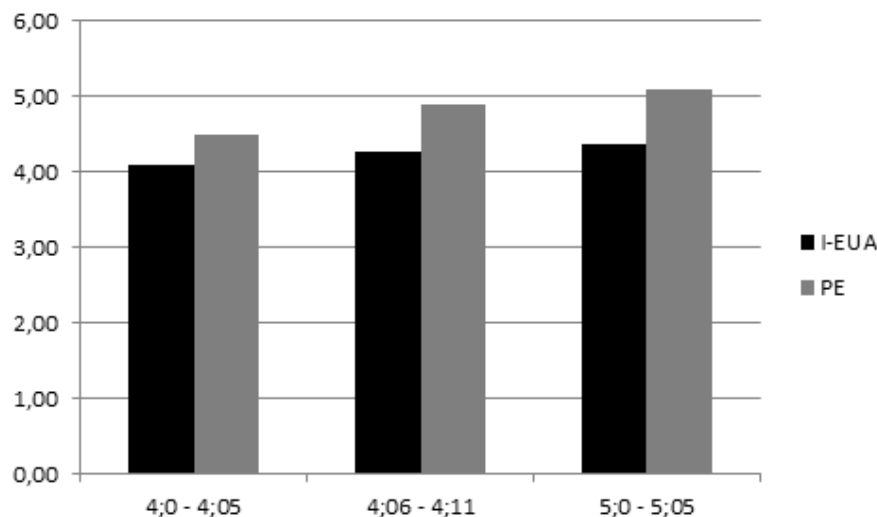


Figura 1 – Valores de EME-p em Inglês (EUA) e em Português Europeu, por grupo etário

Analisados os resultados obtidos por meninos e meninas através do test t de Student para amostras independentes, verificou-se que não havia diferenças entre eles (Média meninos: $4.88 \pm .69$; Média meninas: $4.78 \pm .64$; $t(90) = .755$, $p = .45$).

Na amostra geral, a escolaridade dos pais mostrou correlação significativa com o EME-p (Pearson $r = .289$, $p = .005$), o mesmo se verificou quando foi analisada apenas a escolaridade da mãe (Pearson $r = .292$, $p = .005$). No entanto, quando os dados foram analisados por grupo de idade apenas o grupo intermédio dos 4;06 aos 5;0 apresentava correlação forte, positiva e significativa (Pearson $r = .753$, $p = .000$). No grupo das crianças mais jovens a relação não era significativa (Pearson $r = .179$, $p = .34$), bem como no grupo das crianças mais velhas (Pearson $r = -.125$, $p = .50$).

Verificou-se ainda uma correlação positiva e significativa entre os resultados obtidos na EME-p e os valores no teste de desenvolvimento da linguagem (TALC), tanto para os valores totais deste teste (Pearson $r = .33$, $p = .001$), como para os domínios da compreensão (Pearson $r = .287$, $p = .006$) e da produção da linguagem (Pearson $r = .286$, $p = .006$).

■ DISCUSSÃO

Os dados apresentados constituem o primeiro estudo sobre EME-p para o Português Europeu, abrangendo a faixa etária dos 4;0 aos 5;05. A média obtida para os três grupos, divididos em 6 meses de intervalo, diferem das já obtidas para o Inglês Americano²⁹ sendo para o Português Europeu ligeiramente superiores, como seria de esperar, em resultado das diferenças morfossintáticas entre as duas línguas. Também diferem dos valores encontrados para o Português do Brasil³⁰, contudo esta comparação é difícil de efetuar, pois os intervalos de idade não são iguais. De qualquer forma os valores obtidos no Brasil são bastante inferiores aos obtidos no grupo de crianças com 4 anos, o que poderá ser devido também a diferenças de metodologia. No estudo brasileiro os critérios utilizados para segmentação dos enunciados e para a contagem das palavras não são descritos em pormenor. Assim, os valores resultantes podem não ser comparáveis com os valores obtidos no nosso estudo.

Tendo em conta as características específicas das línguas e/ou das suas variantes é possível compreender as diferenças verificadas. O fato de ter-se obtido médias de EME-p mais elevadas pode ser documentado com os exemplos abaixo, mostrando como a utilização do gerúndio no Inglês e no Português do Brasil, reduz o número de

palavras do enunciado ou que a menor utilização de determinantes em Inglês também faz reduzir esse número.

- Inglês: He told me (that) Ana is eating (6 / 7 words)
- P. Brasil: Ele me disse que a Ana está comendo (8 words)
- P. Europeu: Ele disse-me que a Ana está a comer (9 words)

Além do seu interesse específico para a análise da evolução de cada língua e de aplicações clínicas, os estudos disponíveis em vários países tendem a mostrar uma certa universalidade do número de palavras por enunciado quando uma metodologia semelhante é usada na análise. Por exemplo, na primeira metade do grupo de quatro anos de idade pode ser observada uma variação média de apenas 0,39 entre o Inglês Americano (4.10) e o Português Europeu (4.49).

