A DENSIDADE DE IDEIAS E A COMPLEXIDADE GRAMATICAL NA AVALIAÇÃO DA PRODUÇÃO ESCRITA EM PORTUGUÊS L2

Ângela Filipe Lopes

aflopes@letras.up.pt

Faculdade de Letras da Universidade do Porto, Portugal

Resumo: A avaliação da produção escrita em L2 de nível avançado é de difícil quantificação, como também são os objetivos estabelecidos para esta competência no mesmo nível. A correção não é suficiente. Procura-se a qualidade da escrita, traduzida na precisão e na riqueza lexicais, bem como na complexidade frásica. No sentido de medir estas características da escrita sofisticada, propõem-se a Densidade de Ideias e a Complexidade Gramatical como parâmetros de avaliação da escrita já explorados por Snowdon *et al.* (1996) e por Kemper *et al.* (2001) no âmbito do *Nun Study* (Snowdon 2001). A sua aplicação ao contexto da aprendizagem do português L2 no nível avançado (C1/C2) é feita no sentido de estudar a natureza das frases produzidas em duas tarefas de produção escrita: uma síntese e uma réplica. Verificou-se, em primeiro lugar, uma variabilidade elevada nos resultados que decorre da heterogeneidade do grupo de estudantes em questão. Em segundo lugar, ambas as medidas evidenciaram uma associação significativa em ambas as tarefas de escrita, levando a concluir que quem escrever com riqueza e precisão lexicais produzirá também frases de complexidade gramatical mais elevada. Por último, verificou-se a influência da leitura sobre a escrita, na medida em que os estudantes que leem por hábito são aqueles que obtiveram marcas mais elevadas nos dois parâmetros medidos.

Palavras-chave: escrita em português L2; densidade de ideias; complexidade gramatical; *continuum* leitura-escrita

Abstract: Advanced L2 written production assessment is difficult to quantify as are learning goals established for this competence in the same level. Correction is not enough. Quality in writing is expected to translate into lexical precision and richness, as well as sentence complexity. In order to measure these characteristics of sophisticated writing, Idea Density and Grammatical Complexity are proposed as assessment parameters as they were explored by Snowdon *et al.* (1996) and by Kemper *et al.* (2001) in the *Nun Study* (Snowdon 2001). Its adoption to the advanced Portuguese L2 learning context (C1/C2) underlied the study of sentences produced in two writing tasks: a synthesis and a writing prompt. First, results were highly variable due to the heterogeneity of the group of participants. Both measures were found to be correlated in

both writing tasks, which led to the conclusion that participants who write with lexical precision and richness also produce higher complexity sentences. Lastly, the results confirmed that reading has an influence on writing quality, insofar as the students who usually read as the ones who got the highest scores in both parameters.

Keywords: L2 Portuguese writing; idea density; grammatical complexity; reading-writing *continuum*

1 - Introdução

O domínio da escrita em L2 de nível avançado é, independentemente do registo, uma meta difícil de definir na medida em que os critérios que sustentam a sua avaliação fogem a uma caracterização fácil. Neste nível, já não se procura a correção, mas a riqueza da expressão, a originalidade da palavra e a complexidade criativa da frase. Citando o título de Leki (2015), pode dizer-se “Good writing: I know it when I see it”, mas, a despeito desse instinto, quais são os ingredientes linguísticos que fazem uma boa escrita? E quem são os autores dessa produção superior?

Se, ao ler, o aprendente de português L2 pode debater-se com a incompreensão de frases que, apesar de compostas por palavras familiares, persistem como obstáculos, na escrita essa dificuldade reflete-se habitualmente numa simplificação excessiva, no recurso ao vocabulário mais familiar e na composição menos intrincada da frase por serem mecanismos menos onerosos do ponto de vista psicolinguístico. O *continuum* leitura-escrita (Hirvela 2004) implica que o leitor que experimenta dificuldades na compreensão de frases mais desafiantes seja também aquele que enfrenta mais entraves à escrita do mesmo tipo de frase. A frase mais complexa entende-se como aquela que é construída por meio de encaixes menos convencionais, como as relativas de encaixe à esquerda, ou como as subordinadas múltiplas (Kemper *et al.* 2001; Pinto 1994; Snowdon *et al.* 1996) Por outro lado, a riqueza e precisão lexicais que quem escreve é capaz de imprimir ao texto estão associadas ao grau de complexidade gramatical encontrado significando que frases de configurações menos convencionais se fazem acompanhar de um recheio lexical que foge ao mais comum e que, por isso mesmo, empresta riqueza e profundidade ao texto (Berman 1984; Kemper *et al.* 2001; Snowdon *et al.* 1996), Snowdon *et al.* 1999).

Assim, no sentido de avaliar a produção escrita de 14 estudantes de nível avançado (C1/C2) do Curso Anual de Português para Estrangeiros da Faculdade de Letras da Universidade do Porto (ano letivo 2013-2014), foram sujeitos a uma dupla análise dois textos produzidos pelos mesmos estudantes a partir de leituras anteriores. Esta análise baseou-se na quantificação da Densidade de Ideias (DI) e da Complexidade Gramatical (CG), parâmetros definidos e utilizados por Snowdon *et al.* (1996), Snowdon *et al.* (1999) e Kemper *et al.* (2001) no âmbito de um

estudo centrado na comparação de textos produzidos por uma população idosa em dois momentos da vida (ver também Snowdon 2001). No contexto aqui em apreço estes dois parâmetros foram adaptados ao contexto de aprendizagem do português L2 com o intuito de avaliar a qualidade da escrita.

Colocou-se assim uma questão metodológica que, no caso destes parâmetros de avaliação, se centrou nos fatores que podem fazer variar a competência leitora e escrita e nas consequências visíveis na produção escrita a partir da leitura. Estabeleceu-se como objetivo geral avaliar a natureza das últimas seis frases de cada um dos dois textos produzidos quanto à Densidade de Ideias (DI) e à Complexidade Gramatical (CG). Assim, procurou-se diagnosticar dificuldades e isolar variáveis inter e intraindividuais observadas neste público. Os objetivos específicos orientaram-se no sentido de estudar as frases produzidas pelos estudantes quanto à DI e à CG. Esperava-se que os estudantes que tinham obtido resultados mais elevados num teste *Cloze* anterior (Lopes 2014; Lopes, Martins & Pinto 2015; Lopes 2020) alcançassem marcas mais elevadas no desempenho escrito, evidenciando a ligação leitura-escrita. No campo dos dois parâmetros utilizados na avaliação, antecipava-se uma associação entre DI e CG, confirmando resultados obtidos por Snowdon *et al.* (1996), por Snowdon *et al.* (1999) e por Kemper *et al.* (2001) relativamente à L1.

