

# Diversidade e densidade lexical em textos escritos por alunos recém-alfabetizados: um estudo descritivo de produções individuais e em díades<sup>1</sup>

## Lexical diversity and density in texts written by newly literate students: A descriptive study of individual and dyadic productions

Elian da Silva Santos<sup>2</sup>

elian.santos@cedu.ufal.br

Universidade Federal de Alagoas

Eduardo Calil<sup>2</sup>

eduardocalil@me.com

Universidade Federal de Alagoas

Luísa Álvares Pereira<sup>3</sup>

lpereira@ua.pt

Universidade de Aveiro

Rosa Lídia Coimbra<sup>4</sup>

ricoimbra@ua.pt

Universidade de Aveiro

**RESUMO** – A Diversidade Lexical (DiL) e a Densidade Lexical (DeL) podem caracterizar textos pertencentes a um determinado gênero, relacionando-os com sua riqueza lexical e com a quantidade de informações que contém. Este estudo pretende analisar a presença desses dois aspectos em textos produzidos por alunos portugueses de 7 anos de idade na mesma sala de aula (2º ano), estudando textos escritos individualmente e em díades. A partir de oito propostas de produção textual, nas quais o professor pediu a 12 alunos que inventassem histórias, constituímos um *corpus* com 24 produções individuais e 36 produções em duplas. A unidade de medida adotada para DiL foi a TTR. Para calcular a DeL, foi dividido o número total de itens lexicais pelo total de palavras escritas em cada texto produzido. Os resultados obtidos mostraram que a DeL entre os textos individuais e os textos em dupla é estável, sem variação significativa. No entanto, para DiL, esses textos mostram uma maior variação, e uma relação inversamente proporcional. Comparando apenas as produções individuais, verificou-se uma tendência para o aumento da DeL, especialmente nas produções que tiveram baixa pontuação na 1ª proposta. Foi perceptível uma tendência de aumento da quantidade de palavras entre as duas produções individuais dos alunos. Apesar de esse

**ABSTRACT** – Lexical Diversity (LDi) and Lexical Density (LDe) may characterize texts belonging to a certain genre, relating them to its lexical richness and to the amount of information it contains. This study aims to analyze the presence of these two aspects in texts produced by Portuguese 7-year-old students in the same classroom (2<sup>nd</sup> year), comparing individually written texts and dyad ones. From eight proposals of textual production in which the teacher asked 12 students to invent stories, we constituted a corpus including 24 individual productions and 36 dyadic productions. The unit of measure adopted for LDi was TTR. We calculated the LDe by dividing the total number of lexical items by the total number of words in each text produced. Our results show that the LDe between individual and dyadic texts is stable, with no significant variation among them. However, for LDi, these texts show a greater variation, and an inversely proportional relationship. Comparing the individual productions, we found a tendency to LDe increase, especially in those productions which had a low score in the 1st proposal. Regarding the amount of words per text, there was an increase between the two texts of the same students. In spite of the fact that this suggests that

<sup>1</sup> Este trabalho é financiado pelo CNPq (processo 304050/2015-6), FAPEAL (processo 60030 479/2016) e Fundos Nacionais através da Fundação para a Ciência e a Tecnologia, I.P., no âmbito do projeto UID/CED/00194/2013.

<sup>2</sup> Universidade Federal de Alagoas. Centro de Educação. Laboratório do Manuscrito Escolar (LAME). Av. Lourival Melo Mota, s/n, Tabuleiro dos Martins, 57072-900, Maceió, AL, Brasil.

<sup>3</sup> Universidade de Aveiro. CIDTFF/Protexos. Departamento de Educação e Psicologia. 3810-193, Aveiro, Portugal.

<sup>4</sup> Universidade de Aveiro. Departamento de Línguas e Culturas/Protexos. Departamento de Educação e Psicologia. 3810-193, Aveiro, Portugal.

fato sugerir que os alunos começaram a escrever textos mais longos, a tendência decrescente de DiL permaneceu ao longo das produções.

**Palavras-chave:** texto escolar, escrita colaborativa, densidade lexical, diversidade lexical.

the students started to write longer texts, the decreasing tendency of LDi remained.

**Keywords:** school text, collaborative writing, lexical density, lexical diversity.

## Introdução

A Diversidade Lexical (DiL) e a Densidade Lexical (DeL) são aspectos que podem caracterizar textos pertencentes a um determinado gênero, relacionando-os com sua riqueza lexical e com a quantidade de informações que contêm (Johansson, 2008; Read, 2000). A investigação, neste domínio, procurou, sobretudo, caracterizar estas duas dimensões em textos escritos individualmente.