Estudos destinados a avaliar a EME são relativamente escassos, devido à sua própria dificuldade, pois exigem grandes amostras e critérios de análise rigorosos. Qualquer diferença metodológica influenciará os resultados e os valores obtidos através de contextos e de métodos de eliciação distintos poderão não ser comparáveis³⁵. Esta diferença impede a realização de análises comparativas entre línguas e até dentro da mesma língua, quando se pretende estudar, por exemplo, um largo período de desenvolvimento da linguagem.

Relativamente à influência do género nos valores de EME-p, não foram encontradas diferenças estatisticamente significantes entre meninos e meninas. Este resultado é idêntico ao já constatado por outros autores que analisaram a EME em crianças chinesas dos 2;03 aos 5;08 anos²², em crianças francesas após os três anos de idade²⁵ e em crianças americanas mais velhas dos 6;03 aos 15;02³⁶. As diferenças de género parecem existir apenas até aos 3 e 4 anos de idade, com as meninas a apresentarem valores de EME-p mais elevados e a produzirem formas sintáticas mais complexas^{23,25}. Assim, apesar da discrepância existente, provavelmente também dependente das diferenças de metodologia, é possível defender que por volta dos 4 anos de idade o desenvolvimento sintático atingiu já alguma estabilidade, dissipando as diferenças entre géneros que existirão até cerca dessa idade.

O nível de escolaridade dos pais mostrou influenciar a EME-p na amostra total. Estes resultados são concordantes com os de outros autores²⁵⁻²⁷ que observaram uma forte influência desta variável, de tal forma que as crianças que provinham de meios socioculturais mais altos apresentavam produções

orais mais complexas e um desenvolvimento linguístico mais avançado do que crianças de níveis socioeconómicos mais baixos. No entanto, outros autores não têm constatado relação com o meio de origem das crianças, como por exemplo Rice e colaboradores⁵. Estas diferenças entre os estudos estarão dependentes de diversas variáveis, como as faixas etárias analisadas em cada estudo, o tipo de análise linguística efetuada e a forma de medir o nível sociocultural de origem (nível de escolaridade dos pais, renda familiar e/ou profissão do pai, ...).

Considerou-se na presente análise o número de anos de escolaridade do pai ou da mãe, aquele que fosse mais elevado. Considerou-se ainda apenas a escolaridade da mãe, tal como outros autores²⁹. Em ambas as situações os resultados foram idênticos, contudo esta relação entre escolaridade parental e desenvolvimento da linguagem medida pela EME-p não é linear, pois na análise por grupo de idade apenas o grupo intermédio, dos 4;06 aos 5;0, mostrava relação significativa. Aparentemente nas crianças mais pequenas não há influência da escolaridade dos pais na extensão do enunciado. Na fase seguinte essa influência aumenta em simultâneo com o aumento significativo da extensão do enunciado. De seguida desaparece, provavelmente devido ao desenvolvimento de estruturas linguísticas mais complexas, mantendo-se idêntica a EME-p, pois não há diferença significativa entre o grupo de 4;06 a 5;0 e o grupo de 5;01 a 5;05. Por outro lado, verificou-se que aos 5 anos surge uma tendência de correlação negativa entre a EME e a escolaridade dos pais. De modo semelhante Rice e colaboradores²⁹ constataram que apenas havia relação significativa da EME com a escolaridade parental no grupo de crianças de 5 e 6 anos, sendo essa relação negativa, “suggesting higher MLU levels for the children of less educated mothers”. Assim, não é de excluir a influência da educação parental na EME, contudo esta medida não será a mais indicada para esta observação, uma vez que um enunciado com mais palavras não é necessariamente mais complexo a nível morfológico ou sintático. O progressivo desenvolvimento das capacidades cognitivas e linguísticas permitirão à criança usar a linguagem de forma mais rápida e eficiente, empregando menos palavras para transmitir o mesmo conteúdo, e utilizando estruturas linguísticas mais complexas. O tipo de linguagem usado pelos pais, certamente dependente do seu nível de educação, é um modelo que pode

influenciar a criança nesta fase que precede a entrada na escola. Não é de admirar, assim, que exista uma tendência, como a verificada neste estudo, ou que exista mesmo uma correlação significativa²⁹ mostrando que o uso de enunciados mais curtos se relaciona com níveis mais elevados de escolaridade dos pais.