2 - A complexidade gramatical

Termos como complexidade gramatical ou complexidade sintática são referidos por vários autores (Berman 1984; Bernhardt 1991; 2011; Cheung & Kemper 1992; Grabe 2009; Kemper & Sumner 2001; Snowdon *et al.* 1999) para explicar de que maneira certas frases colocam entraves à compreensão e à produção escritas. É importante, para reduzir ou eliminar este tipo de obstáculo, ter em conta os aspetos que são isolados sob este prisma. A sondagem informal que Berman (1984) levou a cabo com um grupo de estudantes de inglês L2 de nível avançado permitiu à autora destacar dois fatores que explicam bloqueios durante a leitura das frases que estes estudantes indicaram ser as mais problemáticas: “what the parts of a sentence are, and how they interrelate hierarchichally” (Berman 1984: 140). Este tipo de compreensão implica a identificação das componentes básicas da frase, como a oração principal e as subordinadas, a estrutura argumental do verbo, sobretudo a relação sujeito-verbo-complementos, e um triângulo elementar de papéis temáticos agente-ação-paciente. Como Berman (1984) refere, “non-apprehension of SVO relations and basic constituent structure made it impossible for students to get at the propositional core of (...) rather complicated sentences” (p. 152). De facto, o núcleo proposicional remete para o sentido da frase, isto é, para as ideias que estão enraizadas na construção da estrutura cuja compreensão depende da identificação dos grupos de palavras que a constituem. A razão pela qual Berman (1984) prefere apontar este tipo de relação frásica como a mais nuclear prende-se com a observação de dificuldades na compreensão de frases nas quais existe conflito na ordem mais básica esperada pelos leitores quanto às relações

sintático-semânticas, como é o caso de frases passivas, de sintagmas adverbiais antepostos à oração principal ou de outro material linguístico, que não o sujeito, presente à esquerda do verbo. Como constata Berman (1984), a falta de contacto com material linguístico de estrutura variada e gradualmente mais desafiante não só explica o carácter inesperado de frases menos comuns, mas também não favorece a flexibilidade mental do aprendiz. Convém frisar que a ocorrência flexível da ordem das palavras na frase, também em português L2, só se fará sentir através da exposição a materiais de leitura diversos sem a qual não se pode esperar que a produção escrita destes leitores reflita um domínio confortável da língua. Dir-se-ia que, como noutras situações de aprendizagem, o sucessivo “scaffolding” de Bruner (1966) é sempre pertinente.

Berman (1984) não retira, por isso, importância a constituintes secundários na hierarquia da frase e dá antes exemplos das dificuldades que estruturas de subordinação tendem a colocar aos leitores intervenientes no seu estudo. Nestas situações, chama a atenção para a tendência para atribuir um único significado a cada conector. Verificou, por exemplo, que o marcador *while* era sempre visto como temporal em detrimento do seu possível valor concessivo, como aliás sucede também em português com *enquanto* (ou *enquanto que*) que assume um valor temporal, mas também pode ocorrer com valor contrastivo (Lobo 2013: 2006). O valor polissémico conjuncional tem nestes casos um peso considerável em virtude de, no caso de não ser reconhecido, levar a leituras que não se aproximam da função pretendida pelo autor, portanto, do sentido da frase. Koda (2004) defende que “structural complexity typically stems from syntactic maneuvers such as subordination” (p. 96); manobras essas que convocam mais recursos atencionais e mais tempo de armazenamento temporário de material linguístico, levando a um exercício mental continuado de retenção e de integração de elementos na construção de sentido. Mais importante, para além da identificação de recursos conjuncionais na frase, Berman (1984) aponta a complexidade sintática como fator determinante para a compreensão do texto. A autora define a frase complexa como aquela que contém elementos elípticos ou truncados que, por isso, na sua visão, originam alguma confusão referencial. O (re)conhecimento da mecânica da frase a este nível pode ser útil no processo de seleção de funções e de significados de palavras funcionais sem as quais o leitor pode não compreender plenamente o texto ou reutilizar esses recursos em produções linguísticas próprias. Koda (2010), por seu lado, sublinha que frases longas com orações subordinadas podem contribuir para um processamento mais linear e mais lógico da informação quando são processadas de forma sequencial, ao passo que as frases curtas se podem afigurar mais truncadas e compactas, impondo a mobilização de mais recursos atencionais à sua compreensão em L2. Com efeito, a diferença entre o processamento de frases simples e curtas e de frases complexas e longas verifica-se na riqueza de sentido extraído e (re)construído a partir da sua leitura: uma frase simples permite mais frequentemente a compreensão de informação de características explícitas. Pelo contrário, a leitura de uma frase complexa leva à geração de inferências e à consideração de mais informação tácita. É também

por este motivo que um texto, em geral, mais complexo se pode considerar mais hermético e, portanto, mais desafiador para um leitor de L2. Na opinião de Kemper *et al.* (2004) e de Kemper e Herman (2006), a colocação dos encaixes à esquerda determina também o conteúdo proposicional e o grau de complexidade gramatical (Kemper *et al.* 2004: 5). Gibson (1998) aponta como exemplos deste grau acrescido de complexidade as construções relativas com função de sujeito ou com função de complemento direto. O processamento destas duas estruturas não depende de efeitos possíveis de ambiguidade local, mas do patamar mais elevado de esforço de memória necessário para reter e processar as frases relativas do segundo tipo (Gibson 1998: 2-3). Os encaixes à esquerda produzem, de acordo com o autor, uma sobrecarga que é inultrapassável no caso de duas relativas encaixadas entre sujeito e predicado, sobretudo quando as relativas têm função de complemento direto.

