



UNIVERSIDADE CATÓLICA PORTUGUESA | INSTITUTO DE CIÊNCIAS DA SAÚDE

EXTENSÃO MÉDIA DO ENUNCIADO: OBTENÇÃO DE VALORES PADRÃO PARA O PORTUGUÊS EUROPEU EM CRIANÇAS DE 4 ANOS (4;00 – 4;05)

Dissertação apresentada à Universidade Católica Portuguesa para obtenção do grau de mestre em Linguística Clínica

Por
Mariana Cacela

Lisboa, 2013



UNIVERSIDADE
CATÓLICA | INSTITUTO DE
PORTUGUESA | CIÊNCIAS DA SAÚDE

EXTENSÃO MÉDIA DO ENUNCIADO: OBTENÇÃO DE
VALORES PADRÃO PARA O PORTUGUÊS EUROPEU EM
CRIANÇAS DE 4 ANOS (4;00 – 4;05)

Dissertação apresentada à Universidade Católica Portuguesa para obtenção do
grau de mestre em
Linguística Clínica

Por
Mariana Cacela

Sob a orientação de Professora Doutora Ana Mineiro e Professora Doutora
Maria Emília Santos

Lisboa, 2013

Resumo

A Extensão Média do Enunciado (EME) é uma medida de análise de amostras de discurso espontâneo que desde há vários anos tem sido estudada e aplicada nos mais diversos estudos noutras línguas. Através dela tem sido possível auxiliar terapeutas e linguistas na avaliação de perturbações de linguagem, orientar planos de intervenção, estudar o desenvolvimento da linguagem de crianças para uma determinada língua e o desenvolvimento de crianças bilingues. Para o Português Europeu ainda não são conhecidos valores que permitam utilizar esta medida, pelo que se desenvolveu um projeto que envolve o estudo de crianças dos quatro aos cinco anos e cinco meses.

A presente dissertação tem como principal objectivo a obtenção do valor padrão da EME em palavras (EME-p) de crianças de 4;00 a 4;05 anos de idade. Pretende-se, ainda, perceber se estes valores são influenciados pelo género das crianças e pelo nível de escolaridade dos seus pais.

Participaram no estudo 30 crianças residentes e a frequentar jardins de infância da área da grande Lisboa, com desenvolvimento típico da linguagem, cujo discurso foi gravado (um mínimo de 100 enunciados por criança), transcrito e depois analisado. A EME-p de cada criança foi calculada, sendo depois obtida a média das 30 amostras recolhidas, sendo, desta forma, obtido o valor médio de EME-p para esta faixa etária.

Não foram observadas diferenças significativas ao nível do género nem influência da escolaridade dos pais.

Palavras-chave: Linguagem; Infância; Discurso; EME-p, Género; Escolaridade

Abstract

Mean Length of Utterance (MLU) is a spontaneous language sample analysis measure that has been used for many years and applied in a number of different studies in other languages. Through this measure, speech therapists and linguistics have been able to evaluate language disorders, plan their interventions, and study children’s language development in a specific language and in bilingual children. For the European Portuguese language, values haven’t been obtained to realise if this is a good measure to be used with Portuguese children. Therefore, a project was developed to study this measure in children between the ages of 4 and 5 and a half years old.

This dissertation’s main goal is the obtainment of values of MLU in words (MLUw) from 4;00 to 4;05-year-old children. It is also intended to understand if these values are influenced by the child’s gender or its parents’ educational level.

30 children, living and attending kindergartens in the area of great Lisbon participated in this study, all presenting typical language development. Their speech samples were recorded (a minimum of 100 utterances per child), transcribed and then analysed in terms of MLUw. Each child’s individual MLUw was calculated, followed by the obtainment of an average of the 30 samples collected and thus found the mean value of MLUw for this age group.

No significant differences were observed in terms of the child’s gender and the parents’ education level doesn’t seem to influence the child’s MLUw.

Keywords: Language; Childhood, Speech, MLUw, Gender, Parents Education.

Agradecimentos

Agora que concluo este trabalho, e chega ao fim mais uma etapa da minha vida académica, não poderia deixar de agradecer a todos aqueles que contribuíram para isto fosse possível.

Desta forma, gostaria de agradecer à Professora Doutora Ana Mineiro a orientação, disponibilidade e apoio prestados, bem como o seu interesse em melhorar o meu trabalho e contribuir de uma forma construtiva para o mesmo.

À Professora Doutora Maria Emília Santos agradeço por me ter suscitado o gosto pela investigação científica, ainda na época da minha licenciatura, levando-me a querer continuar os meus estudos um bocadinho mais além. Agradeço-lhe também a disponibilidade em que me ajudou a esclarecer as incertezas da análise dos resultados.

Agradeço a todos os jardins de infância que me acolheram e às minhas colegas de investigação, permitindo a nossa presença nas escolas e por terem possibilitado a recolha de dados no terreno.

Agradeço às minhas duas colegas de investigação, Sara e Sofia, que permitiram que este percurso se mantivesse agradável e interessante mesmo quando as coisas não corriam como previsto.

A todos os meus amigos, agradeço por sempre acreditarem nas minhas capacidades, pela confiança que depositam em mim e por me apoiarem, pelos comentários de incentivo e pela forma acolhedora com que me ouviram falar deste trabalho, vezes sem conta.

Por último, não poderia deixar de agradecer à minha família, em especial aos meus pais, pelo carinho, pelas palavras de apoio e por estarem sempre ao meu lado quando mais preciso deles.

Índice

Introdução	1
1. Revisão da Literatura	3
<i>Análise de amostras de discurso espontâneo</i>	3
<i>Extensão Média do Enunciado</i>	12
<i>Possíveis utilizações da EME</i>	20
2. Problemas em estudo	25
3. Metodologia	26
<i>Participantes</i>	26
<i>Procedimentos</i>	28
<i>Análise do corpus linguístico</i>	30
<i>Tratamento de dados</i>	33
4. Resultados	34
5. Discussão	37
6. Conclusões	41
Referências Bibliográficas	43
Apêndices.....	55

Índice de Quadros

Quadro 1. Níveis e valores de EME-m segundo Brown (1973)	13
Quadro 2. Valores de EME-p e EME-m obtidos por Araújo e Befi-Lopes (2003)	19
Quadro 3. Distribuição de rapazes e raparigas por idade	27
Quadro 4. Escolaridade dos pais por género da criança	28
Quadro 5. Valor padrão da EME-p em crianças dos 4;00 aos 4;05 anos	34
Quadro 6. Teste de normalidade	35
Quadro 7. Comparação dos resultados da EME-p entre rapazes e raparigas	35

Índice de Figuras

Figura 1. Brinquedos utilizados para a recolha de amostras de discurso espontâneo....	29
Figura 2. Média de EME-p em cada idade.....	35
Figura 3. EME-p dos rapazes e raparigas.....	36

Índice de Apêndices

Apêndice I. Carta de pedido de autorização aos diretores dos agrupamentos	56
Apêndice II. Carta aos encarregados de educação e consentimento informado	58
Apêndice III. Folha de registo da criança	60
Apêndice IV. Modelo de relatório de avaliação para inclusão no estudo	62
Apêndice V. Modelo de relatório de avaliação para exclusão do estudo	64
Apêndice VI. Valores de EME-p obtidos por cada criança	66

Introdução

A avaliação da linguagem é umas das principais funções dos Terapeutas da Fala que trabalham na área da pediatria e do desenvolvimento infantil, sendo também de extremo interesse para investigadores na área da linguística. Os testes e baterias de avaliação existentes em Portugal são ainda escassos e o processo de avaliação torna-se, por vezes, incompleto. Noutros países como os Estados Unidos da América (EUA) e Inglaterra, onde o papel dos profissionais que trabalham nas áreas de avaliação e intervenção da linguagem é mais alargado, considera-se que uma avaliação de amostras de discurso espontâneo da criança é fundamental para se obter uma noção mais geral da sua *performance*, que será mais próxima da do dia-a-dia, com os seus vários parceiros de comunicação. Existem diversas medidas estudadas e divulgadas ao longo dos anos, tanto na área da morfossintaxe, como da semântica que permitem analisar o discurso produzido pela criança.

A Extensão Média do Enunciado (EME) é uma dessas medidas que tem vindo a ser estudada por diversos investigadores na área da linguística e que é considerada uma fonte de conhecimento fundamental do desenvolvimento normal da linguagem. Em diversas línguas têm sido realizados estudos de forma a serem obtidos valores padrão de EME para as várias faixas etárias. Em Portugal, no entanto, ainda não foi realizado este trabalho. Apesar da primeira referência ter sido feita por Nice (1925), foi o professor e investigador Roger Brown que expandiu o tema, tendo estudado e obtido valores para a população americana. Através da sua obra de 1973, Brown deu a conhecer o seu trabalho e levou a que o interesse por este tema se tivesse expandido. Além dos estudos entre línguas, a EME também tem permitido analisar o comportamento linguístico de crianças bilingues e com perturbações como Síndrome de Asperger, Autismo, Perturbações Específicas do Desenvolvimento da Linguagem (PEDL) entre outros.

Com este estudo pretende-se obter um valor padrão de EME-palavras (EME-p) para crianças dos 4;00 e os 4;05 (48 a 53 meses) e compreender se este valor é de alguma forma influenciado pelo género das crianças e a escolaridade dos seus pais. É importante referir que o presente trabalho se enquadra num projeto de maior dimensão com outras duas investigadoras. Neste trabalho conjunto, além da faixa etária estudada neste estudo, são ainda estudadas outras duas faixas etárias: a dos 4;06 a 4;11 e a dos 5;00 aos 5;05. Através da obtenção dos valores para estas três faixas etárias, será efectuada uma análise global e iniciar-se-á a organização de um *corpus* de linguagem infantil para o Português Europeu.

Inicialmente, as investigadoras tinham considerado a possibilidade de estudar a EME em morfemas (EME-m), seguindo assim a obra de Brown, tendo como referência os valores obtidos para a língua Inglesa. No entanto, como estudos demonstram que existe uma correlação alta entre a EME-m e a EME-p, pela maior facilidade de contabilização e obtenção de valores, pensando já na futura possível utilização por investigadores e terapeutas da fala e uma vez que vários autores sentiram maiores dificuldades em utilizar a EME-m em língua flexionadas, como o Português, mudou-se o objetivo do trabalho.

Apesar da ideia global do trabalho e da metodologia se basear no trabalho de Rice, Smolik, Perpich, Thompson, Rytting e Blossom (2010), foi necessário estabelecer regras mais concretas ao nível do processo de recolha e gravação das amostras de discurso e da definição e contabilização de palavras e enunciados pois a análise realizada no estudo referido, teve por base amostras de linguagem disponíveis numa base de dados já existente. Os brinquedos utilizados bem como o modo de interação com a criança foram, no entanto, semelhantes.

Esta dissertação encontra-se dividida em cinco capítulos. O primeiro capítulo é referente à revisão da literatura onde o tema será contextualizado consoante a informação recolhida em artigos e livros publicados acerca da EME. Neste capítulo são abordados temas como o surgimento do interesse pela análise de amostras de discurso espontâneo e medidas propostas para tal; o estudo e desenvolvimento da EME, tanto para palavras como para morfemas; as diversas utilizações destas medidas como a avaliação linguística, o estabelecimento de objetivos terapêuticos, estudo do bilinguismo na criança, avaliação e comparação de crianças com e sem perturbação linguística, entre outros. No capítulo 2, são apresentados os problemas em estudo, sendo apresentadas as questões que o orientaram e os objetivos propostos para este trabalho. No capítulo 3, é apresentada a metodologia seguida no estudo, nomeadamente o processo de seleção dos participantes, descrição dos procedimentos para recolha de dados, a forma de aplicação dos instrumentos, e o tratamento dos dados.

O capítulo 4, consiste na apresentação dos resultados obtidos cuja discussão é apresentada no capítulo 5. O sexto e último capítulo, é composto pelas conclusões retiradas com o estudo, as limitações verificadas, a aplicação dos valores e resultados obtidos e sugestões para futuros estudos.

1. Revisão da Literatura

A linguagem é uma das características que mais define e distingue a espécie humana, sendo única na sua estrutura e capacidade referencial, crítica para a criação e transmissão de conhecimento cultural (Trask, 2007). Permite a comunicação e interação com as outras pessoas (Chomsky, 2000; Trask, 2007) e é utilizada para expressar necessidades, sentimentos e pensamentos (Trask, 2007). Tendo em consideração todas estas funções, o conhecimento atual acerca da forma como a criança adquire e desenvolve linguagem, baseia-se em dados, surpreendentemente, incompletos e, muitas vezes, recolhidos através de observações tendenciosas (Roy, 2009). Como consequência, alguns autores e investigadores como Roy (2009), consideram que algumas das teorias de aquisição da linguagem se encontram ainda pouco testadas.

Relativamente às alterações de linguagem que podem ser observadas na criança, têm apenas duas origens, de desenvolvimento ou adquiridas, apesar de poderem ser de diversos tipos (Tallal, 1988). Estas alterações são uma das principais razões pelos quais pais/cuidadores de crianças, em idades pré-escolares, procuram um Terapeuta da Fala (Shipley & McAfee, 2009) que é o profissional “responsável pela prevenção, avaliação, tratamento e estudo científico da comunicação humana e das perturbações com ela relacionada” (D/Lei 261/93, de 24 de Julho), o que envolve perturbações e/ou alterações da linguagem.

Segundo Condouris, Meyer e Tager-Flusberg (2003), a avaliação das competências linguísticas da criança é uma área de grande interesse tanto para clínicos, terapeutas e psicólogos, como para investigadores da área da linguística. Os primeiros dependem destas avaliações para diagnosticar as crianças, analisar as suas competências e elaborar e estruturar, adequadamente, planos de intervenção. Por seu lado, na área da investigação, as avaliações permitem documentar o nível linguístico dos participantes em estudo, associar grupos de participantes, investigar características linguísticas em diferentes populações e faixas etárias (Condouris, *et al.* 2003).

Análise de amostras de discurso espontâneo

Para o processo de avaliação da linguagem de crianças em idades pré-escolares são, geralmente, utilizadas duas vias metodológicas, ou seja, dois tipos de medidas: i) testes

formais estandardizados e; ii) medidas de análise de discurso espontâneo que pode ser recolhido em diferentes contextos e de diversas maneiras (Condouris, *et al.*, 2003; McDaniel, McKee & Cairns, 1996). Ambos os tipos de medidas podem ser utilizadas para avaliar uma variedade de competências linguísticas: fonologia, conhecimento lexical, semântica, morfologia, sintaxe e pragmática em crianças de várias idades (McDaniel, *et al.*, 1996). No entanto, as medidas de avaliação formais permitem avaliar tanto competências receptivas como expressivas, enquanto as medidas de discurso espontâneo tendem a ser utilizadas apenas para testar o nível da linguagem expressiva (Roy, 2009).

De acordo com McDaniel, *et al.* (1996), a utilização das duas formas de avaliação permite um conhecimento mais completo das verdadeiras competências linguísticas da criança e, apesar desta ser uma visão partilhada pela grande maioria dos profissionais (Shipley & McAfee, 2009), geralmente, as medidas de discurso espontâneo não são utilizadas (Roy, 2009). São vários os autores que as defendem pois muita da investigação ao nível da linguística, envolve valores e dados qualitativos longitudinais, ou seja, em que o processo de aquisição e desenvolvimento da linguagem é monitorizada tendo em consideração como certos padrões se modificam com o tempo (Gries & Stoll, 2009). Além disso, ao longo dos anos, muitas críticas têm sido levantadas ao processo formal de avaliação, sendo muitos autores a favor de uma observação naturalística (Parker & Brorson, 2005; Roy, 2009). Este é uma observação que se tenta aproximar mais do comportamento linguístico da criança, no seu dia-a-dia, e que permite observar a linguagem em uso, em momentos de conversação (Eisenbeiss, 2010).