Este estudo pretende analisar a presença destes dois aspectos em textos produzidos por alunos portugueses na mesma sala de aula escrevendo o mesmo gênero textual. Portanto, tentamos responder às seguintes perguntas: como esses aspectos lexicais se manifestam em textos escritos individualmente e em pares? A escrita colaborativa favorece o desenvolvimento desses aspectos?

É nossa intenção contribuir para a produção de conhecimento nesta área, numa perspectiva de trabalho colaborativo e com os alunos recém-alfabetizados. A opção por este foco está relacionada ao fato de que há poucos estudos sobre a DiL e a DeL em textos de alunos desta faixa etária e também à escassez de estudos que investigam produções de narrativas ficcionais escritas individual e colaborativamente. Por outro lado, é importante ressaltar a relevância da análise de um conjunto de textos dentro de um mesmo gênero textual, escritos pelos mesmos alunos, a partir da mesma consigna, sob as mesmas condições didáticas.

## Investigação: antecedentes e conceitos

As palavras são os blocos de construção da comunicação humana e permitem que o falante aprenda e expresse conceitos, ideias e conhecimento. Consequentemente, podemos considerar que a riqueza vocabular seja um dos fatores centrais para a qualidade textual. Esta hipótese explica a forte correlação entre o sucesso escolar e o repertório lexical do aluno (Duarte *et al.*, 2011). O desenvolvimento lexical é, portanto, um tema central nos estudos sobre aquisição de linguagem e sobre didática.

Um indicador fundamental usado para descrever o desenvolvimento lexical é o grau de Riqueza Lexical (RL) das produções do aluno. Entretanto, encontrar uma maneira objetiva para medir a RL é uma tarefa complexa: “a revisão da literatura sobre a quantificação da riqueza do vocabulário dá a sensação de uma busca pelo Santo

Graal - a busca por uma medida única da qualidade do vocabulário” (Malvern *et al.*, 2004, p. 3).

Em um estudo sobre a escrita dos alunos na aquisição da segunda língua, Read (2000) identificou alguns componentes que constituem a riqueza lexical: diversidade lexical, densidade lexical, sofisticação lexical e número de erros.

A variação lexical ou diversidade lexical (DiL) é um indicador linguístico associado à quantificação da variedade de palavras presentes em diferentes materiais textuais (Malvern *et al.*, 2004), ou seja, está relacionada à extensão do vocabulário que o escritor conhece e usa ao produzir um texto. Quanto mais lexemas diferentes forem usados, menos repetições serão encontradas. Um texto com alta variação lexical apresenta um leque considerável de vocabulário, e a repetição é evitada usando lexemas semanticamente relacionados como sinônimos, hiperônimos e hipônimos.

A DiL tem sido investigada em estudos sobre a aquisição de língua materna (Le Normand *et al.*, 2008; Scherer *et al.*, 2002; Scherer e Souza, 2011), análises contrastivas entre texto oral e texto escrito (Rodrigues, 2008), ou ainda em estudos sobre textos escolares (Martins, 2016; Johansson, 2008).

Tão variadas quanto as aplicações da diversidade lexical são as fórmulas para a sua aferição. Apesar da literatura especializada reconhecer a importância da medição e propor metodologias para a obtenção da mesma, não é adotado um método unificado, levando a variações de acordo com a natureza das amostras. O método clássico para medir esse indicador foi desenvolvido por Templin (1957), consiste na razão entre a frequência de palavras diferentes (types) e a frequência do total de palavras (tokens); popularmente conhecida com TRR (*Type-Token Ratio*). O resultado da medição da DiL através da TTR pode ser expresso no intervalo de 0 e 1, ou em porcentagens. Isso significa que quanto maior o valor da TTR, mais diversificado será o vocabulário do escrevente (DeBoer, 2014).

Como dissemos, outro indicador da RL é a densidade lexical (DeL), considerada como um dos principais indicadores da trajetória do desenvolvimento rumo a uma escrita mais rica (Colombi, 2000). Segundo Johansson (2009), este indicador representa a concentração de itens lexicais presentes no texto. A extração da DeL pode ser realizada através de alguns métodos, como por exemplo o método de Ure (1971) calculado através da divisão dos

itens lexicais<sup>5</sup> (substantivos, adjetivos e verbos) pelo total de palavras escritas no texto.