Em L2, estes fatores pesam certamente mais do que na L1, na medida em que há limitações articulatórias e lexicais que impõem um abrandamento nas funções da memória operatória (M.O.) das quais depende a recuperação de significado e a correspondência com elementos já armazenados na memória a longo prazo (M.L.P.). Por outro lado, estas limitações podem conduzir adicionalmente a uma diminuição de controlo de atenção e de mecanismos de inibição de informação acessória que pode surgir como um desvio do cerne da tarefa de compreensão em causa, levando o leitor a acomodar estruturas de forma superficial e imprecisa cujos resultados se podem refletir numa abordagem estereotipada ao texto com prejuízo para o seu valor real e para o conteúdo efetivo. Bialystok *et al.* (2012) definem o controlo executivo necessário a este tipo de operações como “the set of cognitive skills based on limited cognitive resources for such functions as inhibition, switching attention, and working memory.” (Bialystok *et al.* 2012: 22).

3 - A densidade de ideias

A densidade de ideias está ligada à qualidade e à precisão da escrita (Snowdon *et al.* 1996; 1999). Como referem Kemper *et al.* (2003), “[t]he number of propositions expressed in a sentence is a measure of how informative it is” (p. 5). Kintsch (1974; 1998) vê a proposição como o produto da interseção entre as memórias pessoais e a reserva de conhecimento geral de cada indivíduo, resultando numa construção que tem tanto de linguística como de conceptual e que tende a constituir uma abstração de referentes reais. Assim, a proposição é, no entender do autor, a expressão de uma ideia presente na relação próxima entre o verbo e os seus argumentos. Há nesta conceção uma visão da arquitetura da ideia que se configura como uma base para a expressão mais precisa e abrangente do pensamento que procuramos traduzir e construir através da escrita.

Pinker (2014) aborda as características de um certo tipo de escrita altamente proficiente, mais informativa e viva, destacando alguns pontos relevantes como “an insistence on fresh writing”, o uso de “concrete imagery” ou “an attention to the reader’s vantage point” e, não menos importante, “the judicious placement of an uncommon word or idiom against a backdrop of simple nouns and verbs”

ou “the occasional planned surprise” (Pinker 2014: 26). Que efeito tem no leitor este conjunto de estratégias presentes no texto, cuidadosamente selecionadas pelo seu autor? Por que motivo fazem, no entender de Pinker (2014), um bom texto? Sendo reflexos de um cuidado que se torna legível, como se obtém um resultado deste tipo? Quem é capaz de gerar este tipo de texto e de fazer funcionar material linguístico a favor da qualidade e da eficácia? E, mais importante, como se desenha a instrução da leitura-escrita de molde a ajudar o aprendente a alcançar este nível de expressão escrita numa L2?

Pinker (2014) atribui este nível de qualidade a uma “reverse-engineering” que se leva a cabo ao longo da escrita, no fundo aludindo à distância necessária entre texto e autor a que Olson (1994) também se refere. A capacidade de ter em conta este tipo de engenharia inversa advém, segundo Pinker (2014), da leitura durante a qual o leitor, depois escritor (ou escrevente), absorve truques retóricos, um repertório lexical considerável, padrões estruturais linguísticos e alguma sensibilidade na combinação destes elementos. Esta reutilização de recursos é feita por leitores ávidos que se apercebem de forma mais ou menos consciente dos mecanismos que compõem a escrita que apreciam.

Talvez leituras anteriores conscientes e continuadas expliquem um determinado grau de qualidade na escrita, mas na possível ausência desta interação produtiva com a leitura, convém saber como se pode dissecar esta capacidade com o intuito de permitir, a quem não lia antes, escrever melhor sobretudo numa L2. Afinal, na linha de pensamento de Stanovich (1986) e de Pinker (2014) quem lê muito, lê melhor e escreve melhor. Não obstante, isto não invalida uma atitude interventiva por parte do docente de português L2 em face de leitores menos ágeis e, portanto, possivelmente autores em face de mais obstáculos.

O processo de criação de pensamento passa por etapas pré-linguísticas, como sublinha Kintsch (1998), ou também Damasio (1989), mas estas etapas rapidamente seguem um curso verbal que obedece a uma organização conceptual e a relações proposicionais que permitem veicular uma ideia. No caso deste estudo, pretendeu-se olhar para a mecânica que subjaz tanto à compreensão escrita como à verbalização escrita em (português) L2. Trata-se de estudar o processo de compreensão das ideias de um autor e a (re)conversão das mesmas ideias para uma escrita própria numa língua que não é a materna e que, por isso mesmo, impõe constrangimentos como a ausência de vocabulário, a falta de fluência na leitura ou na escrita, a inexistência de referentes culturais ou o desconhecimento estrutural da língua. Este interesse justifica-se pela necessidade de apoiar o aprendente de nível avançado que se depara, por necessidade ou interesse pessoal, com textos que fogem aos que são mais quotidianos, imediatos ou concretos, assim como com temas aos quais deve dar resposta através da composição escrita. De acordo com Carson *et al.* (1990), o input intralinguístico é tanto mais relevante quanto mais alto é o nível de proficiência dos estudantes de L2. Os autores apontam duas fontes essenciais de onde deriva a construção do sistema linguístico em L2: o conhecimento prévio da L1 e o input em L2.

Propõe-se em primeiro lugar, no sentido de compreender melhor o que

é transferido do texto lido em L2 para a escrita também em L2, a medição da DI encontrada no texto escrito (Snowdon *et al.* 1999; Kemper *et al.* 2001). Não se preconiza, desta forma, que os textos produzidos pelos estudantes se assemelhem em grau de qualidade àqueles que leram, nem tão pouco se pretende medir os mesmos parâmetros em textos cujos autores escrevem tendo em atenção contornos literários que não podem ser sujeitos a uma avaliação semelhante. Pretende-se antes obter, através desta avaliação, um ponto de referência para os resultados que provenham da análise aos textos dos participantes no estudo.

De acordo com Kemper *et al.* (2001), “idea density is correlated with measures of vocabulary” (p. 227), enquanto “grammatical complexity is correlated with measures of working memory, including digit span and reading span” (p. 227). Sendo certo que a autora se refere sobretudo à densidade de ideias em L1, se este parâmetro se relaciona com a capacidade de definir conceitos e palavras e com a fluência verbal (Kemper *et al.* 2001: 228), esta é uma medida que pode ser útil também em L2, sobretudo nos níveis de proficiência mais avançados, por se prender com a compreensão e com a produção verbais geral. É, realmente, um parâmetro que pode levar a procurar reflexos da boa aplicação da palavra, isto é, em contexto tanto gramatical quanto situacional. A densidade de ideias mede-se, de acordo com Kemper *et al.* (2001), que se basearam em Kintsch (1974) e em Turner e Greene (1977), com base na divisão proposicional simples dos enunciados e no cálculo da média de proposições por cada dez palavras (Kemper *et al.* 2001: 229-230). Chand *et al.* (2012) definem a densidade de ideias como “the only approach that directly measures the ability to use world knowledge to structure propositions in spontaneous speech” (p.5). Os autores seguem a linha de Kintsch (1974; 1998) ao afirmarem que um texto se pode reduzir a uma lista de proposições ou de ideias codificadas pelo autor.