Em Portugal, os testes e baterias de avaliação existentes, para avaliar crianças em idades pré-escolares, são ainda escassos. Noutros países como nos Estados Unidos da América (EUA) e em Inglaterra, apesar dos profissionais terem à sua disposição um número mais elevado, estes testes são duramente criticados pela incapacidade que têm em avaliar a linguagem da criança como um todo (Roy, 2009). Autores como Roy (2009), Paul (2007) e Bachman (1990) consideram que estes testes não têm em consideração a função de comunicação da linguagem, exigindo respostas artificiais a questões muito fechadas e limitadoras. McDaniel, *et al.* (1996) referem que os testes formais são utilizados em situações muito controladas e que a recolha de amostras de discurso permite uma análise do comportamento linguístico da criança em condições mais naturais. Condouris, *et al.* (2003), afirmam ainda que os valores obtidos estão muito dependentes da capacidade de atenção e de motivação da criança para interagir com o avaliador, o que se torna ainda mais difícil de controlar com a aplicação de testes formais, nos quais não se pode alterar a forma de

aplicação nem os materiais (Paul, 2007). Estes autores abordam desta forma a pertinência dos estudos naturalísticos. De acordo com Eisenbeiss (2010), os estudos naturalísticos têm uma validade ecológica elevada pois a gravação do processo de conversação aproxima-se muito às situações da vida real em investigação. Além disso, as amostras naturalísticas podem, em princípio, ser obtidas de qualquer sujeito, independentemente da idade, competências linguísticas e cognitivas (Eisenbeiss, 2010; Lund & Duchan, 1993).

Por este motivo, a análise de amostras do discurso infantil é um método de avaliação que permite, ao avaliador, verificar todos os aspetos da estrutura linguística da criança (Klee, 1992), num processo de compreensão e produção de linguagem (Owens, 1999; Parker & Brorson, 2005), possibilitando assim, observar a forma como esta utiliza a linguagem para comunicar (Hewitt, Hammer, Yont & Tomblin, 2005; Parker & Brorson, 2005).

O contexto em que o discurso é obtido também parece facilitar o processo de avaliação visto que o contexto de brincadeira livre tende a elicitare um tipo de produção verbal mais próximo da produção do dia-a-dia (Parker & Brorson, 2005), enquanto uma situação estruturada e direcionada pelo avaliador, tende a elicitare outro tipo de produção verbal por parte da criança (Borstein, Tamis-Lemonda & Haynes, 1999). Aram, Morris e Hall (1993), consideram que medidas obtidas através de amostras linguísticas possam ter uma sensibilidade e especificidade superior na identificação de crianças por perturbações linguísticas. Stockman (1996) apresentou no seu trabalho evidências que a análise de amostras de linguagem está menos vulneráveis a variações dialetais e culturais do que as formas tradicionais de avaliação.

Ao nível da investigação linguística, a criação e disponibilização de amostras de discurso tem sido fundamental para os estudos realizados nesta área. Hoje em dia, o estudo da linguagem através do recurso a estas amostras é um processo que está muito facilitado pela existência de *corpus* de linguagem infantil computadorizados e processados por *taggers* automáticos, especialmente em países como os EUA. No entanto, o estudo da linguagem espontânea da criança começou muito antes do surgimento de computadores e dos *corpora* linguísticos modernos. As primeiras investigações da linguagem natural da criança são de há mais de cem anos atrás e foram escritos em diários, por pais de crianças onde eram documentados os seus progressos linguísticos, tendo por base as suas observações diárias. Um dos relatos mais antigos, e mais citado, é o estudo de Clara e William Stern (1907, traduzido por Lamiell, 1999), dois alemães, que documentaram com muita precisão e detalhe a fala dos seus dois filhos Hilde e Gunter. Outro estudo descrito num diário, que teve um impacto importante na área da investigação, é o de Leopold (1939-49, citado por Diessel, 2009) que

descreveu o desenvolvimento linguístico bilingue da sua filha Hildegard, desde o nascimento até aos dois anos.

Nos anos 40 e 50, psicólogos comportamentais começaram a recolher dados de linguagem infantil de forma mais sistemática (Braine, 1963). Em contraste com os estudos dos diários, estes psicólogos, em vez de estudarem o desenvolvimento de uma criança em detalhe, recolhiam dados de um número maior de crianças, usando pequenas amostras de linguagem de uma hora de discurso espontâneo. Alguns dos estudos desta época incluem transcrições com mais de 100 crianças, cuidadosamente selecionadas tendo em conta as suas idades, género e níveis socioeconómicos (Diessel, 2009).

Os primeiros estudos longitudinais começaram a surgir nos anos 60. Tal como nos estudos baseados nos diários, estes seguiam o desenvolvimento de uma criança em particular (ou de um número reduzido de crianças) durante um extenso período de tempo. No entanto, em contraste com os primeiros, estes estudos não usam as observações parentais como fonte principal; começaram a ser utilizadas novas tecnologias para recolher dados naturalísticos (Diessel, 2009). Pela primeira vez na história, amostras de linguagem infantil foram gravadas e depois cuidadosamente transcritas e anotadas por um investigador. A utilização de cassetes áudio permitiu aos investigadores a documentação sistemática e análise de fenómenos linguísticos mais complexos que surgem após o estágio de duas palavras. Enquanto os dados anotados nos diários parentais é minimamente fiável para investigar os enunciados de uma ou duas palavras, é impossível manter um diário sistemático das produções de crianças numa fase de desenvolvimento mais avançado (Redington & Chater, 1998). Os trabalhos que surgiram nesta altura incluem as investigações de Martin Braine (1963), Lois Bloom (1970) e Roger Brown (1973) que se tornaram clássicos ao nível da investigação da linguagem infantil. Durante os anos 70 e 80, o número de estudos longitudinais aumentou gradualmente. No entanto, os dados recolhidos estavam apenas à disposição de um pequeno número de investigadores (geralmente os que recolhiam os dados e os seus colegas) (Diessel, 2009). Os avanços na tecnologia informática tornaram possível partilhar amostras de discurso de crianças mais facilmente. No início dos anos 80, Brian MacWhinney e Catherine Snow estabeleceram uma grande base de dados a que chamaram *Child Language Exchange System* (CHILDES) (Diessel, 2009). Esta base de dados inclui mais de cem projetos de investigação em diversas línguas que foram compiladas por MacWhinney e Snow (1985). Outra base de dados muito conhecida é o *Systematic Analysis of Language Transcripts* (SALT; Miller & Chapman, 1991). Ambos são muito utilizados em estudos pois são fontes de informação de desenvolvimento de acordo com a idade das crianças. O primeiro, fornece um arquivo valioso

de transcrições de linguagem recolhidas de diferentes estudos, realizados ao longo do tempo. A principal limitação do CHILDES é o facto das transcrições, como já referido, provirem de diferente estudos, com grandes variações de participantes, métodos e níveis de documentação, tudo o que, como seria de esperar, complica o seu uso com uma fonte de referência. A segunda base de dados foi desenvolvida como uma base de dados de referência. A colecta original foi recolhida de 1984 a 1987, e é constituída por amostras de linguagem espontânea recolhida em escolas públicas no estado americano do Wisconsin em duas zonas: na cidade de Madison e numa área rural próxima. O *corpus* total é o conjunto de amostras de discurso de 266 crianças entre os três e os 13 anos de idade. Apesar de ter muitas vantagens relativamente ao CHILDES, pois foi recolhida para um único propósito por investigadores treinados, e ser muito utilizado para estudos clínicos e de comparação normativa, apresenta também algumas limitações, como o facto das crianças serem todas da mesma região, com níveis socioeconómicos muito próximos.

Com todas as recolhas a serem efetuadas, algumas medidas que as permitissem analisar começaram a ser estudadas e depois apresentadas por diferentes autores. Em 1925, Nice apresentou o *Mean Length of Response* (MLR), utilizado para demonstrar padrões previsíveis no desenvolvimento da linguagem das crianças. Esta medida teria em consideração o número total de palavras produzido pela criança e depois relacionado com o número de enunciados. Mais tarde, esta medida passou a ser conhecida como *Mean Length of Utterance in words* (MLUw), que em português foi traduzido como Extensão Média do Enunciado em palavras (EME-p).

A partir desta primeira fórmula, Brown (1973) sugeriu que o cálculo da extensão dos enunciados tivesse por base o total dos morfemas apresentados, em vez do número total de palavras, apresentado assim a Extensão Média do Enunciados em morfemas (EME-m). Ainda ao nível da análise morfossintática, Scarborough (1990), propôs o IPSyn (*Index of Productive Syntax*) pois considerava que a complexidade sintáctica não pode ser reflectida apenas pela extensão do enunciado. Esta medida foi desenvolvida de forma a ser utilizada como uma ferramenta na investigação e documentação do desenvolvimento sintático de crianças em idades pré-escolares. No entanto, a amostra através da qual foram obtidos os valores-padrão é muito reduzida e nunca foi estendida a uma população maior. Apesar disso, é uma medida bastante utilizada pelos investigadores americanos (Nieminen, 2009).

Ao nível da análise semântica, mais especificamente para determinar o nível de vocabulário da criança, Templin (1957) apresentou o Type-Token Ratio (TTR). O *ratio* obtido reflete a diversidade de palavras utilizadas durante a amostra de linguagem e um valor

de TTR inferior ao esperado é um indicador de um atraso da linguagem expressiva ou uma perturbação linguística. Outra medida de complexidade linguística na área da semântica, que se tornou bastante popular na avaliação de crianças até aos seis anos nos EUA, é o NDW (*Number of Different Words*) proposta por Miller (1991). Mais recentemente, o D (Malvern & Richards, 2000), medida de diversidade lexical, tornou-se muito famoso por representar uma solução para alguns dos problemas do TTR relacionados com a diminuição do *tokens* quando o tamanho da amostra era grande e dificuldades de comparação entre estudos (McCarthy & Jarvis, 2011). Ao nível fonológico, foi proposta a Extensão Média do Enunciado em sílabas (EME-s) que se dedicava à contabilização de todas as sílabas produzidas. No entanto, esta medida acabou por ser posta de parte pelos investigadores pela dificuldade de contabilização, devido ao uso de diminutivos comuns na infância, assim como pela tendência de repetição de sílabas, o que pode ser exacerbado durante o período de gaguez fisiológica (Hickey, 1991).

Apesar de todas as vantagens da utilização de medidas de amostras de discurso espontâneo sobre os testes formais já apresentadas anteriormente, é facilmente observável nos diferentes estudos que aplicam uma destas medidas que uma das principais limitações apresentada pelos autores, ou criticada por outros, é a forma como decorreu o processo de recolha e o número de enunciados recolhidos e analisados (Rice, Redmond & Hoffman, 2006; Scarborough, 1990). Devido a estas dificuldades, alguns autores têm-se dedicado à optimização deste processo, propondo diversas estratégias. De acordo com Retherford (2000), o primeiro passo para a análise da produção linguística da criança é a obtenção fidedigna destas produções, sendo que, quanto melhor for o processo de recolha, mais essa amostra representará as verdadeiras competências linguísticas da criança (Retherford, 2000). Para este autor, tal é fundamental pois a forma como a interação comunicativa é, por vezes, conduzida leva a que a amostra resultante não seja representativa da produção normal da criança. O termo “representativa” tem sido muito utilizado na literatura. Mclean e Snyder-Mclean (1978) sugerem que uma amostra representativa tem de refletir a *performance* ótima da criança, enquanto Miller (1981), suporta a noção de que uma amostra representativa é a que é a fidedigna e válida. Segundo Retherford (2000) o termo “representativo” é usado para descrever a produção linguística usual da criança, incluindo linguagem que possa ser um pouco inferior ou um pouco superior à sua *performance* comum.

Para garantir que esta representatividade se encontra garantida, é necessário compreender e controlar alguns factores (Miller, 1981). Examinando cada factor e tomando medidas para otimizar todas as condições, conseguir-se-á obter uma melhor qualidade das amostras

(Retherford, 2000; Miller, 1981). Estes fatores são a natureza da interação, a localização, os materiais, a dimensão da amostra e o método de gravação:

- **Natureza da Interação:**

Refere-se à forma como a pessoa, com quem a criança interage, se apresenta durante esta interação, ou seja, se participa da conversação durante a brincadeira interativa ou se se limita a efetuar questões como forma de obter discurso. Miller (1981), defende que se deve tentar obter um número variado de amostras de linguagem com a criança interagindo com diferentes pessoas como o Terapeuta da Fala, um dos pais, um irmão ou um par. A interação da criança com cada um destes parceiros de comunicação varia, e tem sido observado que as crianças apresentam diferentes produções frásicas de acordo com quem interagem (Retherford, 2000). As diferenças nas amostras adicionam/somam à imagem das capacidades comunicativas globais da criança. Esta autora refere ainda que, para obterem uma amostra representativa, os clínicos devem fazer o menor número de questões, pois a criança irá produzir enunciados mais longos e complexos quando conversando espontaneamente do que quando respondendo a questões. No entanto, Stalnaker e Craighead (1982) obtiveram resultados inconclusivos relativamente ao estudo realizado em que era pedido às crianças para recontarem uma história. Não observaram grandes diferenças nos enunciados produzidos espontaneamente em comparação com os obtidos quando lhes eram feitas perguntas sobre a história (Stalnaker & Craighead, 1982). Apesar disso, é aparente que a conversação em que um parceiro faz questões e o outro responde não é uma interação natural (Lund & Duchan, 1993). Como parceiros de comunicação, os clínicos devem tentar reduzir o número de questões ao máximo e permitir que a criança tome a liderança na interação. No entanto, a ausência total de questões por parte do clínico ou investigador é praticamente impossível e a dificuldade em gerir uma conversação sem efetuar perguntas pode levar a dificuldades de interação o que levará a uma amostra pouco representativa das capacidades das crianças (Wieczorek, 2010);

- **Localização:**

Para Miller (1981), as gravações não devem ser efetuadas apenas num local, devendo-se variar os contextos como a escola, casa, clínica. Já no estudo de Brown

(1973), as gravações decorreram em mais de um contexto pois diferenças significativas podem surgir na linguagem produzida pela criança de acordo com a sua localização, especialmente pelo à vontade que apresenta;

- **Materiais:**

A terceira variável apresentada por Miller (1981) são os materiais presentes e que são utilizados para motivar a criança para a interação. Diferentes materiais podem levar a produções variadas, mas não é possível prever para cada criança a influência que estes terão na sua produção. Desta forma, é fundamental providenciar materiais adequados ao desenvolvimento da criança e que a encorajem a interagir o mais possível através destes (Retherford, 2000);

- **Dimensão da amostra:**

Brown (1973) defendeu que uma amostra constituída por 100 enunciados era suficiente para que esta fosse representativa. Outros autores têm sugerido um número superior de enunciados, enquanto outros afirmam que 50 são suficientes. A alternativa é obter enunciados durante um determinado período de tempo – 30 minutos, por exemplo – independentemente de quantas produções ocorram nesse período (Retherford, 2000). Geralmente, estes 30 minutos são suficientes para obter entre 100 a 200 enunciados para crianças com idades acima dos dois anos de idade (Miller, 1981). Crianças mais novas poderão precisar de mais tempo para produzirem 100 enunciados (Wieczorek, 2010);

- **Método de gravação:**

O método mais eficaz é a gravação vídeo pois além de permitir analisar o discurso, permite obter informação ao nível da interação não-verbal, o que poderá ser fundamental quando se recolhe discurso de crianças mais novas, nas quais o discurso ainda não é totalmente inteligível (Miller, 1981). Desta forma, é possível obter pistas relativas ao que a criança está a ver ou a mexer que poderão ajudar na compreensão do que está a ser dito (Miller, 1981). Outro método é a gravação áudio, que permite que o clínico possa interagir com a criança de forma livre, mas que leva a que se tenha de fazer pequenas notas acerca das atividades da criança durante os vários momentos da gravação para providenciar o contexto não-verbal para a transcrição (Retherford, 2000).

