



**NOVA FCSH**

FACULDADE DE CIÊNCIAS SOCIAIS E HUMANAS  
UNIVERSIDADE NOVA DE LISBOA

---

**PROPOSTA DE DICIONÁRIO DE ESPECIALIDADE APLICADO**

**À ANÁLISE SENSORIAL ENOLÓGICA**

**SÍLVIA PAIS BARBOSA**

**TESE DE DOUTORAMENTO EM LINGUÍSTICA**

**ESPECIALIDADE EM LEXICOLOGIA, LEXICOGRAFIA E TERMINOLOGIA**

**NOVEMBRO, 2023**

## DECLARAÇÃO

Declaro que esta tese é o resultado da minha investigação pessoal e independente. O seu conteúdo é original e todas as fontes consultadas estão devidamente mencionadas no texto, nas notas e na bibliografia.

A candidata, Sílvia Barbosa



Lisboa, 27 de novembro de 2023

Declaro que esta tese se encontra em condições de ser apreciada pelo júri a designar.

A Orientadora, Raquel Amaro



O Coorientador, William Martinez



Lisboa, 27 de novembro de 2023

## DECLARAÇÃO

Tese apresentada para cumprimento dos requisitos necessários à obtenção do grau de Doutor em Linguística (área de especialidade em Lexicologia, Lexicografia e Terminologia), realizada sob a orientação científica da Professora Doutora Raquel Amaro, Professora Auxiliar do Departamento de Linguística, NOVA-FCSH, e a coorientação do Professor Doutor William Martinez, SYLED - ILPGA, Sorbonne Nouvelle – Paris 3.

Apoio financeiro da Fundação para a Ciência e a Tecnologia (FCT) através da Bolsa de Doutoramento PD/BD/52263/2013, atribuída no âmbito do Programa de Doutoramento FCT Linguistics: Knowledge, Representation and Use (KRUse).

*In memoriam*

*Teresa Lino*

## AGRADECIMENTOS

À Professora Teresa Lino, pela oportunidade de investir neste tema, pelas palavras de incentivo em vários momentos, pelo seu papel significativo no CLUNL, e também fora dele, pela liberdade que sempre me deu para me dedicar a este projeto e por acreditar em mim, quando quase ninguém acreditou. Pelo incentivo de uma temporada em Aarhus e em Paris. Pelas muitas tardes de trabalho no projeto com direito a bolinhos e muitas risadas. Por me ter permitido conhece-la melhor e ter tornado a minha vida mais rica. Tenho muita pena que não veja o resultado final, mas acho que ia gostar da evolução da ideia que lhe apresentei numa tarde de uma Escola de Verão.

Ao Professor William Martinez, porque acreditou desde o começo na ideia de que em português faltavam trabalhos sobre o vinho. Pelas muitas horas em que falámos sobre a tese, pelos comentários pertinentes sobre análises, sobre programas de *software* e sobre a vida. Por me acolher durante os ataques de Paris e me fazer sentir segura. Pelos jantares, almoços, lanches e fins de semanas em Paris em modo família.

À Professora Raquel Amaro, por ter aceitado orientar-me a meio do percurso, numa fase difícil para mim depois da perda da Professora Teresa e durante uma pandemia. Pela disponibilidade sempre manifestada, pelos comentários rigorosos, pelas leituras atentas, pela paciência e pelo encorajamento ao longo do tempo. Obrigada por encontrar a ordem no meio do caos de tantos dados.

Agradeço o apoio financeiro da Fundação para a Ciência e a Tecnologia (FCT) através da Bolsa de Doutoramento PD/BD/52263/2013, atribuída no âmbito do Programa de Doutoramento FCT Linguistics: Knowledge, Representation and Use (KRUse), sem o qual não teria sido possível dedicar-me a este projeto, e que me permitiu as estadias na Dinamarca e em França e a participação em vários eventos académicos.

Aos meus colegas de artigos: Ana, Catarina, Chiara, Matilde, Mariana, Miguel, Susana.

A todos os meus informantes de questionários: aos amigos e amigos dos amigos, a quem fiz mil perguntas!

Aos colegas da Dinamarca: Mari-Liisa, Thomas, Alina, Parisa, Daniela, Deianira, Federico, Lili, Michail, Wafa ... *TAK and skål to everyone!*

Aos colegas de Paris: Anne-Sophie, Carole, Carolina, David, Oscar, Youngmun, Stephanie, Federico ... *merci a tous!*

Às colegas do LAPE: Isabel, Olga e Susana.

Às ilustres colegas do DPEAP: Fátima, Karima, Olga e Catarina.

Aos colegas da T3: Ana, Beatriz, Carolina, Inês, Kleiane, Joana, Mara, Margarida, Miguel, Radovan, Sandro, pelos muitos chocolates e bolos que partilhámos!

Às garotas de sempre da *Powerbase* pelo apoio constante e a todas as horas: Dani, Inês, Jiny, Renata e Sara.

Ao pessoal da MS: Joana, Hugo, Joana, Dave, Filipa, Mark, Lúcia, Ricardo, Mariana, Zé, Sandra, Hugo e Teixeira.

Às amigas de viagem académica ... Chiara e Mariana. Sem vocês, nada disto teria sido igual. Obrigada pelo apoio diário, pela paciência, pelas risadas e, acima de tudo, pela amizade que apenas pessoas que decidiram fazer tese na mesma altura entendem! Seguimos viagem, garotas!

Às amigas de sempre: Alexandra, Cátia, Sara, Catarina, Idelisa. Porque cada uma de vocês esteve sempre lá quando foi preciso lembrar-me de que havia vida além da tese. Agora vou cobrar tudo o que ficou pendente!

Aos meus pais e à minha irmã por me incentivarem. Por acreditarem em mim e por serem o meu porto de abrigo constante. ☺

Ao Jorge, por ser um exemplo de paciência, de companheirismo, de presença constante e de tudo o que eu precisava. Por acreditar que depois *disto* há muito mais para nós. ♥



E, no fim, apesar de dizerem que o doutoramento é um trabalho solitário ... sei que estive sempre rodeada por muitas pessoas que me ajudaram em muitos momentos a não perder o rumo e a provar muitos vinhos ☺ Tudo em nome da Ciência!

E para brindar com o simbolismo que a ocasião pede... sugiro um brinde aos presentes e aos ausentes!

Saúde!

## [RESUMO]

PALAVRAS-CHAVE: dicionário, lexicografia de especialidade, descritor, protótipo, degustação, vinho

Neste trabalho apresenta-se uma proposta de dicionário de especialidade sobre os descritores utilizados nas notas de prova de vinho. A ideia de apresentar uma metodologia de trabalho para elaborar um dicionário de especialidade visa atender às necessidades específicas dos especialistas, para promover uma comunicação eficaz e o correto uso do léxico especializado. Ao mesmo tempo, procura facilitar a comunicação com públicos menos especializados, como o público interessado e os consumidores de vinho, reconhecendo a diversidade do público-alvo que emerge a partir dos textos que serviram de base e foram utilizados nos *corpora*, abrangendo desde não-especialistas até especialistas. Para tal, foram estabelecidos sete objetivos: (i) Observar a organização e os conteúdos relevantes do domínio de especialidade – Análise Sensorial Enológica (ASE); (ii) Compilar um *corpus* de textos sobre a ASE; (iii) Analisar, numa perspetiva monolíngue, quais são os descritores atualmente em uso em Português Europeu (criação do *corpus* CEnoTEXT, com 21 495 textos especializados do tipo nota de prova, recolhidos entre 2012 e 2021, com um total de 780 585 *tokens*); (iv) Identificar, organizar, categorizar e caracterizar os descritores existentes a partir do CEnoTEXT; (v) Catalogar os descritores já estabelecidos no uso dos especialistas e que surgem no domínio de especialidade (criação do *corpus* CEnoLEX, com 26 obras dos seguintes tipos: glossários, dicionários e outros recursos existentes produzidos pelos especialistas, recolhidos entre 2012 e 2021); (vi) Apresentar uma proposta de dicionário, com pressupostos teóricos e metodológicos que têm em conta as especificidades das modalidades sensoriais invocadas, tendo em conta o trabalho de diversos autores que contribuíram no âmbito da Dicionarística; (vii) Fornecer uma base sólida para um recurso de especialidade útil a um público especialista e semiespecialista da ASE. A compilação de dois *corpora* permitiu, por um lado, observar as unidades e as suas definições atestadas e reconhecidas pelos especialistas, no CEnoLEX, e, por outro lado, verificar o uso das unidades lexicais utilizadas em contexto de especialidade real, no CEnoTEXT. A confrontação de ambos permitiu a comparação de uma abordagem partindo dos dados *versus* uma abordagem pela listagem de descritores dos especialistas, permitindo chegar a uma lista de candidatos a descritores e a um conjunto de reflexões sobre o material recolhido, selecionado e analisado. Por fim, é apresentada uma versão em papel e um protótipo de uma aplicação digital, no sentido de disponibilizar o recurso lexicográfico especializado e contribuir para a promoção e valorização da cultura enológica portuguesa.

## [ABSTRACT]

KEYWORDS: dictionary, specialized lexicography, wine descriptor, wine tasting, enological sensory analysis, wine

In this work, we present a proposal for a specialized dictionary on the descriptors used in wine tasting notes. The idea of introducing a working methodology to develop a specialized dictionary aims to meet the specific needs of experts, promoting effective communication and the correct usage of specialized lexicon. Simultaneously, it seeks to facilitate communication with less specialized audiences, such as the interested public and wine consumers, acknowledging the diversity of the target audience emerging from the texts that served as a basis and were used in the *corpora*, encompassing both non-specialists and experts. For this purpose, we delineated seven key goals: (i) Examination of the structure and pertinent content within the specialized domain of Enological Sensory Analysis (ASE); (ii) Compilation of a *corpus* comprising texts related to ASE; (iii) Analysis, from a monolingual perspective, of the currently employed descriptors in European Portuguese, culminating in the creation of the CEnoTEXT *corpus* (this *corpus* comprises 21,495 specialized texts of the tasting note genre, collected between 2012 and 2021, with a total of 780,585 tokens); (iv) Identification, organization, categorization, and characterization of the existing descriptors derived from CEnoTEXT; (v) Compilation and categorization of descriptors already established in expert usage and emerging in the specialized field (creation of the CEnoLEX *corpus*, comprising 26 works such as glossaries, dictionaries, and other existing resources produced by experts, collected between 2012 and 2021); (vi) Presentation of a dictionary proposal, considering theoretical and methodological assumptions that take into account the specificities of the sensory modalities invoked, based on the work of various authors contributing in the field of lexicography; (vii) Provision of a solid foundation for a specialized resource beneficial to both expert and semi-expert audiences within the ASE domain. The compilation of two *corpora* facilitated the examination of units and their definitions attested and acknowledged by experts in CEnoLEX, as well as the examination of lexical unit usage within genuine specialized contexts in CEnoTEXT. This comparative analysis allowed for the juxtaposition of a data-driven approach against an approach centered on listing expert descriptors, leading to a list of descriptor candidates and a set of reflections on the collected, selected, and analyzed material. Lastly, we present a printed version along with a prototype of a digital application to render this specialized lexicographic resource accessible, thereby contributing to the promotion and valorization of Portuguese wine culture.

<b>1.</b>	<b>Introdução</b> .....	1
1.1.	Objetivos .....	6
1.2.	Metodologia Geral.....	8
1.3.	Estrutura da tese .....	9
<b>2.</b>	<b>Enquadramento Teórico</b> .....	11
2.1.	Lexicografia de Especialidade .....	11
2.1.1.	Léxico, Lexicologia, lexicultura.....	11
2.1.2.	Lexicografia .....	18
2.1.2.1.	Contexto histórico europeu .....	21
2.1.2.2.	Contexto histórico português.....	24
2.1.3.	Língua e recursos de especialidade .....	27
2.1.3.1.	Língua e léxico de especialidade.....	30
2.1.3.2.	Sobre o <i>descriptor</i> .....	35
2.1.4.	Dicionário.....	37
2.1.4.1.	Macroestrutura .....	42
2.1.4.2.	Microestrutura .....	48
2.1.4.3.	Tipos de dicionários .....	51
2.1.5.	Dicionário de especialidade .....	57
2.2.	Domínio da ASE .....	64
2.2.1.	Multidisciplinariedade do domínio ASE.....	65
2.2.2.	Intervenientes, nível de especialização e partilha de conhecimento .....	66
2.2.3.	Recursos e construção do conhecimento do domínio .....	71
2.2.4.	Notas de prova.....	73
2.3.	Recursos lexicais na área da enologia e ASE.....	77
<b>3.</b>	<b>Análise Sensorial Enológica</b> .....	79
3.1.	Sentidos e Perceção Sensoriais .....	82
3.1.1.	A visão, o olfato, a audição, o tato e o paladar.....	84
3.1.2.	Sensação e Perceção.....	85

3.2.	Processo da Análise Sensorial Enológica.....	88
3.2.1.	Exame Visual .....	93
3.2.2.	Exame Olfativo .....	96
3.2.3.	Exame Gustativo .....	99
3.2.4.	Exame Final .....	102
3.2.5.	Testes e métodos da ASE.....	105
<b>4.</b>	<b>Apresentação e Parametrização dos Dados .....</b>	<b>107</b>
4.1.	Base de Dados Enológicos .....	108
4.2.	<i>Corpora</i> .....	113
4.2.1.	Tratamento semiautomático .....	114
4.2.2.	<i>Corpus</i> Enológico Lexicográfico .....	116
4.2.2.1.	Fontes lexicográficas e análise preliminar .....	116
4.2.3.	<i>Corpus</i> Enológico Textual .....	129
4.2.3.1.	Fontes textuais e análise preliminar .....	130
4.2.3.1.1.	CENoTEXT em números.....	140
<b>5.</b>	<b>Léxico de especialidade da ASE: análise e tratamento .....</b>	<b>142</b>
5.1.	Seleção e tratamento de Unidades Lexicais de Especialidade da ASE.....	142
5.2.	Análise Categorial.....	148
5.2.1.	Verbos .....	148
5.2.2.	Nomes .....	161
5.2.3.	Adjetivos .....	177
5.2.4.	Empréstimos.....	189
5.2.5.	Síntese Análise Categorial .....	192
5.3.	Análise Contrastiva .....	193
5.3.1.	Verbos .....	194
5.3.2.	Nomes .....	196
5.3.3.	Adjetivos .....	198
5.3.4.	Empréstimos.....	204
5.3.5.	Síntese da Análise Contrastiva.....	204
5.4.	Análise por Modalidade Sensorial .....	205

5.4.1. Coocorrências.....	208
5.4.2. Análise Fatorial .....	217
5.4.2.1. Exemplo de análise de candidatos verbais .....	224
5.4.2.2. Exemplo de análise de candidatos à modalidade visual .....	229
5.4.3. Síntese da Análise por Modalidade Sensorial .....	231
5.4.4. Considerações finais.....	232
<b>6. Proposta do dicionário .....</b>	<b>235</b>
6.1. Objetivos e público-alvo .....	238
6.1.1. Macroestrutura .....	239
6.1.2. Microestrutura .....	242
6.2. Protótipo de uma Aplicação Digital – APP DEASE .....	252
6.2.1. UX Design da APP DEASE.....	255
6.2.1.1. Empatia: Inquérito aos utilizadores.....	255
6.2.1.2. Definição: Análise competitiva e utilizadores prototípicos.....	258
6.2.1.3. Idealização: Exploração do sistema crazy 8.....	262
6.2.1.4. Protótipo: Protótipo da APP DEASE .....	264
6.2.2. Aplicações futuras .....	269
<b>7. Conclusões .....</b>	<b>273</b>
7.1. Objetivos e resultados alcançados.....	275
7.2. Desafios e trabalho futuro .....	278
7.3. Considerações finais.....	284
<b>8. Referências Bibliográficas .....</b>	<b>286</b>
<b>9. Anexo 1 .....</b>	<b>314</b>
<b>Índice de Figuras .....</b>	<b>319</b>
<b>Índice de Tabelas .....</b>	<b>321</b>

## 1. INTRODUÇÃO

O vinho existe desde os primórdios da civilização e a sua origem remonta a cerca de 6 000 a.C., apontando-se para o período Neolítico, de 5 400 a 5 000 a.C. na Cordilheira de Zagros, entre o Irão e o Iraque (McGovern, Fleming, & Katz, 2003)<sup>1</sup>.

O vinho é definido como uma bebida alcoólica produzida a partir da fermentação da uva. O processo de vinificação é bastante complexo e dependente de fatores ambientais, como a composição e a qualidade da matéria-prima, da sua flora microbiana, a região da produção, a cultura da vinha, as castas e os processos de produção do vinho como a fermentação alcoólica e maloláctica (que, quando não controlada, pode originar alterações não desejadas no produto final), que são fatores que influenciam as características do vinho.

A história do vinho é muito antiga, dizendo respeito a uma bebida produzida e ingerida por razões sociais e culturais distintas, como veículo de espiritualidade, como objeto de prazer que é saboreado e desfrutado e, ainda, como símbolo de algo proibido (Lukacs, 2013). A Análise Sensorial (doravante AS) é uma função primária da Humanidade que a tem auxiliado na definição de preferências alimentares, motivando também atividades como o comércio – recorde-se a procura das especiarias, da cana-de-açúcar, da batata, do cacau, entre outros, na época dos Descobrimentos – e até o surgimento de profissões para controlar e avaliar produtos, baseando-se na sua prova para avaliar e ajustar o preço dos mesmos como, por exemplo, os juízes de gado bovino leiteiro para o leite, os mestres-cervejeiros para a cerveja e os escanções para o vinho.

Sendo o vinho uma das bebidas alcoólicas mais consumidas no Mundo, assistiu-se a uma evolução na produção de vinho, acompanhada de uma necessidade de melhorar a sua qualidade. Com esse objetivo, foram criadas “análises físicas, químicas,

---

<sup>1</sup> Esta é uma visão ocidental partilhada da descoberta do vinho, apesar de estudos recentes como Norrie (2019) que refere a existência de vestígios de vinho recolhidas na zona de Jiahu no Nordeste da China que datam de pelo menos 500 anos antes dos registos europeus, pelo facto de a olaria na China ser uma atividade mais antiga do que na Europa e poder ter preservado amostras de uma bebida fermentada (Norrie, 2019:6).

microbiológicas e sensoriais que permitem uma identificação rápida e atempada de defeitos ou problemas e auxiliam na melhoria e manutenção da qualidade dos vinhos” (Rosado, 2013) e que, tipicamente, se relacionam com a prova de vinho ou ocasiões sociais de degustação de vinho, configurando o que se entende por Análise Sensorial Enológica (ASE).

A prova de vinho começa com uma inspeção visual do vinho (exame visual), seguida de uma apreciação do cheiro (exame olfativo), da textura, do sabor, do gosto e do retrogosto (exame gustativo), que resulta diretamente numa sensação agradável ou desagradável. A perceção sensorial, ativada pela utilização dos sentidos da visão, olfato, audição, paladar e tato, transmite ao provador-especialista as várias sensações resultantes dos vários estímulos proporcionados pelo consumo de vinho. A AS é, por sua vez, “usada para evocar, medir, analisar e interpretar as reações às características dos alimentos e materiais tal como são percebidas pelos sentidos da visão, olfato, paladar, tato e audição” (IFT, 1975).

O ato de degustar<sup>2</sup> um vinho pode ou não desencadear a verbalização oral/escrita da experiência ocorrida. Em caso positivo, observa-se um ajuste daquilo que foi percecionado pelo indivíduo e uma posterior adequação do que será proferido para ser compreendido pelo outro (uma “recontextualização” da perceção sensorial). Esta verbalização apoia-se em unidades lexicais de especialidade – os descritores<sup>3</sup> –, que visam permitir a comunicação das sensações provocadas pelo vinho.

---

<sup>2</sup> “Quem degusta um vinho deve ser capaz de comunicar as suas impressões de determinado vinho de modo claro e compreensível. Assim, para que todos falem a mesma língua, reuniu-se em tabela termos considerados indispensáveis, formando um vocabulário de base de um degustador. Esta tabela permite descrever e julgar tecnicamente um vinho de modo simples e linear. Os significados desses termos mudam e tornam-se qualitativos quando não se referem à simples descrição do vinho, mas quando exprimem um juízo de valor” (retirado de <https://enochataprofissional.wordpress.com/2010/11/14/terminologia-ais/>).

<sup>3</sup> A decisão de seleccionar *descriptor* ao invés de *termo*, *atributo* ou outra designação, prende-se essencialmente, com o facto de neste trabalho, *descriptor* ser entendido como: (i) uma unidade que pode pertencer à língua de especialidade e à língua corrente; (ii) uma unidade conhecida dos utilizadores e que surge nas obras de referência do domínio em questão; (iii) uma unidade que não é necessariamente monorreferencial (Larivière 1989:459 *apud* Contente, 2008:41) e, também, a forma como os especialistas designam as unidades lexicais de especialidade – o *descriptor*.

Vários estudos revelam que a verbalização da percepção sensorial não é fácil. Diferentes autores (Croijmans & Majid (2016), Swiegers, Chambers & Pretorius (2005), entre outros) referem que há uma maior facilidade na descrição da cor do que na descrição do aroma e do sabor, por exemplo. Afirmam também que, para o exame visual é usado, regra geral, um número restrito e reduzido de descritores, enquanto os descritores utilizados nos exames olfativo e gustativo são em maior número, com recurso a mais metáforas e/ou analogias de difícil compreensão imediata. Ao que tudo indica, os exames olfativo e gustativo são de uma maior complexidade e, como tal, de menor facilidade de descrição, como referem Morrot, Brochet e Dubourdieu (2001), por exemplo.

A verbalização da degustação do vinho pode refletir-se em pelo menos duas formas: (i) as notas de prova que estão presentes em revistas da especialidade e em *websites* dos produtores e distribuidores de vinho, e (ii) os rótulos disponíveis nas garrafas de vinho. Em ambos os casos, está-se perante pequenos textos, em regra com os seguintes campos:

- a designação oficial do vinho;
- o produtor;
- a denominação de origem;
- a(s) casta(s);
- a vinificação;
- o teor alcoólico;
- a nota de prova;
- a cave;
- o serviço;
- o(s) enólogo(s).

Ao especialista<sup>4</sup>, este texto permite adquirir informação sobre como melhorar as características da vinha ou decidir comercializar ou não um vinho, bem como descrevê-lo aos restantes interessados/consumidores. Ao não-especialista, este texto possibilita um maior conhecimento do vinho que irá degustar, através da partilha da informação de especialidade.

Qualquer domínio de especialidade, compreendendo atividades e entidades empresariais de uma área, utiliza um léxico particular que transmite o conhecimento próprio do domínio, os seus princípios e como os atores/agentes deste domínio se comportam e comunicam. O desconhecimento do léxico usado gera problemas de diferentes ordens, tais como erros de comunicação ou lapsos no desenrolar de tarefas e atividades do domínio. Para evitar essa situação, é comum investir-se na formação dos vários profissionais e intervenientes da área, para potenciar o domínio de um léxico partilhado que diminui o ruído no processo de comunicação, facilitando a partilha e a transmissão inequívoca do conhecimento por parte dos seus utilizadores.

No entanto, como Johnson e Robinson (2007) referem, “Learning to appreciate wine is one thing; communicating about it is another thing entirely (...)”. Nesse sentido, para falar sobre degustação do vinho (e da respetiva perceção sensorial) com clareza e precisão, é necessário dominar, e logo identificar, descrever e categorizar um léxico especializado.

Numa análise exploratória relativamente ao léxico do domínio do vinho no panorama nacional, verificou-se a existência de um texto de referência multilingue (Dalmasso, 1963) utilizado pela *Organisation Internationale de la Vigne et du Vin*<sup>5</sup>, datado de 1963, com uma versão em Português Europeu (doravante PE), de que não existe qualquer versão atualizada. No *Instituto da Vinha e do Vinho* e na *Associação Portuguesa*

---

<sup>4</sup> A decisão de escolher *especialista* ao invés de *enólogo*, *sommelier*, *degustador*, *técnico*, *crítico de vinhos*, *providor*, *produtor de vinhos* ou outra referência, deve-se ao facto de neste trabalho, *especialista* ser entendido como uma designação abrangente dos diversos protagonistas que atuam no domínio da Enologia e que desempenham funções relevantes na área da ASE ou que a utilizam regularmente nas suas tarefas.

<sup>5</sup> Internacionalmente, é a *Organisation Internationale de la Vigne et du Vin* que tem a missão de disponibilizar materiais de referência, tais como dicionários, glossários, terminologias e manuais; a nível nacional, essa tarefa cabe ao *Instituto da Vinha e do Vinho*.

*de Enologia* não foram encontradas quaisquer fontes de referência atualizadas, ao contrário de outros trabalhos desenvolvidos no panorama internacional, que foram sendo atualizados e validados pelos especialistas: Vedel, Chamay e Tourmeau (1972) e Peynaud (1980) para o francês, Lehrer (1975, 2009) e Parker (1998) para o inglês, e López-Arroyo e Roberts (2014) para o espanhol (para citar alguns exemplos).

Em Portugal, o que se verifica é a existência de uma multiplicidade de fontes como dicionários, glossários, manuais, guias, criadas, por vezes, por especialistas da área da enologia (enólogos, jornalistas de especialidade, entre outros), mas com pouca sensibilidade no que respeita à configuração desses mesmos recursos quanto a aspetos como a organização e conteúdo informacional, do ponto de vista linguístico e, em particular, do ponto de vista da Lexicografia de Especialidade e da dicionarística.

A observação da existente escassez de estudos, por um lado, e da profusão de recursos heterogêneos quanto ao conteúdo, existentes no mercado nacional, por outro, levou a considerar-se pertinente a apresentação de uma proposta de um dicionário nesta área de especialidade. A proposta de dicionário que se apresenta neste trabalho teve por base alguns aspetos motivadores, descritos de seguida.

Em primeiro lugar, a constatação de que a enologia em Portugal é uma área em expansão, que movimenta um elevado valor de negócios acima de vários milhões de euros, com inúmeros eventos ligados a esta temática, cursos e *workshops* fornecidos por diferentes entidades, o que, regra geral, é uma situação associada a um cenário de heterogeneidade de textos e problemas de comunicação.

Em segundo lugar, a evidência de que a ASE é uma ferramenta que permite a descrição e avaliação de bebidas, mas que revela problemas de ordem linguística, no que respeita à verbalização e transmissão de conhecimento. Apesar da existência de vários especialistas e de textos especializados, bem como de *websites* e blogues associados a esse domínio de especialidade, que espelham a dinâmica da área e a tentativa de os profissionais a difundirem, há uma lacuna importante: não existe um repositório/base de dados/recurso que, para além de cumprir requisitos de ordem lexicológica e lexicográfica, reúna o consenso dos intervenientes. Tal facto provoca problemas de comunicação entre especialistas e entre estes e o público-alvo a que pretendem chegar.

A construção de uma proposta de dicionário de especialidade da ASE ajuda, assim, não só especialistas (e futuros especialistas) da área, mas também um público não-especialista, mas curioso e com desejo de ser informado e conhecedor, a ter acesso a um conhecimento organizado e descrito de forma sistemática, lógica e coerente.

O objetivo deste dicionário será fornecer uma ferramenta de consulta para uma comunidade de utilizadores com elevado grau de especialização, como, por exemplo, os enólogos, os escanções, os estudantes (e futuros especialistas destas áreas), mas também poderá funcionar como um contributo para dar respostas a necessidades concretas de outros especialistas, como tradutores, jornalistas e enófilos, que necessitem de uma maior eficácia comunicativa e de conhecer e usar corretamente este léxico de especialidade no decorrer das suas funções/tarefas.

Em suma, foram definidos os seguintes aspetos:

- O domínio de especialidade: ASE;
- O tema: o léxico de especialidade da ASE em PE;
- O tipo de recurso: dicionário monolíngue de especialidade;
- Os destinatários do dicionário: especialistas (e futuros especialistas) do domínio (estudantes, escanções, enófilos, tradutores, jornalistas, etc.), mas também semiespecialistas (público interessado no tema e consumidor de vinho);
- Os objetivos: apresentar dados de referência que atestassem o atual estado da ASE para o PE e propor uma classificação sistemática do domínio da ASE, para auxiliar a redação técnica e/ou a tradução de textos especializados, retificar o uso de unidades inadequadas, e esclarecer sobre a forma e o conteúdo dessas mesmas unidades lexicais.

## 1.1. OBJETIVOS

Na sequência do descrito acima, a presente tese de doutoramento tem como objeto de estudo a Língua de Especialidade da ASE e pretende dar um contributo de carácter

lexicológico e lexicográfico para um melhor entendimento e compreensão sobre este domínio e este léxico de especialidade.

A investigação que aqui se apresenta tem como objetivos:

- observar a organização e os conteúdos relevantes do domínio de especialidade – ASE;
- compilar um *corpus* resultante de uma minuciosa pesquisa de textos relativos à ASE;
- analisar, numa perspetiva monolíngue, quais são os descritores atualmente em uso em PE;
- identificar, organizar, categorizar e caracterizar os descritores existentes referentes à ASE;
- catalogar os descritores já estabelecidos no uso dos especialistas e os que surgem no domínio de especialidade;
- apresentar uma proposta de dicionário, com pressupostos teóricos e metodológicos que privilegiam uma abordagem abrangente tendo em conta a especificidade de modalidades sensoriais<sup>6</sup> invocadas;
- fornecer uma base sólida para um recurso de especialidade útil a um público especialista e semiespecialista da área da ASE.

Como objetivo tangível global, pretende-se apresentar uma proposta de dicionário relativo aos descritores da ASE, que constitui um tipo de língua especializada em Enologia, com características próprias e realizadas por especialistas. Deste modo, pretende-se permitir o desenvolvimento de uma linguagem descritiva objetiva e consensual com maior grau de concordância de julgamentos entre os especialistas, visto que estes têm um papel determinante (funcionam como “instrumentos de medida”

---

<sup>6</sup> Entende-se por modalidade sensorial as sensações percebidas pelos sentidos e que de acordo com Viberg (1983) se dividem por modalidade visual, auditiva, olfativa, tátil, gustativa.

calibrados e controlados apenas pela especificidade da linguagem utilizada (Alvelos, 2002)), contribuindo de modo objetivo e inovador para alguma normalização e estabilização do léxico usado, e servindo de instrumento de trabalho na aprendizagem da área, diminuindo o ruído de comunicação entre aqueles que partilham esta língua de especialidade em PE.

## 1.2. METODOLOGIA GERAL

Para alcançar os objetivos supracitados e apresentar uma proposta de modelo de dicionário de especialidade para o domínio da ASE, foi necessário seguir um conjunto geral de procedimentos, como se apresenta de seguida.

Num primeiro momento, foi feito um estado de arte do trabalho relativo à Lexicografia em Portugal, com destaque para o dicionário e a sua estrutura, e como a Linguística de *Corpus* pode ser utilizada para melhorar as obras lexicográficas. Paralelamente, para compreender o domínio do vinho e da ASE, elaborou-se uma pesquisa bibliográfica com o objetivo de identificar e selecionar obras para construção de uma Base de Dados Enológicos. Constatou-se que, apesar dos vários trabalhos elaborados, são poucos os trabalhos de referência desenvolvidos no âmbito do Português Europeu e os mesmos, na maioria, não consideram a Lexicografia de Especialidade na elaboração das obras lexicográficas disponíveis ao público. Considerou-se que havia motivação para elaborar uma proposta de dicionário de especialidade para este domínio – ASE.

Optou-se por se criar uma Base de Dados Enológicos como ponto de partida, para ter uma noção da pesquisa bibliográfica das obras relativas à ASE. Posteriormente, e tendo em conta a tipologia de textos disponível, decidiu-se selecionar dois subconjuntos de dados, o *corpus* lexicográfico, relativo a obras lexicográficas existentes produzidas por especialistas, e o *corpus* linguístico relativo aos textos das notas de prova produzidas pelos especialistas. Ambos com características distintas, mas complementares, para permitir a identificação e análise de um léxico de especialidade da ASE.

Após o tratamento das unidades lexicais e após a identificação das unidades lexicais especializadas da ASE, foi feita uma análise e proposto um modelo de dicionário. Foi ainda elaborado um protótipo de aplicação digital, construído com os princípios do *User Experience Design*, para visualização da proposta.

### 1.3. ESTRUTURA DA TESE

O presente trabalho de investigação, para além da Introdução e das Considerações Finais, integra os seguintes capítulos:

O capítulo *Enquadramento Teórico* apresenta os pressupostos teóricos da investigação realizada, o estado da arte no que respeita à Lexicografia de Especialidade, bem como as linhas orientadoras que fundamentam e justificam as opções metodológicas tomadas e apresentadas nos capítulos subsequentes.

O capítulo *Análise Sensorial Enológica* explora a AS, e em concreto a ASE, detalhando: (i) que sentidos são invocados aquando da experiência do ato de degustar um vinho, (ii) o que se entende como diferença entre perceção e sensação, (iii) como se processa o processo da degustação e quais os exames efetuados (visual, olfativo e gustativo), bem como (iv) que tipo de testes (paramétricos e estatísticos) podem ser efetuados para além do teste de verbalização das sensações provocadas pelo vinho de forma escrita e/ou oral.

O capítulo *Apresentação e Parametrização dos Dados* explicita as opções metodológicas tomadas relativamente ao domínio de especialidade escolhido, apresenta a Base de Dados Enológicos (BDEno), o *Corpus Enológico Textual* (CEnoTEXT) relativo às notas de prova produzidas pelos especialistas, e o *Corpus Enológico Lexicográfico* (CEnoLEX), relativo aos glossários, dicionários e outros recursos existentes produzidos pelos especialistas.

O capítulo *Léxico de Especialidade da ASE: Análise e Tratamento* apresenta o número de candidatos a unidades lexicais de especialidade, a sua distribuição, bem como

o conjunto de análises que foram alvo: análise categorial, análise contrastiva e análise sensorial. O resultado dessas análises permite a identificação, a organização dos mesmos.

O capítulo *Proposta de Dicionário de Especialidade da ASE* explicita o dicionário de especialidade idealizado para apresentar os dados recolhidos sobre este domínio de especialidade, tendo em conta o trabalho de diversos autores que contribuíram de forma significativa para a sua definição e desenvolvimento no âmbito da Dicionarística. Apresenta-se uma proposta para uma versão em papel e um protótipo de uma aplicação digital.

## 2. ENQUADRAMENTO TEÓRICO

Neste capítulo, procede-se ao enquadramento teórico geral da investigação levada a cabo, apresentando-se os princípios teóricos e metodológicos da Lexicografia de Especialidade em que se insere este trabalho, com a exposição dos contributos de vários autores que servem de fundamento para o dicionário apresentado, e refletindo-se sobre as questões de definição da disciplina e das suas delimitações. O objetivo desta secção é contextualizar teoricamente a proposta de criação do *Dicionário de Especialidade da Análise Sensorial Enológica – DEASE* –, uma ferramenta que permite delimitar, organizar e gerir o conhecimento específico desta área de especialidade.

Refletindo o carácter necessariamente multidisciplinar que enforma o estudo lexical de uma área de especialidade, este capítulo contempla dois domínios distintos: a Lexicografia de Especialidade, enquanto subdomínio específico da Linguística, objeto da primeira subsecção, e a ASE, enquanto parte essencial da enologia contemporânea, que agrega conceitos e práticas de múltiplas disciplinas, tratada na subsecção 2.2. A terceira subsecção apresenta o estado da arte relativamente aos recursos lexicais na área da ASE, apresentados na subsecção 2.3.

### 2.1. LEXICOGRAFIA DE ESPECIALIDADE

#### 2.1.1. LÉXICO, LEXICOLOGIA, LEXICULTURA

A Linguística é a ciência que tem por objeto de estudo a faculdade da linguagem e as línguas naturais, nas suas múltiplas dimensões. Crystal (2001) define Linguística como “[the] scientific study of language, [which] has generated a wide range of branches, subdisciplines, schools of thought, and applications.” (Crystal, 2001:8948).

A língua enquanto objeto da Linguística foi definido primeiramente por Saussure (1911), no contexto de uma distinção entre *langue* (língua) e *parole* (fala). *Langue* seria, então, nesta aceção, o sistema gramatical interno e partilhado entre os falantes de uma língua, e *parole*, associado ao desempenho do indivíduo, o produto real do sistema.

Também é a Saussure que se deve a noção de língua como um sistema de signos. O signo, que nos séculos XVII e XVIII designava apenas a componente fonológica das línguas, tem, para Saussure, duas componentes — a fonológica e a conceptual. O signo é assim uma unidade linguística dupla, com significante e significado, cuja relação estabelecida é arbitrária.

Tendo como base esta noção, a língua pode ser definida como um sistema que existe na consciência das pessoas e que não depende do indivíduo que a fala, mas das normas estabelecidas e partilhadas por uma comunidade linguística, a partir da qual os indivíduos a adquirem.

No contexto deste vasto estudo, a Lexicologia é a disciplina da Linguística que tem como objetivo o estudo científico do léxico das línguas, sendo também considerada a ciência que estuda e representa as relações sintagmáticas e paradigmáticas desse sistema (Gaudin & Guespin, 2000:161). Autores como Vilela (cit. por Silva, 2003:17) destacam que a Lexicologia estuda as palavras de uma língua em diversas vertentes, incluindo a etimologia, a fonologia, a morfologia, a sintaxe, a neologia, preservando necessariamente uma estreita relação com a semântica e fornecendo os pressupostos teóricos que permitem a sua análise individual (palavra/unidade lexical) e coletiva (léxico). A Lexicologia, além de se ocupar de problemas teóricos, procura fazer o levantamento, a organização, a exploração exaustiva e a análise da estrutura interna do léxico, bem como das relações e inter-relações das unidades lexicais. Genouvier e Peytard (1973:351-352) asseguram ainda que a Lexicologia, para além de ter um método próprio no que respeita ao estudo científico do léxico, possui ainda duas dimensões: a Lexicologia descritiva (a Lexicologia no domínio da descrição, que se situa numa relação entre a gramática e a semântica) e a Lexicologia aplicada (a Lexicologia no domínio da aplicação relativamente à Lexicografia).

Enquanto parte essencial das línguas, o léxico é também um elemento principal para que sejam compreendidas as mudanças linguísticas de uma comunidade. Neste contexto, as palavras não têm existência real se não estiverem contextualizadas, i.e. relacionadas, com outras palavras e com a cultura da comunidade falante, com as quais estabelecem uma rede dinâmica, contínua e natural, que reflete a fluidez de usos e sentidos, e o

surgimento e desaparecimento de palavras, mudanças próprias do dinamismo natural de uma língua. O léxico não é caótico (como mencionam Carter & McCarthy, 2013:49) nem entendido como sendo rígido, transparente ou unívoco, mas sim flexível, opaco e polissémico, passível de ser organizado e estruturado, necessitando o falante de um processo de aprendizagem dos seus contextos e dos seus usos.

O léxico define-se, então, como o conjunto de unidades lexicais de uma língua que se subdivide em, ou contempla, unidades lexicais da língua corrente e unidades lexicais de língua(s) de especialidade, refletindo domínios de um determinado conhecimento.

Um aspeto que importa referir quando se menciona o léxico é o conceito de “lexicultura” (Galisson, 1970), no qual as unidades lexicais apresentam uma dimensão dual: a dimensão linguística e a dimensão cultural (Guillén Díaz, 2003). O facto de as unidades lexicais possuírem um conjunto de características com uma carga cultural partilhada (Galisson, 1970) leva a que sejam reconhecidas e aceites/utilizadas por um conjunto de falantes de um determinado grupo.

“La culture en dépôt dans ou sous certains mots, dits culturels, qu’il convient de repérer, d’explicitier et d’interpréter. La démarche consiste à mettre au jour des sites culturels [...]. Les expressions imagées, les mots-valises, les mots à charge culturelle partagée, les palimpsestes verbo-culturels, les mots de situations, les noms de marques, les proverbes et dictons, les mots occultants [...] circonscrivent les sites (ou gisements).” (Galisson, 1999:480).

Define-se a carga cultural partilhada como uma unidade que possui “la quantité de culture, c’est-à-dire d’éléments culturels propre à une certaine civilisation, à un certain pays, à une certaine région, que véhicule ce mot et qui font que ce mot n’est pas immédiatement transparent pour un locuteur/auditeur ne connaissant pas cette culture”. (Van Baardewijk-Rességuier (1993:29) *apud* Ligas (2012:14)). De acordo com Barbosa (2009:35-36), para verificar se se está perante uma unidade lexical com carga cultural partilhada, esta deve ter um conjunto de características que explicam de que modo a carga “adicional” cultural se agrega à unidade lexical, nomeadamente, ser resultado de uma

partilha de significado pela comunidade de falantes, que surge veiculada a um aspeto cultural e partilhado pela comunidade.

A noção de que a língua espelha a representação do modo como os falantes observam o mundo, aspeto já mencionado por Wittgenstein (1961 [1921]) “The limits of my language mean the limits of my world” e se relacionam nas (e com as) comunidades linguísticas é uma forma de também considerar o léxico como um património que condensa em si o uso e a representação da cultura onde os falantes comunicam entre si (Isquierdo, 1996:191). Por exemplo, a relação língua, cultura e ensino de língua estrangeira originou vários modelos pedagógicos com preocupações com as competências interculturais nas ciências sociais. Quando se estuda a didática das línguas e das culturas, a lexicultura favorece a introdução da cultura pelo léxico e do léxico pela cultura, contribuindo para uma prática “do princípio didactológico de consubstancialidade do objeto língua-cultura no processo de ensino-aprendizagem integral correspondente” (Silva e Silva, 2013:69).

Assim, para este trabalho, a noção de lexicultura parece ser relevante, pois aplica-se quer no sistema da língua corrente, quer, naturalmente, nos subsistemas das línguas de especialidade (Lino, Chicuna, Grôz & Medina, 2010:188), bem como no contexto de elaboração de produtos lexicográficos, como, por exemplo, um dicionário. Este último aspeto é pertinente, se se considerar que o léxico e a cultura são inseparáveis no modo como espelham a realidade sociocultural de uma determinada comunidade linguística, de forma a facilitar a apreensão/aprendizagem da informação apresentada para o consulente.

Existe um conjunto de trabalhos que unem a visão da lexicultura e da pragmática lexicultural à perspectiva lexicográfica, quer aplicada a dicionários monolíngues (Aragon Cobo, 2005), quer a dicionários bilingues (Souffi, 2009; Le Gall, 2009; Cherifi, 2009). Estes indicam que, em vários casos, o conhecimento lexicultural pode ser entendido apenas num momento posterior à construção do dicionário, mas deverá ser, sempre que possível, considerado no momento de elaboração. Como menciona Pruvost, “Le dictionnaire, œuvre structurée, lié plus que tout autre aux nouvelles techniques, aux besoins et représentations du public, constitue une œuvre lexiculturelle par excellence.” (Pruvost, 2009:137)

De acordo com o trabalho desenvolvido por Wang (2016), em que a autora aplica o conceito de lexicultura e uma interpretação dos conceitos “cultura popular”, “cultura partilhada” e “cultura erudita”<sup>7</sup> à criação de um dicionário bilingue chinês-português, e relembrando o trabalho de Galisson (1979), que demonstra a existência de níveis de língua especializada e de diferentes estratos de uma mesma língua de especialidade para o domínio do futebol, os intervenientes de uma situação de comunicação de especialidade utilizam diferentes níveis ou estratos da língua de especialidade. No exemplo de Galisson, a equipa técnica (treinadores e dirigentes) tende a usar a “língua especializada”, os jogadores utilizam predominantemente a “gíria” e os jornalistas desportivos usam a “língua banalizada”<sup>1</sup>. A lexicultura é, portanto, um conceito a considerar neste trabalho, na medida em que se apresenta uma proposta de dicionário de especialidade, monolíngue, que pretende integrar a categoria de dicionários de especialidade que incorporam aspetos lexiculturais (Galisson & André, 1998:5-9) e com uma vertente “contextuel de langue à vocation pédagogique ou didactique” (Vigner, 1989: 176-177). Pretende ser um protótipo de um dicionário que visa solucionar as dúvidas dos consulentes e, ao mesmo tempo, codificar informação de especialidade do domínio da ASE, com a respetiva organização do conteúdo linguístico (à semelhança, por exemplo, do trabalho desenvolvido por Tsitsa (2004) no *Dictionnaire Polyvalent du Français Culinaire*).

Pretende-se apresentar um conjunto organizado de unidades lexicais, que o especialista domina (ou deve dominar), uma vez que a primeira competência dos especialistas é lexical (Challe, 2000: 79), sem, no entanto, negligenciar a dimensão cultural ou os implícitos culturais (Pruvost, 2009:138), que existem mesmo num domínio de especialidade e que podem ajudar o não-especialista a compreender melhor o mesmo. Por exemplo, quando um especialista descreve um vinho e menciona que ele “tem frutos”, para o especialista é imediatamente perceptível que o vinho tem aromas ou sabores que lembram determinados frutos, mas sem conter nenhum deles na sua composição. Para um não-especialista, esta perceção não é tão óbvia, como apontam os dados de Barbosa (2021), evidenciando que, pelo facto de ser mencionado o verbo “ter”, o não-especialista entende a estrutura “ter frutos” como “o vinho possuiu na sua composição uvas e outros

---

<sup>7</sup> “culture populaire”, “culture partagee” e “culture savante”.

frutos esmagados”, e não como “o vinho possui cheiros e sabores que *lembram* determinados frutos”, mesmo sem que nenhum deles (à exceção da uva) estejam na sua composição.

Outro aspeto que a organização das unidades lexicais permite é ampliar conhecimento, de forma incremental, como, por exemplo, quando é mencionado que o vinho possui aromas a “frutos azuis” ou a “frutos violeta” está a ser transmitido que o fruto em questão é o “mirtilo” que, por sua vez, também faz parte dos “frutos do bosque”, dos “frutos silvestres” ou dos “frutos em/de baga”. Mas o inverso não é necessariamente verdade. Isto é, se for dito que um vinho tem aroma a “frutos silvestres” pode estar a mencionar que lembra o aroma de vários frutos, tais como: “groselha, morango, amora, framboesa, cássis, cereja, arando”. O facto de um recurso permitir pesquisar o descritor “mirtilo” e compreender, além das suas características, onde se posiciona em relação a outras unidades lexicais (exemplo, se consta como pertencente a outras unidades lexicais complexas) pode ser uma mais-valia para o utilizador.

A lexicultura pode estar presente em diversos aspetos da macro e microestrutura do dicionário, como na nomenclatura selecionada, nas marcas a usar (empréstimo, arcaísmo, neologismo), na indicação do (sub)domínio técnico, entre outras, comuns nos dicionários de língua, mas que espelham variados aspetos lexiculturais de dificuldade acrescida para um aprendente estrangeiro, por exemplo (Ferrara, 2009; Souffi, 2009). Infere-se que, apesar de a lexicultura ser mais estudada nos dicionários de aprendizagem e nos dicionários gerais de língua corrente, é possível aplicar o conceito em dicionários de especialidade, onde também é necessário compreender a cultura do domínio que, apesar de incidir sobre aspetos técnicos e especializados, não deixa de ter carga cultural associada.

A integração e a compreensão da dimensão lexicultural nem sempre são fáceis, pois determinados domínios são particularmente dependentes da cultura<sup>8</sup>. O domínio do vinho,

---

<sup>8</sup> Veja-se os domínios do Direito, Tributação, Economia, que são altamente dependentes da cultura – contexto histórico, económico, social – onde se inserem, e que variam consideravelmente de país para país, para citar alguns exemplos.

presente em diversas regiões do mundo, é, também ele, altamente cultural. Apesar de existir uma universalidade de procedimentos (por exemplo, para atingir a qualidade – processos químicos –, a distribuição, o transporte, ou o processo da ASE do vinho), há aspetos que são moldados pela geografia do terreno, pela história do país, pela religião, pela preferência biológica quanto ao cheiro/paladar, assim como pela verbalização do ato de degustar um vinho, influenciada por tradições, cultura da comunidade e preferências pessoais, de produtores e consumidores. A forma como alguém se expressa através de determinados descritores do vinho numa nota de prova pode fazer com que seja ou não totalmente entendido, dependendo da partilha que os falantes fazem do léxico usado.

Relembrando as palavras de Galisson (1999):

“J’appelle “charge culturelle partagée” la valeur ajoutée à leur signification ordinaire et pose que l’ensemble des mots à charge culturelle partagée connus de tous les natifs, circonscrit la lexiculture partagée. Laquelle est toute désignée pour servir de rampe d’accès à la culture omniprésente dans la vie des autochtones et que les étrangers ont tant de mal à maîtriser.” (Galisson, 1999: 119-120)

É possível transpor a situação, substituindo *autóctones* e *estrangeiros* por especialistas e não-especialistas. Se o falante não conseguir aceder ao conhecimento destas unidades culturais, não acede, conhece ou compreende na totalidade o domínio. A título de exemplo, o descritor de vinhos “adamado” tem como definição: “diz-se de um vinho que contém algum açúcar por desdobrar, embora não constitua sinónimo de doce” (Afonso, 2013) ou “vinho para damas, suave e doce” (fontes: *wineboutique*, *adegaalmeirim*). Culturalmente, verifica-se que o descritor “adamado”, derivado de *dama*, com a definição para indicar que seria um vinho adequado para as mulheres beberem, adquiriu uma segunda definição de um período histórico em que os vinhos se caracterizavam como sendo masculinos ou femininos, como se pode conferir no Triângulo de Vedel (1972).

Um outro exemplo de carga cultural é o do descritor “tem aroma de ginja”. Quando se refere, numa nota de prova, que um vinho “tem aroma de ginja”, é preciso lembrar que um consumidor estrangeiro terá dificuldade em aceder a esse aroma, pois a ginja (ou

cereja ácida, *Prunus cerasus*), apesar de ser conhecida em culturas como a russa, a britânica, a croata (*marrasquino*) ou a alemã (*kirschwasser*), é consumida de forma distinta da usada em Portugal, o que necessariamente dará azo a leituras diferentes, uma vez que ativará aroma/sabor distintos, relacionados com a experiência olfativa/gustativa de cada um.

A noção de lexicultura é também particularmente relevante para o enquadramento do capítulo da *Proposta de dicionário*, em particular quando são comentadas as obras (lexicográficas e congéneres) disponíveis no mercado português do domínio da ASE, com o objetivo de apontar aspetos que podem ser melhorados num dicionário de ASE. Além disso, apesar de neste momento a proposta de dicionário que se apresenta ser pensada como monolíngue, reconhece-se a relevância de considerar os aspetos lexicoculturais portugueses na macro e microestrutura do dicionário, na possibilidade de, no futuro, a proposta poder ser adaptada para uma versão bi- ou multilíngue.

### 2.1.2. LEXICOGRAFIA

A Lexicologia e a Lexicografia são disciplinas com campos conceptuais demarcados, como mostram os trabalhos de Genouvrier e Paytard (1973), Lehmann e Martin-Berthet (1998), Gaudin e Guespin (2000), Lino (1979, 2003, 2010), entre outros.

Por um lado, a Lexicologia debruça-se sobre teorias e metodologias com um pendor descritivo e explicativo, encarando, por exemplo, a sinonímia, a homonímia e até a neologia como fenómenos ou processos regulares de organização e enriquecimento do léxico, e que podem surgir na língua para responder a necessidades dos falantes (Rondeau, 1984:62).

No entanto, apesar de ambas terem como objeto de estudo o léxico, assumem métodos e metodologias distintas, muitas vezes complementares, como mencionam Gaudin e Guespin (2000):

“[Elles] constituent deux disciplines soeurs, donc difficiles à distinguer. La première designe l’étude du lexique d’une langue tandis que la seconde renvoie à la confection

de dictionnaires et, plus largement, d'ouvrages présentant un modèle de ce qu'est ce lexique." (Gaudin e Guespin, 2000:15)

A Lexicografia é habitualmente definida como “a ciência dos dicionários” (Biderman, 1998:15), a arte ou técnica de fazer dicionários (Josselson, 1966), o artesanato de fazer dicionários (Landau, 2001), a técnica, a prática ou a arte na produção de dicionários (Borba, 2003) ou como Hartmann & James (1998):

“The professional activity and academic field concerned with dictionaries and other reference works. It has two basic divisions: lexicographic practice, or dictionary-making, and lexicographic theory, or dictionary research. The former is often associated with commercial book publishing, the latter with scholarly studies in such disciplines as linguistics (especially lexicology), but strict boundaries are difficult to maintain and, in any case, are being bridged by such means as professional training, societies, conferences and publications. (...) Depending on the orientation and purpose of such reference works, several branches of lexicography can be distinguished, each with its own practices and theories.” (Hartmann & James, 1998:85-86)

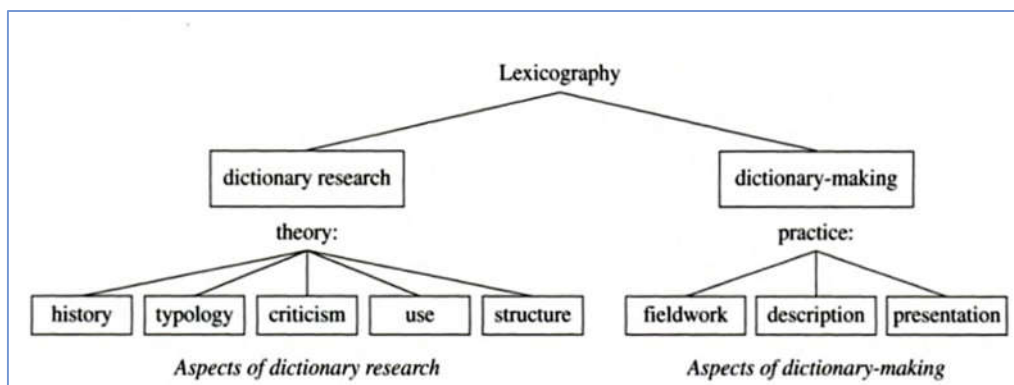


Figura 1. Subáreas da Lexicografia (Hartmann & James, 1998:86)

Hartmann e James (1998) sugerem não só a subárea das práticas lexicográficas (*dictionary-making*), que poderá ser correspondente à noção mais tradicional de “lexicografia”, mas também a da teoria lexicográfica “lexicographic theory” (*dictionary*

*research*) que pode corresponder ao que atualmente se entende por “metalexigrafia”, divisão também defendida por Welker (2006) e Dapena (2002).

Entende-se, portanto, que, enquanto a Lexicografia se ocupa da construção do dicionário, ao uso do *corpus*, à elaboração do dicionário e a como apresentar a informação ao utilizador, a metalexigrafia dedica-se à discussão teórica do conteúdo do dicionário, da tipologia a que pertence, da estrutura interna do dicionário, assim como à crítica do dicionário e do uso que se dele faz.

Se a definição geral de Lexicografia parece mais ou menos consensual, já o seu estatuto enquanto ciência ou subdomínio da Linguística parece reunir menos consenso.

Apesar de muitos lexicógrafos considerarem a Lexicografia uma subdisciplina da Linguística, outros autores possuem outra opinião. Bergenholtz e Gouws (2012) assumem que existe uma relação com a Linguística, mas consideram a Lexicografia como uma disciplina independente, que necessita de especialistas de diferentes áreas, que se dedica à produção de dicionários, mas de um modo geral à produção de ferramentas de informação (Bergenholtz e Gouws, 2012:40).

De acordo com Tarp (2018), a controvérsia em torno da Lexicografia divide-se entre os que consideram a Lexicografia como:

- apenas uma arte ou artesanato de fazer dicionários, uma técnica científica e como uma instância da Linguística aplicada (Landau (2001), Béjoint (2010), Atkins & Rundell (2008), Ten Hacken (2010));
- uma ciência independente que organiza e explica os vocábulos dos dicionários e léxicos (Scerba (1940), Kempcke (1984), Kudashev (2007));
- uma disciplina que não é Linguística aplicada, nem uma subdisciplina da Linguística, mas sim uma ciência prática diferente da metalexigrafia (Wiegand (2013));
- uma ciência independente com a sua teoria e prática e metodologia, com uma tendência para a interdisciplinaridade (Tarp (2008)).

Neste contexto, e respondendo a questões de diferentes ordens, várias teorias foram sendo desenvolvidas no seio da Lexicografia, como é o caso de:

- o esboço da teoria geral da lexicografia de Scerba (1940);
- a teoria da descrição lexicográfica da língua de Wiegand (1984);
- a teoria da lexicografia bilingue de Duda *et al.* (1986);
- a teoria do exemplo lexicográfico de Hausmann (1985);
- a teoria geral da lexicografia de Wiegand (1998);
- a teoria funcional da lexicografia de Bergenholtz e Tarp (2002);
- a teoria da forma do dicionário de Wiegand e Morán (2009).

Independentemente destas questões epistemológicas, e seguindo a definição abrangente de Haensch, Wolf, Ettinger e Werner (1982), a Lexicografia pode ser vista como uma subdisciplina da Linguística aplicada, limitada ou delimitada pelas necessidades dos utilizadores. No contexto do presente trabalho, entende-se a Lexicografia nestas linhas teóricas, como “the science concerned with the theory and practice of dictionaries, that is, dictionaries, encyclopaedias, lexica, glossaries, vocabularies, terminological knowledge bases, and other information tools covering areas of knowledge and its corresponding language” (Fuentes-Olivera, 2018:1) podendo, pelo seu objeto de estudo e métodos, configurar uma ciência independente, ainda que por vezes diretamente relacionada com a Lexicologia.

#### 2.1.2.1. CONTEXTO HISTÓRICO EUROPEU

Embora a Lexicografia seja reconhecida como ciência apenas na segunda metade do século XX, pelo interesse que as obras como dicionários despertaram, a verdade é que

a obras lexicográficas<sup>9</sup> já existem desde tempos remotos, em diferentes pontos do mundo (Hwang 2010, Fernández 2003), havendo inclusive registros e trabalho da possível existência de dicionários orais para determinadas línguas sem sistema escrito (Boisson, 1996). A evolução das línguas e da cultura da comunicação e intercâmbio promoveram o surgimento de materiais lexicográficos, ainda que de aspeto simples e sem as preocupações e “requisito estético e teórico-metodológico, mas que com o tempo se foram aprimorando” (Costa, 2015)

Como se pode constatar, a Lexicografia prática surge antes da Lexicografia teórica, como refere Krieger (2006):

“(…) é milenar, posto que o mundo antigo inaugurou a prática de relacionar palavras e sentidos para atender a necessidades de informação das coletividades linguísticas. É assim, por exemplo, que nascem glossários na Grécia Antiga, nos quais eram listadas e definidas palavras de difícil compreensão de obras literárias. Estas, quando organizadas alfabeticamente ao final dos textos, constituíam os glossários, que representam as primeiras formas de manifestação lexicográfica.” (Krieger, 2006:141).

Compreende-se que a necessidade de explicar e entender determinadas palavras tenha feito surgir glossários, que continham palavras raras e/ou desusadas com explicações para as obras dos escritores gregos e romanos, como é o caso da obra *De linguae latinae* de Varrão (séc. I a.C.), que é por alguns considerado o primeiro glossário<sup>10</sup>. Mas importa notar que a noção de glossário aqui é usada de forma distinta daquela que hoje é convencionalizada.

Na Idade Média, quando se verifica o distanciamento entre o latim culto e o latim vultar, assiste-se a um surgimento dos léxicos que permitiam, por um lado, preservar a

---

<sup>9</sup> Os primeiros dicionários podiam ter diferentes designações, como, por exemplo, glossários, *thesauri*, vocabulários, léxicos, tratados, compêndios, nomenclaturas, entre outros, funcionando em alguns períodos históricos como sinónimos. No entanto, independentemente da sua designação, esta obra lexicográfica é um objeto que ao longo da História evoluiu e se adaptou às necessidades das comunidades linguísticas.

<sup>10</sup> Há também evidências que Marcelo (Nonius Marcellus, séc. III-IV d.C.) foi escritor da obra *De compendiosa doctrina*, uma espécie de léxico que é uma compilação de comentários de outros autores sobre linguagem e gramática, cf. Coradini (1999) e Valenza (2010).

língua culta e a memória dos vocábulos (Bevilacqua, 1993:17) e, por outro, ganhar o caráter normativo e de manutenção de “bons usos” com que os dicionários tendem a ser vistos ainda hoje.

No período do Renascimento e com o advento da imprensa (que possibilita a rápida reprodução das obras anteriormente manuscritas), surgem os primeiros dicionários bilingues e multilingues, com o objetivo principal de integrar económica e culturalmente as comunidades falantes que estavam em contacto. As obras de Palencia (1490) e Nebrija (1492) para o castelhano, de Calepino (1502) para o italiano e de Estienne (1539) para o francês são exemplos de primeiros dicionários, usados para explicar unidades lexicais de obras clássicas.

Nos séculos XVI e XVII, assiste-se ao surgimento de vocabulários e dicionários especializados e com pendor normativo, como o *Dictionnaire de L'Academie Française* 1694 e o *Diccionario de Autoridades de la Real Academia Espanhola* de 1739.

No século XVIII, verifica-se um cuidado e preocupação para integrar mais informação na descrição das unidades lexicais. A Lexicografia de cariz enciclopédico surge com maior destaque, explicando e contextualizando as várias unidades lexicais que surgem com o progresso do conhecimento humano neste período da História.

No século XIX, deu-se o desenvolvimento das ciências e das técnicas e, conseqüentemente, a expansão da Lexicografia às obras de especialidade, proporcionando e visando o intercâmbio de conhecimento.

Por fim, no século XX, o marco mais recente, a Lexicografia relaciona-se com o avanço e o desenvolvimento da informática, da relevância do uso de *corpora* que modificou a forma de criação e elaboração, mas também a velocidade, de construção das obras lexicográficas e a atenção às necessidades dos consulentes/função do dicionário (Bergenholtz & Tarp, 1995). Este avanço tecnológico e metodológico permitiu, por um lado, a proliferação de materiais e, por outro, levou à reflexão, uma vez mais, acerca da conceção da Lexicografia como ciência, questionando os princípios e teorias existentes, à luz deste novo contexto epistemológico.

### 2.1.2.2. CONTEXTO HISTÓRICO PORTUGUÊS

Compreendendo o que se passava no mundo ocidental, considera-se que a Lexicografia portuguesa teve o seu início com Jerónimo Cardoso, com a elaboração do *Breve Dictionarium Vocum Ecclesiasticarum. In Dictionarium latinolusitanicum* (1569-1570). De acordo com Verdelho e Silvestre (2007), o património lexicográfico português<sup>11</sup>, apesar de menor quando comparado com o vasto espólio de trabalhos que se desenvolviam para outras línguas europeias na mesma época, pode dividir-se em períodos, como se apresenta na Tabela 1.

---

<sup>11</sup> A delimitação temporal e espacial da lexicografia portuguesa aqui apresentada circunscreve-se ao contexto ocidental europeu, sem fazer menção e/ou referência a outros contextos mundiais (lexicografia chinesa, árabe e egípcia, por exemplo).

Obras e autores relevantes	Características
<b>Séc. XI-XIII - Pré-lexicografia</b>	
<i>Elementarium</i> (c. 1050), de Papias; <i>Liber derivationum</i> (XII), de Hugúcio de Pisa; <i>Catholicon</i> (1286), de João Balbo.	(i) gênese dos dicionários em escolas medievais; (ii) textos medievais com informação latina, permitindo o aparecimento dos glossários das línguas modernas; (iii) textos jurídicos de Afonso X traduzidos do castelhano com definições lexicográficas.
<b>Séc. XIII-XVII - Lexicografia latino-portuguesa</b>	
<i>Glossários bilingues medievais</i> ; Dicionário medieval alcobacense de verbos - Manuscrito alcobacense (códice CDIV/286); <i>Vocabulário latino-português</i> (1170) de Martinho de Arraiolos.	(i) emergência da escrita vernacular; (ii) sistematização da escrita; (iii) <i>corpus</i> lexical com 1100 verbos; vocabulário ordenado alfabeticamente.
Origens renascentistas da lexicografia portuguesa; <i>Dictionarium ex lusitanico latinum sermone</i> (1562) de Jerónimo Cardoso; <i>Dictionarium lusitanicolatinum</i> (1611) de Agostinho Barbosa.	(i) primeiros dicionários gerais e científicos; (ii) dicionário de aprendizagem com uma secção dedicada à medicina; (iii) lexicografia de especialidade (com terminologias científicas direito, comércio e geografia; (iv) neologismos de medicina, anatomia, botânica.
Lexicografia dos Jesuítas; Dicionaristas das línguas de missão; <i>Vocabulário Lusitanico Latino</i> (1561) de Manuel Barreto; <i>Tesouro da Língua Portuguesa</i> (1647) de Bento Pereira.	(i) produção de manuais, dicionários e gramáticas escolares; (ii) dicionaristas no Brasil e no Oriente; (iii) referência normativa com uso de <i>corpus</i> .
<b>Séc. XVIII - Vocabulário de Bluteau</b>	
<i>Vocabulário português e latino</i> (1712-1728) de Rafael Bluteau	(i) obra monumental com pouca informação em latim, (ii) obra monolíngue com explicitação pormenorizada, normalização lexical e ortográfica; (ii) obra com reflexão teórica e com informação de variedades regionais
<b>Séc. XVIII - Tentativas de elaboração enciclopédica</b>	
<i>Corte na Aldeia</i> (1619) de Rodrigues Lobo; <i>Escola decurial de várias lições</i> (1696) de Fradique Espinola; <i>Recreação filosófica</i> (1751) de Teodoro de Almeida	(i) produção dicionarística com informação enciclopédica (ii) informação em diversos volumes; (iii) súmulas eruditas com informações culturais da diacronia lexical portuguesa.
<b>Séc. XVIII - Lexicografia moderna monolíngue portuguesa</b>	
<i>Dicionário da língua portuguesa</i> (1783) de Bernardo Bacelar, <i>Orthographia, ou arte de pronunciar com acerto a língua portuguesa</i> (1734) de Madureira Feijó; <i>Elucidário</i> (1798) de Viterbo Rosa; <i>Dicionário da Academia Real das Ciências e dicionários especializados</i> (1780).	(i) dicionários monolíngues portugueses modernos; (ii) vocabulários especiais e ortográficos
<b>Séc. XIX-XX - Lexicografia monolíngue</b>	
<i>Dicionário de língua portuguesa</i> (1789) de Morais Silva; <i>O grande dicionário português</i> (1871) de Domingos Vieira; <i>Dicionário contemporâneo da Língua portuguesa</i> (1881) de Caldas Aulete; <i>Novo dicionário da língua portuguesa</i> (1913) de Cândido Figueiredo; <i>Dicionário geral e analógico</i> (1948) de Artur Bivar; <i>Dicionário etimológico</i> (1932) de Antenor Nascentes; <i>Dicionário etimológico</i> (1952) de José Pedro Machado; <i>Português Fundamental</i> (1984-1987) de Centro de Linguística da Universidade de Lisboa.	(i) democratização e a escolarização; (ii) Grandes dicionários, dicionários universais, dicionários paradigmáticos ou morfológicos, dicionários etimológicos; (iii) reconceptualização do conceito de dicionário, dicionarística e nova lexicografia (Quemada); (iv) lexicografia linguístico-literária, aspetos como a ortografia, história da língua; (v) dicionários informatizados; lexicomática (dicionários-máquina), ciberlexicografia.
<b>Séc. XX - XXI - Lexicografia digital</b>	
Porto Editora; Priberam; e-dicionários.	Dicionarística digital, outros recursos multimodais e multifuncionais, relacionados com a evolução da tecnologia.

Tabela 1. Períodos da Lexicografia portuguesa

A Tabela 1 visa um resumo visual de uma enorme quantidade de informação relativa a vários momentos temporais e, também, às diversas tendências lexicográficas em uso. Como afirma Verdelho (1995:137), “a lexicografia nasceu da consciência das fronteiras linguísticas e da necessidade de estabelecer a comunicação entre idiomas diferentes”, o que explica a sua afirmação de que “todos os dicionários das línguas

modernas nasceram bilingues”, para colmatar a necessidade existente de comunicação e entendimento entre culturas e línguas.

Não se pretende neste trabalho entrar em pormenores sobre o aspeto histórico da Lexicografia portuguesa. No entanto, esta contextualização explica algumas opções metodológicas seguidas na proposta do DEASE.

Relativamente ao período da Lexicografia digital, é relevante mencionar que este período começa nos finais do século XX, quando se assiste às preocupações relativas à informática e à forma como se poderia transformar um dicionário impresso em versão eletrónica (disquetes, CD, CD-ROM) ou em *software* educativo, muito em voga e com ampla difusão nos anos 90 do século XX. No entanto, a democratização do acesso à Internet trouxe à tona que a transposição da informação de um dicionário em papel para o ambiente eletrónico não podia ser linear, como refere Dziemianko (2018:13) “electronic lexicography is not just a continuation of print lexicography. Nor does it mean merely making print dictionaries available online or on some data carrier”. Se, num determinado momento, a preocupação era a qualidade da informação lexicográfica apresentada no dicionário – que devia ser correta, fidedigna, recolhida em fontes de referência, para ser considerada uma obra de referência – com as novas tecnologias, as preocupações expandiram-se noutros sentidos.

A versão digital permitiu o aparecimento não só de um espaço maior para o dicionário (as questões de tamanho, número de páginas, número de entradas deixam de ser vistas como impeditivas), mas também o desenvolvimento de ferramentas de consulta (pesquisas avançadas, por exemplo), que antes não existiam e, paralelamente, o surgimento de novos produtos, como redes de dicionários (*dictionary nets*), motores de busca associados a dicionários ou coleções de dicionários (Boelhouwer, Dykstra & Sijens, 2018).

A mudança de paradigma relativo à Lexicografia e ao modo como os seus atores entendem o que é e como poderá ser um dicionário do século XXI lembra que, na verdade, o sentido de comunidade e de partilha de informação deverá ser feito por todos os intervenientes, como sugere Hartmann (2018):

“shows the necessity of having a directory of lexicography where lexicographers, researchers, and publishing houses have access to lexicographers’ names and addresses, lexicographic resources (e.g. journals), lexicographic institutions, research centres, publishers, conferences, and so on.” Hartmann (2018:783)

Esta expansão de produtos, ferramentas e espaço para os dados do dicionário, e a facilidade e maior amplitude de acesso aos mesmos, possibilita que o papel do consulente também se possa alterar. Ou seja, aquele que era antes apenas o utilizador passa a poder ter um papel diferente, seja de forma direta ou indireta, sendo considerado na forma como a informação no dicionário eletrónico é apresentada, na adequação dos conteúdos às suas necessidades, até ao ponto de poder ser colaborador na construção e atualização do dicionário, com contribuições que vão desde comentários, sugestão de inserção de palavras, correção de informação até à elaboração de entradas de dicionário, como é o caso dos dicionários *crowdsourcing*, que se baseiam na Lexicografia cooperativa (Meyer & Abel, 2018; Boelhouwer, Dykstra, & Sijens, 2018).

### 2.1.3. LÍNGUA E RECURSOS DE ESPECIALIDADE

Para fazer uma proposta de um dicionário, é necessário refletir sobre o processo de criação de dicionários e sobre a ciência que os origina, ou seja, a Lexicografia geral, quando nos referimos aos dicionários de língua geral e a Lexicografia de Especialidade, quando nos referimos aos dicionários especializados.

Se por um lado se entende a Lexicografia e a Terminologia como duas disciplinas distintas, com objetivos distintos, mas que em determinados momentos podem apresentar alguma convergência na metodologia de trabalho na elaboração de dicionários, considera-se, por outro, que a Lexicografia de Especialidade e a Terminografia podem ser vistas como “técnicas, ciências ou disciplinas independentes, para outros são técnicas que, embora diferentes, possuem pontos de intersecção, e há ainda os que afirmam que são sinónimos, isto é, variantes denominativas para uma mesma atividade”, como refere (Costa, 2015:44). Ou seja, enquanto a Lexicografia e a Terminologia podem ser

entendidas como áreas complementares e não excludentes, na ótica de Bergenholtz (1995) e de Bergenholtz e Tarp (2010), a Terminografia e a Lexicografia Especializada podem ser vistas como paralelas se não mesmo equivalentes, dado que possuem um mesmo objeto de estudo e um mesmo objetivo, que é a construção de um dicionário de especialidade.

Há inclusivamente autores que consideram sinónimos “Lexicografia de Especialidade”, “Lexicografia de Linguagens Especiais”, Lexicografia Terminológica” como Ciobanu (2003). Já Schierholz (2012) menciona que considera Terminografia ou Lexicografia Terminológica distinta de Lexicografia de Especialidade ou Lexicografia Especializada, ainda que ambas as disciplinas tenham como objetivo a construção de um dicionário de especialidade que pode ser direcionado para diferentes públicos – especialistas ou não-especialistas – com uma apresentação da informação considerada tradicional ao nível da macro e microestruturas, uma nomenclatura com artigos lexicográficos por ordem alfabética com informação lexicográfica, linguística e enciclopédica, conforme as opções feitas para apresentar quer os artigos quer as definições de um domínio de especialidade.

Outros autores, não satisfeitos com as denominações já existentes, sugerem novas designações, como Marzá (2009), que vai mais longe e além de considerar que a Lexicografia de Especialidade possa ser entendida como sinónimo de Terminologia ou de Terminografia, sugere uma nova abordagem “Specialized Lexicographical Approach”, que se baseia na proposta de Cabré (1999) e da Teoria Comunicativa da Terminologia. Esta nova proposta baseia-se em três aspetos: Língua de especialidade, Terminologia e Linguística de *Corpus*.

Neste trabalho, entende-se por Lexicografia de Especialidade:

“A complex of activities concerned with the design, compilation, use and evaluation of specialised dictionaries. There is no uniform framework for this as the nature and the scope of such reference works can range widely, from a brief glossary without definitions, through technical dictionaries aimed at lay persons, to large scale and standardised terminological databases for subject experts and translators.” (Hartmann & James, 1998:129)

Esta disciplina, que permite registar e documentar o léxico de um domínio específico, onde lexicógrafos especializados podem utilizar um conjunto de aspetos metodológicos partilhados em comum com a Terminografia, mas na linha de raciocínio de Bergenholtz e Tarp (2010), que mencionam que ambas se dedicam a um léxico especializado com características particulares, que apresentam regra geral um modelo de macroestrutura sistemático, muitas das vezes organizado por campos semânticos e que se posicionam entre um carácter descritivo e um carácter prescritivo, tentando sempre que o resultado final seja uma fonte de consulta de um domínio específico e que possa servir a codificação e a descodificação daqueles que o consultam.

“We still see terminography as a synonym of specialized lexicography. Not all colleagues agree, nor the majority of lexicographers, who see lexicography as a part of linguistics, and most terminography neither, who claim there are large but unclear differences in relation to specialized lexicography. In reality, it is a discussion about something else, about research funding, about influence and positions at universities, and about defending a position concerning two traditions in making tools to solve exactly the same types of problems” (Bergenholtz & Tarp, 2010: 29).

A atenção ao léxico de línguas de especialidade surge já em Bluteau que, na definição da nomenclatura, expressa uma preocupação de apresentar um “português básico e também de domínios lexicais de especialidade” (Verdelho & Silvestre, 2007:127), bem como em outras obras do século XVIII. Como referem os autores Verdelho e Silvestre (2007), esse período corresponde a um momento em que a renovação do léxico dicionarizado procura acompanhar o desenvolvimento das linguagens de especialidade. A tradução dos tecnoletos torna-se uma área essencial nas preocupações da reflexão da metalinguística, não só porque a descodificação das terminologias é um ponto de partida para a descrição enciclopédica do mundo, mas também porque este léxico se torna progressivamente indispensável para essa mesma descrição (Verdelho & Silvestre, 2007:153).

Ao longo da história, o conceito de dicionário especializado passou por diferentes momentos e abrangeu diversos tipos de recursos lexicais, desde os vocabulários ortográficos, como a obra de Rosa Viterbo (1798/1984), os vocabulários poéticos e para-literários, caracterizados pela originalidade dicionarística e interesse linguístico a recursos que constituem “um excelente repositório para a história da língua e do vocabulário da técnica e da ciência nos meados do séc. XIX” (Verdelho, 1994), como as “terminologias da ciência e da técnica” (Verdelho, 1994) presentes no Dicionário Aulete, por exemplo.

Já Costa (2015) refere que “a fim de atingir seu objetivo, o dicionário de especialidade contém um índice de verbetes especializados (em ordem alfabética ou outra ordem), e para cada verbete há geralmente informações linguísticas ou enciclopédicas. A estrutura tradicional (verbetes com lema e informações lexicográficas, como informações morfológicas e sintáticas, sinonímia, marcas de uso e exemplos) é frequentemente deixada de lado nos dicionários de especialidade.” (Costa, 2015:45)

A discussão sobre como se caracteriza a Lexicografia de Especialidade e onde se posiciona como disciplina é e como se caracteriza está longe de estar terminada. Diferentes autores demonstram ter posições bastante extremadas e, mesmo considerando neste trabalho que se tende para uma visão mais no sentido de considerar a Lexicografia de Especialidade uma disciplina que se dedica ao estudo do léxico de um determinado domínio de especialidade, considera-se que esta segue um conjunto de procedimentos que se aproximam da Terminografia, mas apresentando ao utilizador do dicionário de especialidade um conjunto de informação considerada tradicional, aliada a um conjunto de informação de especialidade.

#### 2.1.3.1. LÍNGUA E LÉXICO DE ESPECIALIDADE

Nesta secção, reflete-se sobre os conceitos de língua geral e de língua de especialidade, elemento diferenciador que justifica, para muitos a oposição entre um dicionário geral e um dicionário especializado (Costa, 2015:62).

Entendendo-se que “a língua é um sistema de signos” com vários sistemas ou subsistemas interdependentes, como mencionou Saussure (1916), é comum mencionar-se que este sistema se divide em língua comum (com uma aproximação de 20 000 a 25 000 unidades lexicais partilhadas entre a generalidade da comunidade linguística), língua corrente (com 50 000 a 100 000 unidades lexicais frequentemente em uso) e língua de especialidade de um domínio científico, técnico e/ou tecnológico (que pode variar entre 1000 e 300 000 unidades lexicais, consoante diferentes estádios ou níveis de especialização).

A língua corrente é o sistema linguístico que o falante utiliza no seu quotidiano, quando comunica com os restantes falantes na sua comunidade linguística (Guilbert, 1971). Autores como Gil (2003) e Bergenholtz e Tarp (1995) referem que não há distinção entre língua geral e língua de especialidade, mas sim inclusão, havendo autores como Cabré (2005) que entendem as línguas de especialidade como um subconjunto da língua geral. Entende-se que a língua de especialidade não é apenas usada por especialistas de um domínio ou de diferentes domínios; também pode ser usada por semiespecialistas e por não-especialistas (Alvarez, 2019:6).

O subsistema da língua corrente (semas mais frequentes) pode ser distinguido do subsistema da língua de especialidade (semas estáveis de um determinado domínio). Ambos possuem unidades lexicais, as unidades de significação. No entanto, na língua de especialidade a unidade lexical é especializada por possuir características semânticas específicas, por um lado, e por ser tendencialmente monossémica, por outro. A língua de especialidade é, assim, um subsistema no âmbito do sistema da uma língua, relativo a uma área do conhecimento (Lino, 2018).

Segundo Alves (2006), o léxico português expandiu-se ao longo da sua existência, quer em relação às unidades lexicais da língua comum, quer em relação às unidades lexicais específicas de um determinado domínio, ou seja, nas línguas de especialidade. Tal sucede devido ao facto de as unidades lexicais refletirem o uso e a realidade de uma comunidade falante (Barros, 2004), que acompanham as alterações, seja na língua comum, seja na língua de especialidade, através do surgimento de unidades, da modificação de significados de acordo com as necessidades da comunidade e das

atividades que pratica, e também através da perda de unidades lexicais que deixam de ser necessárias a essa mesma comunidade. A título de curiosidade, basta lembrar as pesquisas que podem ser feitas através da ferramenta do Google – *Google Trends*, que mostra as palavras mais populares pesquisadas num determinado período, numa determinada região do globo ou até mesmo em diferentes línguas.

São diversas as designações associadas ao conceito de língua de especialidade relevantes para o presente trabalho:

- *língua* – línguas especializadas (Cabré, 1999), línguas científicas e línguas técnicas, língua instrumental, língua profissional, língua para fins específicos, (Galisson e Coste 1983), língua de especialidade (Kocourek, 1991);
- *linguagem* – linguagem de especialidade, linguagem de ofícios, discurso temático (Descamps, 1977), linguagens técnicas (Carvalho, 1973);
- *discurso* – discurso temático, discurso do domínio, discurso científico, discurso funcional, discurso específico (Descamps, 1977), discurso de especialidade (Méric, 2018; Lerat, 2006; Gautier 2019), discurso multimodal (Gautier & Hohota, 2014), discurso especializado (Petit, 2010);
- *vocabulário* – vocabulários técnicos (Phal, 1971; Guilbert, 1973; Ettinger, 1982), vocabulário científico (Guilbert, 1971), vocabulário técnico-científico (Rey, 1986), vocabulário tecnológico (Rey, 1986);
- outros – temas de interesse (Rey-Debove, 1973), noção de especialidade (Méric 2018), léxico de um domínio (Coutier, 1994; Langlois, 2010; Vion, 2009; Zannoni, 1997).

De acordo com Gil (2003), as línguas de especialidade podem ser definidas como o conjunto de elementos linguísticos comuns na comunicação entre especialistas de determinado domínio e também entre especialistas e um público em vias de especialização, isto é, estudantes que pretendem especializar-se num determinado domínio. Podem estar também relacionadas com os “domínios de experiência” e os

“campos de experiência”, que são mencionados como os “laços que indivíduos de uma mesma comunidade linguística estabelecem com o mundo que os rodeia e a sociedade em que estão integrados” (Gil 2003:114-115). Ou seja, as línguas de especialidade estariam sempre relacionadas com um campo de investigação/atuação, isto é, seriam a língua utilizada em determinado âmbito de estudo/atividade.

Como apontam Galisson e Coste (1983), as línguas de especialidade englobam os domínios com objetos de investigação específicos, como a matemática ou a biologia, as línguas técnicas e de aplicação, como a informática ou a engenharia, e as línguas profissionais associadas a atividades como a enfermagem, a advocacia, entre outras. Todos estes domínios “utilizam um fundo lexical que se encontra também na língua comum, mas em aceções específicas, tendo em conta a ciência a que dizem respeito” como ressalta Gil (2003:116).

Kocourek (1991) fala inclusivamente da diferença entre *língua de especialidade* (coesão interna do subsistema linguístico que está associado a um domínio) e das *línguas de especialidade* (vários microssistemas). Chega aliás a usar *língua de especialidade* como hiperónimo de *língua tecno-científica* e de *língua erudita*.

Boutin-Quesnel, Bélanger, Kerpan, & Rousseau (1985) referem que *língua de especialidade* visa a não ambiguidade na comunicação, enquanto Lerat (1988) considera que a diferença entre *língua de especialidade* e a *linguagem de especialidade* não são as regras, mas a especificidade do vocabulário utilizado num determinado domínio do saber.

Entende-se, neste trabalho, portanto, o léxico de especialidade como uma parte (ou subparte) do léxico da língua corrente, geral ou comum, que se especializa de acordo com o contexto em que se insere e é utilizado, adquirindo características e significações próprias desse contexto. Esta distinção ajuda a separar a língua comum ou geral e a língua de especialidade, explicando a diferença entre um dicionário geral e um dicionário de especialidade.

Falar de léxico de especialidade é falar de um conjunto de unidades lexicais especializadas que, como mencionam Krieger e Finatto (2004), são unidades semelhantes

às unidades lexicais comuns da língua, mas que ganham um valor especializado e refletem um valor específico do daquele que possuíam na língua comum.

“Em realidade, os termos técnicos e/ou científicos deixaram de se configurar como uma “língua à parte”; já não são mais facilmente identificados, como ocorria quando, ao modo das nomenclaturas, correspondiam a palavras muito distintas da comunicação ordinária e permaneciam praticamente restritos aos diferentes universos comunicacionais especializados. Hoje, os termos circulam intensamente, porque ciência e tecnologia tornaram-se objeto de interesse das sociedades, sofrendo, conseqüentemente, processo de vulgarização favorecidos pelas novas tecnologias da informação” (Krieger, Maciel & Finatto, 2000:145).

Ou seja, quando um domínio de especialidade precisa de ser comunicado a outros intervenientes, tem duas hipóteses: ou cria unidades lexicais novas, ou apropria-se de unidades lexicais já existentes no léxico comum, com novos significados. Aspeto recorrente em diversos domínios, e no qual, através do contexto e domínio em que as unidades são utilizadas, é possível identificar esse valor especializado que possuem, refere Biderman (2006):

“Um problema teórico de grande relevância relativamente aos vocabulários científicos e técnicos diz respeito ao processo de criação das novas unidades léxicas que, como se disse, é fato que ocorre com enorme frequência nesses domínios. Ora, sucede que as linguagens de especialidade geram novos termos com base no acervo que a língua já possui. De um lado, reutiliza palavras já existentes para criar outras, ou então, serve-se dos processos de formação de palavras que existem no sistema do português. Assim, por exemplo, os processos de derivação e composição são continuamente solicitados para gerar neologismos que se fazem necessários.”  
Biderman (2006:35)

Assim, entende-se que o léxico de uma língua possui unidades lexicais comuns e que estas podem ganhar, através do uso e do contexto em que são utilizadas, o carácter especializado, como refere Cabré (2008). Partindo das ideias de unidade lexical comum

e especializada, estas devem ser entendidas como estruturas que podem ser mono- ou polissêmicas, que têm uma categoria gramatical, que podem pertencer a vários domínios de especialidade e, por fim, cujos significados podem depender de fatores sociais e culturais – princípio da poliedricidade de Cabré (2008). Ao nível do léxico de especialidade, pode então considerar-se que existe a unidade lexical comum (que inclui unidades simples, neológicas e compostas) e a unidade lexical especializada (que inclui unidades simples e compostas, unidades fraseológicas e unidades neológicas).

### 2.1.3.2. SOBRE O *DESCRITOR*

No contexto do presente trabalho, diversos autores usam *descriptor*, *termo*, *atributo*, *descriptor sensorial*, *atributo sensorial*, *termo descritivo*, *termo técnico*, para designar a(s) unidade(s) lexical(is) especializada(s) relacionada(s) com a descrição da sensação e percepção sensorial associada a ASE.

Optou-se por *descriptor*, em vez de outras formas, pois parece ser a forma mais semanticamente transparente, integrando uma série de características diversas, como se apresenta de seguida.