A sofisticação lexical (SL), outro indicador de RL, dependeria do uso de palavras de baixa frequência; incluindo termos técnicos ou vocabulário incomum. Um dos métodos usados para calcular o nível de sofisticação lexical é através da divisão do número de palavras sofisticadas presentes no texto pelo número total de classes de palavras do texto (Read, 2000).

Dentre estes indicadores, a DeL e a DiL têm sido usadas para análise de textos escolares em alguns trabalhos em língua portuguesa (Martins, 2016; Rodrigues, 2008). Martins, por exemplo, em um estudo sobre correlação entre diversidade lexical e progressão escolar, analisou produções textuais de alunos portugueses do quinto, sétimo e décimo ano do ensino básico. O autor constatou a existência de uma correlação positiva entre a DiL e a progressão escolar dos alunos do quinto ao décimo ano, ou seja, quanto mais se avança nos anos de escolaridade, ciclo a ciclo, mais variado se torna o vocabulário dos alunos.

Rodrigues (2008), em um estudo sobre a capacidade da composição da escrita, em termos de produtividade e DiL, analisou textos de alunos do 1º ao 3º ano do ensino básico e identificou, através dos resultados, que o

desempenho dos alunos aumentou progressivamente ao decorrer dos anos. Também foi observado que os textos produzidos pelas alunas possuíam uma DiL maior que os textos produzidos pelos alunos.

Apesar da importância desses indicadores em estudos vinculados ao desenvolvimento linguístico, observa-se que existem poucas investigações que tomam como objeto de análise os textos escolares produzidos por alunos dos anos iniciais do Ensino Fundamental.

## Metodologia

Os textos coletados para o presente estudo foram produzidos por alunos de uma única sala de aula, em uma escola portuguesa de Ensino Fundamental, situada na zona rural de Aveiro. A classe era composta por 21 alunos, com faixa etária de 7 anos de idade, cursando o 2º ano. As propostas de produção textual fizeram parte de um projeto didático de escrita criativa, em que os alunos eram solicitados a escreverem histórias inventadas. Uma vez por semana, durante oito semanas (janeiro a março de 2015), os alunos foram convidados a escreverem histórias inventadas. A 1ª e a 8ª propostas foram escritas individualmente. Estas duas propostas individuais foram



**Figura 1.** Aba do programa Lexicanalytics indicando a TTR do Td\_D1.

**Figure 1.** Tab of Lexicanalytics software showing Td\_D1 TTR value.

<sup>5</sup> Definido por Ure (1971) e Halliday (1985) como palavras de conteúdo, ou seja, palavras que trazem informações, ao invés de desempenhar apenas funções de ligações gramaticais.

consideradas por nós como um pré-teste e um pós-teste, respectivamente. As outras seis propostas (2ª, 3ª, 4ª, 5ª, 6ª, 7ª) foram escritas em duplas, sempre formadas pelos mesmos alunos, que deveriam escrever juntos, um único texto.

Como o projeto didático envolveu todos os alunos da sala, tivemos um agrupamento formado por dois alunos. Além disso, alguns alunos faltaram em algumas propostas. Para este estudo, um dos critérios para a seleção do *corpus* foi o conjunto de oito produções escritas pelas mesmas duplas. Por estas razões, foram excluídas quatro díades (D4, D6, D8 e D10), fazendo parte do *corpus* os textos produzidos por seis díades (D1, D2, D3, D5, D7 e D9), totalizando 60 textos: 24 textos individuais (Ti) e 36 textos escritos em dupla (Td). Vale ressaltar que todos os conjuntos de textos foram produzidos sob as mesmas condições de sala de aula, as propostas foram feitas pela mesma professora para os mesmos alunos, envolvendo o mesmo gênero de texto, a narrativa ficcional. Essas condições didáticas favoreceram tanto a análise da diversidade lexical e da densidade lexical, bem como das características dessas medidas em produções individuais e em díades.

Os problemas relacionados à ortografia e aos limites das palavras foram corrigidos. Essas correções foram necessárias para facilitar no processo do levantamento quantitativo das palavras escritas nos textos. O total de palavras foi contado a partir do título, ou, nos casos em que a história não apresentasse título, consideramos a primeira frase escrita até a última palavra (mesmo quando essa palavra era a palavra “fim”).