O processo de recolha está também muito dependente da forma como se consegue interagir com a criança. Enquanto algumas crianças estão mais à vontade, e assim que veem os brinquedos começam a interagir, outras, mais tímidas, e/ou mais novas, necessitam de algum incentivo, o que nem sempre é fácil de alcançar (Wieczorek, 2010). Por este motivo, alguns investigadores têm sugerido algumas linhas orientadoras acerca da forma como se deve interagir com a criança de forma a obter uma performance aproximada e otimizada. Miller (1981), sugere que uma boa forma de começar a interação e, em vez de se começar a fazer perguntas sobre os brinquedos presentes, é útil começar a brincar e a falar ao mesmo tempo, mesmo que a criança se limite a observar no início (e.g: enquanto se mexe com uma colher dentro da panela ir dizendo “vou fazer uma sopa muito saborosa com esta panela”). De seguida, é fundamental conseguir trazer a criança para a brincadeira, como “eu posso ser a senhora que vai almoçar ao restaurante e tu, queres ser o cozinheiro?”, desta forma encorajando-a a fazer sugestões acerca do que se vai passar (Miller, 1981). Retherford (2000), defende que é fundamental seguir o tópico da criança: se ela se mantém no *role-playing*, o investigador deve também manter-se. Além disso, Wieczorek (2010), é da opinião de que se deve sempre responder às suas perguntas, reconhecer os seus comentários, solicitar mais informações sobre tópicos escolhidos e que se dê opções à criança (e.g: “vamos fazer um piquenique ou preferes estar num restaurante?”). O investigador deve aprender a aceitar pausas, não devendo ser muito rápido a tentar preencher momentos de silêncio pois poderá não estar a dar tempo à criança para responder ou fazer comentários espontâneos (Lund & Duchan, 1993). É fundamental ter uma variedade de materiais disponíveis, mas não se deve mudar de brinquedos de forma abrupta, sendo até importante dar a oportunidade à criança de escolher se quer continuar a brincar com aqueles brinquedos ou se quer mudar (Miller, 1981; Retherford, 2000).

Concluindo, a análise de amostras de discurso espontâneo parece ser um instrumento fidedigno para auxiliar a avaliação do desenvolvimento da linguagem das crianças em idades pré-escolares, tanto para utilizações clínicas como crianças que apresentam perturbações linguísticas como para fins científicos para a obtenção de valores e conhecimento acerca das características e competências nas diferentes faixas etárias. No entanto, o processo através do qual estas medidas são obtidas tem de ser criterioso e específico e, através de um número mínimo de 100 enunciados, com materiais adequados e com um forma de interação que promova a produção linguística da criança. Existem diversas medidas que podem ser

aplicadas, dependendo da área da linguagem que se pretende estudar, sendo que neste estudo, e visto em Portugal não existirem valores de referência nesta área, foi escolhida a EME-p.

Extensão Média do Enunciado

A Extensão Média do Enunciado é uma das medidas de linguagem, já referida anteriormente, que é obtida através da recolha de amostras de discurso espontâneo (Hickey, 1991) e que tem como principal objetivo a obtenção de dados acerca dos aspectos morfológicos e sintáticos de crianças com desenvolvimento típico de linguagem ou com perturbação linguística (Brown, 1973; Hickey, 1991; Parker & Brorson, 2005).

Como referido anteriormente, o termo Extensão Média do Enunciado surgiu como uma evolução da medida, para análise de discurso, proposta por Nice (1925). Segundo esta autora, o comprimento das frases deveria ser um dos critérios mais importantes para julgar o progresso linguístico das crianças, sendo assim um marcador da maturação linguística. A medida original tinha por nome *Mean Length of Response* (MLR) e era calculada através da seguinte fórmula:

$$\text{MRL} = \frac{\text{n}^\circ \text{ total de palavras}}{\text{n}^\circ \text{ de enunciados}}$$

A forma como a extensão de enunciados é calculada tem vindo a evoluir desde a já referida publicação de Nice. Uma das alterações passou pela mudança da nomenclatura de MRL para *Mean Length of Utterance in words* (MLUw), que em português foi traduzido como Extensão Média do Enunciado em palavras (EME-p), tendo-se mantido a fórmula de cálculo, apesar da alteração do nome (Parker & Brorson, 2005):

$$\text{EME-p} = \frac{\text{n}^\circ \text{ total de palavras}}{\text{n}^\circ \text{ de enunciados}}$$

Com o tempo e utilização por vários autores, esta medida foi considerada como útil ao conhecimento do nível do desenvolvimento da linguagem da criança (Wells, 1995). A partir desta, diferentes formas começaram a surgir e as medidas de extensão de enunciados produzidos por crianças ganharam popularidade entre os investigadores com a publicação da

obra de Brown (1973). Roger Brown, na altura professor de Psicologia e investigador na Universidade de Harvard nos EUA, propôs um novo método para calcular a extensão de enunciados tendo por base o total de morfemas presente em cada enunciado, em vez do número total de palavras. O estudo deste autor teve por base o acompanhamento do desenvolvimento da linguagem de três crianças americanas, durante um período de tempo: Eve, Adam e Sarah. Esta nova medida ficou assim conhecida por Extensão Média do Enunciado em morfemas (EME-m), calculada através da seguinte fórmula:

$$\text{EME-m} = \frac{\text{n}^\circ \text{ total de morfemas}}{\text{n}^\circ \text{ de enunciados}}$$

Brown (1973) definiu-a como sendo uma excelente medida de desenvolvimento gramatical, e defendia que a EME-m pode ser considerada como uma medida de evolução progressiva da linguagem. Este autor demonstrou nos seus estudos que a maioria dos avanços ao nível do desenvolvimento da linguagem resulta de um aumento da extensão das frases, que decorre da adição de novas palavras e outros elementos linguísticos às frases (Brown, 1973).

Para a língua Inglesa, e através dos valores obtidos com a fórmula já apresentada, Brown (1973) definiu cinco níveis para o desenvolvimento da EME-m (Quadro 1). A partir do último valor indicado, para Brown (1973) os valores já não poderiam ser considerados como credíveis pois muito do desenvolvimento linguístico a partir dessa fase deve-se à reorganização interna das frases e não ao acrescento de novas estruturas (Brown, 1973). Cada nível apresentado, está associado às diferentes características de desenvolvimento que incluem: morfemas gramaticais, negação, questões sim/não, questões Qu- e frases complexas (Retherford, 2000).

Quadro 1. Níveis e valores de EME-m segundo Brown (1973)

Níveis	Valores EME-m	Idade média em meses
I	1,75	15-30
II	2,25	28-36
III	2,75	36-42
IV	3,5	40-46
V	4,0	42-52+

Além da obtenção de valores e da sua organização em níveis, Brown (1973) fez uma associação entre a EME e a idade cronológica da criança. Este autor considerava que a EME estava altamente relacionada com a idade cronológica da criança, ou seja, os valores de EME deveriam refletir um patamar específico do desenvolvimento da linguagem da criança. Por outras palavras, para Brown (1973), através do valor de EME obtido, poder-se-ia estimar a idade real da criança pois a sua idade deveria aumentar juntamente com o desenvolvimento da sua linguagem. No estudo com as três crianças já referidas: Eve apresentou o mesmo progresso gramatical de 1;7 para 2;3 que o Adam e a Sarah fizeram de 2;2 a 3;6. Desta forma, Brown (1973) foi capaz de criar uma medida de desenvolvimento linguístico que lhe permitiu observar que duas crianças que se encontrassem no mesmo nível de EME, seriam mais próximas ao nível do desenvolvimento linguístico do que duas crianças com a mesma idade cronológica.

Desde a publicação do estudo em 1973 que muitas pesquisas têm sido realizadas com o intuito de obter dados semelhantes e as conclusões apresentadas inicialmente. No entanto, os resultados nem sempre têm sido os esperados.

Relativamente à associação EME-idade cronológica, alguns autores foram capazes de reduplicar os resultados de Brown, como é o caso de De Villiers e De Villiers (1973) bem como Miller e Chapman (1981). Este últimos, por exemplo, estudaram a relação entre a idade de 123 crianças, entre os 17 e os 59 meses, com a EME obtida. Todas as crianças pertenciam à classe média alta e residiam no centro-oeste dos EUA, sendo o seu discurso gravado enquanto brincavam com as suas mães. Estes autores encontraram uma correlação significativa entre a idade e a EME ($r=0,88$). Outro estudo com o mesmo objetivo, mas realizado no Reino Unido, foi realizado por Rondal, Ghiotto, Brédart e Bachalet (1987) que reportaram um forte relação entre a idade cronológica e a EME num grupo de 21 crianças entre os 20 e os 32 meses. Mais recentemente, Parker e Brorson (2005) verificaram uma alta correlação entre os valores de EME-m obtidos na sua investigação e o desempenho morfosintático de crianças com idade compreendidas entre os 3;01 e 3;10, com desenvolvimento típico da linguagem. No entanto, esta relação nem sempre tem sido observada. Klee e Fitzgerald (1985), por exemplo, não encontraram uma correlação significativa, principalmente na faixa etária entre os 24 e os 48 meses.

As publicações da década de 70 e 80 também apresentaram outras críticas ao trabalho de Brown: Klee e Fitzgerald (1985) e Conant (1987) afirmaram que os valores de EME-m obtidos nas suas investigações, não podiam ser utilizados como índice de desenvolvimento gramatical uma vez que não refletiam a qualidade e complexidade que já é observável em

estruturas fráscas das crianças a partir dos cinco anos de idade. Para estes autores, assim como Scarborough, Wyckoff e Davidson (1986), a EME-m poderia ser considerada como uma medida que apenas permite observar a evolução linguística das crianças ao longo do tempo, mas não para realizar diagnósticos nem fazer previsões.

A referência do patamar atingido aos cinco anos também foi abordada por outros autores. A partir dessa faixa etária, foi observado que a EME sofria um alinhamento (Bol, 1996) ou mesmo um declive (Klee, 1992). Para Wiczorek (2000), foi possível observar que a tendência de aumento da EME parece ser inconstante e que de acordo com estes resultados, a EME não parece estar totalmente dependente da idade da criança e que, desta forma, a validade da correlação com a idade cronológica pode desaparecer a partir de certo momento (Wiczorek, 2000).

Para Scarborough (1990) a EME parece ser útil apenas nos primeiros anos de desenvolvimento linguístico mas depois dos quatro anos, deixa de fazer sentido pois as diferenças nos enunciados nessa altura já não são ao nível de extensão mas sim da complexidade destes.

Por seu lado, os estudos dos anos 90 e do ano 2000 parecem suportar os resultados e conclusões de Brown (1973), obtendo resultados que apresentaram uma maior proximidade e semelhança. Blake Quantaro e Onorati (1993), procuraram correlacionar os valores da EME com instrumentos de avaliação formais para a língua Inglesa, instrumentos estes que avaliam e caracterizam o desenvolvimento gramatical. Os resultados obtidos indicaram uma boa correlação entre a EME-p e a Bateria LARSP – *Language Assessment, Remediation and Screening Procedure* (Crystal, Fletcher & Garman, 1976), indicando assim que a avaliação da linguagem, realizada através da medida EME, poderia ser considerada como um bom instrumento para verificar o desenvolvimento gramatical das crianças com um desenvolvimento normal (Blake, *et al.*, 1993). Estes autores puderam então concordar com Brown (1973) na medida em que a EME aumenta através da utilização de formas sintáticas mais elaboradas, com a utilização de formas negativas, de coordenação e com a adição de morfemas obrigatórios. Para Wells (1995), a EME pode ser utilizada para descrever as primeiras fases do desenvolvimento da sintaxe, desde as primeiras combinações de duas palavras, para a génese da sintaxe complexa em crianças com desenvolvimento normal em idade pré-escolares.

Estudos mais recentes continuam a demonstrar o apoio que os investigadores dão a esta medida, continuando a utilizá-la nos seus estudos. Rispoli (2003), realizou um estudo longitudinal, no qual observou um conjunto de crianças, entre os 22 meses e os quatro anos de

idade, com um desenvolvimento típico, tendo como principal objetivo verificar a utilização de estruturas gramaticais mais elaboradas ao longo do tempo. O autor pôde observar que até aos três anos, as crianças produzem frases mais simples e curtas, o que varia com o aumento da idade, começando a produzir frases cada vez mais extensas e complexas (Rispoli, 2003). Neste mesmo estudo, também foi observado um aumento dos valores da EME à medida que as crianças reduziam a quantidade de revisões, repetições silábicas e reinício de frases, o que é típico em crianças mais novas (Rispoli, 2003). Parker e Brorson (2005) também concluíram que a utilização da EME-m pode ser indicada para a verificação das competências gramaticais em idades pré-escolares. Outra vantagem apresentada por Stockman (1996, citado por Hewitt, *et al.*, 2005) é o facto da EME parecer apresentar uma menor vulnerabilidade a diferenças dialetais e culturais, em comparação com a dos testes formais.

Tendo em consideração o intuito deste trabalho, é importante perceber que estudos têm sido realizados acerca da influência do género da criança e do nível sociocultural a que pertence nos valores de EME que apresenta. Apesar da maioria dos estudos não fazer distinção entre géneros, estudando o grupo de crianças na globalidade (Blake, *et al.*, 1993; Hewitt, *et al.*, 2005; Parker & Broson, 2005), é uma questão importante de analisar visto que, ao nível do desenvolvimento da linguagem, alguns resultados demonstram que os rapazes começam a produzir as primeiras palavras (Maccoby, 1966, citado por Özçaliskan & Goldin-Meadow, 2010), e as primeiras frases, mais tarde do que as raparigas e que, mesmo ao nível da utilização de pequenos e simples gestos, as raparigas começam mais cedo (Özçaliskan & Goldin-Meadow, 2010). Além disso, alguns estudos também têm demonstrado que as raparigas têm um vocabulário mais completo e usam uma maior variedade de frases na sua comunicação mais inicial (Ramer, 1976). Em idades escolares, as raparigas também parecem ter mais sucesso em todas as competências verbais que os rapazes (Davies, 2007). Desta forma, Schachter (1979), ao comparar a EME de crianças de acordo com a idade, género e raça, não observou diferenças entre rapazes e raparigas. Por seu lado, Jackson e Roberts (2001), ao analisar a complexidade sintática produzida por 85 crianças afro-americanas entre os três e os quatro anos de idade, observaram que as raparigas produziam formas sintáticas mais complexas que os rapazes e que esta complexidade tinha uma correlação positiva com a EME-p. Num grande estudo realizado com 316 crianças francesas entre os 24 e os 48 meses de idade, Le Normand, Parisse e Cohen (2008) analisaram a EME e o TTR após a gravação do seu discurso enquanto brincavam. Estes autores observaram uma ligeira vantagem na produção linguística das raparigas em relação aos rapazes, mas apenas até aos 36 meses. Observaram também uma forte influência do nível sociocultural dos pais, com crianças de

nível alto a demonstrar produções mais complexas e um nível de desenvolvimento mais avançado. Esta grande influência foi também observada por Walker, Greenwood, Hart e Carta (1994) no seu estudo longitudinal com diferentes níveis socioeconómicos, observando que a EME das crianças que participaram nos estudo estavam extremamente correlacionados com os múltiplos índices de estrato socioeconómico (educação dos pais, estrato ocupacional e ordenado). Estes resultados não foram, no entanto, encontrados por Hoff (2003) nem por Rice *et al.* (2010). O primeiro não encontrou diferenças significativas num conjunto de 63 crianças, entre os 16 e os 31 meses, relativamente à influência do meio na sua EME. Rice *et al.* (2010), investigaram a possível associação entre os valores de EME-p e EME-m de um grupo de 306 crianças e o nível de educação das suas mães, não tendo encontrado uma correlação estatisticamente significativa ($r=0,38$).

Apesar das diversas vantagens e das evidências apresentadas por diversos autores em suporte à utilização da EME, dois fatores têm levado a que nem sempre seja uma medida utilizada na investigação ou para fins clínicos. Estes fatores são: i) o tempo que é necessário despende e a dificuldade de análise tendo por base os morfemas e; ii) as dificuldades de adaptação dos resultados obtidos para outras línguas, que não o Inglês.