Assim, o *descriptor*, neste trabalho:

- é entendido como uma unidade da língua de especialidade, que pode também pertencer à língua corrente (Larivière 1989:459 *apud* Contente 2008:41), compreendida pelos especialistas, além de surgir nas obras de referência deste domínio (Larivière 1989:459 *apud* Contente 2008:41);
- tem valor comunicativo (Langlois 2010, Chollet & Valentin 2000), é entendido como relevante no domínio em questão (Harter & Cheng 1996:312) e engloba o significado técnico e sensorial (Gautier & Bach, 2017);
- permite descrever e qualificar a sensação que é entendida pelos especialistas de forma homogênea (Giboreau, 2009:15), não sendo necessariamente monorreferencial (Larivière, 1989:459 *apud* Contente, 2008:41), mas cuja

denotação permite expressar um caráter multissensorial, mas não ambíguo, e tendencialmente monossêmico (Giboreau *et al.*, 2007:267);

- através de um processo de recategorização (Cayot, Gautier & Soubrier, 2006) e especificação, estabelece relações com hipónimos que permitem verbalizar as propriedades sensoriais (Langlois, 2010:7) e as propriedades sensoriais hedônicas, possibilitam a sua descrição simples e a descrição quantitativa (Nogueira, 2011:24) de diferentes tipologias de descritores – avaliativos (Lehrer, 2009), hedônicos (Gautier 2018), sensoriais (Pagliarini, Laureati & Gaeta, 2013; Silva 2016), metafóricos (Lehrer & Lehrer, 2016; Isani 2018; Wipf, 2010; Amoraritei, 2002) – e permitem avaliar em escala (Nogueira, 2011:11);
- é uma unidade multidimensional (Lesschaeve, 2006), que possui uma dimensão tripartida: emocional, cultural e linguística (Gautier 2018:9), denota a percepção de um produto pelo indivíduo (Nogueira 2011:11) e o recorte de uma experiência sensorial específica (Geeraerts, 2006), e condensa diversas vertentes: semântica, lexicográfica, linguística e documental (Peña, 2000:111).

Os descritores são usados para descrever algo psicofisiológico, através de verbalização, no sentido de clarificar o que foi experienciado. Funcionam como uma relação entre a sensação e a percepção, e a verbalização, que pode incluir a conceptualização, a nomeação e a produção oral/escrita, como menciona Giboreau *et al.* (2007:270). Os descritores são unidades lexicais (simples ou complexas) que permitem aos especialistas referir-se de uma forma tendencialmente objetiva e verdadeira a características do vinho e aos não-especialistas a considerá-los de fácil memorização e com significado partilhado (Dubois & Giboreau 2006:670), facilitando a comunicação.

No entanto, não raras vezes, observa-se falta de clareza nos descritores usados (como mencionam López-Arroyo & Roberts, 2020), tanto em termos do que pretendem descrever, isto é, do que os especialistas consideram que denotam, como em termos da percepção que os não-especialistas deles têm, isto é, do que os não-especialistas interpretam que denotam. É esta falta de clareza que se pretende eliminar através do

estudo apresentado neste trabalho, visando a criação de um dicionário que ajude a estabilizar o significado dos descritores e a promover a sua utilização de uma forma menos ambígua e mais adequada, sem formalizar ou propor formas específicas para a normalização e seguindo o sentido de harmonização, na medida em que a organização do conhecimento parte do sensorial para o descritor e não o contrário. Neste contexto, *descriptor* parece ser a forma mais neutra de maximizar a clareza e diminuir a confusão na comunicação da ASE.

#### 2.1.4. DICIONÁRIO

A palavra *dicionário* tem origem na palavra latina *dictionarĭu*, que, em latim medieval, significava “repertório de «dictiones», isto é, frases ou palavras” (Houaiss, 2004). Para iniciar esta contextualização, de forma simplista, um dicionário pode ser visto como um catálogo, uma coleção de palavras e/ou expressões ordenadas alfabeticamente, seguidas de um pequeno texto informativo relativo ao seu significado e modo de uso expressão numa determinada língua, o que tipicamente origina um livro de dimensão considerável.

Este tipo de recursos, como qualquer outro, foi sofrendo alterações ao longo dos tempos, consoante a época e a evolução do conhecimento científico e técnico. Relembre-se o trabalho de Grimm (1966), que já alertava para a denominação das obras lexicográficas,

“Los griegos y romanos no tenían idea de un diccionario, y las posteriores denominaciones ‘lexicon’, ‘glossarium’, ‘dictionarium’, ‘vocabularium’, usuales en suas lenguas, significan otra cosa. El lexikón (biblion), derivado de léxis, y el dictionarium, derivado de dictio, reúnen locuciones, expresiones; el glossáron interpreta voces antiguas, oscuras, contiene glosas; el vocabulario se refiere tan sólo a unos pocos vocablos, que se han recogido para estudiantes o para otra finalidad.” (Grimm (1966:105)).

Aspeto que Haensch (1982) mais tarde relembra e enfatiza,

“A partir [do século XVI], se usan, en la lexicografía europea, los términos ‘vocabulary’, ‘dictionary’ y también ‘lexicón’; pero, como hemos visto ya, por desgracia, estas denominaciones se aplicaron a obras de tipo muy diferente, y así surgió una gran confusión en cuanto a las denominaciones de los distintos tipos de obras lexicográficas, confusión que aún es mayor en otras lenguas, como en alemán, ‘Lexikon’, ‘Wörterbuch’, ‘vokabular’, etc. Esta confusión se debe al uso arbitrario de estas denominaciones por parte de los distintos autores o casas editoriales y, hasta cierto punto, también a modas de las distintas épocas”. (Haensch, 1982:107)

Tendo por base a proposta de Correia (2009: 25-31) que classifica as obras lexicográficas por tipologia, entende-se:

- *dicionário de língua*: obra que apresenta e descreve um conjunto de unidades lexicais ou palavras com informação gramatical;
- *enciclopédia*: obra de grandes dimensões, com unidades lexicais de diversas áreas do saber, apresentando textos que podem ser acompanhados de ilustrações e com informação extralinguística;
- *glossário*: obra restrita a unidades lexicais de uma área do saber, ou de um autor, por exemplo, pode ser surgir anexada a outra obra, para consulta;
- *tesauro*: obra que apresenta as unidades lexicais através da relação onomasiológica;
- *vocabulário*: obra circunscrita a um conjunto de unidades lexicais de um registo da língua ou a um conjunto de textos.

Enquanto obra e recurso lexical com dimensão social, um dicionário pode ser considerado:

- um objeto educacional e de normalização, onde se preserva, descreve e dissemina uma língua (a par de obras de interesse, como as gramáticas), que contribui para

a sua uniformização, funcionando também como instrumento de apoio à aquisição e aprendizagem quer da língua materna como de línguas estrangeiras;

- um objeto político, na medida em é, muitas vezes, usado como referencial para questões de política de língua e representação/fixação oficial da língua, promovendo estabilidade e coerência relativamente à variação diacrónica e diatópica das línguas e funcionando como instrumentos mais harmonizadores (descritivos) ou mais normalizadores (prescritivos);
- um objeto histórico, social e cultural (Biderman (1984:28)), na medida em que preserva e permite guardar a memória de uma língua e da forma como a comunidade linguística observa e entende o mundo, permitindo uma análise à representação dessa realidade social através da língua e dos seus usos;
- um objeto comercial e editorial, direcionado para diferentes tipos de público e colmatando lacunas de mercado, seja em domínios generalistas, seja em domínios especializados.

Desta forma, e como menciona Gelpí Arroyo (2000), um dicionário é um produto poliédrico porque são múltiplos os pontos de vista sob os quais se pode descrevê-lo.

O dicionário tem como propósito dissipar dúvidas quanto à ortografia de uma palavra ou expressão, à categoria gramatical a que pertence, ao seu significado e ao uso em contexto comunicacional. Pode informar sobre a etimologia da palavra ou, simplesmente, ser consultado por curiosidade intelectual. Permite descodificar palavras consideradas difíceis, científicas ou técnicas, bem como compreender o significado de arcaísmos e/ou neologismos, ainda que nem todos os dicionários contemplem todos estes aspetos. Pode afirmar-se que o dicionário serve como um repositório do léxico da língua de uma dada comunidade linguística, com os significados das várias unidades lexicais, transformando-se “numa fonte de consulta sobre palavras expressões, termos e sentidos desconhecidos” (Krieger, 2006).

Para Lara (1996:102), “o ato de pergunta e de resposta acerca do significado de um signo revela seu carácter social, como um fenómeno distinguido pela sociedade entre as

múltiplas ações que se orientam para o entendimento intersubjetivo; como um verdadeiro gênero da significação.”

O dicionário é, assim, usado de diferentes formas e com diversas intenções, sendo muitas vezes a sua elaboração também distinta consoante os propósitos que visa. Contudo, devido ao facto de poder ser consultado por muitos utilizadores, com formações e interesses distintos, um dicionário é uma obra que muitas vezes é severamente criticada a vários níveis, nomeadamente palavras que estão ou não atestadas, definições em falta, etimologia e/ou exemplos de uso controversos. A sensação de pertença que a língua desperta na comunidade linguística reflete-se na forma como o falante se relaciona com o dicionário, despertando reações fortes, críticas e polémicas<sup>12</sup>. Apesar dos inevitáveis erros e gralhas que obras desta envergadura possam apresentar, o teor e a intensidade das reações que suscita podem explicar-se pelo inquestionável papel de referência, autoridade e legitimidade que é atribuído ao dicionário.

É possível distinguir dois grupos de utilizadores dos dicionários: o público que consulta dicionários no seu quotidiano e o público especialista, que inclui os compiladores de dicionários.

No grupo de quem consulta dicionários estão, à partida, todos os falantes, desde estudantes a profissionais de diversas áreas, que os podem consultar quer por questões profissionais, académicas ou escolares, ou por lazer, como, por exemplo, para validar a ortografia de algumas palavras (por exemplo, no contexto da realização de palavras-cruzadas ou em jogos do tipo Scrabble®). No grupo de quem elabora o dicionário estão os lexicógrafos, que possuem uma formação académica específica, mas também

---

<sup>12</sup> A maioria do público desconhece que, em PE, os dicionários são considerados obras de referência com pendor descritivo e não dicionários normativos e prescritivos, como acontece noutras línguas. Razão pela qual ficam surpreendidos quando verificam pronúncias de outras zonas do país apresentadas ou unidades lexicais sem a marca de regionalismo, por acreditarem que o dicionário deve apenas apresentar a norma padrão. Outros aspetos que são frequentemente alvo de crítica, de censura, ou de indignação são (i) as atestações de empréstimos não adaptados ao PE, quando muitos consideram que já existem no PE unidades lexicais que representam essa realidade (como *bullying* quando já existe *acossamento*, *violência escolar*), (ii) os aportuguesamentos quando se verifica apenas parcialmente a adaptação à grafia do PE (o exemplo de *stresse*), (iii) a unidade lexical não apresentar todas as polissemias existentes na língua, e por fim, (iv) a ideia de que se uma unidade lexical não está atestada, esta não existe, ignorando neste caso que o dicionário é um repositório finito das unidades lexicais de uma língua.

autodidatas, um conjunto de pessoas que se interessam pela prática lexicográfica e que, apesar de não possuírem qualificação específica, desenvolvem uma série de produtos lexicográficos, mostrando a necessidade dos mesmos e também o lado necessariamente cultural dos dicionários (Correia, 2009:17).

Relativamente à formação dos lexicógrafos, houve um período em que, normalmente, eram pessoas com projetos pessoais que se dedicavam exclusivamente a essa tarefa durante vários anos. Contudo, atualmente, é comum ser uma editora a formar uma equipa de lexicógrafos para a criação de um dicionário (Duran & Xatara, 2007). Rundell (2002) chama a atenção para a formação dos lexicógrafos em ambiente académico, importante para a melhoria da qualidade da produção lexicográfica, pois permite aliar a teoria à prática, além de permitir introduzir determinadas descobertas imediatamente em certas produções lexicográficas. No entanto, em muitas editoras, a formação das equipas é fortemente direcionada apenas para o lado prático e de mercado, o que poderá não permitir a aplicação imediata (ou mais rápida) das descobertas e a cooperação académica.

Para a elaboração de um dicionário, entende-se que o lexicógrafo deve (i) estar ciente dos aspetos lexicológicos e metalexicográficos, (ii) definir a macro e a microestrutura da obra, (iii) decidir se utilizará a conceção polissémica ou homonímica no que respeita ao artigo lexicográfico, (iv) determinar a forma de apresentar a sinonímia, as variantes, os contextos, entre outros.

O papel do lexicógrafo será o de executar o projeto, ao passo que o metalexicógrafo terá como preocupação o estudo do dicionário, da metodologia e dos fundamentos teóricos para a sua conceção e elaboração, e compreender que tipo de usos são feitos pelos utilizadores e como estes se podem refletir na estrutura e organização interna do dicionário.

Embora haja distinções entre ser lexicógrafo e metalexicógrafo, na prática, é comum estes dois perfis se sobreporem (Béjoint, 2000). Quando tal acontece eficazmente, observa-se um produto lexicográfico de qualidade e inovador para a comunidade linguística (Duran & Xatara, 2007).

#### 2.1.4.1. MACROESTRUTURA

A elaboração de um dicionário é uma tarefa intrincada, que deve ter em conta aspetos teórico-metodológicos para garantir organização, coerência, sistematicidade e qualidade, por forma a que o produto final seja reconhecido como adequado à comunidade a que se destina. O dicionário possui duas componentes estruturais essenciais: a macroestrutura e a microestrutura (Baldinger, 1960; Rey-Debove, 1971).

A macroestrutura contempla as partes pré-textuais<sup>13</sup>, nomeadamente o prefácio, o guia de utilização, a lista de abreviaturas e símbolos utilizados na obra, e a nomenclatura (corpo do dicionário — número de verbetes — conjunto dos artigos lexicográficos). Alguns dicionários contemplam ainda apêndices, opcionais, como, por exemplo, o apêndice gramatical com os paradigmas de conjugação verbal e/ou apêndices de palavras e expressões estrangeiras. Apesar de não serem obrigatórios, dependendo da sua qualidade, os apêndices podem ser uma mais-valia para o utilizador.

O prefácio (ou a introdução ou a nota de editor) é de maior importância, uma vez que é nesta secção que são explicitadas as informações relativas às fontes consultadas, a finalidade da obra, o público-alvo, os critérios utilizados na sua elaboração, incluindo os usados na seleção da nomenclatura.

No guia de utilização (ou instruções de uso), deve estar explicitado o modo como a obra está organizada, assim como a exemplificação de artigos lexicográficos — uma entrada e o respetivo verbete (ou mais do que um) — com os elementos componentes identificados e explicitados, para uma melhor compreensão da organização interna da obra e da informação nela contemplada, de modo a auxiliar o utilizador a compreender corretamente o artigo lexicográfico, como se pode observar na Figura 2, a título de exemplo.

---

<sup>13</sup> Existem diferentes denominações consoante as línguas e os autores: *textes externes* em francês (Boulanger, 1995), *megastructure* em inglês (Hartmann & James, 1998).

Como usar o dicionário		
<p><b>a entrada e a palavra a definir</b></p> <p><b>acordar</b> v. 1 despertar (alguém) do sono 2 dar voltar a si 3 arr. decidir, resolver 4 para si ficar estupefado</p> <p><b>acorde</b> nm. 1 produção simultânea de três ou mais notas 2 som musical</p> <p><b>acordeão</b> nm. instrumento musical de palhetas livres que vibram por ação de um fole, com teclado num lado e teclado no outro, que se abre e fecha para produzir som</p> <p><b>acordeonista</b> m. 2y músico que toca acordeão</p> <p><b>acordo</b> nm. 1 enquadramento recíproco 2 consentimento, conformidade 3 pacto, tratado, convenção; <b>acordo de cavalheiros</b> pacto verbal, em que as partes envolvidas dispensam formalidades legais; <b>acordo ortográfico</b> convenção que estabelece normas de como escrever corretamente as palavras de uma determinada língua e que, em termos legais, prevê a assinatura dos representantes oficiais dessa mesma língua 4 <b>chegar a um acordo</b> alcançar um consenso; <b>de acordo com</b> segundo a opinião ou informação de; <b>de comum acordo</b> com o consentimento de todas as partes envolvidas; <b>estar de acordo com</b> concordar com, ser da mesma opinião</p> <p><b>açoriano</b> adj. dos Açores (arquipélago português, no Oceano Atlântico) 2 n. natural dos Açores</p> <p><b>acorrer</b> v. 1 ir/veni acudir, socorrer (alguém) 2 ir ir apressadamente</p>	<p><b>categoria gramatical</b></p>	
<p><b>definição</b></p>	<p><b>acupuntor</b> ou <b>acupuntora</b> nm. especialista em acupuntura</p> <p><b>acupuntura</b> ou <b>acupuntura</b> nf. método de origem chinesa para tratar doenças e aliviar dores por meio de picadas com agulhas muito finas, em pontos específicos do corpo</p>	<p><b>separa diferentes categorias gramaticais</b></p> <p><b>grafas duplas introduzidas pela reforma ortográfica</b></p>
<p><b>introduz expressões e locuções no final da entrada</b></p>	<p><b>adjetivo</b> s. nova para s. <b>adjetivo</b></p> <p><b>adjetivação</b> nf. emprego de adjetivos</p> <p><b>adjetival</b> adj. 2y relativo ou pertencente ao adjetivo</p> <p><b>adjetivização</b> nf. processo morfológico que consiste em atribuir a função de adjetivo a uma palavra de outra classe gramatical</p> <p><b>adjetivar</b> v. 1 qualificar com adjetivo(s) 2 empregar adjetivos 3 atribuir função de adjetivo a</p> <p><b>adjetivo</b> nm. palavra que flexiona em gênero, número e em grau + <b>adjetivo numeral</b> o que indica o lugar de ordem ocupado por uma entidade numa sucessão ou conjunto ordenado, ocorrendo geralmente antes do nome</p>	<p><b>separa diferentes categorias gramaticais</b></p> <p><b>grafas duplas introduzidas pela reforma ortográfica</b></p> <p><b>assina grafas novas ou grafas duplas</b></p>
<p><b>as preposições que acompanham o verbo são assinaladas entre parênteses angulares</b></p>	<p><b>amigo</b> nm. 1 pessoa que tem com outra uma relação de amizade; <b>ANT.</b> inimigo; <b>em amigo da oca</b> pessoa considerada amiga que toma uma atitude frívola ou desleal; <b>os amigo de Pêsoche</b> amigo íntimo; <b>amigo do peito</b> pessoa muito chegada 2 pessoa que tem admiração por uma determinada coisa; <b>SM.</b> apertador 3 partidarista, simpático 4 adj. que demonstra atenção + <b>fidelidade, probidade, inocência</b> <b>amigo do ambiente</b> que ou o que evita o impacto negativo no ambiente</p> <p><b>amimar</b> v. dar mimo a, <b>SM.</b> acariciar</p> <p><b>aminação</b> nm. 2y amonidação</p> <p><b>aminoácido</b> nm. nome geral dos compostos orgânicos que são os componentes essenciais das proteínas</p> <p><b>amistoso</b> adj. 1 próprio de amigo; <b>SM.</b> amigável 2 <b>SM.</b> (1990) disputado não para competição, mas para treino, confraternização ou por beneficência</p> <p><b>amizade</b> nm. com frequência</p> <p><b>amizade</b> nf. sentimento de grande afeto ou simpatia por alguém; <b>ANT.</b> amizade 2 lealdade entre amigos; <b>SM.</b> companheirismo 3 intencionalidade e cooperação entre pessoas, entidades, países; <b>SM.</b> aliança 4 pessoa amiga</p> <p><b>amnésia</b> nf. perda total ou parcial da memória</p> <p><b>amnesiar</b> v. causar amnésia a</p>	<p><b>o contexto em que se usa a palavra é indicado entre parênteses curvos</b></p> <p><b>a letra gráfica onde a palavra é usada indica-se de forma abreviada</b></p> <p><b>o nível de uso da palavra indica-se de forma abreviada e em itálico</b></p> <p><b>remete para a palavra onde se encontra a definição</b></p> <p><b>os algarismos destacam diferenças sentidas da entrada</b></p>
<p><b>as preposições que acompanham o verbo são assinaladas entre parênteses angulares</b></p>	<p><b>amnésia</b> nf. perda total ou parcial da memória</p> <p><b>amnesiar</b> v. causar amnésia a</p>	<p><b>separa diferentes categorias gramaticais</b></p>
<p><b>grafas anteriores à reforma ortográfica</b></p>	<p><b>corta-vento</b> nm. 1 blusão de tecido resistente à chuva e ao vento, usado sobretudo em desportos de montanha 2 dispositivo colocado à frente dos veículos de grande velocidade, para atenuar a pressão do ar</p> <p><b>corte</b> m. 1 sentimento feito com objeto cortante; <b>SM.</b> golpe 2 abate (de árvores) 3 quebra (de relações) 4 debaste (de cabelo) 5 talhe (de roupa)</p> <p><b>corte</b> m. 1 residência de um monarca 2 conjunto de pessoas que frequentam a residência de um monarca 3 número, galanteio 4 <b>SM.</b> antiga assembleia dos representantes da nação</p>	<p><b>os algarismos em destaque distinguem palavras com a mesma grafia e pronúncia diferente</b></p> <p><b>a pronúncia das vogais de palavras homógrafas é indicada entre barras oblíquas</b></p>
<p><b>grafas novas introduzidas pela reforma ortográfica</b></p>	<p><b>CPLP</b> sigla da Comunidade de Países de Língua Portuguesa</p> <p><b>CPU</b> nm. unidade central de processamento, componente do computador em que estão os elementos que interpretam e executam todas as tarefas</p> <p><b>CMS</b> Sigla de Central Processing Unit</p> <p><b>Cr</b> símbolo químico de cromo</p> <p><b>crachá</b> nm. 1 insignia 2 chapa identificativa</p>	<p><b>separa diferentes categorias gramaticais</b></p> <p><b>observações que fornecem detalhes sobre a entrada</b></p>

Figura 2. Exemplo de um guia de utilização retirado de Dicionário da Língua Portuguesa (2018).

A nomenclatura — o corpo do dicionário — consiste na listagem das entradas, que podem ser organizadas alfabética ou tematicamente. Esta pode ser maior ou menor, consoante a decisão do lexicógrafo e o tipo de dicionário que se pretende elaborar. Relativamente ao número de verbetes, este pode variar conforme os autores. De acordo com Biderman (1984), um dicionário infantil deve conter aproximadamente 5000 verbetes; um dicionário escolar entre 10 000 e 30 000 verbetes; um dicionário monolíngue padrão cerca de 50 000 verbetes e um *thesaurus* entre 100 000 e 500 000 verbetes. Já Martinez de Sousa (1995) concorda com os números referentes aos dicionários infantil e escolar, mas, em vez de falar em dicionário padrão ou *thesaurus*, refere a existência de um dicionário de língua seletivo e normativo que deve conter entre 50 000 e 100 000 verbetes, e de um dicionário de língua descritivo com mais de 100 000 verbetes. Por sua vez, Welker (2004) refere que um microdicionário deve conter até 5000 verbetes, os

minidicionários entre 5000 e 20 000 verbetes, um dicionário pequeno ou de bolso entre 20 000 e 50 000 verbetes, um dicionário médio ou de mão entre 50 000 e 100 000 verbetes e, por fim, um dicionário grande deve incluir mais de 100 000 verbetes.

É de realçar, no entanto, que, independentemente do número de verbetes, “nenhum dicionário, por mais volumoso que seja, dará conta integral do léxico de uma língua de civilização” (Biderman, 1998:130). Além disso, o número de verbetes (aspecto muito relevante para a publicitação e promoção de um dicionário) nem sempre corresponde ao número de verbetes indicado pelos diversos autores, pois, em algumas circunstâncias, o número divulgado contempla as entradas, outras vezes contempla as subentradas que surgem dentro do verbete principal, outras vezes contempla também locuções.

Outro aspecto que merece alguma reflexão é a forma como se chega à determinação não só do número, mas também das entradas que vão figurar na obra. Alguns autores usam exclusivamente fontes lexicográficas e baseiam a nomenclatura das obras de menor dimensão nas obras maiores da própria editora (por exemplo, caso da nomenclatura de um dicionário monolíngue ser utilizada para construir um dicionário bilingue), outros procuram seguir e/ou incluir as nomenclaturas de obras de referência anteriores, assumem essa posição publicamente como “inevitável” (como no caso do Dicionário Aurélio), existindo também casos em que há evidência de *inspiração* em obras anteriores, ainda que não estas não sejam explicitamente mencionadas como fontes lexicográficas.

Há outros autores que, apesar de não descartarem o trabalho anteriormente feito começam a apoiar o seu trabalho em *corpora*. A Linguística de *Corpus* pode, assim, ser usada como uma metodologia ou como abordagem para a análise de textos de especialidade.

Na época pré-digital, uma das dificuldades em usar *corpora* prendia-se com a morosidade em recolher, organizar e codificar os textos<sup>14</sup>. Se antes se entendia como

---

<sup>14</sup> Recorde-se os trabalhos de Thorndike (1921) que criou manualmente um *corpus* com 4,5 milhões de palavras, ou Quirk (1954) com a sua equipa que criaram o *Survey of English Usage* de forma manual e que após anos de trabalho e modificações ficou conhecido em 1989, como o *London-Lund Corpus*, um *corpus* computadorizado e que foi a base para os futuros etiquetadores.

*corpora* as obras de literatura (ou as obras de referência de um domínio) que eram consultadas pelos lexicógrafos para apresentar os exemplos em alguns dicionários (como atestam o *Corpus* Helenístico e outras obras que citam passagens da Bíblia, indicando que desde a Antiguidade era comum esta prática), hoje, o entendimento do uso de *corpora* é distinto. Com o surgimento da Linguística de *Corpus*<sup>15</sup>, apoiada nas novas tecnologias, revolucionou a área da lexicografia e o uso de *corpora* passou a ter um papel relevante, seja na criação de *corpora* para compreender e descrever os fenômenos de uso da língua ou para auxiliar no ensino da língua materna e não materna.

A história da Linguística de *Corpus* está associada à compilação e disponibilidade de *corpora* e são tidos como referências internacionais os seguintes casos: o *Brown Corpus* (Brown University Standard Corpus of Present-Day American English) de 1964 o primeiro a atingir um 1 milhão de palavras, o *Birmingham Corpus* (Birmingham University International Language Database) de 1987 com 20 milhões de palavras e que deu origem ao Bank of English, e o *British National Corpus* de 1995 com 100 milhões de palavras, considerado um mega-*corpora*. Os *corpora* para a língua inglesa foram pioneiros e abriram espaço para a criação de trabalhos de referência em várias línguas, como é o caso dos seguintes trabalhos para o português europeu, como o Tycho Brahe (com 3 milhões de palavras dos séculos XVI ao XIX, o CetemPÚBLICO (com 180 milhões de palavras do século XX), o *Corpus* do Português (com 45 milhões de palavras dos séculos XV ao XX) só para citar alguns.

O uso de *corpora* pode auxiliar a elaboração de dicionários, seja através (i) da abordagem “corpus-based”, onde o *corpus* serve essencialmente para o lexicógrafo que tem uma intuição em relação a uma palavra verificar se ela ocorre no *corpus* e quais os seus contextos de uso, permitindo a extração de exemplos para os diferentes significados; ou de (ii) da abordagem “corpus-driven”, que permite ao lexicógrafo observar o *corpus* como ponto de partida e através da observação das palavras e das frequências, analisar e

---

<sup>15</sup> Se, num determinado período da história, a Linguística de *Corpus* perdeu algum terreno perante o impacto da perspectiva de Chomsky, o surgimento de novas capacidades do computador veio possibilitar um avanço enorme na Linguística de *Corpus* que se irá desenvolver paralelamente a outras disciplinas e subdisciplinas, como as Ciências da Computação, o Processamento de Linguagem Natural ou a Linguística Computacional.

agrupá-las por categorias e a partir daí elaborar o artigo lexicógrafo de acordo com os dados obtidos.

Neste trabalho, considera-se que o uso de *corpora* e os métodos da Linguística de *Corpus* devem ser combinados para a elaboração de dicionários, mas devem ser considerados alguns aspetos, não devendo ser utilizados como base única para elaboração de um dicionário, mas antes como apoio. Assumindo que o *corpus* reflete um conjunto de características selecionadas pelo lexicógrafo (representar um determinado domínio, ser produzido por especialistas desse domínio, etc.), que funciona como repositório/base de dados (como apontam Biber, Conrad e Reppen 1998:4) e que não é apenas um conjunto de textos recolhidos, como referido por Lino (2006:509): “Un corpus n'est ni un ensemble d'énoncés ni une archive de textes; ils se distingue d'une base textuelle, organisée sans des objectifs précis”. Um *corpus* funcionará, assim, como fonte de informação complementar à dada pela consulta de especialistas como consultores de projetos lexicográficos, da qual é possível extrair informação (exemplos de uso, contextos) que auxilia a produção de conteúdo fiável sobre o domínio em questão. É possível, assim, aliar, por um lado a informação extraída de uma representação textual de um domínio de especialidade, onde se podem observar as unidades lexicais em uso, ao contributo do especialista que auxilia na organização dos dados de forma explícita de acordo com o que se pretende.

Advoga-se neste trabalho que a criação e uso de um *corpus* de especialidade deve refletir “réalité matérielle accessible, analysable et évaluable” (Slodzian, 2000: 78) e permitir a visualização de frequência para determinar unidades lexicais mais frequentes e menos frequentes, critério muito utilizado para a inserção de unidades lexicais na nomenclatura, aspeto aliás bastante comum desde 1998 com o uso de concordanceadores associados aos *corpora* disponibilizados. A utilização de *corpora* de especialidade num projeto desta natureza é também relevante porque, apesar da existência de *corpora*, muitos dicionários nem sempre são elaborados tendo em conta estes recursos, e dependem tendencialmente de especialistas como consultores (Norman, 2002), por um lado, e por outro, um grande número de dicionários de especialidade continua a ser elaborado por pessoas sem formação em Lexicografia ou Linguística, mas por especialistas que nem

sempre têm como preocupação aspetos como a sistematização e organização da macro e microestruturas (Williams, 2002).

O uso de *corpora* é importante e relevante na elaboração de obras lexicográficas e o ambiente digital proporciona maior facilidade para o lexicógrafo, mas também para o utilizador que pode e deve beneficiar de recursos que possibilitam a pesquisa de unidades lexicais nesses *corpora* (Gouws, 2021). O uso de *corpora* permite ao lexicógrafo ir além da sua intuição de falante e observar as unidades lexicais num determinado contexto linguístico com a possibilidade de verificar qual o ambiente sintagmático onde determinadas unidades lexicais ocorrem (Geeraerts, 2009) e quais as relações lexicais entre várias unidades lexicais (Stubbs, 2002:24). Os métodos quantitativos e as ferramentas computacionais permitem análises robustas e potenciam observações das unidades lexicais que nem sempre são possíveis a “olho nu”, reduzindo a complexidade da análise e facilitando a interpretação de aspetos como a frequência de unidades lexicais (as mais frequentes, as menos frequentes ou mesmo a computação de distâncias entre tabelas de frequências), a similaridade semântica entre unidades, a probabilidade de uma determinada unidade se relacionar com outras em determinados contextos, etc., facilitando ao lexicógrafo a compreensão das unidades lexicais e das suas especificidades, e o acesso a informações atualizadas dos dados de uso na língua (McEnery e Wilson, 2004: 106), que podem ser incorporadas num dicionário.

Anteriormente, elaborar um dicionário era uma tarefa morosa e intrincada, porque dependia de um trabalho manual, de cartões/fichas em papel ordenadas e elaboradas por um ou mais lexicógrafos. Atualmente, no entanto, as condições de produção lexicográfica alteraram-se consideravelmente, o que, em certas situações parece poder potenciar ainda mais casos de “criminalidade lexicográfica” (tradução de Welker, 2004:87), descrita por Hausmann (1989) como “dictionary criminality”, e que é definida como “qualquer forma de fraude, incluindo plágio, praticado por um autor, compilador, editor ou editora de um dicionário ou de um trabalho similar; direitos de autor” (Hausmann, 1989; Béjoint, 1994). Se, por um lado, parece haver um fundo de verdade no facto de os trabalhos lexicográficos partirem sempre de uma base existente (ou de, pelo menos, seguirem as mesmas nomenclaturas), que, consoante as necessidades, era adaptada, aumentada ou diminuída,

por outro lado, as questões de direitos autorais e de propriedade intelectual são assuntos delicados e sobre os quais se deve seriamente refletir.

Estas secções pré-textuais são de maior importância. Ainda assim, sabe-se que nem sempre os utilizadores direcionam a sua atenção para esta parte, como mencionam Haensch e Omeñaca (2004), e que, noutros casos, nem sempre os dicionários explicitam convenientemente os critérios e fontes utilizados, o que dificulta a consulta e compreensão do próprio dicionário (Camilotti, 2011). Apesar de todos os autores considerarem que as instruções devem ser claras, há diferenças quanto à dimensão ideal quer do prefácio, quer do guia de utilização, uma vez que alguns autores consideram que estes devem ser textos pequenos porque, regra geral, o utilizador não quer perder tempo na sua leitura (Herbst, 1996), enquanto outros recomendam que deve ser um texto longo e detalhado (Welker, 2004:80).

#### 2.1.4.2. MICROESTRUTURA

Se a macroestrutura diz respeito às decisões tomadas relativamente à obra, tais como quais as entradas (formas de citação, lemas, ou unidades lexicais) que devem constar na nomenclatura (geralmente apresentadas por ordem alfabética), a microestrutura refere-se à estrutura e organização da informação de natureza diversa que o artigo lexicográfico deve apresentar.

Num primeiro momento, são determinados os aspetos relacionados com o tipo de informação que pode figurar dentro do artigo lexicográfico. Consoante as decisões tomadas, apresentar-se-á transcrição fonética, sinónimos e antónimos, marcação de regionalismos, arcaísmos, neologismos, exemplos de uso<sup>16</sup>, ilustração, variantes, registos, entre outras opções possíveis, para além da decisão acerca do tipo de definição a usar.

---

<sup>16</sup> Autores como Sinclair (1991), Martínez de Sousa (1995), entre outros, entendem que “No mundo dos lexicógrafos, polariza-se em dois extremos: por um lado, aqueles “incautos” que se atrevem a incluir exemplos inventados em seus dicionários; e frente a estes, o lexicógrafo ‘competente’, que somente recorre a exemplos reais, extraídos do *corpus* com que trabalha (Sinclair, 1991 *apud* Calderón Campos, 1994:67)”. No entanto, Binon e Verlinde (2000:106) consideram que os exemplos devem ser extraídos de *corpora*,

A definição é um aspeto central nos dicionários e uma das tarefas mais complexas. Trata-se da criação e organização de um pequeno texto com a informação codificada e que requer um profundo grau de conhecimento da técnica lexicográfica e alguma prática na sua elaboração, por forma a adequar-se ao tipo de obra lexicográfica e às necessidades do público-alvo. Na elaboração da definição lexicográfica é, regra geral, tido em conta um conjunto de características (elementos portadores de significação) que ajudam a definir a unidade lexical, muitas vezes apoiado na observação e análise de contextos e exemplos extraídos de *corpora*, e que deve ser redigida de acordo com princípios e regras de forma a permitir uma leitura clara e uma compreensão da unidade lexical em questão.

De acordo com Lino (2018), os dicionários de língua corrente podem ter diferentes conceções (polissémica, homonímica, enciclopédica) e podem apresentar na definição lexicográfica um número significativo de aceções polissémicas, locuções, exemplos.

O texto definitório tende a apresentar uma configuração clássica na qual pode observar o *genus proximum* e o *differentiam specificam*, ou seja, o género próximo e a diferença específica que distingue a unidade lexical de outras da mesma categoria geral. A definição é, regra geral, elaborada a partir dos contextos de uso, extraídos dos *corpora* quer em língua corrente, quer numa língua de especialidade. Estes contextos são portadores ou indicadores de elementos de significação, que ilustram o modo como é usada a unidade lexical e que permitem elaborar uma definição, e, em alguns casos, constituem eles próprios uma forma de definição (Moreira, 2010).

A definição lexicográfica é em si um aspeto de maior relevância, constituindo um texto escrito em língua natural que deve ser claro e objetivo, que deve obedecer a um

---

mas devidamente adaptados, para se tornarem mais esclarecedores. Aspeto que Drysdale (1987) refere como alerta para a dificuldade em encontrar exemplos que reúnam as características pretendidas para o tipo de obra e tipo de público-alvo. Da mesma opinião Laufer (1992), *apud* Calderón Campos (1994:167), acredita que os exemplos inventados, ainda que menos naturais, serão didaticamente mais úteis, pois caracterizam-se por serem mais fáceis de entender, posto que as necessidades dos usuários são levadas em conta. E, por fim, Calderón Campos (1994), que considera que "De todas as formas, a polémica em torno da autenticidade dos exemplos me parece um problema secundário. O que é verdadeiramente importante é que os exemplos sejam representativos, naturais, imitáveis e levem à curiosidade e ao interesse dos alunos e dos usuários, de um modo geral, a utilizar novas palavras. Se conseguirmos atingir esses objetivos, pouco importa que os exemplos sejam inventados ou autênticos." (Calderón Campos, 1994:68).

padrão de formulação uniforme e que dê conta das especificidades de significação da unidade lexical. Vários autores defendem padrões de redação rígidos para definições. A rigidez e lógica de elaboração da definição lexicográfica podem fazer pouco sentido para o utilizador do dicionário, mas são de extrema relevância para o lexicógrafo e a sua equipa, na medida que são um instrumento de garantia de coerência e consistência para o resultado (Finatto, 2003).

Há também uma preocupação na forma como se apresentará toda a informação pretendida dentro do artigo lexicográfico, o que, no entender de Dubois e Dubois (1971), é metaforizado como sendo um enunciado (artigo lexicográfico) com um sujeito (entrada) e um predicado (definição), aspeto que Rey-Debove (1971) define como sendo o conjunto de informações ordenadas de cada artigo após as entradas e que corresponde *grosso modo* à *microestrutura básica* de Barbosa (1995), que pode ser ampliada com um conjunto de informações ordenadas, como se apresenta na Figura 3.

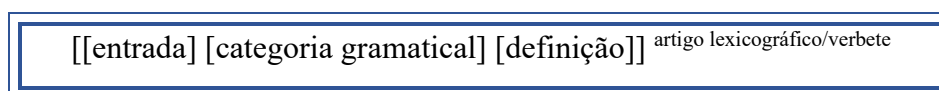


Figura 3. Microestrutura básica essencial/mínima do artigo lexicográfico (baseado em Barbosa, 1995)

Evidentemente, esta estrutura tradicional, além de mínima/essencial, é questionada por vários autores, que consideram que, apesar de ser vantajoso e desejável que exista coerência e recorrência na microestrutura do artigo lexicográfico, pode e deve ser adaptada consoante a categoria gramatical (um artigo lexicográfico de um nome terá uma microestrutura distinta de um verbo, por exemplo).

Neste contexto, pode colocar-se a hipótese de uma microestrutura possível do ponto de vista lexicográfico, mas que, do ponto de vista do utilizador, pode ser confusa quer pelo elevado número de informação disponibilizada, quer pelo grau de descodificação necessário para selecionar e/ou identificar que informação consultar de forma rápida e eficaz, pelo que o estudo e definição da microestrutura de um dicionário é um aspeto essencial.

A organização da informação no interior do artigo lexicográfico (Farias, 2009:129) costuma considerar dois blocos de constituintes; um relativo à informação sobre a entrada,

a categoria gramatical, a transcrição fonética, a informação morfológica, e outro relativo aos aspetos semânticos, como o registo de uso, definição, paráfrases, equivalentes, exemplos, sinónimos, remissões, ilustrações, entre outros elementos.

Num segundo momento, são definidos os aspetos relacionados com a formatação de cada tipo de informação/identificação dos elementos das entradas. Por exemplo, o negrito indica a entrada, o itálico a categoria gramatical, a numeração (árabe ou romana) indica distinção das diferentes definições relativas à mesma entrada, as abreviaturas dos registos/variantes (geográficas, diacrónicas, entre outras) são grafadas em itálico, qual a localização da etiqueta de domínio, quais os símbolos para remissões (consulta de outras entradas), quais os símbolos usados na etimologia (para explicar adição/cruzamento de elementos ou de formas hipotéticas), quais os símbolos ou espaçamento usados para separar categorias gramaticais dentro da mesma entrada, qual o formato (ex.: parêntesis curvos ou retos) para apresentar contextos, entre outras codificações possíveis e existentes e geralmente usadas nos dicionários.

A estrutura da microestrutura do dicionário pode ser variável consoante os critérios estabelecidos pelos autores. No entanto, considerando que o foco é o dicionário (seja ele em papel ou digital) as preocupações do lexicógrafo na era da Lexicografia Digital não são apenas relativas à macro e microestrutura e ao tipo de informação a figurar nos artigos lexicográficos, em termos de qualidade de conteúdo, mas também em igual proporção e preocupação as necessidades e o perfil do utilizador, que se sabe serem igualmente importantes aquando da criação do dicionário de um domínio de especialidade. Atualmente, valoriza-se, por um lado, o conteúdo e, por outro exploram-se as novas tendências lexicográficas adaptadas ao utilizador, tentando perceber quais os aspetos que este considera fundamentais aquando as suas pesquisas, em determinados casos não alinhados com a tradição lexicográfica, podendo levar a um afastamento público-alvo.

#### 2.1.4.3. TIPOS DE DICIONÁRIOS

Falar de dicionários é também falar de tipologias, de classificações e de funções dos dicionários, que dependerão da prática, da metodologia e dos critérios adotados pelo

lexicógrafo. Nesse sentido, há vários autores, como Scerba (1940), Malkiel (1959/1962), Wagner e Quemada (1969), Rey (1970), Al-Kasimi (1977), Haensch (1982), Hausmann (1985), Martinez de Sousa (1995), Béjoint (2000), Hartmann e James (1998), Welker (2004), Krieger (2006) entre outros, que se dedicaram a refletir sobre a diversidade de obras lexicográficas e apresentaram distintas categorizações.

De uma forma sucinta, pode-se classificar um dicionário tendo em conta:

- o número de línguas presentes: monolíngue, bilingue, multilíngue;
- o objetivo de utilização: compreensão ou produção de texto (codificação e descodificação);
- a faixa etária do público-alvo: infantil, adolescentes, jovens, adultos;
- os critérios de extensão considerando o número de artigos lexicográficos presentes na nomenclatura –, do menor para o mais exaustivo: básico, escolar, integral, etc...;
- a nomenclatura selecionada: língua geral, de especialidade (terminológico), especial (neologismos, sinónimo, histórico);
- a organização da macroestrutura: semasiológicos, onomasiológico/analógico, visuais;
- o tipo de informação da microestrutura: comum, pedagógico, especializado); aprendizagem;
- o tipo de definição: língua geral, misto, enciclopédico;
- o intervalo temporal coberto: sincrónico, etimológico, histórico;
- o tipo de instrumentalização/perspetiva: descritivo, normativo, referência;
- o suporte: papel, eletrónico, digital.

De acordo com o cruzamento e a relevância dada a estes fatores, vários autores defendem diferentes tipologias, classificações e funções dos dicionários.

O dicionário pode existir em diferentes suportes, aspeto aliás que não é inovador, bastando recordar o que se designa como “protodicionário” ou “paleolexicografia” (Boisson, Kirtchuk & Béjoint, 1991: 262), uma vez que, em tempos, houve obras lexicográficas em tábuas de argila, pedra ou pele curtida.

Se é um facto que o que se entende por dicionário e suas características se tem alterado ao longo dos séculos, é também verdade que, no século XX, por vários motivos, nomeadamente o desenvolvimento da Linguística de *Corpus* e da Informática, se assiste a um novo período que se entende por Lexicografia e dicionarística contemporâneas.

No início deste período, Quemada (1987) fala sobre a reconceptualização do conceito de dicionário – Dicionarística<sup>17</sup> – e propõe novos termos para realidades recentes, como “Nova Lexicografia” (associada às novas teorias e metodologias de elaboração de dicionários informatizados), “Lexicomática” (associada ao tratamento automático do léxico, como os primeiros corretores ortográficos – ex.: FLIP, dicionários eletrónicos/dicionários-máquina para a tradução automática – SYSTRAN). Segundo Lino (2018)

“o conceito de dicionário mudou nestes últimos anos graças às novas tecnologias informáticas e telemáticas de elaboração e pesquisa hipertextual; existem muitos dicionários apenas em suporte electrónico ou telemático na Internet para diferentes públicos; existem dicionários só para as máquinas: computador ou telemóvel (léxicos computacionais, léxicos fonológicos para o telemóvel)” (Lino, 2018:19).

---

<sup>17</sup> Neologismo cunhado por Quemada (1987) “Conceito que tem por âmbito o domínio complexo do dicionário, constituído por todos os tipos de dicionários e por tudo o que lhes diz respeito, recobrando simultaneamente o campo do "dicionário-realização" e do "dicionário-objecto de estudo". O neologismo "dicionarística" veio colmatar, em determinada ocasião, uma carência nocional e denominativa, isto é, no momento em que a lexicografia deixou de implicar a realização do dicionário.” (Xavier & Mateus, 1992).

Os dicionário-máquina ou dicionários para computador serão, basicamente, linhas de código para serem lidas pela própria máquina, com um conjunto de informações diferentes consoante a finalidade, como forma ortográfica, lema, classe de palavras, sinónimos, informações para desambiguação de categorias e sentidos, etc.

Os dicionários eletrónicos, por sua vez, serão versões de dicionários em papel ou versões concebidas de raiz para o suporte eletrónico (disquetes e CD-ROM, inicialmente, .html, .xml, para disponibilização *online* atualmente) para serem utilizados pelo utilizador no seu computador e outros dispositivos (telemóvel, *tablet*) de uso pessoal.

Num determinado momento, a informatização do dicionário tinha por preocupação converter as edições em papel para o eletrónico, conservando, em grande medida, a apresentação tradicional dos conteúdos (e, por vezes, até a disposição das duas colunas, uso de abreviaturas e símbolos, etc.). Mais tarde, observa-se uma transição para relacionar a nomenclatura/base de dados (conjunto dos artigos lexicográficos) com novas opções de pesquisa (pesquisa avançada, por domínio, por palavras semelhantes, pesquisa nas definições, etc.) e permitir ao utilizador uma relação mais dinâmica com o dicionário. O primeiro exemplo em Portugal surge com a Priberam que, em 1997, lançou o corretor ortográfico FLIP e posteriormente o *Dicionário Priberam da Língua Portuguesa* online.

Assim, assiste-se primeiro à coexistência do dicionário em versão papel e em versão eletrónica (Tarp, 2012:116) e, posteriormente, a uma aposta direcionada para versão digital, com características distintas, iniciando-se o período da Lexicografia na era da *Internet* ou a Lexicografia eletrónica ou e-Lxicografia).

A constatação de que as edições digitais seriam amplamente utilizadas não apenas nos computadores de secretária e portáteis, mas também em dispositivos móveis como os *smartphones* e *tablets* motivou uma modificação naquilo que é o conteúdo e a forma de apresentação do dicionário, demonstrando que a “electronic lexicography is not just a continuation of print lexicography. Nor does it mean merely making print dictionaries available online or on some data carrier” (Dziemianko, 2018:678). E também permitiu o surgimento de uma nova categorização, que Fuertes-Olivera (2009:11) classifica de “dicionários institucionais e dicionários coletivos/colaborativos”. Os dicionários institucionais são, como o próprio nome indica, elaborados por casas editoriais ou

instituições, com uma equipa dedicada, e que podem ter um uso gratuito ou uma subscrição. Os dicionários coletivos/colaborativos são sempre gratuitos e são elaborados por utilizadores, como é o caso do *Dicionário inFormal* (2006–2021).

Vários são os trabalhos que salientam os benefícios dos dicionários digitais (Bergenholtz & Gouws, 2007; De Schryver, 2003; Dziemianko, 2012a, 2012b, 2016; Lew, 2011, 2012; Nesi, 1999, 2015; Prinsloo, 2005), como a questão do acesso fácil e rápido, o tipo de recuperação de informações com a ajuda de hipertexto, o espaço de armazenamento virtualmente ilimitado, a menor condensação textual, a melhor legibilidade das entradas ou as ligações diretas para fontes externas (como *corpora* ou bancos de citações), como refere em detalhe Dziemianko (2018:667).

O trabalho de Fuertes-Olivera e Tarp (2014), por exemplo apresenta uma categorização dos dicionários em cinco categorias dependendo da tecnologia associada:

- os *Copycats* (que seriam os dicionários em versão papel que foram apenas fotocopiados, fotografados ou digitalizados em formato imagem e apresentados em versão eletrónica, sem fazer uso das tecnologias);
- os *Faster Horses* (que seriam dicionários eletrónicos numa versão parecida com a versão impresso, mas que permitiam pesquisas avançadas);
- os *Stray Bullets* (dicionários digitais com novas visões e novas técnicas de utilização oferecendo aos utilizadores mais acesso a dados através da inserção de mais dados na microestrutura);
- os *Model T Fords* (dicionários digitais que possibilitam as pesquisas através do uso da técnicas da ciência da informação, como o uso de maior interatividade e ferramentas personalizadas de rápido acesso e que começam a ter em conta as necessidades do utilizador);
- e os *Rolls Royces* (dicionário digitais semelhantes à categoria anterior, com a particularidade de poderem permitir ao utilizador uma pesquisa personalizada quanto ao tipo de informação a surgir no ecrã, e também ao tipo de pesquisa poder ser também não apenas ao nível da unidade lexical, mas também articulada com

o acesso a bases de dados do dicionário ou de outros repositórios associados permitindo uma adaptação às necessidades do utilizador e da obra).

Verifica-se que os autores categorizam os dicionários consoante o tipo de tecnologia utilizada e não tanto pelo tipo de informação disponibilizada. Além disso, apesar da interessante forma de categorizar os dicionários, os autores falham em apresentar exemplos da última categoria alegando que ainda há não exemplos de nenhuma obra que preencha os requisitos.

Compreende-se que a Lexicografia passou da preocupação com a exaustividade em número de entradas, de definições dependentes, de informações extra associadas, como inventariadora do léxico da língua, para uma preocupação em agregar e transformar informação lexical em algo adequado (a vários níveis) aos seus utilizadores.

Embora haja consciência desta nova dualidade de preocupações por parte do lexicógrafo, parece que, na maioria dos casos, as principais questões se relacionam com o uso da Internet, o utilizador, o tipo de acesso, a utilização do *corpus* (*web-for-corpus* vs. *web-as-corpus*), a necessidade de cooperação, a adequação, mas também a competitividade do dicionário. Os aspetos supracitados são, de resto, apresentados por Lew & Schryver (2014), que os veem como positivos para que haja uma maior e melhor literacia digital. Segundo os autores, trata-se de uma interface necessária para que se possa lutar contra a ideia de que “dictionaries as isolated island of knowledge are dying” (Amsler, 2012), potenciando a revolução digital, na qual os nativos digitais (Prensky, 2001) podem utilizar o dicionário como um assistente digital (Nielsen, 2013:370). Se alguns autores acreditam que se está perante a existência de dicionários digitais para nativos digitais, outros autores acreditam que essa ideia não passa de um mito (Frawley, 2017)<sup>18</sup>.

Na presente tese de doutoramento, coloca-se a hipótese de que, ao longo dos tempos, foi crescendo o número de pessoas com acesso a dispositivos digitais, o que

---

<sup>18</sup> Disponível em <https://educational-innovation.sydney.edu.au/teaching@sydney/digital-native-myth/>

conduziu a um aumento do número de nativos digitais. Não obstante essa realidade, crê-se existir uma diferença entre ser nativo e ser proficiente. Simultaneamente, acredita-se que são os utilizadores nativos e proficientes digitais aqueles que melhor preparados estarão para construir dicionários digitais, ainda que não possuam, ainda, proficiência científica sólida na área da lexicografia. Tal facto origina uma lacuna entre o dicionário digital “ideal” e os seus elaboradores. Acredita-se, desta forma, que a proficiência na elaboração de dicionários digitais possa evoluir na mesma proporção em que tem vindo a evoluir a proficiência de consulta desses dicionários. Para tal, é essencial que novas propostas de dicionários sejam feitas e adequadas a ambos os atores deste cenário: aos utilizadores e aos lexicógrafos.

#### 2.1.5. DICIONÁRIO DE ESPECIALIDADE

Nesta secção será detalhado aquilo que se entende por dicionário de especialidade<sup>19</sup>, as suas características e a sua necessidade.

Elaborar um dicionário implica uma pesquisa aprofundada do objeto de estudo e a definição de um conjunto de aspetos relativos à metalexicografia que devem ser tidos em conta para que a obra lexicográfica seja reconhecida pelos lexicógrafos e pelo público como tal. No que respeita a um dicionário de especialidade, este não é completamente distinto do dicionário de língua geral, como menciona Lerat (1995), pois ambos podem possuir uma macro e microestrutura semelhantes, com campos selecionados em função das opções tomadas pelo lexicógrafo para um público em concreto, como aconteceria com um dicionário de língua geral para adultos ou para crianças, por exemplo, que seria resultado de critérios diferenciados.

Como anteriormente referido, existem diversas tipologias para classificar dicionários, sendo uma delas distinguir os dicionários gerais de língua (utilizados para

---

<sup>19</sup> Existem outras denominações como dicionário especializado, dicionário temático, dicionário terminológico, neste trabalho escolheu-se dicionário de especialidade.

verificar o uso corrente de uma palavra ou significado) dos dicionários de especialidade (que se referem a um domínio particular, como a medicina, a botânica, entre outros).

A necessidade de dicionários de especialidade é antiga e evoluiu bastante fruto da evolução da língua e das comunidades linguísticas, por um lado, e porque houve necessidade de construir obras de referência para melhorar e difundir as novas denominações técnicas e científicas entre os vários profissionais e especialistas destes domínios, por outro.

Diferentes autores definem o dicionário de especialidade e a sua utilidade de diferentes formas.

Hartmann e James (1998:129) definem dicionários especializados como obras de referência relativas a um conjunto restrito de fenómenos, tais como expressões idiomáticas, informação mais restrita de um determinado domínio, o jargão e os nomes técnicos, diferentes dos dicionários de língua geral e corrente.

Schmitt (2016), por exemplo refere que

“La formation de nouvelles langues professionnelles de plus en plus spécifiques due à l’accroissement régulier des métiers techniques et scientifiques contribue à un élargissement constant des dictionnaires actuels (...) le français actuel, est caractérisé par l’imposant développement des langues de spécialités et des langues techniques (technolectes)” et que le “problème sociologique qui prime au XXe siècle est celui des langues spécialisées et des langues techniques qui sont à présent plus intéressantes à étudier que les sous-registres des différentes classes sociales”. (Schmitt, 2016:53)

Por seu lado, Tarp (2012) acredita que

“Specialised lexicography has produced a big and growing quantity of practical products during the past two decades; however, when it comes to the quality of these products and the underlying theory that has to support and guarantee this quality, specialised lexicography – including terminography (...)” (Tarp, 2012:125)

Van Campehnoudt (2000:128), ainda, entende que “chaque dictionnariste de la langue spécialisée tente, avant tout, de concevoir un produit adapte à ses besoins et aux besoins de ses lecteurs potentiels”.

Apesar de todas estas (entre muitas) posições sobre o dicionário de especialidade, existe ainda a necessidade e a preocupação de distinguir o dicionário de especialidade de outras obras de referência de especialidade, como menciona L’Homme (2003): “I distinguish specialized dictionaries from term banks. The former usually bear on a specific field of knowledge or a few related fields (e.g., business and accounting; eletrconics and computing). Term banks are large sets of computerized records describing terms associated with a wide variety of domains.” (L’Homme, 2003:419)

Em qualquer domínio específico, os seus especialistas utilizam um léxico que pode ser específico do domínio, a par de unidades lexicais do léxico comum da língua geral, o que, em determinadas circunstâncias, pode gerar alguma incompreensão ou equívocos por aqueles que não dominam o léxico de especialidade ou que não estejam familiarizados com a comunidade de especialistas.

O dicionário de especialidade, como o nome sugere, está direcionado para contemplar um conjunto de unidades lexicais de um domínio em particular e ser utilizado por um conjunto de utilizadores que pretende informar-se sobre esse domínio, apresentando informação definitória ajustada a ser mais bem compreendida pelo público que o consulta.

Um dicionário possui, regra geral, um conjunto de características para ser considerado de especialidade. Como refere García Palacio (2002), por exemplo, terá (i) ser monolíngue; (ii) ter um público-alvo bem definido (para professores e alunos de uma determinada disciplina, por exemplo); (iii) ser descritivo quanto à nomenclatura apresentada (pois na maioria das vezes a pretensão é apresentar as unidades lexicais utilizadas, com uma possível perspetiva harmonizadora, aconselhar qual a unidade lexical que deve ser utilizada num determinado contexto, mas não de modo completamente prescritivo); (iv) oferecer um conjunto de informação linguística sincrónica e,

potencialmente, alguma informação enciclopédica; (v) apresentar a informação apresentada no sentido de permitir a produção, a comunicação e o uso.

Para além destas características, vários autores indicam que um dos aspetos mais relevantes no processo de elaboração de um dicionário de especialidade é o público-alvo destes dicionários – em particular o seu perfil e as suas necessidades (ou como na Teoria Funcional da Lexicografia é referido como os quatro pilares – os utilizadores, as situações, as necessidades e as funções do dicionário). No entanto, quando confrontados com as questões empresariais, sabe-se que a maior preocupação é que a questão do lucro de um produto disponível ao utilizador se sobrepõe a outras questões. Camacho Niño (2019) propõe ainda que, a par das preocupações relativas às funções, seja considerada uma microestrutura adequada a um dicionário de especialidade, lembrando que deve depender da função do dicionário, mas também dos utilizadores, da situação extralinguística e das necessidades lexicográficas relevantes e servir o propósito de apresentar informação lexicográfica e facilitar a apresentação desta na interface escolhida.

A elaboração do dicionário de especialidade, tarefa morosa, deve ter em atenção o texto definitório. Vários autores já estabeleceram as características que tal texto deve apresentar e quais os defeitos que devem ser evitados. Se, por um lado, a microestrutura básica deve conter um conjunto de informações básica/mínima ordenada, a definição de unidades lexicais especializadas deve conter neste texto informação referente aos traços essenciais e, possivelmente, aos traços acessórios se assim se decidir, como se apresenta na Figura 4, baseada em Barbosa (1995) e de Almeida, Souza e Pino (2007).

[[entrada] [categoria gramatical] [definição [traço essencial + (traço acessório)]] artigo lexicográfico/verbo
--

Figura 4. Estrutura do texto definitório da unidade lexical especializada do artigo lexicográfico

Outro aspeto que deve ser considerado é a definição dos traços essenciais e dos traços acessórios (Almeida, Souza e Pino, 2007). A definição da unidade lexical especializada deve:

- ser diferenciada por ser uma unidade única utilizada no domínio de especialidade, mas relacionada com as restantes;
- conter as características essenciais suficientes para reconhecer o significado da unidade;
- mostrar um compromisso entre a definição lexicográfica e a definição enciclopédica para alcançar a definição de especialidade (como refere Rey, 1979);
- ser redigida em língua natural tendo em conta a natural ambiguidade, polissemia, conotação, entre outros aspetos;
- ser cuidadosa e sistematicamente redigida e revista de modo evitar erros na construção do texto, tais como a definição tautológica, definição circular e a definição negativa (como lembra Dubuc, 1999: 120).

No sentido de elaborar o texto definitório, o lexicógrafo deve considerar que a unidade lexical especializada deve ser definida tendo em conta o domínio de especialidade em que se insere (mesmo que a unidade lexical possa existir noutros domínios ou como unidade lexical não-especializada com outros usos pelos utilizadores (Desmet, 2002: 182)) e pode (e deve) recorrer a um conjunto de princípios que facilitem a redação. Dubuc (1999), por exemplo, menciona que pode não existir um modelo para o texto definitório aplicável a todas as unidades lexicais especializadas, mas deverá haver um modelo que se possa adaptar consoante os traços que devem prevalecer na definição, o léxico e a elaboração da definição considerando o público-alvo. Podem, ainda, ser consideradas à partida algumas estruturas frásicas para início do texto definitório (“fórmulas de início: para os substantivos que expressam uma ação ou resultado dela, utilizam-se com frequência as fórmulas ação de e efeito de. Para definir as locuções dos substantivos e adjetivos que marcam um estado, podem-se utilizar as fórmulas ação, efeito de e estado de” (Almeida, Souza e Pino, 2007)).

Como lembra Desmet (2002), “determinar o que é necessário e suficiente não é uma tarefa aritmética e pode-se cair facilmente na definição hipo-específica, isto é, sem

caráter diferencial, ou hiper-específica, apresentando traços acidentais, não pertinentes ou supérfluos, próxima de definição enciclopédica” (Desmet, 2002: 180). Ter em conta todas as especificidades poderá ter como resultado definições maiores e menos concisas, aspeto que pode agradar aos utilizadores.

Parece consensual que a preferência dos utilizadores relativamente à organização da informação dentro do texto definitório no artigo lexicográfico depende, em certa medida, da combinatória entre o perfil do utilizador e as suas necessidades de consulta. Nesse sentido, apesar da informação que o lexicógrafo considera pertinente e relevante apresentar ao público-alvo tendo em conta as particularidades do domínio em questão, esta informação nem sempre pode ser apresentada de forma sistemática e exatamente com a mesma estrutura. Um dos motivos para que tal não seja necessariamente uma má opção de construção do texto definitório, prende-se com os aspetos comunicativos. Nem sempre é possível definir as unidades lexicais exatamente da mesma forma (uma forma nominal será definida de forma diferente de uma forma verbal), para além do facto de, apenas em casos excepcionais, os utilizadores consultarem todas as unidades de uma categoria para verificar se a microestrutura é exatamente igual.

Tendo em conta o perfil do utilizador e as suas necessidades, as preferências dos utilizadores relativamente ao tipo de definição que procuram e que consideram suficiente e pertinente podem ser distintas, podendo em determinadas obras lexicográficas considerar-se a existência de diferentes níveis de definições, como, por exemplo, um nível de definição adequado para não-especialistas, um nível de definição adequado para semiespecialistas e um nível de definição adequado para os especialistas. Nesta situação, o utilizador teria acesso imediato ao nível pretendido consoante o perfil que escolheu aceder, mas tendo sempre acesso aos restantes níveis (à distância de um clique para abrir na versão digital, por exemplo). Estes níveis de definição diferem quanto à complexidade da informação apresentada no texto e não necessariamente no tipo de informação. Ou seja, os traços essenciais devem estar presentes em todos os níveis, mas a forma pode ser mais clara num nível adequado para o não-especialista e mais complexa no nível para o especialista.

A título de exemplo, e considerando o domínio selecionado para este trabalho, o descritor “frutado” poderia ser apresentado tendo em conta diferentes níveis associados a um público-alvo como se apresenta na Tabela 2 (a noção de nível de especialidade será retomada em 4.2.3.1).

Nível	Definições para “frutado”
<i>não-especialista</i>	Vinho com aroma e/ou sabor que lembra frutos que não apenas a uva (citricos, tropicais, silvestres, ...).
<i>semiespecialista</i>	Vinho com aromas primários e/ou sabor a frutos, característica evidente sobretudo aos vinhos jovens. Nos vinhos tintos os aromas podem invocar frutos vermelhos; nos vinhos rosés frutos do pomar e nos vinhos brancos frutos cítricos ou tropicais, com maior ou menor intensidade.
<i>especialista</i>	Vinho com aromas primários e/ou sabor a frutos, consequência de uma vinificação em que se respeitaram as características da casta. São vinhos leves e delicados no paladar. Se forem provenientes de regiões frias, o percentual alcoólico é menor e a acidez mais viva; se forem de regiões mais quentes, o percentual alcoólico será mais elevado e a carga de taninos baixa.

Tabela 2. Texto definatório de “frutado” consoante os níveis de especialidade

Ter consciência do perfil dos futuros utilizadores do dicionário e daquilo que pretendem pesquisar auxilia os lexicógrafos a orientar o tipo e a forma da informação que deve constar na obra lexicográfica.

A proposta de um modelo de dicionário de especialidade que se apresenta neste trabalho pretende ser um contributo para a Lexicografia de Especialidade em PE e também para o domínio da ASE. A razão para a sua apresentação prende-se com os seguintes factos:

- as várias obras lexicográficas existentes em PE para a ASE não foram elaboradas de acordo com critérios lexicográficos;
- os critérios de elaboração das obras lexicográficas existentes não são apresentados nas obras;

- as obras lexicográficas apresentam algumas lacunas ao nível da nomenclatura e das definições;
- o dicionário resultante da presente proposta terá, por um lado, dados reais recolhidos em *corpora* e, paralelamente, a participação de especialistas como consultores para que os dados apresentados sejam fiáveis para o utilizador.

Por fim, a proposta que se apresenta assenta sobre preocupações lexicográficas e não empresariais, e preocupa-se não só com os aspetos relacionados com a macro e microestrutura, mas também com a disponibilização dos conteúdos lexicográficos num formato digital.

Ao conhecer-se o domínio da ASE (que se apresenta na secção seguinte) compreender-se-á que é:

- um o domínio do saber emergente e em expansão, sobre o qual um conjunto de utilizadores tem interesse em compreender;
- um domínio com um conjunto de informação passível de ser interpretado como conhecimento científico especializado que importa transferir entre especialistas e de especialista para semiespecialistas;
- um domínio com ampla difusão de eventos relacionados com o vinho, e a ASE em particular, sem obras de especialidade que satisfaçam as necessidades de compreensão e conhecimento existentes.

## 2.2. DOMÍNIO DA ASE

A indústria e as atividades implicadas no processo da uva à garrafa incluem diferentes áreas de especialização, o que resulta numa natural diferenciação da língua, discurso e cultura usados e associados à sua descrição, comunicação e avaliação, como é referido por Bruce (1999):

“What makes this particular field interesting is the sheer variety of the discourses of wine that have emerged, and the complex social hierarchies, systems of

classifications, and compendious descriptive vocabulary that have evolved in those discourse communities concerned with the production, marketing, chemical analysis, reviewing and general appreciation of wine.” (Bruce, 1999:150)

Conscientes do impacto económico-financeiro do incremento de conhecimento técnico-científico deste domínio e da importância da comunicação entre os vários intervenientes, várias foram as entidades e os indivíduos que apostaram no apoio e no desenvolvimento de estudos sobre o vinho, não só para facilitar a comunicação e colmatar necessidades de mercado, como também (ainda que de forma inconsciente) para homogeneizar os descritores e sistemas de classificação, mesmo para um público de especialistas<sup>20</sup>. Esta é uma preocupação antiga, dado que “for all winetasters — scientists, professionals, connoisseurs, buffs and yes, most importantly, for everyday wine consumers — there is need to simplify and universalize wine language” (Sharp, 1998:177).

### 2.2.1. MULTIDISCIPLINARIEDADE DO DOMÍNIO ASE

Falar de vinho é falar de um produto conhecido pela sociedade, não apenas porque pode estar presente na casa de qualquer um, por fatores económicos e culturais, mas também pelo peso histórico que tem na sociedade portuguesa. O vinho é também um produto ligado a um conjunto de atividades económico-financeiras distintas, tais como enologia, viticultura, vinicultura, comércio, turismo, gastronomia ou serviços de bem-estar, mostrando a razão pela qual o vinho é uma indústria milionária, nacional e internacionalmente falado.

Demonstrando a sua multidisciplinaridade e também a complexidade de trabalhar um produto tão multifacetado e com múltiplas aplicações em contextos tão diversos,

---

<sup>20</sup> Por exemplo, os projetos OENOLEX Burgundy, VinoLingua Project e South African Trilingual Wine Industry Dictionary, ou o Dicionario Bilingue, *contrastive, basado en areas de conocimiento sobre la cata del vino castellano-manchego*.

depreende-se, intuitivamente, que o vinho é um produto complexo, com a característica de permitir que haja utilizadores que fazem dele o seu objeto de trabalho/estudo e outros o seu alvo de consumo.

A atividade relacionada com o vinho ativa não só um saber (enológico) do domínio vitivinícola, mas também competências de áreas periféricas com contornos difusos, como, por exemplo, as áreas da comunicação, *marketing*, formação ou as áreas de atividades etnográficas e culturais<sup>21</sup>. Dado que a prova de vinho faz uso de uma ritualização do processo já estabelecida, esta permite a ativação de um discurso estereotipado, na medida em que a interação de quem participa na prova (especialistas, técnicos ou consumidores) é limitada, pois o que se verifica não é um diálogo entre quem prova e quem ouve e responde, mas sim uma verbalização entre o agente/experienciador e o paciente/beneficiário, na qual alguém prova e verbaliza na forma oral/escrita e alguém ouve/lê.

### 2.2.2. INTERVENIENTES, NÍVEL DE ESPECIALIZAÇÃO E PARTILHA DE CONHECIMENTO

Em condições normais, qualquer pessoa pode provar um produto (fruta, vegetal, queijo ou vinho) e descrever o que sentiu ao prová-lo, mas, na maioria das vezes, a descrição é algo como “é bom”, “eu gosto” ou “é saboroso”, fazendo com que este ato seja uma avaliação puramente hedónica, ou seja, relativo à expressão do prazer que provoca. Já uma descrição como “tem cor de palha”, “aroma fresco” ou “sabe a frutos vermelhos” é uma avaliação pertencente à área da ASE. Nos dois atos – hedónico e descritivo –, a tarefa de descrição é difícil, na medida em que a influência dos diferentes estímulos sobre os consumidores pode assumir diferentes importâncias, consoante a sensibilidade da pessoa (Azodanlou, 2001:10) e a sua competência de verbalização<sup>22</sup>.

---

<sup>21</sup> “Al hablar de un vino en concreto no solo se hace referencia a sus características organolépticas, su color, su aroma, su sabor, etc., sino que también se hace alusión a todos aquellos rasgos que determinan su identidad, rasgos que van inexorablemente unidos a la tradición, la cultura y la idiosincrasia del lugar de origen” (Feu & de la Cuadra, 2003: 53), cf. também o trabalho de López Zurita (2016).

<sup>22</sup> Assim como nem todos sensores se desenvolvem no organismo humano da mesma forma, também o modo como se aprende a descrever as sensações é igual. A neurofisiologia funciona de forma distinta e

Como tal, não é invulgar constatar que nem todas as pessoas conseguem verbalizar de forma clara o que degustaram, havendo algumas que o fazem com mais facilidade do que outras, que estão mais bem treinadas ou que têm um maior conhecimento do léxico sensorial e hedónico partilhado com quem ouve.

Também se pode colocar a questão da hipótese de Sapir (1949) e Whorf (1956), ou seja, se a forma como se verbaliza o vinho é afetada pelo modo como se percebe e se interpreta a experiência sensorial. Isto é, se alguém provar um produto, mas não tiver ao seu dispor/alcance um léxico disponível para verbalizar o que sentiu, essa tarefa ficará (ou poderá ficar) afetada. Contudo, se essa pessoa tiver acesso a um léxico sensorial e o utilizar convenientemente para descrever o que sentiu na degustação, a compreensão desse ato será mais fácil por parte de quem também ouve/lê.

Focando a atenção nos intervenientes, e existindo uma escala, esta varia entre o não-especialista (leigo/consumidor) e o especialista.

Como o próprio nome indica, o não-especialista é alguém sem um conhecimento profissional ou especializado de um determinado domínio. Por oposição, o especialista é alguém que reúne conhecimentos e competências numa dada área do saber (Honoré-Chedozeau, 2017), tipicamente por formação específica ou experiência prática acumulada.

A compreensão de uma área de especialidade pode ser vista como a articulação entre o “savoir-quoi” e o “savoir-comment”, que culminam, como refere Honoré-Chedozeau (2017:17): “Quelle que soit la discipline, les notions de “connaissance” et de “compétences” ressortent dans la définition de l’expertise”.

---

ativa áreas cerebrais e emocionais distintas. Há autores que defendem que as modalidades sensoriais existem em hierarquia sensoriais (Ullmann, 1959), outros que defendem a existência de hierarquias lexicais consoante os aspetos culturais dominantes nessa comunidade (San Roque et al., 2015) e outros que contestam a existência de hierarquia (Burenhult & Majid, 2011). Já Winter sugere “another way in which across-the-board generalisations such as “modality X is more dominant than modality Y” need to be qualified: The dominance of a sensory modality can only be defined with respect to a certain domain, and not necessarily across the board.” (Winter, 2016:10).

Vários autores afirmam, existir um conjunto de aspetos que é necessário considerar para designar alguém como especialista de um domínio, tais como: conciliar *performance*, competências adquiridas e conhecimentos (Anders Ericsson & Towne, 2010; Gobet, 2015); ter a capacidade para desempenhar tarefas com sucesso, associadas ao produto (Alba & Hutchinson, 1987:411); apresentar, qualitativamente, um nível elevado de conhecimento e de perícia (Jacoby, Troutman, Kuss, & Mazursky, 1986:469); ter facilidade em executar tarefas, devido ao abundante conhecimento adquirido (Scardamalia & Bereiter, 1991:172); manifestar habilidades (*skills*) e possuir conhecimentos resultantes da acumulação de informação ao longo do tempo (Chi, 2006:167).

Na prática da área em causa, o especialista em vinhos tipicamente divulga o vinho numa rotina profissional diária para vender, para o harmonizar com a gastronomia, para melhorar as características da vinha, etc.<sup>23</sup>, porque possui informações e conhecimentos detalhados sobre o produto e o objetivo que visa atingir. No entanto, no lado diametralmente oposto, o não-especialista pode ser um consumidor, um leigo no domínio ou um profissional de uma área correlacionada, mas não necessariamente afim (ex.: hotelaria, venda a retalho).

Frøst e Noble (2002) referem que os “verdadeiros” especialistas do vinho devem ainda possuir ou adquirir duas competências: uma baseada no conhecimento do vinho e outra na capacidade de descrição sensorial dos vinhos. Estas duas competências podem funcionar de forma independente e o facto de dominar melhor uma competência não significa que domine ambas, nem que uma competência possa “compensar” a outra. De acordo com Hughson e Boakes (2002:463), “Wine expertise refers here to a superior ability to discriminate between, recognize and describe different wines”, aspeto confirmado pelo trabalho de Croijmans e Majid (2016) sobre especialistas de vinho, cujos resultados apontam para a ideia de que os especialistas do vinho conseguem detetar

---

<sup>23</sup> Há registos do século VIII a.C. em Atenas, Esparta, Argos e Tebas que referem que os especialistas do vinho são responsáveis por determinar a qualidade e o preço do vinho nessas regiões, como menciona Unwin (1996:81).

melhor os odores do vinho e descrever melhor os vinhos do que os não-especialistas do vinho.

Ainda no mesmo estudo, mas entre especialistas do vinho e do café, verifica-se que o especialista do vinho deteta e descreve melhor os aromas e sabores do que os especialistas do café, perante os mesmos estímulos. Este estudo evidencia que, mesmo perante especialistas de diferentes bebidas que usam a perceção e a verbalização como tarefa principal para descrever o produto, os especialistas do vinho apresentam uma melhor memória (Croijmans & Majid, 2016:145). Tal facto, no entanto, não explica totalmente o resultado, caso contrário, também os especialistas do café teriam melhores resultados. Verifica-se que os especialistas parecem efetivamente ter uma vantagem na capacidade de verbalização quando comparados com os não-especialistas e que os especialistas do vinho possuem um vocabulário mais específico do que os especialistas do café. Provavelmente, como apontam Croijmans e Majid (2016), pela própria experiência e pelo facto, de alguns aromas e sabores serem melhor codificáveis, após treino, mas não necessariamente fruto de uma melhor memória. Caso fosse isso, os especialistas do café teriam desempenho semelhante ao dos especialistas do vinho, o que não se verifica.

Honoré-Chedozeau (2017) refere ainda que:

- o conhecimento sobre os vinhos parece estar organizado hierarquicamente para os especialistas (independentemente do nível de familiarização com o tema);
- o nível de especialização afeta mais as representações conceptuais do que as representações percetivas de vinhos;
- o nível de especialização afeta mais as representações de vinhos para níveis cada vez mais específicos de organização para representações conceptuais.

Os resultados demonstram que a aquisição de conhecimento no domínio do vinho parece seguir um *continuum* exponencial. Ou seja, o *continuum* começa no não-especialista iniciante (que possui apenas uma representação fragmentada do domínio) passando pelo não-especialista familiarizado (com exposição e uma aquisição ativa de

conhecimento), terminando no especialista (com formação técnica e experiência demonstrando maior competência).

Os factos referidos acima seguem o trabalho já realizado por Honoré-Chedozeau *et al.* (2015). Este compara as descrições de especialistas e de dois grupos de não-especialistas (uns mais familiarizados e outros nada familiarizados com o tema do vinho) de vinhos de Borgonha e mostra que, para os especialistas, a representação do conhecimento é flexível e leva diferentes esquemas de categorização (“categorization schemes”), ao passo que os consumidores familiarizados gerem a informação num sistema mais rígido de categorização e os não familiarizados usam outras estratégias para compensar o conhecimento fragmentado que possuem.

Outros autores dedicaram-se ao estudo do uso dos descritores do vinho quer por especialistas quer por não-especialistas, obtendo dados interessantes: perante um mesmo vinho, os especialistas tendem a usar descritores mais específicos e concretos do que os não-especialistas (Lawless, 1984; Solomon, 1990); os especialistas são mais propensos a adicionar informação relativa à casta do que os não-especialistas (Parr, Mouret, Blackmore, Pelquest-Hunt & Urdapilleta, 2011); os especialistas, ao descreverem o mesmo produto e tendo uma escala de descritores mais gerais e mais específicos, tendem a utilizar descritores mais específicos/técnicos; e, por fim, é mais fácil saber-se de que vinho se trata a partir das descrições refletidas nas notas de prova dos especialistas (Valentin, Chollet & Abdi, 2003).

Existem ainda diferentes categorias de especialistas (Lino, 2013) neste domínio, começando no aprendiz e terminando no especialista, categorização atribuída através de treino e tempo dedicado a conhecer e a entender o domínio através de leituras e de experiências associadas à prova de vinho. Estas experiências, por um lado, consolidam a informação lida e, por outro lado, melhoram aspetos como a memória (Humphrey & Underwood, 2011) e a *performance* linguística relativamente à capacidade de compararem mentalmente um determinado vinho com a informação mental armazenada<sup>24</sup>

---

<sup>24</sup> “Wine experts’ ability to describe a wine might in part hinge on their ability to mentally compare a given wine with prototypical wines in terms of color, smell and flavor.” (Croijmans, 2018:99)

(Croijmans, 2018), associada à verbalização das características do vinho (Herdenstam, Hammaren, Ahlstrom, & Wiktorsson, 2009).

Deste modo, pode falar-se de uma situação (a prova) onde estão presentes intervenientes especialistas e intervenientes não-especialistas, com diferentes competências e diversas formas de comunicar (oralmente, por escrito e dentro de cada uma destas tipologias distintas).

### 2.2.3. RECURSOS E CONSTRUÇÃO DO CONHECIMENTO DO DOMÍNIO

O especialista na área da ASE começa a sua formação apoiando-se num conjunto de recursos que o ajudam a organizar o pensamento e a descrever as notas de prova. Estes recursos possuem diferentes formatos, mas incidem essencialmente sobre os descritores que se utilizam. São então:

- listagens de descritores (Lehrer, 2009; Lenoir, 2011; Parker, 2017; Dalmasso, 1963; Sánchez, 2011);
- rodas/esquemas circulares com descritores sobre cores, aromas, sabores (Noble *et al.*, 1984);
- tabelas de descritores (Carvalheira, 2012; Afonso, 2013; Almeida, 2016: etc.);
- dicionários, vocabulários, glossários, léxicos (Pintão & Cabral, 2014; Salvador, 2010; etc.).

Todos estes recursos foram elaborados por especialistas, para colmatar lacunas, de formação ou aplicação prática, que encontraram e que ajudam a compreender o léxico das notas de prova (Solomon, 1990), mas, como se verifica, são recursos variados e de diferentes tipos, além de muitas vezes terem sido criados para diferentes públicos e com diferentes propósitos.

Tendo em conta esta disparidade, vários foram os autores que demonstraram a sua preocupação com a profusão de recursos e de descritores, bem como as apresentaram as vantagens da existência de um repositório organizado, com pendor normalizador (ou pelo menos com uma tendência harmonizadora bem definida), que possa evitar desentendimentos e falhas de comunicação entre os intervenientes sobre uma série de descritores, bem como acerca dos seus usos e das suas definições. Para o inglês, são conhecidos os trabalhos de Lawless e Civile (2013), por exemplo, que defendem que um recurso com estas características facilitaria a comunicação entre especialistas de diferentes empresas e painéis de provadores treinados.

Esta preocupação com a criação e utilização de recursos existe quer ao nível monolíngue quer ao nível multilíngue. Basta recordar o trabalho realizado por Dalmasso (1963) para a construção e organização de um recurso de referência internacional, utilizado pela *Organisation Internationale de la Vigne et du Vin* (OIV), que, na edição de 1963, possuía uma versão em francês, italiano, castelhano, alemão, português, inglês e russo, ou a edição de Escarpit (1995) que se diz uma versão atualizada de Dalmasso (1963), mas de carácter privado, sem a chancela da OIV e com menos línguas apresentadas. No entanto, por vários motivos, nem o Instituto da Vinha e do Vinho, nem a Associação Portuguesa de Enologia continuaram a parceria internacional e interdisciplinar com Dalmasso (1963) ou trabalhos posteriores, no sentido de atualizar a versão portuguesa, facto que se explica pela ausência do PE na última versão<sup>25</sup>.

Havendo esta lacuna em PE, uma vez que não há organismos oficiais que validem os produtos criados/apresentados para posterior utilização e divulgação, parece pertinente o questionamento sobre que tipo de recursos existem ao dispor do especialista e do não-especialista português, e como contribuir para o desenvolvimento e para a produção de um recurso fiável, que garanta a qualidade e facilite às entidades responsáveis o início de um possível processo de harmonização.

---

<sup>25</sup> Apesar de existir outro recurso multilíngue de Glatre (2014) com versão em PE, não é mencionado em nenhuma parte qual a base de realização desse trabalho, ficando a sensação de que se trata de uma versão semiatualizada e adaptada mais para o francês e o inglês do que para as restantes línguas mencionadas (além de não fornecer informação sobre a origem dos dados nem sobre a equipa envolvida).

Neste contexto, observa-se uma multiplicidade de materiais (glossários, léxicos, dicionários, tabelas, listas, escalas) criados por diferentes autores, e tipicamente com objetivos distintos, objeto da subseção 2.3, que surgem na tentativa de colmatar a falha no que respeita à transmissão de informação entre especialistas (e entre especialistas e não-especialistas), mas que parecem gerar algumas dificuldades de compreensão – por exemplo, ao nível das unidades lexicais selecionadas e das suas definições – naqueles que os consultam. No dia a dia, é evidente que o especialista não comunica através de listagens, de tabelas ou de léxicos, mas sim através de outras formas, nomeadamente as notas de prova. Contudo, estes materiais surgem como apoio para o entendimento do que é verbalizado não só entre especialistas, mas também entre especialistas e não-especialistas.

A esta lista de recursos, podem ainda adicionar-se os diversos projetos internacionais que surgiram nos últimos anos e que se dedicaram ao estudo não só do léxico como também do discurso do vinho, como, por exemplo, o *VinoLingua* (Lavric, 2013), o *Minéralité des vins*<sup>26</sup> (Deneulin, Le Fur, Le Bras & Gautier, 2014), o *OENOLEX BOURGOGNE* (Gautier & Leroyer, 2014), entre outros. Estes ajudaram a compreender como ocorre a comunicação entre especialistas e como é feita a passagem de informação relativa ao vinho aos especialistas e aos não-especialistas.

#### 2.2.4. NOTAS DE PROVA

Como referido acima, uma das formas de “comunicar o vinho” é através da verbalização oral<sup>27</sup> das suas características, como é referido nos trabalhos de Gautier e Hohota (2014). No entanto, para uma maior difusão desta informação, são utilizadas

---

<sup>26</sup> O Projeto “Minéralité des vins: étude et valorisation d’un concept en vogue mais méconnu” foi selecionado para financiamento através do INTERREG IV A France-Suisse 2007-2013.

<sup>27</sup> Material de acesso mais difícil, pois as provas não costumam ser filmadas/gravadas. No entanto, a partir de março de 2020 (pandemia do SARS-CoV-2) assistiu-se a uma maior divulgação através de vídeos de especialistas. O tratamento destes dados implica a recolha de material distintos, e respetiva salvaguarda de direitos autorais, bem como a tarefa de transcrição do oral, exercício de interpretação moroso e complexo e que este trabalho não contempla.

estratégias escritas<sup>28</sup>, como as fichas técnicas e, em particular, as notas de prova. Estas consistem em pequenos textos com uma apreciação concisa da experiência sensorial de degustar o vinho, com particular destaque para a cor, o aroma e o sabor, para disseminação na imprensa generalista, em revistas da especialidade, nas páginas das empresas e cooperativas do setor enológico.

O texto<sup>29</sup> da nota de prova permite, através da leitura, provocar no outro uma sensação imaginária da experiência sensorial sem a presença de um estímulo físico (Freeman, 1981; Pylyshyn, 1973) e possibilita uma reconstrução sensorial interna e a antecipação de outra (Arshamian & Larsson, 2014). Por outro lado, permite veicular informação relativa à qualidade do vinho e tabelar preços para a venda do produto, por outro, permite que o consumidor possa escolher de acordo com as suas preferências sensoriais e com a relação qualidade-preço.

A nota de prova possibilita, assim, avaliar e incitar a um comportamento de compra/prova por parte do consumidor (Cardebat & Livat, 2016; Oczkowski & Doucouliagos, 2014), fazendo com que as dúvidas na escolha do produto sejam menores (Gokcekus & Nottebaum (2011)). Aliás, atualmente, o consumidor procura cada vez mais

---

<sup>28</sup> Sabe-se que existe uma competição entre empresas/companhias/cooperativas que pretendem dar a conhecer o seu produto através de campanhas publicitárias apelativas que utilizam diversas técnicas de comunicação verbal, condicionadas e/ou potenciadas pelo facto de, por lei, em nenhuma imagem (estática ou em vídeo) se poder ver alguém a consumir vinho, como por exemplo: (i) a associação de determinadas palavras (tais como “tradição”) e expressões populares combinadas com imagens (de prémios, da natureza, da herdade), apelando ao seguimento de um estilo de vida; (ii) o apelo equilibrado, por um lado, à comercialização, venda e compra do vinho, por outro lado, ao seu consumo com moderação. Apesar de constituírem uma temática muito interessante, nenhum destes aspetos será abordado neste trabalho.

<sup>29</sup> Vários autores possuem diferentes noções do que entendem por género, texto e discurso consoante o quadro teórico que seguem. Neste trabalho, entende-se a nota de prova como um género de texto, na aceção de Bronckart (1997/2003), segundo a qual existe um género que funciona como um modelo a partir do qual diversas manifestações textuais se baseiam e incluindo-se nele. Já Breit (2013) menciona “that the wine tasting sheet may be considered “genre”, as it meets the definitional requirements determined by Swales (1990) and the adequacy of the Appraisal Theory (Martin & White, 2005) for the analysis of this genre within the framework of Systemic Functional Linguistics” e Teixeira (2014:183), “um género de texto associado (...) que corresponde o mencionado registo das propriedades do vinho e respetiva classificação (...) género de texto recensão crítica enológica bem conseguido, segundo a atividade de produção e de comercialização do vinho”.

informação para se sentir esclarecido<sup>30</sup>, e, nesse sentido, a indústria do vinho tem tentado educar o seu público através de provas de vinho, *workshops*, feiras e outros eventos abertos ao grande público e potencial comprador/consumidor.