<p>1. O narrador conta como o amor entre ele e a sua esposa/ namorada vai-se diminuindo. (participante 9)</p> <p>Conta como o amor (1) Entre ele e a sua esposa/namorada (2) Vai-se diminuindo (3)</p> <p>Frase1:3ideias ÷ 15palavras x10 = 2DI</p>	<p>2. Muitas vezes acham que há pessoas divertidas, cultas, atraentes e parvos. (participante 13)</p> <p>Muitas vezes Acham que (1) Há pessoas divertidas (2) (Há pessoas) cultas (3) (Há pessoas) atraentes (4) E (5) (Há pessoas) parvos (6)</p> <p>Frase 2: 6 ideias ÷ 11 palavras x 10 = 5,45 DI</p>
--	---

Tabela 1. Exemplos de análise de frases dos participantes quando à DI

Os três grupos de proposições (predicadores, modificadores e conectores) apontados por Kintsch (1974) são utilizados como a base para a seleção de funções relevantes para a identificação rápida das classes de palavras que se podem associar às ideias: Verbo (V), Adjetivo (Adj), Advérbio (Adv), Conjunção (Conj) ou Sintagma Preposicional (SPrep), elementos que Snowdon *et al.* (1996; 1999) também elencam na explicação da análise dos textos estudados. A título de exemplo,

encontram-se duas frases na Tabela 1: uma com densidade de ideias alta e outra com densidade de ideias baixa. As frases foram produzidas por dois participantes no estudo e servem meramente, neste ponto do trabalho, para ilustrar o método de medição da DI e as diferenças possíveis entre frases mais densas e outras com menos pontuação neste parâmetro. A ideia, ou a proposição, implica um elo entre o pensamento e a sua expressão verbal. A mensagem não resulta de uma codificação de léxico independente, como as frases não se resumem a uma lista de palavras. O pensamento macroplaneado a um nível pré-linguístico é codificado de acordo com relações estabelecidas entre ideias (ou predicções elementares) e por isso mesmo resultam em enunciados que se pautam pelas relações linguísticas que traduzem as relações entre essas ideias. No processo de conversão destas ideias em material verbal, entram o acesso lexical, mas também a formulação intencional das componentes do enunciado. De facto, não basta que a correção gramatical esteja presente para que o discurso verbal faça sentido, mas que estas relações linguísticas traduzam relações entre ideias. Para isso, é também necessário que haja controlo, intenção, sobre a produção linguística, na medida em que um discurso automático ou jaculatório, como o define Sachs (2017), não traduz pensamento mais elaborado. Sustenta-se antes em formulações menos refletidas e em automatismos. A seleção dos elementos linguísticos surge pela sua identificação com conceitos abstratos formados com base em léxico armazenado na memória semântica. Quando estes conceitos são concretizados em linguagem verbal podem ser compostos por um ou mais itens lexicais organizados de acordo com as relações proposicionais que correspondem, em geral, a uma ligação argumental. Brown *et al.* (2008), definem ainda a proposição como a unidade envolvida na compreensão e recuperação de um texto, o que reforça a sua importância tanto para a compreensão leitora como para a produção escrita na dependência da leitura. Completam esta definição dizendo que a DI é uma medida que permite mais do que olhar para a referência a entidades. Abre sim uma porta para um grau variável de questionamento e de afirmação do seu autor. São estas as relações que, no entender dos autores, se veem refletidas nas unidades contabilizadas como ideias: o verbo e os seus argumentos, contanto como uma proposição, e os seus elementos complementares como os adjetivos, os advérbios e as conjunções, não esquecendo sintagmas preposicionais, contabilizados como proposições independentes do verbo. Este sistema, que como já foi referido deriva do estudo de Kintsch (1998), não contempla os nomes por nem sempre serem denunciadores de relações entre ideias, mas parte de outros sistemas integrados nas relações adjacentes ao verbo. No caso do estudo aqui em questão, cumpre sublinhar que o objetivo é medir a DI seguindo os parâmetros apontados por Susan Kemper (Kemper *et al.* 2001) e utilizados posteriormente por Brown *et al.* (2008) ou por Chand *et al.* (2012).

4 - Metodologia

4.1 - Participantes

No grupo de 14 participantes neste estudo, contavam-se 4 do sexo masculino e 10 do sexo feminino, cujas nacionalidades se distribuíam como se descreve em seguida: alemã (2), espanhola (5), italiana (1), russa (1), venezuelana (1), vietnamita (1) e timorense (3). As idades variavam entre os 20 e os 60 anos, com uma média de 29,21 (desvio padrão de 9,916). Nove destes 14 participantes no estudo frequentavam no ano letivo de 2013-2014 cursos de mestrado e de doutoramento na Faculdade de Letras da Universidade do Porto (64,28%).

O português era para todos uma L3, de acordo com a descrição que Hammarberg (2001) faz da língua cuja aprendizagem se encontra em curso, independentemente da ordem cronológica a que obedece a aprendizagem dessa e de outras L2s. As L1s são consequentes com as nacionalidades, à exceção das línguas maternas dos participantes timorenses: tétum, tokodede e quemaq. As duas últimas línguas foram descritas por estes falantes como sendo exclusivamente orais.

Em resposta a um questionário distribuído à entrada no curso de português, todos os participantes isolaram como dificuldades principais a leitura e a escrita, a gramática e a produção oral. Entre as expectativas à entrada no curso, contavam-se o aprofundamento das capacidades de leitura e escrita, o alargamento de vocabulário e o aperfeiçoamento da pronúncia.