A DiL foi medida através da TTR, para isso as palavras diferentes (*types*) foram contadas considerando apenas uma vez a ocorrência de cada palavra. Isto é, “rei” e “reis” ou “o” e “os” foram consideradas, respectivamente, como ocorrências duas palavras diferentes, mas as ocorrências de uma mesma palavra gráfica, como “rei” ou “menino”, por mais de uma vez no texto, foram contadas, em cada caso, como apenas uma palavra diferente. Para extração da DiL, contamos com o auxílio do programa computacional Lexicanalytics<sup>6</sup>, software desenvolvido pelo LAME (Laboratório do Manuscrito Escolar/UFAL), contendo, dentre outras funcionalidades, um algoritmo para calcular a TTR. Esse software foi desenvolvido com o objetivo de otimizar o tempo e para fornecer resultados mais precisos (Figura 1).

A DeL foi extraída através do método de Ure, dividindo os itens lexicais pelo total de palavras escritas no texto. O total de palavras foi levantado com o auxílio do Lexicanalytics. Já os itens lexicais foram classificados morfossintaticamente e contados através do programa Excel. Após esse processo, a DeL foi calculada manualmente.

## Resultados e discussões

Os resultados a seguir, obtidos através do cálculo da DiL e DeL, poderão nos ajudar a compreender mais precisamente a diferença entre as produções individuais dos alunos que participaram das díades. Nossos resultados mostram que a relação entre Ti1 (pré-teste) e Ti2 (pós-teste) não é uniforme, nem linear. Ou seja, um aluno da díade que tinha o maior DiL ou DeL em Ti1 não necessariamente mantém a mesma relação em Ti2.

### Médias de Tokens e Types dos Ti e Td

Após contabilizar o número total de palavras (*tokens*) e o total de palavras diferentes (*types*) das produções em díades e individuais, calculamos as médias expressas na Tabela 1.

Comparando as produções textuais individuais Ti1 e Ti2 (pré-teste e pós-teste), temos um aumento na média de *Tokens* e *Types*. O aumento médio de, aproximadamente, 16 *tokens* e 11 *types* entre o conjunto de Ti1 e Ti2 sugere que os Td podem ter afetado o aumento dessas médias.

Nesta tabela, há também médias de *tokens* e *types* menores nos Td em relação aos Ti. Isso pode estar relacionado ao menor número de palavras escritas nos Td, pois durante as produções colaborativas os alunos precisavam estar de acordo sobre o que escrever e como escrever o texto.

### DiL em textos individuais

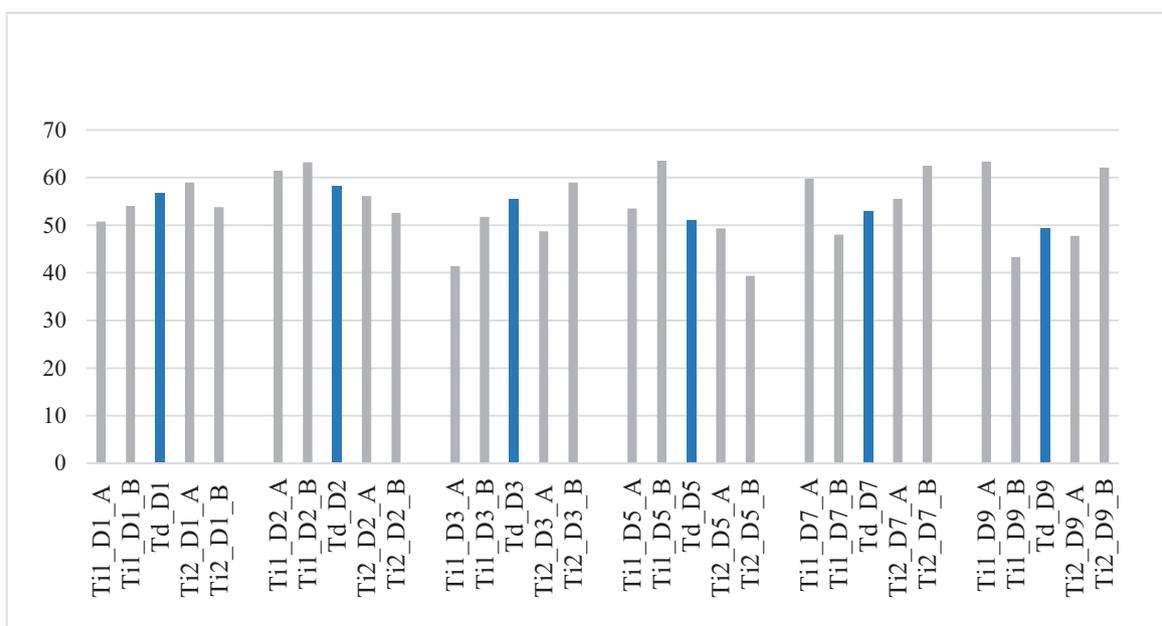
Os resultados obtidos pelo cálculo da DiL e da DeL podem nos ajudar a entender de modo mais preciso a diferença entre Ti1 e Ti2 dos alunos participantes das duplas. Estes resultados mostram que a relação não é constante. Ou seja, um aluno da dupla que teve a DiL ou a DeL mais alta no Ti1, não necessariamente irá manter esse *score* no Ti2, conforme pode ser observado no Gráfico 1.