Além disso, também é mencionado em alguns trabalhos que, de acordo com o léxico utilizado nas notas de prova, existe uma expectativa associada ao preço do produto (Nerlove, 1995). Ou seja, existe uma relação entre o tipo de descritores usados, que os não-especialistas associam a uma experiência sensorial (Smith, Møgelvang-Hansen & Hyldigc, 2010), à qualidade (Charters & Pettigrew, 2006), à marca (Priilaid, Human, Pitcher, Smith & Varkel, 2017) e, por sua vez, a um determinado valor monetário (Lecocq & Visser, 2006) do vinho, que se pode traduzir na aquisição de um bem de luxo (Sjostrom, Corsi & Lockshin, 2014) e que o tecido empresarial tem sabido manejar a seu favor para aumentar a intenção de compra dos seus produtos.

É impossível ignorar a importância das notas de prova e do respetivo léxico no que concerne ao envolvimento dos não-especialistas. A este respeito, sabe-se que a maioria dos não-especialistas sente alguma ansiedade por não conhecer o suficiente sobre vinho, quando confrontada com a necessidade de falar sobre o assunto<sup>31</sup> ou de escolher um vinho para oferecer. Este facto pode ainda relacionar-se com o sentimento de inadequação em utilizar e/ou compreender o léxico específico ou geral para descrever uma experiência sensorial de forma sistemática, mesmo com recurso a metáforas, personificações, antropomorfizações ou analogias. Autores como Lesschaeve (2006) relatam a importância de melhorar a comunicação entre estes intervenientes, no sentido de beneficiar ambas as partes, por considerarem que o léxico utilizado pelos especialistas só é eficaz, neste contexto, se este for entendido pelos não-especialistas.

No entanto, apesar de a nota de prova possuir uma estrutura semirrígida, mais ou menos estável e aceite, é importante ter presente que cada autor tem um estilo próprio (Brochet & Dubourdieu, 2001), sendo que alguns utilizam descritores considerados mais

---

<sup>30</sup> “Consumers are nowadays eager to learn about the background and origin of the products they purchase, and seek for authenticity in brands” (Beverland, Lindgreen & Vink, 2008).

<sup>31</sup> <https://www.wsj.com/articles/wine-anxiety-how-do-you-know-a-bottle-is-ok-1532628703>.

hedônicos (ou menos objetivos) e outros mais técnicos (ou mais objetivos), como frisam Parr *et al.* (2011) e Sauvageot, Urdapilleta e Peyron (2006). Apesar disso, considera-se mais ou menos objetivo o descritor utilizado. Trabalhos como o de Hendrickx *et al.* (2016) indicam que o não-especialista os considera sempre informativos.

Contudo, há quem sustente que as notas de prova são pouco esclarecedoras ou adequadas ou até mesmo inúteis, para informar o não-especialista acerca do vinho. Alguns autores acreditam que se trata de uma forma inadequada, visto dependerem da língua para verbalizar todo um processo cognitivo interno e pessoal, que nem sempre é bem entendido por quem lê as descrições, como refere Gluck (2003:107): “We wine writers are the worst qualified of critical experts. This is largely, though not exclusively, because we are the most poorly equipped. The most important tool at our disposal is inadequate for the job. That tool is the English language.” Outros autores consideram que o especialista escreve de forma muito particular, o que, nalguns casos, se comprova pela utilização de determinados descritores usados somente por um especialista (Lawless, 1984:122).

Outros autores sustentam ainda que a forma como determinados especialistas escrevem sobre o vinho espelha mais aquilo que o especialista sabe sobre o vinho do que as características de um vinho específico propriamente ditas (Silverstein, 2006). Por fim, há quem argumente que as notas de prova estão pejudicadas de metáforas e de palavras “pretensiosas”, “sensuais” e “vigorosas” (Erickson, 2008), concluindo que as notas de prova são de uma profunda inutilidade e que “the wine trade is intrinsically bullshit-prone and therefore attracts bullshit artists”, como refere Quandt (2007:135).

Independentemente da opinião que se possa ter sobre as notas de prova, estas são uma tipologia de texto válida para extrair um conjunto de dados relativos ao uso em contexto real, visto que são redigidas por especialistas com o propósito de descrever um vinho através de um conjunto limitado de frases e descritores, além de serem uma tipologia de texto relativamente acessível para compilar um *corpus* de tamanho razoável e, a partir daí, iniciar uma análise que permita uma proposta de dicionário.

### 2.3. RECURSOS LEXICAIS NA ÁREA DA ENOLOGIA E ASE

No contexto dos recursos lexicais na área da enologia e ASE, e numa perspetiva mais tradicional, há quem fale em discurso científico, discurso profissional dos especialistas e discurso dos consumidores, sendo que este último reduz a dimensão sensorial do vinho a apenas um conjunto de palavras utilizadas durante a degustação. O domínio do vinho pode ser considerado um referencial lexicográfico especializado (Leroyer & Gouws, 2009:100), que implica, simultaneamente, um conhecimento de diversos aspetos técnicos (degustação, conservação, consumo, serviço, etc.) e dos traços da conceção do mundo vinícola, como uma “terminologia cultural” ou uma “etnoterminologia” (Depecker, 2002).

Deste modo, verifica-se que existem diversas formas de abordar as palavras utilizadas no ato da degustação do vinho, a prova de vinho. Como lembra Tsai (2014):

“La terminologie de la dégustation peut être classée selon les registres de perception sensorielle (visuelle, olfactive, gustative et de l’ensemble), ou selon les divers constituants et aspects du vin (sa teneur en extrait sec/acidité/sucre/tanins/alcool, son intensité, sa persistance aromatique, son évolution des qualités gustatives et son équilibre entre les divers éléments constitutifs).” (Tsai, 2014:120).

Ainda a respeito desta grande divisão do léxico da degustação e vinhos, há estudos que referem que descrever aromas e sabores parece ser uma tarefa mais difícil para os povos/línguas faladas no Ocidente (Europa e América do Norte).

Ainda assim, os profissionais habituados a mencionar descrições de café e vinho parecem estar bem preparados, sendo que, simultaneamente, parece haver um léxico (geral) mais restrito para descrever aromas e sabores (Myers, 1904; Sperber, 1975, Croijmans, 2018). Muitas vezes, os falantes apoiam-se em estruturas como “cheira a canela”, “sabe a peixe” ou em metáforas como “cheira a verde”, para descreverem a sua perceção através de um discurso, aparentemente, mais compreensível por outros. Outros estudos indicam que a dificuldade em expressar aromas e sabores ocorre nesta cultura, porque se vive no mundo WEIRD (Western, Educated, Industrialized, Rich, Democratic) (Henrich, Heine & Norenzayan, 2010), isto é, numa parte do mundo em que a importância do aroma e do sabor é enorme, mas apenas para aspetos industriais, como, por exemplo,

para os perfumes, e não tanto em outras atividades quotidianas, onde não é usada. Os especialistas do vinho, os *flavorist* e os perfumistas (Sezille *et al.*, 2014), por seu lado, devido à sua atividade profissional, estão habituados a descrever aromas e sabores (Herdenstam *et al.*, 2009).

No contexto internacional, vários autores desenvolveram produtos relativos ao vinho e existe um conjunto de recursos que visam a transmissão de conhecimento do domínio do vinho, como os trabalhos de Coutier (2007), Lehrer (2009), Gautier (2017), Vedovelli (2016), Machetti (2010), Reckinger (2012), Stengel e Marinescu (2016), Ibanez Rodríguez e Sánchez Nieto (2006), Sacha González (2006), para citar alguns autores que se dedica(ra)m ao domínio do vinho em diversas línguas e com diferentes abordagens.

No entanto, em PE, existe um conjunto limitado de recursos relativos ao vinho e à ASE conforme apresentado na secção 2.2.3 e se apresentará com maior detalhe na secção 4.2.2.1 aquando da apresentação das obras lexicográficas recolhidas para a criação do *corpus*. Entende-se que, tendo em conta o panorama nacional relativo à lexicografia de especialidade e, em concreto, ao dicionário de especialidade, existe uma lacuna no que concerne ao dicionário de especialidade aplicado à ASE.

A lacuna existente poderá ser colmatada se forem tidos em conta aspetos como os princípios da lexicografia de especialidade atuais e adequados ao domínio e ao público em questão. Além disso, o modelo proposto deverá ter em conta os desenvolvimentos da Linguística de *Corpus*, permitindo uma seleção e recolha de *corpora* reais com características específicas para extração de candidatos a unidades lexicais especializadas. Estes candidatos deverão ser apresentados a especialistas permitindo verificar, por um lado, o estado atual dos descritores do domínio e, por outro, a sua instabilidade e variação, promovendo um debate entre os especialistas, que possa ir no sentido de uma harmonização do léxico do domínio e em último lugar, da sua normalização.

### 3. ANÁLISE SENSORIAL ENOLÓGICA

A Análise Sensorial (AS), que consiste na utilização dos diversos sentidos (visão, audição, olfato, paladar e tato), é uma função primária do ser humano, auxiliando, ainda que em muitos casos de forma inconsciente, nas escolhas alimentares desde os primórdios. E os seres humanos, em comunidades distintas, fazem uso dos seus sentidos para experienciar a realidade e construir social, histórica, cultural e individualmente o que percebem e como compreendem o mundo e os outros (Breton, 2016).

Sempre existiu uma preocupação com a qualidade dos alimentos e das bebidas. Num primeiro momento, por questões de sobrevivência (comer uma baga de cor errada poderia significar ingerir um alimento tóxico que podia resultar na morte de membros de uma tribo); posteriormente, por motivações de poder (ter acesso a alimentos de melhor qualidade indicava maior prosperidade e saúde; ter conhecimento acerca de alimentos específicos impedia o envenenamento<sup>32</sup>) e financeiras (vender produtos que agradam aos consumidores para aumentar vendas e lucros).

Ao longo dos séculos, a civilização desenvolveu-se e com ela o consumo, a venda e/ou troca de bens consumíveis, *i.e.* o comércio. No decorrer destas trocas comerciais tornou-se frequente o teste e a avaliação de uma amostra do produto e a respetiva definição do preço em função do resultado da prova (Drake, Drake, Bodyfelt, Clark, & Costello, 2008:1). Surgem assim profissões para controlar e avaliar produtos, como, por exemplo, os mestres-cervejeiros para a cerveja, os escanções para o vinho ou os juizes para o gado bovino leiteiro. Contudo, a partir de 1900 e devido à industrialização, os métodos foram sendo alterados consoante a época, como refere Freitas (2008).

Apesar de as referências indicarem que o termo AS foi utilizado pela primeira vez em 1961 (“il semble avoir été utilisé pour la première fois en 1961 lors d’une conférence donnée au centre technique de la biscuiterie” (Depledge & Sauvageot (2002) *apud* Thomas

---

<sup>32</sup> Casos de figuras históricas envenenadas através de bebidas ou comida são conhecidos desde Sócrates, o filósofo (Dayan, 2009) ou Cleópatra (Brown, 2013), passando pelo período vitoriano, por exemplo, onde várias pessoas foram acusadas de envenenar os seus parentes (Whorton, 2010), até ao caso do General croata Praljak, já em pleno século XXI (Bowcott, 2017).

(2016:24)), é evidente que anteriormente a esta data já existiam preocupações relativas à qualidade dos produtos, como se descreve na Tabela 3.

<b>Período</b>	<b>Descrição</b>
Período anterior a 1940	Época artesanal ou pré-científica da indústria de alimentos, período onde a qualidade sensorial era determinada pelo “dono” ou encarregado da indústria.
Período de 1940 a 1950	Incorporação de pessoal técnico geralmente vindo de outras áreas, como a área da química e da farmacêutica. Houve a introdução dos conceitos de controlo de processo e produto final. No entanto, os métodos padronizados ainda eram químicos e instrumentais.
Período de 1950 a 1970	Definição dos atributos primários que integram a qualidade sensorial dos alimentos e os órgãos sensoriais com eles relacionados. Normatização da forma e das condições da realização da AS, bem como o tratamento estatístico dos dados obtidos.
Período após 1970	Instituição da qualidade sensorial de um alimento como sendo o resultado da interação entre o alimento e o homem, que varia de pessoa para pessoa. Reconhecimento de que a qualidade sensorial é função tanto dos estímulos procedentes dos alimentos como também das condições fisiológicas, psicológicas e sociológicas dos indivíduos que o avaliam.

Tabela 3. Períodos do século XX e respetiva descrição (adaptado de Freitas, 2008)

Vários estudos apontam que a AS enquanto disciplina terá surgido na década de 1940:

- na Escandinávia, em trabalhos relacionados com a cerveja, por Bengtsson e Helm (1946), na Carlsberg Brewery;
- nos Estados Unidos da América, durante a Segunda Guerra Mundial, devido à preocupação de preservar ração militar dos soldados do exército por Peryam (1958) para o Quartermaster Food and Container Institute e o Exército dos Estados Unidos;

- no Brasil, em 1954, para a avaliação de café pelo Instituto Agrônomo de Campinas (Esteves, 2014:1).

Relativamente a Portugal, não foi possível encontrar um registo exato sobre os primeiros trabalhos da AS aplicados.

Atualmente, a AS é definida como a “(...) disciplina da Ciência usada para evocar, medir, analisar e interpretar as reações às características dos alimentos e materiais tal como são percebidas pelos sentidos da visão, olfato, paladar, tato e audição” (IFT, 1975:50), que utiliza os sentidos humanos para avaliar as características ou os atributos de um artigo e que é parte integrante do processo de controlo de qualidade dos produtos, tendo sido mais recentemente definida como a “disciplina que aplica princípios do planeamento experimental e da análise estatística para analisar os resultados da utilização dos sentidos” (Esteves, 2006:1).

A AS é também considerada uma ferramenta intensamente utilizada pela indústria com diferentes aplicações:

- na criação de novos produtos;
- no aperfeiçoamento de produtos já existentes;
- em testes de tempo de vida útil;
- em testes do produto relativamente a concorrentes;
- em controlos de qualidade da matéria-prima, da fabricação e do produto final;
- em testes de mercado para compreender as preferências do consumidor na aprovação do produto.

Nesse sentido, foram desenvolvidos vários procedimentos permitindo, por um lado, o desenvolvimento de ferramentas e protocolos para a avaliação sistemática de um

produto e, por outro lado, o estabelecimento de padrões de qualidade (Lawless & Heymann, 2010).

Se, por um lado, nenhuma ferramenta ou combinação de ferramentas pode substituir totalmente os sentidos humanos que permitem descrever a complexidade da percepção de um determinado produto, por outro lado, a AS pode funcionar como uma ferramenta de “princípios do planeamento experimental e da análise estatística para analisar os resultados da utilização dos sentidos” (Esteves, 2006:1), que permite aos especialistas avaliar e descrever as características e a qualidade de vários produtos (como, por exemplo, iogurte, mel, chá, queijo, carne, azeite, vinho), através da percepção humana.

Nos últimos anos, a investigação alimentar foi alvo de alguns desenvolvimentos, no sentido de melhorar determinadas propriedades dos produtos em geral, nomeadamente no que respeita ao aumento da resistência a agentes patogénicos, à preservação durante o transporte e à extensão do prazo de validade, focando em particular a manutenção da aparência do produto, visto esta ser fundamental para apelar à prova e à possível aquisição pelo consumidor. No entanto, houve um momento em que a indústria alimentar compreendeu que a melhoria do aspeto final do produto não era suficiente por si só, sendo necessário contrabalançar o investimento neste aspeto com investigação relativa às características organoléticas e sensoriais.

### 3.1. SENTIDOS E PERCEÇÃO SENSORIAIS

Para compreender como a AS se processa, é essencial conhecer as características e o funcionamento da percepção humana, bem como a fisiologia e função dos órgãos dos sentidos e identificar quais são utilizados na AS.

A atividade cerebral do ser humano é drasticamente ativada no momento do nascimento e desenvolve-se através de estímulos sensoriais do mundo exterior. A percepção sensorial é a forma de compreender a informação exterior, que permite ao ser humano interagir constantemente com tudo e todos os que o rodeiam, e que se desenvolve ao longo das várias etapas de vida (Marques, 2016).

No entanto, não deixa de ser curioso que “um dos maiores paradoxos da condição humana é que a grande abundância de sensações, que tanto significam para nós, não é percebida directamente pelo cérebro. Este apenas recebe impulsos electrónicos (...). O cérebro é cego, surdo, mudo e insensível” (Ackerman, 1990:313).

Para perceber, ou seja, perceber e compreender a realidade sensível que o rodeia, o ser humano faz uso não só dos sentidos – a visão, o olfato, a audição, o tato e o paladar –, mas também do sistema nervoso central, que recebe e processa a informação sensorial. Isto é, os estímulos sensoriais são inicialmente captados pelos sentidos/órgãos recetores – olhos, ouvidos, nariz, pele e língua –, seguem através das vias nervosas até ao sistema nervoso central e, por fim, chegam ao córtex cerebral (Gimeno, 1986).

Ainda que a informação sensorial externa seja a mesma, esta afeta os hemisférios de forma distinta: o hemisfério esquerdo é tendencialmente analítico, verbal, numérico, sequencial, simbólico, linear e objetivo, enquanto o hemisfério direito não é verbal e global, mas guiado por códigos não verbalizados, como referem Buzan e Buzan (1996). Como tal, a informação sensorial converge para a forma de pensar e de funcionar em harmonia daquilo que experienciou e como foi percebido, potenciando a enorme capacidade associativa do cérebro humano (Ponte, 2011).

O cérebro, enquanto órgão, não é capaz de sentir, reagir e pensar normalmente, se encontrar um vazio sensorial<sup>33</sup>, uma vez que a mente é um conjunto de funções e a riqueza de estímulos e respostas sensoriais beneficia, em geral, a aprendizagem e, em particular, a inteligência (Pereda, 2002). Ou seja, a função cerebral depende da constante existência de estímulos que exerçam uma continuada exercitação sensorial, permitindo conservar e desenvolver a atividade mental.

---

<sup>33</sup> É conhecida como uma técnica de tortura psicológica (ou “white torture”) que inclui a privação sensorial, ou seja, a ausência total de qualquer estímulo sensorial e o isolamento forçado (<https://www.humanrightsfirst.org/senate-report-cia-torture/sensory-deprivationsolitary-confinement>).

### 3.1.1. A VISÃO, O OLFATO, A AUDIÇÃO, O TATO E O PALADAR

O ser humano possui cinco sentidos desenvolvidos – a visão, o olfato, a audição<sup>34</sup>, o tato e o paladar. Alguns desenvolvem-se ainda durante a gestação, como, por exemplo, a audição (vários estudos apontam para a capacidade de ouvir música ainda no ventre materno (Arabin, 2002)); outros imediatamente após o nascimento, como, por exemplo, o paladar (durante a amamentação, onde o leite materno é a primeira bebida dada a provar e que influenciará os julgamentos sensoriais do indivíduo (Ellis (1993), Eckblad (1981)).

Cada um dos diferentes sentidos tem uma localização específica no córtex cerebral, como se observa na Figura 5.

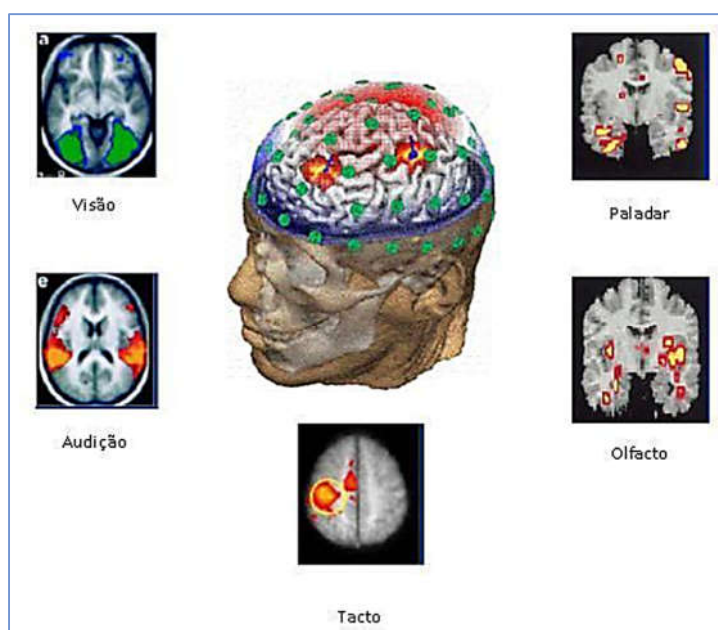


Figura 5. Sistemas sensoriais no Córtex (retirado de Nogueira, 2011:13)

A estes cinco sentidos correspondem diversos recetores (olhos, narinas, língua, ouvidos, mãos e pele), que permitem receber estímulos, tais como, respetivamente, imagens, odores, sabores, sons e sensações táteis, e que enviam a informação para ser organizada e interpretada posteriormente no córtex.

---

<sup>34</sup> Apesar de o sentido da audição poder ser invocado aquando da prova de vinho, importa salientar que este apenas é ativado para examinar a característica da efervescência em vinhos espumantes ou champagnes, pelo que se decidiu não eleger este sentido como relevante.

Cada um dos sentidos desempenha um papel relevante na AS, como sucintamente se descreve de seguida.

O olho, órgão associado à visão, permite obter informação relativa às sensações de cor e das tonalidades, à limpidez, à fluidez e à efervescência, que contribuem conjuntamente para se ter uma ideia da aparência. Através do nariz e do bolbo olfativo em particular, órgão associado ao olfato, captam-se as sensações relativas ao aroma e à fragrância, quer por via nasal quer por via retronasal. O ouvido e o tímpano em particular, órgão associado à audição, possibilitam a aquisição de informação relativa ao som e à melodia. Através da pele, órgão associado ao tato, obtém-se informação relativa à consistência e textura. E, por fim, pela boca, órgão associado ao paladar, captam-se informações relativas ao sabor (Nogueira, 2011).

Com exceção da audição, todos os sentidos são convocados na ASE, como apresentado na secção 2.3 com maior detalhe.

### 3.1.2. SENSACÃO E PERCEÇÃO

Perceber o modo de funcionamento dos sentidos e do cérebro humano é importante para compreender não só o que é uma sensação, mas também o que o ser humano entende como sensação, descrevendo a mesma a si mesmo e aos outros, isto é, fazendo o percurso entre a fisiologia sensorial e a psicologia da percepção.

Diversos autores questionam-se como é que o estímulo é codificado pelos órgãos sensoriais e de que modo o sentido é interpretado pelo cérebro. Partindo da Figura 6, de Nogueira (2011), baseada no trabalho de Depledts (1998), será dada uma breve explicação sobre as etapas psicológicas da percepção, de modo a diferenciar sensação de percepção, distinção de grande relevância para a área em estudo.

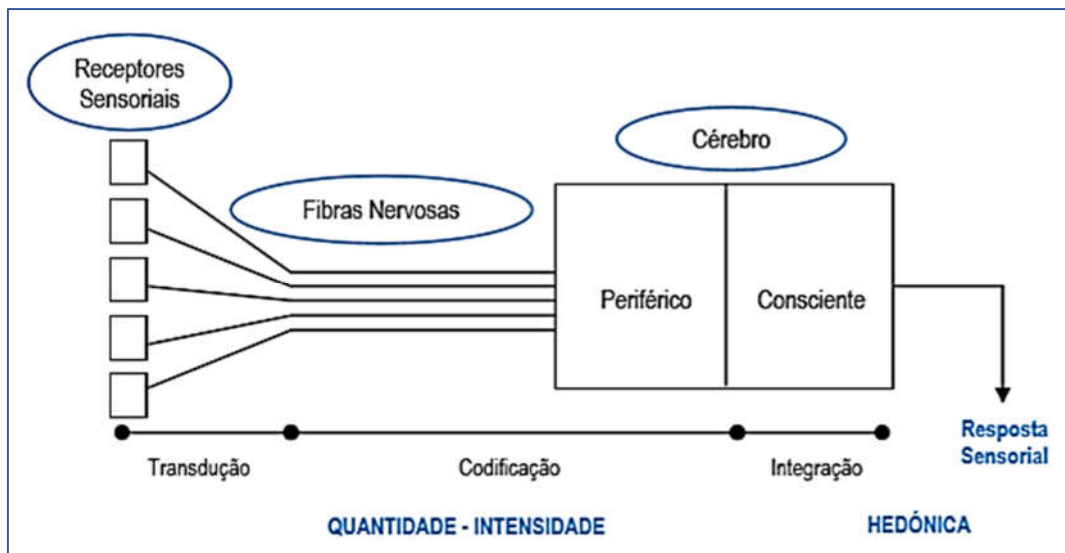


Figura 6. Etapas psicológicas da percepção (retirado de Nogueira, 2011)

Perante a apresentação de um estímulo (seja ele uma imagem, um perfume, um toque numa superfície, um som ou um alimento), o seu reconhecimento – ou seja a indicação física de que este é recebido – é convertido em informação que percorre os recetores sensoriais periféricos, através das fibras nervosas, levando a sensação até ao cérebro. Aí, a sensação é interpretada, isto é, descodificada e integrada na área consciente do cérebro, dando origem à percepção. Esta percepção consciente ocorre em milésimos de segundo (0,1 a 0,2 segundos), após a receção do estímulo e permite evocar um conjunto de informações já existentes no cérebro humano (como o conhecimento não-verbal, a memória, a cultura, a genética, a consciência, entre outras), de modo a criar uma resposta (sensorial) àquilo que experienciou.

Mas Nogueira vai mais longe ao referir que:

“O indivíduo emite espontaneamente o que sente. A descrição de uma percepção é uma operação mais complexa do que o simples reconhecimento. Os homens comunicam entre si por palavras. É possível, por meio desta semântica, fazer compreender a uma pessoa que não se tenha apercebido de um objecto, o que ela experimentaria na percepção deste. A descrição consiste em representar uma forma, por meio de um conjunto de palavras, designado por “descritores” ou “atributos”. Distinguem-se dois níveis de descrições: a descrição simples ou a representação de um objecto por um conjunto de palavras e a descrição quantitativa que é a

representação de um objecto por um conjunto de palavras associadas a números ou a uma escala numérica. Estes números servem para quantificar a importância relativa do descritor ao qual estão associados, de acordo com o objecto descrito. É aprendendo a analisar o que se sente, que se treina a utilizar a palavra adequada e que progressivamente se vem a dar respostas precisas e reproduzíveis.” (Nogueira, 2011:23)

Assim, a verbalização, enquanto parte da resposta sensorial, parece ser uma etapa mais complexa do que aquela inicialmente apresentada na Figura 7. Uma etapa que necessita de um conjunto de unidades lexicais – descritores –, de aprendizagem (de treino explícito) e de ser validada pelo próprio e pelos outros.



Figura 7. Percurso do estímulo apresentado à verbalização do mesmo (adaptado de Barbosa, 2016)

Descrever as sensações que o mundo exterior provoca no ser humano foi objeto de reflexão de diversos autores em várias épocas e em diversos locais, como Yangming, no Oriente, Ki-Zerbo em África, e Platão, Aristóteles, Condillac, e Locke, na Europa, entre outros, que exploraram o que entendiam por sentido, sensação, sensorial, sensível e sensorialidade. Apesar de ser um tema de grande interesse<sup>35</sup>, esta reflexão não será objeto

---

<sup>35</sup> A este propósito ler “La notion de sensorialité, terme dû à Merleau-Ponty, désigne l’ensemble des modalités sensorielles, le sensoriel concernant “la sensation sur le plan psycho-physiologique” (Grand Robert) ou “la sensation considérée sous son aspect objectif ou cognitif” (Dictionnaire Foulquié et Saint-Jean). Paul Valéry (1957, p. 1298) parle du “plaisir d’espèce commune, le fait purement sensoriel”. Sens, sensation et sensibilité obligent à des distinctions conceptuelles selon des filiations disciplinaires croisées que Bernard Valade retrace ici dans un article qui recontextualise également les conditions et conséquences de la prégnance progressive de la vue en Occident.” (...) “La hiérarchisation des cinq sens est un

de análise neste trabalho. Tratar-se-á da verbalização e descrição sensorial apenas no contexto da ASE, como explorada na secção que se segue.

### 3.2. PROCESSO DA ANÁLISE SENSORIAL ENOLÓGICA

O vinho é uma bebida de fabricação complexa e relacionada com condições ambientais diversas que acompanha a Humanidade desde a Antiguidade. Mas falar de vinho é muito mais do que apenas falar de uma bebida alcoólica, uma vez que o vinho é uma indústria de grandes dimensões que movimenta milhares de milhões de euros por ano em todo o mundo e também, em Portugal, país no qual a enologia é uma área em expansão. Os dados estatísticos do Instituto da Vinha e do Vinho, I.P. referem que, em 2021, as exportações de vinho português tiveram um valor de negócios de 856 milhões de euros, para além dos inúmeros eventos ligados a esta atividade e do elevado número de cursos e *workshops* associados a esta área. Estas atividades de disseminação e marketing são fornecidas por diferentes entidades e refletem a preocupação dos vários setores empresariais envolvidos em comunicar o produto – neste caso, o vinho – aos (futuros) consumidores.

A atual e crescente procura por vinho de qualidade, desde a seleção das melhores condições de solo e clima para o cultivo da vinha, passando por preocupações de harmonização de castas e terminando no momento da prova, faz com que a ASE seja vista como uma ferramenta essencial na descrição e avaliação do vinho. Ao especialista (enólogo e produtor), a ASE permite (i) adquirir informação sobre como melhorar as características do vinho e da vinha, (ii) decidir se comercializa ou não um vinho e para

---

phénomène anthropologique, culturel et social fondé sur une conception modulaire du sensorium (la vue perçoit la couleur, l'odorat l'odeur, etc.) qu'invalident les neurosciences : elles ont démontré l'organisation multisensorielle du cerveau qui, grâce à sa "plasticité modulaire" (Howes, 2010, p. 37), peut, en cas de lésion, affecter une zone dédiée à l'usage d'un sens à un autre sens." .... "À grand renfort d'exemples, Benoît Le Blanc et Bernard Claverie montrent ici le mode de fonctionnement global, croisé et multimodal de nos perceptions sensorielles, à la source d'illusions et d'associations au quotidien, elles sont aussi supports d'augmentation humaine au moyen d'instrumentations techniques de plus en plus intrusives, voire inquiétantes. Très importantes, ces avancées n'ont pas encore d'influence sur la conscience ordinaire, traditionnelle et triviale de l'usage de nos sens dont nous voulons ici esquisser l'origine." (Letontunier et Munier (2016:18)).

que público, (iii) informar o consumidor das características do vinho, entre outros. Ao consumidor, a ASE possibilita um maior conhecimento do vinho que irá comprar/degustar.

A prova de vinho é parte integrante das análises de qualidade criadas para identificar defeitos ou características do vinho e auxiliar na sua melhoria ou manutenção (Rosado, 2013). A ASE é definida como um “exame das características organolépticas de um produto pelos órgãos dos sentidos” (Norma EN ISO 5492:2009-en), ou seja, o ato de provar e descrever as características do vinho. Durante este exame, é feita uma percepção sensorial, usando o paladar, o olfato, a visão e o tato, que transmite ao provador-especialista as sensações resultantes dos respectivos estímulos provocados pelo vinho.

No entanto, desde o processo da escolha do local para a plantação da vinha, passando pelas questões climáticas durante o amadurecimento da uva, até à forma como se apresenta o vinho no copo num determinado local para que o especialista faça a prova, sabe-se que há um conjunto de fatores externos que podem influenciar a prova e que podem ser categorizados em três grupos:

- situacionais: pelas sensações, pelas emoções, pela hora da prova, se o provador é do sexo masculino ou feminino, etc.;
- de produção: pelas condições do terreno e do ambiente onde estava a vinha;
- de processo: por aspetos biológicos, como, por exemplo, que castas estão presentes no vinho, que químicos foram usados ou desenvolvidos, etc.

A ASE permite, assim, (i) identificar diferenças, definir e medir atributos sensoriais dos produtos, compreender se as distinções nos produtos são identificadas e aprovadas ou não pelo consumidor; (ii) responder a um conjunto de perguntas referentes a três aspetos importantes: a descrição, a discriminação e a preferência<sup>36</sup>; e (iii) facilitar a catalogação

---

<sup>36</sup> “**Descrição:** A que é que sabe o produto? Quais são as suas características sensoriais apercebidas? De que modo a qualidade do produto difere de outro produto? Quais são as consequências de uma modificação no processo, formulação, embalagem ou condições de armazenamento nos atributos do produto?”

e classificação do vinho provado, sendo este último aspeto altamente valorizado quando se pretende descrever um produto para efeitos de comercialização, quer para as empresas quer para o público em geral.

Esta ferramenta, a ASE, é utilizada por especialistas de diferentes áreas. Tomando como exemplo a área do *marketing*, é necessário descrever o produto como a empresa pretende que o consumidor o identifique, mas também é preciso saber como o consumidor descreve o produto, dado que há casos em que pode não existir acordo entre aquilo que a empresa e o consumidor possuem como imagem/identidade da marca/reputação de um produto e como usa, naturalmente, a língua de especialidade da área. Essencialmente, a ASE é cada vez mais utilizada para caracterizar o produto, evidenciar determinadas propriedades, uma vez que na procura da qualidade é essencial o papel do especialista: “without the winemaker there would be no wine” (Jackson, 2002:227).

Para obter uma caracterização de um produto, na ASE, recorre-se às competências de um painel de provadores-especialistas treinados (Lawless & Heymann, 1998). No caso do vinho, é frequente os provadores-especialistas terem experiências distintas, estando uns atentos a umas características, enquanto outros dedicam mais atenção a outras, mas todos os provadores-especialistas provam e descrevem o produto.

A prova do vinho recorre aos métodos da ASE e tem por finalidade principal determinar e valorizar o conjunto de sensações por ele produzidas. Para a realização de uma prova de vinho, existem vários requisitos a ser respeitados de acordo com a legislação nacional e internacional existente, nomeadamente sobre (i) o local das provas (norma ISO 8589:2010/A1:2014-en), que regula aspetos como a localização, a temperatura, a humidade, o ruído, o odor, a cor das instalações, a iluminação geral, as cabinas; (ii) o tipo de copo utilizado (norma ISO 3591); (iii) a constituição do painel de provadores (normas NP ISO 8586-1:2001-pt e NP ISO 8586-2:2001-pt).

---

**Discriminação:** Será que o consumidor nota a diferença? Será que o consumidor detecta isto? Quantos consumidores detectariam esta diferença? Estes produtos são diferentes? Qual a magnitude da diferença? Será que isto é igual aquilo? **Preferência ou Hedónico:** Quantas pessoas gostam deste produto? O produto é aceitável? Este produto é tão bom como o concorrente? Será que este produto é melhor que o anterior? Quais são as características mais apetecidas? Será o preferido pelo consumidor?” (Noronha, 2006)

A ASE é condicionada por fatores fisiológicos, psicológicos, físicos e de saúde (Noronha, 2006):

- Fatores fisiológicos:
  - (i) Antagonismo: quando os efeitos de dois ou mais estímulos provoca um nível de sensação inferior ao esperado da sobreposição dos efeitos de cada um dos estímulos tomados separadamente.
  - (ii) Ampliação: quando um estímulo aumenta ou modifica a intensidade de outro estímulo.
  - (iii) Sinergismo: quando dois ou mais estímulos provocam um nível de sensação superior ao esperado da sobreposição dos efeitos de cada um dos estímulos tomados separadamente.
  - (iv) Camuflagem: quando um estímulo provoca a modificação da sensação de um outro estímulo.
  
- Fatores psicológicos:
  - (i) Erro de expectativa: quando são fornecidas muitas informações sobre o produto ou sobre o objetivo da avaliação, devem ser apenas indicadas as estritamente necessárias.
  - (ii) Erro de habituação: quando as amostras apresentam estímulos que aumentam ou diminuem gradativamente de intensidade, as avaliações podem não incluir amostras que fujam da tendência.
  - (iii) Erro psicológico: quando existe erro de habituação, uma vez que induz o avaliador a associar uma característica a outra.
  - (iv) Erro do efeito halo: quando, ao invés de as classificar individualmente duas ou mais características, é solicitado ao provador que as avalie simultaneamente. Neste caso, o resultado pode ser diferente daquele obtido da análise individual dos mesmos.

- Fatores físicos:
  - (i) Horário da prova: sempre que possível, a prova deve ocorrer num horário em que os avaliadores não estejam com muita fome ou completamente saciados.
  - (ii) Duração e intervalo entre as provas: é conveniente que a duração de cada prova ou o intervalo entre elas seja o suficiente para que não haja saturação dos sentidos (visuais, olfativos e gustativos) dos avaliadores.
  - (iii) Local da prova: relativamente ao ambiente onde a prova é realizada, deve apresentar temperatura e iluminação adequadas. Deve ser isento de odores externos e deve haver cabine individual para cada avaliador.
  - (iv) Temperatura das amostras: é fundamental que as amostras apresentem uma temperatura compatível com a qual são normalmente consumidas.
  
- Fatores de saúde: sempre que possível, os avaliadores devem apresentar bom estado de saúde, físico e mental, e não divergir muito do seu estado normal, ou seja, não serem pessoas demasiado acanhadas, agitadas, inseguras, entre outros.

Durante a prova do vinho, é feita uma percepção sensorial relacionada com os sentidos – paladar, olfato, visão e tato –, transmitindo ao provador-especialista as várias sensações resultantes de certos estímulos. De acordo com a norma EN ISO 5492:2009-en, as etapas da prova são as seguintes: exame visual, exame olfativo, exame gustativo, sensação tátil.

Na Figura 8, é apresentada a relação entre os órgãos, os sentidos e as características sentidas pelo provador, para resumir os diferentes exames que se realizam durante a prova de vinhos.

Sentidos utilizados na degustação			
Órgão	Sentido e sensações	Características encontradas	
Olhos	<i>Vista</i>	Cor, limpidez, fluidez, efervescência.	Aspecto
	Sensações Visuais		
Nariz	<i>Olfacto</i> (Via nasal directa)	Aroma, »bouquet«	Cheiro
	Sensações olfactivas		
Boca	<i>Olfacto</i> (Via retro-nasal)	Aroma, de boca	Gosto Complexo
	Sensações olfactivas		
	<i>Paladar</i>	Sabor ou gosto propriamente dito	
	Sensações gustativas		
<i>Sensibilidade química</i>	Adstringência. causticidade, borbulhante.		
<i>Sensibilidade táctil</i>	Consistência	Toque	
<i>Sensibilidade térmica</i>	Temperatura		
			«Flaveur»

Figura 8. Quadro dos sentidos utilizados durante uma prova de vinho (retirado de Peynaud, 1993:17)

Para uma melhor compreensão da Figura 8, serão apresentados mais detalhes de cada exame nas secções seguintes. É nesta secção, sobre cada um dos exames, que importa salientar que existem diferentes tabelas/escalas/níveis para cada um dos parâmetros, consoante os autores, que terão necessariamente implicações ao nível dos descritores usados.

### 3.2.1. EXAME VISUAL

O primeiro exame da prova corresponde ao exame visual. De acordo com a norma EN ISO 5492:2009-en, a visão é o sentido que discrimina as diferenças no mundo exterior pelas impressões sensoriais devido às radiações visíveis captadas pelo olho. O exame visual do vinho possibilita a extração de informações relativas a diversos parâmetros, como a fluidez, a presença de gás carbónico (espuma) e os fenómenos de capilaridade (lágrima do vinho), entre outros.

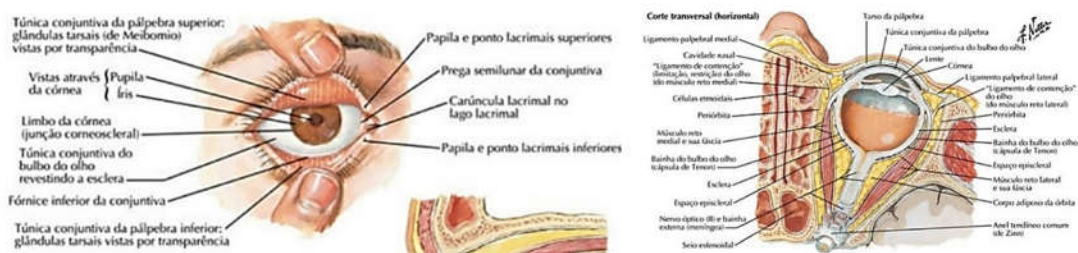


Figura 9. Imagem frontal e transversal relativa ao olho do ser humano (retirado de Netter, 2019)

Regra geral, aquando da ASE, os especialistas têm particular atenção a três parâmetros: limpidez, cor e espuma.

A **limpidez** está relacionada com a estabilidade do vinho e a ausência de turvação. O nível de limpidez está dependente dos processos tecnológicos a que o vinho esteve sujeito (trasfegas, filtrações, centrifugações, colagens, etc.) e à presença (ou não) de alterações químicas e microbiológicas. Um vinho turvo é um vinho instável que foi alvo de uma alteração microbiológica e que não proporcionará uma apreciação gustativa perfeita (no entender dos especialistas). Relativamente a este parâmetro, os especialistas podem classificar o vinho em diversos níveis<sup>37</sup>, consoante o descritor utilizado, podendo ir desde o turvo (nível mais baixo de pontuação), passando pelo encoberto/empoado/opalino, pelo límpido, pelo brilhante, chegando ao cristalino e transparente (nível mais elevado de pontuação).

A **cor** contribui para a discriminação do sabor e do cheiro, influenciando a preferência alimentar. O vinho pode ser tipificado por cor: branco, rosado, tinto. O vinho tinto deve ser analisado com bastante luz e, de preferência, contra um fundo liso; quanto aos vinhos branco e rosado, basta segurar o copo ao nível dos olhos contra um fundo liso. Em relação à cor, há dois aspetos a ter em conta: a tonalidade e a intensidade. Existem diferentes tonalidades associadas a intensidade de cores relacionadas com cada tipo de vinho, como se apresenta de seguida, em algumas listas de diferentes autores.

<sup>37</sup> De novo, importa referir que para cada autor a listagem de descritores utilizados é ligeiramente diferente ou nem sempre é coincidente, quer nos descritores utilizados quer na ordem apresentada. A título de exemplo, vejam-se as diferenças em três autores: Cardoso (2007) refere cristalino, brilhante, límpido, encoberto, empoado ou opalino e turvo. Afonso (2013) apresenta: cristalino, brilhante, limpo, velado e turvo. Carvalheira (2009) refere: brilhante, claro, cristalino, cristalizado, fino, límpido, luminoso, limpo e transparente.

- Para o vinho branco: incolor, citrino, amarelo-pálido, amarelo-esverdeado, amarelo palha, amarelo-dourado, ouro, madeirizado, plúmbeo, âmbar, acastanhado, topázio, caramelo.
- Para o vinho rosado: rosado ligeiro, rosado franco, rosado cereja, rosado framboesa, alperce, casca de cebola, alaranjado, salmão.
- Para o vinho tinto: palhete, rubi, cereja, groselha, sangue, retinto, granada, púrpura, violeta, azulado, cor de tijolo, acastanhado, amarelado.

Para cada uma destas cores (branco, rosado, tinto), a respetiva tonalidade pode estar associada a um ou vários aspetos, como, por exemplo, o tipo de oxidação, a maceração, o tipo de extração de mosto, a intensidade corante, a origem varietal, as castas presentes, entre outros.

A **espuma** é extremamente importante em espumantes ou champanhes. Quanto à bolha, esta pode ser fina, média ou grossa. Relativamente à persistência da espuma, esta pode ser muito ou pouco persistente, duradoura ou fugaz. Idealmente, a espuma deve ser persistente e duradoura, caso contrário dará lugar a uma apreciação negativa.

O exame visual condiciona a restante análise do vinho, pois o aspeto é considerado um indicador seguro do estado de conservação e de estabilidade do vinho.

Na tentativa de facilitar a discriminação das cores do vinho, foram criados instrumentos para auxiliar os especialistas e não-especialistas, como as imagens e a Roda das Cores<sup>38</sup> (que surge posteriormente à Roda dos Aromas, criada por Noble (1990)), como se apresenta na Figura 10.

---

<sup>38</sup> Alerta-se para o facto de que todas as Rodas apresentadas neste trabalho – Roda das Cores, Aromas, Sabores – nas versões em PE podem ser distintas de autor para autor, dado que quem as produz ou traduz pode adotar diferentes listagens de descritores, evidenciando também neste instrumento variabilidade.

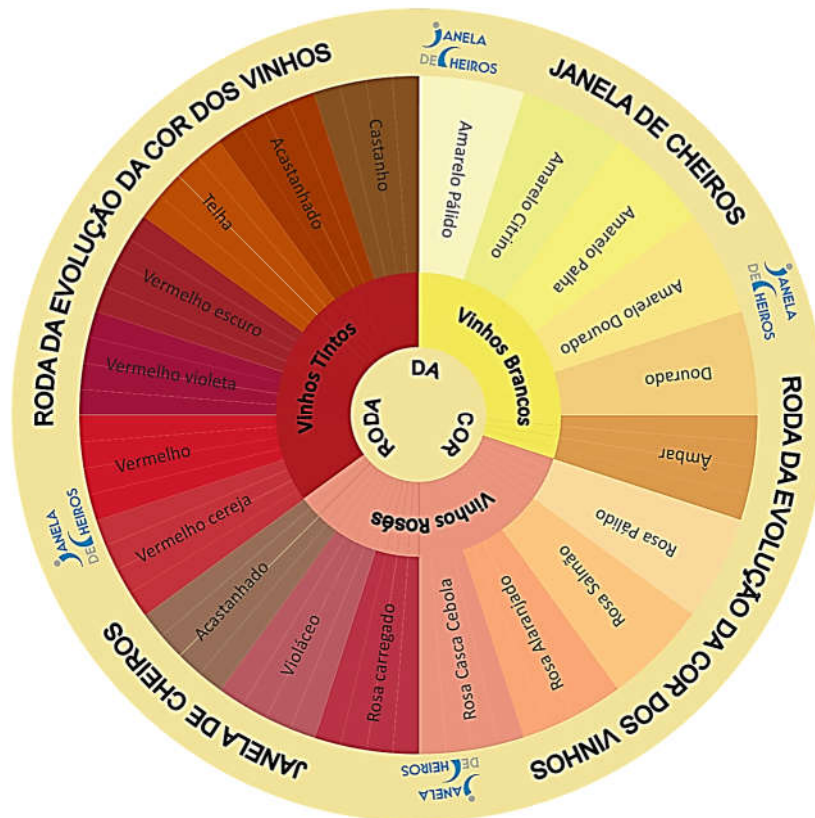


Figura 10. Roda das Cores em: <http://tiny.cc/lur8gz> (disponível e consultado em setembro de 2018)

### 3.2.2. EXAME OLFATIVO

O segundo exame da prova corresponde ao exame olfativo. De acordo com a norma EN ISO 5492:2009-en, o cheiro é definido como o sentido do olfato relativo à percepção de um odor/aroma. Embora este sentido não seja tão desenvolvido no ser humano como em alguns animais, o ser humano consegue detetar um grande espectro de odores e convocar/despertar memórias antigas e variadas. É um sentido complexo e sofisticado que está sempre ativo e que permite não só detetar substâncias maiores ou menores e com uma estrutura molecular simples ou complexa, como também identificá-las mesmo em concentrações baixas (Swiegers, Chambers, & Pretorius, 2005).

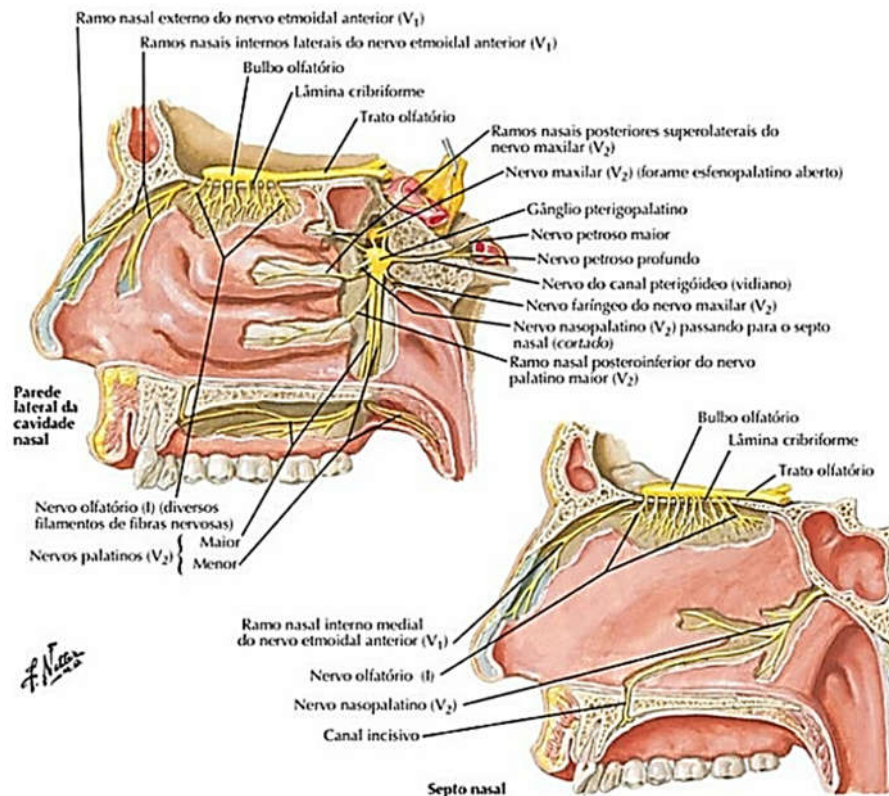


Figura 11. Imagem lateral da cavidade nasal do ser humano (retirado de Netter, 2019)

Sabe-se que o genoma responsável pelo sistema olfatório no ser humano é um dos genes com maior propensão para a variação<sup>39</sup>, explicando, por exemplo, que alguns grupos étnicos consigam sentir com maior intensidade certos aromas ou preferir um certo produto com um determinado aroma a outro, mesmo tendo o mesmo número de receptores olfativos (Dong *et al.*, 2017).

No entanto, segundo Morrot, Brochet e Dubourdiou (2004), a identificação das moléculas odorantes não parece ser totalmente objetiva, dado que, quando um aroma é detetado, existe a tendência de associar mais informação ao mesmo<sup>40</sup> (carga emocional,

<sup>39</sup> A este respeito, sugere-se a consulta e leitura do “The smell survey” de Gilbert e Wysocki (1987) realizado pela National Geographic, no qual foram enviadas amostras de diferentes aromas a voluntários por todo o mundo, perguntando se conheciam o aroma em causa e, em caso afirmativo, se sabiam dizer qual era.

<sup>40</sup> A propósito da possibilidade de um aroma invocar memórias e emoções involuntárias com anos de distância, mas que quando ativadas surgem vividamente, sugere-se as leituras de Proust (1918) *À la*

cultural, fisiológica, entre outras), o que afeta a percepção global. Por exemplo, os autores mencionam que parece existir uma tendência para associar os vinhos brancos a odores de frutas mais claras, como a banana, de mel, de manteiga, e os vinhos tintos a odores de frutos mais escuros, como frutos vermelhos, de madeira e de especiarias.

Os aromas do vinho resultam do *terroir* (casta, viticultura e localização da vinha), da vinificação e do envelhecimento do vinho.

Em termos gerais, o aroma relativo ao vinho classifica-se em três categorias:

- Aroma primário ou varietal – originário das uvas utilizadas;
- Aroma secundário ou de fermentação – originário das fermentações de características vinosas;
- Aroma terciário, de evolução ou *bouquet* – próprio dos vinhos amadurecidos e sujeitos a envelhecimento (em contacto com o ar ou ao abrigo deste).

Normalmente, os aromas dos vinhos são tipificados por afinidade a odores de outras substâncias: aroma animal, balsâmico, a madeira, a torrefação, frutado, floral ou vegetal. Numa tentativa de descrever os aromas com uma maior precisão, Noble (1990) concebeu a Roda dos Aromas, na qual se apresentam os aromas por famílias: vegetais e homólogos, frutos associados aos vinhos brancos, frutos associados aos vinhos tintos, floral, especiarias, madeira, microbiológico, químico, acre, oxidado, terra, caramelo.

---

*recherche du temps perdu* (o caso das *madeleines*) e de Herz, Eliassen, Beland & Souza (2004), no que respeita ao facto de determinados aromas desencadearem respostas negativas em indivíduos com stress pós-traumático.



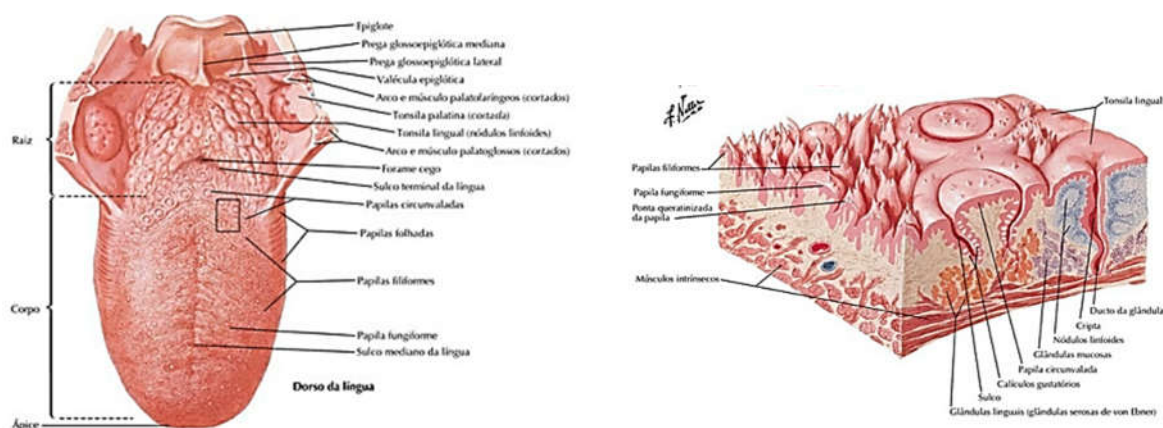


Figura 13. Imagem frontal e de corte transversal da língua do ser humano (retirado de Netter, 2019)

De acordo com a norma EN ISO 5492:2009-en, o sabor é entendido como a sensação detetada pelo órgão gustativo (boca), quando estimulado por certas substâncias solúveis, e o tato é entendido como o reconhecimento pelo contacto direto da pele ou mucosa, da forma e do estado dos corpos que são importantes no julgamento do equilíbrio do vinho.

Este órgão permite experienciar cinco tipos de sensações:

- as sensações olfativas, através da via retronasal, para sentir o aroma de boca;
- as sensações gustativas, relativas aos sabores básicos (doce, salgado, ácido e amargo) e complementares (o alcalino e o metálico)<sup>41</sup>;
- as sensações químicas, como reação das mucosas relativamente à adstringência ou aspereza, a causticidade, a untuosidade e a efervescência;

---

<sup>41</sup> Na cultura asiática existe um quinto sabor. Na literatura chinesa, desde o III século d.C., sempre existiu a noção de um quinto sabor, o picante. No Japão também existia a noção de um quinto sabor, o *umami*, que Ikeda (1909) identificou e isolou como sendo o aminoácido L-glutamato e 5'-ribonucleotídeos, que estaria associado a uma experiência de quase perfeição. Somente em 1985 foi reconhecida a existência do *umami* como quinto sabor reconhecido internacionalmente. Para mais informação, consultar Lindemann, Ogiwara e Ninomiya (2002). Recentemente, também se fala de um sexto sabor no ocidente: o *gosto da água* (Carvalho, 2009)

- as sensações táteis, que permitem sentir a consistência, a liquidez, a fluidez a espessura, o volume, a intensidade taninosa e a opulência (corpo do vinho);
- as sensações térmicas, como o refrescante, o ardente, o frio, o quente e o teor alcoólico (calor intenso na língua).

Para este exame, também existe um instrumento desenvolvido por Gawel, Oberholster e Francis (2000), o *Mouth-feel Wheel*, que representa um conjunto de descritores que podem ser usados para caracterizar o sabor do vinho, como se apresenta a título de exemplo para o vinho tinto.

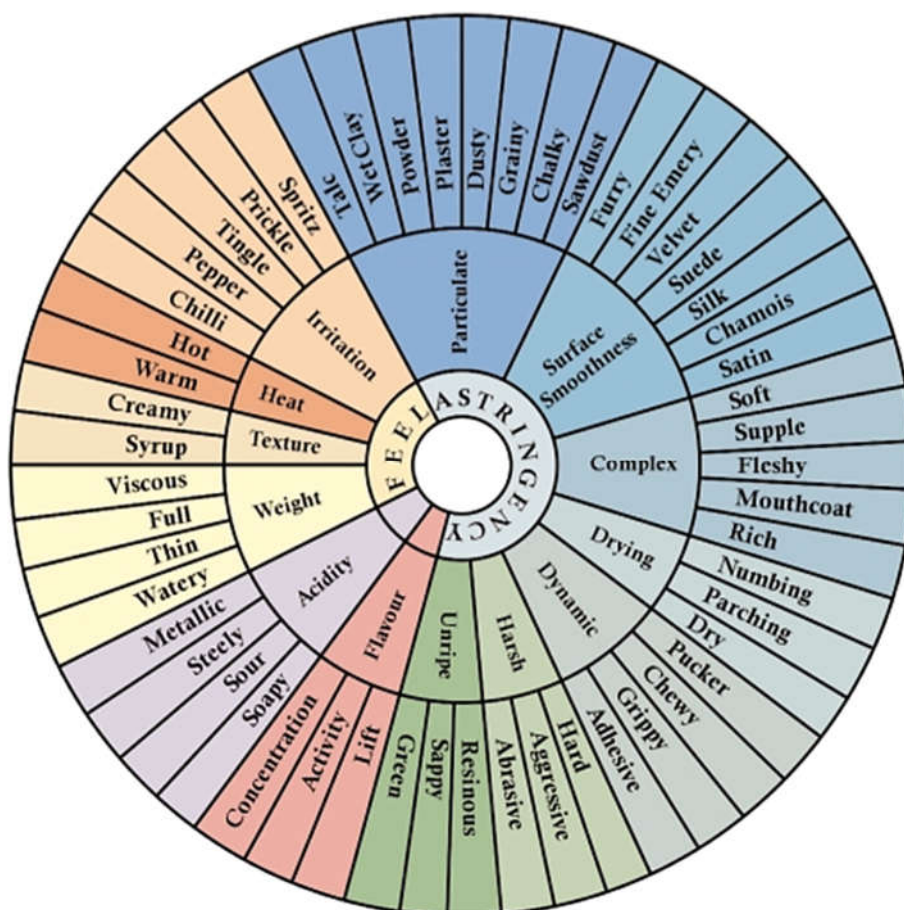


Figura 14. Roda dos Sabores em versão inglesa (retirada de Gawel, Oberholster, & Francis, 2000)

### 3.2.4. EXAME FINAL

A prova de um vinho é um exame complexo e com uma hierarquia predeterminada de ações – observar, cheirar, provar e concluir – que devem ser efetuadas a fim de obter uma avaliação do vinho, como se apresenta na Figura 15.

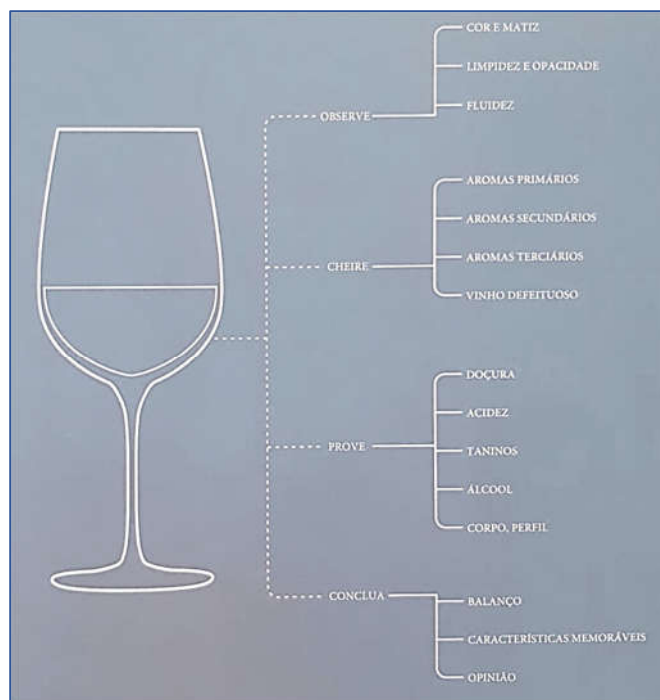


Figura 15. Avaliação do vinho (esquema retirado de Puckette & Hammack (2017))

Durante e após a realização de todos os exames, é preenchida uma ficha descritiva ou uma ficha de pontuação pelo especialista, de modo a ser possível situar a apreciação das múltiplas características e avaliar e quantificar de modo mais objetivo um vinho.

De um modo geral, as fichas avaliam não só os três exames (visual, olfativo e gustativo), mas também os parâmetros de cada um com valor numa escala.

Existem diversos tipos de fichas de prova, como se pode verificar nas Figuras 16 a 19, algumas com opção de resposta aberta, outras de resposta fechada e outras de resposta mista.

Data .....

Nome do provador .....

Dados relativos ao vinho (que não comprometam o anonimato da amostra)

.....

**Apreciação visual**

Cor (intensidade, tonalidade) .....

Limpidez .....

Outras observações .....

**Aroma**

Franqueza (ausência de defeito) .....

Intensidade .....

Descrição .....

Qualidade .....

Defeitos eventuais .....

**Sabor**

Ataque .....

Evolução .....

**Final de boca**

Equilíbrio e estrutura geral .....


Aroma de boca (intensidade e qualidade) .....

Persistência aromática .....

Outras observações .....

Conclusões .....

Figura 16. Exemplo de uma ficha descritiva de resposta aberta (retirado de Cardoso, 2007)

  
**Ficha de prova**

Utilize estas fichas de prova para apreciar e classificar um vinho

**Marca do vinho** \_\_\_\_\_

**Colheita** \_\_\_\_\_

**Produtor** \_\_\_\_\_

**Outros dados** \_\_\_\_\_

\_\_\_\_\_

**Visão**

**Limpidez**       cristalino       límpido       turvo

**Cor**               branco               rosé               tinto

citrino               rosado claro       rubi

palha               rosado escuro     rubi acastanhado

dourado             acastanhado       vermelho rubi

acastanhado       vermelho violeta    vermelho acastanhado

**Fluidéz**               fluído               denso               espesso               viscoso

**Efervescência**       reduzida             moderada           abundante

**Olfacto**

**Aroma\*** \_\_\_\_\_

**Qualidade**           limpo               simples             complexo

**Intensidade**         nula               suficiente           muito intensa

fraca               intensa

**Paladar**

**Aroma de boca\*** \_\_\_\_\_

**Intensidade**         equilibrado       harmonioso       correcto             desequilibrado

**Persistência**       curta               média               longa               muito longa

**Observações** \_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

\*para classificar, utilize a Roda dos Aromas.

[www.infovini.com](http://www.infovini.com)

Figura 17. Exemplo de uma ficha descritiva de resposta fechada (retirado de INFOVINI, 2018)

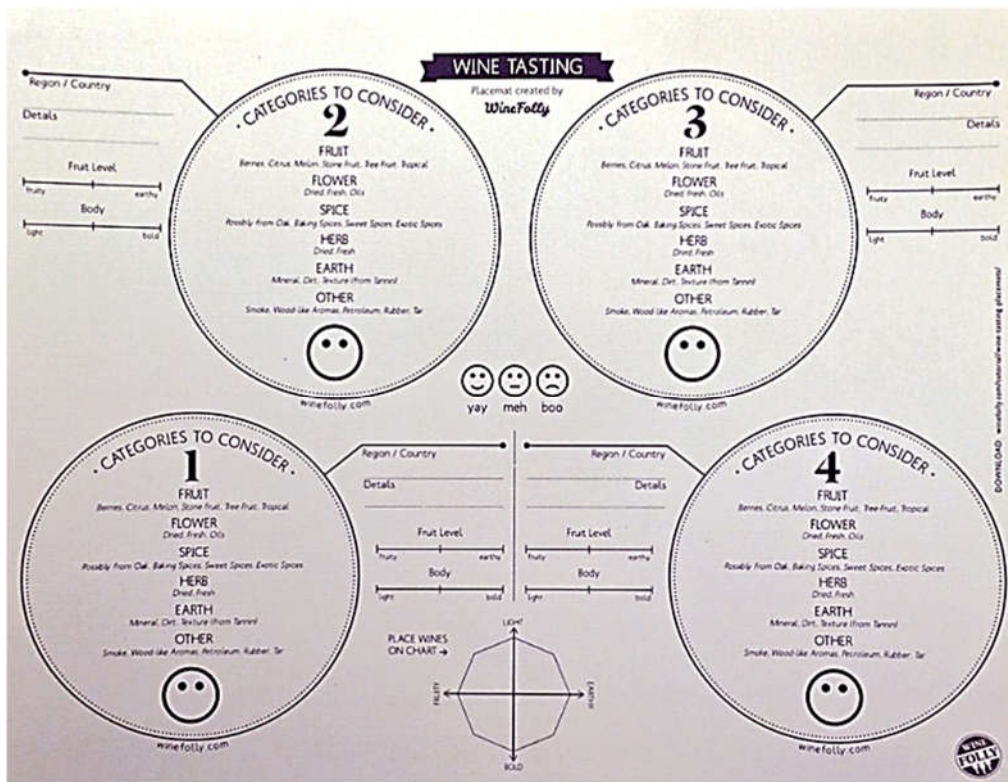


Figura 18. Exemplo de ficha descritiva mista para 4 amostras de vinho (retirada de Wine Folly, 2018)

Existem outras fichas de prova onde o especialista identifica, através de uma pontuação, o valor para cada aspeto. As escalas podem ir dos 0 aos 20 pontos ou até aos 100 pontos, nas quais se consideram *bons* vinhos aqueles que obtêm classificações superiores a 16 e 85 pontos, respetivamente.

**FICHA DE DEGUSTAÇÃO - VINHO TRANQUILO**

NOME DO DEGUSTADOR: \_\_\_\_\_      DEGUSTADOR NO: \_\_\_\_\_      DATA: \_\_\_\_\_  
 CATEGORIA: \_\_\_\_\_      AMOSTRA NO: \_\_\_\_\_      MESANO: \_\_\_\_\_

VINHO TRANQUILO		EXCELENTE	MUITO BOM	BOM	REGULAR	INSUFICIENTE	OBSERVAÇÕES
ASPECTO	LIMPIDEZ	( ) 5	( ) 4	( ) 3	( ) 2	( ) 1	
	ASPECTO	( ) 10	( ) 8	( ) 6	( ) 4	( ) 2	
OLFATO	NITIDEZ	( ) 6	( ) 5	( ) 4	( ) 3	( ) 2	
	INTENSIDADE	( ) 8	( ) 7	( ) 6	( ) 4	( ) 2	
	QUALIDADE	( ) 16	( ) 14	( ) 12	( ) 10	( ) 8	
GOSTO	NITIDEZ	( ) 6	( ) 5	( ) 4	( ) 3	( ) 2	
	INTENSIDADE	( ) 8	( ) 7	( ) 6	( ) 4	( ) 2	
	PERSISTÊNCIA	( ) 8	( ) 7	( ) 6	( ) 5	( ) 4	
	QUALIDADE	( ) 22	( ) 19	( ) 16	( ) 13	( ) 10	
APRECIACÃO GLOBAL		( ) 11	( ) 10	( ) 9	( ) 8	( ) 7	
PONTOS							
ASSINATURA _____							

Ficha de análise sensível do Office Interacademico de la Vigne et du Vin - de l'Union Internationale des Oenologues

Figura 19. Exemplo de uma ficha de avaliação por pontuação (retirado de Clube de Vinhos Portugueses)

Existem, ainda outros, tipos de fichas de prova, que num determinado momento fizeram algum sucesso, pois ajudavam o especialista a situar-se num triângulo (Vedel, 1972), como se pode verificar na Figura 20.

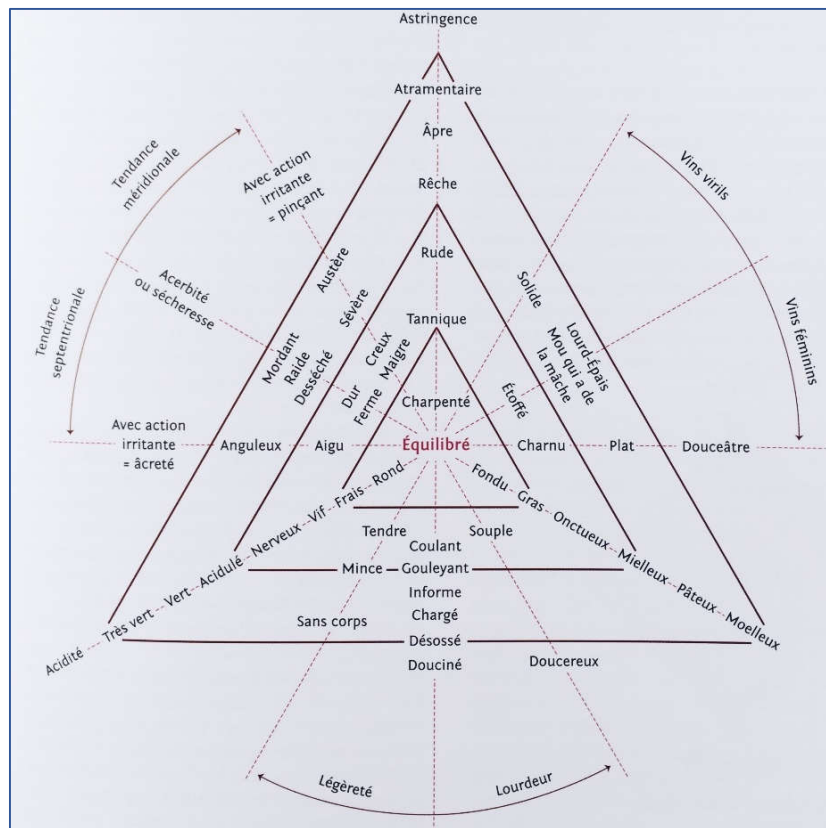


Figura 20. Triângulo de Vedel (retirado de Peynaud & Bloin (2013))

### 3.2.5. TESTES E MÉTODOS DA ASE

Conforme refere Alvelos (2002), existem vários métodos e testes que podem ser aplicados na ASE, tais como:

- Ensaio para a Diferenciação Geral (ex.: teste triangular, teste dois em cinco, teste duo-trio, teste da diferença simples, teste A-não A; teste ABX, teste do padrão duplo; teste da diferença em relação ao controlo, teste sequencial);
- Ensaio para a Diferenciação de Atributos (ex.: comparação por pares direcional; comparação direcional de três amostras; teste de ordenação por pares; teste de

comparação por pares de Scheffé; teste de ordenação simples; teste de cotação de amostras; testes que utilizam desenhos de experiências);

- Testes para determinação de limiares;
- Testes descritivos (perfil de sabor; perfil de textura; análise descritiva quantitativa; método *spectrum*; perfil de escolha livre; análise descritiva de tempo-intensidade);
- Testes afetivos (testes de preferência; testes de aceitação).

Através destes testes, é possível fazer diferentes apreciações do vinho (e de outros produtos) e obter uma avaliação o mais objetiva possível para compreender de que forma o consumidor/utilizador comprova as características do vinho e o que prefere como produto final.

#### 4. APRESENTAÇÃO E PARAMETRIZAÇÃO DOS DADOS

Para identificar o léxico de um domínio de especialidade, considerou-se fundamental a determinação prévia de um conjunto de procedimentos que servisse de modelo e que orientasse essa tarefa. No presente capítulo, descreve-se a realização concreta desses procedimentos, discutindo-se os resultados obtidos sempre que relevante.

Após a delimitação do domínio, considerou-se que o trabalho de recolha dos dados seria sistemático e cobrindo produções monolíngues em PE, uma vez que se pretende estabelecer o léxico de especialidade da ASE nesta língua e ajudar à compreensão de um conjunto de unidades lexicais deste domínio de especialidade. A definição deste primeiro parâmetro geral deveu-se não só à constatação empírica de que a enologia é uma área em expansão no que respeita às técnicas envolvidas, mas também à constatação da existência de poucos trabalhos linguísticos sobre este domínio de especialidade, a que acrescenta a inexistência de uma listagem elaborada de acordo com critérios preestabelecidos e estruturados para o domínio específico da ASE.

Neste contexto, é essencial a recolha e a apresentação de dados reais, que atestem o léxico de especialidade da ASE atual em PE e que permitam um tratamento sistemático do léxico do domínio.

Um trabalho com base nestes dados poderá, assim, contribuir para a melhoria da comunicação entre os especialistas, nomeadamente ao: (i) auxiliar a redação técnica e/ou a tradução de textos especializados para o PE, (ii) fornecer informações relativamente ao uso e ao potencial de ambiguidade de unidades em uso, (iii) esclarecer sobre a forma e o conteúdo das unidades lexicais, através da disponibilização de um recurso atual e informacionalmente rico que permita a consulta ágil da informação e promovendo a estruturação e a sistematização do conhecimento e das unidades lexicais em uso na área.

De acordo com a metodologia seguida, em primeiro lugar foi feito um estado de arte da literatura relativa ao domínio do vinho e da ASE, o que permitiu criar a Base de Dados Enológicos. Posteriormente, e tendo em conta a tipologia de textos disponível em versão papel e digital, e a maior (ou menor) facilidade em selecionar e recolher esses textos, decidiu-se trabalhar com dois subconjuntos de dados: um *corpus* lexicográfico,

relativo a obras lexicográficas existentes produzidos por especialistas, e um *corpus* linguístico relativo aos textos das notas de prova produzidas pelos especialistas. Ambos com características distintas, mas complementares para permitir a identificação e análise de um léxico de especialidade da ASE, como se apresenta nas secções seguintes.

#### 4.1. BASE DE DADOS ENOLÓGICOS

Partindo dos objetivos específicos estipulados para este trabalho, a constituição da Base de Dados Enológicos (BDENO) começou, numa primeira fase, com a recolha de um extenso número de textos sobre o domínio do vinho em vários formatos e várias plataformas, e, numa segunda fase, com a seleção de um conjunto de textos com características particulares, como a seguir se expõe. A designação geral de base de dados, por oposição a uma especificação tipológica de recurso linguístico, advém, como se descreve nas subsecções seguintes, de a base incluir dados linguísticos estruturados (glossários, dicionários, listas de palavras) e não estruturados ou espontâneos (na tipologia de dados de Sinclair (1991) ou Biber (1993)).

Neste contexto, numa primeira perspetiva, consideraram-se como critérios mínimos essenciais de representatividade dos textos selecionados (i) a autenticidade (refletindo a língua em uso), (ii) a autoria por especialistas falantes nativos de PE e (iii) a diversidade, tendo estes de ser de diferentes tipos de textos, de modo a representar uma maior variedade possível dentro do domínio escolhido (Biber *et al.*, 2004; Berber Sardinha, 2000; Sinclair, 2005).

Foram, de seguida, estipulados outros critérios externos para a seleção dos textos no que respeita a:

- Datação: considerou-se apenas os textos produzidos entre 2012 e 2021, embora ainda se tenha considerado a hipótese de alargar o espectro temporal para contemplar alguma obra de referência editada anteriormente, caso se justificasse, o que não veio a acontecer.

- Extensão: com o intuito de obter maior diversidade, foram considerados textos com dimensões distintas (desde artigos de jornal, passando por decretos-lei, até a obras completas).
- Tipologia: foram recolhidos textos dos vários tipos a seguir enumerados, sendo estes representativos da produção nesta área de especialidade.

Relativamente à tipologia de textos, os dados cobrem o seguinte:

- Documentação técnico-científica: dissertações de mestrado, teses de doutoramento, artigos científicos, relatórios; programas de disciplinas escolares e respetiva bibliografia recomendada (relativos a 15 cursos técnicos, licenciaturas, pós-graduações, mestrados e cursos de especialização tecnológica sobre Enologia e Viticultura); glossários, dicionários, listas de palavras, artigos de revistas sobre vinho (seis publicações: *Wine – Essência do Vinho*; *Revista de Vinhos*, *Escanção*, *Enoteca*, *Enovitis* e *Revista da Associação Portuguesa de Enologia*) com publicações mensais, bimestrais e trimestrais;
- Documentação jornalística: textos jornalísticos associados ao domínio do vinho utilizando a base de dados jornalística criada por Martinez<sup>42</sup> (2012);
- Documentação institucional: textos em páginas da Internet de diversos organismos relacionados com o setor do vinho, tais como direções regionais da agricultura, entidades certificadoras, associações, confrarias, comissões, e outras entidades relacionadas com o setor;

---

<sup>42</sup> Martinez (2012) desenvolveu uma base de dados de textos jornalísticos que compila as edições completas desde 1 de janeiro de 2011, com uma média anual de 240 000 artigos e 160 000 000 *tokens*. Esta base de dados reúne também imprensa escrita digital de outros países (Macau, Moçambique, entre outros). No entanto, neste trabalho apenas serão alvo de estudo as seguintes fontes: *Visão*, *A Bola*, *Destak*, *I*, *Sol*, *24 Horas*, *Jornal de Notícias*, *Jornal de Negócios*, *Oje*, *Ciência Hoje*, *Record*, *Diário de Notícias*, *Semanário Económico*, *Portugal Diário*, *Correio da Manhã*, *Expresso*.

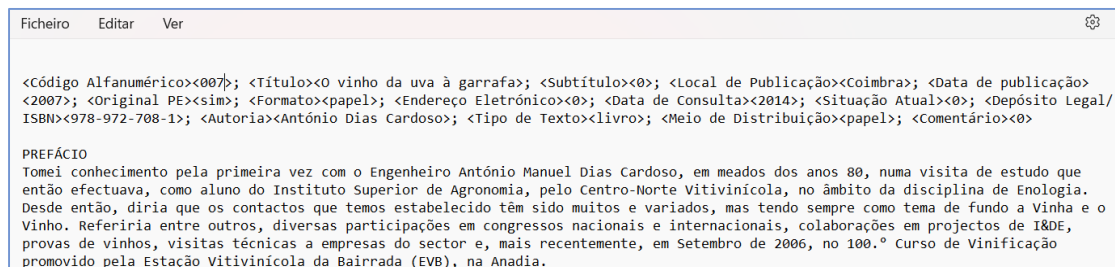
- Documentação legal: textos da legislação nacional, e também da comunitária transposta para ou em uso no país, aplicados aos produtos do setor vitivinícola.

Em termos de tratamento inicial dos dados para integração na base de dados, foram realizados dois passos, tendo em conta as características dos dados recolhidos:

- Conversão: todos os textos em versão papel foram i) digitalizados, objeto de reconhecimento automático de caracteres e manualmente corrigidos, ou ii) digitados (com exceção de tabelas, figuras, imagens e/ou diagramas) e guardados em texto simples, com possibilidade de codificação dos diacríticos portugueses (Unicode UTF-8); todos os textos em versão digital foram convertidos para texto simples, através do programa HTTrack Website Copier para os dados em HTML ou através da conversão direta (“salvar como...”) ou através do procedimento de “copiar e colar” em ficheiro de texto editável para os dados em PDF.
- Codificação: a todos os ficheiros de dados foi atribuído um código alfanumérico, bem como informação respeitante ao nível de especialização, nome do texto e o respetivo número, resultando na criação de fichas de metadados com a seguinte informação:
  - Código Alfanumérico; Título; Subtítulo; Local de Publicação; Data; Original PE (para saber se é ou não uma tradução); Formato; Endereço Eletrónico; Data de Consulta; Situação Atual (relativo aos *websites* e/ou páginas da *Web* – desativado/eliminado); Depósito Legal/ ISBN; Autoria (nem sempre os *websites* disponibilizam as informações pretendidas, pelo que nesses casos foi decidido indicar o nome da entidade responsável); Tipo de Texto; Meio de Distribuição; Comentários (informação adicional).

A informação codificada em cabeçalho não interfere com o texto propriamente dito e permite recuperar um conjunto de anotações que poderão ser importantes no futuro. Esta codificação interna, na Figura 21, permite rápida recuperação e utilização de

qualquer texto da BDEno para este trabalho, mas também permite a sua reutilização por outros utilizadores no futuro.



```
Ficheiro  Editar  Ver  [ícone]
```

```
<Código Alfanumérico><007>; <Título><O vinho da uva à garrafa>; <Subtítulo><0>; <Local de Publicação><Coimbra>; <Data de publicação>  
<2007>; <Original PE><sim>; <Formato><papel>; <Endereço Eletrónico><0>; <Data de Consulta><2014>; <Situação Atual><0>; <Depósito Legal/  
ISBN><978-972-708-1>; <Autoria><António Dias Cardoso>; <Tipo de Texto><livro>; <Meio de Distribuição><papel>; <comentário><0>
```

PREFÁCIO  
Tomei conhecimento pela primeira vez com o Engenheiro António Manuel Dias Cardoso, em meados dos anos 80, numa visita de estudo que então efectuava, como aluno do Instituto Superior de Agronomia, pelo Centro-Norte Vitivinícola, no âmbito da disciplina de Enologia. Desde então, diria que os contactos que temos estabelecido têm sido muitos e variados, mas tendo sempre como tema de fundo a Vinha e o Vinho. Referiria entre outros, diversas participações em congressos nacionais e internacionais, colaborações em projectos de I&DE, provas de vinhos, visitas técnicas a empresas do sector e, mais recentemente, em Setembro de 2006, no 100.º Curso de Vinificação promovido pela Estação Vitivinícola da Bairrada (EVB), na Anadia.

Figura 21. Exemplo de codificação de um texto recolhido para a BDEno

O código alfanumérico do ficheiro permite a identificação de cada texto recolhido (e também de futuras inclusões) para a BDEno e também a organização em categorias, facilmente identificáveis no processo de pesquisa de textos por características específicas.

De acordo com os critérios de seleção dos textos estabelecidos para a criação da BDEno, foram recolhidos textos de manuais, guias de vinhos, enciclopédias, dicionários, glossários, obras de especialidade, páginas *Web* em PE de diversas entidades de referência relativas ao vinho. A diversidade do universo recolhido é apresentada na Tabela 4, abaixo.

	<b>Tipologia de documentação recolhida</b>	<b>n.º de docs</b>
<b>Documentação técnico-científica</b>	dissertações de mestrado	325
	teses de doutoramento	68
	artigos científicos	762
	relatórios	8
	programas de disciplinas	16
	bibliografia recomendada de programas de disciplinas	16
	livros	14
	revistas	43
	artigos de jornal	6 596
	fichas técnicas	21 495
	glossários, dicionários, léxicos	29
guias	19	
<b>Documentação institucional</b>	<i>websites</i> de direções regionais da agricultura	7
	<i>websites</i> de entidades certificadoras	12
	<i>websites</i> de associações	60
	<i>websites</i> de confrarias	26
<b>Documentação legal</b>	legislação nacional	517
	legislação comunitária	242

Tabela 4. Fontes de textos recolhidos para a BDEno

Após a definição dos critérios e da respetiva construção e organização da BDEno<sup>43</sup> (de modo a obter uma visão abrangente do domínio em causa), considerou-se a possibilidade de alimentar a mesma durante um período alargado no tempo (mesmo após

---

<sup>43</sup> Após a definição dos critérios de seleção dos documentos que poderiam ser integrados na BDEno, esta foi organizada de forma a permitir uma constante integração de novos documentos e de novas tipologias (se assim necessário) através da condição de inserir sempre a informação metatextual que permitirá sempre reconhecer a informação interna a ele associado (a tipologia do documento, datação, autores, entre outros). Uma possível/futura disponibilização da base de dados ao público está a ser estudada, visto que grande parte da documentação não é livre de direitos autorais.

o término deste trabalho), com as evidentes limitações que isso pudesse ter, de modo a obter uma Base de Dados estruturada e equilibrada, passível de ser utilizada para este trabalho em particular e para outros trabalhos em curso pela autora ou outros investigadores num futuro próximo.

#### 4.2. *CORPORA*

Como referido acima, a BDEno consiste num acervo de diferentes fontes de textos relativas à temática do vinho, existentes em PE. No entanto, para o desenvolvimento da proposta de elaboração de um dicionário da ASE, e tendo em conta os objetivos específicos diretamente relacionados com a ASE e as opções metodológicas escolhidas, foi necessário selecionar um conjunto de textos para recolha de léxico e exemplos, por um lado, e para validação de informações linguísticas e de especialidade, por outro.

Com efeito, para este trabalho, considerou-se a identificação de dois subconjuntos de dados: um *corpus* lexicográfico (o CEnoLEX), relativo a obras lexicográficas e outros recursos estruturados existentes produzidos por especialistas, no cumprimento dos critérios de seleção e representatividade determinados por estes, e um *corpus* linguístico relativo aos textos das notas de prova produzidas pelos especialistas (o CEnoTEXT).

A necessidade destes dois conjuntos de dados é motivada pelas seguintes razões, de ordem diversa:

- o *corpus* lexicográfico CEnoLEX:
  - inclui um conjunto de descritores que os especialistas consideram importantes apresentar aos não-especialistas;
  - apresenta definições desses descritores redigidas por especialistas;
  - permite a observação de um conjunto de descritores e definições que os especialistas consideram relevantes apresentar em obras de cariz lexicográfico.
- o *corpus* linguístico CEnoTEXT:

- inclui textos autênticos escritos por especialistas;
- é maioritariamente composto por textos concisos, organizados, estruturados;
- inclui texto semirrígido no sentido de ter uma estrutura recorrente, mas que permite liberdade neológica;
- permite perceber as características do vinho consideradas relevantes na descrição.

Se, por um lado, o CEnoLEX permite observar as unidades e as suas definições atestadas e reconhecidas pelos especialistas, o CEnoTEXT permite observar o uso das unidades lexicais efetivamente utilizadas em contexto de especialidade real. A confrontação de ambos e dos seus respetivos dados permitirá definir uma listagem inicial para a elaboração da proposta de dicionário. Os dois *corpora* permitem o confronto de uma abordagem, partindo dos dados (direcionada pelo *corpus*, funcionando como *corpus* de estudo) vs. uma abordagem pela listagem de descritores dos especialistas (funcionando como *corpus* de referência), permitindo uma lista de candidatos a descritores e um conjunto de reflexões em ambas as direções sobre o material recolhido, selecionado e analisado.

O passo seguinte foi preparar os *corpora*, examinando o tipo de tratamento que deveriam ter.

#### 4.2.1. TRATAMENTO SEMIAUTOMÁTICO

Para ambos os *corpora* foi realizada a identificação, a seleção e a recolha dos textos e respetiva preparação dos mesmos, de forma a serem submetidos a um tratamento de

extração de unidades, através de diferentes programas informáticos: o Excel®, o AntConc<sup>44</sup>, o Lexico3<sup>45</sup> e o Sketch Engine<sup>46</sup>.

O Excel® permitiu a limpeza e a organização da informação recolhida de ambos os *corpora* em diferentes folhas com informação de frequências, ordenação, e criação de uma matriz para a Base de dados lexicográfica (BDEnoLEX), que servirá de base para a criação do dicionário.

O AntConc é um programa para tratamento semiautomático de *corpora* que, além identificação da frequência das palavras, permite obter concordâncias e contextos de ocorrência de forma a conseguir uma melhor informação sobre a utilização da unidade escolhida. Um estudo exaustivo das unidades e das suas concordâncias, feito sobretudo com o programa AntConc, permitiu a elaboração de uma lista de colocações, falsos amigos e fraseologias em conformidade com o objetivo inicial deste trabalho.