4.2 - Materiais e descrição do estudo

No contexto de duas unidades letivas, os estudantes leram duas crónicas na dependência das quais lhes foram propostas duas tarefas de escrita, uma Síntese e uma Réplica (ou *writing prompt*), levadas a cabo de forma autónoma. Após a leitura da primeira crónica, *Olhar para ontem* de António Lobo Antunes, escreveram de forma autónoma uma Síntese e na dependência da segunda crónica, *O meu lugar* de José Luís Peixoto, escreveram uma Réplica, isto é, um texto pessoal subordinado ao mesmo título da crónica lida. A variedade de temas e de configurações textuais e retóricas das crónicas foi intencional e destinou-se a promover o contacto com textos que, pertencendo ao mesmo tipo, podiam ainda ser veículos de contextos variados, de temas que oscilassem entre os universos público e pessoal, refletindo-se estas características em registos de linguagem variáveis e em graus de dificuldade que pudessem também beneficiar todos os estudantes.

Os materiais dos quais dependeram as duas tarefas de avaliação da escrita são eles próprios materiais de leitura e estão disponíveis *online*¹. No contexto de unidades letivas diferentes, foram propostas aos estudantes tarefas de leitura e subsequentemente de escrita que resultaram em produções escritas analisadas neste estudo. Os estudantes autorizaram que os materiais que resultassem de

¹ *Olhar para ontem* está disponível em <https://visao.sapo.pt/opiniao/a/antonio-lobo-antunes/2013-10-25-olhar-para-ontem/754335/> (último acesso a 6 de março de 2021). *O meu lugar* está disponível em <http://www.joseluispeixoto.net/80811.html> (último acesso a 6 de março de 2021).

tarefas que lhes foram propostas fossem utilizados como material de análise no contexto de investigação levada a cabo na FLUP de forma anónima.

4.3 - Procedimentos

Na Tabela 2, encontram-se os procedimentos levados a cabo no contexto das unidades letivas de Português L2 com vista à leitura e à posterior escrita de dois textos. As unidades letivas que precederam a produção autónoma dos textos analisados podem ser descritas como momentos de pré-escrita (segundo a perspetiva de Emig 1971) se tivermos em conta que culminam na composição de um texto derivado da leitura, da compreensão e da discussão das duas crónicas. No âmbito da análise estatística dos dados recolhidos junto dos sujeitos participantes no estudo, usou-se o programa IBM SPSS Statistics, versão 26. Descrevem-se em seguida os procedimentos de análise dos textos redigidos pelos estudantes.

<p><i>Síntese</i> <i>Olhar para ontem</i> <i>António Lobo Antunes</i></p>	<p><i>Writing Prompt/ Réplica</i> <i>Ô meu lugar</i> <i>José Luís Peixoto</i></p>
<p>1. Distribuição da crónica; 2. Leitura silenciosa, discussão e análise do mesmo texto durante uma unidade letiva; 3. Exploração do conteúdo e de aspetos linguísticos ou gramaticais; 4. Redação de uma síntese do texto lido como tarefa autónoma, fora do contexto da aula; 5. Após correção e revisão por parte da docente, devolução dos textos comentados aos estudantes.</p>	

Tabela 2. Tarefas de leitura-escrita realizadas com base em duas crónicas.

4.4 - Procedimentos de análise

4.4.1 - Análise quanto à Densidade de Ideias (DI)

Apesar de o estudo de Snowdon *et al.* (1999) recomendar a contagem das dez últimas frases das produções escritas, tal não foi possível por haver apenas 3 textos com 10 frases entre os 13 que foram entregues. Assim, selecionaram-se os textos que contavam pelo menos 6 frases por existirem em maior número (7). Com base em Brown *et al.* (2008), consideram-se proposições aquelas que derivam da classificação lexical de verbos, adjetivos, advérbios, sintagmas preposicionais e conjunções. Assim, os mesmos autores elencam de forma prática os elementos selecionados pela aplicação informática designada por CPIDR, sigla que corresponde a *Computerized Propositional Idea Density Rater*, em Brown *et al.* (2008): conjunções, numerais, determinantes (mas não artigos), preposições, adjetivos, advérbios, possessivos, relativos, interrogativos e verbos. As ideias, como explicam Snowdon *et al.* (1996) de forma sucinta, correspondem a proposições básicas selecionadas com base nos elementos Verbo, Adjetivo, Advérbio, Sintagma Preposicional e também Conjunções por marcarem proposições complexas de ligação temporal, causal ou outras. Assim, subordinou-se a análise à fórmula seguinte: $[N^{\circ} \text{ Ideias} \div N^{\circ} \text{ Palavras}] \times 10 = \text{DI}/10 \text{ palavras}$.

4.4.2 - Análise quanto à Complexidade Gramatical (CG)

Relativamente à medida de complexidade gramatical, com base na escala de Cheung e Kemper (1992) e em Pinto (1994), desenvolveu-se a escala de avaliação na Tabela 3 da qual consta a pontuação atribuída a cada tipo de frase. Há nesta escala uma pontuação que varia entre o zero e o cinco, dependendo do nível de complexidade atribuída a cada frase analisada.

Pontuação	Frases
0	Imperativas SVC SVO SVI SVS'
1	Causais, temporais, consecutivas e conclusivas Opositivas e restritivas comparativas
2	Infinitivas, Participiais
3	Subordinadas conjuncionais condicionais Relativas de encaixe à esquerda (antecedente em SU da oração subordinante)
4	Relativas de encaixe à esquerda (antecedente em SU da oração subordinante)
5	Subordinadas múltiplas

Tabela 3. Escala de complexidade gramatical em Português L2

4.5 - Resultados

Analisou-se a possível existência de uma correlação entre a DI e a CG em ambas as tarefas através do teste de correlação de Spearman. Encontrou-se uma correlação significativa entre variáveis relativas a tarefas de escrita diferentes e parâmetros de análise diferentes, cujos resultados estatisticamente significativos (a um nível de significância de 0,05) se apresentam na Tabela 4. Verifica-se uma associação significativa entre a DI encontrada nas últimas 6 frases da tarefa Síntese e a CG na tarefa Réplica. Constata-se ainda que existe uma correlação entre a CG da tarefa Síntese e a DI e a CG, ambas da tarefa Réplica.