Relativamente ao primeiro ponto, Eisenbeiss (2010) defende que os cálculos da EME-m envolvem difíceis e demoradas decisões ao nível dos critérios de determinação do número de morfemas nas diferentes línguas (Brown, 1973; Eisenberg *et al.*, 2001), pelo que alguns investigadores têm utilizado dados da EME-p, que corresponde à proposta já referida de Nice (1925). Esta é uma proposta que permite facilitar o trabalho dos clínicos aquando da utilização com crianças em processo de avaliação.

Alguns estudos têm encontrado altas correlações entre a EME-m e a EME-p, o que levou a que muitos investigadores a considerassem uma medida mais confiável e de mais fácil contagem (Parker & Brorson, 2005). Gutiérrez-Clellen, Restrepo, Peña e Anderson (2000), sugerem que o cálculo da EME-p seja utilizado, principalmente quando se pretende realizar estudos comparativos entre línguas, uma vez que a contagem de palavras minimiza as diferenças morfológicas que podem interferir no cálculo da EME-m. Arif e Bol (2008) procederam a uma investigação que tinha por objetivo comparar os resultados da EME-m e EME-p. Para tal, utilizaram uma amostra de discurso de uma criança com desenvolvimento de linguagem típico e outra amostra de uma criança com perturbação de linguagem, recolhidas da base de dados CHILDES. O resultado demonstra que existe uma forte correlação entre as duas medidas e, os autores, concluem que, se o objetivo for obter um valor de EME, que é preferível calcular a EME-p pois é mais fácil e menos técnica em termos de complexidade

gramatical (Arif & Bol, 2008). Oosthuizen e Southwood (2009), num estudo realizado com crianças sul africanas, calcularam a EME oito vezes para cada participante e encontraram uma correlação positiva ($>0,96$) entre a EME-p e a EME-m.

Wieczorek (2010) também obteve resultados no seu estudo que sugerem uma alta correlação entre a EME-p e a EME-m em crianças com perturbação de linguagem, assim como os já referidos resultados obtidos por Parker e Brorson (2005). No entanto, Wieczorek (2010) não concordou com Parker e Brorson (2005) na medida em que a EME-p pudesse ser utilizada como uma medida alternativa. Para esta autora, as diferenças encontradas ao nível dos valores obtidos entre a EME-p e EME-m poderão indicar que os dois índices refletem dois tipos diferentes de desenvolvimento linguístico (Wieczorek, 2010). Nomeadamente a EME-p fornece informação acerca do léxico, enquanto a EME-m ilustra o desenvolvimento morfosintático. Desta forma, também é afirmado por alguns autores (Peters, 1999) que a alta correlação entre estes dois parâmetros não é suficiente para se substituírem um ao outro. Este autor afirma também que a EME-p é um contínuo de palavras e que a EME-m é a inclusão de marcadores gramaticais nesse contínuo, defendendo então que ambos fenómenos devem ser avaliados separadamente (Peters, 1999).

Relativamente à adaptação às diferentes línguas, de acordo com Hickey (1991) é necessário conhecer bem as estruturas e funcionamento da língua em estudo para que as adaptações sejam mais corretas. Para tal, o mesmo autor refere que o uso da EME só é indicado para línguas cujos processos de aquisição e desenvolvimento já tenham sido muito estudadas de forma a verificar e confirmar se os valores da EME encontrados realmente representam o estadió de desenvolvimento da criança (Hickey, 1991). Brown (1973) já tinha referido que estas adaptações deveriam ser realizadas em línguas que apresentem, por exemplo, um grande número de flexões como é o caso do Espanhol, Português, Sueco e Italiano.

Vários estudos têm vindo a ser realizados noutras línguas, de forma a compreender se a EME poderá ser utilizada como medida de desenvolvimento linguístico, como é para a língua inglesa. Apesar das diferenças entre as estruturas morfológicas e sintáticas de cada língua, os estudos parecem apontar para a possibilidade de utilização da EME sendo, no entanto, necessário adaptá-la à língua em estudo, sendo estas adaptações sido alvo de críticas por alguns autores. Alguns exemplos de estudos realizados são o de Hickey (1991) que aplicou esta medida para o Irlandês, o estudo de Leonard e Bortolini (1998) que a aplicaram para o Italiano e Klee, Stokes, Wong, Fletcher e Gavin (2004) para o Cantonês.

Apesar destas evidências da elevada qualidade da EME-m para a identificação e avaliação das produções efectuadas por crianças de língua inglesa, e dos diversos estudos noutras línguas que têm sido realizados de forma a perceber se esta medida pode ser considerada universal (Guitérrez-Clellen, Restrepo, Peña, & Anderson, 2000; Hickey, 1991; Leonard & Bortolini, 1998), para o Português Europeu (PE), não têm sido elaborados estudos que possam indicar se esta medida pode ser desenvolvida e aplicada, como o é para a língua Inglesa. Ao nível do Português do Brasil (PB), no entanto, alguns estudos têm vindo a ser realizados como o de Araújo e Befi-Lopes (2003) que descrevem de forma detalhada a observação de um aumento significativo dos valores de EME-m e EME-p conforme o aumento da idade em crianças em idade pré-escolar, com um desenvolvimento normal, entre os dois e os seis anos de idade.

Quadro 2. Valores de EME-p e EME-m obtidos por Araújo e Befi-Lopes para PB (2003)

Idade	EME-p	EME-m
2 anos	2,35	3,21
3 anos	2,83	3,72
4 anos	3,53	4,55
5 anos	4,73	6,01
6 anos	6,85	9,01

As principais diferenças entre as várias faixas etárias ocorreram aos quatro anos, onde, segundo Araújo e Befi-Lopes (2003), as regras linguísticas da morfologia e da sintaxe da língua já estão, relativamente, organizadas nas crianças.

Todas as vantagens apresentadas acerca da EME parecem ser importantes para os Terapeutas da Fala americanos, como comprovado pelo estudo realizado por Loeb, Kinsler e Bookbinder (2000), em que 93% dos Terapeutas afirmavam utilizar uma análise de uma amostra de linguagem espontânea para complementarem o seu diagnóstico, sendo que 91% desses utilizava a EME.

Resumindo, a EME, tanto em palavras como em morfemas, é uma medida de análise de amostras de discurso espontâneo que teve um grande avanço desde a década de 70. Apesar de alguns estudos realizados, na grande maioria dos casos, durante a década de 80, não terem obtido resultados concordantes com os de Brown (1973), a ideia geral é que a EME vai aumentando com o avanço da idade até certo momento em que se dá uma estabilização (isto para a língua Inglesa). É uma medida que permite analisar o desenvolvimento linguístico das

crianças. São vários os estudos que demonstram uma alta correlação entre a EME-m e a EME-p, sendo que esta ganha assim algum suporte por ser mais simples de obter resultados e para comparar línguas. A adaptação a diferentes línguas é um processo complexo, mas possível. Os estudos acerca da influência do género não são conclusivos mas o nível sociocultural dos pais parece influenciar positivamente o desenvolvimento dos seus filhos.

Possíveis utilizações da EME

Além de ter vindo a permitir uma compreensão da forma como as produções orais das crianças, com desenvolvimento típico de linguagem, se manifestam, a EME tem sido aplicada em diversos estudos com outras funções.

A função mais comumente estudada é o auxílio ao diagnóstico de crianças com perturbações da linguagem e conseqüente comparação da sua evolução com crianças com desenvolvimento típico. Dentro destas perturbações, as mais estudadas são as PEDL. Klee, Schaffer, May, Mambrino e Mougey (1989), procuraram correlacionar a idade das crianças com a EME obtida através da gravação dos seus discursos enquanto interagiam com as suas mães. Encontraram uma alta correlação nas 24 crianças com desenvolvimento típico da linguagem e também uma alta correlação entre as 24 crianças com PEDL (Klee, *et al.*, 1989). Observaram ainda que a EME do grupo com PEDL era inferior à do grupo de controlo, com a mesma idade. Estes mesmos autores observaram que a EME evoluía a um ritmo semelhante nos dois grupos estudados, o que vem contrariar a ideia que crianças com perturbação linguística tenham níveis mais lentos de desenvolvimento da linguagem (Klee, *et al.*, 1989). Outro estudo em que a EME foi utilizada como medida a ter em consideração na identificação e diagnóstico de crianças com PEDL foi realizado por Moyle, Karasinski, Weismer e Gorman (2011).

Hewitt, *et al.* (2005), também concluíram que a EME-m pode ser uma boa medida para avaliação de crianças com e sem perturbação de linguagem. No entanto, os seus resultados não demonstraram valores de EME muito baixos para as crianças com PEDL, como seria de esperar, o que poderá ser explicado pela forma como o discurso foi elicitado durante a fase de recolha dos enunciados (Hewitt, *et al.*, 2005). A metodologia escolhida foi a de uma entrevista estruturada por parte dos avaliadores e não uma conversação, o mais espontânea possível, guiada pela criança como é aconselhado por outros estudos (Parker & Brorson, 2005).

Rice, Redmond e Hoffman (2006), apresentaram resultados que indicaram que a EME foi positivamente associada com o crescimento da complexidade das estruturas sintáticas, sugerindo que a EME-m é uma medida válida, tanto para crianças com desenvolvimento típico, tanto para crianças com PEDL. Num estudo de grande dimensão, Rice, *et al.* (2010), calcularam a EME-p e EME-m de um total de 306 crianças, 170 com PEDL e 136 com desenvolvimento típico da linguagem. Apesar de não poderem garantir que as amostras de linguagem eram as mais representativas de cada criança, pelo facto de terem acedido as estas através do CHILDES, estes autores observaram que as crianças com PEDL apresentavam uma EME-p e EME-m inferior à do grupo de crianças com desenvolvimento típico.

Befi-Lopes, Nuñez e Cáceres (2011) encontraram uma correlação alta entre o vocabulário expressivo e a EME de crianças com PEDL.

Outro grupo de crianças com uma perturbação linguística, neste caso com Trissomia 21, foi estudado por Marques e Limongi (2011) no Brasil. Estes autores verificaram a eficácia da utilização da EME-p como medida do desenvolvimento linguístico geral neste grupo de crianças. Observaram que quanto maior a idade da criança, maiores os valores de EME-p que conseguiam obter, apesar destes valores serem sempre mais baixos do que nas crianças com desenvolvimento típico. Outro estudo realizado com a mesma patologia foi realizado por van Bysterveldt, Westerveld, Gillon e Foster-Cohen (2012). Neste caso, no entanto, o principal objetivo foi investigar a qualidade das narrativas pessoais de 25 crianças, com idades compreendidas entre os 5;11 e os 13;01 anos. Narrativas pessoais, de acordo com estes autores, são descrições de eventos passados que foram experienciados pelo falante. As narrativas destas crianças neozelandesas foram elicitadas pelos Terapeutas da Fala que os acompanham, através da apresentação de 11 fotografias. As amostras de discurso foram depois transcritas e analisadas ao nível da EME, NDW e qualidade da narrativa. van Bysterveldt, *et al.* (2012) observaram que a EME de 92% das crianças foi muito baixa, não tendo sido encontrada uma correlação entre a EME obtida e a idade da criança, nem com a qualidade da narrativa.

As perturbações do Espectro do Autismo também têm sido alvo de estudos com base na EME. Foudon, Reboul e Manificat (2007), realizaram um estudo longitudinal relativo ao processo de aquisição da linguagem em nove crianças autistas francesas, entre os três e os nove anos de idade. Diversas amostras de discurso foram recolhidas ao longo de três anos, sendo recolhidas em três contextos: hora do almoço, enquanto brincavam, e durante as horas de terapia. Ao longo deste tempo, e Foudon, Reboul e Manificat (2007) estudaram a evolução da sua EME, observando grandes variações de criança para criança, e mesmo na própria

criança em diferentes ocasiões, bem como fases de estabilização. Por seu lado, Lopes-Herrera e Almeida (2008) dedicaram-se à criação e implementação de uma forma de intervenção específica de forma a aumentar os valores de EME-p em crianças com Síndrome de Asperger e Autismo com alta funcionalidade.

O desenvolvimento da linguagem de crianças prematuras também é de extrema importância para diversos profissionais. Seidman, Allen e Wasserman (1986) compararam o desenvolvimento da linguagem de três grupos de crianças de dois anos de idade: crianças com desenvolvimento típico, crianças prematuras e crianças com alterações físicas. Através da gravação do seu discurso em conversação com as suas mães, foi observado que o segundo e o terceiro grupo apresentavam uma EME inferior às crianças do primeiro grupo.

Ainda relativo à patologia, mas com uma perspectiva um pouco diferente, alguns autores têm procurado testar a eficácia de programas de intervenção tendo em conta a EME, entre outras medidas. Cable e Domsch (2011) na sua revisão acerca da intervenção com crianças com desenvolvimento tardio da linguagem, procuraram estabelecer qual o programa de intervenção que parece ter melhores resultados, tendo por base os resultados dos testes formais e dos valores da EME pré e pós intervenção. Com base nos resultados de todos os artigos e estudos que analisaram e compilaram, estes autores propõem uma forma de intervenção específica para este grupo de crianças. Outro estudo que utilizou a EME como medida para testar a eficácia de programas de intervenção foi elaborado por van Balkom, Verhoeven, van Weerdenburg e Stoep (2010). Neste caso, o programa de intervenção baseava-se na intervenção dos pais envolvendo o *Video Home Training* (PVHT) que fornece estratégias para ajudar os pais a estabelecer uma maior coerência nas conversações com os seus filhos, que neste caso tinham todos um Atraso no Desenvolvimento da Linguagem (ADL). Este tipo de intervenção era então comparado com a de uma intervenção direta do Terapeuta da Fala com a criança. Para tal, participaram neste estudo 11 crianças com os seus pais e outras 11 crianças em intervenção direta com a Terapeuta. Foi observado, através da análise dos resultados, que as crianças na metodologia do PVHT demonstravam efeitos significativos da EME a curto e a longo prazo, melhores que as crianças em intervenção direta, após a intervenção e em *follow-up*.

Com o objetivo de testar também a importância da intervenção da família, ao nível das competências expressivas de crianças com PEDL, Allen e Marshall (2011) foram analisar alguns parâmetros, antes e após o tratamento, para perceber se esta participação influencia, ou não, o seu desenvolvimento linguístico. Desta forma, participaram no estudo 16 crianças com PEDL e os seus pais. Metade destas crianças tinham apenas terapia com os seus Terapeutas

(grupo de controlo) enquanto a outra metade, além dessa terapia, tinha também PCIT: *Parent-Child Interaction Therapy*, ao longo de quatro semanas. Percebeu-se que, apesar do pequeno tamanho da amostra, que as crianças com PCIT e intervenção melhoravam mais ao nível da sua expressão verbal, tanto ao nível da iniciação do tópico de conversação como da EME.

O desenvolvimento da linguagem em crianças bilingues também tem sido alvo de diversos estudos ao longo dos anos. Por todas as vantagens que apresenta, a EME tem sido utilizada como medida de comparação entre as duas línguas em aquisição (Blom, 2010; Yip & Matthews, 2006) e auxílio no processo de avaliação, muitas vezes complexo pela influência de uma língua na outra (Carmen & Wong, 2012; Gutiérrez-Clellen & Simon-Cereijido, 2009). Blom (2010), estudou o desenvolvimento de crianças bilingues, falantes de turco e holandês, com diferentes quantidades de *input* para as duas línguas, comparando-as depois com crianças monolíngues das duas línguas. Para tal, basearam-se na EME-p para perceber que as crianças bilingues apresentavam um atraso nas duas línguas mas com maior problema no Holandês. Por seu lado, Yip e Matthews (2006) basearam-se na EME para analisar o desenvolvimento linguístico de crianças bilingues para Cantonês e Inglês e assim perceber se existe dominância de uma das línguas ou não. Perceberam que a língua para a qual a criança tinha maior EME era a que sintaticamente se encontrava mais desenvolvida e que a criança apresentava preferência por essa língua em especial (Yip & Matthews, 2006).