O Lexico3 é um programa que permite observar as concordâncias, o inventário dos segmentos repetidos, inventários distribucionais, especificidades de formas simples ou de conjuntos de formas, e fazer uma análise fatorial, uma vez que se baseia em Estatística Textual através do cálculo de Especificidades (*Calcul des spécificités*, Lafon 1984) e da Análise de Palavras-Chave (*Keyword Analysis*, Scott & Tribble 2006)).

O Sketch Engine é um programa que permite criar, organizar e explorar *corpora*, através de um conjunto de opções, como observar as frequências, as concordâncias. É, igualmente, possível analisar o *corpus*, bem como o *Word Sketch*, que permite a visualização esquemática das coocorrências com relevância estatística de uma determinada forma ou lema.

Nas secções seguintes, serão apresentados com maior detalhe os procedimentos relativos aos *corpora* e a forma como foram analisados.

---

<sup>44</sup> Disponível em: <https://www.laurenceanthony.net/software/antconc/>, consultado a 19/12/2021.

<sup>45</sup> Disponível em: <http://lexi-co.com/>, consultado a 19/12/2021.

<sup>46</sup> Disponível em: <https://www.sketchengine.eu/>, consultado a 19/12/2021.

## 4.2.2. CORPUS ENOLÓGICO LEXICOGRÁFICO

A necessidade de criar um *corpus* enológico lexicográfico (CENoLEX) deve-se ao facto de não existirem muitas obras de referência em PE relativamente aos descritores, nomeadamente à sua definição. Deste modo, decidiu-se desenhar um *corpus* que permitisse, separadamente, compilar um conjunto de glossários disponíveis para, posteriormente, ser mais fácil e rápida a consulta dos candidatos a unidades lexicais de especialidade, a definição proposta e se os candidatos encontrados no CENoTEXT estariam atestados e adequadamente descritos em algum dos recursos constantes do CENoLEX.

### 4.2.2.1. FONTES LEXICOGRÁFICAS E ANÁLISE PRELIMINAR

A seleção dos textos a figurar neste *corpus* constituiu uma etapa adicional, mas que se considerou fundamental, dado que não só concentra num só local a informação de diversos glossários sobre a ASE, mas também facilita o processo de procura de informação relativa aos descritores.

A primeira tarefa consistiu em pesquisar em versão papel e digital as fontes textuais que fossem glossários/léxicos/dicionários do vinho ou da ASE, originalmente escritas em PE (excluindo-se todas as versões em PB), mas em determinados casos também foram selecionadas traduções para a língua portuguesa de autores estrangeiros.

Foram recolhidas 26 obras lexicográficas (glossários, dicionários, léxicos, vocabulários), com um total de 9 521 entradas e respetivas definições, em versão papel e versão digital, entre 2012 e 2021:

- versão digital: <http://www.abcdosvinhos.com/>, <http://www.ivdp.pt/>, <http://www.aesbuc.pt/>, <http://www.vinhotinto.com/>, <http://www.adegaalmeirim.pt/>, <http://www.cantanhede.com/>, <http://cct.portodigital.pt/>, <http://www.justinosmadeira.com/>, <http://www.esporao.com/>, <http://www.carmim.eu/>, <http://clubedevinhos.com/>, <http://enologoonline.continente.pt/>, <http://www.infovini.com/>, <http://www.ivv.min-agricultura.pt/>, <http://www.vinhos.online.pt/>, <http://www.vinhosdoalentejo.pt/>, <http://www.vinhosdoalentejo.pt/>, <http://www.vinhosdoalentejo.pt/>

w.wineboutique.pt/,<http://www.vinhoverde.pt/>,<http://www.revistadeguste.com/>,  
<http://www.sograpevinhos.eu/>.

- versão papel: Afonso (2013), Dalmasso (1963), Desseauve (1998), Pintão e Cabral (2014), Robinson (2016), Salvador (2010).

Inicialmente, preencheu-se uma ficha de avaliação de cada uma das obras recolhidas, à semelhança do que Faulstich (2011) elaborou, mas com as devidas adaptações, que aqui se apresenta na Tabela 5.

<b>Sobre a obra</b>	<b>Resposta</b>
Título da obra	
Versão papel	
Data de consulta	
Autor(es)	
Editora	
Edição	
Versão digital	
Data de consulta	
Hiperligação	
Assume-se como	
<b>Codificação na CEnoLEX</b>	
Código interno	
Data de exportação	
Tipo de exportação	
Número de entradas inseridas	
Disponível/consultável em 2021?	
<b>Sobre o(s) autor(es)</b>	
Trata-se de pessoa reconhecida na área da Lexicografia e/ou Dicionarística?	
Fez parte de grupo de pesquisa da área d Lexicografia e/ou Dicionarística?	
Qual a formação académica do autor principal e dos participantes do grupo de pesquisa?	
Qual a profissão exercida na época da publicação da obra em análise?	
<b>Sobre a apresentação da obra pelo autor</b>	
Há introdução/prefácio/apresentação?	
Assinada por quem?	
São apresentados os objetivos da obra?	
Qual o público-alvo?	
Tem guia de uso sobre como consultar a obra?	
Há referências à bibliografia de onde foi extraído o <i>corpus</i> ?	
A obra é mono-, bi- ou plurilingue?	
Há menção ao número de entradas, subentradas, aceções?	
<b>Sobre o conteúdo</b>	
Há algum guia que esclareça os componentes do artigo?	

As ilustrações, se houver, estão adequadas à microestrutura informacional?	
O sistema de abreviações e de símbolos aparece corretamente no corpo do texto?	
Os verbetes são apresentados em ordem alfabética? Em ordem temática?	
As entradas apresentam neologismos, palavras derivadas, etc.?	
Há entradas separadas por domínio?	
Os verbetes apresentam:	
Categoria gramatical?	
Género?	
Sinonímia?	
Variante(s) da entrada?	
Variante(s) da definição?	
Crítérios para distinguir homonímia de polissemia? Quais?	
Marcas de uso? Como se classificam?	
Indicação de área ou subárea de especialidade?	
Contexto? (exemplo ou abonação?)	
Equivalente(s)?	
Formação da palavra?	
Indicação de pronúncia?	
Origem e etimologia?	
Divisão silábica?	
Remissivas entre entradas?	
Fontes?	
Notas?	
A definição é constituída de um enunciado de uma só frase?	
A definição leva em conta o nível de discurso do usuário?	

Tabela 5. Modelo de ficha de avaliação das obras seleccionadas

De forma resumida, após o preenchimento deste modelo de ficha de avaliação para as 26 obras recolhidas, verificou-se que há uma ligeira diferença entre as obras em versão digital e em versão papel. Nas 19 obras em versão digital, não há qualquer referência à autoria da listagem apresentada, nem a mesma foi dada após contacto direto sobre o assunto. As mesmas não apresentam qualquer informação relativa à dimensão da nomenclatura apresentada nem aos critérios utilizados na sua elaboração, não facultam um guia de utilização, algumas permitem apenas pesquisa por palavra (nem sempre a pesquisa avançada era possível). As 7 obras em versão papel apresentam informação sobre a autoria, mas apenas duas apresentam informação relativa à dimensão da nomenclatura e facultam um guia de utilização, sendo que o modo de pesquisa é o tradicional alfabético. Em todas, o número de entradas parece ser aleatório (desde 30 a

2003) e a estrutura da definição é também muito diversa (apenas uma palavra, por sinonímia, ou expressa em várias frases).

Em primeiro lugar, foi criado um ficheiro em formato Excel®, no qual foram incorporadas todas as entradas e definições (e em alguns casos informação acessória, como equivalentes noutras línguas, informação enciclopédica, sinónimos, etc.) das 26 obras que foram codificadas de 1 a 26, respetivamente, cada linha correspondendo a uma entrada e respetiva informação associada, num total de 9 521 entradas/linhas recolhidas em todas as obras. Como se apresenta a título de exemplo na Tabela 6, a entrada para “abafado”, que ocorre em seis obras distintas (nas obras 1 (Abcdosvinhos), 10 (Infovini), 16 (Sogrape), 20 (Vinho verde), 26 (Dalmasso), 27 (Afonso)) seguido da definição registada em cada uma dessas obras.

<b>Entrada</b>	<b>Definição</b>	<b>Obra onde ocorre</b>
<b>abafado</b>	Vinho cuja a fermentação foi interrompida através da adição de aguardente ou álcool vínico, permanecendo mais ou menos doce (uma vez que o açúcar natural da uva não se transformou em álcool) e com elevado teor alcoólico.	1
<b>abafado</b>	Vinho a que se interrompeu a fermentação através da adição de aguardente ou álcool vínico, permanecendo doce (uma vez que o açúcar natural da uva não se transformou em álcool) e com elevado teor alcoólico.	10
<b>abafado</b>	Vinho a que se interrompeu a fermentação através da adição de aguardente ou álcool vínico, permanecendo doce (uma vez que o açúcar natural da uva não se transformou em álcool) e com elevado teor alcoólico.	16
<b>abafado</b>	Vinho a que se interrompeu a fermentação através da adição de aguardente ou álcool vínico, permanecendo doce (uma vez que o açúcar natural da uva não se transformou em álcool) e com elevado teor alcoólico.	20
<b>abafado</b>	Mosto cuja fermentação foi «amuada» pela adição de aguardente para lhe conservar a força sacarina.	26
<b>abafado</b>	vinho doce natural	27

Tabela 6. Exemplo de “abafado” nas diferentes obras

Posteriormente, procedeu-se à eliminação de entradas (e em alguns casos definições também) multiplicadas. Ficava apenas uma entrada com diferentes definições associadas e com marcação da obra onde estava atestada. Ou seja, para a entrada “abafado”, verificou-se que a definição era *exatamente* igual nas obras 10, 16 e 20. Como tal, foi

mantida apenas uma linha de entrada e inserida a informação de atestação na coluna correspondente (obra(s) onde ocorre). Este processo resultou num total de 6 004 entradas, como se observa na Tabela 7, para o exemplo das múltiplas entradas de “abafado”.

		Definição	Obra(s) onde ocorre
<i>Abafado</i>	1	Vinho a que se interrompeu a fermentação através da adição de aguardente ou álcool vínico, permanecendo doce (uma vez que o açúcar natural da uva não se transformou em álcool) e com elevado teor alcoólico.	10, 16, 20
	2	Vinho cuja a fermentação foi interrompida através da adição de aguardente ou álcool vínico, permanecendo mais ou menos doce (uma vez que o açúcar natural da uva não se transformou em álcool) e com elevado teor alcoólico.	1
	3	vinho doce natural	27
	4	Mosto cuja fermentação foi «amuada» pela adição de aguardente para lhe conservar a força sacarina.	26

Tabela 7. Exemplo da entrada para “abafado” organizado consoante a definição apresentada nas fontes

Após este procedimento, verificaram-se os casos de entradas onde constavam variantes ortográficas (o caso das siglas “DOC”, “D.O.C.”, que foram associadas à entrada “Denominação de Origem Controlada”), entradas escritas com gralhas (como por exemplo, “água pé”, que foi associada à entrada “água-pé”) ou ainda casos onde foi feita uma harmonização gráfica, visto que havia grafia pré-AO90 e pós-AO90 (o exemplo “decepcionante” e “dececionante”, que foi associado à entrada “dececionante”).

Optou-se, para efeitos de organização, por ter uma entrada/linha como entrada principal e, na mesma linha, outras denominações, as definições, e a(s) obra(s) de onde foi extraída a informação. Desta análise e organização, resultou um total de 3 833 entradas/linhas distintas correspondentes a descritores únicos.

Após a eliminação de linhas com entradas e definições duplicadas, e a organização de entradas com variantes ortográficas, considerou-se que seria igualmente pertinente para a análise dos dados criar e aplicar um conjunto de etiquetas que possibilitasse a identificação dos descritores de acordo com categorias associadas ao domínio da Enologia. Esta etiquetagem (que não existia em nenhuma das obras consultadas) permitiria agrupar os descritores consoante a etiqueta e, num momento futuro, auxiliar na

sua pesquisa de informação relativa a subdomínios da ASE, se tal fosse necessário. Para tal, foi revista manualmente a listagem de 3 833 descritores e chegou-se à conclusão que as seguintes etiquetas seriam as pertinentes para a organização dos dados obtidos, como se apresenta de seguida, acompanhadas de um exemplo: castas (*Moscatel*), tipo de classificação (*reserva*), documentação (*registo de adega*), doença (*podridão negra*), entidades (*Instituto da Vinha e do Vinho*), instrumento/utensílio (*pipa*), local (*vinoteca*), profissão (*tanoeiro*), região (*Douro*), sensorial (*aberto*), substância (*antocianina*), técnica (*aração*), vinha (*peciolo*), vinho (*generoso*).

Como o interesse deste trabalho é a ASE, selecionou-se apenas os descritores marcados com a etiqueta *sensorial* e categorizou-se cada um deles com a etiqueta *modalidade sensorial* (visual, olfativa e/ou gustativa), consoante as definições mencionavam explícita (como o descritor “apagado – de aroma inexpressivo”, que foi etiquetado como pertencente à modalidade sensorial olfativa) ou implicitamente (como o descritor “opaco – Sem limpidez, turvo”, que foi etiquetado como pertencente à modalidade sensorial visual) a modalidade sensorial a que estavam associados. Os 1 054 descritores distribuem-se pelas seguintes modalidades sensoriais: 164 para a modalidade visual, 271 para a modalidade olfativa, 478 para a modalidade gustativa, 44 associados a sensorial, mas sem ser uma modalidade sensorial (como por exemplo, “prova organoléptica”, que não é um descritor de nenhuma modalidade sensorial, mas está associado a ASE) e 97 que estão associados a mais do que uma modalidade sensorial. Esta nova etiquetagem (com a etiqueta *sensorial* e a etiqueta *modalidade sensorial*) surge da verificação de que, em nenhuma das obras consultadas, havia menção desta informação e de esta poder ser útil na organização dos descritores por modalidade, e também ser uma característica possivelmente relevante para apresentar aos utilizadores de um futuro recurso lexicográfico.

Subsequentemente, para cada descritor do CEnoLEX, foi anotado (semiautomaticamente e com revisão manual) a categoria gramatical, visto que, na maioria das obras de onde foram extraídas, essa informação raramente é mencionada ao utilizador. Foram identificados 571 nomes, 461 adjetivos, 14 verbos.

Observou-se, para cada descritor no CEnoLEX, em quantas obras estava atestado. Tal contagem justifica-se para verificar quais as unidades lexicais que muitos especialistas consideram pertinentes para constar nas suas obras e aquelas unidades lexicais que apenas alguns especialistas consideram relevantes, como se observa na Tabela 8. Por exemplo, “persistente” ocorre em apenas uma das 26 obras, já “buquê” ocorre em 24 obras consultadas. Verifica-se que num total de 1 054 descritores, nenhum ocorre em todas as obras e 1 025 descritores surgem em menos de metade das obras, mostrando que há uma variação no número de unidades lexicais que os especialistas consideram que devem constar nas obras lexicográficas que elaboraram.

N.º de descritores	Nº de obras em que surgem	Exemplos
235	1	<i>persistente</i>
112	2	<i>semidoce</i>
245	3	<i>silex</i>
97	4	<i>excecional</i>
140	5	<i>ténue</i>
64	6	<i>mole</i>
39	7	<i>apagado</i>
24	8	<i>tinto</i>
25	9	<i>denso</i>
13	10	<i>torrado</i>
10	11	<i>adstringência</i>
8	12	<i>lágrima</i>
13	13	<i>balsâmico</i>
4	14	<i>equilibrado</i>
5	15	<i>untuoso</i>
4	16	<i>seco</i>
5	17	<i>rolha</i>
2	18	<i>generoso</i>
3	19	<i>frutado</i>
2	20	<i>tanino</i>
2	21	<i>corpo</i>
1	22	<i>redondo</i>
1	24	<i>buquê</i>
<b>1 054</b>	<b>26</b>	<b>totais</b>

Tabela 8. Distribuição dos descritores nas obras consultadas

Relativamente à definição, foi também feita uma análise preliminar. Observou-se que, para 1 054 descritores, existiam 2 578 definições associadas. Como acima mencionado, verificou-se que em alguns casos, havia um número de definições escritas

exatamente iguais em diferentes fontes. Perante tal facto, estas foram contabilizadas e foi decidido apenas manter uma dessas definições, com a indicação das diferentes fontes onde surgiam, como se pode observar na Tabela 9. Por exemplo, o descritor “ranço” tem apenas uma definição e ocorre apenas numa obra (o mesmo acontece com 568 unidades lexicais), já um descritor como “áspero” apresenta seis definições (acontecendo o mesmo para 22 outras unidades lexicais) e por fim, “buquê” apresenta 21 definições diferentes.

Descritores	Nº de definições em que surgem	Exemplos
568	1	<i>ranço</i>
207	2	<i>intensidade</i>
89	3	<i>limpidez</i>
54	4	<i>opaco</i>
40	5	<i>cansado</i>
22	6	<i>áspero</i>
13	7	<i>espumante</i>
12	8	<i>harmonioso</i>
7	9	<i>suave</i>
10	10	<i>macio</i>
9	11	<i>generoso</i>
5	12	<i>persistência</i>
5	13	<i>rolha</i>
2	14	<i>adstringente</i>
6	15	<i>frutado</i>
2	16	<i>aveludado</i>
1	17	<i>acidez</i>
1	18	<i>corpo</i>
1	21	<i>buquê</i>
<b>1 054</b>	<b>26</b>	<b>totais</b>

Tabela 9. Distribuição dos descritores consoante o número de definições distintas encontradas nas obras

Outro aspeto que se pode mencionar em relação às entradas do CEnoLEX é a relação entre o número de ocorrências nas obras e o número de definições para cada descritor: há 78 hipóteses desde 1 descritor em 1 obra com 1 definição, até a 1 descritor atestado em 24 obras com 21 definições diferentes.

Relativamente à dimensão, a definição pode corresponder a apenas uma palavra, ou corresponder a 290 palavras em 15 frases, como se pode observar na Tabela 10.

descriptor	Definição
<b>oleoso</b>	Viscoso.
<b>depósito</b>	Diz-se do vinho que tem corpo e untuosidade, é gordo e produz certa sensação de engrossamento da boca. Depósito comercial onde se armazenam os vinhos para o seu estágio (docks do Porto) ou onde se guardam, já engarrafados, à espera de serem vendidos. Sedimento do mosto ou do vinho. O vinho engarrafado pode apresentar depósitos muito diferentes. É um facto normal nos tintos velhos, em que alguns componentes se tornam insolúveis ao envelhecer e formam um depósito bastante espesso, inclusive algumas vezes colado às paredes da garrafa, que se pode separar com facilidade por meio da decantação. Por vezes, os melhores vinhos tintos apresentam depósitos de pigmentos. Poderiam ter sido eliminados na adega com uma forte filtração, mas é de agradecer que o produtor tenha sido mais cuidadoso e tenha assumido o risco de uma possível sedimentação das borras. O vinho agradecerá, guardando toda a sua substância, como acontece nos velhos portos Vintage, que não se filtram e decantam ao serem servidos. Outro depósito relativamente habitual são os pequenos cristais transparentes que aparecem no fundo de algumas garrafas de vinho branco. Também não têm nada de anormal. Trata-se de tartaros (bitartarato de potássio), sais do ácido tartárico: substância que está presente em todas as uvas e todos os vinhos. Os tartaros depositam-se geralmente quando o vinho sofre um choque térmico durante o transporte ou na conservação por causa da passagem a uma temperatura muito baixa. Existem outros depósitos mais ligeiros, que ficam em suspensão quando se manipula a garrafa e que podem corresponder a algum defeito do vinho, a doenças ou acidentes. São estas as causas da maioria dos vinhos turvos. Sedimento fino, alongado, accidental, que adere à parede da garrafa de vinho espumoso depois de terminada a segunda fermentação.

Tabela 10. Exemplos de definições no CEnoLEX

A organização da informação dentro das 2 578 definições apresenta diversas estruturas e dimensões:

- definições curtas, com apenas uma palavra, funcionando como remissão para outro descriptor (sem que fosse dito explicitamente que era um sinónimo, ou com as marcas *o mesmo que, ver*), assumindo um alto grau de inferência do utilizador para a sua compreensão;
- definições de dimensão intermédia, de várias palavras e uma frase longa ou a frases mais curtas;

- definições longas, com várias frases, com maior detalhe e informação, prevendo um baixo grau de inferência do utilizador, lembrando as definições enciclopédicas.

Observou-se como se apresenta na Tabela 11 a seguinte distribuição:

N.º definições	n.º frases	N.º mínimo de palavras (exemplo)	N.º máximo de palavras (exemplo)
1 756	1	1 (oleoso)	57 (herbáceo)
569	2	6 (generoso)	93 (levedura)
156	3	6 (untuoso)	72 (cor)
54	4	10 (floral)	111 (lágrima)
21	5	25 (licoroso)	121 (véu)
15	6	21 (licoroso)	182 (boca)
3	7	106 (amargor)	132 (aroma)
2	8	152 (tanino)	181 (carvalho)
2	15	290 (depósito)	302 (aroma)

Tabela 11. Dimensão da definição quanto ao número de frases e palavras

Um outro aspeto que mereceu alguma atenção foi a forma como as obras estruturavam a definição propriamente dita. Verificou-se que há variabilidade quanto à sistematicidade (isto é, se todas as definições começavam pelo hiperónimo ou constituíam em si uma frase completa), quanto à sua estrutura sintática (por exemplo, se se tratava de relativas iniciadas por “que” ou frase incompleta iniciada por “é” ou outro verbo), bem como à forma e sistematicidade na ordenação da informação internamente (exemplo, se indicavam características semelhantes no final da definição). É evidente que tal facto se deve em parte por serem resultado da extração de definições de 26 obras. Não obstante, mesmo entradas da mesma obra apresentavam esta variabilidade e esta deve ser mencionada, porque evidencia que os critérios para a elaboração das definições parecem ser incoerentes, inconsistentes ou até mesmo inexistentes, como se pode observar na Figura 22.

ENTRADA	DEFINIÇÃO	OBRA(S) onde ocorre
ABOCADO	significa doce, e a etnologia da palavra é italiana (abbocato).	21
ACIDEZ	Característica essencial dos vinhos que contribui para o seu sabor, frescura e capacidade de conservação, a acidez provoca uma sensação de frescor agradável, motivada pelos ácidos do vinho (citríco, tartárico, málico, láctico, succínico) e que resulta em salivação.	15
ACIDO LACTICO	provém da decomposição do ácido málico na fermentação maloláctica. Dá suavidade ao vinho.	27
ADSTRINGENTE	sensação táctil que um excesso de tanino deixa no interior da boca.	9
ÁGUA-PE	Bebida obtida por maceração ou leixivação dos bagaços prensados ou não a que por vezes se juntam substâncias que lhes comunicam o carácter de um vinho.	16
ARDENTE	um vinho que "arde" na boca, que queima por ser excessivamente alcoólico.	25
AROMÁTICO	Quando o vinho é muito perfumado	19
AVELUDADO	Trata-se de um vinho com sabor rico e sumptuoso, de textura sedosa.	6
AZUL	Dá-se o nome de quebra azul à ligeira turvação que se observa nos vinhos brancos. Esta turvação pode formar uma sedimentação na parede da garrafa durante a fermentação e persiste depois da remoção. Pode observar-se uma cor azul ou arulada em alguns vinhos tintos que têm um baixo conteúdo de ácido tartárico.	10, 20
BRILHANTE	(cor) um vinho brilhante é um vinho de aspeto límpido, sem qualquer tipo de turvação ou casse.	4
BUQUÊ	(do francês bouquet) - aroma complexo, também denominado aroma terciário, resultante do envelhecimento.	15
CAPITOSO	termo aplicado a vinhos muito ricos em álcool.	13, 27
CHEIO	Dotado de qualidades gustativas marcantes e de boa amplitude.	28
CURTO	ver "fim de boca".	29
DOURADO	Diz-se da cor apreciada de certos vinhos brancos	26
DURO	Designação dada a um vinho muito tannoso, nada macio ou aveludado.	8, 12
EMPIREUMÁTICO	aroma a substâncias orgânicas queimadas. Cheiro a grelhados e a tostados.	27
ESPIRITUOSO	Descreve os vinhos que são frescos, frutados, brilhantes e vivazes.	6
GLICERINA	É produzida durante a fermentação e contribui para o encorpar do vinho.	6
GRANDE	Excelente	19
JAMBE	Galicismo (perna) que designa as lágrimas que o vinho deixa no copo.	10, 16, 20
LÁGRIMA	Ou "pernas" que escorrem na parede dos copos depois de beber, resultam da diferença da velocidade de evaporação entre a água e o álcool	19
PESADO	Mede a riqueza do vinho e, mais especificamente, a sua consistência. Ainda mais consistente, tornar-se-ia viscoso. Termo frequentemente utilizado para designar um vinho alcoólico	23
PÓ	Odor característico de alguns vinhos, que recordam o da terra poeirenta.	7, 10, 16, 17, 20
RETROGOSTO	Identifica o aroma e sabor deixado pelo vinho na boca após ser engolido. Grandes vinhos têm retrogosto rico, complexo e prolongado	19
ROLHA	Para definir o vinho atacado pelo TCA (Tricloroanisol), o principal composto responsável pelo aroma a rolha que resulta da degradação de compostos clorados em cloroanisóis pela atividade fúngica existente na rolha, nas madeiras e cartões ou ainda por contaminação atmosférica da adega durante processos em que o vinho está exposto ao ar. Nestes casos, diz-se então que o vinho tem aroma a rolha.	28
RUGOSO	Refere-se a uma textura que carece de finura.	28
SEIVA	(do francês séve), diz-se que um vinho tem seiva quando é um grande vinho	2
SUBTIL	qualifica um vinho fino e delicado.	25
TRANQUILO	vinho sem espuma, ou quase sem qualquer efervescência; vinho não generoso.	25
VERDE	1. Diz-se da cor de alguns vinhos brancos.; 2. Diz-se de um vinho ácido e nervoso.; 3. Nome do vinho produzido na Região dos Vinhos Verdes.	27
VERDURA	Sabor especial que têm os vinhos novos.	26
ZURRAPA	Vinho que não presta.	26

Figura 22. Exemplos de diferentes definições no CEnoLEX

Em síntese, o CEnoLEX é um *corpus* lexicográfico monolíngue do PE constituído por um conjunto de 26 textos especializados do tipo lexicográfico, recolhidos entre 2012-2021. Após um processo de organização das entradas e das variantes ortográficas, bem como um processo de etiquetagem das mesmas quanto à sua categoria gramatical e à

categorização no domínio, chegou-se a um total de 1 054 descritores relativos à modalidade sensorial elencados (mas não explicitamente marcados como tal) nas obras lexicográficas elaboradas pelos especialistas, incluindo 571 nomes, 461 adjetivos e 14 verbos. Destes, 164 dizem respeito à modalidade visual, 271 à modalidade olfativa, 478 à modalidade gustativa, 44 descritores estão associados à ASE e 97 estão associados a mais do que uma modalidade sensorial. Curiosamente, observa-se, neste *corpus* que a modalidade sensorial é aquela com menos descritores explicitamente apresentados pelos especialistas, seguindo uma direção ligeiramente diferente da mencionada por outros autores (Croijmans, 2018), nomeadamente de que a modalidade visual seria uma das mais bem descritas, contrastando com o maior número de descritores olfativos e gustativos.

Este facto pode ter diferentes motivações possíveis:

- a definição da cor e a tonalidade é um processo complexo e que é pouco sistemático nos dicionários, como referem Marelló e Onesti (2016);
- a elaboração da definição relativa a cor e tonalidade pode oscilar entre ser [+/- técnica] e [+/- ambígua], parecendo a definição [+ técnica, - ambígua] direcionada para um público altamente especializado, pois implica um conhecimento de carácter científico; a definição [± técnica, ± ambígua] direcionada para um público vasto e heterogéneo, ainda que pressuponha algum conhecimento científico; e por fim, a definição [- técnica, + ambígua] parece direcionada para o público geral, onde a definição surge por meio de associação com uma entidade do mundo real, como descrito em Barbosa e Barbero (2021).

Nas definições observadas no CEnoLEX relativas à modalidade visual, não se verifica a menção de aspetos técnicos, como, por exemplo, a posição dos nanómetros no comprimento de onda das radiações monocromáticas no espectro solar, mas definições, na sua maioria [- técnica, + ambígua]. Tal pode ser explicado por os especialistas considerarem que os descritores visuais não necessitam de ser definidos com grande rigor, dado que é uma das características sensoriais que depende da observação do vinho por si. Por exemplo, em alguns casos, a definição da tonalidade é vaga (“verde – Diz-se da cor de alguns vinhos brancos.”) sem indicar efetivamente de que tonalidade se trata, visto que não há vinhos verdes na tonalidade; em alguns casos, a definição menciona uma gradação

em escala (“branco – Cor clara que vai desde a cor de palha seca, à cor casca de limão e a tons dourados.”), invocando elementos do mundo real que possuem essa cor, mas que no domínio do vinho é uma cor/tonalidade aproximada, não necessariamente aquela que imediatamente é invocada como prototípica (aspecto mencionado em Barbosa (2021)).

Pode extrapolar-se que, pelos motivos supracitados, os especialistas consideram que os descritores visuais são mais complexos de definir e que, também por isso, há um número menor de descritores, quando comparados com o número de descritores olfativos e/ou gustativos, aspecto que contraria as conclusões de Croijmans (2018).

Relativamente aos descritores olfativos e/ou gustativos, há uma tendência muito marcada para definir o seu significado através da invocação de um referente do mundo real (uma planta, um fruto, ...), ampliando a noção de familiaridade do utilizador com o vinho (exemplo “groselha – Odor e sabor frutado que recorda o do fruto da groselheira. Existe uma groselha preta, também chamada cássis, cujo aroma se encontra sobretudo nos vinhos tintos.”).

Observou-se também que os especialistas têm critérios distintos quanto aos descritores que devem figurar nas obras, pois o número de descritores é diverso consoante as obras. Não há um único descritor que figure em todas as obras. Observa-se, aliás, um conjunto de descritores que apenas surge em menos de metade das obras consultadas.

Além disso, o facto de um descritor poder ter até 21 definições distintas com dimensões variáveis, como corresponder a uma palavra ou corresponder a 290 palavras em 15 frases, mostra a variabilidade do tipo de definição associada, sendo que essa variabilidade é interna quanto à informação disponibilizada nas obras consultadas.

Perante a observação e análise objetivas das obras recolhidas, comprova-se que estas, apesar da maior ou menor dimensão quanto ao número de descritores apresentados e da motivação pessoal na sua criação e disponibilização ao público, não foram elaboradas tendo em conta critérios meta e lexicográficos. Não são dadas informações sobre, por exemplo, quais as fontes para criar a obra, a metodologia utilizada para seleccionar as entradas, como foi definido o número de entradas, como foi elaborada a definição, qual a equipa envolvida, entre outros aspetos que os autores por razões várias, não mencionaram,

mas que, do ponto de vista lexicográfico seriam pertinentes apresentar para que a análise destas obras fosse mais profunda e completa.

O tratamento preliminar dos dados do CEnoLEX permitiu observar as características das obras disponíveis e possibilitar a confrontação das mesmas com os dados CEnoLEX que se apresentam no *Capítulo 5*. Permitiu também constatar as lacunas das obras lexicográficas que existem no mercado, evidenciando uma lacuna no mercado de uma obra que considere (i) as preocupações lexicográficas; (ii) a utilização de uma metodologia coerente nas diferentes fases do processo de elaboração de um dicionário; (iii) as particularidades do domínio da ASE, que envolve uma reflexão sobre como definir e apresentar informação sobre a modalidade sensorial; (iv) as necessidades e expectativas dos utilizadores quando recorrem a um dicionário de especialidade.

#### 4.2.3. *CORPUS ENOLÓGICO TEXTUAL*

O domínio de especialidade sobre o qual o trabalho se debruça é a ASE, que permite aos especialistas, através de uma análise explícita, mas pessoal, avaliar e descrever o vinho recorrendo a um conjunto de etiquetas verbais – os descritores.

O facto de o especialista utilizar um conjunto de unidades lexicais – os descritores – levou à necessidade de criar instrumentos como, por exemplo, a Roda dos Aromas, na tentativa de uniformizar o léxico utilizado.

Os descritores utilizados pelo especialista pretendem evocar um conjunto de sensações, de modo a replicar a prova do vinho na mente do leitor/futuro consumidor. Estes descritores são fonte de informação sobre os aspetos sensoriais em particular. Se, por um lado, há descritores não ambíguos e que correspondem a perceções sensoriais identificadas (como ácido, doce, entre outras), por outro, parece existir um número de descritores que se associam à imagem, à analogia para invocar propriedade de objetos, locais, animais, etc., de diferentes domínios, provocando uma imagem vívida, mas potencialmente diferente, naqueles que leem a ficha técnica com a caracterização de um determinado vinho.

#### 4.2.3.1. FONTES TEXTUAIS E ANÁLISE PRELIMINAR

O CEnoTEXT foi constituído pela totalidade dos textos designados como “ficha técnica” ou “nota de prova”, constantes quer nas secções temáticas das revistas quer das páginas Web dos produtores. Esta designação foi seleccionada por representar prototipicamente o léxico de especialidade do domínio da ASE, uma vez que é nesta ficha técnica/nota de prova – num texto sucinto – que os especialistas descrevem as sensações organoléticas de cada vinho (a cor, o aroma, o tipo de fruto, o grau de complexidade, a textura, a persistência final, etc.) e avaliam o vinho, utilizando os descritores disponíveis no seu repertório, no sentido de funcionar como um guia de orientação para quem lê. Além de funcionarem como uma secção de comunicação especializada nesse domínio, e de veicularem informação de certo modo hermética, não deixam de cumprir a função primordial de informar e descrever o produto em questão.

No entanto, acredita-se que esta secção (ficha técnica), como refere Pearson (1998), representa o seguinte:

“When experts in any given field communicate about their subject, they tend, (...) to use a highly specialized jargon. It is assumed that author and reader share a common language and that when certain words or phrases are used, each understands what is meant. This language differs from general language in that specific meanings have been assigned to the language used, and these have been defined prior to the communication act by an external authority.” (Pearson, 1998:36)

Para integrarem o *corpus*, os textos têm de ser escritos em PE por especialistas (ou redatores técnicos, aspeto realçado por Conceição (2005)). Considerando que se trata de uma BDEno de especialidade, os textos são ainda categorizados de acordo com os seguintes cinco níveis de especialização: (i) altamente especializados; (ii) altamente especializados didáticos; (iii) semiespecializados; (iv) banalizados e (v) como refere Lino (2013). Para este trabalho, não foram tidos em conta os textos banalizados e vulgarizados, mas apenas: (i) altamente especializados – textos provenientes de artigos científicos, teses

de doutoramento, dissertações de mestrado, revistas, obras de referência e legislação; (ii) altamente especializados didáticos – textos procedentes de manuais e programas das disciplinas lecionadas nos cursos de enologia; (iii) semiespecializados – textos jornalísticos, páginas da Web dos diferentes organismos e outras obras de referência.

Tendo sido feita a análise de uma amostra de textos recolhidos para a BDEno, considerou-se que a nota de prova (incluída na documentação técnico-científica) constitui um discurso especializado. Como apresentado anteriormente (na Tabela 2 na secção 4.1. relativa à Base de Dados Enológicos), foram considerados apenas os textos da BDEno relativos a fichas técnicas e a notas de prova produzidas por especialistas, num total de 21 495 notas de prova como as que se ilustram na Figura 23, com um exemplo à esquerda retirado de uma revista e um exemplo à direita retirado de uma página da *Internet*.

**QUINTA DA OLIVEIRINHA**  
**GRANDE RESERVA TINTO 2018**  
*(DOC Douro)*

**PRODUTOR** Alves de Sousa  
**DENOMINAÇÃO DE ORIGEM** DOC Douro  
**CASTAS** Vinhas Velhas predominando Touriga Franca, Touriga Nacional, Tinta Amarela (15 no total)  
**VINIFICAÇÃO** Vindima manual. Desengace total. Fermentação com controlo de temperatura. Estágio de 19 meses em barricas de carvalho francês (50% de 2º ano, 30% de 3º ano, 20% novas)  
**TEOR ALCÓOLICO** 14,5%  
**LONGEVIDADE** 10 a 12 anos  
**SERVIÇO** A 16/17°C com pratos de carnes intensas (vaca ou cabrito). Beneficia decantado antes de servir.  
**ENOLOGIA** Tiago Alves de Sousa



*Cor vermelho rubi. Aroma elegante e equilibrado mas muito complexo, com notas de frutos vermelhos, fruta bem madura, notas de boa barrica, "sous bois". Intenso a mostrar boa evolução. Na boca é equilibrado, redondo com boa estrutura, uma boa acidez a dar-lhe frescura. Elegante com taninos suaves. Final persistente.*

**Conde D'Ervideira Private Selection Branco**

Região	Alentejo
Designação	DOC Alentejo
Castas	Antão Vaz, Arinto
Vindima	Mecânica e noturna
Prova	Este vinho apresenta aromas tropicais, associados à frescura do chá verde. Na boca destaca-se pela elegância, com suaves taninos do carvalho húngaro.
Vinificação	As uvas são desengaçadas e enviadas para vinificadas, onde permanecem durante 6 a 10 horas, a uma temperatura de 12°C. Depois são prensadas e enviadas para a câmara frigorífica, onde fermentam entre 10° e 12°C. A fermentação alcoólica do Antão Vaz decorre em barricas novas de carvalho húngaro durante cerca de 1 mês, seguindo-se um estágio com battonage de 9 meses. Posteriormente, é feito o lote com Arinto.
Estágio	10 Meses   Barricas novas de carvalho húngaro
Longevidade	4 anos
Como Servir	7° a 9°C



Figura 23. Exemplos de fichas técnicas e de notas de prova

Uma vez selecionados os textos, foi criado um ficheiro Excel® no qual se procedeu à inserção de um conjunto de informação relevante sobre cada um deles. Para tal foram identificados oito campos: a codificação de cada nota de prova, a fonte (papel ou digital),

qual a entidade vinícola que apresentou a nota de prova, a que tipologia de vinho correspondia, o ano de colheita e a informação relativa à nota de prova, como se pode observar na Tabela 12, que corresponde a um exemplo de metadados para uma nota de prova do CEnoTEXT.

Campos	Informação
Código	V3944
Origem	Digital – <a href="http://www.essenciadovinho.com/pt/revista-wine/notas-de-prova">http://www.essenciadovinho.com/pt/revista-wine/notas-de-prova</a>
Entidade	Sociedade Agrícola D. Diniz
Denominação	Monte da Ravasqueira
Tipo	Tranquilo
Cor	Branco
Ano de colheita	2016
Nota de prova	Cor citrina. Elegante e convidativo no aroma, mostra a frescura da maçã verde, floral e algum vegetal. Bem agradável, tem firmeza, garra e acidez retemperadora que acalma alguma generosidade. Em crescendo.

Tabela 12. Exemplo de informação recolhida na ficha CEnoTEXT

A razão para a existência destes campos prende-se com a necessidade de garantir que as notas de prova recolhidas para o CEnoTEXT eram abrangentes relativamente a todos estes campos (para evitar ter notas de prova sobre apenas um tipo de vinho, ou um grande número de notas de uma mesma entidade, podendo significar que seria o mesmo especialista a escrever as notas de prova ou de vinhos da mesma data para evitar notas de um determinado ano), evitando enviesamentos em qualquer um dos sentidos e os dados deste *corpus* não serem diversificados.

O processo de seleção, apesar de moroso, revelou-se bastante profícuo, uma vez que se acedeu a um número bastante significativo de textos editados por entidades oficiais, responsáveis, etc. Foi também feita uma harmonização gráfica (visto que havia grafia pré-AO90 e pós-AO90).

Foram selecionadas as seguintes obras para o CEnoTEXT:

- Páginas Web de revistas de especialidade:
  - 356 notas de prova da *Revista Enoteca*;
  - 3 413 notas de prova online da *Revista Wine*;
  - 16 245 notas de prova da *Revista de Vinhos*.
  
- Páginas Web de produtores de vinho nacionais:
  - 1 481 notas de prova de diferentes produtores: *1912 Winemakers, Adega Cooperativa Borba, Adega Cooperativa De Favaios, Adega Cooperativa De Portalegre, Adega Cooperativa De Vila Real, Adega De Alcanhões, Adega De Almeirim, Adega Cooperativa De Mangualde, Adega Mãe, Adega Mayor, Adega De Ponte De Lima, A&D Wines, Agrovia, Ares Do Dão – Comissão Vitivinícola Regional Do Dão, Aveleda S.A., Beyra Vinhos De Altitude, Winemagnum – Carlos Lucas, Vinhos Lda., Casa Agrícola Roboredo Madeira, CARMIM, Casa Agrícola Alexandre Relvas, Casa Agrícola Das Mimosas, Casa Agrícola Santos Jorge, Casa Santa Vitória, Casa Santos Lima, Casal Branco, Caves Do Solar De São Domingos, Centro Agrícola Do Tramagal, Companhia Agrícola Sanguinhal, Companhia Das Lezírias, Companhia Das Quintas, Dão Sul, Encostas De Estremoz, Enoport, Esporão, E.R.T.A. Sociedade Agrícola, Lda., Falua, Fundação Eugénio De Almeida (F.E.A.), Fiúza & Bright, Granacer-Administração De Bens S.A. (Reguengos De Monsaraz), Herdade Grande Wines, Herdade Da Ajuda Nova, CASA AGRÍCOLA HMR, S.A., Casa Agrícola Horácio Dos Reis Simões, João Portugal Ramos Vinhos S.A., João C. Gouveia, José Carlos De Moraes C. Cruz & C<sup>a</sup> Lda., José Maria Da Fonseca Vinho S.A., Lavradores De Feitoria Vinhos De Quinta S.A., Luís Duarte Vinhos, Lda., Malo Tojo Estates Lda., Casal De Valle Pradinhos, Maria Lourdes Osório, O Abrigo Da Passarela, Parras Vinhos, Quinta Da Alorna, Quinta Do Encontro, Quinta Do Quetzal, Quinta Do Sobral, Quinta Da Murta,*

*Quinta Dos Murças, Quinta Dos Roques, Real Companhia Velha, Serrano Mira, Silveira E Outro, Sociedade Agrícola Arruda, Sociedade Agrícola Boas Quintas, Sociedade Agrícola De Santar, Sociedade Agrícola Encosta Do Guadiana, Sociedade Agrícola Gabriel Francisco Dias, Sociedade Agrícola Herdade Vale De Barqueiros, Sociedade Agrícola Monte Novo E Figueirinha, Sociedade Agrícola Quinta Da Lagoalva De Cima, Sociedade Agrícola Quinta Do Conde, Sociedade Quinta Do Portal, Sociedade Vinhos Borges, Sociedade Dos Vinhos Borges, Sogevinus, Sogrape, Symington Family Estates, Tapada Das Gaeiras, Terras De Alter, Terroir d'Origem, Tiago Cabaço, União Comercial Da Beira, VDS, Wiese & Krohn.*

Houve, também, a preocupação de verificar as notas de prova recolhidas quanto à tipologia (branco, rosado, tinto). Observou-se que 35% das notas são relativas a vinho branco, 5% relativas a vinho rosado e 60% relativas a vinho tinto. Apesar do aparente desequilíbrio, há uma relação equilibrada com os dados de produção em Portugal<sup>47</sup> disponibilizados pelo INE/PORDATA, que indicam que, para o ano de 2021, por exemplo, houve um total de produção de 7 358 539hL (vinho branco com 2 300 977 hL (31%) e vinho tinto e rosado 5 057 562 hL (69%)), que motivam a diferença percentual do tipo de vinho contemplado nas notas de prova.

No CEnoTEXT, existem notas de prova referentes a anos de colheita distintos (entre 1863 e 2020), de vinhos provenientes de diferentes regiões de Portugal e de diferentes entidades/marcas (2 447). Esta verificação permite garantir que as notas de prova são heterogéneas e não se direcionam a um determinado tipo de vinho, de uma determinada região ou de uma determinada gama de mercado.

Relativamente à nota de prova em si, observou-se que são pequenos textos. Em termos de dimensão, podia ter o tamanho mínimo de 5 palavras e corresponder a uma

---

<sup>47</sup> Disponível em <https://www.pordata.pt/Portugal/Produ%C3%A7%C3%A3o+de+vinho+por+cor-3364> consultado a 22/04/2022.

frase, até ao tamanho máximo registado de 158 palavras em 15 frases, como se pode observar nos exemplos [00048] e [14145] do CEnoTEXT, como se apresenta na Tabela 13.

Código	Nota de prova
[00048]	Cor citrina, frutos cítricos, seco.
[14145]	Completamente opaco no centro, com curtinha auréola violeta/azulada. Muita lágrima. Promete pela concentração que mostra. No nariz, o aroma é estranho e pouco característico! É de intensidade média. E mostra mamão e papaia, combinado com ligeiro balsâmico. Isto é um vinho do Dão??? Claro que sim! O que está aqui a falar é a agricultura biológica, o que só percebi depois de me ter sido desvendado o produtor! Um pioneiro e uma referência em Portugal neste tipo de agricultura. A boca é de acidez média a elevada, encorpada, com taninos proeminentes. A fruta está em média quantidade, mas continua a mostrar o mesmo que o nariz: mamão e papaia. Claro que também aparece o fruto vermelho silvestre e alguma baunilha, mas em segundo plano. O final é de média persistência, dominado pela força dos taninos, a qual deveria estar mais compensada pela fruta. É um vinho do Dão com personalidade própria, a qual pode ser até muito cativante.

Tabela 13. Exemplos de notas de prova encontrados no CEnoTEXT

A nota de prova tem normalmente a seguinte configuração: começa com a descrição visual, segue-se a descrição olfativa e termina com a descrição gustativa (que envolve a descrição tátil e a descrição do conjunto).



Na análise do CEnoTEXT, verificou-se alguma variação desta configuração. Algumas notas de prova têm a indicação das três descrições, umas mais sucintas e segmentadas e outras mais detalhadas. Como se pode observar através dos exemplos seguintes: (i) há alguns casos em que a descrição da cor é mencionada com detalhe relativo à cor/tonalidade (exemplo 1); (ii) outros casos onde é mencionada apenas a casta

(exemplos 2 e 3); (iii) casos onde é mencionada a tipologia – branco, rosé, tinto – mas sem qualquer menção a outros aspetos visuais (exemplo 4); e outros casos sem especificação em concreto da tonalidade em causa (exemplos 5 e 6).

1. [00013] Cor: Rubi luminoso. Nariz: Aroma a frutos silvestres e com notas de especiarias. Boca: Com volume mas macio, fresco e equilibrado, com fim de boca frutado.
2. [08189] 100% Maria Gomes (Fernão Pires), nota-se o perfil mais floral da casta no aroma elegante, a par de um levíssimo fumado. A boca é surpreendente, na intensidade, leveza e frescura quase crocante, resultando num vinho vibrante e longo, com grande potencial de guarda.
3. [12278] A cor intensa da casta está bem patente neste Vinhão de perfil clássico, cheio e vigoroso, com algum gás a realçar os taninos fortes e a acidez pronunciada. Bom parceiro de pratos gordos de inverno.
4. [15878] Um tinto feito para a mesa, sem excessos de fruto, bons taninos e frescura de conjunto, franco, harmonioso e descomplicado.
5. [17138] De cor aberta, tem bons aromas de compota de ameixa, uva passa, figos secos, num estilo muito "tawny". A boca corresponde a este perfil, suave e aveludada, apontamentos de cereja e cacau no final agradável.
6. [19011] A cor não é muito concentrada, mas tem uma bela profundidade e complexidade aromática, com notas de ameixa e vegetal secos, tabaco, especiarias, tudo bem composto. Bom volume de boca, elegante mas sem perder o carácter alentejano que caracteriza a marca, com uma presença pouco habitual de taninos musculados no final, que apontam para uma evolução segura na garrafa. 165.000 garrafas com esta qualidade, é obra.

Outros exemplos de notas de prova apresentam uma descrição incompleta, no sentido em que não fazem menção à descrição visual, mas apenas às descrições olfativa e gustativa, como se pode verificar nos exemplos seguintes.

7. [16425] Notas fumadas e vidradas num fundo balsâmico anisado com fruto fresco e vibrante, textura suave, especiado, fresco e longo, sem ser vigoroso está muito bem proporcionado.
8. [16290] Muitos fumados a pólvora e fósforo, leve redução, fruto discreto, na boca mostra boa frescura alguma mineralidade, fruto cítrico, ananás verde, um branco curioso.
9. [16524] Simples mas muito limpo com fruto leve e agradável na prova de boca, suave, fresco e com alguma persistência final.

Esta observação inicial permite considerar que, tal como indicado acima, em termos gerais, a nota de prova pode ter a seguinte configuração, mas com as características que a seguir se discriminam:

NOTA DE PROVA  
[[[visual]] + [olfativa] + [gustativa]]

A secção correspondente a [visual] pode ser opcional. Algumas notas de prova apresentam indicação explícita sobre a cor e a tonalidade (*Este é um tinto cheio de cor, rubi intenso e brilhante, com tons mais abertos na periferia do copo.*); outras apresentam indicação explícita relativa à presença de gás carbónico (*Cor dourada. Bolha muito fina, em cordão médio*) e fenómenos de capilaridade (*Amarelo palha de média profundidade, com ligeiríssimos toques dourados. Lágrima presente, mas não muito densa.*); outras, ainda, apresentam informação implícita, quando mencionam apenas a(s) casta(s) e se consegue depreender a cor, mas não a tonalidade (*Com Aragonéz e Alfrocheiro, está muito carregado na cor. (...)*). Note-se que, das 21 495 notas de prova que constituem o CEnoTEXT, verificou-se que 5 950 não apresentavam qualquer menção à descrição referente à modalidade visual (omissão total da descrição), 1 941 apresentavam apenas menção relativa à casta sem qualquer outra referência explícita à cor ou tonalidade e 16 878 notas indicam apenas o descritor relativo à tipologia (*branco, rosé, tinto*), sem explicitar outra informação relativa à cor/tonalidade, para dar conta da modalidade visual.

A secção correspondente à descrição olfativa tem no CEnoTEXT um comportamento quase de presença obrigatória (apenas 55 das notas extraídas não apresentam de forma evidente a informação sobre a modalidade olfativa) e pode ter a seguinte configuração, com indicação explícita sobre diversos parâmetros:

- o aroma primário (*Aromas claramente primários, imediatos, com os morangos e framboesa em destaque, tem um muito subtil floral que lhe assenta bem, mas sem sair de alguma simplicidade.*);

- o aroma secundário (*Bical de vinhas velhas, fermentado em barrica.*);
- o aroma terciário (*Aroma já de bouquet com terciários típicos de evolução, notas licoradas, café*);
- e ainda a apresentação dos aromas por famílias ou por substância (*Aroma a rebuçado de groselha e framboesa, (...)*).

Pode ser apresentada a informação de forma implícita, quando é mencionada a relação com uma casta, que poderá estar mencionada noutra parte que não na nota de prova (*Aroma à casta bem definido.*), ou a menção explícita da casta e qual a sua contribuição para a modalidade olfativa (*Lote de muitas castas, nota-se no aroma uma forte componente floral que tanto pode advir da Touriga Nacional como do Petit Verdot e que dá um tom fresco e vivo ao vinho, afastando-o dos aromas característicos da planície.*). Também há casos em que se funde a informação olfativa e gustativa (*aromas e sabores requintados, laranja cristalizada, madeira, mel, frutos secos, especiarias, elegantíssima, interminável*).

Enquanto a secção da modalidade visual pode ser opcional, a modalidade olfativa pode estar presente de forma independente ou de forma associada à forma gustativa.

A secção correspondente à modalidade gustativa é obrigatória e pode ter diferentes configurações, consoante o tipo de parâmetros descritos de forma explícita, relativamente a:

- sensações olfativas, através da via retronasal (*As notas retronasais confirmam na íntegra as primeiras impressões olfativas.*);
- sensações gustativas, relativas aos sabores básicos (*Na prova tem bom comportamento, com acidez e o perfil amargo.*);
- sensações gustativas, relativas aos sabores complementares (*Ataque desordenado, metálico e amargo, caindo abruptamente ao longo da prova.*);

- sensações químicas, como reação das mucosas relativamente à adstringência ou aspereza, a causticidade, a untuosidade e a efervescência (*Boca com entrada muito equilibrada, evoluindo depois em alguma adstringência, sinal de juventude.*);
- sensações táteis, que permitem sentir a consistência, a liquidez, a fluidez a espessura, o volume, a intensidade taninosa e a opulência (*Cheio e robusto na boca, com acidez integrada e taninos firmes, final sereno e opulento.*);
- sensações térmicas, como o refrescante, o ardente, o frio, o quente e o teor alcoólico (Corpo médio, acidez suficiente, um tom morno e alcoólico marca a prova, final texturado e salino.);
- comentário relativo à sensação final que o vinho poderá deixar no palato (Final muito apelativo.).

Além desta informação, encontra-se ainda, em algumas notas de prova uma secção complementar relativa a aspetos que o especialista considera relevantes, como quando deve ser degustado o vinho, se pode ser imediatamente consumido ou se é preferível aguardar algum tempo (*Creoso, gordo, forte, ainda a precisar de tempo em cave.*), a menção relativa a algum defeito (*A boca é extremamente simples, é um licoroso sem defeitos mas sem grande história para contar.*), comentários do especialista (*é um tinto apenas correto, que precisa de afinamento em próximas edições.*), opiniões pessoais (*Imaginem um grande Vintage novo – este é ainda melhor!*) e sugestões de harmonização<sup>48</sup> (*Acidez pronunciada dá muita frescura mas também um perfil algo “duro” que exige um prato gordo a acompanhar.*).

---

<sup>48</sup> Entende-se por harmonização o ato de combinar um alimento ou prato com um vinho. O ato dessa combinatória deve ser entendido como mais do que a junção dos dois, mas como a busca pelo equilíbrio entre ambos, onde nenhum sobressaia, por forma a alcançar uma experiência que agrade aos sentidos. Apesar da harmonização assentar em questões físico-químicas dos sabores, a experiência sensorial é pessoal e subjetiva e deve ser vista como um guia, dado que existem diferentes formas de harmonização: por similaridade/afinidade (onde as características do vinho e do prato são semelhantes), por contraste (com comida gordurosa contrapor um vinho de grande frescor); e por região (com produtos da mesma região, por exemplo o vinho ser usado para acompanhar o prato que usou o mesmo vinho na confeção).

Apesar desta configuração geral, muitas vezes os especialistas alteram a ordem de apresentação das modalidades ou a ordem da informação. No entanto, pode referir-se que os casos em que a ordem apresentada não ocorre são pouco relevantes.

#### 4.2.3.1.1. CENOTEXT EM NÚMEROS

Para o CEnoTEXT, foram recolhidas 21 495 notas de prova, com 780 585 *tokens* (7 015 *types*, dos quais 2 301 ocorrem apenas uma vez (*hápax*)), apresentando-se na Tabela 14 as 20 unidades lexicais mais frequentes (incluindo palavras funcionais).

Posição	Forma	Freq.	Posição	Forma	Freq.
1	e	53 501	11	fruto	10 416
2	de	34 077	12	final	9 206
3	com	30 107	13	bem	9 148
4	muito	16 863	14	fruta	9 030
5	a	16 562	15	acidez	8 839
6	boca	14 644	16	tanino	7 951
7	na	14 170	17	no	7 433
8	um	13 307	18	o	7 379
9	nota	11 535	19	boa	7 141
10	aroma	10 762	20	ser	7 094

Tabela 14. As 20 unidades lexicais mais frequentes do CEnoTEXT

Ainda relativamente às unidades lexicais verificou-se a sua distribuição consoante o número de ocorrências e a sua percentagem como se pode observar na Tabela 15.

Ocorrências	N.º de UL	%	Ocorrências	N.º de UL	%
1	2 308	32,9	100	236	3,4
2 a 9	2 473	35,3	200	119	1,7
10s	603	8,6	300	50	0,7
20s	304	4,3	400	33	0,5

30s	168	2,4	500	30	0,4
40s	133	1,9	600	27	0,4
50s	83	1,2	700	25	0,4
60s	75	1,1	800	11	0,2
70s	83	1,2	900	17	0,2
80s	53	0,8	1 000-53 501	140	2,0
90	44	0,6	<b>TOTAL</b>	7 015	100%

Tabela 15. A distribuição das unidades lexicais consoante o número de ocorrências no CEnoTEXT

Em síntese, o CEnoLEX é um *corpus* textual monolíngue do PE, constituído por 21 495 textos (com 780 585 *tokens* (7 015 *types*, dos quais 2 301 *hápax*)) de notas de prova recolhidos entre 2012-2021. A restante análise do CEnoTEXT, que tem por objetivo a extração de unidades lexicais da especialidade, será apresentada de seguida no Capítulo 5.

## 5. LÉXICO DE ESPECIALIDADE DA ASE: ANÁLISE E TRATAMENTO

Neste capítulo, será apresentada a análise feita ao CEnoTEXT para extração e tratamento do léxico de especialidade da ASE. Posteriormente, será feita uma comparação com os dados recolhidos no CEnoLEX e uma sugestão de organização das Unidades Lexicais de Especialidade (ULE<sup>49</sup>) por modalidade sensorial, o que contribuirá para a elaboração da proposta de modelo de dicionário – o DEASE.

### 5.1. SELEÇÃO E TRATAMENTO DE UNIDADES LEXICAIS DE ESPECIALIDADE DA ASE

Através da utilização do Sketch Engine, foi extraída uma listagem de unidades (lemas) por categoria gramatical, de acordo com a anotação do sistema: adjetivo, advérbio, nome, verbo, artigo, conjunção, numeral, preposição, pronome. As listas apresentadas foram revistas manualmente e foram corrigidas as etiquetas para a categoria gramatical adequada, quando relevante e consoante as ocorrências no CEnoTEXT.

A razão para esta revisão inicial prende-se com o facto de sistemas como o Sketch Engine etiquetarem automaticamente a classe de palavras, o que implica necessariamente uma taxa de erro e resulta em erros específicos em certas unidades, como, por exemplo, “frutado”<sup>50</sup>.

No sistema, ocorrências de “frutado, frutada, frutados, frutadas” e algumas ocorrências de “fruta, frutas” surgem também etiquetadas com o lema do verbo “frutar”, com 2 445 ocorrências, e não como adjetivo ou como nome feminino. No CEnoTEXT, após revisão manual, verificou-se que não ocorre uma única forma do verbo “frutar”, mas

---

<sup>49</sup> Entende-se ULE como “uma unidade lexical, atómica ou multipalavra, com propriedades sintáticas, combinatórias e de significado específicas, usadas por especialistas de um dado domínio, tipicamente em situação de comunicação entre especialistas desse domínio, que podem ser caracterizadas por critérios estatístico-linguísticos de *unithood* e *termhood*. Essa definição motiva, assim, a relevância da constituição e utilização de *corpora*, bem como a importância e relevância do recurso a ferramentas de exploração de *corpora*, com funcionalidades estatísticas e linguísticas.” (Barbero & Amaro, 2020:73).

<sup>50</sup> São vários os trabalhos que identificam e referem erros da etiquetagem automática do Sketch Engine, relativamente a *corpora* de português; no entanto, não é indicada uma percentagem de erro/acerto (cf. Kuhn (2017), Kilgarriff *et al.* (2014a), Kilgarriff *et al.* (2014b)).

apenas ocorrências de adjetivo e de nome. Tendo sido detetada esta situação atempadamente, decidiu verificar-se a listagem inicial e proceder a revisão manual por forma a reduzir a percentagem de erro de etiquetagem gramatical e lematização feitas automaticamente.

Neste processo, também foram detetadas gralhas e erros ortográficos, consideradas como formas possíveis pelo sistema e posteriormente desconsideradas manualmente, por não haver modo de recuperar uma forma válida<sup>51</sup>.

A título de exemplo, numa listagem inicial do Sketch Engine relativa à categoria verbal, eram apresentadas 1 464 formas (1 464 itens em 83 706 de frequência total). Após a revisão manual, foram mantidos 884 verbos (37 968 de frequência total), como se apresenta na Tabela 16.

Categorial gramatical <sup>52</sup>	Sketch Engine	Revisão
verbos	1 464 itens   83 706 frequência total	884
advérbios	320 itens   45 033 frequência total	260
nomes	3 258 itens   241 671 frequência total	2 919
adjetivos	1 502 itens   164 979 frequência total	2 481

Tabela 16. Valores numerais após revisão manual

Posteriormente, para sustentar a decisão de analisar uma amostra dos dados obtidos, verificou-se a distribuição consoante a frequência de ocorrências, considerando cinco grupos (1: com mais de 100 ocorrências; 2: de 99 a 50 ocorrências; 3: de 49 a 10

---

<sup>51</sup> Em determinados casos, algumas variantes ortográficas foram mantidas, mas com a indicação de variante ortográfica preferida. Por exemplo, no empréstimo “*after eight*”, indicou-se a forma preferida “*after-eight*” como variante gráfica), pois considerou-se que seria importante registar as hesitações dos especialistas quanto à grafia de determinados descritores para análise final.

<sup>52</sup> Foram desconsideradas as categorias funcionais nesta fase.

ocorrências; 4: de 9 a 2 ocorrências; 5: com apenas 1 ocorrência), como se apresenta na Tabela 17.

Categoria gramatical	Total de lemas	Ocorrências				
		(+) de 100	99-50	49-10	9-2	1
<b>verbos</b>	884	6 %	4 %	19 %	37 %	33 %
<b>advérbios</b>	260	5 %	4 %	17 %	41 %	33 %
<b>nomes</b>	2 919	11 %	5 %	18 %	35 %	31 %
<b>adjetivos</b>	2 481	11 %	4 %	19 %	37 %	30 %

Tabela 17. Valores percentuais da distribuição dos lemas

Nas próximas subsecções, apresenta-se a análise dos dados, agrupados pelas categorias sintáticas relevantes no CEnoTEXT. Pretende-se, também, apresentar em alguns casos os candidatos a ULE, tendo em conta o seu comportamento na língua geral vs. língua de especialidade, por forma a verificar quais podem ser consideradas ULE e, por isso, integradas na proposta de dicionário.

Para tal, ao longo da análise, será feita uma comparação sistemática entre as unidades lexicais que constam no CEnoTEXT e no CORLEX<sup>53</sup> (parte do *Corpus* de Referência do Português Contemporâneo, Mendes *et al.*, 2012). A razão para efetuar esta comparação de dados entre um *corpus* de língua de especialidade e um *corpus* de língua geral prende-se com a necessidade de compreender melhor os dados relativos às frequências de unidades lexicais que podem ser candidatas a ULE do DEASE e aquelas que são apenas unidade lexicais de uso comum, que serão potencialmente excluídas. Seguiu-se uma adaptação da proposta de Barbero e Amaro (2020:81-82) para a extração de dados de frequência com monitorização.

---

<sup>53</sup> Para mais informação consultar <https://clul.ulisboa.pt/projeto/lexico-multifuncional-computorizado-do-portugues-contemporaneo>.

Tendo em conta que os dois *corpora* (CEnoTEXT e CORLEX) são numericamente distintos (em número de *tokens* e frequências absolutas), optou-se por torná-los comparáveis através da aplicação de uma frequência relativa normalizada por 100<sup>54</sup>.

De seguida, observou-se as unidades em ambos os *corpora* e verificou-se se o seu comportamento seria de candidato a ULE ou a excluir. Foram considerados três critérios gerais. Se uma unidade lexical:

- (i) ocorrer com a mesma frequência relativa igual ou semelhante em ambos os *corpora*, é excluída como candidata a unidade;
- (ii) ocorrer com frequência relativa inferior no CORLEX do que no CEnoTEXT, é considerada como candidata;
- (iii) não ocorrer no CORLEX, então é imediatamente considerada candidata.

Para executar esta tarefa, foi necessário observar o CORLEX (após importação para o Sketch Engine), extrair a listagem individual (por categoria gramatical – adjetivos, nomes, verbos), aplicar a fórmula de frequência relativa, comparar com a listagem individual extraída do CEnoTEXT e observar os dados, extraídos para o formato Excel. Esta observação permitiu uma revisão e comparação manual, para além da verificação do diferencial<sup>55</sup> de percentagem entre ambos os *corpora*, como se apresenta na Tabela 18.

---

<sup>54</sup>  $\frac{\text{frequência absoluta da ocorrência absoluta} \times 100}{\text{corpus total}} = \text{frequência relativa}$

<sup>55</sup> O diferencial de frequência entre as frequências relativas obtidas do CEnoTEXT e do CORLEX permite verificar se uma determinada unidade lexical ocorre com elevada frequência no CEnoTEXT e não ocorre, ou ocorre com a mesma ou menor frequência no CORLEX e se pode ser considerado candidato a ULE.

CEnoTEXT		LEMA	CORLEX		DIF. %
FREQ. ABS	FREQ. REL		FREQ. ABS	FREQ. REL	
14 644	1,8760	boca	1 602	0,0242%	1,8618%
11 535	1,4777	nota	656	0,0099%	1,4757%
10 762	1,3787	aroma	186	0,0028%	1,3833%
8 839	1,1324	acidez	11	0,0001%	1,1388%
7 951	1,0186	tanino	8	0,0001%	1,0244%
7 094	0,9088	ser	13 936	0,2108%	0,7032%
3 516	0,4504	vinho	1 044	0,0157%	0,4373%
1 180	0,1512	vivo	1 542	0,0233%	0,1287%
559	0,0716	sentir	4 780	0,0723%	-0,0003%
65	0,0083	dizer	19 096	0,2889%	-0,2806%
55	0,0070	perfume	575	0,0086%	-0,0016%
14	0,0018	pinho	73	0,0011%	0,0007%
11	0,0014	preferir	741	0,0121%	-0,0107%
2	0,0003	vulgar	363	0,0054%	-0,0052%

Tabela 18. Excerto da listagem comparativa entre CE<sub>no</sub>TEXT e CORLEX

Na Tabela 18, é possível analisar os valores relativos ao diferencial de frequência (em valores percentuais) nos dois *corpora* que permitem categorizar os candidatos em diferentes grupos para verificar se se enquadram no DEASE. Para tal, foram considerados, de novo, três critérios gerais:

- se a frequência relativa no CE<sub>no</sub>TEXT é maior do que a frequência relativa no CORLEX, a unidade é considerada candidata, tal como sucede com “boca, nota, aroma, acidez, tanino, vinho, ser, vivo”;
- se a frequência relativa no CE<sub>no</sub>TEXT é menor do que a frequência relativa no CORLEX, a unidade não é considerada um candidato, tal como sucede com “dizer, preferir, vulgar”;
- se o diferencial entre as frequências relativas de ambos os *corpora* é muito ténue (menos de 0,0020%), a unidade não é considerada um candidato, tal como sucede com “sentir, perfume, pinho”, mas as ocorrências são analisadas caso a caso.

Estes critérios, para além de permitirem o contraste entre um *corpus* de especialidade e um *corpus* de língua geral, foram estabelecidos para permitir um outro conjunto de análises. Ou seja, a metodologia usada assenta, em primeiro lugar, no aspeto quantitativo para a seleção de candidatos, mas admite também o aspeto qualitativo para resgatar candidatos interessantes que, neste *corpus*, tenham uma frequência menor.

Deste modo, os dados recolhidos no CEnoTEXT, que são candidatos a ULE serão alvo do seguinte procedimento metodológico:

- (i) na análise categorial – observar somente os verbos, os nomes e os adjetivos com mais de 100 ocorrências no CEnoTEXT;
- (ii) na análise semântica – contrastar os dados da frequência relativa CEnoTEXT e CORLEX, em particular, se o valor do diferencial é ou não relevante para justificar a inserção dos candidatos a ULE; verificar as polissemias dos candidatos em dicionários de língua geral e de língua de especialidade;
- (iii) na análise contrastiva – observar os dados recolhidos no CEnoTEXT e no CEnoLEX e analisar o que é mencionado pelos especialistas e o que é atestado nas obras recolhidas;
- (iv) na análise sensorial – verificar a distribuição das ULE por modalidade sensorial (visual, olfativa e gustativa) para ver quais são as ULE tipicamente de uma modalidade e aquelas que são potencialmente multissensoriais.

A metodologia seguida no presente trabalho respeitará, em termos gerais, a estrutura apresentada. No entanto, aquando da análise de cada categoria, sempre que se justifique, proceder-se-á a alguns ajustes.

Os comportamentos das unidades lexicais em ambos os *corpora* serão tratados por categoria gramatical, como se apresentará nas secções seguintes.

## 5.2. ANÁLISE CATEGORIAL

### 5.2.1. VERBOS

Na análise de léxico de um domínio de especialidade há, regra geral, a tendência para dar destaque às categorias nominais e adjetivais. Neste trabalho, devido ao facto de ser um domínio pouco explorado no PE, considerou-se analisar, também, os verbos, à semelhança do que já foi feito para o PB, o castelhano e o inglês (Abreu & Lara, (2013); San Roque *et al.*, (2018); Caballero, (2007).

Nesta secção, serão apresentados os verbos encontrados no CEnoTEXT, tipicamente nas notas de prova, e que ajudam a descrever o vinho aquando da degustação. Pretende-se observar como são comunicadas as experiências sensoriais dos vinhos através dos verbos usados e que aspetos sensoriais estes descrevem, como já feito para outras línguas (Caballero 2007; Lehrer 2009).

A primeira observação é que, no geral, quando o especialista descreve o vinho, utiliza poucos ou nenhuns verbos, enumerando apenas um conjunto de atributos e características através de nomes e adjetivos, como se pode observar na seguinte nota de prova:

10. [00007] Nariz fresco e exuberante com evidentes notas de limão, lima, pêssego, frutos tropicais e alguns florais, tudo num fundo com evidente mineralidade. Corpo muito cremoso e balanceado. Final longo e persistente.

Estudos na área da psicologia indicam que o ser humano, para descrever algo, nem sempre necessita de recorrer ao uso de verbos; pode simplesmente elencar adjetivos e nomes para ativar a perceção cognitiva daquilo que pretende descrever (Paradis & Hommerberg, 2016:183). O facto de a nota de prova poder ser escrita sem recorrer à utilização de verbos não é uma novidade, nem um problema. No entanto, existem,

evidentemente, exemplos de notas de prova com presença de verbos, ainda que com alguma flutuação quanto ao modo – impessoal/pessoal, tempo – presente/passado/futuro, e pessoa, aspetos, de resto, verificados também para o francês (cf. Gautier, 2014). Vejam-se os seguintes exemplos:

11. [14892] O vinho **apresenta-se** amarelo brilhante, de cor clara com laivos ligeiramente esverdeados. O nariz **tem** notas de maracujá e baunilha, ligeiro toque floral e alguma doçura. Na boca, **sente-se** estruturado, com notas tostadas a par do seu lado mais frutado e fresco. **É** um vinho agradável, sem peso excessivo, com o álcool e a acidez elevados. **É** simples e honesto, com um final ácido e alegre, povoado de citrinos.
12. [01547] Muito aberto na cor, aroma a **desprender** algumas notas de groselha e morango, apontamentos vegetais. Muito ligeiro na boca, a **pedir** pratos muito leves, **podendo** mesmo **substituir** o branco como vinho do quotidiano.
13. [08464] Bom aroma, fruta densa e ligeiras notas licoradas, a **mostrar** que **está** a **evoluir** bem e a **chegar** ao topo. O ligeiro floral **combina-se** com a barrica e **está** todo bem equilibrado. Alguns taninos secos **dizem** que **precisa** de comida por perto que o **fará mostrar-se** bem.

Assim, como referido anteriormente, no CEnoTEXT detetou-se a presença de 884 verbos com um total de 37 867 ocorrências (num universo de 780 585, correspondendo a 4% do total das formas), que se distribuem do seguinte modo: 55 verbos com mais de 100 ocorrências; 204 com 10 a 99 ocorrências; 333 com 2 a 9 ocorrências e 292 verbos com apenas uma ocorrência. Na Tabela 19, apresentam-se os 20 verbos mais frequentes.

POSIÇÃO	VERBO	FREQ. ABS.	POSIÇÃO	VERBO	FREQ. ABS.
1	ser	7 094	11	manter	559
2	ter	3 255	12	sentir	559
3	mostrar	2 920	13	pedir	538
4	dar	2 102	14	conferir	440
5	terminar	1 258	15	resultar	406
6	revelar	1 115	16	lembrar	370
7	poder	668	17	ir	306
8	apresentar	648	18	crescer	305
9	beber	624	19	haver	279
10	surgir	621	20	marcar	254

Tabela 19. Listagem dos 20 verbos (lemas) mais frequentes de CEnoTEXT

Observou-se também quais as formas verbais mais frequentes, como se apresenta na Tabela 20.

POSIÇÃO	VERBO	FREQ. ABS.	POSIÇÃO	VERBO	FREQ. ABS.
1	é	4 722	11	dá	442
2	tem	2 953	12	mostrar	397
3	mostra	1 563	13	pedir	393
4	está	1 403	14	apresenta	386
5	dar	1 273	15	pode	327
6	termina	1 120	16	mantém	327
7	ser	622	17	sente-se	316
8	revela	597	18	surge	277
9	mostra-se	555	19	resulta	266
10	beber	480	20	há	257

Tabela 20. Listagem das 20 formas verbais mais frequentes

Tendo em conta que o CEnoTEXT é um *corpus* de notas de prova de vinhos com descrições de um conjunto de características de vários vinhos, é expectável que os verbos utilizados sejam também eles candidatos a ULE da ASE e alguns sejam considerados verbos sensoriais ou verbos com polissemias relacionadas para a modalidade sensorial.

Existem várias formas de utilizar verbos para expressar a modalidade sensorial. As estratégias comuns são as que se apresentam de seguida, com exemplos também observados no CEnoTEXT:

- utilização de um verbo sensorial (aquele cujo significado primeiro é associado a alguma modalidade sensorial, como por exemplo “cheirar, ver, saborear”);

14. [08464] Bom aroma, fruta densa e ligeiras notas licoradas, a **mostrar** que **está** a **evoluir** bem e a **chegar** ao topo. O ligeiro floral **combina-se** com a barrica e **está** todo bem equilibrado. Alguns taninos secos **dizem** que **precisa** de comida por perto que o **fará mostrar-se** bem.

15. [09224] **Cheira** a violetas e lavanda, com pequenas bagas silvestres, ervas aromáticas, tudo muito fresco.

- utilização de um verbo não sensorial (que no domínio da ASE ganha uma polissemia específica);

16. [05949] Boca vibrante, perfeita acidez, incisivo, crocante, longo. Vai **crescer** na garrafa.

- utilização de um verbo copulativo (verbos *ser* e *estar*) numa construção predicativa, com o sentido de enumerar ou para avaliar;

17. [13709] A acidez é muito boa, o vinho é gordo, tem corpo e um grande sentido gastronómico.

18. [13102] Na boca **está** definido, com acidez fresca, doçura frutada, um ligeiro herbáceo que lhe aumenta o volume e prolonga o final.

- utilização do verbo *ter* numa construção predicativa (*ter/apresentar característica*) (verbo *ter*).

19. [16556] Um aroma algo artificial, fumado meio aborrachado, com uma prova doce e um final diluído. O vinho está bem, é consensual, mas **tem** um estilo muito trabalhado.

Uma tendência comum nas línguas é a extensão semântica de algumas formas que podem passar a descrever outros sentidos. Em várias línguas, alguns verbos perceptivos alargaram o seu significado para abranger a experiência cognitiva. Por exemplo, verbos que incorporaram o significado de *intuição*, *sentimento*, *apreensão* como no inglês (“The police have been **sniffing** around here again”, Sweetser 1990:28), no castelhano (“Si le he **oido** bien, usted quiere decir que no hay solucion”, Ibarretxe-Antuñano, 2002), no basco (“Le han ordenado que **husmee** las cuentas”, Ibarretxe-Antuñano, 2002) e como em português europeu no exemplo “Não **vejo** problemas”.

No sentido de compreender a possível extensão semântica de alguns dos verbos detetados como candidatos a ULE da ASE, considerou-se aplicar a metodologia de tratamento de verbos que se apresenta de seguida.

Para efeitos de delimitação do objeto de estudo, optou-se por considerar apenas os verbos com mais de 100 ocorrências no CEnoTEXT. Estes correspondem aos seguintes 55 verbos: “abrir, acompanhar, aparecer, apresentar, assentar, beber, combinar, compensar, conferir, confirmar, conjugar, contribuir, crescer, dar, deixar, destacar, dominar, envolver, esconder, estar, evidenciar, evoluir, faltar, fazer, ficar, funcionar, ganhar, gostar, haver, impressionar, ir, juntar, lembrar, manter, marcar, melhorar, misturar, mostrar, notar, parecer, pedir, perder, poder, prolongar, resultar, revelar, sentir, ser, sobressair, sugerir, surgir, ter, terminar, tornar, trazer”.

A partir desta listagem, pretende-se verificar:

- (i) se existem ULE verbais e quais são;
- (ii) se existem verbos sensoriais ou verbos não sensoriais, mas com uma polissemia específica do domínio da ASE.

Esta verificação, para além de permitir distinguir as unidades lexicais de língua geral das potenciais unidades lexicais de especialidade, contribui também para apurar se um não-especialista acederia facilmente a um significado sensorial de especialidade, ou não, apresentando evidências a ter em conta aquando da elaboração dos textos definitórios (que se apresenta no capítulo 5).

Seguindo a metodologia geral estabelecida, começou-se por contrastar as frequências dos candidatos no CEnoTEXT e num *corpus* de língua geral, o CORLEX, conforme se apresenta na Tabela 21, e por observar, em particular, se o valor do diferencial será ou não relevante como critério de seleção dos candidatos a ULE.

N	CEnoTEXT	VERBO	CORLEX	DIF.	N	CEnoTEXT	VERBO	CORLEX	DIF.
	FREQ. REL.		FREQ. REL.			FREQ. REL.			
1	0,914	ser	0,2108	0,7032	28	0,0323	acompanhar	0,0219	0,0104
2	0,3762	mostrar	0,0318	0,3444	29	0,0171	destacar	0,0067	0,0104
3	0,162	terminar	0,0108	0,1512	30	0,0176	sugerir	0,0084	0,0092
4	0,1436	revelar	0,0144	0,1292	31	0,0136	assentar	0,0055	0,0081
5	0,2708	dar	0,2046	0,0662	32	0,0144	misturar	0,0074	0,007
6	0,0804	beber	0,0154	0,065	33	0,015	contribuir	0,0095	0,0055
7	0,08	surgir	0,023	0,057	34	0,0149	notar	0,0096	0,0045
8	0,0566	conferir	0,0041	0,0525	35	0,0139	funcionar	0,0095	0,0044
9	0,0834	apresentar	0,0398	0,0436	36	0,0184	juntar	0,0143	0,0041
10	0,072	manter	0,0312	0,0408	37	0,0141	esconder	0,0141	0
11	0,0523	resultar	0,0125	0,0398	38	0,072	sentir	0,0723	-0,0003
12	0,0693	pedir	0,0349	0,0344	39	0,015	faltar	0,0184	-0,0034
13	0,0304	sobressair	0,0011	0,0293	40	0,0135	ganhar	0,0211	-0,0076
14	0,0277	conjuguar	0,0019	0,0258	41	0,0213	aparecer	0,0355	-0,0142
15	0,0279	evoluir	0,0036	0,0243	42	0,0247	gostar	0,0397	-0,0150
16	0,0393	crescer	0,0165	0,0228	43	0,0132	trazer	0,0348	-0,0216
17	0,0278	confirmar	0,0061	0,0217	44	0,0273	perder	0,0506	-0,0233
18	0,0282	dominar	0,01	0,0182	45	0,014	abrir	0,0411	-0,0271
19	0,0208	evidenciar	0,0026	0,0182	46	0,0306	tornar	0,0586	-0,0280
20	0,0207	impressionar	0,003	0,0177	47	0,0228	deixar	0,1022	-0,0794
21	0,0329	marcar	0,0168	0,0161	48	0,0176	ficar	0,133	-0,1154
22	0,0203	melhorar	0,0054	0,0149	49	0,0359	haver	0,2697	-0,2338
23	0,0476	lembrar	0,0345	0,0131	50	0,086	poder	0,3361	-0,2501
24	0,026	envolver	0,0137	0,0123	51	0,0282	fazer	0,3254	-0,2972
25	0,0188	prolongar	0,0065	0,0123	52	0,4194	ter	0,7172	-0,2978
26	0,0173	combinar	0,0052	0,0121	53	0,0315	estar	0,4007	-0,3692
27	0,0139	compensar	0,0031	0,0108	54	0,0394	ir	0,5911	-0,5517

Tabela 21. Lista de verbos com as freq. relativas e o valor diferencial de frequência entre CEnoTEXT e CORLEX

Observa-se que os candidatos verbais apresentam a seguinte distribuição:

- Grupo com valores de frequência relativa mais elevados no CEnoTEXT do que CORLEX, como “ser, mostrar, terminar, revelar, dar, beber, surgir, conferir, apresentar, manter, resultar, pedir, sobressair, conjugar, evoluir, crescer, confirmar, dominar, evidenciar, impressionar, marcar, melhorar, lembrar, envolver, prolongar, combinar, compensar, acompanhar, destacar, sugerir, assentar, misturar, contribuir, notar, funcionar, juntar” (a verde na tabela);
- Grupo com valores semelhantes em ambos os *corpora*, como “esconder, sentir”;
- Grupo com valores mais baixos no CEnoTEXT do que no CORLEX, como “faltar, ganhar, aparecer, gostar, trazer, perder, abrir, tornar, deixar, ficar, haver, poder, fazer, ter, estar, ir”.

Esta divisão inicial, tendo em conta os critérios<sup>56</sup> apresentados anteriormente (ver Tabela 18), permite tomar a decisão de considerar o primeiro grupo de verbos como candidatos inequívocos a ULE.

No entanto, observando os candidatos “rejeitados” com mais de 100 ocorrências, ocorre considerar que estes não devem ser liminarmente rejeitados, não só pela alta frequência no CEnoTEXT, mas também pela relação diferencial entre os valores de CEnoTEXT e do CORLEX. Isto é, apesar de tudo parecer indicar que estes elementos seriam candidatos a exclusão, considera-se que se deve olhar para eles mais uma vez. A justificação desta opção é o facto de estes 16 candidatos serem muito frequentes na língua corrente, o que, considerando a ASE e o curto texto da nota de prova, pode, em certas circunstâncias, levar a equívocos de interpretação por um não-especialista que faça a leitura de uma nota de prova.

---

<sup>56</sup> Esta é uma (das várias) formas de seleccionar candidatos quando perante um grande número de candidatos. Esta seleção permite maximizar a estratégia (frequência absoluta; frequência relativa comparada com um corpus de língua geral) de escolha, permitindo, posteriormente, e em casos excepcionais, observar os contextos (as concordâncias e coocorrências) de outros candidatos, permitindo avaliar se é ou não um potencial candidato a ULE.

Assim, optou-se por olhar para os mesmos dados (os 55 verbos com mais de 100 ocorrências, anteriormente apresentados) de uma outra perspectiva, para apurar se há verbos sensoriais, verbos com uma polissemia relativa ao domínio da ASE, ou verbos com uma polissemia sensorial.

Para compreender melhor quais verbos são sensoriais e quais tinham uma polissemia específica do domínio da ASE, optou-se pela seguinte metodologia:

- pesquisar no dicionário de língua geral as polissemias dos 55<sup>57</sup> verbos extraídos do CEnoTEXT;
- verificar no dicionário de língua geral quais apresentam polissemias associadas ao domínio da ASE, no CEnoTEXT;
- determinar quais as polissemias, de acordo com os exemplos no CEnoTEXT;
- selecionar as ULE.

Após a realização dos diferentes passos da metodologia, apresentam-se os verbos analisados, juntamente com uma pequena apresentação das suas particularidades e de exemplos ilustrativos.

Num primeiro momento, selecionaram-se os dicionários *Priberam* e *Infopédia* (versões digitais) como referências para dicionários de língua geral e observaram-se as polissemias apresentadas para os 55 verbos.

A razão para esta dupla pesquisa prende-se com os dois motivos principais:

- (i) após uma análise de várias obras relativas ao vinho, ter-se verificado que as definições oferecidas eram vagas ou inexistentes;

---

<sup>57</sup> Uma vez que os verbos “ter”, “ser”, “estar” e “haver”, nos contextos em análise, traduzem uma ideia de predicação ou existência, e sendo esses usos comuns a outros domínios que não só o da ASE, não serão contemplados na análise que visa identificar o léxico de especialidade da área em apreço.

- (ii) não haver informação atestada sobre a possibilidade de estes verbos terem polissemias sensoriais em língua comum e em língua de especialidade.

A título de exemplo, de seguida apresentam-se os verbos “aparecer” e “evoluir” recolhidos nos dicionários de língua geral – Priberam e Infopédia ((informação recolhida a 10 de janeiro de 2022).



verbo		
<b>aparecer</b>	1. <u>Mostrar-se</u> . 2. Apresentar-se. 3. Revelar-se. 4. Ser posto à venda, publicar-se (impressos). 5. Declarar-se. 6. Manifestar-se o que estava oculto.	1. <u>Tornar-se visível</u> ; mostrar-se de repente. 2. apresentar-se. 3. começar a manifestar-se; surgir. 4. acontecer.
<b>evoluir</b>	1. Passar por transformações sucessivas (ex.: o latim evoluiu para as línguas latinas; o estado do paciente não evolui desde ontem). 2. Executar movimentos graduais e harmoniosos; fazer evoluções (ex.: os bailarinos evoluíam na pista).	1. passar por uma sucessão gradual de transformações. 2. passar para uma situação melhor. 3. fazer movimentos consecutivos.

Tabela 22. Entradas de “aparecer” e “evoluir” nos dicionários Priberam e Infopédia

Observa-se que o artigo lexicográfico de língua geral para o verbo “aparecer” contempla a aceção ligada à sensorialidade, neste caso, sensorialidade visual, no sentido de tornar visível (*mostrar-se, tornar-se visível*). Já para o verbo “evoluir”, não há menção em nenhum dos dicionários que indique polissemia sensorial.

Desta observação, apurou-se que, dos 55 verbos, apenas 14 verbos apresentam informação nos dicionários de língua geral de polissemias associadas ao sensorial. São eles: “abrir, aparecer, apresentar, beber, gostar, mostrar, notar, revelar, sentir, trazer, assentar, esconder, evidenciar, sobressair”.

Numa segunda instância, pesquisou-se quais os verbos que apresentavam polissemias no CEnoTEXT. A título de exemplo apresentam-se os mesmos verbos “aparecer” e “evoluir”.

O verbo “aparecer” surge em contextos do tipo “X aparece/aparece X”. No entanto, não parece ser no sentido da modalidade visual, uma vez que os sabores e aromas não aparecem visualmente, mas sim ao nível olfativo e gustativo. Os vários exemplos sugerem que o verbo possa ser usado como sinónimo de “existir” e de “sentir(-se)”.

20. [15818] Vinho de cor citrina, com aroma intenso onde predominam notas de frutos cítricos maduros conjugado com notas florais de mimosa e laranjeira. No sabor **aparecem** nuances amanteigadas e ligeira baunilha muito suaves e persistentes.

O verbo “evoluir” surge nos exemplos não com o sentido de “passar por diversas transformações” ou no sentido de “passar para uma situação melhor”, mas num sentido mais específico de “melhorar as suas características após um determinado período de tempo conservado em determinadas condições”, parecendo ser o argumento relativo ao local de conservação obrigatório.