Variáveis correlacionadas		Resultados ρ de Spearman
Densidade de ideias nas últimas 6 frases produzidas na tarefa de Síntese	Densidade de ideias nas últimas 6 frases na tarefa de Réplica (Writing Prompt)	$\rho = ,559^*$, $p = 0,038$
	Complexidade gramatical nas últimas 6 frases na tarefa de Réplica (Writing Prompt)	$\rho = ,606^*$, $p = 0,022$
Complexidade gramatical nas últimas 6 frases na tarefa de Síntese	Densidade de ideias nas últimas 6 frases produzidas na tarefa de Réplica (Writing Prompt)	$\rho = ,654^*$, $p = 0,011$
	Complexidade gramatical nas últimas 6 frases na tarefa de Réplica (Writing Prompt)	$\rho = ,584^*$, $p = 0,028$

Tabela 4. Associação entre os parâmetros de análise da produção escrita: densidade de ideias e complexidade gramatical

5 - Discussão e conclusões

Em primeiro lugar, verifica-se que a DI na Síntese apresenta uma associação com o mesmo parâmetro na Réplica. Assim, a possibilidade de encontrar diferenças de desempenho em tarefas de escrita com características diferentes não se verifica no campo da DI. Em geral, é possível afirmar que quem alcança uma pontuação mais alta no campo da DI na Síntese, atinge uma marca semelhante na Réplica quanto ao mesmo campo.

Não há, portanto, discrepância entre os resultados de uma tarefa mais condicionada por uma leitura anterior e outra sujeita apenas a um título coincidente com o do texto lido, mas mais livre. Pode dizer-se assim que, neste caso, a DI tende a ser estável porquanto se mantém em níveis semelhantes nas duas tarefas de escrita que são, por natureza, diferentes quanto ao grau de liberdade criativa tanto do ponto de vista conceptual como da perspetiva da evocação e da utilização de material linguístico.

Adicionalmente, verifica-se que existe uma associação estatisticamente significativa entre a CG na Réplica e o mesmo parâmetro da Síntese, como aliás acontece no caso da DI. Também no caso do grau de CG não existe variabilidade quanto ao tipo de texto produzido. Pode afirmar-se, conseqüentemente, que os participantes que revelam um desempenho alto do ponto de vista da CG das suas frases são constantes na qualidade que imprimem à sua produção linguística escrita deste ponto de vista.

Assim em resposta à questão metodológica sobre a possibilidade de desempenhos variáveis consoante as tarefas de escrita, fossem estas mais ou menos guiadas, no caso aqui em análise, não se pode concluir que exista uma diferença de desempenho na escrita da Réplica ou da Síntese propostas. Por outro

lado, confirma-se a associação encontrada por Kemper *et al.* (2001) entre CG e DI, na medida em que se verificou igual correlação em ambas as tarefas de escrita. A DI na Síntese revela uma associação estatisticamente significativa com a CG avaliada na Réplica, ao passo que a CG verificada na Síntese também se encontra estatisticamente associada à DI na Réplica, como se pode confirmar na Tabela 5.

Participantes	Cloze*	DI** Síntese	DI Réplica	CG*** Síntese	CG Réplica
1	22,58%	4,28	0	13	0
3	45,16%	3,65	4,53	19	9
9	0	4,40	0	3	0
11	0	4,81	5,80	17	11
12	0	5,16	4,23	9	13
13	29,03%	5,22	0	19	0
14	0	4,78	4,40	13	21

*. Percentagem de palavras corretas no teste Cloze.

** Média da Densidade de Ideias por frase.

***. Pontuação total atribuída à Complexidade Gramatical calculada para a tarefa de Síntese e para a tarefa de Réplica nas últimas 6 frases de cada texto.

Tabela 5. Número de respostas corretas no teste Cloze (tarefa de leitura); média de Densidade de Ideias e cálculo da Complexidade Gramatical na Síntese e na Réplica (tarefas de escrita)

Verificou-se uma associação significativa dos resultados encontrados na tarefa de compreensão leitora anterior, o teste *Cloze* (Lopes 2014; Lopes, Martins & Pinto 2015; Lopes 2020), com resultados de uma das tarefas de escrita. Os resultados do teste *Cloze* apresentaram uma correlação com a produção escrita no parâmetro que respeita a orações conjuncionais, no campo da CG tanto na tarefa Síntese como na tarefa Réplica. Significa esta associação que quanto mais palavras corretas tiverem sido produzidas no universo de 31 possíveis no teste *Cloze*, mais orações conjuncionais se encontram nos textos produzidos pelos participantes. Este resultado é independente do tipo de tarefa de escrita, na medida em que se verifica quer na tarefa Síntese, quer na tarefa Réplica. Pode então dizer-se que, no caso dos dados que estão aqui em causa, um melhor leitor, à luz do teste *Cloze* efetuado anteriormente, produz em concreto mais orações conjuncionais. É possível afirmar que quem compreender frases mais complexas, em virtude de ser capaz de as completar com material em falta, é também capaz de evocar recursos que lhe permitem construí-las.

Destes resultados, pode inferir-se que um texto que revela maior índice de estruturas gramaticais complexas, seguindo Kemper *et al.* (2001) na sua definição de complexidade gramatical, tende a estar também mais apetrechado do ponto de vista da precisão e da riqueza frásicas. Quanto ao autor dos textos e à sua capacidade de escrita, pode dizer-se que quem escreve com economia, precisão e riqueza é capaz de imprimir à sua produção a correção e a complexidade linguísticas que servem o sentido pretendido, tornando-se mais conciso sem

prejuízo para a codificação do pensamento.

Por outro lado, dando eco ao “better reader” de Stanovich (2000), questionou-se se este poderá escrever de forma mais proficiente, mesmo em L2. Nesse sentido, tratava-se de compreender que variáveis poderiam intervir de forma determinante para o sucesso do aprendente nas tarefas de leitura-escrita propostas.