A EME também pode ser utilizada para auxiliar o processo de obtenção de valores para outras medidas que permitem analisar algum parâmetro linguístico da criança. Um exemplo de uma destas medidas é a produção de narrativas. Balčiūnienė (2012), analisou as narrativas produzidas por 24 crianças lituanas, entre os seis e os sete anos de idade, através do relato da “História do Gato” (Hickmann 1993). Após as gravações, transcrições e anotações através do *software* CHILDES, foram investigadas a complexidade sintática, a diversidade lexical e produtividade geral através das medidas EME-p e TTR, obtendo assim elementos acerca das tendências microestruturais das narrativas nesta língua. Esta autora também observou que os valores de EME-p durante a narrativa eram mais altos do que os valores obtidos durante o seu discurso espontâneo e que um valor de EME-p muito alto na narrativa não significa que as suas capacidades narrativas sejam boas (Balčiūnienė, 2012).

Outra possível utilização desta medida é a análise de diferenças entre dialetos. Nos EUA, por exemplo, país de grande diversidade cultural e linguística, a EME tem sido utilizada para analisar o desenvolvimento da linguagem de crianças afro-americanas pois os testes formais raramente são aplicáveis a esta população, o que dificulta gravemente a avaliação e diagnóstico destas crianças (Craig & Washington, 2000; Oetting, 2005).

Em suma, a EME é uma medida que pode ter diversas aplicações, a partir do momento em que são obtidos e estabelecidos os valores-padrão para cada língua e faixas etárias. Dentro destas possíveis utilizações estão o auxílio ao diagnóstico e as comparações entre crianças com desenvolvimento típico da linguagem e crianças com perturbações linguísticas como PEDL, autismo, ADL, crianças prematuras. Pode ser também útil no estudo do desenvolvimento das línguas faladas por crianças bilingues e nas diferenças entre dialetos de um mesmo país. Tem permitido também obter novas medidas de análise do desenvolvimento linguístico na criança como as características da narrativa, e ainda analisar a eficácia de programas de intervenção com crianças com determinadas perturbações.

2. Problemas em estudo

Após a decisão de qual o tema a investigar e através da informação obtida na revisão da literatura, foram colocadas três questões orientadoras para este estudo:

- **Questão 1:** Será que a EME-p é uma medida fidedigna da aquisição da linguagem pela criança, para o Português Europeu?
- **Questão 2:** Será que o género das crianças influencia a EME-p?
- **Questão 3:** Será que a escolaridade dos pais influencia a EME-p dos seus filhos?

Deste forma, foram formulados os seguintes objectivos de estudo:

- **Objetivo 1:** Estabelecer o valor padrão da EME-p para a faixa etária dos 4,00 aos 4,05 anos.
- **Objetivo 2:** Verificar se existem diferenças significativas na EME-p entre crianças do género masculino e feminino.
- **Objetivo 3:** Verificar se existem diferenças significativas na EME-p de crianças, de acordo com o grau de escolaridade dos pais.

De forma a cumprir os objectivos referidos, planeou-se um estudo de design quasi-experimental do tipo descritivo-comparativo visto ter sido utilizada uma amostra não-probabilística de crianças entre os 4;00 e os 4;05 anos tendo por principal objetivo estudar e descrever a sua EME-p. Além disso pretende-se também comparar os resultados obtidos em ambos os géneros, verificando-se assim se estes apresentam alguma influência assim como o grau de escolaridade dos pais das crianças nos resultados apresentados.

Este estudo está, como já referido, inserido num estudo mais alargado com outras duas investigadoras, no qual se pretende estabelecer valores da EME-p para crianças dos 4 aos 5 anos e cinco meses.

3. Metodologia

Participantes

Como já referido, foi utilizada uma amostra não-probabilística de conveniência, sendo constituída por 30 crianças entre os 4;00 e os 4;05 anos. Estas 30 crianças frequentavam jardins de infância da rede pública da área da grande Lisboa, mais especificamente dos concelhos de Sintra, Amadora, Cascais, Lisboa e Loures. Apesar de se ter previsto também a inclusão de crianças do conselho de Oeiras, nenhum jardim de infância deste conselho se mostrou interessado em participar no estudo.

Numa primeira fase, a seleção dos jardins de infância foi feita de forma aleatória, selecionando-se dois jardins de infância por cada concelho. No entanto, após uma série de dificuldades em obter respostas positivas por parte dos diretores destes jardins de infância e a um número insuficiente de crianças nas condições previstas, procedeu-se a uma seleção por conveniência dos jardins, tendo em consideração o número de crianças em falta para cada faixa etária, o seu género, o concelho no qual residia e o número de crianças por sala de aula.

De forma a selecionar os participantes para o estudo, foram definidos os seguintes critérios de inclusão para cada criança:

- Apresentar uma idade compreendida entre os 4;00 e os 4;05 anos de idade;
- Ter como única língua materna o Português Europeu;
- Não apresentar perturbações do desenvolvimento que pudessem afetar a sua linguagem;
- Ausência de défices sensoriais (visuais e/ou auditivos);
- Não ser, ou nunca ter sido seguido em Terapia da Fala;
- Apresentar um discurso inteligível;
- Obter um resultado no Teste de Avaliação da Linguagem na Criança (TALC) dentro dos valores normais para a idade (entre -1 e +1 DP).

De forma a garantir que os critérios anteriormente mencionados se encontravam presentes, procedeu-se a uma entrevista aos educadores das crianças dos diversos jardins de infância. No casos em que persistia alguma dúvida relativamente a algum dos critérios, foi estabelecido um contacto mais direto com os pais. Foi aplicado o TALC (Sua-Kay, E. & Tavares, M. D., 2008) a cada uma das crianças.

Pelo não cumprimento de algum destes critérios, foram excluídas 17 crianças desta faixa etária, de um total de 47 (11 rapazes e 6 raparigas):

- dois rapazes e três raparigas por serem bilingues para o Português Europeu e outra língua (Crioulo, Ucrainiano, Mandarim, Moldavo e Romeno);
- um rapaz por falar português do Brasil;
- duas raparigas e dois rapazes por estarem a ser seguidas por um Terapeuta da Fala por apresentaram perturbações linguísticas;
- um rapaz e uma rapariga por suspeita de défices cognitivos por parte das educadoras e que se encontravam em processo de espera para avaliações de desenvolvimento;
- cinco rapazes por apresentaram resultados inferiores a um desvio-padrão no TALC.

Desta forma, após todo o processo de avaliação, recolha e análise do discurso espontâneo, participaram no estudo 30 crianças, mais especificamente 12 rapazes e 18 raparigas dos 4;00 aos 4;05, ou seja, dos 48 aos 53 meses. A distribuição das crianças por meses, e em géneros encontra-se representada no Quadro 3.

Quadro 3. Distribuição de rapazes e raparigas por idade

Faixa etária	Género Masculino	Género Feminino	Total por faixa etária
4;00 (48 meses)	1	1	2
4;01(49 meses)	2	3	5
4;02 (50 meses)	2	3	5
4;03 (51 meses)	4	3	7
4;04 (52 meses)	2	3	5
4;05 (53 meses)	1	5	6
Total	12	18	30

A distribuição das crianças pelo nível de escolaridade dos pais encontra-se apresentada no quadro abaixo (Quadro 4). Para a análise, foi utilizada a escolaridade mais elevada, do pai ou da mãe.

Quadro 4. Escolaridade dos pais por género da criança

Género	4 anos	6 anos	9 anos	12 anos	Ensino Superior
Rapazes	-	-	2	4	6
Raparigas	1	2	5	4	6
Total	1	2	7	8	12

Procedimentos

Todo o processo foi iniciado através do envio de cartas à direção dos agrupamentos de escolas aos quais os jardins de infância selecionados pertenciam (ver Apêndice I). Nesta carta eram descritos os objetivos do estudo, pedindo-se autorização dos jardins de infância para a presença das investigadoras no espaço escolar, para a interação com as crianças, bem como da utilização de uma sala disponível. Uma vez a autorização dada, foram distribuídas pelas educadoras de infância uma série de cartas para serem entregues aos (ver Apêndice II). Estas cartas para os pais explicavam o objetivo do estudo e pediam a sua colaboração através da autorização da participação dos seus filhos no estudo. Esta autorização era fornecida através do preenchimento do consentimento informado no final da carta. De forma a facilitar o processo de organização da informação, foi construído pela investigadora um documento onde constassem todas as informações de cada criança (Apêndice III).

As crianças, cujos pais autorizavam a participação no estudo, eram então avaliadas através do TALC que foi escolhido por permitir a avaliação linguística de crianças nas faixas etárias pretendidas. Após esse processo de avaliação, os discursos das crianças que obtiveram os resultados dentro dos valores previstos eram gravados no processo de conversação com a investigadora. Esta gravação áudio foi realizada através do gravador Olympus WS-650S escolhido para o efeito pelas suas características de redução do ruído e boa captação de som a 50cm de distância, o que era fundamental para que não tivesse de estar muito próximo da criança.

De forma a elicitar o discurso espontâneo que seria gravado, foram utilizados quatro conjuntos de brinquedos e miniaturas que serviram de instrumentos para a interação entre a criança e a investigadora, que eram compostos por:

- i. loiças: panelas, frigideira, pratos, talheres, espátula, copos;
- ii. alimentos: bolos, chocolate, pão e croissant, frutos, legumes, carne, hambúrguer no pão, cachorro quente;
- iii. animais: gorila, leão, girafa, hipopótamo, tubarão, porco, cabra, cavalo, vaca;

- iv. objetos de médico: termómetro, seringa, óculos, pinça, mala de médico, estetoscópio, otoscópio e martelo de reflexos.

Os brinquedos escolhidos e utilizados podem ser observados na Figura 4.



Figura 4. – Brinquedos utilizados para a recolha de amostras de discurso espontâneo

Importa ainda referir que, a pedido de várias educadoras e de alguns pais, foi elaborado um pequeno relatório de avaliação para cada criança avaliada com o TALC. Desta forma, foram criados dois modelos de relatórios: um de inclusão no estudo, quando a criança obteve um resultado dentro dos valores esperados (ver Apêndice IV) e um de exclusão, quando tal não acontecia (ver Apêndice V).

Como já referido anteriormente, este estudo teve início através da autorização dos diretores dos agrupamentos de escolas. Para tal, foram sempre necessárias reuniões com o diretor ou outro elemento da direção de forma a complementar a informação fornecida na carta. Nesta reunião era então pedido que a sala em que se processassem as gravações fosse uma sala em que não houvesse interrupções em que a investigadora pudesse estar apenas com a criança. De seguida, foi acordado com as diferentes educadoras de infância a organização de todo o processo (datas e horários convenientes).

Relativamente ao processo de avaliação das crianças, as investigadoras foram previamente treinadas de forma a que o processo de aplicação do teste decorresse sempre de

forma a garantir a melhor aplicação do teste e garantir que não eram cometidos erros metodológicos de avaliação. Sempre que decorreram dúvidas ao nível da cotação de algum item do TALC, as três investigadoras reuniam-se de forma a que as decisões tomadas fossem sempre no mesmo sentido.

No dia da recolha das amostras de discurso espontâneo, a investigadora ia buscar cada criança à sua sala e seguia até à sala da gravação. No percurso entre os diferentes ambientes, a investigador conversava com a criança e explicava-lhe que iriam brincar com alguns brinquedos. Quando a criança se encontrava sentada, a investigadora ligava o gravador e convidava a criança para brincar. Para recolher a amostra de discurso, a investigadora interagiu com a criança num contexto lúdico que teve a duração de 30 minutos, esperando-se assim que este tempo fosse suficiente para que produzisse, pelo menos, 100 enunciados válidos. As investigadoras foram previamente instruídas para que evitassem ao máximo as questões de resposta “sim” / “não” e para que evitassem o domínio das interações verbais com a criança.

Análise do corpus linguístico

De forma a que as três investigadoras praticassem uma igual contabilização de palavras e enunciados, foi necessário adoptar e estabelecer algumas regras prévias. Estas regras basearam-se em propostas de outros autores e, no caso de alguma regra não se adaptar ou não ter sido mencionada em algum estudo, as investigadoras, em conjunto com as suas orientadoras, estabeleceram o que consideraram necessário. Importa ainda referir que nem todas as regras do mesmo autor foram aplicadas, sendo, desta forma seleccionadas as regras que melhor se adaptavam ao estudo, mesmo que tal significasse a utilização de regras propostas por mais de um autor. As regras estabelecidas para a gravação e contagem dos enunciados foram as seguintes:

- Seriam contabilizados os primeiro 100 enunciados (Brown, 1973);
- Seriam excluídas do estudo todas as crianças que não produzissem 100 enunciados (Parker & Brorson, 2005);
- A gravação decorreria de 20 a 30 minutos, o que seria suficiente para obter esse número de enunciados (Osthuizen & Southwood, 2009);
- Não seriam excluídos os primeiros cinco minutos de conversação / gravação como proposto por alguns autores (Brown, 1973; Marques & Limongi, 2011), pois as investigadoras iam buscar cada uma das crianças à sua sala e os primeiros momentos

de conversação decorriam durante o percurso percorrido. Desta forma, foi respeitado o período de cerca de 5 minutos antes de ser iniciada cada gravação.

Para se definir como delimitar um enunciado, foi utilizada a definição proposta por Lund e Duchan (1993) que um enunciado é delimitado pela sua curva entoacional, que o final de uma frase é o final do enunciado e que quando é efectuada uma pausa de dois segundos, o enunciado termina. Além disso, estes mesmos autores sugerem que: um grupo de palavras pode ser dividido sem perder o significado essencial, um enunciado pode não ser uma frase e que uma frase com duas orações independentes, unidas por uma conjunção de coordenação, é contabilizada como um enunciado (Lund & Duchan, 1993). Foi ainda estabelecido que:

- São seriam contabilizadas imitações de enunciados, produzidos imediatamente após a sua produção por parte da investigadora (Lund & Duchan, 1993);
- No caso de enunciados repetidos exatamente da mesma forma, este seria contabilizado apenas uma vez (Brown, 1973);
- São seriam contabilizadas músicas, enumerações, contagens e palavras isoladas (Lund & Duchan, 1993);
- Se num enunciado, o número de palavras ininteligíveis fosse mais de uma, tornando assim a frase imperceptível, este enunciado seria excluído (Brown, 1973);
- A expressão “e depois”, extremamente utilizada pela grande maioria das crianças gravadas, seria utilizada para dividir enunciados;
- Seriam contabilizados todos os enunciados mesmo que contivessem erros morfossintácticos, pois estes são comuns nas faixas etárias estudadas.

Os dois últimos pontos foram definidos pelas investigadoras em conjunto com as suas orientadoras após a realização das gravações por se ter percebido que se repetiam em quase todas as amostras recolhidas.

Relativamente às palavras produzidas em cada enunciado, foi determinado que:

- Caso a criança produzisse uma palavra ininteligível, o enunciado seria contabilizado mas aquela determinada palavra não entraria na contagem para o número de palavras (Wieczorek, 2010);
- Palavras onomatopéicas foram contabilizadas como palavras (Wieczorek, 2010);
- Os auxiliares de discurso, como exclamações, não foram contabilizadas (Brown, 1973);

- Palavras compostas eram contabilizadas como uma só palavra;
- Repetição de palavras no mesmo enunciado só seriam contabilizadas se esta repetição tivesse como objectivo enfatizar alguma ideia (Brown, 1973).

Após a recolha das amostras de discurso espontâneo das 30 crianças, procedeu-se à sua transcrição. Primeiramente, esta foi realizada em formato Word, sendo, mais tarde, adaptada para o *software* de transcrição ELAN. O ELAN é uma ferramenta disponível para a criação de anotações complexas em documentos áudio e vídeo, sendo uma das ferramentas que constituem o *The Language Archive* do Max Planck Institute for Psycholinguistics.