**Tabela 1.** Média de *tokens* e *types*.

**Table 1.** Average tokens and types.

Textos	Produções	Médias de <i>tokens</i>	Média de <i>types</i>
Individual (pré-teste)	12	122,17	63,75
Díades	36	112,74	59,12
Individual (pós-teste)	12	137,83	73,33

<sup>6</sup>O programa Lexicanalytics foi desenvolvido por Glauber Leite e Elian Santos, a partir do tema de pesquisa proposto por esse estudo.



**Gráfico 1.** Média da DiL (%).

**Graph 1.** Average LDi (%).

A DiL entre os Ti1 e Ti2 de um mesmo aluno não é progressiva e homogênea. Por exemplo, para 5 alunos (D1\_A, D3\_A, D3\_B, D7\_B e D9\_B) houve aumento na DiL. Mas para 7 alunos (D1\_B, D2\_A, D2\_B, D5\_A, D5\_B, D7\_A e D9\_A) houve diminuição. Na comparação entre os Ti1 e Ti2 de uma mesma dupla, identificamos um resultado curioso. Somente o aluno D3\_B manteve uma DiL maior que seu parceiro nas duas produções individuais. A DiL de todos os outros alunos foi invertida, de forma que, quem teve maior DiL em um texto, teve menor DiL em outro, e vice-versa.

A DiL dos textos escritos em duplas apresentou aumento em relação a DiL do Ti1 dos alunos da dupla D1 e D3. Para a D2 e D5, a DiL do Td foi menor que a DiL Ti1. E para as duplas D7 e D9 a DiL do Td foi intermediária, ficando entre a DiL do Ti1 do aluno A e do Ti1 do aluno B.

Também não houve correspondência progressiva entre a DiL do Td e a DiL dos Ti2 de cada dupla. Para as duplas D2 e D5, a DiL do Td foi um pouco maior do que os Ti2 destas duplas. Para a dupla D7, a DiL do Td foi menor que a DiL do Ti2. Para as duplas D1, D3 e D9, a DiL do Ti2 maior para um aluno e menor para outro.

No Gráfico 2, pôde ser observado que a DeL dos Ti1 e Ti2 apresenta característica semelhante à DiL, progressiva e homogênea. Ou seja, para os alunos D2\_A, D2\_B, D3\_A, D5\_B e D9\_A houve uma variação negativa entre Ti1 e Ti2, mas para os alunos D1\_A, D1\_B, D3\_B, D5\_A, D7\_A, D7\_B e D9\_B a variação entre os textos foi positiva, ainda que em alguns casos tenha sido mínima. Na comparação entre a DiL e DeL das produções individuais, somente quatro alunos (D1\_A, D3\_B, D7\_B e D9\_B)