As correlações significantes entre a EME-p e um teste formal de desenvolvimento de linguagem foram elevadas, confirmando que se trata de uma medida válida de desenvolvimento, tal como já constatado¹⁸. Durante os primeiros anos de vida da criança o vocabulário aumenta progressivamente e o mesmo acontece na produção de frases que têm um cada vez maior número de palavras. Assim, é esperada uma correlação positiva e significativa entre a EME-p e os testes formais de avaliação do desenvolvimento da linguagem, até ao final da idade pré-escolar. Depois deste período esta correlação pode não ser adequada, pois a EME-p já não aumenta de forma evidente, sendo a criança capaz de produzir frases mais complexas, com utilização de menos palavras para transmitir o mesmo significado.

■ CONCLUSÃO

Do estudo realizado com 92 crianças portuguesas, salienta-se o seguinte: (1) aos 4 anos de idade as crianças produzem, em média, entre 4 a 5 palavras por enunciado e aos 5 anos de idade esse valor situa-se claramente nas 5 palavras; (2) não foram verificadas diferenças decorrentes do género, sendo idêntico o desempenho de meninos e de meninas; (3) não foi constada uma influência nítida da escolaridade dos pais, pelo que não é possível afirmar a influência desta variável; (4) verificou-se uma correlação positiva e significativa com os resultados de um teste formal de desenvolvimento da linguagem, tanto para a compreensão como para a expressão, o que acentua a validade da EME-p como um instrumento de medida do desenvolvimento da linguagem.

Apesar do número de crianças de cada faixa etária ser relativamente pequeno considera-se que os dados obtidos podem servir de referência para as crianças portuguesas, pois o número de enunciados recolhidos foi bastante elevado e os critérios de inclusão dos participantes, de elicitación de discurso e de análise da produção verbal foram rigorosos.

ABSTRACT

Purpose: to analyze the Mean Length of Utterance-words (MLU-w) in children aged 4;00-5;05 years. **Methods:** ninety two Portuguese children with normal development were observed: 49 girls and 43 boys, divided in age range groups of six months. A sample of 100 utterances produced in spontaneous discourse was collected from each child. The utterances were transcribed and analyzed. **Results:** MLU-w was shown to vary between 4,5 to 5 words, progressing with age. This progression had been previously observed in US English and in Brazilian Portuguese speaking children, although in European Portuguese the number of words is overall a little higher. Both boys and girls performed similarly. Years of parents formal education showed some influence, but not in all age groups. Results showed a positive and significant correlation with a formal test for language assessment, both in comprehension as in language production. **Conclusion:** the MLU-w is a good measure of language development up to 5 years. The values found can serve as a normative reference for Portuguese children, but also in comparative studies on the development of spontaneous language.

KEYWORDS: Language; Development; Child; Evaluation

■ REFERÊNCIAS

1. Gries ST, Stol S. Finding developmental groups in acquisition data: variability based neighbor clustering. *Journal of Quantitative Linguistics*. 2009;16:217-42.
2. Parker MD, Brorson, K. A comparative study between mean length of utterance in morphemes (MLUm) and mean length of utterance in words (MLUw). *First Language*. 2005;25:365-76.
3. Roy D. New horizons in the study of child language acquisition. Proceedings of INTERSPEECH 2009, 10th Annual Conference of the International Speech Communication Association, September 6-10; Brighton, United Kingdom; 2009.
4. Eisenbeiss S. Production methods in language acquisition research. In: Blom E, Unsworth S, editors. *Experimental methods in language acquisition research*. Amsterdam: John Benjamins Publishing Company; 2010. P. 11-34.
5. Rice ML, Redmond SM, Hoffman, L. Mean Length of Utterance in children with Specific Language Impairment and in younger control children shows current validity and stable and parallel growth trajectories. *Journal of Speech, Language, and Hearing Research*. 2006;49:793-808.
6. Retherford, KS. *Guide to analysis of language transcripts*. 3rd edition. Eau Claire, WI: Thinking Publications; 2000.
7. Brown R. *A first language: the early stages*. Cambridge, MA: Harvard University Press; 1973.
8. Miller JF. *Assessing language production in children*. Baltimore, MD: University Park Press; 1981.
9. Hickey T. Mean length of utterance and the acquisition of Irish. *Journal of Child Language*. 1991;18:553-69.
10. Nice MM. Length of sentences as a criterion of a child's progress in speech. *Journal of Educational Psychology*. 1925;16:370-9.
11. Eisenberg SL, Fersko TM, Undgreen C. The use of MLU for identifying impairment in preschool children: a review. *American Journal of Speech-Language Pathology*. 2001;10:323-42.
12. Arif H, Bol GW. Counting MLU in morphemes and MLU in words in a normally developing child and child with language disorder: a comparative study. *Dhaka University Journal of Linguistics*. 2008;1:167-82.
13. Oosthuizen H, Southwood, F. Methodological issues in the calculation of mean length of utterance. *South African Journal of Communication Disorders*. 2009;56:76-87.
14. Gutiérrez-Clellen VF, Restrepo MA, Peña LB, Anderson R. Language sample analysis in Spanish speaking children: methodological consideration. *Journal of Language, Speech, and Hearing Services in Schools*. 2000;31:88-98.
15. Bol GW. Optimal subjects in Dutch child language. In: Koster C, Wijnen F, editors. *Proceeding of the Groningen assembly on language acquisition*. Groningen: Center for Language and Cognition; 1996. P. 125-33.
16. Klee T. Developmental and diagnostic characteristics of quantitative measures of children's language production. *Topics in Language Disorders*. 1992;12:28-41.
17. Wiczorek R. Using MLU to study early language development in English. *Psychology of Language and Communication*. 2010;14:59-69.