Para as transcrições, foram utilizadas algumas convenções de transcrição do Projeto BIBIBI (Projeto Bimodal Bilingue Bi-nacional)* que também são baseadas nas convenções usadas no CHILDES:

- Os enunciados começam sempre com letra minúscula, sendo que apenas nomes próprios começam com letra maiúscula;
- O símbolo “//” foi utilizado quando o fala foi interrompida de forma abrupta, ou seja, quando um dos interlocutores foi interrompido pelo outro, quando a conversação foi interrompida por algum barulho repentino, por outra pessoa na sala, etc.;
- O símbolo “...” foi utilizado quando a fala foi interrompida mas não de forma abrupta, ou seja, quando a criança não terminou o enunciado por mudar de tema, por começar a responder a uma questão à qual não sabe a resposta, quando está a descrever o que vai executando com os brinquedos;
- Quando a palavra não foi reconhecida por ser ininteligível ou devido a ruído de fundo que não permite a sua descodificação, recorreu-se ao símbolo “XXX”.

Sempre que surgiu alguma dúvida ao nível da separação de enunciados, inteligibilidade, repetição de palavras, entre outras, as investigadoras reuniram-se para assim retirar a dúvida ou tomar uma decisão sobre o que fazer. Os valores de EME-p obtidos por cada criança, assim como todos os seus dados encontram-se apresentados de forma mais organizada em Apêndice VI.

* Projeto da autoria de Lilli-Martin, Quadros e Pichler (2011, não publicado)

Exemplos de delimitação de enunciados e contagem de palavras

Exemplo 1

Investigadora – “e onde vai ser o passeio?”

Criança – “é em Coruche” (3 palavras)

Investigadora – “e já sabes o que vais ver?”

Criança – “hm hm animais” (1 palavra: não contabilizado) *pausa de 3 segundos* “e vamos lanchar lá” (4 palavras)

Exemplo 2

Investigadora – “então e agora?”

Criança – “os burros comem XXX os XXX depressa” (não contabilizado) *pausa de 2 segundos* “os leões XXX aqui XXX” (não contabilizado)

Investigadora – “que animal é que queres ser?”

Criança – “eu sou dois” (3 palavras) *pausa de 2 segundos* “e tu...” (não contabilizado) *pausa de 4 segundos* “estes são da equipa destes” (5 palavras)

Tratamento de dados

Com o intuito de analisar os dados obtidos em função dos objetivos delineados, fez-se uma análise estatística dos resultados com recurso ao software SPSS (*Statistical Package for the Social Sciences*, v. 20.0) para Machintosh.

Iniciou-se a análise de dados através da definição dos valores médios da EME-p, concordantes com a faixa etária dos 4;00 aos 4;05 anos através da estatística descritiva.

Para se proceder à análise dos outros dois objetivos estabelecidos, recorreu-se à estatística inferencial devido à distribuição normal dos resultados verificada.

Assim, nesta investigação, para se proceder à comparação entre os valores da EME-p entre géneros (objetivo 2), recorreu-se ao teste paramétrico *t-student* para amostras independentes, uma vez que a variável independente é nominal (género) e a variável dependente é intervalar (EME-p) (Maroco, 2007).

Para o último objetivo em estudo (objetivo 3), procedeu-se à utilização de estatística não paramétrica de forma a perceber se existia uma relação entre os valores de EME-p obtidos pelas crianças e o nível de escolaridade dos seus pais (escolaridade mais alta do pai ou da mãe). Para tal utilizou-se a Correlação de Spearman pois, apesar de existir normalidade nas duas variáveis, uma é métrica e a outra ordinal (Maroco & Bispo, 2003).

4. Resultados

Os resultados obtidos encontram-se apresentados de acordo com os objetivos estabelecidos e já apresentados. Para facilitar a leitura dos dados, serão apresentados quadros com os dados mais relevantes.

O primeiro objetivo deste trabalho consistia na determinação do valor padrão de EME-p para a faixa etária dos 4;00 aos 4;05 anos. Para tal, foi calculada a média dos valores de EME-p individuais de cada criança que participou no estudo (Apêndice VI). Desta forma, é possível concluir que o valor padrão da EME-p em crianças monolíngues do Português Europeu, que não apresentam perturbação de linguagem, na faixa etária em estudo é de 4,49, variando entre 3,80 e 5,18 (Quadro 5).

Quadro 5. Valor de EME-p em crianças dos 4;00 aos 4;05 anos (n=30)

	Amplitude de respostas		Média±DP
	Mínimo	Máximo	
EME-p	3,80	5,18	4,49±0,38

Procedeu-se, de seguida, à análise dos valores de EME-p em função da idade (em meses) das crianças pertencentes à faixa etária em estudo. Para tal, foi efectuada uma correlação de Spearman, tendo-se verificado que não existe correlação estatisticamente significativa entre a idade (os vários meses da faixa etária em estudo, em escala ordinal) e a EME-p ($\rho = 0,38$; $p = 0,04$).

No gráfico seguinte (Figura 2), verifica-se que a EME-p é mais elevada no subgrupo das crianças mais velhas (4;05), no entanto não há um crescendo contínuo ao longo dos meses.

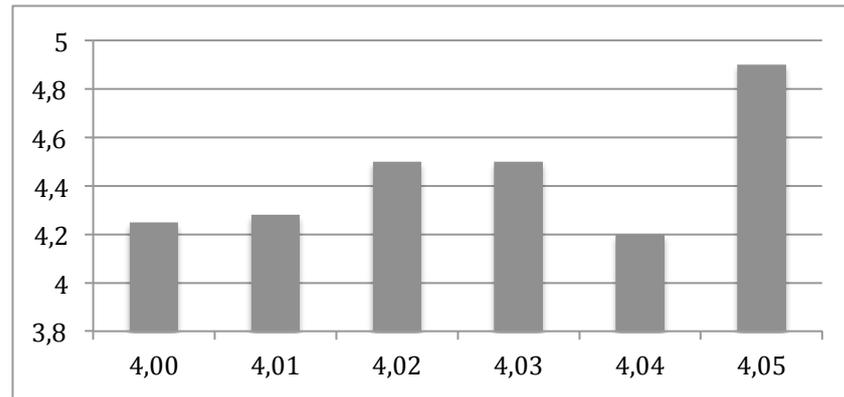


Figura 2. Média da EME-p em cada idade (em meses)

De forma a comparar os resultados de EME-p dos rapazes e das raparigas, foi necessário recorrer a um teste de normalidade para perceber se a distribuição dos resultados era normal e para determinar que tipo de teste seria mais adequado para a comparação das duas amostras (Quadro 6). A necessidade de testar a normalidade resulta da pequena dimensão das amostras.

Quadro 6. Teste de normalidade

	Género	Kolmogorov-Smirnov ^a		
		Statistic	df	Sig.
EME-p	Masculino	,117	12	,200*
	Feminino	,106	18	,200*

Verificou-se que ambos os grupos apresentam uma distribuição normal na variável dependente (EME-p). Assim, utilizou-se o teste paramétrico *t-student* para amostras independentes de forma a proceder-se à comparação dos valores de EME-p obtidos pelos rapazes e pelas raparigas. Através desta comparação, verificou-se que não existe uma diferença estatisticamente significativa (Quadro 7).

Quadro 7. Comparação dos resultados da EME-p entre rapazes e raparigas (*t-student* para amostras independentes)

Género	n	Média ± DP	t(gl)	p
Masculino	12	4,42 ± 0,44	0,86(28)	0,40
Feminino	18	4,54 ± 0,32		

No gráfico seguinte (Figura 3), é possível observar a distribuição dos valores de EME-p ao longo dos seis meses que compõe esta faixa etária, relativamente ao género das crianças. Os valores correspondem à média dos obtidos pelas crianças em cada faixa etária.

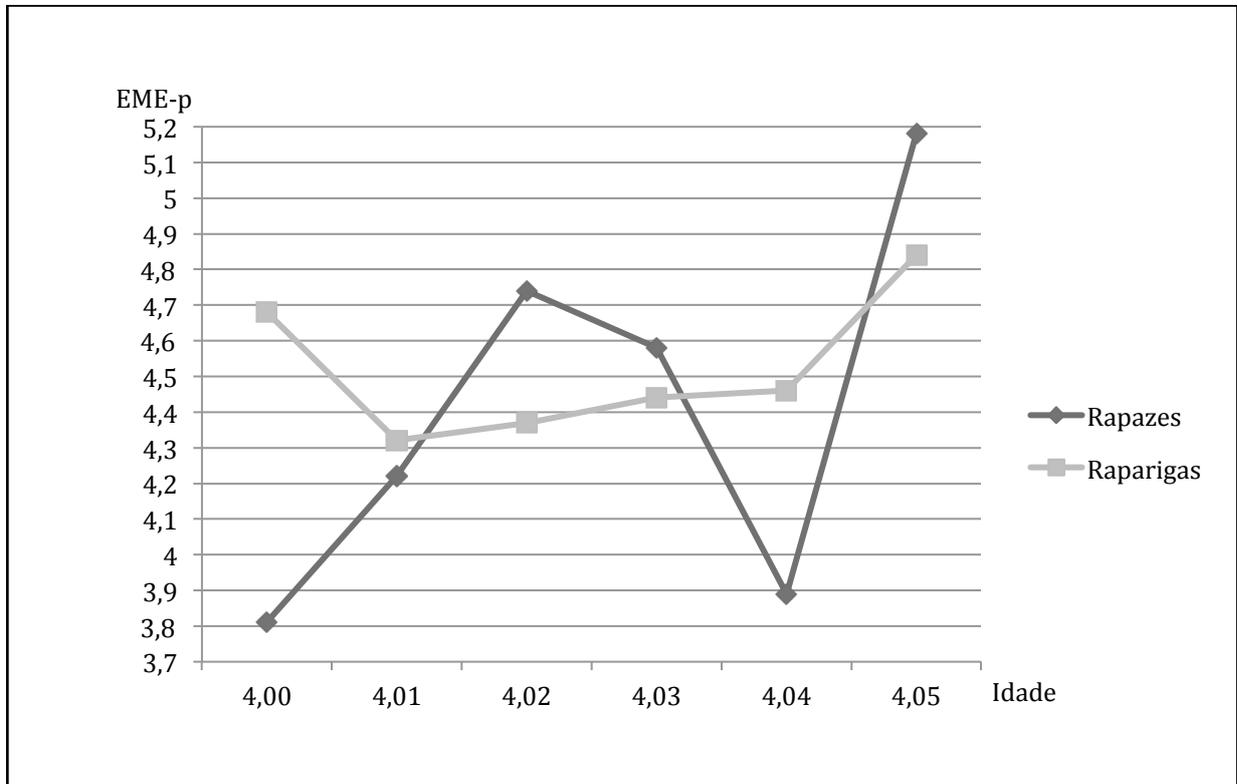


Figura 3. EME-p dos rapazes e das raparigas

Embora não haja diferenças estatisticamente significativas nas médias gerais, verifica-se uma maior variabilidade ao longo dos meses por parte dos rapazes. Os valores dos rapazes sofrem mais alterações do que os das raparigas, estes variam entre 4,32 e 4,84, enquanto que os rapazes apresentam maior variação, entre 3,81 e 5,18.

Por fim, verificou-se se existia relação entre a EME-p das crianças e o grau de escolaridade dos seus pais, ou seja, a escolaridade mais elevada do pai ou da mãe (variável ordinal). Para tal procedeu-se a uma Correlação de Spearman cujos resultados indicam que a correlação é baixa e não significativa ($\rho = 0,22$, $p = 0,24$).

5. Discussão

Neste capítulo propõe-se uma interpretação mais minuciosa dos dados apresentados no capítulo anterior. Os resultados serão analisados e interpretados e será efectuada uma reflexão sobre as suas implicações para o conhecimento acerca da linguagem infantil bem como para o processo de avaliação das crianças na faixa etária estudada.

A EME, tanto em palavras como em morfemas, é uma medida considerada como extremamente útil para a avaliação de crianças de língua Inglesa (Brown, 1973; Parker & Brorson, 2005; Rispoli, 2003). Ao longo dos anos, as suas inúmeras utilizações têm permitido alargar o conhecimento dos investigadores acerca do desenvolvimento da linguagem da criança, tendo possibilitado uma importante utilização clínica. Além da língua Inglesa, estudos realizados noutras línguas, com características muito diferentes desta, têm obtidos resultados que permitem a utilização da EME com a mesma segurança que para o Inglês, como é o caso do Irlandês (Hickey, 1991), do Cantonês (Klee *et al.*, 2004), do Italiano (Leonard & Bortolini, 1998) e do Português do Brasil (Araújo & Befi-Lopes, 2003).

O primeiro passo para o estudo e aplicação da EME, é a obtenção de valores de acordo com a idade das crianças. Este corresponde ao primeiro objectivo do presente trabalho, ou seja, a obtenção do valor padrão da EME-p em crianças de 4;00 a 4;05 anos. Desta forma, pretende-se responder à questão se a EME-p pode ser utilizada como medida fidedigna da aquisição da linguagem pela criança, para o português europeu.

De acordo com os resultados obtidos, verifica-se que o valor padrão de EME-p para a faixa etária estudada é de 4,49, variando entre 3,80 e 5,18. Esta variação demonstra que não existiu uma grande discrepância nos resultados obtidos.

Como já referido, para o PE, ainda não tinham sido calculados quaisquer valores de EME-p noutros estudos, pelo que não é possível realizar uma comparação. Apesar de se tratarem de línguas e populações com características diferentes e dos procedimentos de recolha e critérios metodológicos poderem ser também diferentes, observa-se pela análise dos estudos publicados que os valores de EME-p para línguas como o PB, para o Inglês e para o Cantonês não se encontram muito longe dos valores obtidos neste trabalho. Para o PB, Araújo e Befi-Lopes (2003), indicaram que o valor de EME-p para a faixa etária dos 4;00 aos 5;00 é de 5,53, o que é ligeiramente inferior aos valores obtidos. O mesmo foi observado no estudo com crianças falantes de cantonês, no qual o valor médio de EME-p de crianças de 48 e 49

meses (4;00 e 4;01 anos) era de 3,42 e de crianças de 52 e 53 meses (4;04 e 4;05 anos) era 4,01. Resultados um pouco mais próximos, foram obtidos por Wieczorek (2010) que, no seu estudo gravações e analisou o discurso da mesma criança com 4;01 e 4;04 anos, cuja EME-p era 4,8.

Como já referido, as línguas têm características muito diferentes mas, ao contrário da EME-m, parece ser possível realizar algumas comparações entre as várias línguas com recurso às EME-p, o que também já tinha sido referido por Gutiérrez-Clellen, *et al.* (2000).

Este estudo permitiu também observar que não parece existir uma correlação estatisticamente significativa entre a EME-p e a idade (em meses) das crianças. Apesar de se verificar que a EME-p é mais elevada no subgrupo das crianças mais velhas, de 4;05 anos, não se verificou um crescendo contínuo da EME-p ao longo dos meses. Efetivamente, o valor mais elevado de EME-p foi obtido por duas crianças de 4;05 anos mas o valor mais baixo não foi obtido, como se poderia esperar, por uma criança de 4;00 anos mas sim de 4;04. Esse crescendo contínuo poderá não ter sido alcançado por um problema de amostra pois o valor obtido por essa criança de 4;04 anos é mais baixo do que os das outras crianças do mesmo subgrupo, o que baixou de forma significativa a média deste. Por este motivo, será fundamental aumentar a amostra em estudos futuros de forma a confirmar estes resultados, garantindo também que existe um número igual de crianças em cada subgrupo.

Este desigual número de crianças por cada mês também poderá ter levado a que o valor médio de EME-p para esta faixa etária tenha sido mais elevado do que é na realidade, uma vez que participaram mais crianças nos subgrupos mais velhos de 4;03, 4;04 e 4;05 anos, enquanto que no subgrupo das crianças mais novas (4;00) participaram apenas duas crianças, uma de cada género.

O valor padrão obtido também poderá ter sido influenciado por outros factores que não foram tidos em consideração neste estudo. A frequência em jardim de infância não é ainda obrigatória em Portugal, pelo que muitas crianças ficam em casa com familiares ou empregadas até à entrada para a escola. Todas as crianças que participaram no estudo encontravam-se a frequentar alguma instituição infantil, pelo menos desde os três anos de idade. A escola, bem como os educadores de infância, são um fator importante no desenvolvimento cognitivo e linguístico das crianças, pelo que em estudos futuros seria interessante recolher e analisar amostras de discurso espontâneo de crianças que não frequentam o jardim de infância e, desta forma perceber se a sua EME-p é semelhante à das crianças que participaram neste trabalho.