De facto, pode dizer-se que um leitor mais fluente é capaz de tratar um texto a um nível que implica mais do que a decifração palavra a palavra ou a identificação de mecanismos gramaticais, mas integra estas duas funções da compreensão num ato mais elaborado de pensamento que re(cria) (como frisa Emig 1977; 1983) o sentido do texto à luz não só da agilidade linguística, mas também da sua experiência prévia. Quando Goodman *et al.* (1979) aludem ao conceito de biliteracia, o que permanece do seu postulado é esta via dupla de contributo para a compreensão do texto escrito em L2: o conhecimento dos instrumentos linguísticos ao serviço das ideias do leitor tanto por via da sua L1 como das L2s que utiliza e que podem partilhar com a L3 (Hammarberg 2001) em causa conhecimento metalinguístico, por exemplo, que o auxilie a tomar decisões de cariz mais estrutural na organização do texto. Aliás, como lembra Paradis (2009), a proficiência advém da fluência, isto é, da facilidade em aceder a elementos gramaticais que sustentem e alimentem a comunicação de pensamento, traduzida frequentemente numa compreensão ou produção aceleradas. Mas resulta também da precisão, uma vez que a fluência nem sempre denota uma aproximação aos termos mais adequados a cada situação. Não é raro que um falante de L2 se faça entender sem recorrer a linguagem adequada ao registo ou a vocabulário mais usual, por exemplo. A capacidade de atingir um patamar elevado de fluência e de precisão implica por sua vez o controlo sobre recursos atencionais e inibitórios que regulam a intromissão de itens linguísticos menos adequados, como também referem Bialystok (2001) e Bialystok *et al.* (2012). Este controlo não é sempre automático, no caso de um falante de uma L2. É, pelo contrário, intencional e construído pela prática até ser substituído por processos automáticos, segundo Paradis (2009).

De facto, pode afirmar-se que, no que se refere aos resultados obtidos, tanto a DI como a CG se acompanham e complementam independentemente das tarefas de produção escrita em causa, o que corrobora a integração sistemática de redes neuronais responsáveis pelo processamento da leitura e da escrita, com base na visão de Damasio e Damasio (2000), de Damasio e Tranel (1993) e sobretudo de Geschwind (1965; 1979). E confirma, adicionalmente, a pesquisa de Snowdon *et al.* (1996), Snowdon *et al.* (1999) e de Kemper *et al.* (2001).

No campo da variação interindividual, talvez as diferenças residam no grau de à vontade que os participantes sentiram neste contexto com a leitura em português L2, o que, com base em Stanovich (1986) pode derivar de um contacto variável com o registo escrito ao longo da vida tanto em L1 como nas várias L2s que utilizam.

É possível ainda que a relutância relativa à escrita, bem como à leitura, por parte de alguns participantes advenha da dificuldade em abstrair-se e adotar uma posição de leitor distanciado, ou “mock reader” a que aludem Gibson

(1950) e também Olson (1994). Esta dificuldade pode refletir um conhecimento metalinguístico insuficiente para fazer face às tarefas propostas nos casos dos estudantes que sentiram menos motivação e conforto na realização das tarefas.

De facto, constataram-se desempenhos diferentes que parecem decorrer do grau variável de exposição ao registo escrito e de níveis de competência linguística também variáveis. Os estudantes que obtiveram resultados mais elevados nos parâmetros DI e CG são também aqueles que revelam ler por hábito, gostar de ler e ainda ler em português². O confronto destes dados leva a crer que também a escrita beneficia de um efeito Mateus, como aquele que Stanovich (1986) descreve para a leitura, isto é, quem lê mais, não só lê melhor, mas também escreve melhor. Um desempenho deste tipo reforça o postulado de Emig (1983) quando a autora diz que se aprende a escrever sobretudo através da exposição à leitura e nem sempre através da instrução explícita. Isto significa que se pode questionar a utilidade de teorizar em torno da boa escrita em detrimento de fomentar oportunidades de escrita efetivas. É por ser um processo psicolinguístico recursivo (Pinto 2014) que a escrita lucra com a prática, sem esquecer o contacto com a leitura e com o que ela encerra de sugestivo, tanto do ponto de vista tácito, como metalinguístico. Como lembra Hirvela (2004), quem não contacta com determinados tipos de frase e de texto não os produz, perpetuando um ciclo de leitura-escrita pobre. Convirá, a fim de evitar um tal ciclo, que a abordagem a textos com uma ampla variedade de tipos e organizações frásicas seja uma preocupação do docente. Mas não só. Importa ainda fazer o que Hirvela define como “writerly reading” (2004), isto é, uma leitura analítica com vista a uma escrita informada. Preconiza-se por isso uma visão complementar da leitura-escrita que otimize o seu potencial recursivo e criativo através da influência recíproca que ambas exercem (Pinto 2014). Esta visão considera, naturalmente, o processo de leitura-escrita como a chave presciente relativamente ao produto escrito. É ainda fundamental expor os estudantes, independentemente de fatores individuais, a leituras que vão ao encontro dos seus interesses, mas que não se lhes sujeitem ao ater-se meramente ao que se supõe que já dominam.

² Os estudantes responderam a um inquérito sobre hábitos de leitura à entrada no curso (ver Lopes 2014 ou Lopes 2020).

REFERÊNCIAS

- Bernhardt, E. (2011). *Understanding Advanced Second Language Reading*. New York: Routledge, Taylor and Francis Group.
- Bialystok, E.; Craik, F.; Luk, G. (2012). Bilingualism: consequences for mind and brain. *Trends in Cognitive Science*. 16 (4): 240-250.
- Brown, C.; Snodgrass, T.; Kemper, S.J.; Herman, R.; Covington, M.A. (2008). Automatic Measurement of Propositional Idea Density from Part-of-Speech Tagging. *Behavior Research Method*. 40 (2):540-545.
- Bruner, J. (1966). *Toward a theory of instruction*. Cambridge, MA: The Belknap Press of Harvard University Press.
- Carson, J.E.; Carrell, P.L.; Silberstein, S.; Kroll, B.; Kuehn, P.A. (1990). Reading-writing relationships in first and second language. *TESOL Quarterly*. 24 (2). 245-266.
- Chand, V.; Baynes, K.; Bonnici, L.; Farias, S. (2012). *Analysis of Idea Density (AID): a manual*.
- Conselho da Europa (2001). *Quadro Europeu Comum de Referência para as Línguas*. Porto: Edições Asa.
- Damasio, A. (1989). Concepts in the brain. *Mind & Language*. 4 (1-2): 24-28.
- Damasio, A. & Damasio, H. (2000). Language and the brain. In Emmorey, K.; Lane, H. (Eds.) *The signs of language revisited. An anthology to honour Ursula Bellugi and Edward Klima*. New Jersey: Lawrence Earlbaum Associates, Publishers, 477-491.
- Damasio, A.; Tranel, D. (1993). Nouns and verbs are retrieved with differently distributed neural systems. *Proceedings of the National Academy of Sciences of the United States of America*. 90: 4957-4960.
- Geschwind, N. (1965). Disconnexion syndromes in animals and man. *Brain*, 88 (2): 237-294.
- Geschwind, N. (1979) Specializations of the human brain. *Scientific American*, 241 (3): 180-199.
- Goodman, K.; Goodman, Y.; Flores, B. (1979). Reading in the bilingual classroom: literacy and biliteracy. InterAmerica Research Associates, Inc.
- Emig, J. (1977). Writing as a mode of learning. *College Composition and Communication*. 28(2): 122-128.
- Emig, J. (1983). *The web of meaning*. Upper Montclair, N.J.: Boynton/ Cook Publishers, Inc.
- Gibson, E. (1998). Linguistic complexity: locality of syntactic dependencies. *Cognition*. 68: 1-76.
- Hammarberg, B. (2001). Roles of L1 and L2 in L3 production and acquisition. In Cenoz et al. (Eds.). 2001. *Cross-linguistic influence in third language acquisition: Psycholinguistic perspectives*. Bristol, UK: Multilingual Matters, 21-41.
- Hirvela, A. (2004). *Connecting reading and writing*. The University of Michigan Press.
- Kemper, S.; Sumner, A. (2001). The structure of verbal abilities in young and older