21. [10768] Fruto muito maduro, compotas, químico de tinta-da-china, austero no aroma. Na boca mostra corpo mediano, taninos secos, num estilo aguerrido e muito firme que gostará de **evoluir** em garrafa.
22. [09997] Aroma com boa densidade, com notas de fruta madura e ligeira percepção floral, enquadrada pela tosta da barrica. Na boca sente-se um corpo mediano, com taninos ainda bem firmes e um perfil que mostra ser um tinto que pode ainda **evoluir** em cave.

Dos 55 verbos extraídos, 18 verbos apresentam polissemias associadas ao domínio da ASE, em concreto com exemplos relacionados com a modalidade sensorial (“abrir, aparecer, apresentar, beber, gostar, mostrar, notar, revelar, sentir, trazer, crescer, evoluir, ganhar, ir, pedir, perder, poder, sugerir”) e 12 verbos parecem ter uma polissemia relativa ao domínio da ASE, mas não necessariamente uma polissemia associada à modalidade sensorial (“acompanhar, combinar, conjugar, dar, funcionar, melhorar, misturar, prolongar, assentar, esconder, evidenciar, sobressair”).

Explicitando melhor esta diferença, apresentam-se alguns exemplos como os verbos “abrir” e “crescer” para o primeiro caso, e o verbo “acompanhar” para o segundo caso.

Um verbo como “abrir” surge em diversos exemplos, não com o significado de “destapar ou descerrar”, mas apontando para uma polissemia de especialidade da ASE. Quando se menciona “abrir”, o significado é o de dar um tempo ao vinho na garrafa ou em copo, antes de degustar ou beber, para permitir a libertação das suas características olfativas e gustativas na plenitude.

23. [17942] Discreto mas muito limpo, sem sinais de evolução, fruta algo contida. Mostra muito equilíbrio na boca, há uma boa relação entre corpo e acidez, pode beneficiar com algum tempo de cave para **abrir** o aroma.

O verbo “crescer” surge não com o sentido de desenvolver ou com o sentido de aumentar de volume ou de grandeza, mas, sim, com o sentido de tornar-se melhor num determinado momento da prova – aquando do exame olfativo, durante o exame gustativo, por exemplo – ou ainda no sentido de melhorar as suas características após um determinado período de tempo no copo.

24. [21454] Discreto no aroma, palha seca e fruta escondida, resultando ainda pouco expressivo. **Cresce** bastante na prova de boca, com bom volume e acidez, resulta agradável para a mesa. Pode também melhorar com o tempo em garrafa.

O verbo “acompanhar”, que na língua geral não parece ter polissemias associadas a ASE, mas sim de “X faz companhia a/vai junto com”, surge no CEnoTEXT em vários exemplos com o sentido de recomendação ou de instrução de harmonização. O que parece ser uma polissemia do domínio da ASE, mas não sensorial, como nos casos anteriores.

25. [01811] Feito de Jaen, Tinta Roriz e Alfrocheiro. Distingue-se do outro vinho por ter maior percentagem de Jaen e isso deu-lhe um bom carácter a que um toque de boa barrica daria outra expressão. Boa expressão na boca, com fruta viva e um estilo muito apto para **acompanhar** uma boa refeição.

A Tabela 23 apresenta os resultados da recolha da informação relativa às polissemias na língua geral, da observação das ocorrências no CE<sub>no</sub>TEXT e da verificação das premissas consideradas relevantes (i.e., (i) era um verbo com polissemia associada ao domínio da ASE; (ii) era um verbo sensorial; (iii) era um caso verbo com polissemia sensorial).

verbos	ASE		verbos	ASE	
	Língua geral	CE <sub>no</sub> TEXT		Língua geral	CE <sub>no</sub> TEXT
abrir	✓	✓	dar	x	?
aparecer	✓	✓	funcionar	x	?
apresentar	✓	✓	melhorar	x	?
assentar	✓	✓	misturar	x	?
beber	✓	✓	prolongar	x	?
esconder	✓	✓	compensar	x	x
evidenciar	✓	✓	conferir	x	x
gostar	✓	✓	confirmar	x	x
mostrar	✓	✓	contribuir	x	x
notar	✓	✓	deixar	x	x
revelar	✓	✓	destacar	x	x
sentir	✓	✓	dominar	x	x
sobressair	✓	✓	envolver	x	x
trazer	✓	✓	faltar	x	x
crescer	x	✓	fazer	x	x
evoluir	x	✓	ficar	x	x
ganhar	x	✓	impressionar	x	x
ir	x	✓	juntar	x	x
pedir	x	✓	lembrar	x	x
perder	x	✓	manter	x	x
poder	x	✓	marcar	x	x
sugerir	x	✓	resultar	x	x
acompanhar	x	?	surgir	x	x
combinar	x	?	terminar	x	x
conjugar	x	?	tornar	x	x

Legenda: tem polissemia [✓]; não tem polissemia [x]; dúvida se tem polissemia [?]

Tabela 23. Detecção da existência de polissemia sensorial na língua geral ou no CE<sub>no</sub>TEXT

Analisando os dados apresentados na tabela, verificou-se a seguinte distribuição:

- Grupo 1 – os verbos que têm polissemia associada ao domínio da ASE na língua geral e no CenoTEXT (“abrir, aparecer, apresentar, assentar, beber, gostar, esconder, evidenciar, mostrar, notar, revelar, sentir, sobressair, trazer”);
- Grupo 2 – os verbos que não têm polissemia associada ao domínio da ASE na língua geral, mas têm uma polissemia no CenoTEXT (“crescer, evoluir, ganhar, ir, pedir, perder, poder, sugerir”);
- Grupo 3 – os verbos que não têm polissemia associada ao domínio da ASE na língua geral e parecem ter um sentido mais específico no CenoTEXT (“acompanhar, combinar, conjugar, dar, funcionar, melhorar, misturar, prolongar”);
- Grupo 4 – os verbos que não têm polissemia associada ao domínio da ASE nem em língua geral nem no CenoTEXT (“compensar, conferir, confirmar, contribuir, deixar, destacar, dominar, envolver, faltar, fazer, ficar, impressionar, juntar, lembrar, manter, marcar, resultar, surgir, tornar, terminar”).

Observando os dados com particular atenção ao CEnoTEXT, verificou-se que alguns verbos podem ser de interpretação mais difícil (por exemplo, os verbos do grupo 2 “crescer, evoluir, ganhar, ir, pedir, perder, poder, sugerir”), aspeto abordado já em Barbosa (2021) relativamente a nomes e adjetivos. Apesar das baixas frequências em comparação com um *corpus* de língua geral, há uma necessidade de observar estes verbos e considerar se estes devem ser incorporados na lista de candidatos a figurar no DEASE, para evitar que a leitura destes verbos em contexto de nota de prova não seja equivocada.

Conforme é possível concluir a partir dos dados anteriormente apresentados, inúmeros são os usos não correntes de formas verbais, criando aquilo que se intitula neste trabalho de polissemia da ASE. Assim, e à semelhança do verificado na Tabela 23, existem evidências de diversas categorias em que se pode distribuir os verbos (neste caso quatro grupos), umas que ativam polissemias específicas diferentes daqueles que se lhes associam na língua comum. Complementarmente, verifica-se que nem sempre o contexto

de especialidade obriga a ativação dessa polissemia, apresentando-se uma alternância de significados, própria da noção de polissemia, que é recuperada pelo contexto e permite ao leitor apropriar-se do conteúdo das notas de prova e outros documentos de especialidade. Prova-se, desta forma, existir uma clara necessidade de sistematização desses usos, a fim de facilitar o processo de interpretação, organizando as possibilidades e tornando, paulatinamente, mais homogênea a ativação dessas formas e respetivos significados.

Assim, de acordo com esta metodologia são considerados candidatos a ULE os seguintes 26 verbos: “abrir, aparecer, apresentar, acompanhar, beber, combinar, conjugar, crescer, dar, evoluir, funcionar, ganhar, gostar, ir, melhorar, misturar, mostrar, notar, pedir, perder, poder, prolongar, revelar, sentir, sugerir, trazer”.

### 5.2.2. NOMES

Na presente secção, serão apresentadas as unidades lexicais nominais encontradas nas notas de prova e que ajudam a descrever o vinho aquando da degustação. Pretende-se observar que nomes (simples e compostos) são utilizados na descrição da experiência sensorial da prova dos vinhos. Quando o especialista descreve o vinho, utiliza no geral, muitos nomes, para denominar diferentes aspetos, tais como as castas utilizadas, as cores observadas, os aromas e os sabores sentidos, como se pode observar na seguinte nota de prova:

26. [00487] **Cor vermelho-rubi** profundo quase opaco. **Aromas** complexos, muito frutado (**amoras, mirtilos**) e floral (**violeta**), **pimento** verde e um ligeiro **toque** de **couro** com boas **notas** de **madeira**. Na **boca** é fresco com **taninos** suaves, bem estruturado, frutado e equilibrado. **Final** longo e elegante.

No CEnoTEXT, detetou-se a presença de 2 919 nomes com um total de 243 585 ocorrências (num universo de 780 585, correspondendo a 31%), que se distribuem da seguinte forma 272 nomes com mais de 100 ocorrências; 153 com 50 a 99 ocorrências;

517 com 10 a 49 ocorrências, 993 com 2 a 9 ocorrências e 906 nomes com apenas uma ocorrência. Na Tabela 24 apresenta-se os 20 nomes mais frequentes.

POSIÇÃO	LEMA	FREQ. ABS.	POSIÇÃO	LEMA	FREQ. ABS.
1	boca	14 633	11	vinho	3 516
2	nota	10 456	12	frescura	3 304
3	fruto	10 123	13	especiaria	2 796
4	aroma	9 538	14	prova	2 725
5	acidez	8 837	15	conjunto	2 716
6	final	8 620	16	perfil	2 558
7	fruta	8 220	17	nariz	2 037
8	tanino	7 881	18	madeira	1 910
9	corpo	4 798	19	estilo	1 874
10	cor	4 466	20	sabor	1 669

Tabela 24. Listagem dos 20 nomes mais frequentes de CEnoTEXT

Nesta tabela, é possível observar que os lemas que surgem como mais frequentes dão indícios sobre que unidades são candidatas a ULE. No CEnoTEXT há uma tendência para a menção da “nota” de “prova” do “vinho”, para a nomeação de alguns órgãos sensoriais em causa (“boca, nariz”), qual a modalidade em causa (“cor, aroma, sabor”), que características são comumente nomeadas para depois adjetivar de acordo com opinião do especialista (“fruto, acidez, fruta, tanino, corpo, frescura, especiaria, madeira”) e por fim, que aspetos relativos à sensação global da degustação (“conjunto, perfil, estilo”).

Com base nos dados acima apresentados, neste trabalho, de acordo com a metodologia geral optou-se por considerar apenas os nomes com mais de 100 ocorrências no CEnoTEXT, que correspondem a 244 nomes: “acidez, açúcar, afinação, alcatrão, álcool, alperce, ameixa, amêndoa, amora, amplitude, ananás, ano, aperitivo, apontamento, aptidão, arejamento, aresta, aroma, árvore, ataque, atração, auréola, austeridade, avelã, banana, barrica, barro, baunilha, biscoito, boca, bolha, bolo, bosque, branco, cacau, café, calda, caramelo, caráter, caroço, casa, casca, casta, cave, centro, cereja, chocolate, citrino, classe, colheita, complexidade, componente, compota, comprimento, concentração, conjunto, consumo, contenção, copo, cor, corpo, couro, cremosidade, definição, delicadeza, densidade, dia, doce, doçura, edição, elegância, entrada, equilíbrio, erva, escuro, especiaria, espumante, estágio, esteva, estilo, estrutura,

eucalipto, evidência, evolução, excesso, exemplo, expressão, extração, exuberância, figo, final, finura, firmeza, flor, folha, força, forma, framboesa, frescura, fruta, fruto, fumo, fundo, futuro, garra, garrafa, gás, geleia, ginja, graça, grafite, granada, groselha, harmonia, húmus, idade, impacto, impressão, intensidade, iodo, jovem, juventude, lado, laivo, laranja, laranjeira, leite, lenho, leveza, licor, lima, limão, limoeiro, limonada, linha, longo, lote, maçã, madeira, manga, marca, marmelo, mato, média, mel, melão, menta, mentol, mês, mesa, mineral, mineralidade, momento, morango, moscatel, nariz, nervo, nível, nota, notas, noz, palha, par, passa, peixe, pendor, pera, percepção, perfil, persistência, personalidade, peso, pêssego, pimenta, pimento, plano, poder, polimento, pólvora, ponto, porte, porto, potência, prato, prazer, preço, presença, profundidade, proporção, prova, qualidade, quantidade, químico, raça, rebuçado, recorte, redução, reflexo, região, registo, resina, riqueza, rosa, rubi, rugosidade, rusticidade, sabor, salmão, segura, sensação, silvestre, simplicidade, suavidade, sugestão, tabaco, tangerina, tanino, tempo, terra, textura, tinta, tinto, tipo, tom, tonalidade, toque, toranja, torrefação, tosta, traço, uva, verão, vida, vigor, vinha, vinho, violeta, vivacidade, volume”<sup>58</sup>.

Esta listagem inicial permite verificar quais as unidades lexicais mais frequentes de modo a ser possível distinguir as unidades lexicais de língua geral das potenciais ULE.

Imediatamente, é possível verificar que há um conjunto de UL associadas a vários campos: (i) temporalidade, (ii) elementos naturais, (iii) elementos não naturais, (iv) alimentos, (v) bebidas, (vi) sensações, (vii) características psicológicas (viii) características físicas (ix) características de antropomorfização (x) elementos geográficos, (xi) entre outros.

Seguindo a metodologia estabelecida, contrastou-se as frequências dos candidatos no CEnoTEXT e no *corpus* de língua geral, o CORLEX, conforme se apresenta na Tabela 25, e observou-se se o valor do diferencial será ou não relevante como critério de seleção dos candidatos a ULE.

---

<sup>58</sup> As seguintes unidades lexicais apesar de apresentarem mais de 100 ocorrências no *corpus* foram excluídas, por se considerar que possuem características distintas das UL apresentadas na Tabela 25: (i) nomes de castas: *alfrocheiro*, *alicante*, *alvarinho*, *aragonês*, *arinto*, *bouschet*, *cabernet*, *castelão*, *chardonnay*, *baga* (foram verificadas as ocorrências em que *baga* se referia a bagas de frutos e essas constam da tabela), *fernão*, *malvasia*, *pires*, *roriz*, *sauvignon*, *syrah*, *touriga*, *trincadeira*, *vaz*, *verdelho*, *viosinho*.

N.º	CEnoTEXT	NOME	CORLEX	DIFF
	FREQ. REL.		FREQ. REL.	
1	0,0486	mineralidade	0	0
2	1,8746	boca	0,0202	1,8544
3	1,3395	nota	0,0083	1,3312
4	1,2968	fruto	0,0070	1,2898
5	1,2219	aroma	0,0024	1,2196
6	1,1321	acidez	0,0001	1,1320
7	1,1043	final	0,0250	1,0793
8	1,0531	fruta	0,0032	1,0499
9	1,0096	tanino	0,0001	1,0095
10	0,6147	corpo	0,0491	0,5656
11	0,5721	cor	0,0227	0,5495
12	0,4504	vinho	0,0132	0,4372
13	0,4233	frescura	0,0022	0,4210
14	0,3582	especiaria	0,0008	0,3573
15	0,3491	prova	0,0087	0,3404
16	0,3479	conjunto	0,0197	0,3283
17	0,3277	perfil	0,0033	0,3244
18	0,2610	nariz	0,0059	0,2551
19	0,2352	tinto	0,0014	0,2338
20	0,2362	toque	0,0034	0,2329
21	0,2401	estilo	0,0079	0,2322
22	0,2447	madeira	0,0194	0,2253
23	0,2138	sabor	0,0040	0,2098
24	0,2013	silvestre	0,0018	0,1994
25	0,1991	barrica	0,0002	0,1989
26	0,2104	estrutura	0,0121	0,1982
27	0,1987	equilíbrio	0,0070	0,1917
28	0,1838	chocolate	0,0021	0,1817
29	0,1814	doçura	0,0021	0,1793
30	0,1897	presença	0,0131	0,1767
31	0,1787	elegância	0,0021	0,1766
32	0,1795	sugestão	0,0034	0,1761
33	0,1662	complexidade	0,0016	0,1646
34	0,1622	compota	0,0005	0,1617
35	0,1548	ameixa	0,0002	0,1546
36	0,1612	intensidade	0,0075	0,1537
37	0,1573	volume	0,0101	0,1472
38	0,1464	textura	0,0015	0,1450
39	0,1540	caráter	0,0096	0,1444
40	0,1592	flor	0,0163	0,1429

N.º	CEnoTEXT	NOME	CORLEX	DIFF
	FREQ. REL.		FREQ. REL.	
41	0,1363	casta	0,0009	0,1354
42	0,1717	branco	0,0367	0,1350
43	0,1323	citrino	0,0001	0,1322
44	0,1391	tom	0,0104	0,1287
45	0,1276	secura	0,0014	0,1262
46	0,1294	vegetal	0,0050	0,1244
47	0,1191	garra	0,0010	0,1181
48	0,1175	baunilha	0,0004	0,1170
49	0,1171	cereja	0,0008	0,1163
50	0,1234	evolução	0,0105	0,1129
51	0,1257	qualidade	0,0157	0,1100
52	0,1120	maçã	0,0028	0,1092
53	0,1086	persistência	0,0007	0,1079
54	0,1085	garrafa	0,0035	0,1050
55	0,1034	tosta	0,0001	0,1033
56	0,1003	amora	0,0007	0,0996
57	0,0994	cacau	0,0003	0,0991
58	0,0974	rubi	0,0005	0,0969
59	0,0949	framboesa	0,0005	0,0945
60	0,0933	ananás	0,0004	0,0928
61	0,0925	mineral	0,0024	0,0901
62	0,0899	baga	0,0002	0,0897
63	0,1108	mesa	0,0230	0,0878
64	0,0852	apontamento	0,0008	0,0844
65	0,0838	morango	0,0006	0,0832
66	0,0869	concentração	0,0043	0,0825
67	0,0843	profundidade	0,0031	0,0812
68	0,0826	laranja	0,0022	0,0804
69	0,0790	comprimento	0,0055	0,0736
70	0,0732	geleia	0,0002	0,0729
71	0,0771	erva	0,0050	0,0721
72	0,0740	casca	0,0025	0,0716
73	0,0734	limão	0,0020	0,0714
74	0,0703	resina	0,0007	0,0696
75	0,0694	pêssego	0,0005	0,0690
76	0,0655	nuance	0,0002	0,0653
77	0,0696	sensação	0,0062	0,0633
78	0,0669	pimenta	0,0037	0,0632
79	0,0751	folha	0,0130	0,0621
80	0,0658	harmonia	0,0038	0,0620

N.º	CEnoTEXT	NOME	CORLEX	DIFF
	FREQ. REL.		FREQ. REL.	
81	0,0616	groselha	0,0002	0,0614
82	0,0644	registro	0,0033	0,0611
83	0,0630	lima	0,0020	0,0611
84	0,0596	austeridade	0,0004	0,0592
85	0,0529	bolha	0,0004	0,0525
86	0,0529	bosque	0,0017	0,0512
87	0,0519	rebuçado	0,0009	0,0510
88	0,0553	traço	0,0053	0,0501
89	0,0514	mel	0,0020	0,0494
90	0,0469	tabaco	0,0023	0,0446
91	0,0450	leveza	0,0009	0,0440
92	0,0437	toranja	0,0000	0,0437
93	0,0478	gás	0,0049	0,0429
94	0,0460	álcool	0,0036	0,0424
95	0,0421	alperce	0,0002	0,0419
96	0,0455	palha	0,0041	0,0414
97	0,0471	prato	0,0062	0,0410
98	0,0410	limonada	0,0001	0,0409
99	0,0404	finura	0,0004	0,0399
100	0,0407	uva	0,0011	0,0396
101	0,0430	ataque	0,0037	0,0393
102	0,0433	excesso	0,0041	0,0392
103	0,0393	violeta	0,0009	0,0384
104	0,0377	auréola	0,0005	0,0372
105	0,0377	aptidão	0,0011	0,0366
106	0,0366	granada	0,0008	0,0359
107	0,0369	couro	0,0015	0,0354
108	0,0356	vivacidade	0,0007	0,0349
109	0,0656	fundo	0,0310	0,0346
110	0,0351	pera	0,0015	0,0336
111	0,0343	delicadeza	0,0010	0,0334
112	0,0380	personalidade	0,0051	0,0330
113	0,0346	perceção	0,0017	0,0329
114	0,0331	espumante	0,0004	0,0327
115	0,0410	consumo	0,0084	0,0326
116	0,0329	caramelo	0,0004	0,0325
117	0,0331	cave	0,0010	0,0320
118	0,0339	vigor	0,0021	0,0319
119	0,0346	mato	0,0028	0,0317
120	0,0302	menta	0,0002	0,0300

N.º	CEnoTEXT	NOME	CORLEX	DIFF
	FREQ. REL.		FREQ. REL.	
121	0,0301	notas	0,0002	0,0299
122	0,0311	aresta	0,0015	0,0297
123	0,0336	juventude	0,0039	0,0296
124	0,0402	café	0,0109	0,0293
125	0,0296	exuberância	0,0003	0,0293
126	0,0288	marmelo	0,0001	0,0287
127	0,0282	mentol	0,0000	0,0282
128	0,0304	nervo	0,0028	0,0276
129	0,0296	evidência	0,0020	0,0276
130	0,0284	pólvora	0,0009	0,0276
131	0,0281	esteva	0,0005	0,0276
132	0,0291	impacto	0,0016	0,0275
133	0,0383	classe	0,0108	0,0275
134	0,0272	estágio	0,0010	0,0262
135	0,0261	grafite	0,0002	0,0260
136	0,0259	cremosidade	0,0000	0,0259
137	0,0264	tangerina	0,0006	0,0258
138	0,0286	proporção	0,0039	0,0246
139	0,0259	porte	0,0013	0,0246
140	0,0252	figo	0,0010	0,0242
141	0,0292	componente	0,0051	0,0241
142	0,0341	prazer	0,0101	0,0239
143	0,0250	tonalidade	0,0011	0,0239
144	0,0246	alcatrão	0,0008	0,0238
145	0,0342	quantidade	0,0117	0,0225
146	0,0270	vinha	0,0048	0,0222
147	0,0223	ginja	0,0001	0,0222
148	0,0268	fumo	0,0049	0,0219
149	0,0227	lote	0,0008	0,0219
150	0,0227	amêndoa	0,0013	0,0214
151	0,0233	manga	0,0023	0,0210
152	0,0222	recorte	0,0012	0,0210
153	0,0204	limoeiro	0,0003	0,0201
154	0,0204	banana	0,0007	0,0197
155	0,0225	densidade	0,0029	0,0197
156	0,0201	biscoito	0,0005	0,0196
157	0,0196	rugosidade	0,0001	0,0195
158	0,0196	eucalipto	0,0011	0,0185
159	0,0218	bolo	0,0035	0,0183
160	0,0268	doce	0,0085	0,0183

N.º	CEnoTEXT	NOME	CORLEX	DIFF
	FREQ. REL.		FREQ. REL.	
162	0,0184	aperitivo	0,0005	0,0179
163	0,0306	peso	0,0127	0,0179
164	0,0181	melão	0,0004	0,0177
165	0,0178	polimento	0,0002	0,0176
166	0,0250	marca	0,0075	0,0175
167	0,0187	firmeza	0,0015	0,0172
168	0,0174	pimento	0,0005	0,0169
169	0,0169	calda	0,0005	0,0164
170	0,0168	contenção	0,0005	0,0163
171	0,0163	arejamento	0,0002	0,0160
172	0,0160	moscatel	0,0002	0,0158
173	0,0164	laranjeira	0,0006	0,0158
174	0,0325	expressão	0,0168	0,0158
175	0,0160	caroço	0,0005	0,0156
176	0,0164	suavidade	0,0009	0,0155
177	0,0259	rosa	0,0107	0,0152
178	0,0208	riqueza	0,0056	0,0151
179	0,0154	lenho	0,0005	0,0149
180	0,0190	definição	0,0043	0,0147
181	0,0155	licor	0,0008	0,0147
183	0,0146	laivo	0,0003	0,0143
184	0,0223	copo	0,0081	0,0142
185	0,0151	noz	0,0009	0,0142
186	0,0142	rusticidade	0,0001	0,0141
187	0,0150	pendor	0,0010	0,0140
188	0,0142	passa	0,0004	0,0138
189	0,0137	húmus	0,0003	0,0134
190	0,0135	salmão	0,0003	0,0132
191	0,0132	torrefação	0,0000	0,0132
192	0,0174	redução	0,0043	0,0131
193	0,0135	iodo	0,0004	0,0130
194	0,0129	vintage	0,0001	0,0129
195	0,0129	afinação	0,0001	0,0128
196	0,0128	avelã	0,0002	0,0126
197	0,0147	simplicidade	0,0022	0,0126
198	0,0167	reflexo	0,0042	0,0125
199	0,0159	amplitude	0,0035	0,0124
200	0,0137	extração	0,0013	0,0124
201	0,0257	graça	0,0134	0,0123
202	0,0146	colheita	0,0026	0,0120

N.º	CEnoTEXT	NOME	CORLEX	DIFF
	FREQ. REL.		FREQ. REL.	
203	0,0151	tinta	0,0032	0,0119
204	0,0183	média	0,0067	0,0116
205	0,0220	par	0,0105	0,0115
206	0,0145	barro	0,0032	0,0113
207	0,0141	atração	0,0028	0,0112
208	0,0155	açúcar	0,0047	0,0108
209	0,0496	ponto	0,0389	0,0107
210	0,0132	edição	0,0039	0,0093
211	0,0170	leite	0,0083	0,0087
212	0,0304	centro	0,0217	0,0086
213	0,0133	raça	0,0049	0,0084
214	0,0146	químico	0,0062	0,0084
215	0,0138	impressão	0,0066	0,0073
216	0,0135	potência	0,0063	0,0072
217	0,0730	lado	0,0659	0,0071
218	0,0229	plano	0,0172	0,0057
219	0,0156	peixe	0,0101	0,0055
220	0,0164	verão	0,0119	0,0045
221	0,0202	árvore	0,0173	0,0030
222	0,0159	escuro	0,0141	0,0018
223	0,0146	preço	0,0129	0,0017
224	0,0169	jovem	0,0168	0,0001
225	0,0145	entrada	0,0166	-0,0021
226	0,0147	futuro	0,0171	-0,0024
227	0,0128	porto	0,0181	-0,0053
228	0,0177	linha	0,0233	-0,0056
229	0,0314	tipo	0,0391	-0,0077
230	0,0131	idade	0,0215	-0,0085
231	0,0188	nível	0,0278	-0,0090
232	0,0129	região	0,0245	-0,0116
233	0,0245	força	0,0366	-0,0121
234	0,0514	terra	0,0646	-0,0132
235	0,0140	mês	0,0285	-0,0145
236	0,0145	momento	0,0350	-0,0205
237	0,0965	tempo	0,1175	-0,0210
238	0,0129	exemplo	0,0405	-0,0275
239	0,0372	forma	0,0714	-0,0343
240	0,0172	vida	0,1069	-0,0897
241	0,0138	casa	0,1080	-0,0941
242	0,0380	ano	0,1479	-0,1099

N.º	CEnoTEXT	NOME	CORLEX	DIFF
	FREQ. REL		FREQ. REL	
243	0,0165	dia	0,1285	-0,1119

N.º	CEnoTEXT	NOME	CORLEX	DIFF
	FREQ. REL		FREQ. REL	
244	0,0138	poder	0,2808	-0,2670

Tabela 25. Lista de nomes com as freq. relativas e o valor diferencial de frequência entre CEnoTEXT e CORLEX

Observa-se que os candidatos nominais apresentam a seguinte distribuição:

- Grupo de unidades que não ocorrem no CORLEX, como “mineralidade”;
- Grupo com valores de frequência relativa mais elevados no CEnoTEXT do que CORLEX, como “boca, nota, fruto, aroma, acidez, final, fruta, tanino”;
- Grupo intermédio com valores de frequência maiores no CEnoTEXT;
- Grupo com valores mais baixos no CEnoTEXT do que no CORLEX, como “jovem, entrada, futuro, porto, linha, tipo, idade, nível, região, força, terra, mês, momento, tempo, exemplo, forma, vida, casa, ano, dia, poder”.

Numa análise inicial, é lógico colocar a hipótese de que a maioria dos nomes aqui listados remete para elementos presentes no vinho ou que, de alguma forma, ajudam a caracterizá-lo, como determinadas cores, aromas ou sabores. É, igualmente, possível verificar a presença de nomes que indicam uma sensação na prova, como “prazer”, “suavidade”, “rugosidade”, “intensidade”, “persistência”, etc. Ao observar-se a listagem acima, e uma vez conscientes do domínio, parece previsível e natural convocar tais elementos, que acabam por funcionar como elementos do mundo real e que ajudam o leitor a associar uma cor, aroma ou sabor a um produto prototípico. Apesar de haver, naturalmente, um número elevado de palavras que pertencem, igualmente, à língua geral, outras unidades são, por outro lado, muito mais específicas, como é o caso de “mineralidade”. Alguns dos nomes listados, ainda que, no eixo sintagmático, assumam o papel de nomes, o facto é que substituem, provavelmente, alguns adjetivos. Veja-se, por

exemplo, o caso de “secura” ou de “persistência”, que parecem substituir estruturas copulativas como “ser/estar/parecer seco” ou “ser/estar/parecer persistente”. Há, por esse motivo, um grande número de nomes deadjetivais, que podem ser utilizados como uma estratégia para que as notas de prova não se tornem numa lista de adjetivações e tornar a leitura diversificada.

Esta divisão, tendo em conta os critérios<sup>59</sup> já apresentados (ver Tabela 18), permite tomar a decisão de considerar os três primeiros grupos de nomes como candidatos a ULE.

No entanto, ao contrário do que sucedeu com os candidatos rejeitados nos verbos (onde alguns candidatos “rejeitados” com mais de 100 ocorrências foram alvo de uma outra revisão relativa ao valor diferencial entre CEnoTEXT e CORLEX) neste caso, optou-se pela exclusão dos candidatos com valores negativos de diferencial (“entrada, futuro, porto, linha, tipo, idade, nível, região, força, terra, mês, momento, tempo, exemplo, forma, vida, casa, ano, dia, poder”), pois não apresentam contextos polissémicos no CEnoTEXT ou concordâncias/colocações específicas, ou mesmo dúvidas de interpretação na leitura das notas de prova que justificassem a sua repescagem. A exceção foi o candidato “jovem”.

Considerando os 20 candidatos com as frequências absolutas mais altas no CEnoTEXT e os 20 candidatos com o diferencial mais elevado entre CEnoTEXT e CORLEX, 17 candidatos são exatamente iguais e com a mesma disposição, exceto as três últimas posições, onde há duas alterações: (i) a alteração de candidato (“madeira” vs. “tinto”) e (ii) a alteração de ordem (“estilo” e “toque”), como se verifica na Tabela 26.

---

<sup>59</sup> Esta é uma (das várias) formas de selecionar candidatos quando perante um grande número de candidatos. Esta seleção permite maximizar a estratégia (frequência absoluta; frequência relativa comparada com um *corpus* de língua geral) de escolha, permitindo, posteriormente, e em casos excepcionais, observar os contextos (as concordâncias e coocorrências) de outros candidatos, permitindo avaliar se é ou não um potencial candidato a ULE.

Freq. Absoluta		Valor Diferencial	Freq. Absoluta		Valor Diferencial
boca	1	boca	vinho	11	vinho
nota	2	nota	frescura	12	frescura
fruto	3	fruto	especiaria	13	especiaria
aroma	4	aroma	prova	14	prova
acidez	5	acidez	conjunto	15	conjunto
final	6	final	perfil	16	perfil
fruta	7	fruta	nariz	17	nariz
tanino	8	tanino	<b>madeira</b>	<b>18</b>	<b>tinto</b>
corpo	9	corpo	<b>estilo</b>	<b>19</b>	<b>toque</b>
cor	10	cor	<b>toque</b>	<b>20</b>	<b>estilo</b>

Tabela 26. Lista de contraste dos 20 candidatos com frequência absoluta e os valores diferenciais

De uma forma geral, os nomes desempenham um papel crucial na comunicação, ajudando a identificar, a classificar e a expressar uma ampla variedade de informações. Relativamente aos candidatos observados eles servem para: (i) denominar o produto, (ii) categorizar o produto, (iii) mencionar a zona de produção (iv) referir as castas(s) envolvidas, (v) apresentar as características que invocam quando o vinho é ingerido.

Pode dividir-se os candidatos por partes do corpo (“boca, corpo, nariz”); sentidos (“cor, aroma, toque”); avaliação psicológica (“perfil, estilo”); tipo de vinho (“tinto”); objetos e alimentos (“fruto, fruta, especiaria, madeira”), entre outras categorizações possíveis.

Para uma análise mais fina cruzaram-se análises. Por um lado, foi usada, a funcionalidade *Concordance* que facilita a observação dos contextos à esquerda e à direita da UL, como se apresenta na Figura 24.



Figura 24. Excerto de concordâncias de “boca”

Por outro, foi usada uma funcionalidade do Sketch Engine, o Word Sketch<sup>60</sup> (combinatórias estatisticamente relevantes), que permite ver facilmente, e de modo imediato, os contextos sintáticos de ocorrência das formas pesquisadas, dando pistas objetivas acerca de sentidos especializados. Esta funcionalidade permite (i) identificar colocações, (ii) analisar coocorrências, (iii) observar os contextos onde ocorrem as unidades lexicais, (iv) ver informações estatísticas e (v) observar gráficos de distribuição. Em síntese, o Word Sketch é uma ferramenta importante para a análise detalhada do uso de unidades lexicais em contextos linguísticos, ajudando a compreender como estas se relacionam umas com as outras nos textos selecionados.

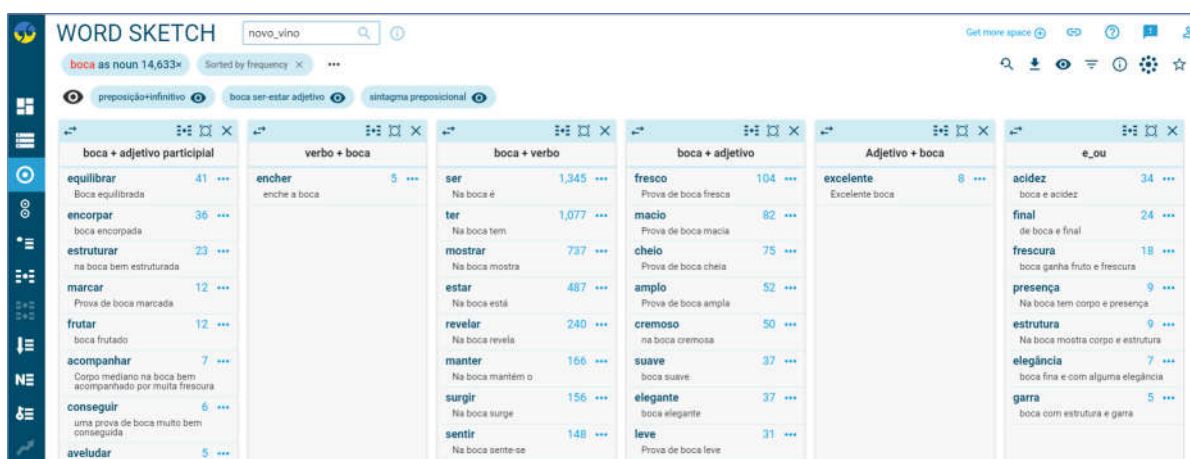


Figura 25. Word Sketch de “boca”

Observando a Figura 25, verifica-se que do ponto de vista da análise lexical, o facto de “boca” surgir junto de verbos e adjetivos é o esperado do ponto de vista sintático de

<sup>60</sup> “Word sketches are defined as automatic corpus-derived summaries of a word’s grammatical and collocational behavior (Kilgarriff et al., 2004). Rather than looking at an arbitrary window of text around the headword—as occurs in previous corpus tools – Sketch Engine is able to look for each grammatical relation that the word participates in (Kilgarriff et al., 2004). The default word sketches provided by Sketch Engine represent diferente relations, such as verb-object, modifiers or prepositional phrases. However, except for the recently implemented generic-specific word-sketches, they only represente linguistic relations.”(León-Araúz, & Martín, (2018).)

um nome, no entanto, as coocorrências que apresenta permitem ter uma ideia geral dos contextos relevantes onde “boca” surge e que outras unidades coocorrem com ela. A análise de Word Sketch permite, também, extrair informação, que pode ser pertinente, para determinar particularidades das UL candidatas, pois indica sentidos específicos das UL. Essa mesma informação, pode ser mais tarde relevante para inserir no DEASE.

De seguida, apresentam-se alguns exemplos de algumas particularidades, que se observam alguns dos candidatos mais frequentes.

As UL “boca” e “nariz” são, muitas vezes, usadas em contextos que permitem interpretá-las como agentes ou protagonistas da experiência,

27. [09369] Aroma de boa intensidade com notas de vegetal seco, fruta madura mas fresca, bons apontamentos minerais e citrinos. A **boca** tem volume e frescura, a fruta está presente de forma austera e a acidez refresca o conjunto.
28. [05776] O **nariz** mostra laranja suave, minerais, folha de limoeiro, num todo composto e alegre. Na boca revela corpo médio, cremoso, com muito boa acidez, tudo bastante completo e equilibrado.

em vez de partes do corpo.

29. [20037] Muito limpo e jovem mas com carácter menos marcado, cítrico com algum tropical leve toque de líchias. **Na boca** tem corpo mediano, mas é todo equilibrado, suave e com elegante secura especiada final.
30. [16952] Bonita cor rubi. **No nariz** mostra morangos e rosas, num todo fresco e apelativo. Muito certinho **na boca**, redondo e gordo, com alguma doçura compensada por acidez fresca, final com comprimento, mas um pouco quente e pesado.

Na expressão “boca ampla”, não se espera que o leitor interprete tratar-se de uma zona bucal maior, mas de uma sensação provocada pelo vinho e que diz respeito à ideia de volume/corpo e/ou de sensação de textura, viscosidade e estrutura que se traduzem numa perceção global de preenchimento e de equilíbrio na boca.

31. [03121] Syrah, Touriga Nacional e Cabernet. Apontamentos vegetais, de frutos maduros, ameixa, cacau. **Boca ampla**, com taninos firmes a dar garra, num tom seco e austero, fresco e gastronómico.

Em ASE, é comum o uso de “nota” que ao contrário do que é sugerido na língua corrente, em que é compreendido como registo escrito de algo, ou como um aspeto musical que indica um som musical, no domínio da ASE é regularmente utilizado como indicador de presença de aroma “nota de X” ou “nota X”.

32. [13676] Um tinto sério, algum fruto, bastante mineral, leve **nota** de tabaco, todo cheio de intenção. Na boca é vigoroso de taninos, com adstringência fina, complexo, fresco e cheio de corpo e elegância.
33. [08526] Uma **nota** madura, compotada, num fundo de fruta vermelha e do bosque, nuances tostadas. Corpo médio, boa frescura, muito jovem, com notas de bagas maceradas, final apimentado e longo.

No entanto, também há exemplos de “nota” como indicador de sabor, em casos de retrogosto olfativo (sensação olfativa sentida na boca).

34. [14305] A cor amarelo palha claro e luminoso, com reflexos de limão, anuncia uma frescura intensa que se vai confirmar no nariz. É um vinho que apresenta um perfil aromático muito simples, frutado e agradável. Na boca, predominam as **notas** de limão, a acidez é elevada, está cheio de juventude e frescura. O seu final é ditado pela acidez, é seco, frutado e descontraído.
35. [03429] Deste vinho existem 5.000 litros em casco. Tem uma cor de aguardente velha, um aroma austero, muitas notas de caixa de charutos, figos, mel e muito fruto seco. O aroma abre muito no copo. Perfeito na boca, **notas** de farinha e fruto seco, viscoso, é seda pura. Muito rico e fino. Final muito longo.

Para “fruto”, é interessante verificar que há um conjunto de coocorrências de:

- “Fruto de X (tipo de fruto)”: essa categoria envolve expressões que descrevem o tipo específico de fruto, indicando a origem ou variedade do fruto em questão. Por exemplo, “fruto de amora” se refere ao tipo de amora (outros exs. “de amora, de cereja, de ameixa, de maçã, de polpa branca, de framboesa, de caroço (de pêsego), de groselha, de morango, de bagas do bosque, de bagas silvestres, de lima, de ananás, de ginja, de groselha, de manga, de meloa, de alperce, de citrino, de clementina, de líchias”);
- “Fruto de qualidade”: aqui, “fruto” está relacionado a uma avaliação de qualidade. Por exemplo, “fruto de qualidade” indica que o fruto é considerado de alta qualidade, acima do esperado.
- “Fruto em X (processo de preparação culinária)”: esta categoria envolve a descrição de como o fruto é preparado ou usado na culinária. Por exemplo, “fruto em compota” significa que o fruto foi transformado em compota (outros exs. “em passa, em geleia, em calda, em camadas, em rebuçado”)
- “Fruto + COR”: nessa categoria, a cor do fruto é explicitada. Por exemplo, “fruto encarnado” se refere a um fruto de cor vermelha (outros exs. “amarelo, vermelho, preto, azul, branco”);
- “Fruto + adjetivo”: esta categoria descreve o estado, sabor ou características olfativas do fruto.
  - Por exemplo, “fruto macerado” indica que o fruto foi alvo de um processo não natural (outros exs. “cozido, confitado, escondido”);
  - enquanto “fruto verde” descreve um fruto que ainda não amadureceu (outros exs. “doce, verde, maduro” (de acordo com o sabor);
  - “expressivo, exuberante, polido, contido, vibrante” (de acordo com a sensação que produz na boca).
- “fruto silvestre” descrição da tipologia a que pertence o fruto (outros exs. “fruto cítrico, fruto tropical, fruto do bosque”).
  - ainda há a especificação de tipo concreto de silvestre (“fruto silvestre de amora, fruto silvestre de morango”).
- “Fruto maduro”: Esta categoria descreve frutos que estão maduros e prontos para consumo. Pode haver subcategorias relacionadas ao processo de preparação ou ao tipo de fruto.

- “em compota, em passa, em geleia”;
- “de cereja, de cereja, de polpa branca, de pêssego, de maçã, de pêssego”;
- “sóbrio, fresco, bonito, leve, limpo, elegante, delicado, austero, discreto, franco, jovem, intenso, macerado, cozido”

Estas coocorrências indicam que sempre que é mencionado “fruto” a preocupação é descrever o tipo de cor, ou fruto em particular, que tipo de processo foi alvo, que tipo de aroma/sabor está associado.

Relativamente à UL “fruta”, esta assemelha-se em muitos contextos e coocorrências ao detetado para “fruto”, apenas com destaque para as combinações de cores que não ocorrem com “fruto”, como é o caso de “fruta preta e vermelha, fruta preta e azul, fruta preta e violeta, fruta violeta”, onde não ocorre a discriminação de que frutas estão contempladas.

36. [13672] Impressionante no aroma, profundo e rico, denso, com **fruta preta e vermelha**, notas tostadas, ervas aromáticas. Grande estrutura de boca, taninos poderosos mas polidos, surpreendentemente elegante.
37. [10018] **Fruta preta e azul**, violetas, tostados da madeira um pouco em evidência, um toque de couro. Corpo médio, taninos um pouco saídos, acidez alta a dar boa frescura, bom comprimento, mas um pouco áspero.
38. [07937] **Fruta preta e violeta** muito densa, com trazo mineral, ervas aromáticas, chocolate, alcaçuz. Bonito e profundo, complexo. Na boca entra discreto, corpo médio, depois explode em sabores, texturado, fresco, profundo, rugoso, sedutor, longo. Belíssimo.
39. [07939] Tostados, terra húmida, **fruta violeta**, madeiras exóticas e especiarias doces. Corpo médio, com alguma secura a vir na frente, depois a acidez gulosa e os taninos firmes suportam a prova de boca, até ao final longo e vibrante.

Ao falar de “estilo” é sempre de um determinado tipo: (muito) atrativo, consensual, gastronómico, jovem, limpo, acessível, afirmativo, definido, diferente, direto, discreto, doce, elegante; (mais) leve, fresco, vigoroso, tradicional, amplo, arcaico, clássico

A UL “estilo”, quando associada a adjetivos, geralmente, descreve características/qualidades subjetivas, aspeto reforçado pela coocorrência elevada de “estilo + muito” e de “estilo + mais”. Algumas das coocorrências parecem indicar que há uma preferência em associar “estilo” a:

- adjetivos descritivos associados a personificação à semelhança do que se verificou com “nariz” e “boca” (“atrativo, gastronómico, jovem, limpo, acessível, afirmativo, definido, direto, discreto, elegante”);
- adjetivos associados a sabor ou qualidade sensorial “leve, fresco, vigoroso, amplo”;
- outros adjetivos “tradicional, arcaico, clássico”.

O nome “acidez” é frequentemente usado para descrever a qualidade de ser ácido, sensação associada ao exame gustativo. As coocorrências obtidas fornecem informações adicionais sobre essa característica específica:

- “acidez + a + verbo”: nesse contexto, a “acidez” está associada a ações ou características específicas. Por exemplo, “acidez a casar” sugere que a acidez combina bem com algo, enquanto “acidez a compensar” indica que a acidez é capaz de equilibrar ou compensar outras características. (outros exs. “a comandar, a compensar, a conferir, a contribuir, a dar, a equilibrar, a permitir, a potenciar, a preencher”)
- “acidez + adjetivo”: aqui, a acidez é qualificada por adjetivos que descrevem sua natureza ou intensidade. Por exemplo, “acidez agradável” significa que a acidez é agradável ao paladar, enquanto “acidez alta” indica uma acidez pronunciada ou intensa (outros exs. “aceitável, acertada, acutilante, adequada, afiada, agradável, agridoce, ajustada, alegre, alta, baixa, acidez citrina, correta, crispante, crocante, cítrica, delicada, discreta, elevada, fina, firme, fresca, frutada, harmoniosa, integrada,

limonada, macia, marcada, mediana, moderada, média, notável, perfeita, pronunciada, quanto baste, refrescante, suave, suficiente, vibrante, viva”)

- “acidez + que”: essa construção sugere que a acidez é descrita ou definida em termos específicos. Por exemplo, "acidez que refresca" indica que a acidez tem a qualidade de ser de um tipo específico que é refrescante.
- “acidez no final”, “acidez no ponto”: estas expressões indicam que a “acidez” é evidente no paladar após outras características do sabor, ou que o vinho atingiu um determinado grau de acidez que se considera não só aceitável, como desejável para poder ser ingerido.

40. [20575] Feito apenas com Touriga Nacional, está muito perfumado no aroma, com sugestões de framboesa, morango, delicado toque floral. Tem um bom volume de boca mas sem ser pesado, revela grande limpeza de fruto, é um rosé vivo, muito fresco, quase limonado na bela **acidez final**.

41. [06098] Muita cor, frutos do bosque, notas tostadas, minerais, muito sério e com boa contenção. Cheio, encorpado, com taninos firmes e sólidos, **acidez no ponto**, final sereno, com austeridade que indicia longa vida em garrafa.

Pelo acima exposto, facilmente se compreende que os significados assumidos em contexto pelas UL listadas são diversos e distintos dos que são usualmente associados a essas UL na língua corrente. Compreende-se desta feita, que esses usos revelam tratar-se de polissemias associadas a um domínio específico – o da ASE –, conduzindo à hipótese de que essas estruturas são fortes candidatas a constar numa compilação do léxico de especialidade.

Em conclusão, verifica-se que o uso de Word Sketch proporciona uma pista sobre as coocorrências das UL indicando pistas para a inserção de ULE complexas, bem como auxiliar à escolha das ULE complexas com critérios estatisticamente relevantes.

Assim, de acordo com esta metodologia são considerados candidatos a ULE os seguintes 219 nomes (i) aqueles que não ocorrem no CORLEX, (ii) aqueles com valores de frequência relativa mais elevados no CEnoTEXT do que CORLEX, (iii) aqueles que após uma observação das particularidades indicadoras de domínio de especialidade:

“acidez, açúcar, afinação, alcatrão, álcool, alperce, ameixa, amêndoa, amora, amplitude, ananás, aperitivo, apontamento, aptidão, arejamento, aresta, aroma, árvore, ataque, atração, auréola, austeridade, avelã, baga, banana, barrica, barro, baunilha, biscoito, boca, bolha, bolo, bosque, branco, cacau, café, calda, caramelo, caráter, caroço, casca, casta, cave, centro, cereja, chocolate, citrino, classe, colheita, complexidade, componente, compota, comprimento, concentração, conjunto, consumo, contenção, copo, cor, corpo, couro, cremosidade, definição, delicadeza, densidade, doce, doçura, edição, elegância, equilíbrio, erva, especiaria, espumante, estágio, esteva, estilo, estrutura, eucalipto, evidência, evolução, excesso, expressão, extração, exuberância, figo, final, finura, firmeza, flor, folha, framboesa, frescura, fruta, fruto, fumo, fundo, garra, garrafa, gás, geleia, ginja, graça, grafite, granada, groselha, harmonia, húmus, impacto, impressão, intensidade, iodo, juventude, lado, laivo, laranja, laranjeira, leite, lenho, leveza, licor, lima, limão, limoeiro, limonada, lote, maçã, madeira, manga, marca, marmelo, mato, média, mel, melão, menta, mentol, mesa, mineral, mineralidade, morango, moscatel, nariz, nervo, nota, notas, noz, nuance, palha, par, passa, peixe, pendor, pera, percepção, perfil, persistência, personalidade, peso, pêssego, pimenta, pimento, plano, polimento, pólvora, ponto, porte, potência, prato, prazer, presença, profundidade, proporção, prova, qualidade, quantidade, químico, raça, rebuçado, recorte, redução, reflexo, registo, resina, riqueza, rosa, rubi, rugosidade, rusticidade, sabor, salmão, segura, sensação, silvestre, simplicidade, suavidade, sugestão, tabaco, tangerina, tanino, textura, tinta, tinto, tom, tonalidade, toque, toranja, torrefação, tosta, traço, uva, vegetal, verão, vigor, vinha, vinho, violeta, vivacidade, volume”.

### 5.2.3. ADJETIVOS

Na presente secção, serão apresentadas as unidades lexicais adjetivais encontradas nas notas de prova e que ajudam a descrever o vinho aquando da degustação. Pretende-se, desta forma observar que adjetivos são utilizados na descrição da experiência sensorial da prova dos vinhos. Quando o especialista descreve o vinho, utiliza adjetivos para caracterizar diferentes aspetos, tais como a intensidade das cores observadas, os aromas e os sabores sentidos, como se pode observar na seguinte nota de prova:

42. [06235] Cor amarelo **mediano**. Aroma **rico** e **cheio**, fruta **atrativa**, **refinada**, notas de maçã, pomar, floral, **subtil** complexidade e profundidade. Um deleite de aroma. **Volumoso**, **amplo** e **fresco** na boca, uma acidez **impecável** de notoriedade. Mostra uma **boa** integração da madeira. **Persistente**, complexidade **evidente**, mas a garrafa será conselheira **ideal**.

No CEnoTEXT, detetou-se a presença de 2 481 adjetivos com um total de 164 979 ocorrências (num universo de 780 585, correspondendo a 22%), que se distribuem da seguinte forma 181 nomes com mais de 100 ocorrências; 203 com 50 a 99 ocorrências; 467 com 10 a 49 ocorrências, 917 com 2 a 9 ocorrências e 713 adjetivos com apenas uma ocorrência. Na Tabela 27, apresenta-se os 20 adjetivos mais frequentes.

POSIÇÃO	LEMA	FREQ. ABS.	POSIÇÃO	LEMA	FREQ. ABS.
1	bom	11405	11	suave	2994
2	maduro	5945	12	médio	2765
3	fresco	5266	13	fino	2728
4	leve	4731	14	cheio	2683
5	seco	3443	15	mineral	2572
6	elegante	3316	16	intenso	2496
7	floral	3169	17	vivo	2262
8	longo	3084	18	vermelho	2116
9	vegetal	3079	19	preto	2103
10	ligeiro	3015	20	firme	2102

Tabela 27. Listagem dos 20 adjetivos mais frequentes de CEnoTEXT

Nesta tabela, é possível observar que os lemas que surgem como mais frequentes dão indícios sobre que unidades são candidatas a ULE. No CEnoTEXT, os dados deixam antever que as unidades adjetivais em análise focam, essencialmente, atributos qualitativos (como ser “bom”), de durabilidade (“longo, médio”), associados ao olfato (“seco, floral, vegetal”), à dimensão (“fino, cheio”), à intensidade (“intenso, vivo”), à cor (“preto, vermelho”), à textura (“leve, suave, firme, fino”), ao sabor (“maduro, fresco,

seco, mineral”), entre outras categorizações, o que nos leva a concluir que de facto se está perante um conjunto de candidatos a UL do domínio da ASE.

Neste trabalho, optou-se por considerar apenas os adjetivos com mais de 100 ocorrências no CEnoTEXT, que correspondem a 181 adjetivos (“bom, maduro, fresco, leve, seco, elegante, floral, longo, vegetal, ligeiro, suave, médio, fino, cheio, mineral, intenso, vivo, vermelho, preto, firme, aromático, citrino, macio, discreto, delicado, grande, redondo, belo, doce, agradável, austero, limpo, tropical, balsâmico, atrativo, cremoso, complexo, cítrico, mediano, expressivo, franco, verde, jovem, gastronómico, envolvente, bonito, presente, persistente, excelente, simples, branco, correto, saboroso, amarelo, harmonioso, sólido, nacional, exuberante, rico, sério, gordo, amargo, direto, fácil, profundo, vigoroso, denso, final, forte, alegre, evidente, vibrante, curto, químico, negro, clássico, atraente, sedoso, amplo, apelativo, sumarento, aberto, volumoso, sóbrio, refrescante, ácido, novo, poderoso, alto, subtil, guloso, generoso, herbáceo, crocante, sedutor, morno, rosé, tinto, moderno, opaco, brilhante, primeiro, interessante, enorme, puro, claro, violeta, largo, mentolado, pronto, acessível, escuro, quente, geral, consensual, azul, melhor, rugoso, imediato, untuoso, húmido, certo, velho, meio, exótico, apetecível, afirmativo, picante, robusto, distinto, tenso, pleno, típico, super, salino, compacto, versátil, vinoso, resinoso, rústico, capaz, seguro, potente, pálido, notável, diferente, delicioso, groselha, incisivo, convincente, alcoólico, amanteigado, prazenteiro, polivalente, suficiente, musculado, preciso, positivo, terroso, lácteo, afável, ácido, consistente, notório, habitual, impositivo, anterior, imponente, curioso, cristalino, alentejano, fragrante, completo, imenso, aperitivo, residual, excessivo, límpido”).

Seguindo a metodologia estabelecida, contrastou-se as frequências dos candidatos no CEnoTEXT e no *corpus* de língua geral, o CORLEX, conforme se apresenta na Tabela 28, e observou-se se o valor do diferencial será ou não relevante como critério de seleção dos candidatos a ULE.

N	CEnoTEXT	NOME	CORLEX	DIFF	N	CEnoTEXT	NOME	CORLEX	DIFF
	FREQ.REL		FREQ.REL			FREQ.REL			
1	0,0217	vinoso	0	0	39	0,1434	mediano	0,0006	0,1428
2	0,0161	acidulo	0	0	40	0,1419	expressivo	0,0012	0,1408
3	1,4611	bom	0,0686	1,3925	41	0,1475	complexo	0,0081	0,1394
4	0,7616	maduro	0,0018	0,7598	42	0,1290	franco	0,0035	0,1255
5	0,6746	fresco	0,0079	0,6667	43	0,1206	gastronómico	0,0003	0,1203
6	0,6061	leve	0,0105	0,5956	44	0,1194	envolvente	0,0012	0,1182
7	0,4411	seco	0,0120	0,4290	45	0,1273	verde	0,0173	0,1100
8	0,4248	elegante	0,0030	0,4218	46	0,1102	persistente	0,0010	0,1092
9	0,4060	floral	0,0006	0,4053	47	0,1255	jovem	0,0168	0,1088
10	0,3944	vegetal	0,0050	0,3894	48	0,1191	bonito	0,0116	0,1076
11	0,3862	ligeiro	0,0042	0,3820	49	0,1189	presente	0,0148	0,1041
12	0,3836	suave	0,0063	0,3773	50	0,1070	excelente	0,0046	0,1024
13	0,3951	longo	0,0371	0,3580	51	0,1020	saboroso	0,0008	0,1012
14	0,3495	fino	0,0092	0,3403	52	0,1031	correto	0,0042	0,0989
15	0,3542	médio	0,0154	0,3388	53	0,0989	harmonioso	0,0011	0,0978
16	0,3295	mineral	0,0024	0,3271	54	0,1004	amarelo	0,0066	0,0938
17	0,3437	cheio	0,0203	0,3234	55	0,2193	grande	0,1279	0,0914
18	0,3198	intenso	0,0053	0,3145	56	0,0913	exuberante	0,0006	0,0907
19	0,2898	vivo	0,0195	0,2703	57	0,0943	sólido	0,0042	0,0901
20	0,2671	aromático	0,0008	0,2663	58	0,1057	simples	0,0200	0,0857
21	0,2693	firme	0,0044	0,2649	59	0,0890	sério	0,0098	0,0792
22	0,2694	preto	0,0109	0,2586	60	0,0912	rico	0,0140	0,0773
23	0,2711	vermelho	0,0130	0,2581	61	0,0796	amargo	0,0024	0,0772
24	0,2458	citrino	0,0001	0,2457	62	0,0783	vigoroso	0,0013	0,0770
25	0,2375	macio	0,0026	0,2349	63	0,0776	denso	0,0027	0,0749
26	0,2296	discreto	0,0037	0,2259	64	0,0801	gordo	0,0065	0,0735
27	0,2293	delicado	0,0043	0,2250	65	0,0929	nacional	0,0195	0,0733
28	0,2178	redondo	0,0057	0,2121	66	0,0716	vibrante	0,0009	0,0707
29	0,2139	belo	0,0146	0,1993	67	0,0788	direto	0,0090	0,0698
30	0,1874	austero	0,0012	0,1862	68	0,0733	alegre	0,0054	0,0678
31	0,1874	agradável	0,0038	0,1836	69	0,1044	branco	0,0367	0,0677
32	0,1906	doce	0,0085	0,1821	70	0,0784	profundo	0,0125	0,0659
33	0,1770	limpo	0,0044	0,1727	71	0,0788	fácil	0,0134	0,0654
34	0,1683	balsâmico	0,0001	0,1682	72	0,0712	químico	0,0062	0,0650
35	0,1682	atrativo	0,0010	0,1672	73	0,0630	sedoso	0,0003	0,0628
36	0,1686	tropical	0,0027	0,1659	74	0,0632	atraente	0,0013	0,0619
37	0,1542	cremoso	0,0002	0,1541	75	0,0619	apelativo	0,0001	0,0618
38	0,1457	cítrico	0,0001	0,1455	76	0,0717	evidente	0,0101	0,0617

N	CEnoTEXT	NOME	CORLEX	DIFF	N	CEnoTEXT	NOME	CORLEX	DIFF
	FREQ.REL		FREQ.REL			FREQ.REL			
77	0,0716	curto	0,0105	0,0611	118	0,0252	tenso	0,0009	0,0243
78	0,0629	amplo	0,0040	0,0589	119	0,0297	húmido	0,0056	0,0241
79	0,0579	sumarento	0,0001	0,0578	120	0,0301	imediatO	0,0060	0,0241
80	0,0638	clássico	0,0067	0,0571	121	0,0246	super	0,0007	0,0239
81	0,0544	volumoso	0,0007	0,0537	122	0,0234	salino	0,0002	0,0233
82	0,0539	sóbrio	0,0013	0,0527	123	0,0232	compacto	0,0015	0,0217
83	0,0529	refrescante	0,0003	0,0526	124	0,0218	versátil	0,0002	0,0216
84	0,0744	forte	0,0234	0,0510	125	0,0254	distinto	0,0040	0,0213
85	0,0749	final	0,0250	0,0500	126	0,0352	escuro	0,0141	0,0211
86	0,0524	ácido	0,0034	0,0490	127	0,0213	resinoso	0,0004	0,0209
87	0,0673	negro	0,0191	0,0481	128	0,0249	típico	0,0042	0,0206
88	0,0475	guloso	0,0005	0,0470	129	0,0333	quente	0,0132	0,0201
89	0,0492	subtil	0,0027	0,0465	130	0,0208	rústico	0,0013	0,0194
90	0,0459	herbáceo	0,0001	0,0457	131	0,0377	largo	0,0185	0,0192
91	0,0456	crocante	0,0000	0,0456	132	0,0201	potente	0,0011	0,0190
92	0,0474	generoso	0,0021	0,0453	133	0,0186	groselha	0,0002	0,0184
93	0,0496	poderoso	0,0046	0,0450	134	0,0251	pleno	0,0071	0,0180
94	0,0451	sedutor	0,0012	0,0439	135	0,0183	convicente	0,0004	0,0180
95	0,0438	rosé	0,0001	0,0437	136	0,0179	amanteigado	0,0000	0,0179
96	0,0436	tinto	0,0014	0,0422	137	0,0178	prazenteiro	0,0002	0,0176
97	0,0441	morno	0,0019	0,0422	138	0,0183	incisivo	0,0008	0,0176
98	0,0416	opaco	0,0008	0,0409	139	0,0177	polivalente	0,0002	0,0175
99	0,0576	aberto	0,0179	0,0398	140	0,0182	alcoóico	0,0011	0,0171
100	0,0373	mentolado	0,0001	0,0372	141	0,0172	musculado	0,0004	0,0168
101	0,0378	violeta	0,0009	0,0369	142	0,0187	delicioso	0,0021	0,0166
102	0,0400	interessante	0,0050	0,0350	143	0,0323	azul	0,0159	0,0164
103	0,0364	acessível	0,0015	0,0349	144	0,0165	terroso	0,0002	0,0163
104	0,0415	brilhante	0,0068	0,0347	145	0,0164	lácteo	0,0004	0,0160
105	0,0433	moderno	0,0092	0,0341	146	0,0196	notável	0,0038	0,0158
106	0,0331	consensual	0,0002	0,0328	147	0,0163	afável	0,0007	0,0156
107	0,0306	rugoso	0,0005	0,0302	148	0,0199	pálido	0,0043	0,0155
109	0,0393	puro	0,0095	0,0299	149	0,0150	impositivo	0,0001	0,0149
110	0,0298	untuoso	0,0002	0,0296	150	0,0158	consistente	0,0010	0,0148
111	0,0284	apetecível	0,0003	0,0282	151	0,0155	notório	0,0011	0,0144
112	0,0369	pronto	0,0091	0,0278	152	0,0136	fragrante	0,0000	0,0136
113	0,0274	afirmativo	0,0004	0,0270	153	0,0149	pires	0,0013	0,0135
114	0,0284	exótico	0,0017	0,0267	154	0,0205	seguro	0,0072	0,0133
115	0,0272	picante	0,0006	0,0266	155	0,0132	aragonês	0,0000	0,0132
116	0,0266	robusto	0,0012	0,0255	156	0,0141	imponente	0,0010	0,0131
117	0,0398	enorme	0,0155	0,0244	157	0,0131	aperitivo	0,0005	0,0125

N	CEnoTEXT	NOME	CORLEX	DIFF	N	CEnoTEXT	NOME	CORLEX	DIFF
	FREQ.REL		FREQ.REL			FREQ.REL		FREQ.REL	
158	0,0138	cristalino	0,0013	0,0125	170	0,0206	capaz	0,0164	0,0042
159	0,0131	residual	0,0007	0,0124	171	0,0132	imenso	0,0098	0,0033
160	0,0386	claro	0,0266	0,0120	172	0,0331	geral	0,0353	-0,0022
161	0,0138	alentejano	0,0020	0,0118	173	0,0142	anterior	0,0199	-0,0057
162	0,0128	límpido	0,0014	0,0114	174	0,0170	preciso	0,0257	-0,0087
163	0,0170	positivo	0,0065	0,0106	175	0,0188	diferente	0,0375	-0,0187
164	0,0131	excessivo	0,0034	0,0097	176	0,0314	melhor	0,0505	-0,0191
165	0,0141	curioso	0,0050	0,0091	177	0,0292	velho	0,0542	-0,0250
166	0,0133	completo	0,0049	0,0084	178	0,0297	certo	0,0558	-0,0261
167	0,0174	suficiente	0,0096	0,0079	179	0,0290	meio	0,0614	-0,0325
168	0,0154	habitual	0,0079	0,0075	180	0,0509	novo	0,0991	-0,0483
169	0,0493	alto	0,0443	0,0050	181	0,0410	primeiro	0,0974	-0,0564

Tabela 28. Lista de contraste dos 20 candidatos com frequência absoluta e os valores diferenciais

Observa-se que os candidatos adjetivais apresentam a seguinte distribuição:

- Grupo de unidades que não ocorrem no CORLEX “vinoso, ácido”;
- Grupo com valores de frequência relativa mais elevados no CEnoTEXT do que CORLEX, do exemplo “bom” a “imenso”;
- Grupo com valores mais baixos no CEnoTEXT do que no CORLEX, como “geral, anterior, preciso, diferente, melhor, velho, certo, meio, novo, primeiro”.

Esta divisão inicial, tendo em conta os critérios apresentados anteriormente (aquando da Tabela 18), permite tomar a decisão de considerar os dois primeiros grupos de dados como candidatos a ULE. Uma vez mais, a maior parte dos adjetivos listados não é específica da língua de especialidade, sendo que existem algumas formas que, à semelhança do que já acontecia com os nomes, parecem remeter de forma mais clara para esse domínio, como “rosé” ou “vinoso”.

Outra observação interessante é a de que o número de adjetivos com mais de cem ocorrências é menor do que o número de nomes com mais de cem ocorrências, o que leva a colocar duas hipóteses: (1) de que a utilização de nomes deadjetivais substitui, talvez por motivos estéticos do texto, o uso de adjetivos; (2) de que a diversidade/originalidade no uso de adjetivos é superior à do uso de nomes, razão pela qual estes últimos apresentam um maior número de formas com mais repetição.

Aplica-se a esta secção a mesma metodologia aplicada aos nomes, ou seja, a utilização das funcionalidades *Concordance* e *Word Sketch* do *Sketch Engine* para observar com detalhe as particularidades de alguns dos candidatos apresentados anteriormente.

Através da análise das concordâncias e coocorrências, é possível identificar que certos adjetivos apresentam comportamentos singulares. Este aspeto já foi inicialmente explorado em Barbosa (2021), apontando que determinados adjetivos utilizados nas notas de prova podiam ser interpretados de forma equívoca ou errónea, devido ao facto de serem adjetivos que ocorrem com frequência na língua quotidiana.

Começando pelos candidatos que apenas surgem no CEnoTEXT e não ocorrem no CORLEX, como é o caso de “vinoso” e “ácido”.

“Vinoso” surge maioritariamente em posição pós-nome [x + vinoso], onde X pode ser associado à modalidade olfativa (“aroma vinoso, perfil vinoso, fundo vinoso”).

43. [13022] Muita cor, aroma vinoso, exuberante, cheio de notas de frutos do bosque, jovem e vivaz. É um tinto sumarento, cheio de fruto muito limpo, com raça duriense no final longo a que uma boa acidez empresta frescura e irreverência. Um perfeito vinho jovem.
44. [20127] Um rosé de cor muito viva e intensa e perfil **vinoso**, sólido, nos aromas a bagas esmagadas e morangos. Fruto de boa qualidade na boca, acidez generosa, um rosé para a mesa, com alguma garra.
45. [05798] Um tinto musculado. Aroma com fruto bem maduro, algum em passa, nota suave de chocolate num fundo **vinoso**. Taninos muito aguerridos na boca, seco, intenso e pujante final.

Mas, na sua maioria, é associado à modalidade gustativa, seja a aspetos táteis ou de paladar (“toque vinoso, tanino vinoso”), seja a sensação de conjunto e de sensação final (“final vinoso”).

46. [19133] Aroma atractivo, entre o floral e o fruto, o mentol, o chocolate e a apara de lápis. Toque **vinoso** na boca, sobressai o corpo alcoólico, tanino com leve aresta, bom centro de boca, final seco e morno.
47. [20573] Muito atractivo no aroma, com a fruta presente, jovem e fresca, mas sem excessos, num estilo que continua a ser vinho e não refresco. Na boca mantém o mesmo perfil, austero e frutado, sentindo-se mesmo um certo tanino **vinoso**, a boa acidez a equilibrar o conjunto e a prolongar o final. Um rosé gastronómico, feito para a mesa.
48. [09650] Centro escuro, de grande profundidade, com auréola semi-alargada de cor rubi. Nariz em evolução, de intensidade média, com notas balsâmicas, fruta silvestre confitada e sabonete. Boca de acidez média, alcoólico, corpo médio, fruta em média quantidade, mas a eclipsar-se. Taninos fortes, que apesar de não estarem em enorme quantidade, sobressaem no conjunto. Final **vinoso**, não mais que médio em persistência, e algo amargo.

“Acídulo” surge sempre associado à modalidade gustativa, explicitando que aspeto em concreto é acídulo, se é ao início (“garra, entrada”), se é durante (“postura”), se durante o final da prova (“nervo, final”).

49. [20640] Especiado num aroma a limão, maçã com amargo de caroço de ameixa. Boca com garra **acídula**, bom corpo, segura elegante, muito fruto e especiaria, final longo e saboroso.
50. [01218] Cor granada e aberta. Nariz com notas frutadas ligeiras, sentindo-se o couro e a noz-moscada, num conjunto que, mesmo ao fim de alguma agitação do vinho no copo, exige benevolência. Boca desequilibrada, com entrada **acídula**, nunca conseguindo ao longo da prova encontrar um foco.

51. [08161] Nota cítrica casada com nota de aço, leve austeridade, com alguma tropicalidade associada. Na boca tem uma postura **ácidula**, cheia de nervo e vivacidade, arrepiante na frescura, bom fruto, alguma mineralidade, final seco.
52. [16578] Um aroma com alguma pólvora e sulfuroso, fruto de maçã verde, acidez viva e impositiva, final longo e com nervo **ácidulo**.
53. [14378] Nariz eminentemente mineral, pouco intenso. Boca simples, sem sobressaltos nem surpresas, em cima de notas citrinas. Final **ácidulo**, pouco confortável.

O candidato “bom” é um adjetivo que, regra geral, está associado à avaliação de qualidade, subjetiva, tratando-se de uma avaliação pessoal e variável consoante o sujeito, no sentido de gosto/não gosto na língua geral. No CEnoTEXT, no entanto, este adjetivo surge em ocorrências do tipo [vinho + bom], como indicador da qualidade geral e intrínseca do vinho e, regra geral, sempre em contexto de sugestão de harmonização.

54. [13935] Cor amarela pálida, com reflexos de tonalidades citrinas. Aromas florais, complementados por ligeiros toques de pêssego e nêspera. A boca é fresca e mais sugestiva do que o nariz, revelando salva, raspas de lima e um mineral que enquadra o conjunto. Comprimento razoável, com bom foco no final de boca. Um vinho **bom** para utilização quotidiana, sem compromissos.

Ou, em alternativa, associado à qualidade ou avaliação positiva, no sentido de indicar que algo possui um conjunto de características positivas e adequadas – como uma avaliação objetiva de qualidade intrínseca do vinho em causa. Esta avaliação pode ser em relação a um aspeto em particular (“volume, corpo, equilíbrio, comprimento, fruto, final, perfil, tanino, impacto”) e surge sempre em posição [bom + X], ou seja, naquela característica em especial a avaliação é positiva e não subjetiva.

55. [13986] Muita fruta vermelha, algo mentolado, num estilo irresistível e rico. A melhorar a cada passo. **Bom** volume de boca, fresco, amplo de sabor e taninos bem desenhados. Belo vinho.

56. [12886] Cor granada aberta. Nariz floral, com sugestões de pétalas de rosa. Boca muito frutada e fresca, a lembrar cerejas maduras. **Bom** comprimento de prova e persistência.

O candidato “mineral” é, numa primeira leitura, interpretado como nome (associado à geologia como substâncias inorgânicas naturais, ou à nutrição, como nutrientes essenciais ao funcionamento do corpo humano). No entanto, no CEnoTEXT ocorre como adjetivo [X + mineral], comportando-se como um adjetivo relacional que numa primeira interpretação leitura se entende como nome.

Ao contrário de ocorrências da língua geral, como de “água mineral”, onde a água é descrita como sendo proveniente de uma fonte natural e contendo minerais dissolvidos, nos exemplos encontrados no CEnoTEXT, a interpretação é associada não apenas a uma substância que pode ser sentida na boca, mas essencialmente como um aroma.

Como se verifica nos exemplos associados a modalidade olfativa.

57. [10864] Alguma evolução na cor, aroma com compota de ameixa, erva seca, algum fumado e sensação **mineral**. É um tinto muito suave mas fresco, acetinado, com taninos cooperantes e muito boa aptidão gastronómica.
58. [13161] Atractivo no aroma **mineral**, com apontamentos florais e frutados de qualidade. Leve e delicado, apesar da estrutura sólida, com boas notas citrinas a surgirem na boca, vivo e persistente.
59. [02920] Nariz **mineral**, com fruta cítrica discreta, um pouco de folhas. Ligeiro e algo discreto, mas com boa secura, tem acidez crocante e boa definição.

Mas também ocorrem exemplos onde a “mineral” está associado a modalidade gustativa.

60. [14584] Cor rubi intensa, boa profundidade com reflexo vermelho-violeta. Desde logo a fruta sobressai, com elegância e alguma complexidade. Notas de groselha preta bem madura, "cassis", tudo bem fresco e em integração com a barrica. Especiarias e algum mineral. Ataque firme, apresenta solidez na estrutura, sempre com finura. Acidez

fresca e **mineral**, taninos elegantes, proporcionando bom desenvolvimento ao longo da prova. Um vinho bem feito que promete evoluir.

61. [15646] Citrino aberto na cor. Espalha frescura e citrinos, é vivaz, mostra suave madureza, mas uma contenção a conferir sobriedade e estilo. Mantém essa postura na boca, com correção e balanço notáveis, num vinho que cresce ao longo da prova. Final **mineral** e campestre, a merecer relação de confiança.

O candidato “complexo” surge num conjunto de estruturas descritivas da qualidade do vinho no seu todo, ou seja, “complexo” não se aplica a um aspeto em particular.

62. [05086] Sem ser um vinho **complexo** ou muito ambicioso está muito bem feito. O fruto vermelho intenso, leve alicorado e a tosta estão bem casados. Na boca os taninos sobressaem sobre o fruto mostrando um conjunto seco e austero.

Existem, no entanto, outros contextos onde “complexo” surge associado à modalidade olfativa (“nariz, aroma”) ou à modalidade gustativa (“boca, acidez”).

63. [17529] Nariz **complexo**, com frutos brancos e amarelos, notas de lima, minerais e muito ligeiros tostados. Harmonioso e sedutor na boca, com acidez perfeita, saboroso mas contido, com uma grande elegância, e muito poder de sedução.
64. [07119] O vinho destaca-se pela sua elegância, pela finura da fruta citrina e pelas notas de ameixa e muito leve nota de fruto tropical que aqui lhe dá classe. Texturado e com excelente acidez, mostra-se na boca **complexo**.

O que também se verifica é que o uso de “complexo” é muitas vezes em situação de coocorrência com outro adjetivo, como estratégia sistemática de explicar que se é complexo é simultaneamente outra qualidade (“complexa e atraente/elegante/estruturada/intensa”, “complexo e alegre/apimentado/atraente/bonito” ou “complexo e distinto/elegante/equilibrado/exuberante/poderoso/rico/sedutor/tenso”) dando pistas que “complexo” implica uma leitura interpretação. Ou seja, quando é mencionado que é “um vinho/nariz complexo” imediatamente associa-se uma dificuldade em explicar em que

consiste essa complexidade. Mas ao mesmo tempo, ao utilizar a estratégia de mencionar “complexo” seguido de outro adjetivo com conotação positiva “alegre/bonito” transmite a ideia de que apesar de não ser fácil verbalizar em que sentido é “complexo” esta característica é encarada como positiva e apreciada.

65. [02853] É um **vinho complexo** e rico, com elegantes notas vegetais entrelaçadas com ameixa madura, cacau, nuances minerais. Bela textura de taninos sedosos, encorpado mas não pesado, com muito equilíbrio, frescura e carácter que lhe garantem longevidade. Excelente relação qualidade-preço.
66. [11709] Bolha fina, em cordão rápido e solto. Cor amarela pálida. Aromas clássicos, com sugestões amanteigadas e pão torrado. A boca é **complexa e atraente**, desde a entrada vigorosa até ao final de prova, longo e persistente. Tem muito ainda para dar.

Pelo exposto acima, e de acordo com a metodologia utilizada, são considerados candidatos a ULE os seguintes 171 adjetivos: (i) grupo de unidades que não ocorrem no CORLEX; (ii) grupo com valores de frequência relativa mais elevados no CEnoTEXT do que CORLEX, que após uma observação das ocorrências e coocorrências apresentem particularidades indicadoras de domínio de especialidade: “aberto, acessível, ácido, acídulo, afável, afirmativo, agradável, alcoólico, alegre, alentejano, alto, amanteigado, amarelo, amargo, amplo, apelativo, aperitivo, apetecível, aragonês, aromático, atraente, atrativo, austero, azul, balsâmico, belo, bom, bonito, branco, brilhante, capaz, cheio, cítrico, citrino, claro, clássico, compacto, completo, complexo, consensual, consistente, convincente, correto, cremoso, cristalino, crocante, curioso, curto, delicado, delicioso, denso, direto, discreto, distinto, doce, elegante, enorme, envolvente, escuro, evidente, excelente, excessivo, exótico, expressivo, exuberante, fácil, final, fino, firme, floral, forte, fragrante, franco, fresco, gastronómico, generoso, gordo, grande, groselha, guloso, habitual, harmonioso, herbáceo, húmido, imediato, imenso, imponente, impositivo, incisivo, intenso, interessante, jovem, lácteo, largo, leve, ligeiro, límpido, limpo, longo, macio, maduro, mediano, médio, mentolado, mineral, moderno, morno, musculado, nacional, negro, notável, notório, opaco, pálido, persistente, picante, pires, pleno, poderoso, polivalente, positivo, potente, prazenteiro, presente, preto, profundo, pronto, puro, quente, químico, redondo, refrescante, residual, resinoso, rico, robusto, rosé,

rugoso, rústico, saboroso, salino, seco, sedoso, sedutor, seguro, sério, simples, sóbrio, sólido, suave, subtil, suficiente, sumarento, super, tenso, terroso, tinto, típico, tropical, untuoso, vegetal, verde, vermelho, versátil, vibrante, vigoroso, vinoso, violeta, vivo, volumoso”.

#### 5.2.4. EMPRÉSTIMOS

Como em qualquer domínio, também na ASE se observa a presença e uso de léxico de origem não portuguesa, para designar algumas características do vinho. Considerou-se pertinente tratar das unidades lexicais denominadas *empréstimos* à parte das outras classes gramaticais, por apresentarem características próprias: grafia não adaptada, grafia variável, variação (na anotação) morfossintática. Este subgrupo de análise surge como curiosidade por aspetos neológicos neste domínio de especialidade que se encontra pouco explorado.

Optou-se por considerar as diferentes grafias apresentadas no CEnoTEXT para um mesmo empréstimo. Por exemplo, para *bâtonnage* não foram normalizadas as formas *batonnage*, *battonage*, *battonnage* também encontradas. Esta decisão prende-se com o facto de que será interessante verificar a instabilidade ao nível da grafia. Na Tabela 29 apresentam-se todas as grafias sem indicação da forma preferida (em função da frequência ou, em alternativa, no caso de frequência iguais ou semelhantes em função da sua ocorrência em dicionários de referência das línguas de origem).

<b>língua</b>	<b>exemplos</b>
alemã	<i>beerenauslese, blaufraenkisch, delicatessen, fuchsia, fuschia, gewurstraminer, gewurtzstraminer, gewürz, gewurztraminer, gewürztraminer, riesling</i>
árabe	<i>humus</i>
castelhana	<i>Alicante, Aragonéz, jaen, Jerez, manzanilla, mencia, Puente Alto, ribera del duero, rueda, tempranillo, tempranillo do Duero, terrantez, Vinho de Jerez, xerês</i>
chinesa	<i>lichies</i>
francesa	<i>Alicante Bouchet, Alicante Bouschet, armagnac, baton, batonnage, bâtonnage, battonage, batonnage, blanc de blanc, blanc de noir, blanc de noirs, bouquet, Bouschet, brioche, brut nature, nature, Cabernet, Cabernet Franc, Cabernet Sauvignon, Cabernet Sauvignon, cadaloc, caladoc, carignan, cereja confit, cerise, chadonnay, champagne, champagne millésimé, Chardonay, Chardonnay, chic, cognac, cointreau, condrieu, confit, Côte du Rhone, Côtes du Rhône, coulis, creme brulé, crème brulé, creme brulée, crème brûlée, creme fraiche, croissants, degorgement, dégorgement, dégradé, dégradé, Provence, finesse, foie gras, foie-gras, fruit confit, ganache, goudron, gourmand, gourmet, Grand Noir, grenache, grenat, loire, Marsanne, marselan, médoc, Merlot, mousse, naif, nougat, pain grillé, grillé, parfait, paté, patina, patine, pattine, perlage, petillant, pétillant, petit Chablis, Petit manseng, Petit maseng, Petit Syrah, Petit Verdot, Pinot, Pinot Blanc, Pinot Gris, Pinot Meunier, Pinot Noir, plateau, pot pourri, pot-pourri, praliné, provence, pudim flan, Rhône, cerise, roussane, sauternes, Sauvignon, Sauvignon Blanc, sauvignon, semillon, sémillon, sommelier, sous bois, sous-bois, souvignon blanc, subois, sur lie, Syrah, tannat, tapenade, terroir, vale de maule, vin de paille, viogner, Viognier, zest, zeste</i>
neerlandesa	<i>Niepoort</i>
inglesa	<i>after eight, after-eight, appeal, approach, aromas-off, bacon, best-seller, blend, blockbuster, block-buster, blush, bold, brandy, Brandy Constantino, brandy de Jerez, brownie, butterscotch, caterpillar, celebration, cocktail, darjeeling, chuney, clean, cocktail, crispy, cup cake, design, drop, drops, earl Grey, easy, easy going and very good sparkling, fashion, field blend, fizz, golden, granny smith, hard, indeed a different white, jam, juicy, kick, kirsch, late harvest, licorice, light, liquorice, long drink, must, must have, off-flavour, offley, orange wine, performance, pink, white, rosa choc, ruby, ruby-violeta, sexy, shiraz, soft, spicy, stock, swing, tawnies, tawny, toffee, trendy, twist, vintage, whiskies, Williams, years old</i>
italiana	<i>aceto, brunello, capuccino, magnum, moka, mokka, muscat, pizza, retrogusto, sottobosco</i>
japonesa	<i>tsunami, umami</i>
latina	<i>botritys, botritys, botrytis, brett, silex, sui generis, sui-generis, sylex</i>
neozelandesa	<i>kiwi</i>

Tabela 29. Listagem de empréstimos do CEnoTEXT

No CEnoTEXT foram detetados 247 empréstimos não adaptados com 3 311 ocorrências, de 11 línguas (alguns com resultados residuais) a saber: 1 da língua árabe, 1

da língua chinesa, 1 da língua neerlandesa, 1 da língua neozelandesa, 2 da língua japonesa, 8 da língua latina, 10 da língua italiana, 11 da língua alemã, 14 da língua castelhana, 74 da língua inglesa e 124 da língua francesa.

Os dados parecem evidenciar poucas atestações de empréstimos quando em comparação com os dados do CEnoTEXT, e que os empréstimos que os especialistas se preocuparam em atestar no CEnoLEX são de pouca ou nenhuma frequência no CEnoTEXT. Observa-se que a maioria dos empréstimos mais frequentes no CEnoTEXT não são alvo da atenção dos especialistas no CEnoLEX.

Outro aspeto que pode explicar a ausência dos empréstimos do CEnoLEX pode ser a opção de os especialistas preferirem em alguns casos as formas adaptadas, veja-se o caso de bruto em vez de *bruit* ou de *brut*; ou de espumante em vez de *mousseux*.

Os dados confirmam também o que já tinha sido referido em Barbosa e Lino (2020), evidenciando um predomínio da língua francesa face à língua inglesa numa amostra dos dados, que se confirma com a análise da totalidade do CEnoTEXT, ainda que nos últimos anos haja um evidente domínio de produção de conteúdos em língua inglesa, quer por parte dos mercados anglo-saxónicos produtores de vinho (Austrália, Canadá, EUA, Nova Zelândia e Reino Unido), quer pela ideia de produzir conteúdos em língua inglesa de outros mercados para funcionar como atrativo comercial e turístico e como instrumento de internacionalização.

Outro aspeto particularmente interessante é a escrita de grafias estrangeiras (“cássis” (“cassis, casis”), “rubi” (“ruby”), “tofe” (“toffee”), “mirtilo” (“mirtilho”), “líchia” (“lícha”), “moca” (“moka, mokka”)) que, em casos excepcionais, podiam ser considerados gralhas, quando as ocorrências são diminutas. Noutros casos, porém, a frequência demonstra a instabilidade entre o uso do equivalente em português/da grafia adaptada e o empréstimo.

Verifica-se que a criatividade neológica existe neste domínio e que pode ser feito um trabalho de contraste, por exemplo, entre a frequência de uso de empréstimos não adaptados, os empréstimos adaptados ou semiadaptados e as suas formas traduzidas equivalentes, ou um estudo das tendências de uso deste tipo de unidades na redação das notas de prova por parte dos especialistas.

A manter a mesma metodologia anteriormente usada para selecionar os candidatos a ULE dos verbos, nomes e adjetivos, numa primeira fase apenas seriam contemplados os candidatos com mais de 100 ocorrências, isso significaria selecionar apenas 10 candidatos: “syrah, alicante, bouschet, sauvignon, cabernet, roriz, chardonnay, mousse, vintage, nougat”. Uma vez que sete desses são nomes de castas (que funcionam como nomes próprios) e, por uma questão de coerência metodológica, foram eliminados da seleção inicial de nomes, resultando a análise na seleção de apenas três candidatos, “mousse, vintage e nougat”.

#### 5.2.5. SÍNTESE ANÁLISE CATEGORIAL

Nesta análise categorial, identificou-se três grupos de candidatos: verbos (884), nomes (2 919) e adjetivos (2 481). Para saber quais eram candidatos a ULE, como primeiro critério foram selecionados de cada categoria apenas os candidatos com mais de 100 ocorrências do CEnoTEXT, a saber: verbos (55), nomes (244), adjetivos (181), empréstimos (10). Posteriormente, a metodologia geral empregada começou com a comparação das frequências dos candidatos do CEnoTEXT *vs.* um *corpus* de língua geral, o CORLEX. Tal, permitiu uma análise detalhada para determinar se o diferencial de frequência era ou não relevante como critério na seleção das candidatas a ULE.