Relativamente à influência do género das crianças nos valores de EME-p obtidos, não foram encontradas diferenças estatisticamente significativas entre rapazes e raparigas. Este resultado está de acordo com o estudo de Schachter (1979) que não encontrou diferenças significativas entre rapazes e raparigas ao nível da EME obtida, por crianças de dois, três e quatro anos de idade. Le Normand, Parisse e Cohen (2008), que gravaram e analisaram o discurso de 316 crianças, também não observaram diferenças entre géneros a partir dos 36 meses de idade, ou seja, dos três anos. Por seu lado, o mesmo não foi verificado por Jackson e Roberts (2001) que analisaram o discurso de 85 crianças de três e quatro anos. Neste estudo, as raparigas apresentaram valores de EME-p mais elevados do que os rapazes, produzindo formas sintáticas mais complexas. Apesar do resultado obtido, observou-se uma maior variabilidade dos valores obtidos nos seis meses estudados por parte dos rapazes. Isto é, enquanto os valores das raparigas variam pouco de mês para mês, não apresentam os mesmos picos de resultados que os rapazes.

O facto de participarem no estudo mais raparigas do que rapazes pode justificar esta discrepância pelo que este é outro motivo pelo qual a amostra deverá ser alargada em estudos futuros, garantido também que são incluídos o mesmo número de crianças de ambos os géneros.

Tal como se verificou que o género não parece influenciar a EME-p das crianças, também o nível de escolaridade dos seus pais não está correlacionada com a EME-p obtida. Estes resultados também foram obtidos por Rice *et al.* (2001) que, num grupo de 306 crianças, não encontraram uma correlação significativa entre os valores de EME-p e o nível de escolaridade das suas mães. Hoff (2003) também não encontrou relações significativas nas 63 crianças que estudou, apesar deste autor não se referir apenas ao nível de escolaridade dos pais, mas sim ao meio como fator de influencia, o que pode não estar completamente associado.

Alguns autores têm apresentado estudos em que referem que existe uma grande influência do nível educacional dos pais no discurso dos seus filhos e, conseqüentemente, na sua EME-p. Walker, *et al.* (1994) e Le Normand, *et al.* (2008), por exemplo, observaram uma forte influência desta variável, de tal forma que as crianças que provinham de níveis socioeconómicos mais altos apresentavam produções orais mais complexas e um desenvolvimento linguístico mais avançado do que crianças de níveis socioeconómicos mais baixos. Estes resultados foram obtidos em estudos com base na análise do discurso de

crianças com as mesmas idades das crianças que participaram no presente trabalho, ou seja, com quatro anos.

É importante referir que apenas uma rapariga cujo discurso foi analisado tinha uma mãe com o 4º ano de escolaridade, sendo que a maior parte das crianças tinha pais com um nível de escolaridade igual ou superior ao 12º ano.

Em suma, o valor padrão de EME-p obtido neste trabalho corresponde a um importante ponto de partida para o estudo e aplicação desta medida com crianças falantes do Português Europeu. Apesar de não se ter verificado grande dispersão dos valores de EME-p, será importante alargar a amostra de forma a ser conseguido um valor normativo seguro e que melhor represente a as possíveis diferenças entre géneros e meios culturais de origem.

É importante mencionar que nenhum dos estudos analisados e referidos neste trabalho utilizam exatamente as mesmas regras metodológicas para recolher e calcular o número de enunciados e de palavras. Esta pode ser uma das razões pelo qual se observam tantas diferenças, pois é um dos factores que mais interfere nos resultados.

6. Conclusões

Neste trabalho foram recolhidas amostras de discurso espontâneo de 30 crianças, com idades compreendidas entre os 4;00 e os 4;05 anos, que frequentam jardins de infância da área da grande Lisboa. Através dessas amostras, foram calculadas as EME-p de cada criança, podendo concluir-se que:

1. O valor padrão de EME-p obtido neste estudo para crianças na faixa etária dos 4;00 aos 4;05 anos de idade é de 4,49, com um desvio-padrão de 0,38.
2. Os valores obtidos de EME-p para esta faixa etária não parecem ser influenciados pelo género das crianças.
3. Os valores de EME-p das crianças que participaram no estudo, não mostraram influência do nível de escolaridade dos seus pais.

Após todo o trabalho ter sido concluído, é importante refletir acerca de alguns detalhes que poderiam ter corrido melhor. O facto de ter sido necessário recolher as amostras de discurso propositadamente para este trabalho levou a que muito tempo fosse despendido e assim o tamanho da amostra não ter sido mais alargado, não tanto pelo processo de avaliação e de gravação mas pelo processo de pedido e espera de autorização das escolas. Desta forma, considera-se que seria importante alargar a amostra e também ter em consideração outros fatores.

Num estudo futuro, seria importante garantir que o número de crianças de ambos os géneros é semelhante e que existe uma distribuição equilibrada em função do nível de escolaridade dos pais, de forma a permitir uma análise mais fiável. Além disso, a inclusão de crianças de outros concelhos da área da grande Lisboa como é o caso do Concelho de Oeiras também seria importante. Inicialmente tinha-se previsto que crianças deste concelho participassem no estudo, no entanto, tal não foi possível visto nenhuma escola ter autorizado a nossa presença no espaço escolar.

Seria também interessante calcular os valores relativos à EME-m e relacioná-los com os valores de EME-p obtidos neste estudo.

Apesar das limitações apontadas, considera-se que os objectivos estabelecidos neste estudo foram atingidos e espera-se que com os resultados obtidos mais estudos acerca da EME possam surgir e que outras investigações com crianças com perturbações linguísticas utilizem esta medida como forma de avaliação e de comparação.

As amostras de discurso recolhidas são uma mais valia para uma futura criação de um *corpus* de linguagem infantil, que será possível através da junção dos três trabalhos realizados neste âmbito. Este *corpus* permitirá a realização de diversas investigações que poderão auxiliar o estudo do desenvolvimento da linguagem de crianças portuguesas.

Referências Bibliográficas

- Allen, J. & Marshall, C. R. (2011). Parent-Child Interaction Therapy (PCIT) in school-aged children with specific language impairment. *International Journal of Language & Communication Disorders, 46* (4), 397-410.
- Aram, D., Morris, R. & Hall, N. (1993). Clinical and research congruence in identifying children with specific language impairment. *Journal of Speech and Hearing Research, 36* (3), 580–591.
- Araújo, K. & Befi-Lopes (2003). 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, 9* (1), 156-163.
- Arif, H. & Bol, G. W. (2008). 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, 1*, 167-182.
- Bachman, L. F. (1990). *Fundamental considerations in language testing*. Oxford: Oxford University Press.
- Balčiūnienė, I. (2012). Lithuanian narrative language at preschool age. *Estonian Papers in Applied Linguistics, 8*, 21-36.
- Befi-Lopes, D. M., Nuñez, C. O. & Cáceres, A. M. (2011). Correlação entre vocabulário expressivo e extensão média do enunciado em crianças com alteração específica de linguagem. *Revista CEFAC*, (online) ISSN 1982-0216.
- Bishop, D. V. M. & Norbury, C. F. (2002). Exploring the borderlands of autistic disorder and specific language impairment: a study using standardized diagnostic instruments. *Journal of Child Psychology and Psychiatry, 47* (3), 917-929.

- Blake, J., Quartaro, G. & Onorati, S. (1993). Evaluating quantitative measures of grammatical complexity in spontaneous speech samples. *Journal of Child Language*, 20 (1), 139-152.
- Blom, E. (2010). Effects of input on the early grammatical development of bilingual children. *International Journal of Bilingualism*, 14 (4), 422-446.
- Bloom, L. (1970). *Language Development: Form and function in emerging grammars*. Cambridge, MA: MIT Press.
- Bol, G. W. (1996). Optimal subjects in Dutch child language. In C. Koster & F. Wijnen (Eds.), *Proceeding of the Groningen assembly on Language acquisition* (pp. 125-133). Groningen: Center for Language and Cognition.
- Bornstein, M., Leach, D. & Haynes, O. M. (2004). Vocabulary competence in first and second born siblings of the same chronological age. *Journal of Child Language*, 31 (4), 855-873.
- Bornstein, M. H., Tamis-LeMonda & Haynes, M. O. (1999). First words in the second year: continuity, stability, and models of concurrent and predictive correspondence in vocabulary responsiveness across age and context. *Infant Behavior and Development*, 22 (1), 65-85.
- Braine, M. (1963), The ontogeny of English phrase structure: The first phase. *Language*, 39, 1-13.
- Brown, R. (1973). *A first language: the early stages*. Cambridge, MA: Harvard University Press.
- Cable, A. L. & Domsch, C. (2011). Systematic review of the literature on the treatment of children with late language emergence. *International Journal of Language & Communication Disorders*, 46 (2), 138-54.

- Casby, M. W. (2011). An examination of the relationship of sample size and mean length of utterance for children with developmental language impairment. *Child language Teaching and Therapy*, 27 (3), 286-293.
- Cherry, L. & Lewis, M. (1978). Differential socialization of girls and boys: implications for sex differences in language development. In N. Waterson & C. Snow (Eds.), *The development of communication* (pp. 189-197). New York: Wiley.
- Chomsky, N. (2000). *The Architecture of Language*. Oxford: Oxford University Press.
- Conant, S. (1987). The relationship between age and MLU in young children: a second look at Klee and Fitzgerald's data. *Journal of Child Language*, 14 (1), 169-173.
- Condouris, K., Meyer, E. & Tager-Flusberg, H. (2003). The relationship between standardized measures of language and measures of spontaneous speech in children with autism. *American Journal of Speech-language Pathology*, 12 (3), 349-58.
- Craig, H. K. & Washington, J. A. (2000). An assessment battery for identifying language impairments in African American Children. *Journal of Speech, Language, and Hearing Research*, 43 (2), 366-379.
- Crystal, D., Fletcher, P. & Garman M. (1976). *The Grammatical Analysis of Language Disability*. London: Arnold.
- Davies, A. (2007). *An introduction to Applied Linguistics – from practice to theory* (2nd edition). Edinburgh: Edinburgh textbooks in applied linguistics.
- De Villiers, J. G. & De Villiers, P. A. (1973). Development of the use of word order in comprehension. *Journal of Psycholinguistic Research*, 2 (4), 331-341.
- Diessel, H (2009). Corpus linguistics and language acquisition. In A. Ludeling & M. Kyto (Eds.), *Corpus Linguistics: An International Handbook* (pp. 1197-1212). Berlin: Mouton de Gruyter.

- Eisenbeiss, S (2010). Production methods in language acquisition research. In E. Blom & S. Unsworth (Eds.), *Experimental Methods in Language Acquisition Research* (pp. 11-34). Amsterdam: John Benjamins Publishing Company.
- Eisenberg, S. L., Fersko, T. M. & Undgreen, C. (2001). The use of MLU for identifying impairment in preschool children: a review. *American Journal of Speech-Language Pathology*, 10 (4), 323-342.
- Feldman, H. M., Janosky, J. E., Scher, M. S. & Wareham, N. L. (1994). Language abilities following prematurity, periventricular brain injury, and cerebral palsy. *Journal of Communication Disorders*, 27 (2), 71-90.
- Foudon, N., Reboul, A. & Manificat, S. (2007). Language acquisition in autistic children: a longitudinal study. In N. Hilton, R. Arscott, K. Barden, A. Krishna, S. Shah & M. Zellers (Eds.), *CamLing 2007: Proceedings of the Fifth University of Cambridge Postgraduate Conference in Language Research* (pp. 72-79). Cambridge: Cambridge Institute of Language Research.
- Gries, S. T. & Stoll, S. (2009). Finding developmental groups in acquisition data: variability-based neighbor clustering. *Journal of Quantitative Linguistics*, 16 (3), 217-242.
- Grunau, R. V. E., Kearney, S. M. & Whitfiel, M. F. (1990). Language development at 3 years in pre-term children of birth weight below 1000g. *International Journal of Language & Communication Disorders*, 25 (2), 173-182.
- Gutiérrez-Clellen, V. F., Restrepo, M. A., Peña, L. B. E. & Anderson, R. (2000). Language sample analysis in Spanish speaking children: methodological consideration. *Journal of Language, Speech, and Hearing Services in Schools*, 31 (1), 88-98.
- Hart, B. & Risley, T. R. (1995). *Meaningful differences in the everyday experience of young American children*. Baltimore US: Paul H Brooks Publishing.

- Hewitt, L. E., Hammer, C. S., Yont, K. M. & Tomblin, J. B. (2005). Language sampling for kindergarten children with and without SLI: mean length of utterance, IPSYN, and NDW. *Journal of Communication Disorders*, 38(3), 197-213.
- Hickey, T. (1991). Mean Length of Utterance and the acquisition of Irish. *Journal of Child Language*, 18 (3), 553-569.
- Hoff, E. (2003). The specificity of environmental influence: socioeconomic status affects early vocabulary development via maternal speech. *Child Development*, 74 (5), 1368-1378.
- Jackson, S. C., & Roberts, J. E. (2001). Complex syntax production of African American preschoolers. *Journal of Speech, Language, and Hearing Research*, 44 (5), 1083-1096.
- Kern, S. & Frédérique, G. (2007). Influence of preterm birth on lexical and grammatical acquisition. *First language*, 27 (2), 159-173.
- Klee, T. (1992). Developmental and diagnostic characteristics of quantitative measures of children's language production. *Topics in Language Disorders*, 12 (2), 28-41.
- Klee, T. & Fitzgerald, M. (1985). The relation between grammatical development and mean length of utterance in morphemes. *Journal of Child Language*, 12 (2), 251-69.
- Klee, T., Schaffer, M., May, S., Membrino, I. & Mougey, K. (1989). A comparison of the age-MLU relation in normal and specifically language-impaired preschool children. *Journal of Speech and Hearing Disorders*, 54 (2), 226-33.
- Klee, T., Stokes, S. F., Wong, A. M. Y., Fletcher, P. & Gavin, W. J. (2004). Utterance length and lexical diversity in Cantonese-speaking children with and without Specific Language Impairment. *Journal of Speech, Language and Hearing Research*, 47 (6), 1396-1410.
- Le Normand, M.-T., Parisse, C. & Cohen, H. (2008). Lexical diversity and productivity in French preschoolers. *Clinical Linguistics and Phonetics*, 22 (1), 47- 58.

- Leonard, L. B. & Bortolini, U. (1998). Grammatical morphology and the role of weak syllables in the speech of Italian-speaking children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 41* (6), 1363-1374.
- Loeb, D. F., Kinsler, K. & Bookbinder, L. (2000). *Current language sampling practices in preschools*. Annual Convention of the American Speech-Language-Hearing Association, Washington D.C.
- Lopes-Herrera, S. A. & Almeida, M. A. (2008). O uso de habilidades comunicativas verbais para aumento da extensão de enunciados no autismo de alto funcionamento e na síndrome de Asperger. *Pró-Fono Revista de Atualização Científica, 20* (1), 37-42.
- Lund, N. J. & Duchan, J. F. (1993). *Assessing children's language in naturalistic contexts*. New Jersey: Prentice Hall.
- Marques, S. F., Limongi, S. C. O. (2011). A extensão média do enunciado (EME) como medida do desenvolvimento de linguagem de crianças com síndrome de Down. *Jornal da Sociedade Brasileira de Fonoaudiologia, 23* (2), 152-157.
- McCarthy, P. M. & Jarvis, S. (2011). MTL, vocd-D, and HD-D: a validation study of sophisticated approaches to lexical diversity assessment. *Behavior Research Methods, 42* (2), 381-392.
- McDaniel, D., McKee, C. & Cairns, H. S. (1996). *Methods for assessing children's syntax*. New York: The MIT Press.
- MacWhinney, B. (2001). *The CHILDES Project – tools for analyzing talk – Electronic Edition* (3ª ed.). Mahwah, NJ: Lawrence Erlbaum Associates.
- MacWhinney, B. & Snow, C. (1985). The child language data exchange system. *Journal of Child Language, 12* (2), 271-296.