18. Blake J, Quartaro G, Onorati S. Evaluating quantitative measures of grammatical complexity in spontaneous speech samples. *Journal of Child Language*. 1993;20:139-52.
19. Özçalışkan S, Goldin-Meadow S. Sex differences in language first appear in gesture. *Developmental Science*. 2010;13:752-60.
20. Ramer ALH. Syntactic styles in emerging language. *Journal of Child Language*. 1976;3:49-62.
21. Davies A. An introduction to applied Linguistics – from practice to theory. 2nd edition. Edinburgh: Edinburgh textbooks in applied Linguistics; 2007.
22. Klee T, Stokes SF, Wong AMY, Fletcher P, Gavin WJ. Utterance length and lexical diversity in Cantonese-speaking children with and without specific language impairment. *Journal of Speech, Language and Hearing Research*. 2004;47:1396-410.
23. Jackson SC, Roberts JE. Complex syntax production of African American preschoolers. *Journal of Speech, Language, and Hearing Research*. 2001; 4:1083-96.
24. Tse SK, Chan C, Kwong SM, Li H. Sex differences in syntactic development: evidence from Cantonese-speaking preschoolers in Hong-Kong. *International Journal of Behavioral Development*. 2002;26:509-17.
25. Le Normand M-T, Parisse C, Cohen H. Lexical diversity and productivity in French preschoolers. *Clinical Linguistics and Phonetics*. 2008;22:47-58.
26. Bornstein MH, Hahn CS, Haynes OM. Specific and general language performance across early childhood: stability and gender consideration. *First Language*. 2004;24:267-304.
27. Walker D, Greenwood C, Hart B, Carta J. Prediction of school outcomes based on early language production and socioeconomic factors. *Children and Poverty*. 1994;65:606-21.
28. Hoff E. The specificity of environmental influence: socioeconomic status affects early vocabulary development via maternal speech. *Child Development*. 2003;74:1368-78.
29. Rice ML, Smolik F, Perpich D, Thompson T, Rytting N, Blossom M. Mean Length of Utterance levels in 6-month intervals for children 3 to 9 years with and without language impairments. *Journal of Speech, Language, and Hearing Research*. 2010;53:333-49.
30. Araújo K, Befi-Lopes. Extensão média do enunciado de crianças entre 2 e 4 anos de idade: diferenças no uso de palavras e morfemas. *Revista da Sociedade Brasileira de Fonoaudiologia*. 2004;9:156-63.
31. Sua-Kay E, Tavares MD. Teste de avaliação da linguagem na criança – TALC. Lisboa: Oficina Didáctica; 2008.
32. Lund NJ, Duchan JF. Assessing children's language in naturalistic contexts. New Jersey: Prentice Hall; 1993.
33. Miller JF, Chapman RS. Systematic analysis of language transcripts [Computer software]. Madison: University of Wisconsin; 1991.
34. Villalva A. Morfologia do Português. Lisboa: Universidade Aberta; 2008.
35. Heilmann JJ. Myths and realities of language sample analysis. *Perspectives on Language Learning and Education*. 2010;17:4-8.
36. McEwen S. Learning to speak like girls and boys: a developmental study in gender and narrative style. *Language, Information and Computation*. 1996;11:449-58.

<http://dx.doi.org/10.1590/1982-021620151741315>

Recebido em: 27/01/2015

Aceito em: 23/04/2015

Endereço para correspondência:

Maria Emília Santos

Instituto de Ciências da Saúde,

Universidade Católica Portuguesa,

Palma de Cima, 1649-023 Lisboa, Portugal

E-mail: me.psantos@ics.lisboa.ucp.pt