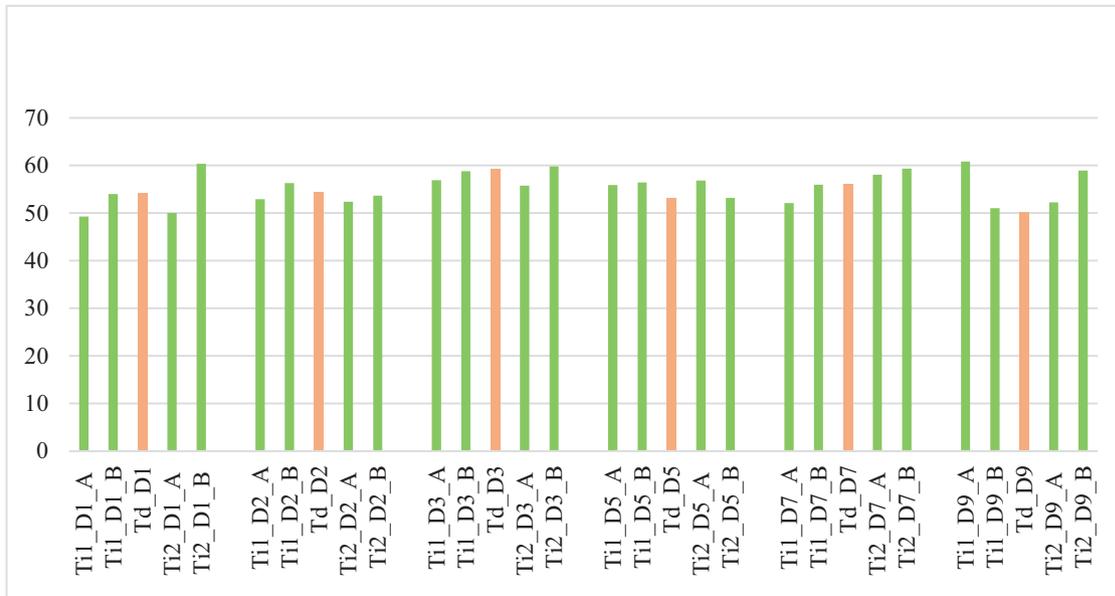
mantiveram variação positiva em ambas as medidas. Nas produções em díades, a média da DeL não mostrou variância significativa entre as duplas, porém, na comparação com a DeL das produções individuais, observamos que a do Ti1 foi menor ou igual a dos Td em D1\_A, D1\_B, D2\_A, D3\_A, D3\_B, D7\_A, D7\_B e maior para D2\_B, D5\_A, D5\_B, D9\_A, D9\_B.

### **DiL e DeL dos textos em díades**

A relação entre as médias de DiL e DeL nos textos em dupla também é diversa, havendo maior DiL nas duplas D1, D2 e D3 e maior DeL nas duplas D5, D7 e D9 (Gráfico 3).

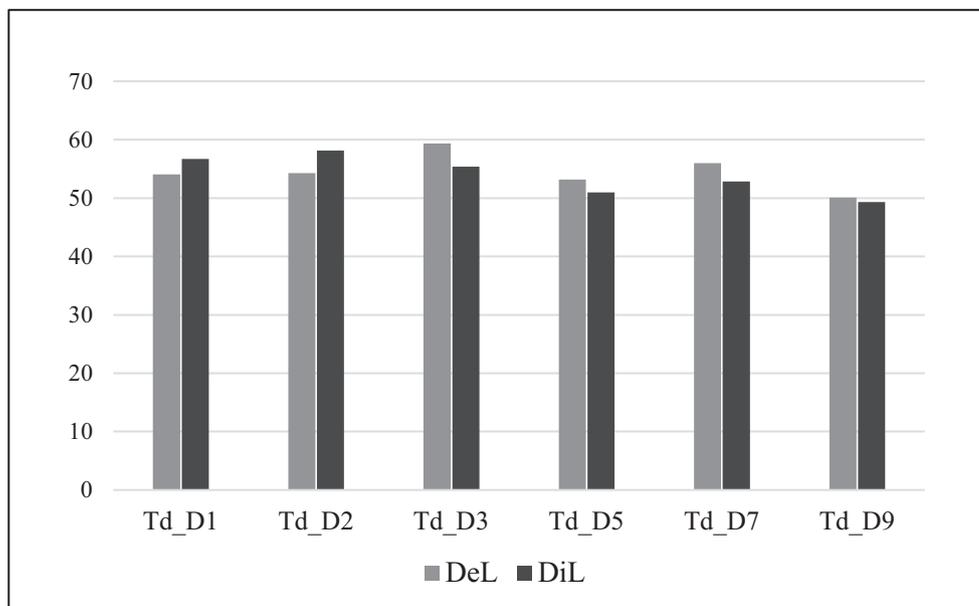
Apesar das diferenças entre DiL e DeL, podemos observar que a variação entre as médias dos textos em duplas de cada díade é pequena. Apenas a D9 apresenta uma diferença maior em relação aos seus pares. Sobre as duplas que obtiveram as maiores médias na DiL e DeL, destacam-se a D2 com maior DiL e a D3 com maior DeL. Já as menores médias foram apresentadas pela D9, apesar de ter sido a dupla que escreveu os textos mais longos.