- Malvern, D. D. & Richards, B. J. (2002). Investigating accommodation in language proficiency interviews using a new measure of lexical diversity. *Language Testing*, 19 (1), 85-104.
- Maroco, J (2007). *Análise estatística com utilização do SPSS (3ª edição)*. Lisboa: Sílabo.
- Maroco, J. & Bispo, R. (2003). *Estatística aplicada às ciências sociais e humanas (2ª edição)*. Lisboa: Climpesi Editores.
- McLean, J. E & Snyder-McLean, L. K. (1978). *A transactional approach to early language training*. Columbus, Ohio: Charles E. Merrill.
- Mercy, J. A. & Steelman, L. C. (1982). Familial influence on the intellectual attainment of children. *American Sociological Review*, 47 (4), 532-542.
- Miller, J. F. (1981). Eliciting procedures for language. In J. F. Miller (Ed.), *Assessing language production in children*. London: Edward Arnold.
- Miller, J. F. (1991). Quantifying productive language disorders. In J. F. Miller (Ed.), *Research on child language disorders: a decade of progress* (pp. 211-220). Austin, Texas: Pro-Ed.
- Miller, J. F. & Chapman, R. S. (1981). The relation between age and Mean Length of Utterance in morphemes. *Journal of Speech and Hearing Research*, 24 (2), 154-161.
- Miller, J. F. & Chapman, R. S. (1991). *Systematic Analysis of Language Transcripts* [Computer software]. Madison: University of Wisconsin
- Moyle, M. J., Karasinski, C., Weismer, S. E. & Gorman, B. K. (2011). Grammatical morphology in school-age children with and without language impairment: a discriminant function analysis. *Language, Speech, and Hearing Services in Schools*, 42 (4), 550-60.
- Nice, M. M. (1925). Length of sentences as a criterion of a child's progress in speech. *Journal of Education Psychology*, 16(6), 370-379.

- Nieminen, L. (2007). MLU and IPSYN measuring absolute complexity. *Estonian Papers in Applied Linguistics*, 5, 173-185.
- Oetting, J. B. (2005). Assessing language in children who speak a nonmainstream dialect of English. In M. J. Ball (Ed.), *Clinical Sociolinguistics* (pp. 181-192). Malden, MA: Blackwell Publishing.
- Ooi, C. C-W & Wong, A. M-Y. (2012). Assessing bilingual Chinese-English young children in Malaysia using language sample measures. *International Journal of Speech-Language Pathology*, 14 (6), 499-508.
- Oosthuizen, H. & Southwood, F. (2009). Methodological issues in the calculation of mean length of utterance. *South African Journal of Communication Disorders*, 56 (1), 76-87.
- Owens, R. (1999). *Language disorders: a functional approach to assessment and intervention* (3rd ed.). Boston, MA: Allyn and Bacon.
- Ozcaliskan, S. & Goldin-Meadow, S. (2010). Sex differences in language first appear in gesture. *Developmental Science*, 13 (5), 752-760.
- Pan, B., Rowe, M. L., Spier, E. & Tamis-Lemonda, C. (2004). Measuring productive vocabulary of toddlers in low-income families: concurrent and predictive validity of three sources of data. *Journal of Child Language*, 31 (3), 587-608.
- Parker, M. D. & Brorson, K. (2005). A comparative study between Mean Length of Utterance in morphemes (MLUm) and Mean Length of Utterance in words (MLUw). *First Language*, 25 (3), 365-376.
- Paul, R. (2007). *Language disorders from infancy through adolescence: assessment and intervention* (3rd edition). St Louis, MO: Mosby.
- Peters, A. (1999). From Info-Childes. *IASCL, Child Language Bulletin*, 19 (1).

Portugal, Ministério da Saúde, Decreto-Lei n.º 261/93, publicado em Diário da República n.º172, | Série I-A, de 24 de Julho

Prieto, P., Estrella, A., Thorson, J. & Vanrell, M. (2001). Is prosodic development correlated with grammatical and lexical development? Evidence from emerging intonation in Catalan and Spanish. *Journal of Child Language*, 28 (1), 1-37.

Ramer, A. L. H. (1976). Syntactic styles in emerging language. *Journal of Child Language*, 3 (1), 49–62.

Redington, M. & Chater, N. (1998). Connectionist and statistical approaches to language acquisition: A distributional analysis. *Language and Cognitive Processes*, 13 (2), 129-191.

Retherford, K. S. (2000). *Guide to analysis of language transcripts* (3rd edition). Eau Claire, WI: Thinking Publications.

Rescorla, L., Dahlsgaard, K. & Roberts, J. (2000). Late-talking toddlers: MLU and IPSyn outcomes at 3:0 and 4:0. *Journal of Child Language*, 27 (3), 643-664.

Rice, M. L., Redmond, S. M. & Hoffman, L. (2006). 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*, 49 (4), 793-808.

Rice, M. L., Smolik, F., Perpich, D., Thompson, T., Rytting, N. & Blossom, M. (2010). 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*, 53 (2), 333-349.

Rispoli, M. (2003). Changes in the nature of sentence production during the period of grammatical development. *Journal of Speech, Language, and Hearing Research*, 46 (4), 818-830.

- Rodal, J.-A. (2000). Análisis del lenguaje espontáneo. In M. P. Sancllemente, J.-A. Rodal & E. H. Wig (Eds.), *Evaluación del lenguaje* (pp. 131-171). Barcelona: Masson.
- Rondal, J. A., Ghiotto M., Bredart S. & Bachelet J. F. (1987). Age-relation, reliability and grammatical validity of measures of utterance length. *Journal of Child Language*, 14 (3), 433-446.
- Roy, D. (2009). New horizons in the study of child language acquisition. *INTERSPEECH-2009*, 13-20.
- Scarborough, H. S. (1990). The Index of Productive Syntax. *Applied Psycholinguistics*, 11 (1), 1-22.
- Scarborough, H. S., Rescorla, L., Tager-Flusberg, H., Fowler, A. E. & Sudhalter, V. (1991). The relation of utterance length to grammatical complexity in normal and language-disordered groups. *Applied Psycholinguistics*, 12 (1), 23-45.
- Scarborough, H. S., Wyckoff, J. & Davidson, R. (1986). A reconsideration of the relation between age and Mean Utterance Length. *Journal of Speech, Language, and Hearing Research*, 29 (3), 394-399.
- Schachter, F. F. (1979). *Everyday mother talk to toddlers: early intervention*. San Diego, CA: Academic Press.
- Seidman, S., Allen, R. & Wasserman, G. A. (1986). Productive language of premature and physically handicapped two-year-olds. *Journal of Communication Disorders*, 19 (1), 49-61.
- Shipley, K. G. & McAfee, J. G. (2009). *Assessment in Speech-Language Pathology: a resource manual* (4th edition). Delmar, NY: Cengage Learning.
- Stalnaker, L. & Craighead, N. (1982). An examination of language samples obtained under three experimental conditions. *Language, Speech, and Hearing Services in Schools*, 13 (2), 121-128.

- Stern, C. & Stern, W. (1907). *Recollection, testimony, and lying in early childhood (Erinnerung, aussage, und lueg in der ersten kindheit)*, traduzido por Lamiell, J. T, 1999. Washington DC: American Psychological Association.
- Stockman, I. (1996). The promises and pitfalls of language sample analysis as an assessment tool for linguistic minority children. *Language, Speech, and Hearing Services in Schools*, 27 (4), 355–366.
- Tager-Flushber, H., Paul, R. & Lord, C. (2005). Language and communication in autism. In F. R. Volkmar, R. P., A. Klin & D. J. Cohen (Eds.), *Handbook of autism and pervasive developmental disorders: assessment, intervention and policy* (vol. 2, pp. 335-364). New Jersey: Wiley
- Tallal, P (1988). Developmental language disorders: part I – definition. *Human Communication Canada/Communication Humaine Canada*, 21 (2), 7-22.
- Templin, M. C. (1957). *Certain language skills in children: their development and interrelationships*. Westport, CT: Greenwood.
- Tomblin, J. B. (1990). The effect of birth order on the occurrence of developmental language impairment. *The British Journal of Disorders of Communication*, 25 (1), 77-84.
- Trask, R. L. (2007). *Language and Linguistics: the key concepts* (2nd edition). London: Routledge.
- van Balkom, H., Verhoeven, L., van Weerdenburg, M. & Stoep, J. (2010). Effects of parent-based video home training in children with developmental language delay. *Child Language Teaching and Therapy*, 26 (3), 221-237.
- van Bysterveldt, A. K., Westerveld, M. F., Gillon, G. & Foster-Cohen, S. (2012). Personal narrative skills of school-aged children with Down syndrome. *International Journal of Language & Communication Disorders*, 47 (1), 95-105.

Walker, D., Greenwood, C., Hart, B. & Carta, J. (1994). Prediction of school outcomes based on early language production and socioeconomic. *Children and Poverty*, 65 (2), 606-621.

Wells, G. (1985). *Language development in the pre-school years*. Cambridge, England: Cambridge University Press.

Wieczorek, R. (2010). Using MLU to study early language development in English. *Psychology of Language and Communication*, 14 (2), 59-69.

Yip, V. & Matthews, S. (2006). Assessing language dominance bilingual acquisition: a case for Mean Length Utterance differentials. *Language Assessment Quarterly*, 3 (2), 97-116.

Apêndices

Apêndice I

Carta de pedido de autorização à direção dos Agrupamentos



Lisboa, ___/___/_____

Exmo.(a) Diretor(a),

Somos três alunas do Mestrado em Linguística Clínica da Universidade Católica Portuguesa. Vimos, por este meio, solicitar a sua colaboração num projeto que estamos a realizar sobre o desenvolvimento da linguagem de crianças de quatro e cinco anos, de forma a podermos obter valores padrão da Extensão Média do Enunciado para a língua portuguesa. Pretendemos gravar o discurso de cada criança através de gravação áudio, durante um período de cerca de 30 minutos de conversação e jogos. Para o efeito, agradecemos que autorize a participação dos seus alunos neste estudo.

Desde já manifestamos a nossa disponibilidade para responder a qualquer questão sobre esta investigação.

Contactaremos em breve no sentido de obtermos resposta ao nosso pedido.

Contatos:

Mariana Cacela	96.....	mariana.cacela@gmail.com
Sara Carvalho	96.....	sarafcarvalho@yahoo.com
Sofia Lynce de Faria	96.....	sofia.lyncedefaria@gmail.com

Gratas pela atenção dispensada.

Cordialmente,

(Mariana Cacela)

(Sara Carvalho)

(Sofia Lynce de Faria)

Apêndice II

Carta aos encarregados de educação e consentimento informado



Lisboa, ___/___/_____

Exmo. encarregado de educação,

Somos três alunas do Mestrado em Linguística Clínica da Universidade Católica Portuguesa. Vimos, por este meio, solicitar a sua colaboração num projeto que estamos a realizar sobre o desenvolvimento da linguagem de crianças de quatro e cinco anos, de forma a podermos obter valores padrão da Extensão Média do Enunciado para a língua portuguesa. Pretendemos gravar o discurso de cada criança através de gravação áudio, durante um período de cerca de 30 minutos de conversação e jogos. Para o efeito, agradecemos que autorize o seu filho/a a participar neste estudo.

Desde já manifestamos a nossa disponibilidade para responder a qualquer questão sobre esta investigação.

Contatos:

Mariana Cacula	96.....	mariana.cacula@gmail.com
Sara Carvalho	96.....	sarafcarvalho@yahoo.com
Sofia Lynce de Faria	96.....	sofia.lyncedefaria@gmail.com

Gratas pela atenção dispensada.

Cordialmente,

(Mariana Cacula) (Sara Carvalho) (Sofia Lynce de Faria)

Caso considere que deve autorizar, por favor preencha o espaço abaixo.

Muito Obrigada

✂-----

Eu, _____, encarregado de educação de _____, declaro que autorizo a sua participação no estudo acima referido e que os dados linguísticos a recolher poderão ser utilizados para fins académicos ou científicos.

Assinatura: _____

Data: ___/___/_____

Apêndice III

Folha de registo da criança

FOLHA DE REGISTO

NOME: _____ CÓDIGO: _____

DATA DE NASCIMENTO: ___/___/___ IDADE: _____

JARDIM DE INFÂNCIA: _____ CONSELHO: _____

ESCOLARIDADE: PAI _____ MÃE: _____

DATA DA AVALIAÇÃO: ___/___/___ DATA GRAVAÇÃO: ___/___/___

INFORMAÇÃO

	Sim	Não
Idade entre 4;00 e 4;05		
Língua materna portuguesa		
Sem Perturbação de desenvolvimento		
Sem défice sensoriais		
Sem Perturbação da linguagem		
Discurso inteligível		

RESULTADO NO TALC

- Compreensão: _____
- Expressão: _____

Apêndice IV

Modelo de relatório de avaliação para inclusão no estudo

Relatório de Avaliação da Linguagem

Nome: _____

Data de Nascimento: ____/____/____

Idade: _____ anos e _____ meses

Data de avaliação: ____/____/____

O/A menino/a _____ realizou uma avaliação da linguagem, através do Teste de Avaliação da Linguagem na Criança (TALC), no âmbito do trabalho de investigação realizado no Mestrado em Linguística Clínica, do Instituto de Ciências da Saúde da Universidade Católica Portuguesa.

Através da análise dos resultados obtidos, conclui-se que o/a _____ obteve uma cotação dentro dos valores esperados para a sua faixa etária, tanto ao nível da compreensão como da expressão.

Lisboa, ____ de Maio de 2012

Pel’A equipa de investigação

_____, Terapeuta da Fala

Apêndice V

Modelo de relatório de avaliação para exclusão do estudo

Relatório de Avaliação da Linguagem

Nome: _____
Data de Nascimento: ____/____/____
Idade: _____ anos e _____ meses
Data de avaliação: ____/____/____

O/A menino/a _____ realizou uma avaliação de linguagem, através do Teste de Avaliação da Linguagem na Criança (TALC) no âmbito do trabalho de investigação realizado no Mestrado em Linguística Clínica do Instituto de Ciências da Saúde da Universidade Católica Portuguesa.

Da análise efectuada, verificou-se que em algumas provas de linguagem, a _____ apresentou valores inferiores aos esperados para a idade. Recomenda-se que, se possível, seja realizada uma avaliação mais aprofundada por um terapeuta da fala.

Lisboa, ____ de Maio de 2012
Pel’A equipa de investigação,

_____, Terapeuta da Fala

Apêndice VI

Valores de EME-p obtidos por cada criança

Código Participante	Género	Idade (meses)	Escolaridade mais elevada do pai ou da mãe	EME-p
1M	F	48	Ensino Superior	4,68
2M	M	48	9 anos	3,81
3M	F	49	Ensino Superior	4,36
4M	F	49	Ensino Superior	4,16
5M	F	49	Ensino Superior	4,44
6M	M	49	Ensino Superior	4,31
7M	M	49	12 anos	4,12
8M	M	50	Ensino Superior	4,82
9M	F	50	9 anos	4,28
10M	M	50	Ensino Superior	4,66
11M	F	50	12 anos	4,33
12M	F	50	12 anos	4,50
13M	M	51	12 anos	4,42
14M	F	51	6 anos	4,06
15M	F	51	9 anos	4,54
16M	M	51	12 anos	4,99
17M	M	51	9 anos	4,42
18M	M	51	Ensino Superior	4,47
19M	F	51	6 anos	4,71
20M	M	52	12 anos	3,80
21M	M	52	Ensino Superior	3,98
22M	F	52	Ensino Superior	4,81
23M	F	52	4 anos	4,13
24M	F	52	9 anos	4,46
25M	F	53	9 anos	4,90
26M	F	53	12 anos	4,22
27M	F	53	12 anos	5,04
28M	F	53	Ensino Superior	4,84
29M	F	53	9 anos	5,18
30M	M	53	Ensino Superior	5,18