No sentido de identificar que verbos eram sensoriais e quais tinham uma polissemia específica do domínio da ASE, optou-se pela seguinte metodologia: (i) pesquisar no dicionário de língua geral as polissemias dos 55 verbos extraídos do CEnoTEXT; (ii) verificar no dicionário de língua geral quais apresentam polissemias associadas ao domínio da ASE, no CEnoTEXT; (iii) determinar quais as polissemias, de acordo com os exemplos no CEnoTEXT. Desta resultou uma divisão dos dados em 4 grupos: (i) os verbos que têm polissemia associada ao domínio da ASE na língua geral e no CenoTEX; (ii) os verbos que não têm polissemia associada ao domínio da ASE na língua geral, mas têm uma polissemia no CenoTEXT; (iii) os verbos que não têm polissemia associada ao domínio da ASE na língua geral e parecem ter uma sentido mais específico no CenoTEXT; (iv) os verbos que não têm polissemia associada ao domínio da ASE nem em língua geral nem no CenoTEXT. Deste processo foram selecionados 26 verbos.

Para identificar que nomes e adjetivos eram candidatos ao domínio da ASE, optou-se por observar os resultados de concordâncias e coocorrências dos mesmos e anotar as particularidades encontradas. Foram considerados apenas (i) aqueles que não ocorrem no CORLEX, (ii) aqueles com valores de frequência relativa mais elevados no CEnoTEXT do que CORLEX, (iii) aqueles que após uma observação das particularidades indicadoras de domínio de especialidade. Deste processo foram selecionados 219 nomes e 171 adjetivos.

### 5.3. ANÁLISE CONTRASTIVA

Após a análise categorial, foi feita uma análise contrastiva entre os dados recolhidos no CEnoTEXT e no CEnoLEX para analisar o que é mencionado pelos especialistas nas notas de prova e o que é atestado nas obras recolhidas e, assim, verificar o que está já atestado no CEnoLEX. Esta análise permite também perceber que falhas são mostradas pelos dados de uso refletidos no CEnoTEXT.

De modo a motivar os dados em contraste, recorda-se, de forma sucinta os *corpora*:

- O CEnoLEX é um *corpus* monolingue de PE. Contém 26 textos especializados do tipo lexicográfico, recolhidos entre 2012-2021 e anotados gramaticalmente (519 nomes, 429 adjetivos, 64 adjetivos e nomes, 14 verbos). Tem um total de 257 581 *tokens* (9 997 *types*, dos quais 1 790 ocorrem apenas uma vez (*hápax*)), correspondendo na contagem final a 1 054 descritores, com a seguinte distribuição por modalidade sensorial: 164 na modalidade visual, 271 na modalidade olfativa, 478 na modalidade gustativa. Contém 44 descritores associados à ASE e 97 descritores associados a mais do que uma modalidade sensorial.
- o CEnoTEXT é um *corpus* monolingue PE. Contém 21 495 textos especializados do tipo nota de prova, recolhidos entre 2012-2021 e anotados gramaticalmente tem um total de 780 585 *tokens* (7 015 *types*, dos quais 2 301 ocorrem apenas uma vez (*hápax*)).

O que se pretende com esta análise contrastiva é verificar quais os candidatos do CEnoTEXT estão atestados no CEnoLEX e verificar se os descritores que surgem nas obras contidas no CEnoLEX espelham a utilização dos especialistas e as necessidades dos utilizadores. Isto é, se os descritores muito frequentes no CEnoTEXT, com polissemias específicas do domínio da ASE, estão atestados no CEnoLEX. Além disso, a análise permite verificar quais os descritores que os especialistas usam e que, por tal, seria relevante constarem em recursos lexicográficos de especialidade.

A análise contrastiva permite compreender que unidades estão em uso efetivo (os dados do CEnoTEXT) e os descritores que os especialistas consideram pertinentes ocorrerem nas obras lexicográficas que eles próprios produzem (os dados do CEnoLEX). Observa-se, de forma evidente que há uma percentagem de unidades que não constam como descritores atestados, apesar de serem UL com características e polissemias associadas ao domínio da ASE, ou a aspetos sensoriais em particular, que ocorrem em textos de especialidade e que surgem também na língua geral. Observa-se também o contrário: descritores pouco frequentes e não coincidentes com a língua geral que não se encontram atestados em recursos de especialidade.

Ou seja, parece ser possível concluir que, se é um descritor que ocorre com muita frequência e não é compreendido na totalidade pelos utilizadores, não está atestado. O que se traduz num desajuste entre a realidade do uso e a realidade daquilo que se considera pertinente constar numa obra lexicográfica, evidenciando que as necessidades dos utilizadores não estão satisfeitas pelas obras existentes.

De igual modo, apresenta-se a análise por categoria gramatical: verbos, nomes e adjetivos.

### 5.3.1. VERBOS

Relembrando os dados obtidos no CEnoLEX relativamente a descritores sensoriais da categoria verbal atestados, no CEnoLEX foram apenas atestados 14 verbos e no CEnoTEXT 884 verbos.

Dos verbos no CEnoLEX, ocorreram no CEnoTEXT os seguintes 8:

- “sentir”, com 316 ocorrências;
- “melhorar”, com 70 ocorrências;
- “beber” e “abrir”, com 30 ocorrências cada;
- “refrescar” e “perfumar”, com 3 ocorrências cada;
- “suavizar” e “exalar”, apenas com uma ocorrência cada.

Os verbos “avinhar, borbulhar, corar, degustar, porvir” e “sedimentar” não foram encontrados no CEnoTEXT.

Observa-se assim que, no CEnoLEX, os autores das várias obras atestam poucos verbos e que os poucos verbos que são atestados são de pouca ou nenhuma frequência no CEnoTEXT.

Noutra instância, analisando as definições no CEnoLEX dos verbos (“sentir, melhorar, beber, sentir, abrir, perfumar, suavizar, exalar”), garantiu-se que estas são confirmadas pelos vários exemplos dos verbos no CEnoTEXT, exceto no caso de “refrescar”.

No CEnoLEX, o verbo “refrescar” apresenta a seguinte definição:

	Definição	Obra(s) onde ocorre
1	Rejuvenescer um determinado vinho adicionando-lhe outro mais novo.	10, 16, 20

Tabela 30. Definição de “refrescar” no CEnoLEX

No entanto, no CEnoTEXT, há diferentes exemplos que indicam uma diferente polissemia associada e que não consta atestada no CEnoLEX.

Neste exemplo, “refrescar” apresenta uma polissemia de língua corrente, que não apresenta novidade.

67. [14153] Aroma frutado vivo, com notas de citrinos, tudo bem exposto e claro. Na boca, a presença é simples, com destaque para a acidez e o caráter revigorante. Um vinho feito para **refrescar**.

Nos exemplos seguintes, observa-se que “refrescar” não é usado no sentido de explicar que tenha sido adicionado outro vinho, mas sim como efeito da presença da

“acidez”. Aliás, o que se verifica em quase todos os exemplos de “refrescar” à exceção do anterior é a relação entre a existência de “acidez” que tem como efeito refrescar ao nível gustativo o vinho em questão.

68. [01161] Floral com notas de moscatel e laranjeira, tudo na dose certa e sem exageros. Boa proporção na boca, está com uma boa acidez que corta uma certa doçura da fruta e equilibra e **refresca** o conjunto.

69. [04858] Fechado, com folhagens frescas, notas cítricas de limão maduro, frutos amarelos, minerais, boa complexidade. Fresco e macio, com acidez viva a **refrescar** o corpo médio, onde os ligeiros amargos dão garra e comprimento.

Observa-se também, que os 55 verbos com mais de 100 ocorrências no CEnoTEXT não são mencionados pelos especialistas no CEnoLEX, mesmo quando apresentam usos e polissemias próprias na ASE, como são o caso de “crescer”, “abrir”, por exemplo, já anteriormente apresentados.

### 5.3.2. NOMES

Relembrando os dados obtidos no CEnoLEX relativamente a descritores nominais, foram apenas atestados 519 nomes e no CEnoTEXT foram identificados 2 919 nomes.

Confrontando esses verbos com os dados do CEnoTEXT, constatou-se que os seguintes nomes surgiam no CEnoTEXT:

- “boca” com 14 633 ocorrências;
- “aroma, acidez, tanino” com mais de 5 000 ocorrências;
- “corpo, cor, frescura, vegetal, nariz, madeira, mineral, tinto, estilo, estrutura, equilíbrio, sabor, chocolate, nota, complexidade, intensidade, carácter” com 1 000 a 4 999 ocorrências;
- “vivo, baunilha, tanino, maçã, persistência, compota, cacau, verde, rubi, silvestre, cereja, laranja, profundidade, gordo, limão, granada, pêssego, groselha” com 500 a 999 ocorrências;
- “pimenta, framboesa, seca, rebuçado, mel, terra, violeta, folha, tabaco, sólido, químico, morango, rosé, tinta, personalidade, vivacidade, mato, resina, vigor,

percepção, tangerina, nervo, menta, pêra, marmelo, pólvora, manga, rosa, riqueza, raça, sílex” com 100 a 499 ocorrências.

Os restantes 488 descritores de categoria nominal não foram encontrados no CEnoTEXT.

É possível ainda fazer um conjunto de contrastes que podem ajudar a decidir que outros candidatos podem ser selecionados. Apresenta-se a título de exemplo alguns casos de candidatos que demonstram as possibilidades de análise contrastiva:

- Se os descritores no CEnoLEX de unidades complexas estão atestados e se há outros candidatos que pela frequência de ocorrências, deviam ser considerados a figurar no DEASE.
  - No CEnoLEX surgem atestados “frutos vermelhos, frutos secos, frutos tropicais, frutos de polpa branca, frutos de caroço” e “silvestre”, mas não “frutos silvestres”. No CEnoTEXT surgem com mais de 100 ocorrências “frutos vermelhos, frutos secos, frutos tropicais, frutos silvestres, frutos pretos, frutos amarelos”, já “frutos de polpa branca” ocorrem apenas 31 vezes e “fruto(s) de caroço” ocorrem 48 vezes.
  - Por uma questão de coerência metodológica, pode-se sugerir que deve ser considerada a inserção dos candidatos “frutos pretos, frutos amarelos”.
  
- Se os candidatos no CEnoTEXT apresentam contextos de uso que não são encontrados definidos nos descritores do CEnoLEX.
  - No CEnoLEX “nota” é definido como:

	<b>Definição</b>	<b>Obra(s) onde ocorre</b>
<b>1</b>	Matiz aromático subtil, apreendido durante a prova de um vinho.	10, 16, 20

Tabela 31. Definição de “nota” no CEnoLEX

- No CEnoTEXT são encontrados vários exemplos onde “nota” surge, também, associado ao gustativo (sabor ou ao retrogosto) como apresentado anteriormente na secção 2.2.4..
  - Por uma questão de coerência metodológica, pode-se sugerir que deve ser considerada a inserção do candidato “nota” com uma diferente polissemia, para dar conta deste uso.
- 
- Se candidatos que surgem no CEnoTEXT (com frequência superior a 100 ocorrências) e não ocorram no CORLEX e não estejam atestadas no CEnoLEX.
    - Como o candidato “mineralidade”.
    - Por uma questão de coerência metodológica, pode-se sugerir que deve ser considerada a inserção do candidato “mineralidade”.

Para esta análise, consideram-se candidatos a ULE todos os descritores que (i) estão no CEnoLEX e que têm mais de 100 ocorrências, (ii) têm diferentes polissemias das atestadas no CEnoLEX.

### 5.3.3. ADJETIVOS

Relembrando os dados obtidos no CEnoLEX relativamente a descritores adjetivais, foram apenas atestados 429 adjetivos e foram identificados 2 481 adjetivos no CEnoTEXT.

Registou-se que os seguintes adjetivos surgiam no CEnoTEXT:

- “leve, fresco, longo, elegante, seco, cheio, maduro, suave, intenso, floral, equilibrado, grande, encorpado, fino, ligeiro, redondo, macio, limpo, delicado, doce, complexo, firme, jovem” com 1 000 a 5 000 ocorrências cada;
- “cremoso, expressivo, discreto, franco, excelente, concentrado, fumado, frutado, fechado, saboroso, harmonioso, rico, profundo, cítrico, perfumado, fácil” com 500 a 999 ocorrências cada;
- “estruturado, vigoroso, denso, refrescante, forte, sedutor, subtil, guloso, poderoso, tostado, pronto, opaco, pleno, sofisticado, generoso, novo, sedoso, picante,

untuoso, pesado, quente, robusto, plano, vinoso, mentolado, rústico, potente, requintado, salino, torrado” com 100 a 499 ocorrências cada.

Nenhum dos restantes 2 052 descritores presentes no CEnoLEX foram encontrados no CEnoTEXT;

Na observação dos dados, é evidente que alguns candidatos a ULE apresentam desafios interpretativos. À semelhança do que foi feito para os nomes, também foi verificado um conjunto de adjetivos que foram alvo desta análise. No entanto, para os adjetivos o foco foi observar os candidatos no CEnoTEXT e verificar se as polissemias apresentadas no CEnoLEX se verificavam ou apresentavam lacunas.

Um candidato como “mineral” é definido no CEnoLEX como apenas associado à modalidade olfativa:

	<b>Definição</b>	<b>Obra(s) onde ocorre</b>
<b>1</b>	Conjunto de odores que evocam os de alguns minerais (pederneira) ou de algumas rochas (giz, pedra de amolar, silício, tufo, grafite).	7, 17
<b>2</b>	qualificativo muito usado e muito discutido, genérico para aromas estranhos aos espectros vegetal, animal e da fruta, e que antes lembram minerais ou produtos químicos.	17
<b>3</b>	Aroma característico de alguns vinhos brancos de grande qualidade e finesse e que lembram o cheiro do siléx quando raspado ou de algumas rochas (giz, pedra de amolar, silício, tufo, grafite).	10, 16, 20

Tabela 32. Definição de “mineral” no CEnoLEX

Os dados da secção anterior demonstram que há também “mineral” associado à modalidade gustativa (pelo retrogosto, sabor).

Um candidato como “cheio” é definido no CEnoLEX como:

	<b>Definição</b>	<b>Obra(s) onde ocorre</b>
<b>1</b>	Carácter agradável dum vinho.	26
<b>2</b>	Dotado de qualidades gustativas marcantes e de boa amplitude.	28
<b>3</b>	o mesmo que encorpado	2

4	Termo usado para designar um vinho com muito corpo, de estrutura forte e com carácter. Diz-se de um vinho pujante, pastoso, espesso e que enche a boca.	8, 12
5	vinho amplo, equilibrado, denso, que enche a boca.	25

Tabela 33. Definição de “cheio” no CEnoLEX

Há ocorrências no CEnoTEXT que comprovam estas definições:

70. [16575] Um aroma com algo de tradicional, fenólico, fumado, com fruto sóbrio, muito bom tanino na boca, seco e com alguma altivez, **corpo médio a cheio**, um tinto com muito carácter.
71. [19776] Muito atractivo de aroma, ameixa, compotas, especiarias, um toque floral delicado. Revela **corpo cheio** e taninos acetinados e redondos, acidez q.b. a dar frescura a um vinho maduro e sumarento.

No entanto, há outras ocorrências que diferem das definições dadas:

72. [16564] Um aroma bastante tostado e especiado, fruto preto bem maduro, volume e largura de prova, taninos musculados, equilibrado e algo vigoroso, com final longo e **cheio**. Um excelente companheiro de mesa.
73. [21203] O tom acentuado da cor e as notas tropicais de toranja, laranja e ananás no aroma, dizem-nos de imediato que é um branco para dias frios. Muito **cheio** na boca com um perfil maduro e bastante apelativo. Exuberante, bem feito e capaz de muito boa prova. Merecerá copos largos que o façam respirar.
74. [17634] Trincadeira, Aragonês, e Alicante Bouschet. Evolução clara mas com muita qualidade, nota de couro, fruto vermelho, algum pimento, conjunto clássico. Prova de boca com sabor, perfil seco e leve, todo **cheio** de carácter e complexidade. Em óptimo momento de consumo.

Nestes casos, enquadram-se ainda exemplos que são distintos das definições apresentadas, pois estão associados às modalidades sensoriais: visual e olfativa.

75. [20954] Mostra-se **cheio de** cor, profundo de aroma, combinando vegetais maduros, especiarias, bagas maceradas, tudo ainda fechado. Está concentrado e vigoroso na boca, mas os taninos são finos, moldados pelo corpo untuoso, o perfil polido, a fruta franca e brilhante. O conjunto denota carácter num final longo cheio de frescura atlântica.
76. [16255] Muita cor para um aroma **cheio de** fruto e vegetal com alguns drops à mistura, muito tanino apertado e bem polido, bom corpo, alguma especiaria num final longo e seco. Um tinto do tipo "duro, mas com bons modos".
77. [04455] Um tinto muito escuro com um **aroma cheio**, compacto de fruto em passa, leve fumado e tabaco. Na boca tem potência tânica, textura redonda e gorda, extraído com conta, peso e medida. Final longo a especiaria e tabaco.

Finalmente, encontram-se ainda exemplos em que se verifica a coocorrência de “cheio” e de “encorpado”, contrariando a indicação dada no CEnoLEX de que estas unidades podem ser sinónimas.

78. [01967] Cor intensa, rubi atijolada. Fruta vermelha muito madura, como cerejas e ameixas, algumas folhas de árvore. **Encorpado e cheio**, com algum peso, acidez integrada, textura glicerinada, termina com alguma doçura frutada.

Um candidato como “vivo” é definido no CEnoLEX como:

	<b>Definição</b>	<b>Obra(s) onde ocorre</b>
<b>1</b>	Diz-se dum vinho fresco, ácido e claro	26
<b>2</b>	Qualificativo que indica grande vitalidade de força e de cor. Utiliza-se também para designar um certo domínio de agradável acidez, que geralmente corresponde a um vinho leve. Numa prova, o adjectivo vivo pode resultar pejorativo quando, por exemplo, um provador diz: "Este vinho possui um bouquet no nariz, mas no paladar é demasiado vivo". Neste caso, a palavra vivo refere-se a um excesso de acidez relativamente ao equilíbrio.	10, 16, 20
<b>3</b>	vinho cuja acidez fixa actua agradavelmente no paladar, sem se acusar excessivamente.	13

<b>4</b>	Vinho dominado pelo ácido, mas de um nível aceitável que lhe confere um aspecto fresco e leve.	23
----------	--	----

Tabela 34. Definição de “vivo” no CEnoLEX

Os exemplos do CEnoTEXT que comprovam estas definições:

79. [14564] Amarelo **vivo** na cor, com reflexos esverdeados. Toranja, diospiro e nêspera surgem tanto no ataque de nariz como nos retronasais. Vinho muito coerente, cheio e encorpado, com bom comprimento de prova. No final, fica uma impressão fresca de melaço. A pedir desafio com comida.
80. [00804] Cor granada, aroma **vivo**, amplo, aberto, com fruto silvestre maduro e floral. Compacto e sólido, taninos de belo efeito e bastante frescura. Alguma irreverência mas a crescer.

Exemplos que são distintos das definições apresentadas, pois estão associados a outros aspetos ou modalidades.

81. [10299] Cor de ouro velho. Aroma fresco, com alperce, mel e couve fermentada. Requentado. Muito redondo, espicaçado por uma acidez fresca, quase citrina, salgada que torna o conjunto **vivo** e perfeito de versatilidade.
82. [19264] Cor rosada bonita, levemente aberta, discretamente frutado no aroma, contido. Muito mais expressivo de boca, com boas notas de cereja, num estilo **vivo** e seco, com agradável frescura de fruta.

Um candidato como “opaco” é definido no CEnoLEX como:

	<b>Definição</b>	<b>Obra(s) onde ocorre</b>
<b>1</b>	Diz-se de um vinho com muita cor, quase sem transparência.	27
<b>2</b>	Sem limpidez, turvo.	19

3	Turvo, velado, sem límpidez; diz-se igualmente dos vinhos que têm aroma e sabor apagados.	10, 16, 20
4	turvo; velado; sem limpidez	2, 13

Tabela 35. Definição de “opaco” no CEnoLEX

Os exemplos do CEnoTEXT que comprovam estas definições:

83. [20610] **Opaco** na cor, com notas de fruta em passa, alfarroba, chocolate, frutos pretos com um traço mineral que lhe dá seriedade. Imponente e concentrado, com textura sedosa pontuada por taninos picantes, termina apimentado e longo.
84. [00313] cor rubi **opaca**. aromas terciários de boa evolução em garrafa com grande complexidade (balsâmicos, pinhal, sândalo, caixa de tabaco). na boca é fresco, revelando uma excelente acidez. taninos redondos, elegantes e muito bem integrados no conjunto. final longo e persistente.
85. [02754] Mostra um aroma onde se nota muita barrica, concentrado mas não **opaco**, químico, com frutos pretos de muito boa qualidade, tudo apontando para uma boa vida em cave. O equilíbrio na boca é notório e deixa muito boa impressão pela excelente harmonia de conjunto. Merece claramente algum tempo de garrafa.
86. [21494] É um vinho de porte aristocrático, profundo e opaco, que alia uma faceta musculada e viril à sutileza que reconhecemos aos grandes. É intenso e explosivo na boca, fino e elegante. Ao mesmo tempo, austero. Um vinho muito completo.

Esta análise de comparação entre o que é registado nas diferentes definições atestadas no CEnoLEX permite, por um lado, verificar o grau de concordância entre os autores e, por outro, pesquisar no CEnoTEXT os usos desses candidatos e verificar se todas as polissemias estão descritas.

Para esta análise consideram-se candidatos a ULE todos os descritores que (i) estão no CEnoLEX e que têm mais de 100 ocorrências, (ii) têm diferentes polissemias das atestadas no CEnoLEX.

#### 5.3.4. EMPRÉSTIMOS

Foram apenas atestados 33 empréstimos no CEnoLEX, sendo detetados 247 no CEnoTEXT. Confrontando os dados, verificou-se que as seguintes formas ocorriam no CEnoTEXT:

- *vintage*, com 124 ocorrências
- *finesse*, com 55 ocorrências
- *bouquet*, com 22 ocorrências
- *degorgement*, com 6 ocorrências
- *perlage*, com 4 ocorrências
- *petillant*, com 3 ocorrências;
- *brut, sur lie*, com 2 ocorrências cada;
- *brettanomyces, manzanilla*, com apenas com uma ocorrência cada.

Os 22 empréstimos atestados no CEnoLEX (*abocado, bee-wings, bouchonné, bourru, bruit, crémant, demi-sec, dolce, dry, extra brut, fine, flat, jambe, lias, macela, moelleux, mousseux, moustillant, nature, racé, vendanges tardives, vin d'une nuit*) não ocorrem no CEnoTEXT.

Observa-se assim que, no CEnoLEX, os autores das várias obras atestam poucos empréstimos, e que os que são atestados são de pouca ou nenhuma frequência no CEnoTEXT.

#### 5.3.5. SÍNTESE DA ANÁLISE CONTRASTIVA

A análise contrastiva permite:

- (i) identificar que descritores do CEnoLEX os especialistas consideram relevantes constarem em obras lexicográficas;
- (ii) verificar quais descritores estão (e quais não estão) em uso no CEnoTEXT;
- (iii) analisar as definições atestadas para os descritores no CEnoLEX
- (iv) observar quais definições são coincidentes e quais são discrepantes entre si (para determinar o nível de variabilidade nas definições);
- (v) comprovar se a(s) polissemia(s) registadas o seu uso no CEnoTEXT;

- (vi) identificar que candidatos estão no CEnoTEXT e não surgem no CEnoLEX;
- (vii) encontrar através da observação dos seus contextos que outras polissemias podem ser acrescentadas.

O que se verifica, em primeiro lugar, é um desfazamento entre o que está atestado no CEnoLEX e a realidade de usos, que se verifica no CEnoTEXT, como se pode observar na Figura 26.

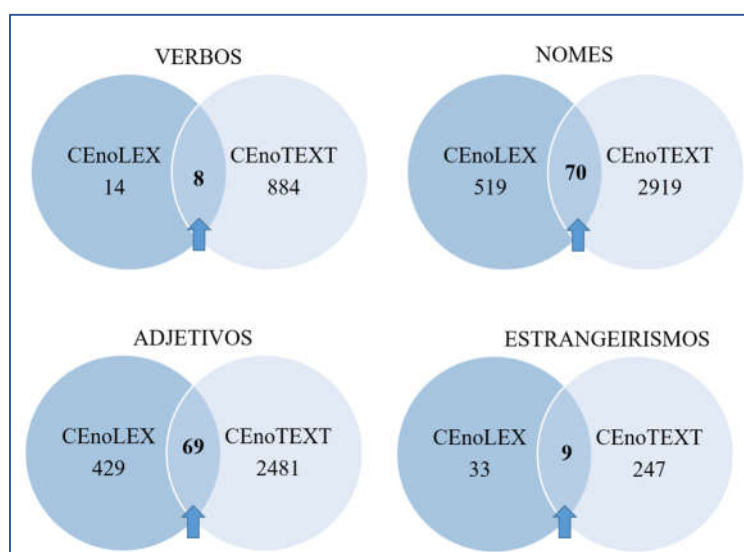


Figura 26. Diagrama Venn do número de UL de cada categoria gramatical no CEnoLEX e no CEnoTEXT

Em segundo lugar, há lugar à constatação de que existe variabilidade de definições de um mesmo descritor (aspeto já anteriormente mencionado na secção 4.2.2.1.).

O resultado deste contraste permite obter uma listagem de candidatos a ULE que pode ser utilizada como primeira nomenclatura de descritores a apresentar aos especialistas para validação.

#### 5.4. ANÁLISE POR MODALIDADE SENSORIAL

Considerando o domínio da ASE, a modalidade sensorial é um dos aspetos mais relevantes para a análise semântica das ULE. No contexto do presente trabalho, esta análise foi elaborada considerando a distribuição das UL no CEnoTEXT, com base numa

análise fatorial e tendo em conta as modalidades sensoriais. Nesta subsecção, é apresentado o trabalho desenvolvido.

A razão para contemplar esta análise prende-se com o facto de nenhuma das obras recolhidas no CEnoLEX associar de forma clara ao descritor apresentado a modalidade sensorial este pode estar associado, para facilitar a sua organização dentro do domínio da ASE. Aspeto que outras obras internacionais já contemplam e prova ser do agrado dos utilizadores. No entanto, a forma como essas obras associam o descritor a uma das modalidades sensoriais não é explicitada de forma clara nas obras.

Neste trabalho partiu-se da intuição de segmentar o texto da nota de prova em secções, para verificar se seria uma forma possível de comprovar que podem existir descritores que surgem apenas numa das modalidades ou que podem surgir em mais do que uma modalidade. E se surgem em mais do que uma modalidade se há particularidades e quais são que podem ajudar a dar pistas para a categorização da modalidade sensorial.

Nesse sentido, considerou-se, observar a lista de candidatos obtida da análise categorial e da análise contrastiva e fazer uma análise às definições e consoante o que é mencionado associar uma modalidade sensorial ao descritor. Mas, para ir mais além, ou para tornar a decisão mais robusta e sólida, considerou-se acrescentar um passo final na metodologia e fazer uma análise sensorial. Validando a anotação manual feita para cada descritor com os resultados da análise de coocorrências e de análise fatorial e ter dessa forma uma noção mais real da notação de modalidade sensorial nos descritores.

Como mencionado na secção 4.2.3.1, a nota de prova tem normalmente a seguinte configuração: inicia-se com a descrição visual, segue-se a descrição olfativa e termina com a descrição gustativa (que envolve a descrição tátil e a descrição do conjunto). Partindo desta configuração, procedeu-se à divisão em três blocos das 21 495 notas de prova que compõem o CEnoTEXT. Decidiu-se segmentar a nota de prova em secções para criar uma lista para cada modalidade sensorial (visual, olfativa, gustativa), que continha a porção de texto correspondente, devidamente identificada e etiquetada, como se apresenta na Figura 27. A categorização de partes do texto em secções é essencial para

a análise fatorial<sup>61</sup>, na medida em que permite ter em conta as ocorrências das UL em cada secção, ou seja modalidade.

<b>Código</b>	<b>00573</b>
<b>Origem</b>	ENOTECA_N.92_2014
<b>Entidade</b>	VINHA GRANDE 2011 DOC Douro
<b>Denominação</b>	Tinto
<b>Tipo</b>	2011
<b>Cor</b>	ENOTECA_N.92_2014
<b>Ano de colheita</b>	VINHA GRANDE 2011 DOC Douro
<b>Nota de prova</b>	Cor rubi profundo. Aromas intensos de frutos vermelhos (amora, groselha) e pretos (ameixa) com notas balsâmicas de cedro e caixa de tabaco. Na boca apresenta boa acidez e equilíbrio, taninos firmes e boa estrutura. Final longo.

↓

<b>Código</b>	<b>V0573</b>
<b>Fonte</b>	ENOTECA_N.92_2014
<b>Denominação (do vinho)</b>	VINHA GRANDE 2011 DOC Douro
<b>Categoria</b>	Tinto
<b>Data de Colheita</b>	2011
<b>Nota de prova original</b>	Cor rubi profundo. Aromas intensos de frutos vermelhos (amora, groselha) e pretos (ameixa) com notas balsâmicas de cedro e caixa de tabaco. Na boca apresenta boa acidez e equilíbrio, taninos firmes e boa estrutura. Final longo.
<b>Mod. Visual</b>	Cor rubi profundo.
<b>Mod. Olfativa</b>	Aromas intensos de frutos vermelhos (amora, groselha) e pretos (ameixa) com notas balsâmicas de cedro e caixa de tabaco.
<b>Mod. Gustativa</b>	Na boca apresenta boa acidez e equilíbrio, taninos firmes e boa estrutura. Final longo.

Figura 27. Exemplo da divisão em secções relativas às modalidades sensoriais

A análise categorial permite um entendimento das UL da sua frequência e das suas polissemias associadas no domínio da ASE (entre outras: distribuição de frequência, associações semânticas, relações de oposição, tendências de uso, identificar palavras-chave, similaridades, visualizar dados). A Análise Fatorial é um método estatístico para resumir dados complexos através da sua visualização num gráfico e é, comumente, usada na análise de dados para simplificar conjuntos de dados complexos, nos quais há muitas variáveis a serem processadas. O método reduz as numerosas linhas e colunas de dados

---

<sup>61</sup> O programa Lexico3 pede como requisito a divisão em partições (secções/modalidades) do *corpus* para uma melhor apresentação dos dados relativos à Análise Fatorial.

(daí ser multidimensional) a um pequeno conjunto de variáveis subjacentes que explicam a essência do *corpus*.

No caso do CEnoTEXT, isto traduz-se pela existência de centenas de *word types* (linhas) distribuídas por três modalidades (colunas) que são reduzidas a dois fatores, que correspondem às duas maiores tendências lexicais dentro do vocabulário. Essas tendências são consideradas como “super-variáveis” que contrastam o léxico. Como o método analisa todas as correlações entre linhas e colunas, os fatores identificados refletem a coocorrência de tipos de palavras nos diferentes tipos de secção. Ao fornecer indicações das *word types* mais usadas em cada modalidade, os resultados da Análise Fatorial confirmam os de outros métodos, como coocorrências, segmentos repetidos e concordâncias.

Neste trabalho, pretende-se identificar a distribuição das UL consoante a modalidade sensorial, bem como determinar quais as UL que atraem outras UL de acordo com especificidades identificadas, extraindo mais características das UL analisando os dados de uma perspectiva diferente das anteriores.

O Lexico3 é um programa que permite observar as concordâncias, o inventário dos segmentos repetidos, inventários distribucionais, especificidades de formas simples ou de conjuntos de formas, e fazer uma análise fatorial, uma vez que se baseia em Estatística Textual através do cálculo de Especificidades (*Calcul des spécificités*, Lafon 1984) e da Análise de Palavras-Chave (*Keyword Analysis*, Scott & Tribble 2006)).

#### 5.4.1. COOCORRÊNCIAS

Partindo do trabalho inicial de Martinez e Barbosa (2018), foi feita uma análise das coocorrências no CEnoTEXT e identificou-se as formas que mais contribuem para o texto e que têm tendência para atrair outras formas, de acordo com as probabilidades estatísticas. Quando estas surgem associadas a uma rede de unidades, há sempre unidades que funcionam como polos ou como âncoras (que têm um forte poder de atração lexical e desempenham um papel relevante para a compreensão da estruturação do texto) e que criam uma malha do léxico do domínio em causa. Permitem observar com clareza quais

as UL que potenciam a criação de um *thesaurus* distribucional de coocorrências para uma determinada domínio.

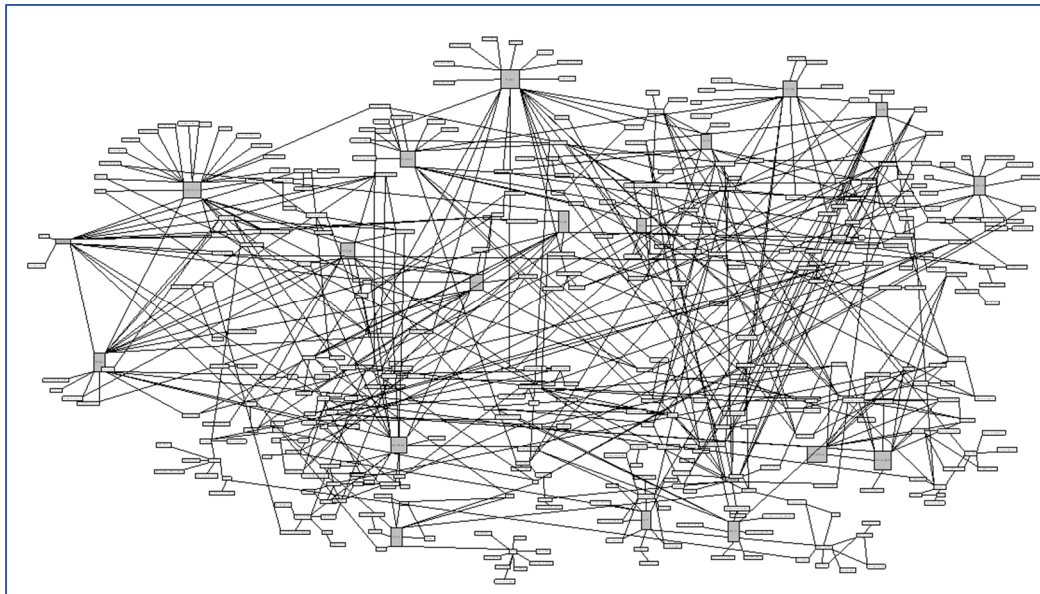


Figura 28. Imagem abstrata da rede de coocorrências retirada de Martinez e Barbosa (2018)

Este trabalho permitiu verificar que havia um conjunto de 300 formas (*word types*) com alta produtividade de coocorrências que ajudavam a estabelecer evidências de associações de formas, quer fossem associações positivas ou negativas.

Seguindo a abordagem de Greffentette (1994), a análise de coocorrências de segunda ordem (coocorrências de coocorrências) pode identificar conjuntos de “sinónimos” (palavras que aparecem no mesmo tipo de contexto que a *keyword*). Uma extensão dessa lógica seria explorar coocorrências negativas para uma palavra-chave e detetar suas coocorrências de segunda ordem para construir conjuntos de antónimos dentro do vocabulário de um domínio.

A ideia de que pode haver léxico específico em cada modalidade sensorial (visual, olfativa, gustativa) pode ser medido por meio da especificidade estatística. A escolha do

modelo hipergeométrico<sup>62</sup> permite ter em consideração a dimensão global do *corpus* e a dimensão de determinados subconjuntos do *corpus* (cada modalidade específica) e permite observar a frequência global e a frequência local e extrair formas de cada subconjunto do *corpus*, para observar que unidades surgem em cada modalidade mais ou menos consoante o que é esperado pela probabilidade.

Das 300 formas apresenta-se um excerto, Tabela 36, com as 20 formas mais frequentes que apresentam cada uma delas mais do que 5000 ocorrências.

N.	Type	Freq.	N.º	Type	Freq.
1	e	40 548	11	bem	7 377
2	de	24 747	12	final	7 013
3	com	23 712	13	acidez	7 006
4	muito	14 318	14	fruta	6 726
5	boca	11 815	15	fruto	6 494
6	na	11 091	16	no	5 429
7	a	10 887	17	taninos	5 401
8	um	10 315	18	boa	5 327
9	aroma	7 993	19	mas	5 315
10	notas	7 816	20	o	5 003

Tabela 36. As formas mais frequentes (frequência  $\geq$ 5000 ocorrências).

Através do modelo hipergeométrico compara-se a frequência global de uma formas por exemplo, “brilhante” ou “acidez” (respetivamente, com 309 e 7 006 ocorrências) com sua frequência local no subconjunto (respetivamente, com 234 e 2). Ou seja, “acidez” surge no *corpus* com 7006 ocorrências e no subconjunto relativo à modalidade com visual

---

<sup>62</sup> O modelo hipergeométrico fornece uma medida da especificidade de cada unidade num subconjunto contextual, através de cálculos sobre a frequências absolutas e relativas de uma unidade em relação a outras num determinado conjunto de dados (*corpus*). Para saber mais consultar: <https://www.statisticshowto.com/hypergeometric-distribution-examples/>

com 2 ocorrências. É feito um cálculo de um índice<sup>63</sup> entre 2 e 50<sup>64</sup> que fornece uma medida da especificidade de cada unidade num subconjunto contextual. Determina-se que “brilhante”, com um índice positivo de 50 (+50), tinha uma probabilidade de 1/10<sup>50</sup> de aparecer 234 vezes em descrições visuais, enquanto “acidez”, com um índice negativo de menos 50 (-50), tinha uma probabilidade igualmente muito baixa de aparecer apenas 2 vezes em descrições visuais quando possui uma frequência global de 7 006.

Através daquilo que se designa por valência lexical – que é a propensão de uma ou mais *keywords* de atrair outras formas – é possível medir a atração de coocorrências dessa(s) *keyword(s)* num determinado *corpus*. Por exemplo, a *keyword* “boca” surge com uma frequência de 11 815, tendo sido estabelecidas 4 538 coocorrências relevantes, das quais 2 259 são positivas e 2 279 são negativas.

Quando as formas são classificadas de acordo com seu índice de especificidade, obtém-se uma lista das unidades mais representativas para cada categoria; tanto aquelas que são privilegiadas (índices positivos), como as que são mais evitadas (índices negativos).

De acordo com esta análise, na modalidade visual verifica-se uma tendência para mencionar cores, tonalidades, claridade (“reflexo, salmonada, granada, escura, pálido”, etc.) e uma ausência de características não visuais (“fruta, frescura, corpos, seco, taninos, acidez”, etc.). Estas últimas seriam apresentadas somente noutras modalidades (olfativa ou gustativa).

---

<sup>63</sup> O cálculo é feito através da seguinte fórmula:

$$\Pr(X = k) = \frac{\binom{K}{k} \binom{N-K}{n-k}}{\binom{N}{n}}$$

, onde cada parêntesis contém uma combinação de cálculos, onde se compara o tamanho do *corpus* (N), o subconjunto (n), a frequência global (K) e a frequência local (k).

<sup>64</sup> Onde 2 e 50 correspondem ao número de zeros antes da vírgula decimal e que são apresentados em posição expoente para facilitar a leitura. Este conceito da notação científica ou exponencial pode ser expresso de diferentes formas, por diferentes autores, mas o número de zeros indica a magnitude da “pequenez” ou da “precisão”.

Além de constituírem o léxico mais utilizado na modalidade visual, algumas das unidades mais utilizadas neste subconjunto de textos (secções relativas à modalidade visual) descrevem aspetos ou tipos de vinho mais propensos a serem descritos de maneira visual (“castas, granada, alfocheiro, trincadeira”).



Os índices extremos (+50 e -50) apresentam o antagonismo na especificidade das unidades: algumas são privilegiadas, outras são totalmente evitadas. No caso das especificidades positivas, essas são as formas mais características para cada categoria. Ao contrário das unidades gramaticais (“a, de, o”, etc.), as unidades com especificidades positivas são as mais discriminativas do ponto de vista estatístico para cada modalidade sensorial.

Assim, utilizando as secções dos textos das notas de prova consegue-se obter uma percepção do léxico utilizado em cada modalidade sensorial.

<b>N.</b>	<b>visual</b>	<b>Freq.</b>	<b>olfativa</b>	<b>Freq</b>	<b>gustativa</b>	<b>Freq</b>
<b>1</b>	cor	4183	aroma	16 902	boca	13072
<b>2</b>	rubi	741	nota	15770	muito	9832
<b>3</b>	touriga	698	fruto	13762	final	8291
<b>4</b>	muito	595	muito	12512	acidez	7892
<b>5</b>	nacional	579	fruta	10908	tanino	6299
<b>6</b>	granada	544	bem	5310	bem	5520
<b>7</b>	fazer	535	leve	4696	bom	4302
<b>8</b>	tinto	467	algum	4130	corpo	4289
<b>9</b>	amarelo	433	vegetal	3972	fresco	2602
<b>10</b>	arinto	427	maduro	3500	frescura	2366

Tabela 37. Os 10 lemas mais frequentes em cada modalidade sensorial

Tendo em conta que a “boca” é a primeira UL não funcional com mais frequências a surgir na Tabela 37 na modalidade gustativa, esta será usada como exemplo. Observa-se que “boca” surge na modalidade gustativa e apresenta um conjunto de especificidades positivas e negativas, como se apresenta na tabela 38.

<b>Posição</b>	<b>Type</b>	<b>Positiva</b>	<b>Posição</b>	<b>Type</b>	<b>Negativa</b>
1	acidez	+100	1	fruta	-100
2	final	+100	2	aroma	-100
3	taninos	+100	3	madura	-100
4	corpo	+100	4	frutos	-100
5	bom	+100	5	especiarias	-100
6	macio	+100	6	cor	-100
7	volume	+100	7	minerais	-100
8	redondo	+82	8	flores	-100
9	frescura	+68	9	preta	-100
10	doçura	+63	10	vermelha	-87
11	estrutura	+62	11	floral	-86
12	sabor	+60	12	citricos	-85
13	textura	+52	13	vegetais	-84
14	secura	+48	14	nariz	-73
15	viva	+47	15	barrica	-72
16	longo	+46	16	tostados	-70
17	mediano	+46	17	aromática	-65
18	equilíbrio	+40	18	silvestres	-64
19	sabroso	+35	19	folhas	-60
20	cheia	+32	20	chocolate	-57

Tabela 38: Especificidades positivas e negativas das coocorrências de “boca”

Na Tabela 38, excluídas as palavras funcionais apresentam-se os índices hipergeométricos que são suficientes para demonstrar a relevância probabilística das unidades que coocorrem com “boca”, seja pelo excesso ou déficit estatístico. Os índices de [+/-]100 indicam improbabilidades extremas para essa coocorrência e fornecem uma lista de termos privilegiados ou, inversamente, evitados em torno da unidade “boca”.

Com esses resultados, é possível compreender que, na modalidade gustativa, a partir da “boca”, as unidades que têm especificidades positivas são “acidez, taninos, corpo, suavidade, redondeza, frescor e doçura”. Inversamente, com especificidades negativas relativamente a “boca”, surgem “frutas, maturidade, especiarias, flores, cítricos, vegetais e chocolate”. O mesmo é dizer que estas unidades não são usadas quando se usa “boca”. Observando a imagem da análise de “boca”, torna-se mais fácil compreender as especificidades positivas (Figura 30).



Na Figura 30, observa-se o universo de coocorrências de “boca”. Num momento imediato, a tendência imediata e natural é olhar para “boca” e construir um caminho de unidade em unidade e encontrar e detetar unidades complexas ou frases (por exemplo, “boca boa a dar frescura”). No entanto, o método de cálculo envolvido aqui é sempre entre coocorrências de duas unidades. A associação entre mais do que duas palavras é possível, mas a imagem não representa isso. A Figura apresenta, assim, malhas lexicais do tipo *constelações lexicais*, onde certas unidades podem logicamente associar-se num determinado contexto, mas não obrigatoriamente criar estruturas. A Figura apresenta um mapeamento do léxico, que serve de guia para a aplicação de outros métodos (como procurar segmentos repetidos, concordâncias, entre outros). Assim, observa-se um conjunto de unidades que são as mais específicas positivamente, como “prova, na, macio, redondo, corpo, final, bom, boa, taninos”, que funciona como uma espécie de malha lexical mais restrita e imediata, *vs.* outra malha lexical, também ela específica e positiva, mas mais distante criada a partir de “aroma”. Assim, “boca” e “aroma” são as unidades polo/âncoras que funcionam na imagem como principais atratoras de coocorrências criando as suas constelações lexicais, onde “acidez” funciona como unidade de transição, que por acaso ocorre em ambas as constelações.

#### 5.4.2. ANÁLISE FATORIAL

Ao contrário da análise de coocorrências, em que se parte de uma *keyword* específica, a Análise Fatorial busca identificar fatores subjacentes que explicam relações entre unidades, proporcionando uma compreensão mais profunda e abrangente dos dados. Além disso a Análise Fatorial permite uma análise de quantidades grandes de dados e formular hipóteses. Por exemplo, partindo do CE<sub>no</sub>LEX e da lista de candidatos a ULE é possível, de forma neutra, isto é, sem manipular os dados, colocar algumas questões.

Lembrando os dados da Tabela 37, existem 3 divisões que equivalem a 3 modalidades sensoriais predefinidas e um conjunto de dados em diferentes linhas que correspondem às formas. A Análise Fatorial permite verificar (i) como as UL se distribuem pelas modalidades, (ii) quais as associações possíveis entre elas, e (iii) quais as semelhanças de distribuição de unidades entre as modalidades.

A Análise Fatorial parte de um método estatístico *redução da dimensionalidade* (*dimensionality reduction*), que nos últimos anos se rebatizou como *Deep Learning* (associada à Inteligência Artificial), onde é recorrente rejeitar alguns dados ou algumas variáveis consideradas menos importantes, para preservar o que é mais importante, neste caso as relações entre unidades das modalidades.

Esta operação de redução permite resumir os fatores que são “super-variáveis” que resumem as variáveis de uma tabela de dados e calcular uma menor quantidade de fatores que representam os principais itens ou tendências presentes no *corpus*. Este cálculo não é manipulado por escolhas humanas, mas sim reflexo de correlações entre unidades e variáveis. Os fatores são ordenados pela quantidade de informação (“variância” nas estatística) e ajudam a compreender qual a estrutura subjacente do léxico e como este se apresenta, permitindo observar a organização por áreas, por tendências.

Na Figura 31 é apresentada uma projeção factorial das unidades mais usadas em cada modalidade sensorial, obtida através do programa Lexico3.

Para compreender a figura, apresentam-se algumas indicações:

- **Eixos:** num gráfico de análise factorial, as unidades são representadas num espaço bidimensional, marcado com linhas que se interseccionam. Na Figura 31 o eixo vertical corresponde ao fator 1 e o eixo horizontal representa o fator 2<sup>65</sup>.
- **Etiquetas/polos:** as áreas do gráfico são etiquetadas de acordo com os fatores considerados na análise. Neste caso, as etiquetas [VISUAL], [OLFATIVA] e [GUSTATIVA] identificam as diferentes modalidades.

---

<sup>65</sup> Os dois fatores resultam da redução da dimensionalidade dos dados do *corpus*, no caso relacionados com as três modalidades em análise. O fator 1 contém 58.6% da informação do *corpus* e o fator 2 representa 41,3%, o que corresponde a uma observação do que acontece em todo o *corpus*. Permitem assim a análise e a visualização dos dados bidimensionalmente.

- **Elementos da rede:** cada unidade no gráfico representa uma palavra específica. Se surgirem mais perto das etiquetas, significa que são unidades fortemente associadas a esse fator.
- **Clusters:** A proximidade de unidades a etiquetas ou a outras unidades indica que partilham características semelhantes. Unidades agrupadas ou muito próximas indicam uma forte relação e apontam para um possível padrão de comportamento distribucional.
- **Outliers:** elementos que surgem muito distantes dos eixos ou de *clusters*, que indicam unidades que são distintas das restantes.



Assim, a Figura 31 deve ser lida da seguinte forma:

- (i) no centro do gráfico, (onde os eixos se intersectam) é a zona neutra, onde ocorre o léxico comum às modalidades e onde estão as unidades neutras, ou seja, que ocorrem em todas as modalidades;
- (ii) nas margens do gráfico, posicionam-se as unidades pouco comuns de cada modalidade, ou seja, as unidades com comportamentos mais específicos de cada modalidade (*outliers*);
- (iii) a proximidade gráfica de *clusters* corresponde a proximidade lexical (ou contextual) de unidades entre modalidades.

Após uma primeira observação, pode inferir-se que há três polos dominantes, mas não coincidentes e não sobrepostos, i.e., distribuídos no espaço, que correspondem às três modalidades referidas (visual, olfativa e gustativa). Cada etiqueta reúne junto de si o que se designa como uma constelação lexical própria, com poucas interseções de vocabulário entre elas. A disposição espacial das formas comprova que existe efetivamente uma distribuição distinta consoante a modalidade em causa.

Também é perceptível uma delimitação relativamente clara entre cada constelação de cada modalidade. Isto é, o eixo vertical opõe a modalidade visual às modalidades olfativa e gustativa, e o eixo horizontal opõe a modalidade olfativa às modalidades visual e gustativa). É visível que as modalidades olfativa e gustativa estão mais próximas, pela partilha de elementos como “limpo, maduro, suave, frutado”. Pode-se inferir que estes podem ser utilizados para considerações sobre aromas ou sobre sabores.

Conclui-se, pela distribuição do léxico, que a modalidade visual é mais original quanto ao léxico e coocorrências e que a modalidade gustativa é a menos original. Para uma melhor leitura, apresenta-se cada modalidade de forma independente.

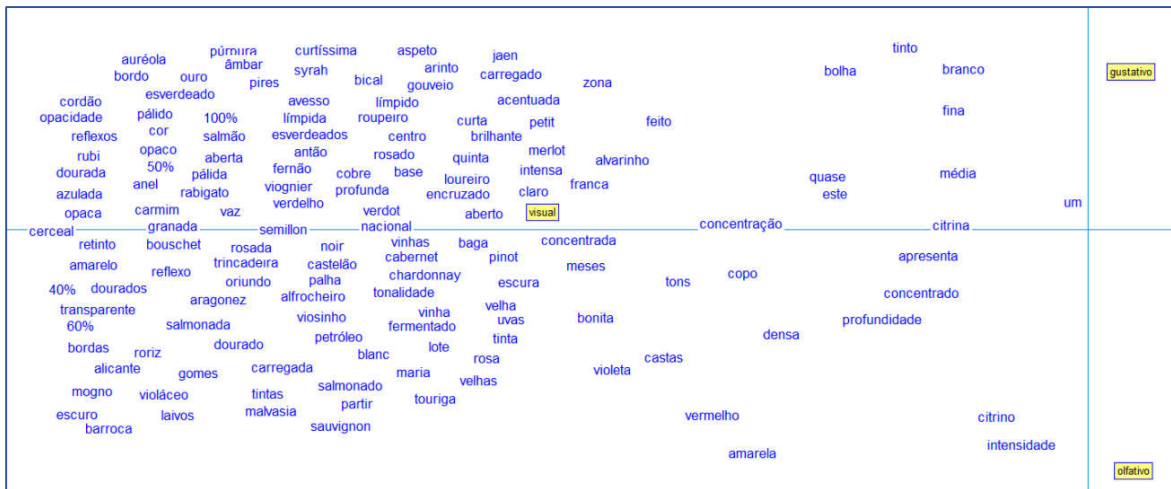


Figura 32. Projecção fatorial – modalidade visual

A Figura 32 apresenta a modalidade visual. De acordo com o explicado acima, as unidades que circundam a etiqueta [VISUAL] são as mais frequentes nesta modalidade no CEnoTEXT. Observa-se o nome das castas (“alvarinho, pinot, merlot”), que surgem quando se fala de cor, porque há relação entre a utilização da casta e as possíveis cores e tonalidades associadas, sem que, em muitos casos, a cor exata seja mencionada.

As unidades “cordão, mogno, retinto, auréola” são *outliers* da modalidade visual, ou seja, são unidades que são específicas quanto ao uso e às outras unidades que atraem.

As unidades “vermelho, amarela, citriño, intensidade” indicam uma partilha de uso entre a modalidade visual e a modalidade olfativa, enquanto que “branco, fina, tinto, bolha” indicam uma partilha entre as modalidade visual e gustativa.

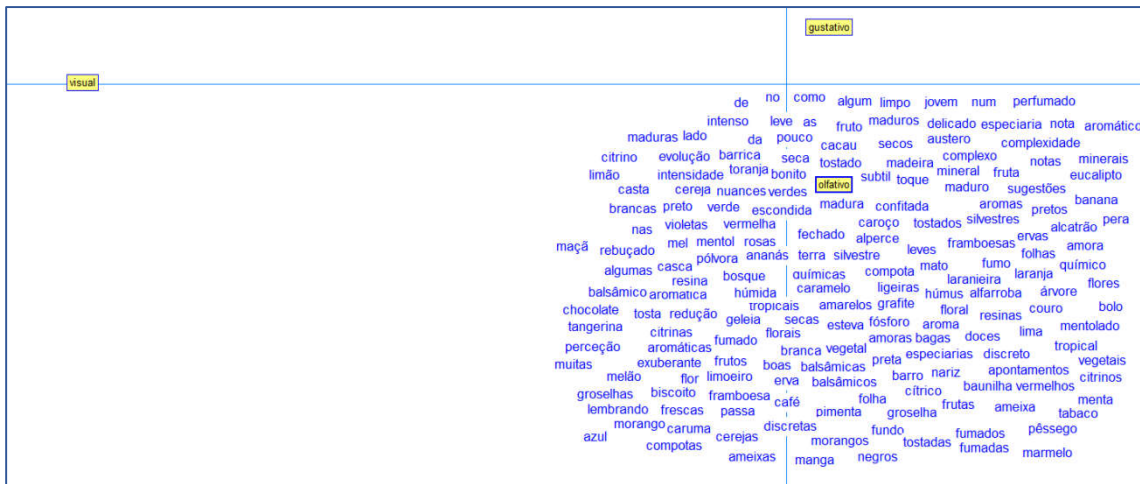


Figura 33. Projeção fatorial – modalidade olfativa

A Figura 33 apresenta a modalidade [OLFATIVA]. É interessante observar que os elementos que denotam nomes de frutas, ou de objetos do mundo real, como “marmelo, pêssego, tabaco, manga, ameixas, compotas” para invocar os aromas, se encontram na área externa dos eixos, demonstrando que são unidades típicas apenas nesta modalidade. O contrário ocorre com “fruto, especiarias, intenso, jovem”, que são partilhadas – estão na área interna dos eixos, e logo mais próxima dos outros fatores –, e logo, com maior probabilidade de serem usadas para mencionar questões olfativas, mas também gustativas.

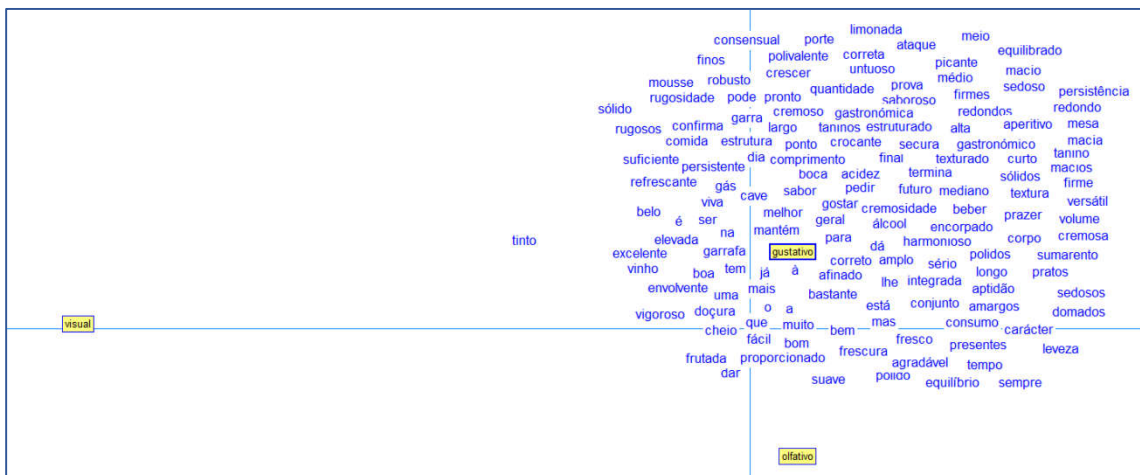


Figura 34. Projeção fatorial – modalidade gustativa

Finalmente, a Figura 34 apresenta a modalidade [GUSTATIVA]. Nesta modalidade é de realçar que “rugosos, rugosidade, firme, macio, sedoso,sólido, textura, texturado”, para citar alguns exemplos, são unidades associadas à modalidade gustativa no domínio da ASE, mas que tipicamente são associadas ao tato. Verifica-se também que estas não são partilhadas com outras modalidades.

#### 5.4.2.1. EXEMPLO DE ANÁLISE DE CANDIDATOS VERBAIS

Um exemplo da aplicabilidade de vantagens da análise anterior, é a análise dos verbos candidatos a ULE. Considerou-se pertinente analisar a sua distribuição nas diferentes secções da nota de prova (secção da modalidade visual, olfativa e/ou gustativa) para verificar se um verbo é um verbo sensorial usado apenas em relação a uma modalidade ou a mais do que uma (multissensorial).

Para efeitos ilustrativos, descreve-se de seguida a análise de alguns verbos candidatos a ULE. Recuperando os dados da análise categorial e observando as definições em dicionário de língua geral, foi possível verificar que apenas alguns verbos são imediatamente reconhecidos como sensoriais (quer no domínio da ASE, quer na língua geral), como se pode ver na Tabela 39.

<b>Modalidade sensorial</b>	<b>verbos sensoriais</b>
visual	ver
olfativa	cheirar
gustativa	beber
multissensorial	sentir

Tabela 39. Verbos sensoriais

Detetou-se, ainda, outra categoria de verbos que são multissensoriais, isto é, são verbos que podem ser associados a mais do que uma modalidade sensorial, consoante a posição que ocupam na nota de prova, como é o caso do “sentir”<sup>66</sup>. Este verbo pode surgir

---

<sup>66</sup> Aspeto já notado por Sweetser (1990:35) em diferentes línguas indoeuropeias.

associado às modalidades olfativa (*sentir um aroma*), gustativa (*sentir o sabor de algo*) e tátil (*sentir a rugosidade na cavidade bucal*):

- na modalidade olfativa

87. [10428] Denso na cor e ainda um pouco fechado nos aromas, **sentem-se** notas florais de bergamota, aqui equilibradas com fruta madura mas bastante elegante.

- na modalidade gustativa (quanto ao sabor)

88. [10874] Na boca **sente-se** uma acidez correta, um perfil gordo, com notas de fruto tropical, num conjunto bem ajustado.

- na modalidade gustativa (quanto ao tato)

89. [01338] Na boca **sente-se** um tinto muito sólido, em que os taninos arredondados quase se mastigam e a fruta de qualidade está sempre presente.

Analisando a análise fatorial relativa aos verbos sensoriais, e observando os contextos onde estes surgem no CEnoTEXT, foi possível fazer a seguinte distribuição nas modalidades sensoriais (listas obtidas através dos contrastes nas projeções fatoriais), como se apresenta na Tabela 40, onde se apresenta um excerto da lista total com os 10 formas verbais mais frequentes.

N.	Visual	Freq.	Olfativa	Freq	Gustativa	Freq
1	feito	515	mostra	1190	está	883
2	é	358	tem	1062	dar	854
3	tem	206	está	718	mostra-se	332
4	apresenta	97	dar	536	dá	325
5	mostra	82	revela	464	mostrar	192
6	apresenta-se	50	lembrando	338	sente-se	181
7	revela	45	mostrar	320	surge	167
8	mostra-se	33	mostra-se	286	gostar	158
9	está	28	apresenta	260	apresenta	152
10	partir	28	surgem	208	há	149

Tabela 40. Os 10 verbos mais frequentes em cada modalidade sensorial

Observando a totalidade das listas obtidas em cada modalidade, verificou-se que, no universo de 884 verbos, com 2 877 formas verbais correspondentes a um total de 29 901 ocorrências, obtém-se a seguinte projeção factorial das formas verbais mais usadas em cada modalidade sensorial.



Figura 35. Projecção factorial – verbos por modalidade sensorial

Num primeiro momento, observa-se a existência de três grupos, que correspondem às três modalidades referidas (visual, olfativa e gustativa) e apenas relativa a verbos com mais de 100 ocorrências, para efeitos de permitir uma visualização das formas. A disposição espacial das formas comprova que existe uma distribuição distinta das formas verbais consoante a modalidade em causa. A observação imediata é verificar que são escassas as formas verbais com mais de 100 ocorrências na modalidade visual. Pode ser uma pista que esta categorial gramatical não é tão relevante para descrever a modalidade sensorial visual, quanto as categorias nome e adjetivo. Relembrando a projeção dos dados da Figura 35, comprova-se que efetivamente a estratégia para a verbalização da modalidade visual assenta sobretudo em nomes e adjetivos. Já as modalidades olfativa e gustativa usam mais diversidade verbal.

A segunda observação é que uma vez mais as modalidades olfativa e gustativa partilham um conjunto de formas verbais. No entanto é preciso observar o uso dos verbos em contexto para verificar se se trata de um verbo sensorial, ou com uma polissemia relativa à ASE ou se remetem para uma modalidade sensorial em particular, pois o uso do verbo em si pode não ser associado a nenhuma destas situações e o uso do verbo não ter qualquer particularidade de especialidade. Assim, os dados obtidos na análise categorial e na análise contrastiva permitem olhar para os candidatos e organizá-los com a informação relativa à análise por modalidade.

A observação dos dados verbais por modalidade permite a criação de uma tabela mais ampla com informação sobre os verbos e as respetivas modalidades a que se associam.

<b>modalidade</b>	<b>verbos sensoriais</b>	<b>verbos com uma polissemia ASE/remetem para uma modalidade sensorial</b>
<b>visual</b>	<i>ver</i>	<i>mostrar, esconder, aparecer, revelar, surgir, apresentar,</i>
<b>olfativa</b>	<i>cheirar</i>	<i>abrir, perfumar, exalar</i>
<b>gustativa</b>	<i>beber, provar</i>	<i>refrescar</i>
<b>multissensorial</b>	<i>sentir</i>	<i>suavizar, notar, surgir, sugerir, misturar, juntar, lembrar</i>
<b>hedónicos</b>		<i>gostar, melhorar, misturar, impressionar, marcar</i>

Tabela 41. Distribuição de verbos por modalidades

A tabela permite observar a coluna dos verbos que são considerados sensoriais (na língua geral e na língua de especialidade) e que não levantam dúvidas sobre a sua significação (“ver, cheirar, beber, provar, sentir”) e a coluna dos verbos que possuem uma polissemia associada ao domínio da ASE (em que pode ser um verbo usado na ASE ou um verbo sensorial). Este último grupo é aquele em que se depreende que, por exemplo, um não-especialista possa ter dificuldade em usar ou interpretar alguns verbos, pois o grupo inclui verbos que são usados em língua geral e em língua de especialidade, mas com particularidades específicas.

Olhando para outros casos particulares de verbos, verifica-se que o uso de “poder”, “precisar”, por exemplo, parecem estar associados a um grau de comprometimento do especialista que faz a prova e varia consoante o tipo de verbo que utiliza que poderá ser relevante na compreensão do texto. Os exemplos ilustram estes casos.

[07143] Duas garrafas. Muito escuro, concentrado, notas de fruta madura e algum torrado, tudo volumoso e por abrir. Cheio na boca, muito corpo e estrutura, taninos polidos mas presentes. **Precisa** de tempo.

[09136] Varietal de Loureiro. Fino e muito vegetal no aroma, boas notas minerais, austero e vivo no recorte citrino. Acidez marcante, elegante na boca, é branco que **precisa** de tempo e terá longa vida em cave.

Nestes casos, há uma noção implícita de, se o vinho envelhecer, isto é, for guardado nas condições ideais ele melhorará (mas pode não melhorar, mas a culpa não pode ser atribuída a ninguém). Isto é, o especialista provou e considera que possui as características e que pode melhorar se considerada a sua sugestão técnica.

O verbo “pedir” também tem uma leitura particular de incentivo ao não consumo imediato. Ou seja, o vinho como produto não “pede” nada, mas a sugestão do especialista surge como indicação que deve ser tida em conta.

[06790] Aroma potente mas em contenção e algo fechado, muita esteva, minerais e fruto azul no fundo. Na boca dá já grande prova apesar da juventude, boa frescura, corpo cheio, fruto fino e leve vegetal seco, não muito doce. Um belo Vintage, do estilo focado e tenso, que apenas **pede** garrafa e paciência.

#### 5.4.2.2. EXEMPLO DE ANÁLISE DE CANDIDATOS À MODALIDADE VISUAL

Outro exemplo de aplicabilidade da análise fatorial é na observação de nomes e adjetivos na modalidade visual. Considerou-se pertinente analisar a sua distribuição nas diferentes secções da nota de prova, com particular interesse pela modalidade visual, para verificar que candidatos são usados nesta modalidade e que podem ajudar a determinar com melhor detalhe as submodalidades associadas na organização dos descritores.

Lembrando, o CEnoLEX possui 1 054 descritores, dos quais 164 foram manualmente associados à modalidade sensorial visual (271 na modalidade olfativa, 478 na modalidade gustativa. Contém 44 descritores associados à ASE e 97 descritores associados a mais do que uma modalidade sensorial).

Os candidatos nomes e adjetivos do CEnoTEXT que já foram apresentados na Figura 36 e uma análise mais detalhada da lista de cada modalidade, neste caso na modalidade visual, permitiram uma anotação manual da sua distribuição e chegar a uma segmentação das secções em secções menores.

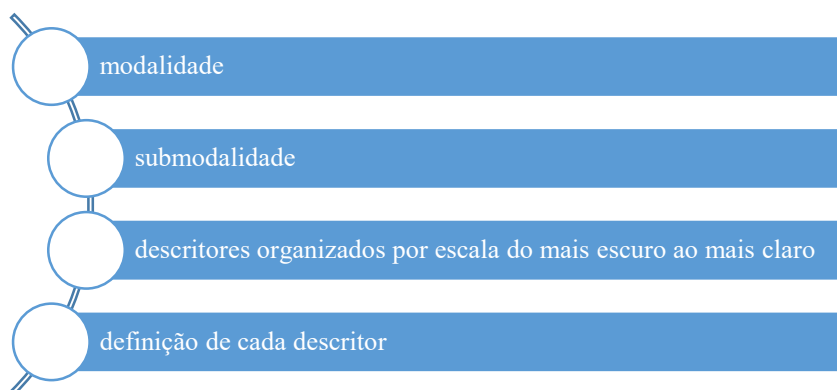


Figura 36. Esquema da segmentação da nota de prova

Este tipo de análise, apesar de morosa, no sentido em que exige um tratamento mais detalhado dos textos (a divisão em secções consoante o tipo de texto ou do domínio pode não ser muito fácil e tem de ser feita manualmente), facilita a visualização de campos lexicais, ajuda a compreender as outras análises anteriores, e permite a associação quase imediata de unidades a modalidades sensoriais.

<b>Código</b>	00573	<b>Código</b>	00573
<b>Origem</b>	ENOTECA_N.92_2014	<b>Fonte</b>	ENOTECA_N.92_2014
<b>Entidade</b>	VINHA GRANDE 2011 DOC Douro	<b>Denominação (do vinho)</b>	VINHA GRANDE 2011 DOC Douro
<b>Denominação</b>	Tinto	<b>Categoria</b>	Tinto
<b>Tipo</b>	2011	<b>Data de Colheita</b>	2011
<b>Cor</b>	ENOTECA_N.92_2014	<b>Nota de prova original</b>	Cor rubi profundo. Aromas intensos de frutos vermelhos (amora, groselha) e pretos (ameixa) com notas balsâmicas de cedro e caixa de tabaco. Na boca apresenta boa acidez e equilíbrio, taninos firmes e boa estrutura. Final longo.
<b>Ano de colheita</b>	VINHA GRANDE 2011 DOC Douro	<b>Mod. Visual</b>	Cor rubi profundo.
<b>Nota de prova</b>	Cor rubi profundo. Aromas intensos de frutos vermelhos (amora, groselha) e pretos (ameixa) com notas balsâmicas de cedro e caixa de tabaco. Na boca apresenta boa acidez e equilíbrio, taninos firmes e boa estrutura. Final longo.	<b>Cor/Tonalidade</b>	Rubi
		<b>Intensidade da cor</b>	Profundo
<b>Código</b>	V0573	<b>Mod. Olfativa</b>	Aromas intensos de frutos vermelhos (amora, groselha) e pretos (ameixa) com notas balsâmicas de cedro e caixa de tabaco.
<b>Fonte</b>	ENOTECA_N.92_2014	<b>Família(s), subfamília</b>	Frutado (frutos vermelhos: amora, groselha; frutos pretos: ameixa), madeira (árvore: cedro), vegetal (herbáceo seco: caixa de tabaco)
<b>Denominação (do vinho)</b>	VINHA GRANDE 2011 DOC Douro	<b>Intensidade</b>	Intenso
<b>Categoria</b>	Tinto	<b>Mod. Gustativa</b>	Na boca apresenta boa acidez e equilíbrio, taninos firmes e boa estrutura. Final longo.
<b>Data de Colheita</b>	2011	<b>Nervo</b>	(boa) acidez
<b>Nota de prova original</b>	Cor rubi profundo. Aromas intensos de frutos vermelhos (amora, groselha) e pretos (ameixa) com notas balsâmicas de cedro e caixa de tabaco. Na boca apresenta boa acidez e equilíbrio, taninos firmes e boa estrutura. Final longo.	<b>Harmonia</b>	Equilíbrio, (boa) estrutura
<b>Mod. Visual</b>	Cor rubi profundo.	<b>Intensidade taninosa</b>	Taninos firmes
<b>Mod. Olfativa</b>	Aromas intensos de frutos vermelhos (amora, groselha) e pretos (ameixa) com notas balsâmicas de cedro e caixa de tabaco.	<b>Fim de boca</b>	Final longo
<b>Mod. Gustativa</b>	Na boca apresenta boa acidez e equilíbrio, taninos firmes e boa estrutura. Final longo.		

Figura 37. Exemplo do processo de inserção informação relativa às modalidades sensoriais no CEnoTEXT

Na Figura 37 observa-se como o bloco inicial de texto de nota de prova (que surge no CEnoTEXT como versão original) se vai desdobrando em três secções associadas às modalidades sensoriais e finalmente pode ainda ser desdobrados em submodalidades sensoriais, aumentando o nível de detalhe de informação sensorial associado às notas e, consequentemente aos descritores envolvidos. Será, assim, possível fazer a divisão por modalidades e por submodalidades de todas as 21 495 notas de prova até se chegar ao grau de pormenor que se pretende.

A combinatória das várias análises permite um maior aprofundamento da análise individual de cada descritor, permitindo fazer uma categorização por modalidade sensorial das ULE mais detalhada, de forma mais fácil e à semelhança do trabalho que já foi feito para a língua francesa, por Gautier (2020), como se ilustra na Figura 38 para a modalidade sensorial visual.

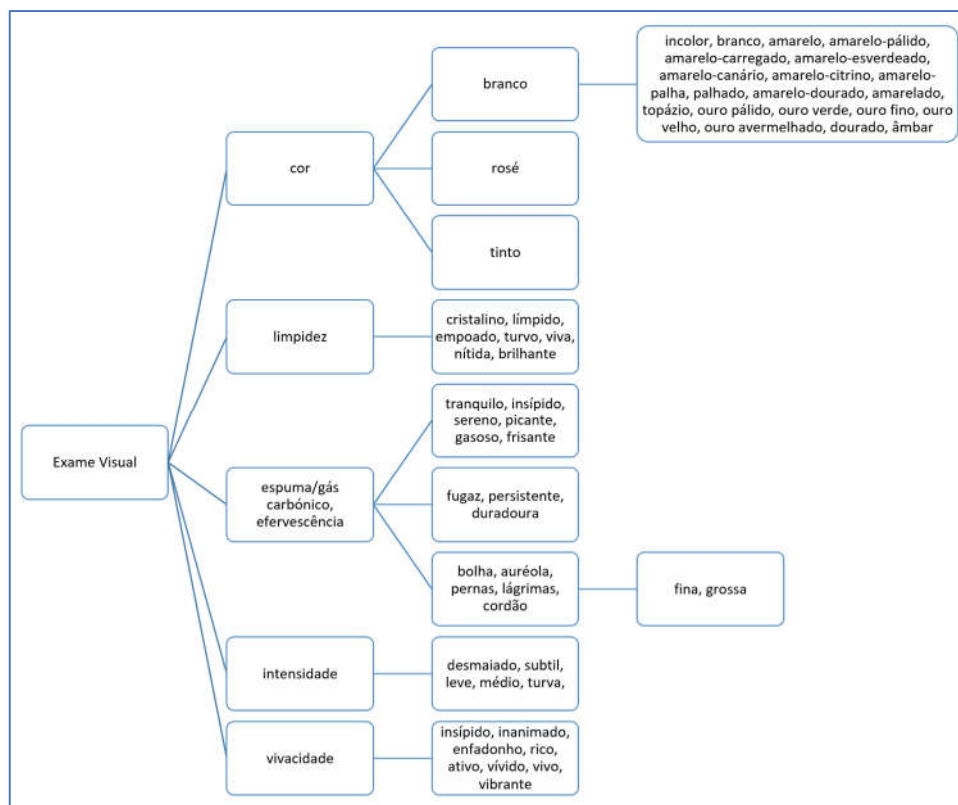


Figura 38. Possível categorização da modalidade sensorial visual

### 5.4.3. SÍNTESE DA ANÁLISE POR MODALIDADE SENSORIAL

A análise por modalidade através da análise fatorial, e da análise das coocorrências permite verificar com mais rapidez se (i) há UL que apenas surgem numa modalidade ou se em várias modalidades; (ii) se há UL que surgem em diferentes modalidades mas com contextos diferentes.

Relativamente à primeira situação, há de facto candidatos que apenas surgem numa das modalidades, que são as unidades que surgem como *outliers* nas projeções fatoriais, enquanto que as unidades que são comuns entre modalidades estão na zona neutra e central dos eixos. As projeções possibilitam uma leitura das listagens das secções das notas de prova que permite observar que a modalidade visual possui um léxico diverso das modalidades olfativa e gustativa. Quanto à segunda situação é um facto que existem várias unidades que são partilhadas por todas as modalidades, como é o caso de palavras funcionais, incluindo verbos de cópula e leves como “ser, estar, haver”, comuns em todas

as modalidades e por não serem específicos de uma das modalidades. Outro aspecto relevante e logo não específicos de uma delas.

Outro aspecto relevante é compreender que as projeções permitem uma leitura visual que facilita a observação quase imediata da dispersão lexical entre modalidades. Como se verificou há um contraste maior entre modalidade visual *versus* modalidade olfativa e gustativa e outro contraste entre olfativo *versus* visual e gustativo (de acordo com as posições que as suas malhas lexicais ocupam nas projeções. Assim, apenas pela observação lexical compreende-se que o léxico associado à modalidade visual é mais específico dessa mesma modalidade e não há tantos cruzamentos de unidades como se verifica entre olfativo e gustativo. Este é um aspecto que se aproxima da intuição de que o léxico visual tem um conjunto de particularidades, como os nomes de cor e de tonalidades, e que o léxico olfativo e gustativo, que em certa medida se cruza, usa a estratégia de mencionar um conjunto de substâncias com determinado aroma ou um determinado sabor.

Apesar do Lexico3 funcionar com formas e não com lemas, ainda assim possibilita um conjunto de observações como a análise por coocorrências, que facilita a compreensão de do que se entende como especificidades positivas, que podem ser um auxiliar para verificar conjuntos de candidatos a ULE e para determinar se há interesse em observar os contextos de uso.

#### 5.4.4. CONSIDERAÇÕES FINAIS

Neste capítulo apresentou-se o conjunto de análises realizado para tratar os dados relativos ao CEnoTEXT e, em certa medida, ao descritores do CEnoLEX. O objetivo em vista era o de extrair e tratar o léxico de especialidade da ASE e chegar à nomenclatura do DEASE.

Para chegar a um conjunto de ULE representativas do léxico da ASE para figurar no DEASE, foi considerada uma metodologia diversificada, que tem por base três análises: gramatical, contrastiva e por modalidade sensorial. A diversidade de análises e de metodologias utilizadas serviu o propósito de gerar listas de candidatos de formas distintas e concorrentes. Dentro de cada análise, foi considerado um processo

metodológico específico de seleção de unidades para construção de uma nomenclatura de um dicionário de especialidade.

Com a análise gramatical pretendeu-se, determinar as categorias gramaticais que importavam destacar, neste caso: verbos, nomes, adjetivos e empréstimos. A listagem de candidatos que se extraiu resulta de um processo (i) de comparação dos candidatos com mais de 100 ocorrências com um *corpus* de referência. Em seguida, as listagens de UL por categoria gramatical (adjetivos, nomes, verbos) foram extraídas e as frequências relativas foram calculadas e comparadas com as listagens do CEnoTEXT. Esse processo permitiu uma revisão e comparação manual, juntamente com uma análise do diferencial percentual entre os dois *corpora*. Identificou-se um conjunto de candidatos de acordo com os seguintes pressupostos: (i) aqueles que não ocorriam no CORLEX, (ii) aqueles com valores de frequência relativa mais elevados no CEnoTEXT do que CORLEX, (iii) aqueles que após uma observação das particularidades indicadoras de domínio de especialidade. Deste processo foram selecionados: 26 verbos, 219 nomes, 171 adjetivos e 2 empréstimos.

Com a análise contrastiva visou-se detetar as discrepâncias entre o que está atestado por especialistas em obras publicadas – o CEnoLEX – e compreender o que está atestado, como estava definido e o que efetivamente está em uso – os dados do CEnoTEXT. Observou-se que há um número considerável de descritores atestados que não constam no CEnoTEXT e que há descritores pouco frequentes e não coincidentes com a língua geral que não se encontram atestados em recursos de especialidade. Com este processo, foram selecionados todos os verbos, nomes, adjetivos e empréstimos que (i) estão no CEnoLEX e que têm mais de 100 ocorrências, (ii) têm diferentes polissemias das atestadas no CEnoLEX.

Com a análise por modalidade sensorial ambicionou-se identificar e categorizar as unidades obtidas das análises anteriores e verificar (i) quais UL apenas surgiam numa modalidade ou em várias modalidades; (ii) quais UL que surgiam em diferentes modalidades mas com contextos diferentes. Com o uso de listagens de coocorrências e da aplicação de análise fatorial, a atribuição de modalidade sensorial aos candidatos a ULE teve por base não só dados de frequência, mas também de dispersão lexical das modalidades visual, olfativa e gustativa. Estes dados facilitaram também, a compreensão

da diferença entre uma unidade surgir numa secção de modalidade sensorial e ser uma unidade com uma modalidade sensorial atribuída. Com este processo, foram selecionados todos os verbos, nomes, adjetivos e empréstimos extraídos anteriormente e etiquetados com a modalidade sensorial.

O tratamento e a análise efetuados aos candidatos a ULE retirados do CEnoTEXT permitem chegar a uma nomenclatura sólida e motivada para o DEASE. Este incluirá, assim, um conjunto de descritores com uma maior precisão relativamente ao seu significado, ao seu uso efetivo, à sua frequência e aos seus contextos de ocorrência. Considera-se que a metodologia adotada, incluindo diferentes análises, é essencial para se chegar a uma categorização dos descritores de forma mais científica do que a maioria das tabelas difundidas por muitas obras, para as quais não se sabe os critérios de seleção e/ou criação. É, assim, uma forma de utilizar critérios lexicográficos de seleção (cruzar vários *corpora*), com a noção de que este é um domínio com particularidades (sensorial) e que pode ser replicado com outros domínios da AS.

## 6. PROPOSTA DO DICIONÁRIO

Neste capítulo, apresenta-se a proposta de Dicionário de Especialidade da Análise Sensorial Enológica (DEASE), bem como as etapas de análise que sustentam os critérios escolhidos e os procedimentos adotados na organização deste recurso.

Como discutido na secção 2.1.4., um dicionário pode ser entendido como um produto que apresenta um inventário de formas lexicais de uma língua natural, organizado de uma forma mais ou menos convencional e que será consultado por um conjunto de falantes que procuram esclarecer dúvidas, ou saber mais, sobre a grafia de uma dada forma (ou melhor, do seu lema), as definições, os sinónimos (ou quase-sinónimos) que possam estar atestados, entre outras possibilidades. O dicionário é, assim, um produto que é construído tendo em mente um conjunto de pressupostos de quem o elabora e por parte de quem o irá (potencialmente) consultar, e que veicula não só língua (palavras e significados), mas também cultura – aspectos culturais, históricos e sociais da língua e da sociedade, (ainda que muitas vezes de forma não consciente).

Construir um dicionário é uma tarefa complexa que exige uma profunda reflexão teórico-metodológica, por forma a garantir padrões de organização, de coerência, de sistematicidade e de qualidade suficientes para que o produto final reflita procedimentos específicos do trabalho lexicográfico e seja considerado e reconhecido pela comunidade à qual se destina – neste caso, especialistas e semiespecialistas do vinho.

Pretende-se, assim, aqui apresentar uma proposta de dicionário que motive e discuta as opções lexicográficas tomadas e que permita a sua ampliação posterior a todo o domínio da ASE, bem como a outros subdomínios ou domínios em que a modalidade sensorial possa estar presente, e não o resultado de uma compilação de um dicionário propriamente dito.

Para a prossecução dos objetivos acima identificados, importa perceber que a elaboração de uma proposta de dicionário prevê uma metodologia específica, conforme se apresenta em seguida:

1. Objetivo do dicionário e o público-alvo: Na linha do exposto discutido no Capítulo 2, é necessário definir o objetivo do dicionário e o público-alvo a

que se destina. Estes dois aspetos são de importância vital na projeção do produto final, pois são eles que subjazem à maioria das decisões a tomar. Os objetivos e o público-alvo do dicionário ajudam a definir que unidades lexicais serão incorporadas na nomenclatura, qual a informação a incluir no recurso, de que forma é que as definições nele apresentadas devem ser construídas e qual o modo adequado de organização e apresentação das unidades lexicais e das restantes informações a elas associadas.

2. Extração de dados e ferramentas: Numa segunda etapa (apresentada e discutida no capítulo 4), é necessário extrair os dados dos *corpora* e estabelecer quais as ferramentas de auxílio lexicográfico a usar. No presente caso, para a importação e adaptação das unidades lexicais foram usados os sistemas Lexico3 (2021) e o Sketch Engine (2021).
3. Definição da macroestrutura: Numa terceira etapa, à semelhança do que foi referido anteriormente, há que estabelecer a macroestrutura, tendo em conta as necessidades do público-alvo, determinando a listagem de lemas a incluir e averiguando se há necessidade de inserção de unidades de áreas conexas para solucionar dúvidas ou colmatar lacunas, como assinalado aquando dos comentários sobre as fontes textuais.
4. Definição da microestrutura: Numa quarta etapa, é necessário definir a microestrutura, ou seja, que campos devem existir em cada um dos artigos lexicográficos e quais os princípios a adotar na elaboração da definição.
5. Idealização da interface: Por fim, numa quinta etapa, e considerando o projeto aqui apresentado, será preciso idealizar um protótipo de uma aplicação digital de interface com o dicionário. Para esta etapa, considerou-se pertinente ter em conta, além das questões lexicográficas, outras necessidades práticas dos utilizadores que permitam construir soluções com vista à facilitação de acessos mais intuitivos a um conteúdo complexo e promover uma experiência de usabilidade positiva, razão pela qual se optou por uma metodologia de *User Experience (UX) Design* centrada no utilizador. A interface a desenvolver deverá considerar as fases desde o esboço até ao *wireframe*

(imagem de ecrã), que, através de um processo iterativo, permitem a criação de um protótipo ajustado ao utilizador.

Apesar de existir um conjunto considerável de obras no mercado nacional com preocupações relativas ao léxico do vinho (como apresentado em 2.3), estas são, na sua maioria, insuficientes para os objetivos que se propõem cumprir, na medida em que apresentam listagens de descritores associadas a uma definição, ou algo com função similar, sem referência à competência e ao processo de autoria, ainda que se possa perceber a vontade dos autores de esclarecer o público relativamente ao léxico em uso.

Observando as características das obras, e como discutido na secção 2.3, poucas são aquelas que se podem enquadrar na tipologia de dicionário de especialidade. Por um lado, as concebidas por lexicógrafos (indicado em 4.2.3.1.) estão desatualizadas; por outro, as restantes não disponibilizam informação sobre o modo de organização e estruturação interna, isto é, de informação relativa à macro e microestruturas, não permitindo assim perceber os objetivos e avaliar o cumprimento dos mesmos.

Perante estas constatações, a proposta do DEASE surge como resultado de uma pesquisa numa área de especialidade, direcionada por critérios lexicográficos, que pretende ser usada como ferramenta a considerar no processo de aprendizagem e no ensino-aprendizagem de especialistas e semiespecialistas no domínio da ASE e na atualização de informação por especialistas, visando a transmissão de informação, que não é apenas lexical, mas que pode potenciar o conhecimento deste domínio de especialidade. Em termos mais objetivamente lexicográficos, a proposta que aqui se apresenta visa também servir como instrumento de normalização e autorregulação do léxico de especialidade da área, fornecendo informações acerca de variantes em uso, ambiguidades, frequências e inconsistências de utilização, mas também como proposta de procedimento ou método para a criação de outros recursos que envolvam áreas de especialidade em que a análise sensorial faça parte.

## 6.1. OBJETIVOS E PÚBLICO-ALVO

O DEASE visa preencher uma lacuna existente no mercado português relativamente ao léxico (e à sistematização desse léxico) utilizado aquando da descrição das provas de vinho em língua portuguesa, dada a diversidade de interpretações e de formas de explicar os descritores que são apresentados nas obras existentes.

Embora se admita a possibilidade de variação num léxico de especialidade, em termos de eficácia comunicacional, a variação não pode ser tão grande quanto a verificada sob pena de o conteúdo não ser compreendido ou ser erradamente compreendido, em particular considerando a amplitude do público-alvo da ASE.

Neste contexto, o público do DEASE tem necessariamente diferentes níveis de especialização sobre o domínio em causa, conforme já foi mencionado anteriormente. Para o DEASE, é de referir que se pretende apresentar um instrumento de consulta que possa ser consultado por todos aqueles que necessitam de ampliar os seus conhecimentos relativamente ao domínio da ASE.

Assim, entende-se que o DEASE tem como público-alvo (i) aqueles que atuam neste domínio, sejam eles profissionais ou especialistas envolvidos na produção e comunicação deste domínio; e (ii) estudantes de disciplinas associadas à enologia; formando, por um lado, um grupo de especialistas e, por outro lado, um grupo de semiespecialistas, que também podem beneficiar deste recurso. Complementarmente, o DEASE pode visar também aqueles que procuram solidificar ou aumentar a sua formação de base e/ou procuram saber mais sobre o tema (consumidores conhecedores e potenciais compradores de outras áreas de atividade afins), mas que têm já algumas noções sobre o domínio em questão.