Comparando as duplas que obtiveram menores e maiores médias, temos uma variação de 49,56 a 57,80 (diferença de 8,24%) na DiL, enquanto que para a DeL uma variação de 50,31 a 56,11 (diferença de 5,8%), resultando em uma diferença de 2,44% entre as medidas de DiL e DeL. Isso mostra a interdependência entre elas e, ao mesmo tempo, certa estabilidade entre os textos produzidos em dupla.



**Gráfico 2.** Média da DeL (%).

**Graph 2.** Average LDe (%).



**Gráfico 3.** Média da DeL e DiL dos textos em díades.

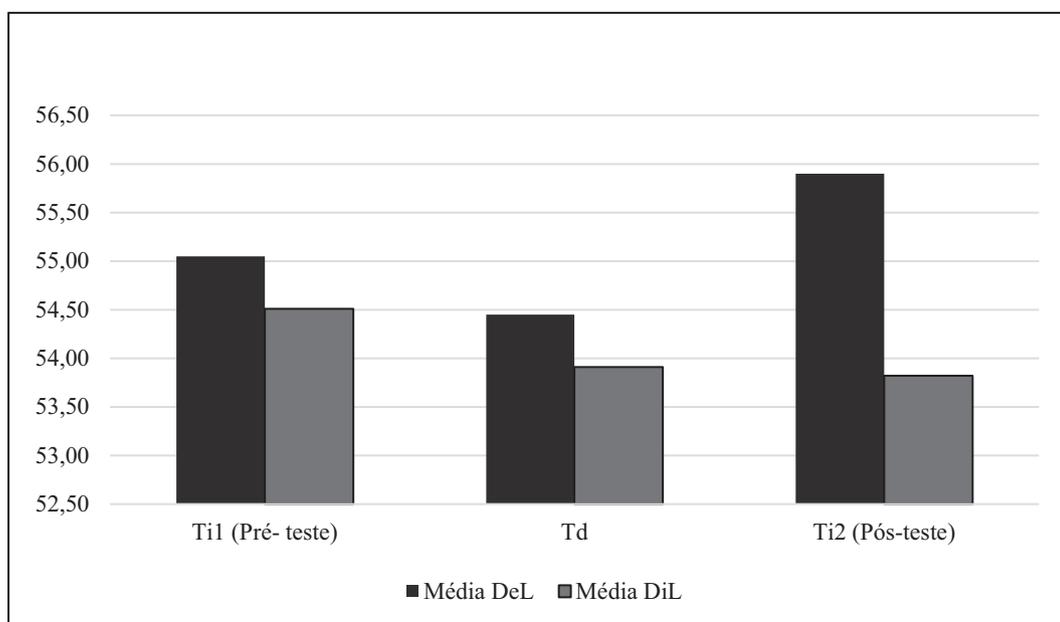
**Graph 3.** Average LDe and Average LDi on dyadic texts.

### Resultados gerais da DeL e DiL em Ti e Td

A comparação entre as médias gerais de DeL e DiL obtidas nos 3 momentos (Ti1, Td e Ti2) fornece informações importantes.

Por fim, foi observado que os textos produzidos em duplas obtiveram menor DiL e DeL em relação às produções individuais (Gráfico 4). Contudo, nas produ-

ções individuais, o Ti2 aponta dois movimentos opostos: enquanto a DiL mostra uma média menor que as produções anteriores, a DeL encontra-se maior, inclusive que a do Ti1. Isso sugere que as produções anteriores possam ter contribuído para a riqueza lexical, levando os alunos a ampliarem os itens lexicais usados, apesar de terem diminuído a diferença entre palavras (*tokens*) e palavras diferentes (*types*) por texto.



**Gráfico 4.** Comparação entre as médias de DeL e DiL.

**Graph 4.** Comparison between LDe and LDi average.

## Conclusão

Este trabalho dedicou-se a investigar a presença de dois indicadores linguísticos (DeL e DiL) associados à riqueza lexical do texto, com o objetivo de observar como eles se comportam em produções de textos escritos individualmente e em dupla. Nossos resultados indicam que, na amostra analisada, as produções individuais e as produções em dupla apresentaram um número de palavras escritas, aproximadamente, duas vezes maior que o do número de palavras diferentes. Este resultado é semelhante ao já identificado no estudo de Templin (1957).