- adults. *Psychology and Aging*. 16(2): 312-322.
- Kemper, S., Greiner, L., Marquis, J. & Mitzner, T. (2001) Language decline across the life span: Findings from the Nun Study. *Psychology and Aging*, 16 (2): 227-239.
- Kemper, S., Herman, R.; Lian, C. (2003). Age differences in sentence production. *Journals of Gerontology: Psychological Sciences*, 58: 260-269.
- Kemper, S., Herman, R.E., Liu, C. J. (2004). Sentence production by younger and older adults in controlled contexts. *Journals of Gerontology: Psychological Sciences*, 58B: 220-224.
- Kemper, S.; Herman, R.E. (2006). Age differences in memory-load interference effects in syntactic processing. *The Journals of Gerontology Series B Psychological Sciences and Social Sciences*. 61(6): 327-332.
- Kintsch, W. (1974). *The representation of meaning in memory*. New York: Erlbaum Kintsch.
- Kintsch, W. (1998). *Comprehension. A paradigm for cognition*. Cambridge: Cambridge University Press.
- Koda, K. (2004). Reading and language learning: Crosslinguistic constraints on second language reading development. *Language Learning*. 57 (1): 1-44.
- Koda, K. (2010). *Insights into second language reading. A Cross-linguistic approach*. New York: Cambridge University Press.
- Leki, I. (1995). Good writing: I know it when I see it. In Belcher, D.; Braine G. (Eds.). *Academic writing in a second language: Essays on research and pedagogy*. Norwood. N.J.: Ablex Publishing Corporation, 23-46.
- Lobo, M. (2013). Subordinação adverbial. In Paiva Raposo, E.B; Nascimento, M.; Mota, M.; Segura, L.; Mendes, A.; Vicente, G.; Veloso, R. (Orgs.). *Gramática do Português*. Lisboa: Fundação Calouste Gulbenkian.
- Lopes, Â. F. (2014). *A compreensão leitora em português L2 nos níveis C1 e C2: contributos para a compreensão de sentido não literal*. Porto: Universidade do Porto, Faculdade de Letras. Dissertação de Mestrado.
- Lopes, Â. F. (2020). *O continuum leitura-escrita. Efeitos da compreensão leitora sobre a produção escrita em português L2 no nível C*. Porto: Universidade do Porto, Faculdade de Letras. Tese de Doutoramento.
- Lopes, Â.F.; Martins, F; Pinto, M. da G. L. C. (2015). Aprofundar a compreensão leitora em português L2 no nível avançado à luz do QECRL partindo do Teste Cloze. *Lingvarum Arena*. 6. 71-90.
- Luria, A.R. (1980). *Higher cortical functions in man*. New York: Basic Books, Inc., Publishers.
- Murray, D. M. (2013). *The craft of revision*. Boston: Wadsworth.
- OEDC (2016). *Skills matter: Further results from the survey of adult skills*. OEDC Skills Studies. Paris: OEDC Publishing.
- Olson, D. R. (1994). *The world on paper*. Cambridge: Cambridge University Press.
- Paradis, M. (2009) *Declarative and procedural determinants of second languages*. Amsterdam/ Philadelphia: John Benjamins Publishing Company.
- Pinker, S. (2014). *The sense of style. The thinking person's guide to writing in the 21st*

- century. UK: Penguin Books.
- Pinto, M. da G. L.C. (2014). *A escrita. O papel da universidade na sua otimização*. Porto: Faculdade de Letras da Universidade do Porto.
- Pinto, M. da G. L. C. (2017). Da revisão na escrita: uma gestão exigente requerida pela relação entre leitor, autor e texto escrito. *Revista Observatório Palmas*, 3 (4), 488-517.
- Sachs, O. (2017). *O Rio da consciência*. Lisboa: Relógio D'Água.
- Snowdon, D. (2001). *Aging with grace. The nun study and the science of old age: how we can all live longer, healthier and more vital lives*. London: Fourth Estate.
- Snowdon, D., Kemper, S. J., Mortimer, J. A., Greiner, L. H., Wekstein, D. R.; Markesbery, W. R. (1996). Linguistic ability in early life and cognitive function and Alzheimer's disease in late life: Findings from the Nun Study. *JAMA* 275, 528-532.
- Snowdon, D.; Greiner, L.; Kemper, S.; Nanayakkara, N.; Mortimer, J. (1999). Linguistic ability in early life and longevity: Findings from the Nun Study. In Robine, J.-M.; Forette, B; Franceschi, C.; Allard, M.. (Eds). *The paradoxes of longevity*. Heidelberg: Springer-Verlag, 528-532.
- Stanovich, K. E. (1986). Matthew effect in reading. Some consequences of individual differences in the acquisition of literacy. *Research Quarterly*. 21 (4): 360-407.
- Stanovich, K. E. (2000). *Progress in understanding reading: scientific foundations and new frontiers*. New York: The Guilford Press.
- Turner, A.; Greene, E. (1977). *The construction and use of a propositional text base*. Boulder: University of Colorado. Psychology Department.