O facto de o público-alvo ser relativamente abrangente e heterogéneo poderá implicar opções de microestrutura, por forma a garantir uma descrição coerente e coesa no plano da definição e na apresentação da restante informação associada, aspeto que foi considerado na organização e estruturação dos artigos lexicográficos.

### 6.1.1. MACROESTRUTURA

Planear com rigor o conjunto de especificações lexicográficas que um dicionário deve apresentar ao seu público é uma tarefa fundamental e à qual o lexicógrafo deve dedicar algum tempo, como sugere Gouws (2012). Assim, é possível garantir uma estrutura coerente ao longo de todo o dicionário, mas também possibilitar uma possível adaptação da estrutura, caso o lexicógrafo tenha a necessidade de acrescentar informação posteriormente.

Ao criar um dicionário, vários aspetos – teóricos e práticos – devem ser tidos em conta. Um aspeto prático que deve ser considerado é a quantidade de trabalho envolvido, ou seja, o custo em termos temporais e financeiros. Ainda que não estejam sob o escopo deste trabalho, alguns destes aspetos devem ser tidos em conta, por forma a minimizar o tempo despendido em algumas tarefas. Um destes aspetos pertinentes a ser considerado diz respeito ao tamanho da nomenclatura e à opção de construir apenas um dicionário, ou vários dicionários, num determinado período.

A organização do processo de trabalho de compilação do dicionário permite que este possa ser replicado para outros domínios que apresentam especificidades semelhantes, como é o caso da AS aplicada a outras bebidas, para além do vinho, por exemplo.

Uma vez que um dicionário também deve ser adequado às necessidades dos potenciais futuros utilizadores, isto é ser *user-friendly*, foram ainda considerados os seguintes aspetos:

- Perfil do utilizador: de acordo com (Bergenholtz & Tarp, 1995:21), os potenciais utilizadores de um dicionário pretendem adquirir mais competências. No entanto, uma vez que alguns podem não estar habituados a utilizar dicionários, é preferível que as definições sejam escritas de forma simples e que as abreviaturas e os símbolos usados sejam autoexplicativos, para que o artigo lexicográfico não exija descodificação.
- Função linguística: sendo um dicionário de produção/compreensão, com ênfase na ajuda para entender a significação, cada artigo lexicográfico terá uma

configuração canónica, com uma entrada seguida de categoria gramatical e definição, mas com o acrescento de um conjunto de outras informações próprias do domínio e úteis para a aquisição de informação de especialidade.

A nomenclatura a apresentar no DEASE foi extraída a partir dos dados recolhidos nos *corpora* CEnoLEX e CEnoTEXT. A escolha da lista de lemas/descriptores a apresentar é, segundo vários autores (Gouws & Prinsloo, 2005), uma das mais importantes tarefas na elaboração de um dicionário. Esta deve resultar não só da avaliação e análise de informação de frequência (indicadora de unidades recorrentes ou raras nos textos de especialidade e que convém ao utilizador compreender), mas também da análise de um conjunto de outras unidades lexicais de diferentes categorias gramaticais e unidades lexicais compostas, que permitam também aceder a outras unidades relevantes, de igual importância para a compreensão da ASE.

A nomenclatura a apresentar no DEASE compreende assim um conjunto de unidades lexicais que poderão ser pesquisadas por ordem alfabética ou por modalidade sensorial. Esta dupla possibilidade permite ao utilizador uma organização tradicional (a ordenação alfabética), comum em dicionários de língua, a que os utilizadores estão habituados desde cedo e simplificando a pesquisa por letra. Esta modalidade de organização é considerada um requisito quase universal de organização e de segura e fácil consulta para qualquer dicionário (Hausmann & Wiegand, 1989:337). A ordenação por modalidade sensorial permitirá uma pesquisa com maior complexidade. Implica que o utilizador compreenda de antemão qual a modalidade a que o descritor em causa pertence, mas também permite ao utilizador não familiarizado com o domínio e com as modalidades sensoriais ampliar o seu conhecimento através da consulta de uma lista alfabética de descritores organizados dentro de cada modalidade sensorial.

Considerou-se possibilitar a pesquisa da nomenclatura através destas duas possibilidades para que o utilizador possa escolher a forma como pesquisar: (i) se a partir de um descritor que encontrou numa nota de prova e que desconhece (pesquisa alfabética); (ii) se através da curiosidade de conhecer os vários descritores associados a uma modalidade em particular (pesquisa por modalidade).

Esta proposta segue a preocupação com os princípios lexicográficos de apresentação dos artigos lexicográficos e com os métodos lexicográficos de possibilidade de pesquisa pelos utilizadores (Wiegand, 1984:563).

A organização das fichas lexicográficas dos descritores, apresentada na secção 5.2., permite a esquematização dos componentes macro que o DEASE deve apresentar, a saber:

- **Índice:** Apresentação ao utilizador dos diferentes conteúdos internos do DEASE, funcionando como um referente da estrutura do dicionário;
- **Introdução:** Pequena descrição acerca da obra, da sua autoria, função e escopo.
- **Guia de utilizador:** Secção onde se explica ao utilizador como usar o DEASE e como compreender a informação apresentada, bem como a sua ordem. É também aqui que se indica ao utilizador como interpretar o artigo lexicográfico para que possa retirar o maior partido do DEASE. Inclui a explicação de cada campo do artigo lexicográfico, a informação referente a outros descritores associados, as remissões, a fonte das obras selecionadas, a apresentação do alfabeto fonético (para o utilizador compreender as transcrições fonéticas), entre outras informações consideradas pertinentes.
- **Lista de abreviaturas**<sup>67</sup>: Local onde se apresentam por extenso as abreviaturas utilizadas ao longo do dicionário e se refere o seu significado, para que o utilizador as compreenda.
- **Nomenclatura:** Apresentação ao utilizador das unidades lexicais ordenadas alfabeticamente (Parte I) e por modalidade sensorial (Parte II).

---

<sup>67</sup> Apesar de a ideia ser um produto digital, e como tal sem necessidade de utilização de abreviaturas por motivações de espaço, considerou-se também a possibilidade de ser um produto em versão papel. Assim a ideia de utilização de abreviaturas é desde logo considerada para essa eventualidade. Além disso, também se considera o uso de abreviaturas na versão digital, considerando o seu potencial na utilização de versões digitais para formatos pequenos (*smartphone*) em vez de *tablet* ou computador.

Uma boa definição da macroestrutura permite tornar a informação interna do DEASE perceptível (cumprindo os princípios de adequação e da qualidade). Como refere Costa (2015:67), o utilizador, na eventualidade de não compreender algum artigo lexicográfico aquando da sua utilização, pode consultar estas páginas e ficar esclarecido.

### 6.1.2. MICROESTRUTURA

Após a seleção dos candidatos a ULE foram criadas fichas lexicográficas que constam da BDEnoLex. Esta ficha tem vários campos, como se apresenta de seguida na Tabela 42.

Secções	Campos da ficha lexicográfica BDEnoLex
	entrada / descritor
Secção 1	código
	data
	criado por
Secção 2	transcrição fonética
	categoria gramatical
	género
Secção 3	modalidade sensorial
	submodalidade sensorial
Secção 4	definição
	fonte da definição
	sinónimo (quase-sinónimo)
	nota adicional interna
Secção 5	marcas
	variante ortográfica
Secção 6	exemplo real
	fonte do exemplo (associar ao CenoText)
	exemplo construído
	fonte exemplo
Secção 7	imagem
	nota adicional externa

Tabela 42. Estrutura base da ficha lexicográfica da BDEnoLex

Neste sentido, a ficha lexicográfica da BDEnoLex que se apresenta reúne um conjunto de diferentes tipos de informação passível de considerar no DEASE, como se mostra de seguida:

- O descritor<sup>68</sup> (cabeça de entrada, a ULE).
- A secção 1, referente aos dados da ficha lexicográfica da BDEnoLex que refletem informações sobre atualidade e autoria.
- A secção 2, referente a aspetos gramaticais, como categoria gramatical, a transcrição fonética<sup>69</sup>.
- A secção 3, referente à identificação da modalidade e submodalidade sensorial (esta secção pode ser replicada, no sentido em que uma mesma ULE pode estar associada a várias modalidades (as unidades multissensoriais de que se falou na secção 5)).
- A secção 4, relativa à definição (que pode ser replicada consoante a necessidade – como no caso dos descritores multissensoriais, como acima mencionado), a sinónimos (com a particularidade de diferentes sinónimos poderem estarem associados a diferentes definições, consoantes as modalidades), a notas com observações do lexicógrafo (que podem mais tarde serem utilizadas para considerar uma abordagem prescritiva e ajudar, juntamente com os especialistas envolvidos, na escolha de uma direção para a normalização do léxico).

---

<sup>68</sup> Aquando da organização das ULE considerou-se que para efeitos de organização interna, estas deviam ser agregadas a uma única ULE, mesmo com o potencial de terem diferentes modalidades e diferentes definições associadas, isto é, uma organização que facilitasse a pesquisa e o entendimento do utilizador, em vez de diferentes ULE para cada modalidade sensorial e definição. Assim na BDEnoLex decidiu-se um descritor, um código numérico interno e toda a informação associada ao mesmo (por exemplo, “001 “groselha” que possui “001-01 “groselha” associado a modalidade olfativa e “001-02 “groselha” associado a modalidade gustativa).

<sup>69</sup> A decisão de introduzir a transcrição fonética em vez de um ficheiro áudio prende-se com a maior facilidade de inserção da informação. Além disso alguns programas, como o *Phonetic Sounds* permite a leitura dessa transcrição. A outra razão para a preferência inicial pela transcrição fonética é que a introdução de um ficheiro áudio implica a escolha de uma voz, a gravação em estúdio para garantir a qualidade e com tudo isso uma despesa financeira, que a transcrição fonética evita numa fase inicial, e que em nada impede a inserção de áudio numa fase posterior.

- A secção 5, relativa a variantes ortográficas, permitindo a remissão para as mesmas dentro do dicionário, existindo também espaço para marcas de uso (neologismo, arcaísmo, empréstimos, por exemplo).
- A secção 6, relativa ao exemplo de utilização real do descritor, através da apresentação de extratos de texto retirados do CenoText ou da apresentação de exemplos construídos, quando necessário, que facilitem a compreensão da definição, permitindo ao utilizador familiarizar-se com o descritor e o(s) seu(s) uso(s).
- A secção 7, que contempla uma possível integração com bases de dados de imagens/ilustrações (o que pode ser benéfico por possibilitar uma melhor compreensão do descritor em causa, permitindo ao utilizador uma informação cultural associada para compreender o que é mencionado na definição) e notas de carácter mais geral, que, por vezes, compreendem uma estrutura equivalente a uma definição enciclopédica, já que descrevem as relações do objeto com o mundo real, para auxiliar o público semiespecialista.

Evidentemente, a ficha contém, por defeito, campos específicos previamente estabelecidos. O importante desta ficha é a possibilidade de permitir à equipa lexicográfica uma organização da informação coerente e sistemática e, também, a inserção de outros casos consoante as necessidades futuras do DEASE.

A configuração de um artigo lexicográfico de especialidade pode variar de acordo com o tipo de domínio de especialidade em causa, bem como os objetivos e público-alvo da obra. A escolha dos campos obrigatórios e opcionais de uma ficha lexical depende do objetivo do dicionário e do público-alvo. De um modo geral, os campos obrigatórios num dicionário são aqueles que são considerados essenciais para a identificação e compreensão da entrada (considerando os objetivos e o público-alvo do dicionário), enquanto os campos opcionais são aqueles que fornecem informações complementares ou adicionais. No caso do DEASE, considerou-se a seguinte configuração:

1. A transcrição fonética é um campo obrigatório importante, pois permite que o leitor identifique a pronúncia correta da palavra.

2. A categoria gramatical é outro campo obrigatório relevante, pois é essencial para a correta utilização da palavra na frase e para identificar possíveis restrições de uso.
3. A modalidade sensorial também é um campo obrigatório importante em um dicionário de especialidade, pois ajuda a identificar o tipo de experiência sensorial que a ULE representa.
4. A definição e o sinónimo (ou quase-sinónimo, quando estes existam) também são campos obrigatórios essenciais, pois são os principais elementos que permitem ao leitor compreender o significado da palavra.
5. As marcas e a variante ortográfica também são considerados campos obrigatórios importantes, pois fornecem informações relevantes sobre a palavra, como a sua grafia e se é um termo técnico ou especializado.
6. A nota adicional interna, o exemplo real e o exemplo construído, são campos opcionais pois podem ser considerados relevantes para oferecer ao leitor informações adicionais que possam esclarecer o uso da palavra em determinado contexto. A imagem também é uma opção que pode complementar a descrição da palavra, mas não é necessária em todos os casos.

Portanto, a escolha dos campos obrigatórios e opcionais teve em consideração o objetivo do dicionário, o público-alvo e o tipo de informação que se pretende fornecer. Cada campo é justificado em função desses fatores, visando oferecer ao leitor uma descrição completa e precisa da entrada lexical.

Pelos motivos vários apresentados anteriormente, para a microestrutura do artigo lexicográfico do DEASE não serão apresentados todos os campos da ficha. A razão para tal escolha prende-se com o facto de algumas secções não serem pertinentes para o utilizador, como por exemplo a secção 1. Além disso, a escolha de apenas algumas secções permite uma leitura clara da informação que é relevante por parte do utilizador.

Como se apresenta na Tabela 43, os campos marcados com “sim” são sempre apresentados na microestrutura; os campos marcados com “±” só serão incluídos na microestrutura se for pertinente; enquanto os campos assinalados com “não” apenas constam na ficha lexicográfica, não sendo pertinente em incluí-los na microestrutura do DEASE.

Secções	Campos da ficha lexicográfica BDEnoLex	Campos a surgir na microestrutura DEASE
	entrada / descritor	sim
Secção 1	código	não
	data	não
	criado por	não
Secção 2	transcrição fonética	sim
	categoria gramatical	sim
	género	±
Secção 3	modalidade sensorial	sim
	submodalidade sensorial	±
Secção 4	definição	sim
	fonte da definição	não
	sinónimo (quase-sinónimo)	±
	nota adicional interna	±
Secção 5	marcas	sim
	variante ortográfica	sim
Secção 6	exemplo real	±
	fonte do exemplo (associar ao CenoText)	não
	exemplo construído	±
	fonte exemplo	não
Secção 7	imagem	±
	nota adicional	±

Tabela 43. Comparação da estrutura base da ficha lexicográfica da BDEnoLex e da microestrutura do artigo lexicográfico do DEASE

A possibilidade de a ficha lexicográfica conter mais secções do que a microestrutura do artigo lexicográfico do DEASE ajuda, num primeiro passo, a separar a informação recolhida e a organizá-la por secções que determinados autores denominam como Paradigma Informacional, Paradigma Definicional e Paradigma Pragmático (Messias, 2003). Determinou-se que a ficha lexicográfica seria mais ampla e com mais secções para que no futuro pudesse ser extraída informação que se determine relevante inserir à disposição do utilizador do DEASE e para manter registo atualizado de informação para estudos linguísticos/lexicográficos futuros.

A Tabela 44 lista os campos obrigatórios do DEASE.

<b>Campos da ficha lexicográfica BDEnoLex</b>	<b>Campos a surgir na microestrutura DEASE</b>
entrada / descritor	sim
transcrição fonética	sim
categoria gramatical	sim
modalidade sensorial	sim
variante ortográfica	sim
definição	sim
marcas	sim

Tabela 44. Elementos obrigatórios do artigo lexicográfico do DEASE

A transcrição fonética é obrigatória para auxiliar o utilizador a produzir o descritor e, também, por se considerar que se a transcrição fonética existir na versão papel, esta pode ser usada como base na versão digital para auxiliar um sistema externo a produzir áudio para os utilizadores que assim o pretendam, ou utilizadores com alguma deficiência visual.

A categoria gramatical considera-se obrigatória porque auxilia o utilizador a fazer um uso adequado da mesma aquando o uso do descritor em contexto. Este campos contém apenas os valores adjetivo, nome e verbo.

Relativamente à modalidade sensorial esta é considerada obrigatória, pois permite ao utilizador organizar os descritores por modalidade: visual, olfativa, gustativa. Esta é também a ordem preferencial de apresentação das modalidades. Quando um descritor estiver associado a mais do que uma modalidade sensorial, a ordem de apresentação é

sempre: 1.º) definição associada a modalidade visual, 2.º) definição da modalidade olfativa e 3.º) definição da modalidade gustativa. Esta ordem reflete a ordem usada no exame sensorial. Considera-se que permite uma organização familiar ao domínio de especialidade que permite ao utilizador ao ler um descritor numa nota de prova, compreender melhor a modalidade em causa dependendo da posição que ocupa na nota de prova e perceber mais facilmente qual das definições melhor se adequa ao seu caso.

A variante ortográfica é importante para estabelecer relações entre diferentes formas gráficas associadas, que na versão papel devem aparecer duplicadas, mas que na versão digital são diretamente relacionadas e apresentadas.

Quanto à forma e ordem de apresentação da definição, apesar das possíveis ordenações das definições num dicionário (frequência de uso, importância, uso cronológico, ordem alfabética, entre outras), a escolha no DEASE prendeu-se de novo com a modalidade sensorial a que está associada, como explicado anteriormente.

Os restantes campos são campos opcionais, isto é, apenas serão incorporados quando se encontre dados relativos a estes campos associados aos descritores e quando relevante. A ausência destes campos preenchidos não representa falta de informação essencial.

<b>Campos da ficha lexicográfica BDEnoLex</b>	<b>Campos que podem surgir na microestrutura DEASE</b>
género	±
submodalidade sensorial	±
nota adicional interna	±
exemplo real	±
exemplo construído	±
imagem	±
nota adicional	±
sinónimo (quase-sinónimo)	±

Tabela 45. Elementos opcionais do artigo lexicográfico do DEASE

Após a definição dos campos que devem constar obrigatoriamente (em fundo azul) e aqueles que são opcionais, foi elaborada uma estrutura ideal do artigo lexicográfico<sup>70</sup>, como se apresenta na Figura 39.

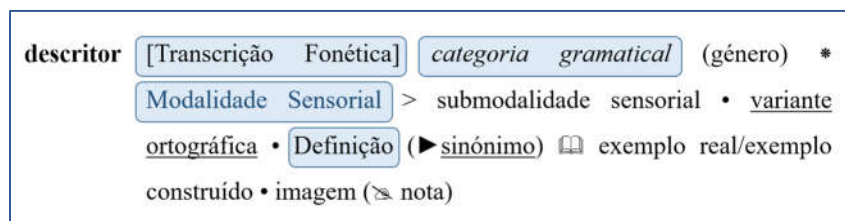


Figura 39. Estrutura básica dos constituintes obrigatórios e opcionais do artigo lexicográfico

Esta configuração de microestrutura contempla também a criação de *layout* (ao nível da fonte, do tamanho, da cor, do espaçamento, negrito, itálico, sublinhado e outros caracteres especiais (\*, •, ▶, 📖, 🔗), para que o utilizador possa identificar com clareza e sistematicidade os elementos da microestrutura. A Figura 40 ilustra esta microestrutura para a entrada “suor de cavalo”.

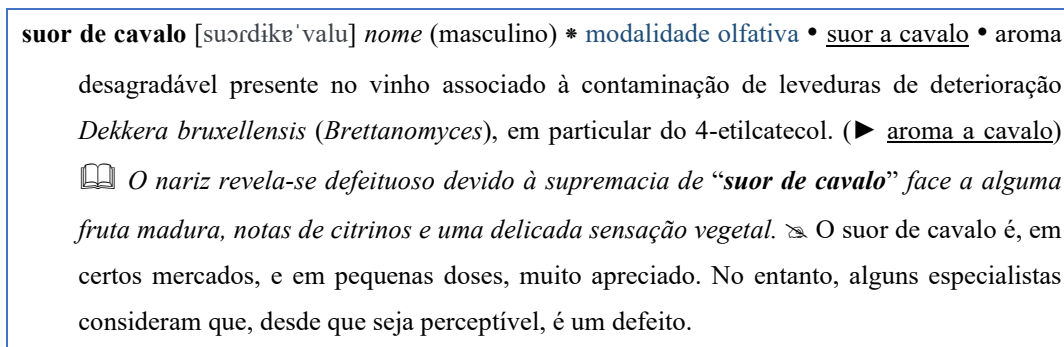


Figura 40. Proposta de artigo para “suor de cavalo”

Como mencionado anteriormente, algumas secções podem surgir duplicadas, isto é, um descriptor pode estar associado a mais do que uma modalidade sensorial, como o

<sup>70</sup> A versão que aqui se apresenta considera, num primeiro momento, uma versão tradicional de artigo lexicográfico para versão em papel por motivos de clareza de apresentação. A fase de adaptação do artigo lexicográfico consoante as especificidades de tamanho de ecrã e de formato digital serão retomadas aquando da apresentação do protótipo da App DEASE.

exemplo “opaco”, tal como a um descritor podem estar associadas mais do que uma definição. Assim, perante um caso como “opaco”, decidiu apresentar-se o descritor desdobrado quanto à modalidade sensorial e definição, como se apresenta na Figura 41.

**opaco** [ɔ'paku] *adjetivo masculino*. \* 1. modalidade visual > limpidez • que é escuro ou sem limpidez. (▶ turvo; velado) 📖 *Cor rubi profunda, quase opaca*. \* 2. modalidade olfativa > intensidade • que tem aromas pouco intensos. (▶ apagado); 📖 *Mostra um aroma onde se nota muita barrica, concentrado, mas não opaco, químico, com frutos pretos de muito boa qualidade*. \* 3. modalidade gustativa > intensidade • que tem pouco sabor. (▶ deslavado). 📖 *Opaco na boca, com acidez viva, gás a dar volume, final um pouco curto. Pouco ambicioso*.

Figura 41. Proposta de artigo para “opaco”

Observa-se que o descritor “opaco” apresenta três modalidades sensoriais, que se apresentam por ordem de exame, primeiro a definição do descritor associada ao visual, seguido da modalidade olfativa e por fim, a modalidade gustativa.

A modalidade sensorial apenas surge explicitamente mencionada em nomes e adjetivos uma vez que estes podem estarem associados a uma ou mais modalidades sensoriais e se considera importante que a explicitação sensorial seja mencionada para auxiliar o utilizador a contextualizar o descritor que procura no DEASE.

Evidentemente que, consoante o tipo de classe gramatical do descritor, algumas secções também podem ser variáveis, como se pode verificar no exemplo seguinte, onde se apresenta o descritor “crescer”, verbo.

**crescer** [kɾiʃ'seɾ] *verbo* • modificar características (olfativas e gustativas) de forma positiva (um aroma mais harmonioso, por exemplo) sob o efeito do arejamento ou do envelhecimento, nas condições adequadas de conservação. (▶ evoluir; melhorar) 📖 *Vinho muito jovem, ainda fechado, o tempo em garrafa vai fazê-lo crescer*. ∞ São comuns as expressões: *crescer em garrafa; crescer com arejamento*.

Figura 42. Proposta de artigo para “crescer”

Quando se fala de vinho, o verbo “crescer” pode ser definido como a evolução positiva das características do vinho ao longo do tempo, seja na garrafa ou no barril. Essa evolução pode estar relacionada a aspectos como o aroma, o sabor, a cor e a complexidade, que se desenvolvem e se aprimoram com o passar do tempo. O crescimento pode ser influenciado por diversos fatores, como as condições de armazenamento, a colheita da uva, as técnicas de vinificação, entre outros. Para efeitos de apresentação de uma definição curta, concisa e clara, entendeu-se que há uma modificação de características e a inserção de exemplo real retirado do CenoTEXT, a par de uma nota adicional sobre a existência de algumas expressões comuns e frequentes com o descritor “crescer” ampliariam a informação prestada ao utilizador.

Embora a organização da informação possa ser considerada tradicional, na medida em que os dicionários de especialidade, por norma, apresentam informação em forma de etiqueta do domínio ou subdomínio a que a unidade lexical se refere, a forma de apresentação dos artigos lexicográficos foi decidida tendo em consideração que os utilizadores de um dicionário deste tipo estão familiarizados com uma determinada configuração estética, tanto em dicionários digitais como em formato papel. Essa configuração “tradicional” inclui o descritor em negrito, seguido da transcrição fonética, da categoria gramatical, da definição e de um exemplo, o que facilita a leitura e compreensão da informação por parte do utilizador.

Também se entende que a maioria dos utilizadores (frequentemente ou não) de dicionários está familiarizado com essa configuração (Gottlieb & Mogensen (2007). Gruyter. (2013), Müller-Spitzer (2014)), de modo que nem sempre aceitam ou compreendem quando produtos lexicográficos alteram de forma muito evidente essa estrutura em formato papel. No sentido de apresentar uma proposta de dicionário de um domínio de especialidade com destaque para as modalidades sensoriais, considerou-se relevante manter em parte essa estrutura tradicional de apresentação dos dados em formato papel, mas apresentar algumas diferenças na apresentação da mesma informação no formato digital. Nesse sentido, tendo por base toda a informação recolhida e analisada para a criação de uma proposta de dicionário em versão papel, também se considerou

pertinente apresentar uma proposta de uma aplicação digital, tendo em conta que o público de um produto lexicográfico como este seriam, fundamentalmente, utilizadores que iriam ter dúvidas ou hesitações aquando do momento de comprar um vinho e poderiam consultar mais facilmente uma ferramenta digital.

## 6.2. PROTÓTIPO DE UMA APLICAÇÃO DIGITAL – APP DEASE

Nesta secção, apresenta-se o trabalho desenvolvido para o planeamento e a construção de uma aplicação digital para o DEASE, tendo em conta não só os critérios lexicográficos anteriormente mencionados, mas também os aspetos relativos ao utilizador digital, às suas necessidades e às suas expectativas. O protótipo possibilita testar soluções digitais direcionadas para o utilizador, que permitem facilitar o alargamento da difusão do DEASE, potenciando a sua utilização em ambiente digital.

Se o DEASE pode ser visto como um contributo ao nível da construção, divulgação e valorização do saber deste domínio de especialidade, a APP DEASE espelha a nova tendência de adequação deste tipo de ferramentas e recursos para satisfazer as necessidades do utilizador digital. Deste modo, o processo de criação da APP DEASE segue a metodologia do Design Thinking, em contexto de User Experience Design (UX Design).

Segundo Morville (2004), pioneiro na área de UX Design, existem sete fatores que afetam a experiência do utilizador ao interagir com um produto digital, como se pode observar na Figura 43 (imagem retirada de <https://tinyurl.com/2p9b2ffp>).

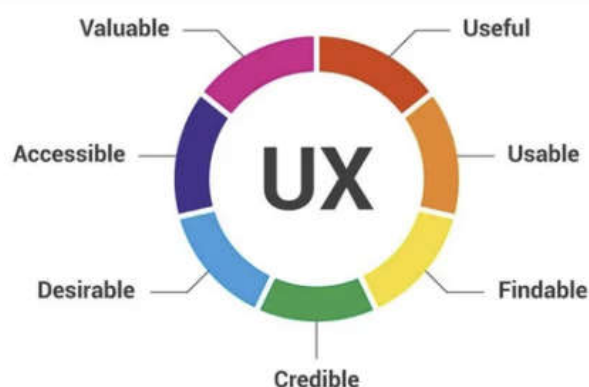


Figura 43. Os sete fatores *útil/usável/achável/creível/desejável/acessível/valioso* que afetam o utilizador

Cada um destes fatores serve para lembrar o desenvolvedor de qualquer aplicação digital de que há um conjunto de aspetos que levam o utilizador a escolher descarregar, utilizar e preferir uma aplicação em vez de outra.

Assim, para a construção do protótipo APP DEASE, foi seguido o percurso do UX *Design*, que permite tomar de modo mais seguro as decisões que influenciam a forma como a APP DEASE funcionará, e que se apresenta na Figura 44.

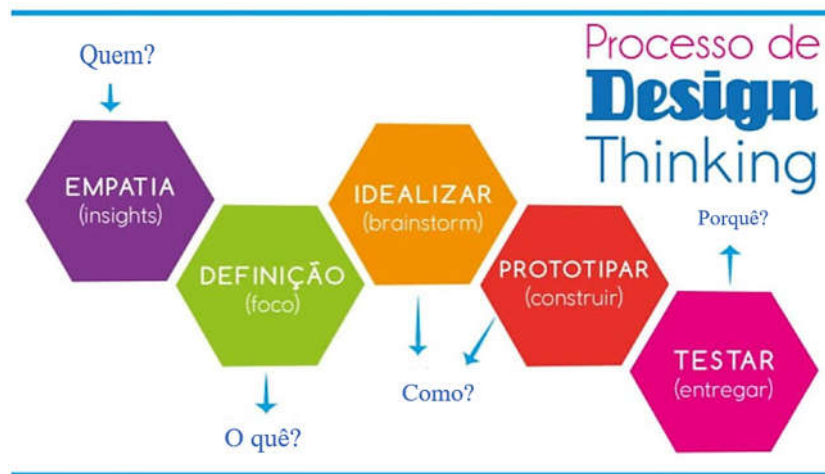


Figura 44. Descrição do processo do UX Design (retirado de <https://tinyurl.com/332mb3ad>)

A primeira etapa do processo de UX *Design* diz respeito à Empatia, em que se procura conhecer o potencial utilizador do DEASE. Vários métodos de investigação podem ser usados para definir qual o perfil do futuro utilizador digital, nomeadamente através de entrevistas, questionários, entre outros.

Na segunda etapa – Definição –, procura definir-se o problema que a solução digital pretende resolver através da aplicação de uma de várias estratégias possíveis, neste caso, será a estratégia 5W1H (um conjunto de 6 perguntas – *who* (*quem*), *what* (*o quê*), *where* (*onde*), *when* (*quando*), *why* (*porquê*) e *how* (*como*)), para definir um utilizador, numa determinada situação (espaço e tempo), que utilizaria a aplicação para um determinado

fim e como isso aconteceria. Paralelamente, faz-se um estudo de mercado (nacional e internacional, se houver necessidade) para conhecer outras aplicações existentes que têm função e público semelhantes (análise competitiva) e, por fim, consideram-se os pontos fortes do produto, demonstrando as razões pelas quais este produto é inovador.

Na terceira etapa – Idealização –, concebem-se várias ideias que possam resolver os diferentes problemas e necessidades identificados nas fases anteriores, privilegiando ainda nesta fase a quantidade de ideias sobre a qualidade das mesmas. Uma das técnicas usadas nesta fase é designada de *Crazy 8*, onde se desenham oito ideias num período de 5 a 8 minutos. Esta técnica pode ser repetida tantas vezes quantas forem necessário para chegar aos *wireframes* adequados.

Posteriormente, na quarta etapa – Protótipo –, após a exploração de várias ideias e da escolha das melhores soluções, pesquisa-se qual a arquitetura de base e criam-se os primeiros ecrãs sem detalhes (*wireframes*), que se focam somente na disposição dos diferentes elementos (menus, botões, entre outros) no ecrã. Tendo por base a investigação feita anteriormente, constrói-se um protótipo com um conjunto de *wireframes* que começam por ser de baixa-fidelidade, até evoluir para um protótipo de alta-fidelidade que se assemelhará ao produto real final a disponibilizar ao utilizador.

Na quinta e última etapa – Teste –, são configurados vários testes de usabilidade, que posteriormente são apresentados a potenciais utilizadores de modo a recolher *feedback* relativamente à compreensão e facilidade de navegação na aplicação (às páginas, à ligação entre páginas, à disposição da informação no ecrã, entre outros aspetos). Esta é a fase em que se identifica se existem bloqueios na utilização da APP e quais as oportunidades de melhoria. Este ciclo de teste é feito e repetido até se chegar a uma solução que esteja alinhada com as necessidades do utilizador.

O *UX Design* permite conhecer o utilizador, as suas necessidades e expectativas, as dificuldades de utilização com outros produtos semelhantes, e compreender como o utilizador usa uma aplicação e de que forma é que as suas emoções influenciam essa experiência, por forma a adequar a solução digital aos futuros utilizadores.

### 6.2.1. UX DESIGN DA APP DEASE

De seguida, apresenta-se o processo do *UX Design* desenvolvido para criar a APP DEASE.

#### 6.2.1.1. EMPATIA: INQUÉRITO AOS UTILIZADORES

O primeiro passo foi construir um questionário para conhecer os perfis dos potenciais utilizadores e extrair um conjunto de informação relativa a vários aspetos, tais como:

- características sociodemográficas;
- hábitos de consumo de bebidas alcoólicas;
- hábitos relacionados com atividades associadas ao vinho (degustações, vindimas...);
- práticas de resolução de dúvidas, isto é, o que fazem quando não compreendem certas palavras associadas ao vinho;
- conhecimento e utilização de recursos sobre o vinho, ou seja, que obras em versão papel conhecem, que obras em versão digital conhecem, se consultam essas obras em versão papel ou digital, quais as razões que os levam (ou não) a consultar essas obras, que aplicações sobre o vinho conhecem, que aspetos gostam (e não gostam) nas aplicações digitais, entre outras questões.

Elaborar um questionário requer o cumprimento de alguns requisitos, nomeadamente a definição do número de perguntas a apresentar, a escolha do tipo de pergunta – se são perguntas de resposta aberta ou fechada, a estimativa do tempo de resposta, a escolha do tipo de imagens a apresentar, a definição de quantas obras da especialidade apresentar, etc. Neste caso específico, optou-se por mais perguntas com

resposta múltipla de escolha fechada, poucas perguntas de resposta aberta e que o preenchimento do questionário não ultrapassasse os dez minutos, no total (consultar Anexo 1, onde se apresenta o formulário do questionário apresentado).

Após a validação do questionário por um pequeno grupo de utilizadores (para despistar alguma pergunta mal elaborada ou com interpretação errada), o questionário foi aplicado a um conjunto de 65 utilizadores portugueses com as seguintes características sociodemográficas:

- 57% do sexo masculino e 43% do sexo feminino;
- 59% na faixa etária dos 40 aos 49 anos e 26% dos 30 aos 39 anos;
- 59% com licenciatura, 25% com ensino secundário completo e 17% com mestrado;
- 90% trabalhadores a tempo inteiro.

Os resultados do questionário são analisados de seguida.

Relativamente à questão “Qual a sua relação com o vinho?” e às respetivas hipóteses apresentadas, 43% considera-se um apreciador de vinho que tem prazer em saber mais sobre o tema, 43% consome regularmente vinho em eventos sociais, 37% interessa-se em saber mais sobre o tema, 70% já participou em degustações e 43% já participou em vindimas.

Quanto à questão “Quando não compreende alguma informação sobre o vinho (palavras ou expressões), como resolve a dúvida? Pode escolher mais do que uma hipótese.”, 75% pesquisa na Internet para esclarecer as dúvidas, 54% conversa com alguém que considera que sabe mais sobre o assunto e 11% tenta decifrar o significado pelo contexto.

Sobre as obras disponíveis no mercado, foram apresentadas 11 obras, de autores nacionais e internacionais, estes últimos traduzidos em PE, de nível iniciante a

especializado, com preço acessível (entre os 5 e os 10 euros), com preço intermédio (entre os 11 e os 49 euros) e com preço elevado (acima de 50 euros).

<b>Relação entre nível / preço</b>	<b>Obras selecionadas</b>
Obra de nível iniciante a especializado, com preço acessível (entre os 5 e os 10 euros)	<i>Especialista de Vinhos em 24 Horas</i> , de Jancis Robinson <i>Como apreciar Vinhos</i> , de Hugh Johnson <i>Branco ou Tinto? E Outras Respostas às Dúvidas que Todos Temos sobre como Escolher, Comprar e apreciar o Vinho</i> , de Joana Maçanita
Obra de nível e com preço intermédio (entre os 11 e os 49 euros)	<i>Entender de Vinho</i> , de João Afonso <i>O Vinho na Ponta da Língua</i> , de Maria João de Almeida <i>O Abecedário do Vinho</i> , de José Salvador <i>Saber Beber Vinho</i> , de Rodolfo Tristão <i>O Vinho da Uva à Garrafa</i> , de António Dias Cardoso
Obra de nível e com preço elevado (acima de 50 euros)	<i>Dicionário Ilustrado do Vinho</i> , de Jorge Böhm <i>O Gosto do Vinho: O Grande Livro da Prova</i> , de Émile Peynaud <i>Dicionário Ilustrado do Vinho do Porto</i> , de Manuel Pintão e Carlos Cabral

Tabela 46. Lista de obras apresentadas no Inquérito

Cerca de 85% dos informantes nunca ouviu falar das obras em questão, apenas 11% já ouviu falar, mas não consultou nenhuma das obras referidas, e apenas 4% conhece e consulta essas obras.

No que diz respeito às revistas de especialidade, foram apresentadas quatro obras nacionais (*Revista de Vinhos – A Essência do Vinho*, *Enoteca*, *Grandes Escolhas e Paixão pelo Vinho*), sendo que 58% não as conhece, 21% já ouviu falar delas, mas não as consultou, e 17% conhece e consulta essas obras, seja em versão papel ou digital.

Com as respostas obtidas neste questionário, pretendia-se compreender como as pessoas se identificavam relativamente à sua relação com o vinho, que obras em versão papel e digital conheciam e consultavam e, em caso de dúvida, como se esclareciam. Tais informações permitiram perceber melhor como seriam os potenciais utilizadores não-especialistas e em que versão prefeririam consultar informação sobre o vinho.

De seguida, fez-se uma análise de todas as respostas obtidas e verificou-se que estes informantes consomem vinho regularmente e/ou em eventos sociais, gostam do tema, mas que, perante dúvidas, consultam a Internet (sem necessariamente consultar um dicionário/léxico/glossário/vocabulário do vinho: digitam a palavra ou expressão e veem os resultados que aparecem no motor de busca), preferem obras em versão digital, por uma questão de facilidade e de rapidez de acesso a informação, apesar de considerarem que muitas vezes o ecrã de consulta é pequeno (telemóvel) e de se queixarem que nem sempre têm confiança na informação que obtém a partir da Internet.

A aplicação deste questionário permitiu observar um conjunto de respostas de um público heterogéneo (que pode englobar especialistas e não-especialistas) e que funciona como ponto de partida para comparar com os resultados da análise competitiva e chegar a uma definição inicial de um utilizador prototípico para o qual será desenvolvido o protótipo da aplicação.

#### 6.2.1.2. DEFINIÇÃO: ANÁLISE COMPETITIVA E UTILIZADORES PROTOTÍPICOS

Simultaneamente à aplicação do questionário, foi feita uma análise competitiva (*competitive audit*) relativamente a aplicações já existentes que pudessem resolver o problema do utilizador, para perceber o que funciona e não funciona nessas soluções.

Foram pesquisadas na loja *Google Play* aplicações com a designação “dicionário de vinho”, tendo-se verificado que no mercado nacional não surgia nenhuma aplicação com essa mesma denominação, mas em que eram apresentadas sugestões de outras aplicações relacionadas com vinho, como se pode verificar na Figura 45.

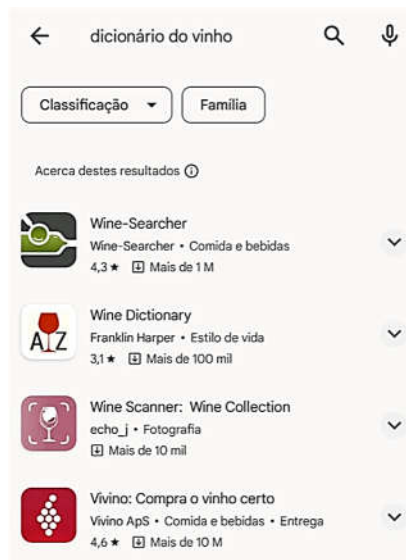


Figura 45. Exemplo de pesquisa de aplicações digitais relacionadas com vinho

Perante a falta de resultados em PE, pesquisaram-se aplicações noutras línguas (inglês, francês, italiano e espanhol) com as palavras-chave “dicionário”, “léxico” e “vocabulário” para abranger um maior número de possíveis aplicações com informação do tipo que se pesquisava. Foi selecionado como adequadas um conjunto de aplicações (por exemplo, a *Wine Dictionary*, da Franklin Harper, a *Wine Dictionary*, da Faith More ou a *Wine Dictionary*, da GreenLife Apps), que foram descarregadas e usadas para se observar e analisar:

- a organização da informação no sistema;
- a arquitetura das páginas (fontes, cores, navegação entre páginas);
- número de descritores e respetivas definições apresentadas;
- outra informação adicional associada;
- número de *downloads*;
- a avaliação dada pelos utilizadores;
- os comentários dos utilizadores.

Através dos comentários dos utilizadores, foi possível verificar os pontos fortes e pontos fracos das aplicações já existentes no mercado. Como exemplos de pontos fracos, os utilizadores apontaram: a dificuldade em pesquisar; a falta de acessibilidade (não possuir áudio ou imagens); a falta de categorização das palavras apresentadas; e a informação disponibilizada ser considerada básica e incompleta. Esta pesquisa permite compreender o que existe, o que funciona e o que uma futura proposta deve ter em conta para ser mais efetiva.

Adicionalmente, foi aplicada a estratégia 5W1H para pensar em utilizadores prototípicos (que representam um grupo de público com características semelhantes), que em determinados cenários, utilizariam a aplicação e em que circunstâncias isso aconteceria, como se apresenta nas Figura 46.



<b>Who / Quem:</b>	A Rita - consumidora de vinho	O Mário - empregado num restaurante	A Marta - provadora de vinho
<b>What / O quê:</b>	A Rita pretende comprar um vinho para consumir/oferecer e não compreende as características descritas na nota de prova do rótulo/folheto.	É esperado que o Mário sugira vinhos aos seus clientes e que tenha conhecimento nesta matéria.	A Marta está a estudar para ser provadora de vinhos, mas ainda não compreende todas as descrições associadas ao vinho.
<b>Where / Onde:</b>	Em casa ou no supermercado.	Principalmente no trabalho.	Em casa ou no curso.
<b>When / Quando:</b>	Durante as compras.	Quando tem algum tempo para estudar os vinhos que o restaurante vende.	Durante os estudos.
<b>Why / Porquê:</b>	A Rita é muito atarefada, mas interessa-se por saber mais sobre vinhos e pretende esclarecer-se rapidamente.	O Mário quer aconselhar os seus clientes e esclarecer convenientemente as dúvidas que os clientes possam ter.	A Marta quer pesquisar um recurso confiável para se esclarecer.
<b>How / Como:</b>	Atualmente, a Rita lê os rótulos e, em caso de dúvida, pergunta a alguém que perceba mais de vinho do que ela.	Atualmente, o Mário estuda os rótulos e, em caso de dúvida, procura na <i>Internet</i> , sem ter certeza da fidedignidade desses recursos.	Atualmente, a Marta lê os rótulos e em caso de dúvida pergunta ao formador do curso.

Figura 46. Cenários e utilizadores hipotéticos

Assim, perante a análise do *feedback* dos utilizadores, das respostas obtidas da aplicação do questionário e dos cenários de três utilizadores hipotéticos chegou-se à definição do problema – quando o utilizador tem uma dúvida sobre uma palavra ou expressão relativa ao vinho que quer pesquisar, mas tem dificuldade em encontrar uma

fonte de referência que o esclareça – e à *identificação do valor da proposta* (*value proposition*), ou seja, em como a proposta é inovadora e pode servir um conjunto heterogéneo de utilizadores.

Apesar de parecer que os dados iniciais deste questionário estão direcionados para um público não-especialista, diferente do público especialista e semiespecialista para o qual o DEASE foi concebido, este não é considerado excludente de outros cenários e de outros utilizadores hipotéticos. O que se verifica é que, para o conjunto de utilizadores que participou no inquérito, este foi o cenário prevalente, que espelha uma distribuição populacional esperada (isto é, uma população de não-especialistas superior à população de especialistas). No entanto, tratando-se de um método de desenho de aplicações digitais, sabe-se que todos os utilizadores partilham preferências e necessidades relativas a um conjunto de aspetos, tais como a disposição da informação a nível visual, preferências de menus, entre outros aspetos da aplicação digital. Assim, os resultados e os futuros testes de usabilidade que foram aplicados incluem, num primeiro momento, apenas os utilizadores do cenário 1 (a *Rita*), devido à prevalência de número de respostas condizentes com este perfil de utilizador. Esta estratégia (comum em *UX Design*) permite o desenvolvimento e a rápida implementação de um produto digital com os mínimos considerados aceitáveis (a adequação de um produto tendo em conta pelo menos um perfil de utilizador) no mercado para o público com maior representatividade. Isto permite, por um lado, uma disponibilização da aplicação no mercado para colmatar a lacuna existente e, paralelamente, o desenvolvimento de adaptações consoante os outros cenários e utilizadores que são feitos de acordo com os posteriores testes de usabilidade, sem prejuízo do resultado.

Em suma, a análise competitiva do DEASE revela que este se distingue dos seus potenciais concorrentes por:

- Resultar de um trabalho de investigação em Lexicografia de Especialidade;
- Permitir a inclusão de diferentes definições para diferentes públicos;
- Possibilitar a pesquisa por modalidade sensorial, além da alfabética;
- Considerar incorporar áudio e imagem (para efeitos de acessibilidade);

- Facilitar a descoberta da aplicação no mercado das aplicações pelos utilizadores (comparativamente às versões em papel);
- Não ter competição direta em língua portuguesa.

Após a observação e a análise dos dados, que nos conduziram à constatação da viabilidade e razoabilidade da nossa proposta, fez sentido avançar para a etapa seguinte: a idealização.

### 6.2.1.3. IDEALIZAÇÃO: EXPLORAÇÃO DO SISTEMA CRAZY 8

Nesta fase, é comum a adoção do sistema *Crazy 8*, que significa configurar oito *wireframes* de forma informal, como se apresenta na Figura 47.



Figura 47. Crazy 8 aplicado ao DEASE

Nesta tarefa, começam a ser configuradas as características das páginas da aplicação, por exemplo: (1) página principal; (2) apresentação das diferentes definições; (3) número e tipos de menus a apresentar; (4) pesquisa por modalidade sensorial; (5) pesquisa alfabética; (6) adequação do artigo lexicográfico à versão digital; (7) outras

informações adicionais (por exemplo, hiperligações a outros repositórios externos); e (8) outras informações (secção onde o utilizador pode introduzir as suas notas pessoais).

Posteriormente, exploram-se hipóteses para a organização da informação na arquitetura do sistema, o número de páginas a criar, a relação entre elas no que respeita à navegabilidade e a sua interligação no que respeita à informação a apresentar em cada página. Começa-se pela elaboração de um rascunho em papel e seguido da construção de wireframes de baixa-fidelidade, que vão sendo melhorados consoante vão sendo apresentados e avaliados por utilizadores, passa-se por uma configuração de alta-fidelidade, até se chegar ao protótipo final, como esquematizado na Figura 48.

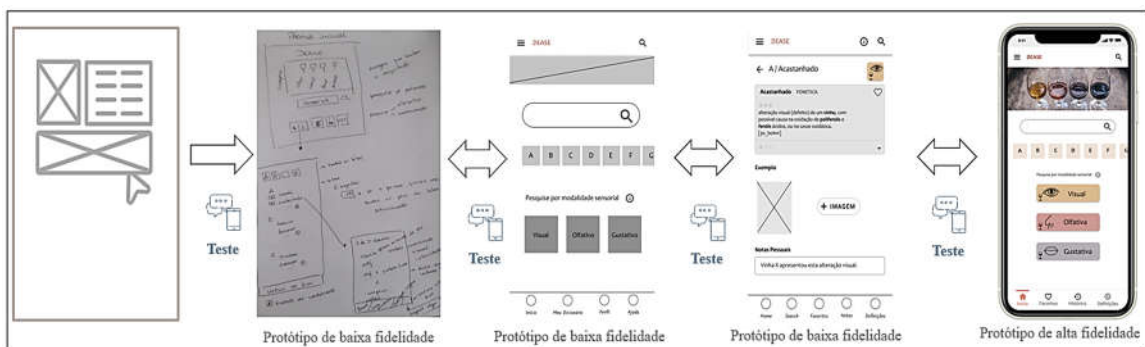


Figura 48. Evolução dos *wireframes* para construção do protótipo

Nesta fase de experimentação de hipóteses, afinam-se os vários *wireframes* e apresentam-se a grupos de utilizadores (no mínimo 5 utilizadores por iteração de teste, no total de 6 testes). De acordo com o *feedback* obtido, passa-se ao protótipo de baixa fidelidade ou reformula-se o *wireframe* até obter uma avaliação positiva. O processo repete-se assim, sucessivamente, até chegar ao protótipo de alta fidelidade (*wireframe* final apresentado em versão beta a um número alargado de utilizadores), que é a imagem final dos *wireframe* (correspondente à imagem final do protótipo de alta fidelidade da Figura 48).

#### 6.2.1.4. PROTÓTIPO: PROTÓTIPO DA APP DEASE

Após o conjunto de testes referido anteriormente, concebeu-se o protótipo da APP DEASE. Esta apresenta um conjunto de características visuais recorrentes nas aplicações digitais, como, por exemplo, a apresentação do menu desdobrável no canto superior esquerdo, a opção de pesquisar em toda a aplicação no canto superior direito e a imagem representativa do DEASE imediatamente abaixo.

Como é possível ver na Figura 49, o ecrã inicial tem uma barra de pesquisa por palavra, a ordenação alfabética apresentada na horizontal – permitindo deslizar para ambos os lados (*swipe*) –, e a apresentação da pesquisa por modalidade sensorial, com fixação de imagem representativa da modalidade, a designação escrita da modalidade em questão e a cor associada a essa modalidade sensorial: um olho e cor alaranjada para modalidade visual um nariz e cor rosada para modalidade olfativa e uns lábios e cor acinzentada para modalidade gustativa. Por fim, há ainda um total de quatro menus (Início, Favoritos, Histórico e Definições) na barra inferior da aplicação.

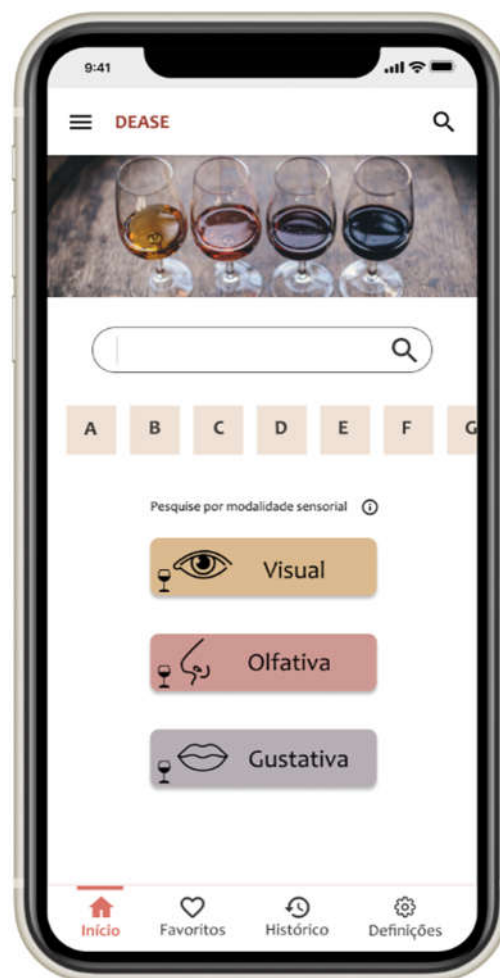


Figura 49. Imagem do ecrã inicial da APP DEASE

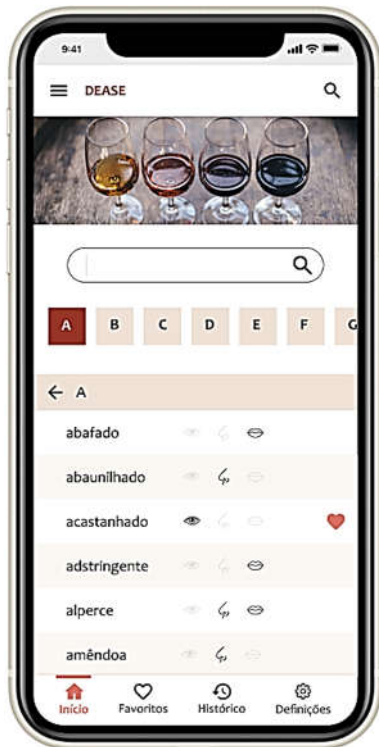


Figura 50. Ecrã de pesquisa alfabética

Ao clicar em cada letra do alfabeto, observa-se o conjunto de entradas (listagem de descritores, como por exemplo “abafado, abaunilhado, acastanhado, adstringente, alperce, amêndoa”) iniciadas pela letra selecionada, a par da(s) modalidade(s) sensorial(ais) que podem estar associadas a cada entrada. Por exemplo, “abafado” apresenta o ícone relacionado com a modalidade gustativa; “abaunilhado” o ícone relacionado com a modalidade olfativa; “acastanhado” o ícone relacionado com a modalidade visual; “alperce” surge seguido de um nariz e de uma boca (pois pode estar associado a duas modalidades, a olfativa e a gustativa – é um descritor multissensorial), ver Figura 50.

No ecrã inicial, o utilizador pode também clicar num ícone relativo a uma modalidade sensorial (visual, olfativa ou gustativa) e observar a listagem de entradas que apresentam essa modalidade associada, ordenadas alfabeticamente. A modalidade surge no topo do ecrã, na respetiva cor, forma gráfica e com o ícone correspondente. A Figura 51 ilustra esta funcionalidade para a modalidade sensorial visual, apresentando os descritores “aberto, acabado, acastanhado, (...) amarelo”






Figura 51. Ecrã da pesquisa por modalidade sensorial visual

Ao carregar na entrada (no descritor “acastanhado”, por exemplo na Figura 51), o utilizador acede a uma página relativa à entrada (Figura 52) onde se observa:



Figura 52. Ecrã do descritor “acastanhado”

- a entrada “acastanhado”, seguida do ícone relativo à modalidade sensorial (neste caso , porque está associado à modalidade visual), um botão que permite marcar este descritor como favorito () e um botão  (para ouvir a produção oral, que se baseia na transcrição fonética inserida no sistema).
- a explicitação gráfica da modalidade sensorial, abaixo, para o caso de não ficar claro para o utilizador pelo ícone.
- a categoria gramatical,
- a definição,
- o exemplo,
- a imagem,
- espaço para notas.
- No caso de ser um descritor da modalidade visual, o utilizador pode ver uma imagem associada



Através do botão “mais” (+) (que surge a vermelho na Figura 53), a aplicação permite ainda ao utilizador inserir outras informações, tais como novas imagens associadas àquele descritor (como uma fotografia de um vinho que degustou), associar notas pessoais, partilhar a definição do descritor com outros utilizadores, e ainda sugerir alguma alteração ou sugestão à própria aplicação.



Figura 53. Ecrã a que se acede a partir do botão (+)



Figura 54. Ecrã do descritor “alperce”

No caso de o descritor poder denotar propriedades associadas a mais do que uma modalidade sensorial, apresenta-se essa informação devidamente separada por uma linha horizontal, com uma listagem inicial dos símbolos das modalidades sensoriais em causa, como se pode observar no exemplo de “alperce”, que está associado à modalidade olfativa  (na parte superior da linha horizontal) e à modalidade gustativa  (na parte inferior da linha horizontal), ver Figura 54.

Para os efeitos do presente projeto, foi concebida uma versão beta do protótipo (cf. Barbosa & Santos, 2021), nomeadamente a nível estético, a nível da apresentação da definição e a nível da arquitetura do sistema no que respeita ao tipo de relações de descritores relacionados.

Por fim, apresenta-se a estrutura do artigo lexicográfico nas suas versões em papel e digital.

**descritor** [Transcrição Fonética] *categoria gramatical* (género) \*  
Modalidade Sensorial > submodalidade sensorial • variante ortográfica • Definição (▶ sinónimo) 📖 exemplo real/exemplo construído • imagem (🔗 nota)

Figura 55. Artigo lexicográfico em versão papel



Figura 56. Artigo lexicográfico em versão digital

## 6.2.2. APLICAÇÕES FUTURAS

O desenvolvimento e melhoramento da aplicação que aqui se apresenta poderá, no futuro, responder ou contribuir para a resposta a questões mais ou menos imediatas, como, por exemplo:

### 1.1. Que imagens associadas à modalidade sensorial visual apresentar?

- Sabe-se que aspetos como tonalidade, limpidez e contraste são altamente importantes no que respeita à visualização de uma cor, tendo em conta o brilho inerente do próprio dispositivo digital utilizado.
- É necessário assegurar a salvaguarda de direitos de autor no caso de remissão para bases de dados de imagens.

### 1.2. Que arquitetura de sistema deverá ser contemplada para permitir a ligação a bases de dados de odores e/ou sabores?

- Atualmente, os dispositivos digitais móveis não permitem a libertação de odores, mas sabe-se que dispositivos como narizes (e línguas) eletrónicos, atualmente utilizados em contexto laboratorial, podem ser, num futuro próximo, disponibilizados ao grande público. No contexto de auxiliar as tarefas de Inteligência Humana<sup>71</sup> ou aprimorar sistemas de Inteligência Artificial Geral de nível humano (IAG)<sup>72</sup>, é possível que a identificação automática de aroma ou sabor possa, no futuro, ajudar a apresentar novas sugestões de vinhos ao utilizador, com base nas suas preferências.

---

<sup>71</sup> À semelhança do que já acontece em ambientes corporativos e empresariais, em que os sistemas auxiliam na identificação de um determinado conteúdo em imagens ou vídeos, para verificação final por parte dos colaboradores humanos (por exemplo na Amazon)

<sup>72</sup> IAG é a designação popularizada por investigadores de Inteligência Artificial, tais como Legg, Gubrud e Goertzel, para designar “a capacidade de realizar qualquer objetivo pelo menos tão bem como os seres humanos” (Tegmark, 2017:80).

- Num futuro próximo, quando os dispositivos digitais forem capazes de funcionar como partes que possam ser acopladas ou mesmo incorporadas no corpo humano (funcionando como os implantes cocleares que auxiliam o ser humano a captar uma sensação auditiva), narizes e línguas eletrónicos (e/ou biónicos)<sup>73</sup> podem funcionar como uma ferramenta que permita ao utilizador experienciar o odor e/ou sabor referido por um determinado descritor (à semelhança do que se passa atualmente com as imagens), e, no sentido oposto, converter a informação molecular de um dado odor e/ou sabor numa descrição verbal que permita ao utilizador compreender que tipo de odor ou sabor se trata.

A definição de um protótipo do DEASE em formato de aplicação digital permite mostrar o produto final da proposta e as suas potencialidades. A elaboração de um dicionário de especialidade nesta área é uma tarefa demorada e complexa e o trabalho num domínio multissensorial pode acrescentar dimensões de complexidade à tarefa. No entanto, como apresentado anteriormente, o produto final, além dos objetivos lexicográficos, poderá funcionar como ponto de partida para outros trabalhos no campo sensorial e trazer benefícios a um conjunto alargado de utilizadores com diversos perfis e necessidades.

A utilização do *UX Design* permite melhorar a interação entre o produto (o dicionário) e o utilizador, aperfeiçoando a forma como este pode aceder de modo eficaz aos conteúdos disponibilizados e perspetivando futuros desenvolvimentos e inovações.

A criação de um dicionário especializado em análise sensorial poderia ser justificada pela necessidade de compilar e sistematizar o vocabulário técnico específico dessa área. Embora já existam obras (dicionários, glossários, léxicos, entre outros) relacionadas com vinho, gastronomia e alimentos em geral, a análise sensorial é uma disciplina com um léxico próprio e com características que justificam um dicionário dedicado exclusivamente a essa área, tendo por base uma metodologia lexicográfica.

---

<sup>73</sup> Como se pode observar nos trabalhos de Persaud (2017) e de Fang (2022).

Os pontos fortes de um dicionário de análise sensorial seriam a sua precisão e o seu rigor técnicos, ao fornecer definições e exemplos específicos relacionados com as modalidades e submodalidades sensoriais. Além disso, poderia ser uma ferramenta útil para estudantes, pesquisadores, profissionais da área e entusiastas que desejam aprofundar os seus conhecimentos nessa disciplina.

As vantagens para o utilizador são a possibilidade de consulta rápida e fácil dos descritores, além da garantia de que as definições e os exemplos apresentados são precisos e relevantes para a análise sensorial. Também poderia servir como referência para a padronização do léxico utilizado em artigos científicos, relatórios técnicos e outras publicações relacionadas com a análise sensorial.

Um dicionário de especialidade em análise sensorial pode ser aplicado a diversas bebidas, como vinhos, cervejas, destilados, licores, entre outras. Complementarmente, também pode ser aplicado noutras áreas, como alimentação, perfumaria e cosmética, por exemplo, uma vez que a análise sensorial é uma técnica ampla que pode ser aplicada a diferentes produtos para avaliar as suas características sensoriais, como o aroma, o sabor, a textura, entre outros. Portanto, um dicionário de especialidade em análise sensorial enológica pode ser útil, no sentido de permitir a sua replicação com as devidas adaptações em diversas áreas que envolvem a avaliação sensorial de produtos, como, por exemplo, para o chá, o leite, a água, o café, bebidas espirituosas<sup>74</sup>, entre outros.

Existem estudos que indicam que o público em geral pode beneficiar do uso de dicionários de especialidade em áreas específicas (Demers, Kernerman, & L'Homme (2012), Fathi (2014), Gromann & Schnitzer (2016)), como, por exemplo, na área da medicina, do direito, das ciências e das tecnologias, entre outras. O uso de ULE dificulta a compreensão de textos por parte de leitores leigos, e o uso de dicionários de

---

<sup>74</sup> Sabe-se que existem já algumas obras criadas em inglês tendo em conta esta particularidade da AS, por exemplo, (i) *Tea: Bioactivity and Therapeutic Potential* de Zhen (2002), que inclui um capítulo com um dicionário de sabor e aroma para chá; (ii) *The Sensory Evaluation of Dairy Products* de Clark et al. (2008), que contém informações sobre análise sensorial aplicada a vários produtos, incluindo leite; (iii) *Water: A Comprehensive Guide for Brewers* de Palmer e Kaminski (1963), que inclui informações sobre análise sensorial de água para cervejarias.

especialidade pode esclarecer o significado dessas unidades lexicais e, conseqüentemente, melhorar a compreensão do texto e do discurso utilizado.

Por outro lado, pode considerar-se ainda que a proposta de dicionário apresentada é, por si só, uma fonte de dados para treinar sistemas de IA, pela organização da informação dos descritores na Base de Dados e pelas definições precisas das ULE, que seriam difíceis de encontrar noutras fontes, por exemplo. Adicionalmente, os exemplos de uso e as informações sobre contextos em que as unidades são usadas são dados de grande utilidade para sistemas de IA.

Um dicionário de especialidade para um público humano pode ter uma apresentação visual mais apelativa e com mais informações adicionais, tais como ilustrações, notas culturais, etc. Além disso, pode ser organizado de forma a permitir uma pesquisa mais fácil, seja por ordem alfabética, por temas, por áreas específicas do conhecimento, ou de outra forma que seja mais adequada para o público a que se destina.

Para um sistema de IA, algumas das secções relevantes seriam:

- Descritor: palavra ou expressão a ser definida;
- Definição: explicação clara e sucinta do(s) significado(s);
- Categoria gramatical: classe morfossintática do descritor;
- Sinónimos e quase-sinónimos: ULE com significados semelhantes;
- Modalidade e submodalidade: sentido da ULE e a subcategoria específica da percepção sensorial;
- Marcas e variantes ortográficas: diferentes formas de grafia do descritor.

Essas secções, juntamente com outras informações pertinentes, como contexto de uso e os exemplos de aplicação, podem ser utilizadas por um sistema de IA para auxiliar na compreensão e no processamento de linguagem natural em análise sensorial.

## 7. CONCLUSÕES

Neste capítulo final, pretende destacar-se um conjunto de considerações sobre a investigação desenvolvida no presente trabalho, sobre as estratégias metodológicas utilizadas e sobre os desafios encontrados, bem como refletir sobre trabalhos futuros.

Observando o panorama português, concluiu-se que pouco trabalho tinha sido feito no sentido de compreender e desenvolver estudos no domínio da ASE, em particular intersetando as áreas de ASE e da Lexicologia e Lexicografia. Neste sentido, o objetivo maior deste trabalho foi elaborar uma proposta de dicionário de especialidade sobre os descritores (ULE) utilizados nas notas de prova provenientes das provas de vinho.

A razão para selecionar este domínio da Enologia teve como motivação dois fatores principais: o facto de o vinho ser um produto ligado a um conjunto de atividades distintas, tais como a enologia, a viticultura, a vinicultura, o comércio, o turismo, a gastronomia ou os serviços de bem-estar, muito presente em vários domínios da vida; e porque a Enologia é uma área em expansão e uma indústria de enorme impacto económico-financeiro, nacional e internacionalmente falado.

O domínio da ASE é um domínio com um discurso especializado em Enologia, com características próprias e produzido por especialistas, onde a degustação pode ser feita por um conjunto de diferentes especialistas com diferentes competências e diversas formas de comunicar (oralmente, por escrito e dentro de cada uma destas tipologias distintas).

Apesar de existirem alguns recursos lexicográficos, estes apresentam lacunas em diversos sentidos (na conceção, na apresentação, nos descritores que apresentam, como os definem, entre outros). Em alguns casos, não se apresenta um resultado descritivo, objetivo e consensual entre os intervenientes, apontando para alguma instabilidade do léxico utilizado e, principalmente, para a falta de metodologia e enquadramento científicos.

Por estes fatores, entende-se que a ASE é um domínio de especialidade que permite a descrição e avaliação do vinho e tem sido pouco ou nada explorado do ponto de vista lexicográfico. A proposta de construção de um dicionário de especialidade beneficiaria

um público diversificado (os especialistas e os futuros especialistas) da área, mas também um público não-especialista, mas curioso e com desejo de ser informado e conhecedor, permitindo-lhes ter acesso a um conhecimento organizado e descrito de forma sistemática, lógica e coerente.

A ideia de criar uma proposta de dicionário passou pela definição dos seguintes aspetos:

- a ASE como domínio de especialidade;
- o léxico de especialidade da ASE em PE;
- o recurso lexicográfico para um público-alvo de especialistas do domínio (estudantes, escanções, enófilos, tradutores, jornalistas, etc.), mas também de semi-especialistas (público conhecedor e interessado no tema e consumidor de vinho);
- a apresentação de dados de referência que atestassem o atual estado da ASE para o PE;
- a proposta de uma classificação sistemática do domínio da ASE, para auxiliar a redação técnica e esclarecer sobre a forma e o conteúdo dessas mesmas unidades lexicais.

Assim, o objetivo desta proposta de dicionário é fornecer uma metodologia de trabalho que permita a elaboração de dicionários de especialidade de domínios que envolvam uma área de especialidade que utilize a Análise Sensorial. Além disso, o dicionário atua como uma ferramenta fundamental para uma comunidade de utilizadores com diferentes níveis de especialização, abrangendo desde enólogos, escanções e estudantes (potenciais especialistas nessas áreas) até outros profissionais como tradutores, jornalistas e entusiastas de vinho. Esta abordagem visa atender às necessidades específicas dos especialistas, promovendo uma comunicação eficaz e o correto uso do léxico especializado nas suas atividades diárias. Ao mesmo tempo, procura facilitar a comunicação com públicos menos especializados, como o público interessado e os consumidores de vinho, reconhecendo a diversidade do público-alvo que emerge a partir

dos textos que serviram de base e foram utilizados nos *corpora*, abrangendo desde não-especialistas até especialistas.

## 7.1. OBJETIVOS E RESULTADOS ALCANÇADOS

Para refletir sobre as questões que serviram de mote à proposta de um dicionário de especialidade, importa recuperar os objetivos inicialmente apresentados e como estes foram alvo de tratamento.

### Objetivo 1. *Observar a organização e os conteúdos relevantes do domínio de especialidade – ASE*

Esta observação foi feita no capítulo *Análise Sensorial Enológica* onde se detalha: (i) que sentidos são invocados aquando da experiência do ato de degustar um vinho, (ii) o que se entende como diferença entre perceção e sensação, (iii) como se processa o processo da degustação e quais os exames efetuados (visual, olfativo e gustativo), bem como (iv) que tipo de testes (paramétricos e estatísticos) podem ser efetuados além do teste de verbalização das sensações provocadas pelo vinho de forma escrita e/ou oral.

### Objetivo 2. *Compilar um corpus resultante de uma minuciosa pesquisa de textos relativos à ASE*

Esta compilação foi descrita no capítulo *Apresentação e Parametrização dos Dados* onde se explicita as opções metodológicas tomadas relativamente ao domínio de especialidade escolhido e se apresenta a BDEno, que consiste num acervo de fontes textuais, relativas à temática do vinho escritas em PE, divididas em três tipologias: documentação técnico-científica (29 391), documentação institucional (105) e documentação legal (759), que auxiliaram a compreender que material existe, funcionando como dados de referência que atestam o estado atual da literatura na ASE em PE.

**Objetivo 3.** *Analisar, numa perspetiva monolíngue, quais são os descritores atualmente em uso em PE*

Esta tarefa foi realizada no capítulo *Apresentação e Parametrização dos Dados*, com base no *corpus* CEnoTEXT, relativo a notas de prova produzidas pelos especialistas, que contém 21 495 textos especializados do tipo nota de prova, recolhidos entre 2012 e 2021 e anotados gramaticalmente com um total de 780 585 *tokens*. Este *corpus* inclui textos autênticos escritos por especialistas e é maioritariamente composto por textos concisos, organizados, estruturados, incluindo texto semirígido no sentido de ter uma estrutura recorrente, mas que permite liberdade neológica e perceber as características do vinho consideradas relevantes na descrição.

**Objetivo 4.** *Identificar, organizar, categorizar e caracterizar os descritores existentes referentes à ASE*

Este processo foi apresentado no capítulo *Léxico de Especialidade da ASE: Análise e Tratamento* que expõe o número total de unidades lexicais de especialidade obtidas, a sua distribuição e respetiva análise e a sua organização por modalidade sensorial. Partindo de um total de 780 585 *tokens* registados no CEnoTEXT, foram identificados três grupos gramaticais de candidatos – verbos (884), nomes (2 919) e adjetivos (2 481) e 247 empréstimos não adaptados. Como primeiro critério foram selecionados de cada categoria apenas os candidatos com mais de 100 ocorrências do CEnoTEXT, a saber: verbos (55), nomes (244), adjetivos (181), empréstimos (10).

**Objetivo 5.** *Catalogar os descritores já estabelecidos no uso dos especialistas e que surgem no domínio de especialidade*

Este objetivo foi explanado no capítulo *Apresentação e Parametrização dos Dados*, onde se apresenta as características do *corpus* CEnoLEX, relativo aos glossários, dicionários e outros recursos existentes produzidos pelos especialistas, que contém 26 textos especializados do tipo lexicográfico, recolhidos entre 2012 e 2021 e anotados

gramaticalmente (519 nomes, 429 adjetivos, 64 adjetivos e nomes, 14 verbos). Este *corpus* inclui os descritores que os especialistas consideram relevantes apresentar em obras, apresenta as definições redigidas por especialistas e permite a observação de um conjunto de descritores e definições que os especialistas consideram relevantes apresentar em obras de cariz lexicográfico.

**Objetivo 6.** *Apresentar uma proposta de dicionário, com pressupostos teóricos e metodológicos que privilegiam uma abordagem abrangente tendo em conta a especificidade de modalidades sensoriais invocadas*

Esta tarefa foi descrita no capítulo *Proposta de Dicionário de Especialidade da ASE*, que explicita o dicionário de especialidade idealizado para apresentar os dados recolhidos sobre este domínio de especialidade, tendo em conta o trabalho de diversos autores que contribuíram de forma significativa para a sua definição e para o desenvolvimento no âmbito da Dicionarística. Apresenta-se uma proposta para uma versão em papel e um protótipo de uma aplicação digital, com base em procedimentos científicos e testados. A proposta apresentada surge, assim, como o resultado da ponderação do estado de arte relativo à Lexicografia de Especialidade, bem como do trabalho de análise que fundamenta e justifica as opções tomadas.

**Objetivo 7.** *Fornecer uma base sólida para um recurso de especialidade útil a um público especialista e semiespecialista da área da ASE*

A prossecução deste objetivo foi explanada em todos os capítulos da tese, na medida em que todos os passos seguidos contribuem para ele. A compilação de dois *corpora* permitiu, por um lado, observar as unidades e as suas definições atestadas e reconhecidas pelos especialistas, no CEnoLEX, e, por outro lado, verificar o uso das unidades lexicais efetivamente utilizadas em contexto de especialidade real, no CEnoTEXT. A confrontação de ambos permite a comparação de uma abordagem partindo dos dados (direcionada pelo *corpus*, funcionando como *corpus* de estudo) *versus* uma abordagem pela listagem de descritores dos especialistas (funcionando como *corpus* de referência),

permitindo chegar a uma lista de candidatos a descritores e a um conjunto de reflexões sobre o material recolhido, selecionado e analisado.

Adicionalmente, a opção de não só fazer um estado de arte sobre o domínio, mas também recolher, de facto, as obras disponíveis no mercado e compilar um *corpus* para utilizar em análises contrastivas é um trabalho exaustivo que permite contrastar não só os descritores e as definições, mas também observar e tentar compreender que critérios têm sido utilizados na criação dessas obras de especialidade. Ainda que neste trabalho a análise tenha incidido com maior destaque nos descritores, a sua compilação possibilitou um olhar sobre vários aspetos – as definições que são apresentadas, as falhas, as circularidades, as categorias usadas, os subdomínios listados, o tratamento da sinonímia, entre outros.

O trabalho de compilação e análise realizado permitiu, ainda, uma sólida reflexão sobre preocupações lexicográficas a considerar, sobre o que foi feito nas obras existentes e sobre como uma nova metodologia, aplicada em diferentes momentos do processo de criação de um dicionário de especialidade, pode ser benéfica. A proposta de dicionário de especialidade apresentada surgiu assim de uma metodologia que inclui três tipos de análises: gramatical, contrastiva e com base na modalidade sensorial. Cada análise gerou uma lista de candidatos a ULE, e o processo de seleção desses candidatos segue uma abordagem não tradicional para a escolha da nomenclatura do dicionário especializado. O protótipo apresentado, além de seguir procedimentos e critérios lexicográficos rigorosos, reflete ainda a utilização de técnicas e melhores práticas de *design* atuais, com vista a uma melhor interação entre o produto (o dicionário) e o utilizador e perspetivando futuros desenvolvimentos e inovações.

## 7.2. DESAFIOS E TRABALHO FUTURO

Ao refletir sobre os resultados alcançados, é importante reconhecer as limitações que este trabalho enfrentou e detetou. O domínio relativo ao vinho é um tema vasto e como tal, estabelecer fronteiras e delimitar o escopo do trabalho é desde logo o primeiro desafio, pois há sempre aspetos que seria interessante trabalhar, mas que, em alguns casos, acabam por dispersar o foco do trabalho. Como tal, assumir opções metodológicas para

circunscrever o *corpus* e a própria análise parece redutor, mas é indispensável. Um trabalho que associa a Lexicografia a um domínio como a ASE é um trabalho complexo e desafiador, que envolveu diversas questões e dificuldades:

- Especialistas: a primeira dificuldade está relacionada com a colaboração de especialistas. A extração, a conversão e a codificação de informação provenientes de especialistas em vinho requer um alto nível de conhecimento técnico e de especialidade. Definir critérios precisos para categorizar e subcategorizar as unidades candidatas para as fichas lexicográficas também é um desafio, uma vez que as nuances da terminologia do vinho podem ser subtis e variáveis entre especialistas. O trabalho aqui presente pretendeu elaborar todo o processo de organização de *corpora*, extração e tratamento de candidatos, determinação de macro e microestruturas, mas não envolveu especialistas durante estas tarefas. Acredita-se que poderia ter sido benéfico o seu envolvimento, mas também se considera que apresentar a listagem final a validação por especialistas pode ser o passo seguinte a considerar na elaboração do DEASE.
- Definições lexicográficas: a evolução das definições lexicográficas ao longo do tempo é uma questão crítica. Unidades lexicais como “escanção” e “sommelier” podem ter definições antigas e novas, tornando a interpretação subjetiva e/ou hedónica. Além disso, a comparação entre definições antigas e novas, como se verificou para o descritor “tranquilo” cuja definição sofreu alterações drásticas de 1967 para 2013, revela mudanças na perceção e no uso das palavras ao longo do tempo. Em nenhuma das obras consultadas no CEnoLEX havia marca ou data de atestação, o que pode dificultar a compreensão de uma definição e alterar a sua apropriação ao momento presente. No entanto, no objetivo deste trabalho não era visada a validação diacrónica, apesar de se considerar esta que poderá ter um impacto relevante na compreensão das diferentes definições que foram encontradas e que deverá ser sempre considerada no futuro. Outro aspeto em relação às definições relaciona-se com a dificuldade de estabelecer com maior clareza a diferença entre níveis de especialização na definição (aspeto ligeiramente abordado neste trabalho), ou seja, quais os aspetos a ter em consideração ao criar uma definição para um público especialista,

semiespecialista ou não-especialista. Considerando que se está perante um recurso de especialidade, há que refletir sobre a complexidade ou simplicidade de definição para que esta possa ser acessível sem que se perca o carácter de especialidade.

- Modalidade sensorial: a categorização sensorial é uma área especialmente complexa. Neste trabalho, optou-se por não se tratar de forma exaustiva questões como a ambiguidade sensorial. Por exemplo, para descritores como “rubi” ou “baunilha”, é clara a sua associação à modalidade visual e olfativa, respetivamente. Mas em casos como “harmonia” ou “elegante” parece haver um nível diferente de categorização além da modalidade. Outro aspeto importante na categorização dos descritores é como tratar a multissensorialidade. Se nos verbos este pareceu um processo mais ou menos fácil, nos nomes e nos adjetivos surgiram exemplos que apresentam maior complexidade ao nível da definição, mas também da ligação ou da dependência sensorial. Por exemplo, um descritor como “citrino” está relacionado com as modalidades visual, olfativa e gustativa, mas as relações de marcação de outras palavras relacionadas gera dificuldade, principalmente na forma como apresentar essa informação ao público.
- Uso: a categorização de ULE com base no seu uso é crucial, mas determinar com exatidão o que deve constar na definição é sempre um desafio. O que é considerado uma definição para especialista, para semiespecialista e não-especialista pode variar daquilo que se considera informação obrigatória e informação necessária, sendo que informação adicional pode ser de extrema importância para os diferentes públicos. Aspetos como o nome químico associado a um dado descritor ou a apresentação de uma fórmula química são obrigatórios, necessários ou adicionais? Como apresentar essa informação, sem extravasar os limites da informação lexical *versus* enciclopédica num recurso de especialidade? Por exemplo, determinar, à partida, listagens de tonalidades ou de nomes de famílias de aromas pode ser uma solução, mas apresenta outros desafios relacionados com fontes e variação (quer teórica no domínio, quer linguística).

- Escalas e representação do significado: a definição de escalas é uma tarefa complexa. A seleção de uma escala para um dado aspeto depende necessariamente do conhecimento que se tem dessa realidade. Por exemplo, há diferentes tipos de escalas que podem ser usadas: (i) as escalas de polaridade positiva/negativa que envolvem a avaliação da existência ou ausência de uma característica e que tem dois extremos opostos, um positivo e outro negativo. Por exemplo, a escala relativa à presença ou ausência de teor açúcar residual (“um vinho doce” com grande percentagem de teor açúcar residual *versus* “um vinho seco” com pouco ou nenhum açúcar residual); (ii) as escalas de grau ou intensidade, que incluem vários pontos intermediários entre os extremos opostos, para permitir uma avaliação mais precisa da característica. Por exemplo, a escala também relativa à presença de doçura “vinho doce, meio doce, seco, extra seco”, a par de outras estratégias como o uso de “muito, ideal, pouco”. Assim, determinar que escalas utilizar associadas a determinados descritores e estabelecer uma consistência de uso adequada à sua apresentação no dicionário obriga a reflexões e análises específicas.
- Fatores culturais: garantir a adequação da informação é um desafio. A adequação de sinónimos, exemplos, analogias e metáforas na descrição dos descritores pode ser subjetiva e complexa. A referência a objetos do mundo real, como “olho-de-perdiz”, “limão”, “jasmim” ou “caramelo”, para explicação das sensações requer cuidado, uma vez que essas referências podem variar culturalmente, como no exemplo “ginja”. Ao descrever a perceção, a cultura (a lexicultura) deve ser tida em conta, pois é fundamental para a descrição de cor, tonalidade, sabor e/ou intensidade do vinho. Os elementos descritivos podem variar de acordo com fatores culturais e requerem um cuidado extra que também, necessariamente, delimita o público-alvo do dicionário.

A proposta de DEASE apresentada, apesar de não pretender responder cabalmente às questões listadas anteriormente, reflete sobre a forma como apresentar o conteúdo adequado de um determinado domínio (ASE) a um público para que este o consiga

descodificar com facilidade. Apresentar uma unidade lexical de especialidade seguida da polissemia adequada ao contexto do domínio implica saber que determinadas unidades podem ser mal interpretadas. No entanto, o que se sabe é que as pessoas não pesquisam nos dicionários o que acham que sabem, mas apenas aquilo que lhes provoca dúvidas ou não sabem. Neste contexto, e direcionando o DEASE também para o público especialista, a proposta apresentada permite sensibilizar os especialistas produtores de textos para essas polissemias e permite-lhes reduzir a ambiguidade e ser mais eficazes na comunicação.

Os resultados do trabalho que aqui se apresenta podem servir de base sólida e robusta a um conjunto de outros trabalhos a desenvolver no domínio do vinho. São várias as extensões e aplicações diretas dos resultados tangíveis deste trabalho:

- A BDEno pode servir de base para (i) a compilação de *corpora* paralelos com outras línguas; (ii) a elaboração de outros produtos lexicográficos especializados, utilizando a mesma metodologia, para, por exemplo, criar um dicionário de neologismos, e/ou de empréstimos.
- O CEnoLEX permite (i) o desenvolvimento de estudos sobre a variação lexical diacrónica (permitindo estudos diacrónicos para conseguir atestações dos descritores, das alterações das definições, etc.), (ii) a construção de recursos orientados para diferentes tipos de utilizador (utilizador especialista, semiespecialista e não-especialista) e, tendo em conta esse aspeto, o ajuste do tipo de informação a apresentar (o tipo de descritores, as definições, a organização dos descritores e seus sinónimos, as escalas de intensidade, etc.).
- o CEnoLEX e o CenoTEXT permitem (i) a utilização dos *corpora* para treinar a interpretação de padrões de reconhecimento sensorial a aplicar em narizes e línguas eletrónicos (e/ou biónicos); (ii) a utilização dos dados em estudos sobre discurso sensorial em PE e verificar como este se constrói e se manifesta, por oposição a todo um conjunto de descritores que indicam castas, denominações de origem demarcada e que têm constrições legais internacionais e nacionais associadas.

- O protótipo de Aplicação Digital do DEASE permite desenvolver trabalho dedicado à noção de que informação deve (e o modo como) deve estar configurada (utilização, visualização, acessibilidade), considerando, (i) do ponto de vista do utilizador, a utilização de um modelo de dicionário que tem de se adaptar a uma imagem estereotípica de dicionário digital para fazer a ponte com novos utilizadores, mas que permite uma boa organização dos dados e a flexibilidade da interface; (ii) do ponto de vista da criação de recursos de especialidade, a criação de uma plataforma digital alargada aos domínios da vinha e do vinho, para disponibilização de recursos (dicionários de profissões e ofícios, bases de dados de objetos e utensílios associados ao vinho, etc.). As duas possibilidades integrariam um trabalho multidisciplinar, tendo como parceiros museus e cooperativas, e impulsionando a ligação do trabalho científico e metodológico com as instituições que facilitam a preservação/disseminação da cultura para o cidadão.

O estudo aqui apresentado demonstra, assim, que uma estratégia metodológica mais abrangente aplicada ao léxico de especialidade permite a sua replicabilidade a outros domínios, que também utilizam a AS. Por exemplo, e considerando o caso português, produtores de produtos criados e comercializados – vinho do Porto, vinho da Madeira, vinho Verde, ou Licores (Fedrisco, Ginja de Alcobaça, Ginja de Óbidos, Granito Montemorense, Jeropiga, Licor de Amora, Licor de Poejo, Medronho do Algarve, Poncha da Madeira, Rum da Madeira) – com um potencial cultural muito grande, que se apoiam em notas de prova para classificar e descrever os seus produtos.

Finalmente, o trabalho desenvolvido constitui, ainda, uma base de dados para treinar sistemas de IA na compreensão e processamento de linguagem natural em AS pela organização da informação dos descritores; pela indicação de marcas e variantes ortográficas dos descritores; pelas definições precisas das ULE, que seriam difíceis de encontrar noutras fontes; pelos exemplos de uso e pelas informações sobre contextos em que as unidades são usadas, que são dados de alta utilidade para sistemas de IA; pela relação de ULE com significados semelhantes; e pela categorização dos sentidos das ULE por modalidade e submodalidade sensorial.

### 7.3. CONSIDERAÇÕES FINAIS

Apesar de vivermos hoje num mundo digital que cada vez mais se apoia na pesquisa de conteúdo em *browsers* e por assistentes virtuais, a criação de dicionários desempenha um papel fundamental no desenvolvimento e na preservação da língua, da cultura e do conhecimento. Os dicionários são ferramentas essenciais que não só definem palavras, como também capturam a evolução da língua ao longo do tempo e desempenham um papel crucial em várias áreas: na comunicação, na educação, na preservação cultural, no desenvolvimento de campos de estudo, na acessibilidade ao conhecimento linguístico, na inovação linguística. Eles são, por isso, recursos versáteis e essenciais em várias áreas da vida quotidiana e da sociedade. Já os dicionários de especialidade desempenham um papel vital nos domínios de especialidade, fornecendo um conjunto de benefícios específicos e essenciais para campos de conhecimento especializados.

Com o propósito de apresentar uma proposta de dicionário de especialidade, selecionou-se um domínio de especialidade pouco estudado que constitui um tipo de língua especializada em Enologia. Pretendeu-se contribuir para o desenvolvimento de uma linguagem descritiva objetiva e consensual, com maior grau de concordância de julgamentos entre os especialistas, contribuindo de modo objetivo e inovador para uma futura normalização e estabilização do léxico usado e servindo de instrumento de trabalho na comunicação e na aprendizagem da área, diminuindo o ruído de comunicação entre aqueles que usam esta língua de especialidade em PE.

Em síntese, esta tese de doutoramento apresenta-se como um contributo para várias áreas. Na Lexicografia, ela resulta na criação de um produto voltado para um público diversificado, com o objetivo de ser implementado e utilizado por especialistas, mas também por não-especialistas. Além disso, envolve a organização dos dados, análise e comentários de obras publicadas no mercado, permitindo a identificação de lacunas, pontos fortes e pontos fracos. Uma proposta para a criação de um dicionário especializado em enologia pode ter um impacto na compreensão desse domínio. A conceção e desenvolvimento de um dicionário dedicado à enologia representa um avanço significativo na organização do léxico específico desse domínio. Esse dicionário pode atender às necessidades de especialistas, estudantes e amantes do vinho, fornecendo definições precisas, contextos de uso e informações detalhadas sobre os descritores

relacionados à vinificação e degustação de vinhos. Além disso, ao preencher uma lacuna no léxico técnico, a proposta ajuda a contribuir para a padronização e para a disseminação do conhecimento na enologia, facilitando a comunicação e o compartilhamento de informações.