Outro resultado a destacar é o decréscimo da DiL nas produções individuais Ti2. Esse decréscimo parece estar associado ao crescimento do número de *tokens* e *types*, visto que o crescimento do número de *types* não acompanhou o crescimento do número total de *tokens*, resultando em uma relação inversamente proporcional entre a DiL e o número de *tokens*. Alguns estudos sobre a DiL (Deboer, 2014) associam essa relação inversa à repetição do vocabulário funcional, e.g. preposições e artigos.

A variação da DiL e da DeL nos textos individuais foi maior do que nos textos em duplas, mas se considerarmos somente as produções em dupla, os textos são mais homogêneos em relação a estes dois aspectos, apesar da DiL ter aumentado mais do que a DeL.

Por fim, temos duas tendências no comportamento da DiL e da DeL entre os Ti e Td. Na relação entre Ti1 e Td, tanto a DiL, quanto a DeL diminuíram. Já na relação entre Td e Ti2, houve um aumento na DeL, porém uma di-

minuição da DiL. Considerando que a riqueza lexical está relacionada à qualidade textual, tais resultados sugerem, desta perspectiva, que os textos individuais apresentaram maior qualidade que os textos escritos em dupla.

## Referências

- COLOMBI, M.C. 2000. Academic language development in Latino student's writing. In: M.J. SCHLEPPGREGEL; M.C. COLOMBI (eds.), *Developing advanced literacy in first and second languages*. Mahwah, Lawrence Erlbaum Associates, p. 67-86.
- DEBOER, F. 2014. Evaluating the comparability of two measures of lexical diversity. *System*, 47(1):139-145. <https://doi.org/10.1016/j.system.2014.10.008>
- DUARTE, I.; COLAÇO, M.; FREITAS, M.; GONÇALVES, A. 2011. *O conhecimento da língua: desenvolver a consciência lexical*. 1ª ed., Lisboa, Estremadura/Ministério da Educação, 99 p.
- HALLIDAY, M.A.K. 1985. *Spoken and written language*. Geelong Vict.: Deakin University, p. 59-86.
- JOHANSSON, V. 2008. *Lexical density and lexical density in speech and writing: a developmental perspective*. Working Papers, Lund University, 53(2):61-79.
- JOHANSSON, V. 2009. *Developmental aspects of text production in writing and speech*. Lund. Tese de Doutorado. Lund University, 242 p.
- LE NORMAND, M.-T.; PARISSE, C.; COHEN, H. 2008. Lexical diversity and productivity in French preschoolers developmental and biosocial aspects by developmental, gender and sociocultural factors. *Clinical Linguistics & Phonetics*, 22(1):47-58. <https://doi.org/10.1080/02699200701669945>
- MALVERN, D.; RICHARDS, B.J.; CHIPERE, N. 2004. *Lexical diversity and language development: Quantification and assessment*. Basingstoke, Palgrave Macmillan, 182 p. <https://doi.org/10.1057/9780230511804>
- MARTINS, M. 2016. *Complexidade textual e progressão escolar em dois registos: um estudo de correlação baseado em um corpus quasi-longitudinal*. Lisboa, Portugal. Tese de Doutorado. Universidade de Lisboa, 211 p.

- SCHERER, S.; SOUZA, A. 2011. Types e tokens na aquisição típica de linguagem por sujeitos de 18 a 32 meses falantes do português brasileiro. *Revista CEFAC*, 13(5):838- 845.  
<https://doi.org/10.1590/S1516-18462011005000058>
- SCHERER, S.; CASARIN, F.; ZART, P.; RAMOS, A. 2002. Perfil evolutivo da relação type/token de crianças de 2 a 5 anos de idade. *Revista CEFAC*, 4(1):223-228.
- READ, J. 2000. *Assessing vocabulary*. Cambridge/Londres. University Press, 271 p. <https://doi.org/10.1017/CBO9780511732942>
- RODRIGUES, S.B.P. 2008. *Escrita espontânea: Desenvolvimento das capacidades de composição escrita de em crianças do 1º ao 4º ano de escolaridade*. Porto, Portugal. Dissertação de Mestrado. Universidade Fernando Pessoa, 107 p.
- TEMPLIN, M. 1957. *Certain language skills in children: their development and interrelations*. Minnesota, Estados Unidos. University of Minnesota Press, 138 p.
- URE, J. 1971. Lexical density and register differentiation. In: G.E. PERREN; J.L.M. TRIM (eds.), *Applications of linguistics. Selected papers of the Second International Congress of Applied Linguistics*. Cambridge/Londres, Cambridge University Press, p. 443-452.

Submetido: 17/04/2017  
Aceito: 23/01/2018