Na área da Linguística de *Corpus*, a tese utiliza uma metodologia estruturada para recolher e selecionar fontes textuais, criando *corpora* que podem servir como diretrizes para outros investigadores na criação de recursos semelhantes. Os *corpora* gerados são únicos, apresentando características especiais e uma riqueza de textos recolhidos de diversas fontes, tornando-os heterogêneos, sincrônicos e autênticos. Além disso, esses *corpora* podem desempenhar um papel pedagógico, auxiliando estudantes e interessados no domínio, como ferramentas de ensino e apoio em atividades de aprendizagem. Eles podem, também, ser usados como produtos de comparação com outros *corpora* que representam diferentes variedades do português.

Na sociedade, esse dicionário especializado pode promover o interesse e a apreciação do vinho, tornando-o mais acessível a um público amplo. Ao fornecer informações claras e acessíveis sobre o mundo dos vinhos, ele pode enriquecer a experiência de degustação. Além disso, a disponibilidade de recursos lexicográficos especializados pode fomentar o respeito e a valorização da cultura enológica portuguesa.

## 8. REFERÊNCIAS BIBLIOGRÁFICAS

- Abreu, S. P. de, & Lara, L. Z. (2013). Considerações sobre a formação lexical de verbos denominais em linguagem de especialidade. *Letras De Hoje*, 35(3). Recuperado de <https://revistaseletronicas.pucrs.br/ojs/index.php/fale/article/view/14787>
- Ackerman, D. (1990). *Uma História Natural dos Sentidos*. Lisboa, Portugal: Círculo de Leitores.
- Adorne, F. C. (2014). *Terminologia da gestão pública da cultura no Brasil: Proposta de glossário*.
- Afonso, J. (2013). *Curso de vinho para verdadeiros apreciadores*. Lisboa, Portugal: A Esfera dos Livros.
- Alba, J. W., & Hutchinson, J. W. (1987). Dimensions of Consumer Expertise. *Journal of Consumer Research*, 13(4), 411-454. <https://doi.org/10.1086/209080>
- Al-Kasimi, A. M. (1977). *Linguistics and Bilingual Dictionaries*. Leiden: Brill.
- Almeida, G. M. B., Souza, D., Pino, D. (2007). A definição nos dicionários especializados: proposta metodológica. *Debate Terminológico*, v. 03, p. 01-20.
- Almeida, M. J. (2016). *O Vinho na Ponta da Língua*. (1ª ed.) Porto Salvo, Portugal: Saída de Emergência.
- Álvarez, E.L. (2019). *A Corpus Analysis of English Phraseology in Court Judgments of the European Union*. Universidad de Valladolid. Facultad de Filosofía y Letras, España.
- Alvelos, H. (2002). *Análise, Desenvolvimento e Teste de Métodos e Técnicas para Controlo Estatístico em Análise Sensorial*. Tese de Doutoramento, Faculdade de Engenharia da Universidade do Porto. Disponível em: <https://hdl.handle.net/10216/12427>
- Alves, I. M. (2006). A observação sistemática da neologia lexical: subsídio para o estudo do léxico. *Alfa*, 50(2), 131-144. Disponível em <http://seer.fclar.unesp.br/alfa/article/view/1415/1116>
- Amoraritei, L. (2002). La métaphore en Oenologie. *METAPHORIK.DE* 3, 1-12. Disponível em <http://www.metaphorik.de/03/amoraritei.htm>
- Amsler, R. (2012). *FW: Macmillan's recent announcement*. Disponível em <https://www.freelists.org/post/euralex/The-future-of-our-field-Part-1>
- Anders Ericsson, K., & Towne, T.J. (2010). Expertise. *WIREs Cognitive Science*, 1(3), 404-416. <https://doi.org/10.1002/wcs.47>
- Arabin, B. (2002). Music during pregnancy. *Ultrasound Obstetric Gynecology*, 20, 425-430.
- Aragon Cobo, M. (2005). La contribution de la pragmatique lexicoculturelle dans un dictionnaire du tourisme. *Meta*, 50(4). <https://doi.org/10.7202/019908ar>
- Arshamian, A., & Larsson, M. (2014). Same same but different: the case of olfactory imagery. *Frontiers in Psychology*, 5(34), 1-8. <https://doi.org/10.3389/fpsyg.2014.00034>

- Atkins, B., & Rundell, M. (2008). *The Oxford Guide to Practical Lexicography*. New York: Oxford University Press.
- Aurélio. (2022). *Dicionário Aurélio da Língua Portuguesa*. 5. ed. Curitiba: Positivo.
- Azodanlou, R. (2001). *A methodology for assessing the quality of fruit and vegetables*. Doctoral Thesis. Swiss Federal Institute of Technology Zurich.
- Baldinger, K. (1960). Alphabetisches oder begrifflich gegliedertes Wörterbuch? *Zeitschrift für romanische Philologie* 76, 521-536.
- Barbero, C., & Amaro, R. (2020). Exploração de *corpora* para extração e descrição de léxico de especialidade: para uma metodologia sólida e sustentada. *Linha D'Água*, 33(1), 69-104. <https://doi.org/10.11606/issn.2236-4242.v33i1p69-104>
- Barbosa, L. M. (2009). O conceito de lexicultura e suas implicações para o ensino-aprendizagem de português língua estrangeira. *Filologia e Linguística Portuguesa*, n.º 10-11, 31-41. <https://doi.org/10.11606/issn.2176-9419.v0i10-11p31-41>
- Barbosa, M. A. (1995). Contribuição ao estudo de aspectos da tipologia de obras lexicográficas. *Ciência Da Informação*, 24(3). Disponível em <http://revista.ibict.br/CIINF/article/view/572>
- Barbosa, S. (2016). *EnoTerm: Da Lexicografia de Especialidade à elaboração de um dicionário de degustação do vinho*. Poster session presented at Encontro CIÊNCIA'16 – Encontro com a Ciência e a Tecnologia em Portugal, Lisboa, Portugal. <http://www.ciencia2016.pt/img/posters/158.pdf>
- Barbosa, S. (2021). An exploratory study about the wine tasting terminology to non-expert wine drinkers. *Academic Journal of Modern Philology*, Vol. 13 (2021) Special Issue, 15–30. ISSN 2299–7164. <https://doi:10.34616/ajmp.2021.13>
- Barbosa, S. & Barbero, C. (2021). Ver as cores através das definições dos dicionários de português. In C. Barbero, & M. Tomaz (Eds.), *Textos Seleccionados do XIII e XIV Fórum de Partilha Linguística* (pp. 85-99). NOVA FCSH/CLUNL. <https://doi.org/10.34619/j1un-mh49>
- Barbosa, S., & Lino, M. T. R. F. (2019). *An exploratory study about the wine tasting terminology to non-expert wine drinkers*. Abstract from TERMOS 2020 Second Wrocław Terminological Meeting, Wrocław, Poland.
- Barbosa, S. & Santos, C. (2021). *UX Design Aplicado ao Dicionário de Especialidade da Análise Sensorial Enológica*. Jornadas dos Dicionários 2021 – Lexicografia e dicionarística de especialidade. CLUNL NOVA FCSH, online, 12 de novembro de 2021.
- Barros, L. A. (2004). *Curso básico de Terminologia*. São Paulo: Editora da Universidade de São Paulo.
- Béjoint, H. (2000). *Modern Lexicography: An Introduction*. New York: Oxford University Press.
- Béjoint, H. (2010). *The Lexicography of English*. Oxford: Oxford University Press.

- Bengtsson, K., & Helm, E. (1946). *Principles of taste tasting*. New York: Wallerstein.
- Berber Sardinha, T. (2000). Lingüística de Corpus: histórico e problemática. *Documentação de Estudos em Lingüística Teórica e Aplicada*, 16(2), 323-367. Disponível em [http://www.scielo.br/scielo.php?script=sci\\_arttext&pid=S0102-44502000000200005&lng=en&tlng=pt.10.1590/S0102-44502000000200005](http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0102-44502000000200005&lng=en&tlng=pt.10.1590/S0102-44502000000200005).
- Bergenholtz, H. (1995). Wodurch Wodurch unterscheidet sich Fachlexikographie von Terminographie? *Lexicographica*, 11, 37-46.
- Bergenholtz, H., & Gouws, R. (2007). The Access Process in Dictionaries for Fixed Expressions. *Lexicographica. International Annual for Lexicography* 23, 237-260.
- Bergenholtz, H., & Gouws, R. H. (2012). What is Lexicography? *Lexikos*, 22, 31-42. doi: <https://doi.org/10.5788/22-1-996>
- Bergenholtz, H., & Tarp, S. (1995). *Manual of Specialised Lexicography. The preparation of specialised dictionaries*. Amsterdam-Philadelphia, John Benjamins.
- Bergenholtz, H., & Tarp, S. (2002). Die moderne lexikographische Funktionslehre. Diskussionsbeitrag zu neuen und alten Paradigmen, die Wörterbücher als Gebrauchsgegenstände verstehen. *Lexicographica* 18, 253-263.
- Bergenholtz, H.; Tarp, S. (2010). LSP Lexicography or Terminography? The lexicographer's point of view. En: Fuerte – Oliveira, P. A. (org). *Specialized Dictionaries for learners*. Berlin, New York: Walter de Gruyter GMBH. S. CO-KG.
- Beverland, M. B., Lindgreen, A., & Vink, M. W. (2008). Projecting Authenticity Through Advertising: Consumer Judgments of Advertisers' Claims. *Journal of Advertising*, 37(1), 5-15. <https://doi.org/10.2753/JOA0091-3367370101>
- Bevilacqua, C. R. (1993). Tipologia de dicionários. *Cadernos do IL Porto Alegre*, n.º 10, 17-22.
- Biber, D. (1993). Representativeness in Corpus Design. In *Literary and Linguistic Computing*, vol. 8, n. 4. Disponível em <http://staff.um.edu.mt/albert.gatt/teaching/dl/biber93.pdf>
- Biber, D., Conrad, S., Reppen, R., Byrd, P., Helt, M., Clark, V., Cortes, V., Csomay, E., & Urzua, A. (2004). *Representing language use in the university: Analysis of the TOEFL 2000 spoken and written academic language corpus*. Test of English as a Foreign Language.
- Biderman, M. T. C. (1984). *A ciência da Lexicografia*. São Paulo: Alfa.
- Biderman, M. T. C. (1998). As ciências do léxico. In A.M.P.P. Oliveira & A. N. Isquendo, (Eds), *As ciências do Léxico: Lexicologia, Lexicografia e Terminologia* (vol.1, pp. 13-22). Campo Grande, MS: Editora UFMS.
- Biderman, M. T. C. (2006). O conhecimento, a terminologia e o dicionário. *Ciência e Cultura*, v. 58, p. 35-37.

- Binon, J., & Verlinde, S. (2000). Les langue(s) de spécialité(s): mythe ou réalité? *Lexicographie et langue(s) de spécialité(s). Des mots aux dictionnaires: travaux de la section Lexicologie, lexicographie, onomastique, toponymie* (pp. 616-628). Actes du XXIIe Congrès international de linguistique et philologie romanes (Bruxelles, 23-29 juillet 1998), Tübingen: Niemeyer.
- Boelhouwer, B., Dykstra, A., & Sijens, H. (2018). Dictionary portals. In P. A. Fuertes-Olivera (Ed.), *The Routledge Handbook of Lexicography* (pp. 754-766). London/New York: Routledge.
- Böhm, J., Curvelo-Garcia, A., Cabrita, M.J., Barros, P., Gato, O., & Catarino, S. (2017). *Glossário ilustrado do vinho. Ed. quadrilingue em português, alemão, inglês e francês*. 1.<sup>a</sup> ed. Lisboa: Dinalivro.
- Boisson, C. (1996). L'antiquité et la variété des dictionnaires bilingues. In H. Béjoint (Ed.), *Les dictionnaires bilingues* (pp. 17-30). Louvain-la-Neuve, Belgique: De Boeck Supérieur. <https://doi.org/10.3917/dbu.bejoi.1996.01.0017>
- Boisson, C., Kirtchuk, P., & Béjoint, H. (1991). Aux origines de la lexicographie : les premiers dictionnaires monolingues et bilingues. *International Journal of Lexicography*, 4/4, 261-315.
- Borba, F. S. (2003). *Organização de dicionários: uma introdução à lexicografia*. São Paulo: Editora UNESP.
- Boulanger, J.-C. (1995). *Lexicographie générale*. [Material de apoio do Curso de Especialização ministrado na Universidade de Brasília.]
- Boutin-Quesnel, R., Bélanger, N., Kerpan, N., & Rousseau, L.-J. (1985). *Vocabulaire systématique de la terminologie*. Québec: Publications du Québec, 1985. (Cahiers de l'Office de la Langue Française).
- Bowcott, O. (2017). *Bosnian Croat war criminal dies after taking poison in UN courtroom*. The Guardian. Disponível em <https://www.theguardian.com/law/2017/nov/29/un-war-crimes-defendant-claims-to-drink-poison-at-trial-in-hague-slobodan-praljak>
- Breit, W. B. (2013). The wine tasting sheet as a linguistic genre and its appraisal theory based analysis. *Revista Española de Lingüística Aplicada*, 1, 71-90.
- Breton, D. (2016). *Antropologia dos Sentidos*. Tradução Francisco Morás. Petrópolis, RJ: Vozes.
- Brochet, F., & Dubourdieu, D. (2001). Wine descriptive language supports cognitive specificity of chemical senses. *Brain and Language*, 77(2), 187-196. <https://doi.org/10.1006/brln.2000.2428>
- Bronckart, J.-P. (1997/2003) *Atividade de linguagem textos e discursos: Por um interacionismo sócio-discursivo*. Trad. A. R. Machado & P. Cunha. São Paulo: EDUC.
- Brown, P. (2013). *The murder of Cleopatra: history's greatest cold case*. Amherst, NY: Prometheus Books.
- Bruce, N. (1999). Classification and hierarchy in the discourse of wine: Émile Peynaud's *The Taste of Wine*. *ASp. la revue du GERAS*, 23, 149-164.

- Burenhult, N., & Majid, A. (2011). Olfaction in Aslian ideology and language. *The Senses & Society*, 6, 19-29. doi:10.2752/174589311X12893982233597
- Buzan, T., & Buzan, B. (1996). *El libro de los mapas mentales*, Urano, Barcelona.
- Caballero, R. (2007). Manner-of-motion verbs in wine description. *Journal of Pragmatics*, 39(12), 2095-2114.
- Cabré, M. T. (2005) La terminología, una disciplina em evolución: pasado, presente y algunos elementos futuros. *Revista Debate Terminológico*. n.1-3. Disponível em [www.riterm.net/revista/n\\_1/cabre.pdf](http://www.riterm.net/revista/n_1/cabre.pdf)
- Cabré, M. T. (2008). El principio de poliedricidad: la articulación de lo discursivo, lo cognitivo y lo lingüístico en Terminología. *Pt. 1. Ibérica*, Castelló de la Plana, v. 16, p. 9-36.
- Cabré, M. T. (1999). *La Terminología: representación y comunicación: elementos para una teoría de base comunicativa y otros artículos*. Girona: Documenta Universitaria. Barcelona.
- Calderón Campos, M. (1994). *Sobre la elaboración de diccionarios monolingües de producción, las definiciones, los ejemplos y las colocaciones léxicas*. Tese de Doutoramento. Granada: Universidad de Granada.
- Calepino, A. (1502). *Vocabularius: thesaurus copiosissimus. Toscolano: apud Benacum in aedibus Alexandri Paganini*.
- Camacho Niño, J. (2019). Propuesta microestructural para un diccionario de especialidad lexicográfica. *ELUA*, 33: 55-72. doi: 10.14198/2019.33.3
- Camilotti, F. (2011). *Inclusão e tratamento de termos técnico científicos em dicionários escolares: um estudo crítico*. Dissertação de Mestrado, Universidade do Vale do Rio dos Sinos, São Leopoldo.
- Cardebat, J., & Livat, F. (2016). Wine experts' rating: a matter of taste? *International Journal of Wine Business Research*. Vol. 28(1), 43-58. <https://doi.org/10.1108/IJWBR-04-2015-0011>
- Cardoso, A. D. (2007). *O Vinho – Da Uva à Garrafa*. Coimbra: Âncora Editora.
- Carter, R., & McCarthy, M. (2013). *Vocabulary and Language Teaching*. Routledge: New York.
- Carvalho, J. (2009). *Análise Sensorial*. [PowerPoint de apoio ao cidadão disponibilizado em Direção Regional de Agricultura e Pescas do Centro]. Disponível em [http://www.drapc.min-agricultura.pt/base/geral/files/fermentacoes\\_2009.pdf](http://www.drapc.min-agricultura.pt/base/geral/files/fermentacoes_2009.pdf)
- Carvalho, J. (2012). *Introdução à prova de vinhos*. [PowerPoint de apoio ao cidadão disponibilizado em Direção Regional de Agricultura e Pescas do Centro]. Disponível em [http://www.drapc.min-agricultura.pt/base/geral/files/fermentacoes\\_2009.pdf](http://www.drapc.min-agricultura.pt/base/geral/files/fermentacoes_2009.pdf)
- Carvalho, J. (1973). *Teoria da Linguagem*, tomo I, 11-26. Coimbra: Atlântida Editora.

- Cavero, M. E. (2010). Sémantisme autour du vin: représentations symboliques et lexicoculturelles. *Anales de Filología Francesa*, 18, 183-202. Disponível em <http://dialnet.unirioja.es/descarga/articulo/3709848.pdf>
- Cayot, P., Gautier, L., & Soubrier, J. (2009). Définir, dénommer et traduire l'onctuosité: regards croisés sur une propriété gustative et ses dénominations en français, anglais et allemand. In D. Beltran-Vidal (Ed.), *Les mots de la santé (2): affaire(s) de goût(s)* (pp. 189-222), Dec 2006, Lyon, France. Centre de Recherche en Terminologie et Traduction. [\(hal-02047063\)](#)
- Challe, O. (2000). *Le français de spécialités*, Paris: CLE International.
- Charters, S., & Pettigrew, S. (2006). Conceptualizing product quality: the case of wine. *Marketing Theory*, 6(4), 467-483. <https://doi.org/10.1177/1470593106069932>
- Cherifi, N. (2009). L'écart culturel dans les dictionnaires bilingues: dictionnaires français-arabe, arabe-français. *Études de linguistique appliquée*, 2(2), 237-248. <https://doi.org/10.3917/ela.154.0237>
- Chi, M. (2006). Laboratory Methods for Assessing Experts' and Novices' Knowledge. In K. A. Ericsson, N. Charness, P. J. Feltovich & R. R. Hoffman (Eds.), *The Cambridge handbook of expertise and expert performance* (pp. 167-184). New York, NY, US: Cambridge University Press.
- Chollet, S., & Valentin, D. (2000). Expertise level and odour perception: What can we learn from red burgundy wines? *Annee Psychologique*, 100(1), 11-36.
- Ciobanu, G. (2003). Peculiarities of Terminography. *Boletim científico*, Tom 2. (Série Limbi moderne)
- Clark, S., Costelle, M., Drake, M., & Bodyfelt, F. (2008). *The Sensory Evaluation of Dairy Products* (2nd ed.). New York: Springer Science + Business Media.
- Conceição, M. C. (2005). Termos em discurso. In *Des(a)fiando discursos: Homenagem a Maria Emília Ricardo Marques* (pp. 247-254). Lisboa: Universidade Aberta.
- Contente, M. (2008). *Terminocriatividade, Sinonímia e Equivalência Interlinguística em Medicina*, Edições Colibri, Lisboa.
- Coradini, H. (1999). *Metalinguagem na obra "De lingua Latina" de Marcos Terêncio Varrão*. Tese de Doutorado, Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo, São Paulo. doi:10.11606/T.8.1999.tde-22082013-103553
- Correia, M. (2009). *Os Dicionários Portugueses*. Lisboa: Caminho.
- Costa, L. A. C. (2015). *Reflexões sobre a variação terminológica na lexicografia corrente no Brasil e a construção das bases teórico-metodológicas para o dicionário de Lexicografia brasileira*. Tese de Doutorado, Universidade Estadual Paulista, São José do Rio Preto; Universidad Pompeu Fabra, Barcelona.

- Costa, L. A. C. (2015). Terminografia *versus* lexicografia especializada: questões concernentes à produção de dicionários especializados e as bases teórico-metodológicas do dicionário de lexicografia brasileira. *Debate Terminológico*. No. 13; pp. 43-53.
- Coutier, M. (1994). Tropes et termes: le vocabulaire de la dégustation du vin. *Meta: Journal des traducteurs*, 39(4), 662-675.
- Coutier, M. (2007). *Dictionnaire de la langue du vin*. Paris: CNRS.
- Croijmans, I. (2018). *Wine expertise shapes olfactory language and cognition*. Tese de Dutoramento, Radboud University, Nijmegen.
- Croijmans, I., & Majid, A. (2016). Not All Flavor Expertise Is Equal: The Language of Wine and Coffee Experts. *PLoS ONE*, 11(6), e0155845.
- Crystal, D. (2001). Linguistics: Overview. In *International Encyclopedia of the Social & Behavioral Sciences* (pp. 8948-8954). Elsevier. <https://doi.org/10.1016/b0-08-043076-7/02936-3>
- Gromann, D., Schnitzer, J. (2016). Where do Business Students Turn for Help? An Empirical Study on Dictionary Use in Foreign-language Learning, *International Journal of Lexicography*, Volume 29, Issue 1, March 2016, Pages 55–99, <https://doi.org/10.1093/ijl/ecv027>
- Dalmasso, G. (1963). *Lexique de la vigne et du vin : français- italieno- español- deutsch- portugês- english- pycckn*. Office International de la vigne et du vin, Paris.
- Dapena, J.A.P. (2002). *Manual de técnica lexicográfica*. Madrid: Arco/libros, S.L.
- Dayan, A.D. (2009). What killed Socrates? Toxicological considerations and questions. *Postgraduate Medical Journal* 85:34-37.
- De Gruyter. (2013). *Dictionaries. an international encyclopedia of lexicography : supplementary volume: recent developments with focus on electronic and computational lexicography*. <https://doi.org/10.1515/9783110238136>
- De Schryver, G-M. (2003). Lexicographers' Dreams in the Electronic-Dictionary Age. *International Journal of Lexicography* 16.2: 143–199.
- Demers, M.-C., Kernerman, I., & L'Homme, M.-C. (2012). Lexicographic interchange between a specialized and a general language dictionary. In R. Vatvedt Fjeld & J. M. Torjusen (Eds.), *Proceedings of the 15th EURALEX International Congress* (pp. 750-757). Department of Linguistics and Scandinavian Studies, University of Oslo. ISBN: 978-82-303-2228-4.
- Deneulin, P., Le Fur, Y., Le Bras, G., Bavaud, F., & Gautier, L. (2014). La minéralité du vin: représentations mentales de consommateurs suisses et français. *Revue Suisse de Viticult., Arboricult., Horticulture* 46, 174-180.
- Depecker, L. (2002). *Entre signe et concept: éléments de terminologie générale*. Paris: Presses Sorbonne nouvelle.

- Depledt, F. (1998). *Évaluation sensorielle: manuel méthodologique* (2e édition). Technique & Documentation Lavoisier, Paris.
- Depledt, F., & Sauvageot, F. (2002). *Évaluation sensorielle des produits alimentaires. Techniques de l'ingénieur traité agroalimentaire*, F4000, vol.1, 7pp. Editions T.I.
- Descamps, J.-L. (1977). *Contribution à l'analyse des discours fonctionnels (Pédagogie des langues de spécialité et lexicographie contextuelle)*. Mémoire de synthèse concernant les travaux présentés en soutenance pour le Doctorat d'État. Université Paris III (Sorbonne Nouvelle).
- Desmet, I. (2002). A análise do sentido em terminologia: teoria e prática da definição terminológica. *TradTerm*, v. 8, p. 169-188.
- Desseauve, T. (1998). *O livro do vinho*. Lisboa: Chaves Ferreira.
- Dicionário de A-Z. (n.d.). Disponível em [http://www.vinhosdoalentejo.pt/detalhe\\_conteudo.php?id=53](http://www.vinhosdoalentejo.pt/detalhe_conteudo.php?id=53)
- Dong, J., Wyss, A., Yang, J., Price, T. R., Nicolas, A., Nalls, M., ... Chen, H. (2017). Genome-Wide Association Analysis of the Sense of Smell in U.S. Older Adults: Identification of Novel Risk Loci in African-Americans and European-Americans. *Molecular neurobiology*, 54(10), 8021-8032. doi:10.1007/s12035-016-0282-8
- Drake, M.A., Drake, S., Bodyfelt, F., Clark, S., & Costello, M. (2008). History of Sensory Analysis. In S. Clark, M. Costello, M. Drake & F. Bodyfelt (Eds.), *The Sensory Evaluation of Dairy Products* (pp. 1-6). Springer, New York, NY. DOI: 10.1007/978-0-387-77408-4
- Drysdale, P. D. (1987). The role of examples in a learner's dictionary. In A. P. Cowie & European Association for Lexicography (Ed.), *The dictionary and the language learner: Papers from the EURALEX Seminar at the University of Leeds, 1-3 April 1985* (pp. 213-223). Berlin, Boston: Max Niemeyer Verlag. <https://doi.org/10.1515/9783111340500-018>
- Dubois, D., & Giboreau, A. (2006). Descriptors: attributes? labels? terms? names?, a contribution of psycho-linguistics to sensory evaluation. *Food Quality and Preference*, 17, 669-672, Elsevier.
- Dubois, J., & Dubois, C. (1971). *Introduction à la lexicographie: Le dictionnaire*. Paris: Larousse.
- Dubuc, R. (1999). *Manual práctico de terminología*. Chile: RiL Editores.
- Duda, W. (1986). Zu einer Theorie der zweisprachigen Lexikographie. Überlegungen zu einem neuen russisch-deutschen Wörterbuch. *Linguistische Studien. Reihe A. Arbeitsberichte*, 142, 1-122.
- Duran, M. S., & Xatara, C. (2007). Lexicografia pedagógica: atores e interfaces. *DELTA: Documentação de Estudos em Lingüística Teórica e Aplicada [online]*, vol. 23, n. 2, 203-222. Disponível em: <<https://doi.org/10.1590/S0102-44502007000200002>>.
- Dziemianko, A. (2012a). On the use(fulness) of paper and electronic dictionaries. In S. Granger and M. Paquot (Eds.), *Electronic Lexicography*, 319-342. <https://doi.org/10.1093/ACPROF:OSO/9780199654864.003.0015>

- Dziemiánko, A. (2012b). Why One and Two Do Not Make Three: Dictionary Form Revisited. *Lexikos* 22, 195-216.
- Dziemiánko, A. (2016). An Insight into the Visual Presentation of Signposts in English Learners' Dictionaries Online. *International Journal of Lexicography* 29: 490-524.
- Dziemiánko, A. (2018). Electronic dictionaries. In P. A. Fuertes-Olivera, (Ed.), *The Routledge Handbook of Lexicography* (pp. 663-683). London/New York: Routledge.
- Eckblad, G. (1981). *Scheme Theory: a Conceptual Framework for Cognitive-Motivational Processes*. London: Academic Press, Inc., Ltd.
- Ellis, H. C., & Hunt, R. R. (1993). *Fundamentals of Cognitive Psychology* (5th ed.). Iowa: Wm. C. Brown Publisher.
- Erickson, R. P. (2008). A study of the science of taste: On the origins and influence of the core ideas. *Behavioral and Brain Sciences* 31, 59-105.
- Escarpit, D. (1995). *La vigne et le vin: Eurolexique*. Presses Universitaires de Bordeaux.
- Esteves, E. (2006). *Notas sobre a seleção e treino de um painel de provadores para análise sensorial de produtos alimentares*. Escola Superior de Tecnologia da Universidade do Algarve. Disponível em <http://w3.ualg.pt/~eesteves>.
- Esteves, E. (2014). *Introdução à Análise Sensorial*. Instituto Superior de Engenharia da Universidade do Algarve, Departamento de Engenharia Alimentar, Faro. Disponível em [http://www.academia.edu/2993383/Introducao\\_a\\_Analise\\_Sensorial](http://www.academia.edu/2993383/Introducao_a_Analise_Sensorial)
- Estienne, R. (1539). *Dictionnaire Francoislatin*.
- Ettinger, S. (1982). La variación lingüística en lexicografía. In G. Haensch, L. Wolf, S. Ettinger & R. Werner (Eds.), *La lexicografía. De la lingüística teórica a la lexicografía práctica* (pp. 359-394). Madrid, Editorial Gredos.
- Fang, C., Li, H.-Y., Li, L., Su, H.-Y., Tang, J., Bai, X., & Liu, H. (2022). Smart Electronic Nose Enabled by an All-Feature Olfactory Algorithm. *Advanced Intelligent Systems*, 2200074. <https://doi.org/10.1002/aisy.202200074>
- Farias, V. S. (2009). O exemplo como informação discreta e discriminante em dicionários semasiólogos de língua portuguesa. *ALFA: Revista De Linguística*, 52(1). Disponível em <https://periodicos.fclar.unesp.br/alfa/article/view/1469>
- Farias, V. S. (2013). *Sobre a definição lexicográfica e seus problemas: Fundamentos para uma teoria geral dos mecanismos explanatórios em dicionários semasiológicos*. UFRGS, Porto Alegre.
- Fathi, B. (2014). Experts and specialised lexicography: perspectives and needs. *Terminàlia*. 9. 12-21.

- Faulstich, E. (2011). Avaliação de dicionários: uma proposta metodológica. *Organon: Revista da Faculdade da Filosofia da Universidade Federal do Rio Grande do Sul*, Porto Alegre, v. 25, n. 50.
- Fernández, D. A. (2003). La lexicografía como disciplina Lingüística. In A. M. Guerra (Ed.) *Lexicografía española* (pp. 33-52). Espanha: Editorial Ariel.
- Ferrara, A. (2009). Le Dictionnaire d'un pays, le pays d'un dictionnaire: Un type particulier de dictionnaire lexiculturel. *Études de linguistique appliquée*, 2(2), 191-203. <https://doi.org/10.3917/ela.154.0191>
- Feu Guijarro, M.J., & Cuadra García, M.T. (2004). Propuesta de un corpus oral bilingüe sobre la cata. In M.P. Battaner & J. A. DeCesaris (Eds.) *De lexicografía: actes del I Symposium Internacional de Lexicografía* (pp. 395-402). (Barcelona, 16-18 de maig de 2002).
- Finatto, M. (2003). A definição de termos técnicocientíficos no âmbito dos estudos de terminologia. *Revista De Estudos Da Linguagem*, 11(1), 197-222. doi:<http://dx.doi.org/10.17851/2237-2083.11.1.197-222>
- Frawley, J. (2017). *The myth of the 'digital native'*. Disponível em <https://educational-innovation.sydney.edu.au/teaching@sydney/digital-native-myth/>
- Freeman, W. J. (1981). A Physiological Hypothesis of Perception. *Perspectives in Biology and Medicine*, 24(4), 561-592. <https://doi.org/10.1353/pbm.1981.0036>
- Freitas, M. (2008). Análise sensorial de alimentos. In *III Simpósio de Controle do Pescado – Segurança Alimentar, Inovação Tecnológica e Mercado*. São Vicente (SP).
- Frøst, M. B., & Noble, A. C. (2002). Preliminary study of the effect of knowledge and sensory expertise on liking for red wines. *American Journal of Enology and Viticulture*, 53(4), 275-284.
- Fuertes-Olivera, P. A. (2009). The Function Theory of Lexicography and Electronic Dictionaries: Wiktionary as a Prototype of Collective Free Multiple-Language Internet Dictionary. In H. Bergenholtz, S. Nielsen & S. Tarp (Eds.), *Lexicography at a Crossroads: Dictionaries and Encyclopedias Today, Lexicographical Tools Tomorrow* (99-134). Bern: Peter Lang.
- Fuertes-Olivera, P. A. (2018). Introduction: lexicography in the Internet era. In P. A. Fuertes-Olivera (Ed.) *The Routledge Handbook of Lexicography* (pp. 1-16). London/New York: Routledge.
- Galisson, R. (1970). *L'apprentissage systématique du vocabulaire*. Paris: Hachette-Larrouse.
- Galisson, R. (1979). *Lexicologie et enseignement des langues*, Paris, Hachette.
- Galisson, R. (1999). La pragmatique lexiculturelle pour accéder autrement à une autre culture par une autre culture, par un autre lexique. *Études de linguistique appliquée*, 116, 477-496.
- Galisson, R., & André, J.-Cl. (1998). *Dictionnaire de noms de marques courants*, Paris: Didier Érudition.

- Galisson, R., & Coste, D. (1983). *Dicionário de Didáctica das Línguas*. Coimbra: Livraria Almedina.
- García Palacios, J. (2002). Dictionnaires y lenguas de especialidad. En I. Ahumada (ed.), *Dictionarios y lenguas de especialidad* (V Seminario de Lexicografía Hispánica. Jaén, 21-23 de noviembre de 2001) (pp. 21-47). Jaén: Universidad de Jaén.
- Gaudin, F., & Guespin, L. (2000). *Initiation à la Lexicologie Française de la Néologie aux Dictionnaires*. Paris: Champs Linguistiques Duculot.
- Gautier, L. (2014). « Des langues de spécialité à la communication spécialisée : un nouveau paradigme de recherche à l'intersection entre sciences du langage, info-com et sciences cognitives ? », in : *Etudes Interdisciplinaires en Francophonie. Sciences humaines* (EIFSH) ; 1, 225-245
- Gautier, L. (2018). Les descripteurs sensoriels à l'épreuve de l'altérité: dimensions linguistiques et culturelles à partir de l'exemple minéral/mineralisch. *Vin et Altérité*, Université de Haute-Alsace, Oct 2018, Mulhouse, France. {halshs-01898585}
- Gautier, L. (2019). *Des constructions à tous les niveaux des discours spécialisés ? Propositions méthodologiques*. Master. Séminaire de linguistique allemand moderne, Paris, France. {halshs-02371684}
- Gautier, L. (2020). Initier à la dégustation ou... enseigner une terminologie de dégustation ? Les termes de la dégustation dans les outils en ligne. In Stengel. *Terminologies gastronomiques et œnologiques. Aspects patrimoniaux et culturels.*, L'Harmattan, pp.137-156, 2020, 978-2-343-19610-7. {hal-02613842}
- Gautier, L., & Bach, M. (2017). La terminologie du vin au prisme des corpus oraux de dégustation/présentation (français-allemand): entre émotions, culture et sensorialité. *Études de linguistique appliquée: revue de didactologie des langues-cultures* (pp. 485-509). Paris: Klincksieck. <halshs-01839256>
- Gautier, L., & Hohota, V. (2014). Construire et exploiter un corpus oral de situations de dégustation: l'exemple d'Oenolex Bourgogne. *Studia Universitatis Babeş Bolyai – Studia Philologia, Universitatea Babeş-Bolyai*, LIX (4), 157-173.
- Gautier, L., & Leroyer, P. (2014). Construction, communication, représentation et réappropriation des discours vitivinicoles dans un 'nuancier' lexicographique en ligne. In C. Condei et al. (Eds.), *Situations professionnelles, discours et interactions en traduction spécialisée*. Berlin: Frank und Timme, GmbH Verlag für wissenschaftliche Literatur [in press]
- Gawel, R., Oberholster, A., & Francis, L. (2000). Terminology of communicating the mouth-feel characteristics of red wine. *Australian Journal Grape and Wine Research*, 6(3), 203-207.
- Geeraerts, D. (2006). *Words and Other Wonders. Papers on Lexical and Semantic Topics*. Berlin: Mouton de Gruyter.

- Geeraerts, D. (2009) *Theories of Lexical Semantics*. Oxford: OUP.
- Gelpí Arroyo, C. (2000). *La Lexicografía*. Barcelona: Grupo Santillana de Editores, S.A.
- Genouvrier, E., & Peytard, J. (1973). *Linguística e ensino do Português*. Coimbra, Livraria Almedina.
- Giboreau A, Dacremont, C., Egoroff, C., Guerrand, S., Urdapilleta, I., Candel, D., & Dubois, D. (2007). Defining sensory descriptors: Towards writing guidelines based on terminology. *Food Quality and Preference*, 18, 265-274.
- Giboreau, A. (2009). *De l'analyse sensorielle au jugement perceptif: l'exemple du toucher*. Habilitation à diriger des recherches, Université Claude Bernard Lyon 1.
- Gil, I. (2003). Algumas considerações sobre línguas de especialidade e seus processos lexicogénicos. *Máthesis*, 12, 113-130, Lisboa.
- Gilbert, A.N., & Wysocki, C.J. (1987). The Smell Survey results. *National Geographic*, 172, 514-525.
- Gimeno, J. R. (1986). *La educación de los Sentidos*. Santillana, Madrid.
- Glatre, E. (2014). *Lexique de la vigne et du vin – Wine lexicon*. Hors Collection Dunod.
- Glossário de A-Z. (n.d.). Disponível em <http://cct.portodigital.pt/gen.pl?p=vinhos&op=vinhos:glossario&sid=cct.sections/19>
- Glossário de A-Z. (n.d.). Disponível em <http://clubedevinhos.com/glossario>
- Glossário de A-Z. (n.d.). Disponível em <http://enologoonline.continente.pt/conhecer/dicionario>
- Glossário de A-Z. (n.d.). Disponível em <http://enologoonline.continente.pt/conhecer/glossario>
- Glossário de A-Z. (n.d.). Disponível em [http://pt.wikipedia.org/wiki/Anexo:Gloss%C3%A1rio\\_do\\_vinho](http://pt.wikipedia.org/wiki/Anexo:Gloss%C3%A1rio_do_vinho)
- Glossário de A-Z. (n.d.). Disponível em [http://www.abcdosvinhos.com/index.php?option=com\\_glossary&Itemid=526](http://www.abcdosvinhos.com/index.php?option=com_glossary&Itemid=526)
- Glossário de A-Z. (n.d.). Disponível em <http://www.adegaalmeirim.pt/curiosidades-glossario.php>
- Glossário de A-Z. (n.d.). Disponível em <http://www.aesbuc.pt/twt/etgi/myfiles/meussites/enologia/glossario.html>
- Glossário de A-Z. (n.d.). Disponível em <http://www.cantanhede.com/index.php?lang=po&menu=glossario>
- Glossário de A-Z. (n.d.). Disponível em <http://www.carmim.eu/content/1/38/glossario>
- Glossário de A-Z. (n.d.). Disponível em <http://www.esporao.com/PT/Rodape/Pages/Glossario.aspx>
- Glossário de A-Z. (n.d.). Disponível em <http://www.infovini.com/pagina.php?codNode=18640#tab0>
- Glossário de A-Z. (n.d.). Disponível em <http://www.ivv.min-agricultura.pt/np4/155>

- Glossário de A-Z. (n.d.). Disponível em [http://www.justinomadeira.com/index.php?option=com\\_content&view=article&id=124&Itemid=133&lang=pt](http://www.justinomadeira.com/index.php?option=com_content&view=article&id=124&Itemid=133&lang=pt)
- Glossário de A-Z. (n.d.). Disponível em <http://www.revistadeguste.com/glossario/?tipo=vinho>
- Glossário de A-Z. (n.d.). Disponível em <http://www.sograpevinhos.eu/glossario>
- Glossário de A-Z. (n.d.). Disponível em <http://www.vinhos.online.pt/dtp.asp>
- Glossário de A-Z. (n.d.). Disponível em <http://www.vinhotinto.com/glossario/>
- Glossário de A-Z. (n.d.). Disponível em <http://www.vinhoverde.pt/pt/recursos/glossario/glossario.asp>
- Glossário de A-Z. (n.d.). Disponível em <http://www.wineboutique.pt/blog/glossario-enologico/>
- Gluck, M. (2003). Wine language. Useful idiom or idiot-speak. In J. Aitchison & D. Lewis (Eds.), *New Media Language* (pp. 107-115). London, United Kingdom: Routledge.
- Gobet, F. (2015). *Understanding Expertise: A Multi-Disciplinary Approach*. Palgrave Macmillan.
- Gokcekus, O., & Nottebaum, D. (2011). The buyer's dilemma — whose rating should a wine drinker pay attention to? *American Association of Wine Economists Working Paper* No. 91.
- Gottlieb H. & Mogensen, J. E. (2007). *Dictionary visions research and practice : selected papers from the 12th International Symposium on Lexicography Copenhagen*. J. Benjamins Publishing.
- Gouws, R. H. (1999). *The general theory of lexicography*. Unpublished Textbook for MPhil in Lexicography. Stellenbosch: University of Stellenbosch
- Gouws, R. H. (2004). Monolingual and bilingual learner's dictionaries. *Lexikos*, 14:264–274.
- Gouws, R. H. (2012). Who can Really be Called a Lexicographer?. *Lexikos*, 22(1). <https://doi.org/10.5788/22-1-1004>
- Gouws, R. H. (2021). A critical appraisal of corpus linguistics: With reference to African languages. *Lexikos*, 31, 78-95.
- Gouws, R.H. & Prinsloo, D.J. (2005). Left-expanded article structures in Bantu with special reference in isiZulu and Sepedi. *International Journal of Lexicography*, 18: 25–46.
- Gouws, R.H. & Prinsloo, D.J. 2005. *Principles and practice of South African lexicography*. Stellenbosch: AFRICAN SUN MeDIA. doi:10.18820/9781919980911.
- Grefenstette G. (1994). *Corpus-derived first-, second- and third-order word affinities*. In Proceedings of Euralex, pp 279–290, Amsterdam, Holland. 1994.
- Guilbert, J. (1971). *Les dictionnaires français de synonymes et de contraires: Étude lexicologique*. Les Presses de l'Université Laval.

- Guillén Díaz, C. (2003). Une exploration du concept “lexiculture” au sein de la didactique des langues-cultures. *Didáctica (Lengua y Literatura)*, vol. 15, 105-119. Madrid.
- Haensch, G., & Omeñaca, C. (2004). La crítica de diccionarios. In G. Haensch & C. Omeñaca, *Los diccionarios del español en el umbral del siglo XXI* (329-337). Salamanca, Universidad de Salamanca.
- Haensch, L., Wolf, S. Ettinger, & R. Werner (1982). *La Lexicografía: de la Lingüística teórica a la Lexicografía práctica*. Madrid: Gredos.
- Harter, S. P., & Cheng, Y.-R. (1996). Colinked descriptors: improving vocabulary selection for end-user searching. *Journal of American Society for Information Science* 47(4), 311-325.
- Hartmann, R. R. K., (2018). Towards an international directory of lexicography. In P. A. Fuertes-Olivera (Ed.) *The Routledge Handbook of Lexicography* (pp. 767-786). London/New York: Routledge.
- Hartmann, R. R. K., & James, G. (1998). *Dictionary of lexicography*. London and New York: Routledge.
- Hausmann, F. J. (1985). Lexikographie. In C. Schwarze, & D. Wunderlich (Ed.), *Handbuch der Lexikologie* (pp. 367-411). Königstein/Ts.: Athenäum.
- Hausmann, F. J. (1989). Le dictionnaire de collocations. In F. J. Hausmann, O. Reichmann, H. E. Wiegand, L. Zgusta. (Eds.), *Wörterbücher: ein internationales Handbuch zur Lexicographie. Dictionaries* (pp. 1010-1019). Berlin/New-York, De Gruyter.
- Hausmann, F.J. & Wiegand, H.E. (1989). Component parts and structures of general monolingual dictionaries: A survey. In F.J. Hausmann et al. (Eds.). *Wörterbücher. Dictionaries. Dictionnaires. International encyclopaedia of lexicography*. Berlin: De Gruyter. 328–360.
- Hendrickx, I., Lefever, E., Croijmans, I., Majid, A., & Van den Bosch, A. (2016). Very quaffable and great fun: Applying NLP to wine reviews. In *54th Annual Meeting of the Association for Computational Linguistics, ACL 2016 – Short Papers*.
- Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences*, 33(2-3), 61-83. <https://doi.org/10.1017/S0140525X0999152X>
- Herbst, T. (1996). On the way to the perfect learner’s dictionary: a first comparison of OALD5, LDOCE3, COBUILD2 and CIDE. *International Journal of Lexicography* 9.4, 321-357.
- Herdenstam, A. P. F., Hammarén, M., Ahlström, R., & Wiktorsson, P.-A. (2009). The Professional Language of Wine: Perception, Training and Dialogue. *Journal of Wine Research*, 20(1), 53-84. <https://doi.org/10.1080/09571260902978543>
- Herz, R.S., Eliassen, J., Beland, S., & Souza, T. (2004). Neuroimaging evidence for the emotional potency of odor-evoked memory. *Neuropsychologia* 42, 371-378.

- Honoré-Chedozeau, C. (2017). *Faire la différence entre deux crus: du tout cuit? Effet de l'expertise sur les représentations conceptuelles et perceptuelles des vins du Beaujolais*. Université Bourgogne Franche-Comté.
- Honoré-Chedozeau, C., Ballester, J., Chatelet, B., & Lempereur, V. (2015). Wine competition: From between-juries consistency to sensory perception of consumers. *BIO Web of Conferences*, 5, 03009-03015. doi: 10.1051/bioconf/20150503009
- Houaiss, A. (Ed.). (2004). *Dicionário Houaiss da língua portuguesa*. Objetiva.
- Hughson, A. L., & Boakes, R. A. (2002). The knowing nose: the role of knowledge in wine expertise. *Food Quality and Preference*, 13(7), 463-472. [https://doi.org/10.1016/S0950-3293\(02\)00051-4](https://doi.org/10.1016/S0950-3293(02)00051-4)
- Humphrey, K., & Underwood, G. (2011). See What I'm Saying? Expertise and Verbalisation in Perception and Imagery of Complex Scenes. *Cognitive Computation*, 3(1), 64-78. <https://doi.org/10.1007/s12559-010-9065-0>
- Hwang, A. (2010). Lexicografia: dos primórdios à nova Lexicografia. In A. D. Hwang & O. L. Nadin (Eds.) *Linguagens em Interação III: estudos do léxico* (pp. 33-44). Maringá: Clichetec.
- Ibáñez Rodríguez, M. & Sánchez Nieto, M. T. (coor.) (2006). *El lenguaje de la vid y el vino y su traducción*. Valladolid: Universidad, Valladolid.
- Ibarretxe-Antunano, I. (2000) *Polysemy and metaphor in perception verbs : a cross-linguistic study*. PhD Thesis. (PhD diss., University of Edinburgh, 1999).
- IFT (1975). *Minutes of Division Business Meeting*. Institute of Food Technologists – Sensory Evaluation Division, Chicago.
- Ikeda, K. (1909). New seasonings. *J. Tokyo Chem. Soc.*, 30, 820-836.
- Isani, S. (2017). Of synaesthetic wine metaphors, discourse communities, addressers and addressees: the specialised language of wine descriptors. *ASp*, 72, 93-112. DOI: [10.4000/asp.5713](https://doi.org/10.4000/asp.5713)
- Isquierdo, A. (1996). *O fato lingüístico como recorte da realidade sócio-cultural*. Tese (Doutorado em Letras) - Universidade Estadual Paulista "Júlio de Mesquita Filho", Araraquara - SP, 1996
- Jackson, R. S. (2002). *Wine Tasting: A Professional Handbook*. Elsevier: Academic Press.
- Jacoby, J., Troutman, T., Kuss, A., & Mazursky, D. (1986). Experience and Expertise in Complex Decision Making. *Advances in Consumer Research*, 13(1), 469-472.
- Johnson, H., & Robinson, J. (2007). *The World Atlas of Wine*. London: Mitchell Beazley.
- Josselson, H. H. (1966). Automatization of lexicography. *Cahiers de lexicologie*, 2(9), 77-87. Paris: Classiques Garnier.
- Kempcke, G. (1984). *Handwörterbuch der deutschen Gegenwartssprache in zwei Bänden* (Vol. 1). Akademie Verlag.

- Kempcke, G. (2000). Paradigmatische und syntagmatische Relationen in einem Wörterbuch, Deutsch als Fremdsprache'. In U. Kramer (Ed.), *Lexikologisch-lexikographische Aspekte der deutschen Gegenwartssprache* (pp. 195-204). Tübingen: Niemeyer.
- Kilgarriff, A., Baisa, V., Bušta, J. et al. (2014a). The Sketch Engine: ten years on. *Lexicography ASIALEX* 1, 7–36 (2014). <https://doi.org/10.1007/s40607-014-0009-9>
- Kilgarriff, A., Jakubíček, M., Pomikálek, J., Sardinha, T.B., & Whitelock, P. (2014b). PtTenTen: A corpus for Portuguese lexicography.
- Kocourek, R. (1991). Textes e Termes, *Meta*, v.36, n.1, 1991, p.71-76.
- Krieger, M. G. (2006). Lexicografia: o léxico no dicionário. In M. C. Seabra. (Ed.). *O léxico em estudo* (pp. 157-171). Belo Horizonte: UFMG.
- Krieger, M.G., Finatto, M.J.B. (2004). *Introdução à Terminologia: teoria e prática*. São Paulo: Contexto.
- Krieger, M. G., Maciel, A. M. B., Finatto, M. J. B. (2000). Terminografia das leis do meio ambiente: princípios teórico-metodológicos. *TradTerm*, São Paulo, v. 6, p. 143- 169.
- Kudashev, I. (2007). Terminography vs. Lexicography Opposition Revisited. In N. Niemelä & E. Lehtinen (Eds.), *Theory of Translation, LSP and Multilingualism* (pp. 157-166). Vasa: Vaasan yliopisto.
- Kuhn, T. Z. (2017). *A design proposal of an on-line corpus driven dictionary of Portuguese for university students*. Tese de doutorado. Universidade de Lisboa.
- Lafon, P. (1984). *Dépouillements et Statistiques en Lexicométrie*, Genève, Slatkine, Champion.
- Landau, S. I. (2001). *Dictionaries: the art and craft of Lexicography*. New York: Cambridge University Press.
- Langlois, J. (2010). *Les expertises dans le domaine du vin: cas du concept de vin de garde*. Tese de Doutorado, Université de Bourgogne.
- Lara, L. F. (1996). *Teorías del Diccionario monolingüe*. Mexico: El Colegio de México.
- Lavric, E. (2013). Wie Europas Winzer/innen Fremdsprachen lernen: Präsentation des EU-Projekts VinoLingua. In V. Atayan, L. Sergo & U. Wienen (Eds.), *Akten der Sektion "Fachsprache(n) in der Romania – Entwicklung, Verwendung, Übersetzung" auf dem Deutschen Romanistentag* (Berlin, 25-28 September 2011) (pp. 309-331). Berlin: Frank & Timme.
- Lawless, H. T. (1984). Flavor description of white wine by 'expert' and nonexpert wine consumers. *Journal of Food Science*, 49(1), 120-123.
- Lawless, H. T., & Heymann, H. (1998). *Sensory Evaluation of Food: Principles and Practices*. Chapman & Hall, New York, N.Y.
- Lawless, L., & Civille, G. V. (2013). Developing lexicons: A review. *Journal of Sensory Studies*, 28(4), 270-281. <https://doi.org/10.1111/joss.12050>

- Le Gall, E. (2009). La lexiculture dans le Dictionnaire québécois d'aujourd'hui (1992) de Jean-Claude Boulanger: étude des lettres A et B. *Études de linguistique appliquée*, 2(2), 177-189. <https://doi.org/10.3917/ela.154.0177>
- Lecocq, S., & Visser, M. (2006). What determines wine prices: objective *versus* sensory characteristics. *Journal of Wine Economics* 1, 42-56.
- Lehmann, A., & Martin-Berthe, F. (1998). *Introduction à la lexicologie. Sémantique et morphologie*. Paris, Dunod, Coll. Lettres.
- Lehrer, A. (1975). Talking about wine. *Language* 51, 901-923.
- Lehrer, A. (2009). *Wine & Conversation*. New York: Oxford University Press.
- Lehrer, K., & Lehrer, A. (2016). The language of taste, *Inquiry*, 59:6, 752-765. DOI: 10.1080/0020174X.2016.1208925
- Lenoir, J. (2011). *Le Nez Du Vin*. Editions Jean Lenoir, France.
- León-Araúz, P., & Martín, A. S. (2018). The EcoLexicon semantic sketch grammar: from knowledge patterns to word sketches
- Lerat, P (1988). Les langues spécialisées. *Que sais-je?* Presses Universitaires de France.
- Lerat, P. (1995). *Les langues spécialisées*. Paris: Presse Universitaire de France.
- Lerat, P. (2006). Dénominations spécialisées, connaissances professionnelles et connaissances linguistiques en terminologie. Le cas du vin. In I. Rodriguez & S. Nieto (Eds.), *El lenguaje de la vid y el vino y su traducción* (pp. 201-224). Soria, 3-6 mars 2004, Universidad de Valladolid.
- Leroyer, P., & Gouws, R. (2009). En termes de vin : lexicographisation du guide oenotouristique en ligne. *Revue française de linguistique appliquée* 2009/2 (Vol. XIV), 99-116.
- Lesschaeve, I. (2006). The use of sensory descriptive analysis to gain better understanding of consumer wine language. In *3rd International Wine Business & Marketing Research Conference*, 6-8 July, Montpellier, France, ENSAM.
- Letontunier, E., & Munier, B. (2016). Introduction. La sensorialité, une communication paradoxale. *Hermès, La Revue* 2016/1 (n° 74), 17-24.
- Lew, R. (2011). Online Dictionaries of English. In P. A. Fuertes-Olivera & H. Bergenholtz (Eds.), *E-Lexicography: The Internet, Digital Initiatives and Lexicography* (pp. 230-250). London/New York: Continuum.
- Lew, R. (2012). How Can We Make Electronic Dictionaries More Effective? In S. Granger & M. Paquot (Eds.), *Electronic Lexicography* (pp. 343-361). Oxford: Oxford University Press.

- Lew, R., de Schryver, G-M. (2014). Dictionary Users in the Digital Revolution, *International Journal of Lexicography*, Volume 27, Issue 4, December 2014, Pages 341–359, <https://doi.org/10.1093/ijl/ecu011>
- Léxico de A-Z. (n.d.). Disponível em <http://www.ivdp.pt/pagina.asp?codPag=14&codSeccao=2&idioma=0>
- Ligas, P. (2012). Registres, stéréotypes et charge culturelle des mots dans le discours sportif non normatif. *Études de linguistique appliquée*, 1(1), 13-27. <https://doi.org/10.3917/ela.165.0013>
- Lindemann, B., Ogiwara, Y., & Ninomiya, Y. (2002). The Discovery of Umami. *Chem. Senses* (2002) 27 (9): 843-844 doi:10.1093/chemse/27.9.843.
- Lino, M. T. R. F. (1979). Importância de uma lexicologia contrastiva. *Revista da FCSH*, 1, 137-152.
- Lino, M. T. R. F. (2003). *La lexicografía plurilingüe en lenguas latinas: actas del Tercer Seminario de la Escuela Interlatina de Altos Estudios en Lingüística Aplicada*. San Millán de la Cogolla, 22-25 de octubre de 2003. En (Eds.), Título completo de los actas del seminario (págs. 95-104). ISBN 84-935340-4-8.
- Lino, M. T. R. F. (2006). *Contextes et néologie terminologique dans le domaine médical*. In: Actes des Journées Scientifiques «Mots, Termes et Contextes», Réseau Lexicologie, Terminologie et Traduction, Bruxelles.
- Lino, M. T. R. F. (2013). *Material de apoio apresentado em diapositivos* [PowerPoint de apoio à disciplina Seminário de especialização de Lexicologia, Lexicografia e Terminologia, lecionada na NOVA FCSH].
- Lino, M. T. R. F. (2018). Portuguese lexicography in the era of Internet. In P. A. Fuertes-Olivera (Ed.) *The Routledge Handbook of Lexicography* (pp. 608-618). London/New York: Routledge.
- Lino, M. T. R. F., Chicuna, A., Gröz, A., Medina, D. (2010). Neologia, Terminologia e Lexicultura. A língua portuguesa em situação de contacto de línguas. *Revista de Filologia e Linguística Portuguesa*, 2(12), 187-201. e-ISSN 2176-9419; ISSN 1517-4530.
- López Zurita, P. (2016). La terminología del vino y la viña jerezana en inglés. Un caso concreto. *Estudios Filológicos*, (58), 141-160. Disponível em <https://www.redalyc.org/articulo.oa?id=173449342007>
- López-Arroyo, B. L., & Roberts, R. P. (2014). English and Spanish Descriptors in Wine Tasting Terminology. *Terminology* 20(1), 25-49. doi: 10.1075/term.20.1.02lop
- López-Arroyo, B., & Roberts, R. (2020). What wine descriptors really mean A comparison between dictionary definitions and real use. *Journal of Wine Research*, 31(4), 301-321. DOI: 10.1080/09571264.2020.1854701

- Lukacs, P. (2013). *Inventing Wine: A New History of One of the World's Most Ancient Pleasures*. W. W. Norton & Company.
- Machetti, S. (2010). La lingua italiana del vino nella comunicazione pubblica e sociale italiana e globale. Analisi e prospettive di ricerca. *Studi Italiani di Linguistica Teorica e Applicata*. XXXIX, 1: 163-180.
- Malkiel, Y. (1959/1962). *A typological classification of dictionaries on the basis of distinctive features*. Washington, D.C.: Distributed by ERIC Clearinghouse.
- Marello, C., & Onesti, C. (2016). Colourful Microstructures: How Italian Dictionaries See Colour Terms. In *Colour and Colour Naming: Crosslinguistic Approaches* (pp. 91-104). Centro de Linguística da Universidade de Lisboa - Universidade de Aveiro.
- Marques, I. (2016). *Perceção sensorial: A importância dos cinco sentidos na marca. Uma análise no setor vitivinícola*. Tese de Doutoramento, Instituto Superior de Gestão, Lisboa.
- Martinez de Sousa, J. (1995). *Diccionario de Lexicografía Práctica*. Barcelona: Bibliograf, S.A.
- Martinez, W. (2012). *SIA, a corpus building and neologism tracking system*. Unpublished tool.
- Martinez, W., & Barbosa, S. (2018). Building a Portuguese oenological dictionary: from corpus to terminology via co-occurrence networks. In J. Čibej, V. Gorjanc, I. Kosem, & S. Krek (Eds.), *Proceedings of the XVIII EURALEX International Congress: Lexicography in Global Contexts* (1<sup>a</sup> ed., pp. 351-361). Ljubljana University Press, Faculty of Arts. <https://e-knjige.ff.uni-lj.si/znanstvena-zalozba/catalog/book/118>
- Martins, S. de C., & Zavaglia, C. (2016). As cores da fauna e da flora: um dicionário especial composto por cromônimos. *Estudos Linguísticos* (São Paulo. 1978), 42(1), 245–256. Recuperado de <https://revistas.gel.org.br/estudos-linguisticos/article/view/1103>
- Marzá, N. E. (2009). *The Specialised Lexicographical Approach: A Step further in Dictionary Making*. (Linguistic Insights 102. Studies in Language and Communication.) Bern: Peter Lang. ISBN: 978-3-0343-0043-8.
- McArthur, T. (1998). What is 'reference science'? *Lexicon*, 28, 135-140. Tokyo: Iwasaki Linguistic Circle. [https://globalex.link/wp-content/uploads/2019/08/Lexicon-28\\_007.pdf](https://globalex.link/wp-content/uploads/2019/08/Lexicon-28_007.pdf)
- McEnery, T., & Wilson, A. (2004). A corpus-based approach to discourse markers. *International Journal of Corpus Linguistics*, 9(1), 37-52. <https://doi.org/10.1075/ijcl.9.1.03mce>
- McGovern, P. E., Fleming, S. J., & Katz, S. H. (2003). *The Origins and Ancient History of Wine: Food and Nutrition in History and Antropology*. London/New York: Routledge.
- Mendes, A.; Génereux, M.; Hendrickx, I.; Pereira, L.; Bacelar do Nascimento, M. F.; Antunes, S. (2012) CQPWeb: uma nova plataforma de pesquisa para o CRPC. *Textos Seleccionados do XXVII Encontro Nacional da APL* p. 466–477.

- Méric, O. (2018). ¿Discurso de especialidad, discurso especializado o discurso situado? Ejemplos del cacao y del vino. *Revista Digital Internacional de Lexicología*, 2018(1).
- Messias, M. F. (2003). Para uma Teoria da Definição Lexical: Paradigma Informacional, Paradigma Definicional e Paradigma Pragmático. In *Atas do VII Encontro da Associação Portuguesa de Linguística* (pp. 275-288). Associação Portuguesa de Linguística.
- Meyer, C.M., & Abel, A. (2018). User Participation in the Internet Era. In P. A. Fuertes-Olivera (Ed.) *The Routledge Handbook of Lexicography* (pp. 735-753). London/New York: Routledge.
- Moreira, A. (2010). *Terminologia e Tradução: Criação De Uma Base De Dados Terminológica Do Turismo Baseada Num Corpus Paralelo Português/Inglês*. Dissertação de Doutoramento, Universidade de Vigo, Faculdade de Filoxía e Tradución, disponível em: [http://webs.uvigo.es/sli/arquivos/Tese\\_Adonay\\_Moreira.pdf](http://webs.uvigo.es/sli/arquivos/Tese_Adonay_Moreira.pdf)
- Morrot, G., Brochet, F., & Dubourdieu, D. (2001). The color of odors. *Brain and Language*, 79(2), 309-320. <https://doi.org/10.1006/brln.2001.2493>
- Morville, P. (2004). User Experience Design. Recuperado de [https://semanticstudios.com/user\\_experience\\_design/](https://semanticstudios.com/user_experience_design/)
- Morville, P. (2004). *User Experience Honeycomb*. [http://semanticstudios.com/user\\_experience\\_design/](http://semanticstudios.com/user_experience_design/)
- Müller-Spitzer, C. (2014). *Using online dictionaries*. De Gruyter.
- Myers, C. S. (1904). The taste-names of primitive peoples. *British Journal of Psychology, 1904-1920*, 1(2), 117-126. <https://doi.org/10.1111/j.2044-8295.1904.tb00153.x>
- Naumis Peña, C. (2000). Análisis de la confluencia entre término y descriptor en la elaboración de tesauros. *Investigación Bibliotecológica: archivonomía, bibliotecología e información*, 14(29). doi:<http://dx.doi.org/10.22201/iibi.0187358xp.2000.29.3947>
- Nebrija, A. (1492). *Dictionarium Aehi Antonii Nebrissensis (...) Hispan seu vulgan nomenclatura*. Antuerpiae [Bélgica]: Iohannes Steelsius, 1545.
- Nerlove, M. (1995). Hedonic price functions and the measurement of preferences: the case of Swedish wine consumers. *Eur. Econ. Rev.*, 39, 1697-1716
- Nesi, H. (1999). A User's Guide to Electronic Dictionaries for Language Learners. *International Journal of Lexicography* 12(1), 55-66.
- Nesi, H. (2015). The Demands of Users and the Publishing World: Printed or Online, Free or Paid for? In P. Durkin (Ed.), *The Oxford Handbook of Lexicography* (pp. 579-589). Oxford: Oxford University Press.
- Netter, F. H. (2019). *Atlas de anatomia humana*. 7ª ed. RIO DE JANEIRO: Elsevier, 602 p.

- Nielsen, S. (2013). The impact of digital dictionaries on the lexicographic process and product. In S. Granger & M. Paquot (Eds.), *Electronic lexicography* (pp. 357-373). Oxford: Oxford University Press.
- Noble, A. C. (1990). *The Wine Aroma Wheel*. Davis. California.
- Noble, A. C., Arnold, R. A., Masuda, B. M., Pecore, S. D., Schmidt, J. O., & Stern, P. M. (1984). Progress towards a standardized system of wine aroma terminology. *American Journal of Enology and Viticulture*, 35(2), 107-109.
- Nogueira, C. (2011). *Análise Sensorial de produtos têxteis*. Tese de Doutoramento, Universidade do Minho. Disponível em <https://repositorium.sdum.uminho.pt/bitstream/1822/19619/1/NOGUEIRA.Clarinda%20phD%20ThesisVF.pdf>
- Norman, G. R. (2002). *Expertise in Dictionary Use: An Experimental Investigation*. Springer Netherlands.
- Noronha, J. F. (2006). *Análise Sensorial* [PowerPoint de apoio à disciplina Análise Sensorial, lecionada na Escola Superior Agrária de Coimbra].
- Norrie, P. (2019). *The History of Wine as a Medicine: From its Beginnings in China to the Present Day*. Newcastle-upon-Tyne: Cambridge Scholars Publishing.
- Oczkowski, E., & Doucouliagos, H. (2015). Wine prices and quality ratings: A meta-regression analysis. *American Journal of Agricultural Economics*, 97(1), 103-121. <https://doi.org/10.1093/ajae/aa057>
- Pagliarini, E., Laureati, M., & Gaeta, D. (2013). Sensory descriptors, hedonic perception and consumer's attitudes to Sangiovese red wine deriving from organically and conventionally grown grapes. *Frontiers in Psychology*, 4:896. doi:10.3389/fpsyg.2013.00896
- Palencia, A. (1490). *Vocabulário universal en latin y en romance*. Sevilla: Paulus de Colónia.
- Palmer, J. J. & Kaminski, C. (1963). *Water: A comprehensive guide for brewers. Includes bibliographical references and index*. ISBN 978-0-937381-99-1.
- Paradis, C., & Hommerberg, C. (2016). We drink with our eyes first: The web of sensory perceptions, aesthetic experiences and mixed imagery in wine reviews. In R. W. Gibbs, Jr. (Ed.), *Mixing metaphor* (pp. 179–201). John Benjamins Publishing Company. <https://doi.org/10.1075/milcc.6.09par>.
- Parker, R. (1998). *Bordeaux: A Comprehensive Guide*. New York: Simon & Schuster.
- Parker, R. (2017). *Glossary Terms*. Disponível em <https://www.robertparker.com/resources/glossary-terms>
- Parr, W. V, Mouret, M., Blackmore, S., Pelquest-Hunt, T., & Urdapilleta, I. (2011). Representation of complexity in wine: Influence of expertise. *Food Quality and Preference*, 22(7), 647-660. <https://doi.org/10.1016/j.foodqual.2011.04.005>

- Pearson, J. (1998). *Terms in Context*. John Benjamins Publishing Co., Amsterdam.
- Pereda, D. B. (2002). *Diseño Sensorial, Las Nuevas pautas de innovacion, especializacion y personalizacion del producto : Los Sentidos Humanos y el Producto*. Dissertação de Mestrado, Univertat Politècnica de Catalunya.
- Persaud, K. (2017). Towards Bionic Noses. *Sensor Review*, 37(2), 165-171. <https://doi.org/10.1108/SR-10-2016-0238>
- Persaud, K. (2017). Towards Bionic Noses. *Sensor Review*, 37(2), 165-171. <https://doi.org/10.1108/SR-10-2016-0238>
- Peryam, D. R. (1958). Sensory difference tests. *Food Technology*, 12(5), 31-236.
- Petit, M. (2010). Le discours spécialisé et le spécialisé du discours: repères pour l'analyse du discours en anglais de spécialité. *E-rea, revue électronique d'études sur le monde anglophone* 8(1), 1-15.
- Peynaud E. & Bloin, (2013). *Knowing and making wine*. Wiley-Blackwell.
- Peynaud E. (1993). *Conhecer e Trabalhar o Vinho*. Litexa Editora.
- Peynaud, E. (1980). *Le goût du vin*, Edition Bordas/Dunod, Paris.
- Phal, A. (1971). *Vocabulaire Général d'Orientation Scientifique (V.G.O.S.). Part du lexique commun dans l'expression scientifique*. Paris, CREDIF.
- Pintão, M. & Cabral, C. (2014). *Dicionário Ilustrado do vinho do Porto*. Porto: Porto Editora.
- Ponte, S. (2011). *Os cinco Sentidos do Vestuário*. Tese de Doutoramento, Faculdade de Arquitetura, Universidade Técnica de Lisboa.
- Prensky, M. (2001). Digital Natives, Digital Immigrants Part 1, *On the Horizon*, Vol. 9 No. 5, pp. 1-6. <https://doi.org/10.1108/10748120110424816>
- Priilaid, D., Human, G., Pitcher, K., Smith, T., & Varkel, C. (2017). Are consumers' quality perceptions influenced by brand familiarity, brand exposure and brand knowledge? results from a wine tasting experiment. *South African Journal of Business Management*, 48(2):45-54, doi:10.4102/sajbm.v48i2.27.
- Prinsloo, D. J. (2005). Electronic Dictionaries Viewed from South Africa. *Hermes* 34, 11-35.
- Proust, M. (1918). *À la recherche du temps perdu*.
- Pruvost, J. (2009). Quelques perspectives lexicographiques à mesurer à l'aune lexicoculturelle. *Études de linguistique appliquée*, 2(2), 137-153. <https://doi.org/10.3917/ela.154.0137>
- Puckette, M., & Hammack, J. (2017). *Doidos por Vinho. O guia de vinhos essencial*. Saída de Emergência.
- Pylyshyn, Z. W. (1973). What the mind's eye tells the mind's brain: A critique of mental imagery. *Psychological Bulletin*, 80(1), 1-24. <https://doi.org/10.1037/h0034650>

- Quandt, R. E. (2007). On wine bullshit: some new software? *Journal of Wine Economics*, 2(2), 129-135.
- Quemada, B. (1987). Notes sur lexicographie et dictionnaire. *Cahiers de Lexicologie*, 2(51), 229-242. Paris: Didier.
- Quirk, R. (1954). *Survey of English usage*. Routledge.
- Reckinger, R. (2012). *Parler vin: entre normes et appropriations*. Coll. Tables des hommes. Rennes: Presses universitaires de Rennes
- Rey, A. (1970). *La terminologie: noms et notions*. Paris: Presses Universitaires de France.
- Rey, A. (1986). *Dictionnaire culturel en langue française*. Éditions du Seuil.
- Rey-Debove, J. (1971). *Étude linguistique et sémiotique des dictionnaires français contemporains*. La Haye, Mouton.
- Robinson, J. (2016). *Especialista de vinhos em 24 horas*. Casa das Letras.
- Rondeau, G. (1984). *Introduction à la Terminologie*. Paris: Gaëtan Morin éditeur.
- Rosado, A. R. (2013). *Evolução de parâmetros físicos, químicos e controlo microbiológico em vinhos brancos e tintos da Adega Cooperativa de Palmela*. Dissertação de Mestrado, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa.
- Rundell, M. (2002). Good Old-Fashioned Lexicography: Human Judgment and the Limits of Automation. In M.-H. Corréard (Ed.), *Lexicography and Natural Language Processing. A Festschrift in Honour of BTS Atkins* (pp. 138-155). Euralex.
- Sacha González. C. (2006). La cata del vino y su terminología, en *El lenguaje de la vid y el vino y su traducción*. (coor. M. Ibáñez Rodríguez y M. T. Sánchez Nieto). Valladolid: Universidad de Valladolid, pp. 77-84.
- Salvador, J. A. (2010). *ABeCedário dos Vinhos*. Lisboa: Edições Afrontamento.
- San Roque, L., Kendrick, K., Norcliffe, E., Brown, P., Defina, R., Dingemanse, M., Dirksmeyer, T., Enfield, N., Floyd, S., Hammond, J., Rossi, G., Tufvesson, S., van Putten, S., & Majid, A. (2015). Vision verbs dominate in conversation across cultures, but the ranking of non-visual verbs varies. *Cognitive Linguistics*, 26(1), 31-60. <https://doi.org/10.1515/cog-2014-0089>
- Sánchez, J. (2011). *Glossário del vino*. Disponível em <https://glosarios.servidor-alicante.com/glosario-del-vino>
- Sapir, E. (1949). *Language*. Harvest Book. Paris, Payot.
- Saussure, F. de. (1911). *Course in general linguistics* (C. Bally & A. Sechehaye, Eds.; W. Baskin, Trans.). Philosophical Library.
- Saussure, F. de. (1916). *Les cours de linguistique générale* (C. Bally & A. Sechehaye, Eds.). Payot.

- Sauvageot, F., Urdapilleta, I., & Peyron, D. (2006). Within and between variations of texts elicited from nine wine experts. *Food Quality and Preference*, 17(6), 429-444. <https://doi.org/10.1016/j.foodqual.2005.05.007>
- Scardamalia, M., & Bereiter, C. (1991). Literate expertise. In K. A. Ericsson & J. Smith (Eds.), *Toward a general theory of expertise: Prospects and limits* (pp. 172-194). New York, NY, US: Cambridge University Press.
- Scerba, L. V. (1940). Towards a General Theory of Lexicography. *International Journal of Lexicography*, 8(4), 315-350.
- Schierholz, S. (2012). Lexicografia de especialidade e terminografia. En: Isquierdo, A. N.; Seabra, M. C. T. C. (orgs.) *As Ciências do Léxico. Lexicologia, Lexicografia e Terminologia*, VI. Campo Grande - MS: Editora UFMS.
- Scott, M., & Tribble, C. (2006). *Textual Patterns: keyword and corpus analysis in language education*. Amsterdam: Benjamins.
- Sezille, C., Fournel, A., Rouby, C., Rinck, F., & Bensafi, M. (2014). Hedonic appreciation and verbal description of pleasant and unpleasant odors in untrained, trainee cooks, flavorists and perfumers. *Frontiers in Cognitive Science*, 5, 12. <https://doi.org/10.3389/fpsyg.2014.00012>
- Sharp, A. (1998). *Winetaster's secrets*. Warwick Publishing.
- Silva e Silva, M. (2013). La lexiculturologie... en didactologie des langues-cultures. *Synergies Portugal* 1, 69-89.
- Silva, H.R. (2016) "*Minéralité des vins: parlons-en!*" : la conceptualisation d'un descripteur sensoriel mal défini. Alimentation et Nutrition. Université de Bourgogne, 2016. Français. (NNT : 2016DIJOS044). (tel-01661322)
- Silverstein, M. (2006). Old wine, new ethnographic lexicography. *Annual Review of Anthropology*, 35, 481-496.
- Sinclair, J. (1991). *Corpus. Concordance. Collocation*. Oxford: Oxford University Press.
- Sinclair, J. (1996). EAGLES Preliminary recommendations on Corpus Typology. *EAGLES Document EAG TCWG CTYP/P*. Pisa: Consiglio Nazionale delle Ricerche. Istituto di Linguistica Computazionale. Unpublished manuscript. Disponível em <http://www.ilc.cnr.it/EAGLES/corpus/corpus.html>
- Sinclair, J. (2005). Corpus and Text - Basic Principles. In M. Wynn (Ed.) *Developing Linguistic Corpora: a Guide to Good Practice* (pp. 1-16). Oxford: Oxbow Books. Disponível em <http://www.ahds.ac.uk/creating/guides/linguistic-corpora/chapter1.htm>.
- Sjostrom, T., Corsi, A.M., & Lockshin, L. (2014). Consumer perceptions of premium and luxury wine brands. *Wine Vitic. J.* 29, 68-70.

- Slodzian, M. (2000) L'émergence d'une terminologie textuelle et le retour du sens. In: Henri Béjoint et Philippe Thoiron, dir. *Le Sens en terminologie*. Lyon: Presses Universitaires de Lyon, 61-85.
- Small, D. M., & Prescott, J. (2005). Odor/taste integration and the perception of flavor. *Exp Brain Res*, 166, 345-357.
- Smith, V., Møgelvang-Hansen, P., & Hyldig, G. (2010). Spin versus fair speak in food labelling: a matter of taste? *Food Quality and Preference*, 21, 1016-1025.
- Solomon, G. E. (1990). Psychology of novice and expert wine talk. *The American Journal of Psychology*, 103(4), 495-517. <https://doi.org/10.2307/1423321>
- Souffi, S. (2009). Le Dictionnaire de l'académie française: entre bon usage et culture. *Études de linguistique appliquée*, 2(2), 155-176. <https://doi.org/10.3917/ela.154.0155>
- Spence, C. (2015). Just how much of what we taste derives from the sense of smell? *Flavour*, 4(1), 30. <https://doi.org/10.1186/s13411-015-0040-2>
- Sperber, D. (1975). *Rethinking Symbolism* (English translation by Alice L. Morton). Cambridge University Press.
- Stengel, K. & Marinescu, A-H, (2016). Wine Tasting Discourse: Traditional Knowledge, and Practice. *Journal of Social Sciences, Science Publications, Open Journal of Social Sciences, Vol.4 (N°5)*, pp.124-134.
- Stubbs, M. (2002). *Words and Phrases: Corpus Study of Lexical Semantics*. Oxford: Blackwell Publishing.
- Swales, J. M. (1990). *Genre Analysis: English in Academic and Research Setting*. Cambridge: Cambridge University Press.
- Swiegers, J.H., Chambers, P.J., & Pretorius, I.S. (2005). Olfaction and taste: Human perception, physiology and genetics. *Australian Journal of Grape and Wine Research*, 11: 109-113. doi:10.1111/j.1755-0238.2005.tb00282.x
- Tarp, S. (2008). *Lexicography in the Borderland Between Knowledge and Non-Knowledge: General Lexicographical Theory With Particular Focus on Learner's Lexicography*. Tübingen: Niemeyer.
- Tarp, S. (2010). Functions of Specialised Learners' Dictionaries. In: FUERTE-OLIVEIRA, P. A. (coord.) *Specialised Dictionaries for learners*. Walter de Gruyter GMBH. S. CO-KG, Berlin – New York.
- Tarp, S. (2012). Theoretical Challenges in the Transition From Lexicographical P-Works to E-Tools. In S. Granger e M. Paquot (Eds.), *Electronic Lexicography* (pp. 107-118). Oxford: Oxford University Press.
- Tarp, S. (2018). Lexicography as an independent science. In P. A. Fuertes-Olivera (Ed.) *The Routledge Handbook of Lexicography* (pp. 19-33). London/New York: Routledge.

- Tegmark, M. (2017). *Life 3.0: Ser-se Humano na Era da Inteligência Artificial*. Dom Quixote.
- Teixeira, C. (2014). *A indução e a formulação de experiências. Análise linguística de textos da área do vinho*. Tese de Doutoramento, Faculdade de Ciências Sociais e Humanas, Universidade Nova de Lisboa.
- Ten Hacken, P. (2009). The Tension between Definition and Reality in Terminology. In *Proceedings of the xiv Euralex International Congress* (pp. 915-927). Ljouwert: Fryske Akademy.
- Thomas, A. (2016). *Analyse sensorielle temporelle descriptive et hédonique*. Tese de Doutoramento, Université de Bourgogne.
- Thorndike, E. L. (1921). *Teacher's Wordbook*. New York: Columbia Teachers College.
- Tsai, C. (2014). La langue spécialisée du vin : étude comparative de comptes rendus de dégustation. *Revue française de linguistique appliquée* 2014/1 (Vol. XIX), 117-131.
- Tsitsa, M. (2004). Le français de l'art culinaire un projet dictionnaire. *Études de linguistique appliquée*, 3(3), 299-309. <https://doi.org/10.3917/ela.135.0299>
- Ullmann, S. (1959). *The Principles of Semantics* (2nd Ed). Glasgow: Jackson, Son & Co.
- Unwin, T. (1996). *Wine and the vine: an historical geography of viticulture and the wine trade*. London: Routledge.
- Valentin, D., Chollet, S., & Abdi, H. (2003). Les mots du vin: experts et novices diffèrent-ils quand ils décrivent des vins? *Corpus*, (2), pp.183-200.
- Valenza, G. (2010). *De Lingva Latina, de Marco Terêncio Varrão: Tradução dos Livros VIII, IX e X*. Dissertação de Mestrado, Universidade Federal do Paraná.
- Van Baardewijk-Rességuier, J. (1993). Les mots de la culture. In A. Hulk (dir.), *Du lexique à la morphologie : du côté de chez Zwaan* (pp. 29-31). Amsterdam-Atlanta, Rodopi.
- Vedel, A., Chamay, C. G., & Tourmeau, P. (1972). *Essai sur la dégustation des vins*. Societe d'Edition et d'Informations viti-vinicoles, Macon.
- Vedovelli, M. (2016). Parlare e scrivere di vino: l'enogrammatologia. *Cultura e comunicazione*, VI, n. 8: 9-25.
- Verdelho, T. (1995). *As origens da Gramaticografia e da Lexicografia Latino-Portuguesas*. Aveiro: Instituto Nacional de Investigação Científica.
- Verdelho, T., & Silvestre, J.P. (2007). *Dicionarística portuguesa. Inventariação e estudo do património lexicográfico*. Aveiro: Universidade do Aveiro.
- Viberg, Å. (1983). The verbs of perception: a typological study. *Linguistics* 21(1), 123-162.
- Vigner, G. (1989). Thèmes, champs lexicaux et activités discursives. *Recherches et Applications, n° spécial du Français dans le Monde* (pp. 134-145). Coll. Recherches et Applications, Rennes, Hachette.

- Vilela, M. (1994). *Estudos de Lexicologia do Português*. Coimbra: Livraria Almedina.
- Vion, R. (2009). De la subjectivité en discours aux discours sur la subjectivité. In D. Dubois, *Le sentir et le dire*. (pp. 357-392). Paris: L'Harmattan.
- Viterbo, R. (1789/1984). *Diccionario da lingua portugueza: abreviado para uso das escolas*. Imprensa Nacional - Casa da Moeda.
- Wagner, R.-L., & Quemada, B. (1969). Pour une analyse des français contemporains. *Le Français dans le monde*, 69, 61-63.
- Wang, S. (2016). *Lexicultura na Lexicografia Bilingue de Chinês-Português*. Tese de Doutorado, Faculdade de Ciências Sociais e Humanas, Universidade Nova de Lisboa. Disponível em <https://run.unl.pt/handle/10362/17164>
- Welker, H. A. (2004). *Dicionários. Uma Pequena Introdução à Lexicografia*. Brasília, Thesaurus.
- Welker, H. A. (2006). Breve histórico da metalexigrafia no Brasil e dos dicionários gerais brasileiros. *Matraga*, vol. 13, n. 19, p. 69-84. Disponível em <http://www.pgletras.uerj.br/matraga/matraga19/matraga19a04.pdf>
- Whorf, B. L. (1956). *Language, thought, and reality. Selected writings of Benjamin Lee Whorf*. (J. Carroll, Ed.). Cambridge, Massachusetts: The MIT Press.
- Whorton, J. C. (2010). *The arsenic century: how Victorian Britain was poisoned at home, work, and play*. Oxford: Oxford University Press.
- Wiegand, H. E. (1984a). On the Structure and Contents of a General Theory of Lexicography. In R.R.K. Hartmann (Ed.), *LEXeter '83* (pp. 13-30). Tübingen: Max Niemeyer.
- Wiegand, H.E. (1984b). Prinzipien und Methoden historischen Lexikographie. In W. Besch et al. (Eds.). *Sprachgeschichte: Ein Handbuch zur Geschichte der deutschen Sprache und ihrer Erforschung*. Berlin: Handbücher zur Sprachund Kommunikationswissenschaft 2.1. 557-620.
- Wiegand, H.E. (1998). *Wörterbuchforschung. Untersuchungen zur Wörterbuch-benutzung, zur Theorie, Geschichte, Kritik und Automatisierung der Lexikographie*.1. Teilband. Berlin/New York: de Gruyter
- Wiegand, H. E. (2013). Lexikographie und Angewandte Linguistik. *Zeitschrift für angewandte Linguistik* 58(1), 13-39.
- Wiegand, H. E., & Morán, M.T.F. (2009). *Estructuras Lexicográficas: Aspectos Centrales de una Teoría de la Forma del Diccionario*. Granada: Tragacanto.
- Williams, H. (2002). *Méthodes d'analyse des dictionnaires: application au Dictionnaire de l'Académie française*. Presses universitaires de France.

- Winter, B. (2016). Taste and smell words form an affectively loaded and emotionally flexible part of the English lexicon. *Language, Cognition and Neuroscience*, 31(8), 975-988. doi: 10.1080/23273798.2016.1193619
- Wipf, B. (2010). *Wine writing meets MIPVU: Linguistic metaphor identification of tasting notes*. Tese de doutoramento, Vrije Universiteit, Amsterdam.
- Wittgenstein, L. (1961 [1921]) *Tractatus-Logico Philosophicus*. Routledge, London
- Xavier, M. F. & Mateus, M.H. (1992). *Dicionário de Termos Linguísticos*. Volume II. 1ª edição. Lisboa: Edições Cosmos.
- Zannoni, M. (1997). Approaches to translation problems of sensory descriptors. *Journal of Sensory Studies*, 12(3), 239-253.
- Zhen, Y.-S. (Ed.). (2002). *Tea: Bioactivity and Therapeutic Potential* (1st ed.). CRC Press. <https://doi.org/10.1201/b12659>
-

## 9. ANEXO 1

### Questionário

Olá,

O presente questionário, para o qual solicitamos a sua colaboração, tem como objetivo recolher informação sobre a descrição dos vinhos. Os resultados deste inquérito serão utilizados para fins exclusivamente académicos.

O preenchimento do questionário demora cerca de 10 minutos. Pedimos-lhe que responda o mais sinceramente e espontaneamente possível: não existem respostas certas ou erradas, pois todos temos opiniões diferentes.

Todos os dados recolhidos são confidenciais. A participação é voluntária. Para que o estudo seja fidedigno, pedimos que responda apenas uma vez.

Agradecemos desde já a sua inestimável colaboração e a atenção dispensadas.

Silvia Barbosa  
PhD Student, Departamento de Linguística  
Faculdade de Ciências Sociais e Humanas - FCSH/NOVA  
E-mail: [silviabarbosa@fcsb.unl.pt](mailto:silviabarbosa@fcsb.unl.pt)

Se concorda em participar no estudo, por favor avance para a primeira questão.

*\* Indica uma pergunta obrigatória*

---

#### 1. Idade \*

*Marcar apenas uma oval.*

- 20 a 29
- 30 a 39
- 40 a 49
- 50 a 59
- 60 ou mais

#### 2. Sexo \*

*Marcar apenas uma oval.*

- Feminino
- Masculino

#### 3. Nacionalidade \*

---

4. situação profissional \*

*Marcar apenas uma oval.*

- Trabalho a tempo inteiro
- Trabalho em part-time
- Estudante
- Desempregado/a
- Reformado/a

5. Habilitações literárias \*

*Marque todas que se aplicam.*

- Primário incompleto
- Primário completo
- Secundário completo
- Licenciatura completa
- Mestrado completo
- Doutoramento completo

6. Consome vinho? \*

*Marcar apenas uma oval.*

- Sim
- Não

7. Com que frequência consome vinho? \*

*Marcar apenas uma oval.*

- várias vezes por dia
- uma vez por dia
- algumas vezes por semana
- algumas vezes por mês
- menos de uma vez por mês
- nunca

Apresente um exemplo de um produto (objeto, alimento, ...) que apresenta as seguintes características.

8. .... tem cor amarelo citrino com ligeiros tons esverdeados \*

---

9. ... tem cor limpa \*

---

10. ... tem aroma jovem \*

---

11. ... tem aroma cativante \*

---

12. .... tem frescura mineral \*

---

13. ... tem elegância \*

---

14. Defina, por palavras suas, o que entende por vinho carregado? \*

---

---

---

---

---

15. Defina, por palavras suas, o que entende por vinho aberto? \*

---

---

---

---

---

16. Defina, por palavras suas, o que entende por vinho limpo? \*

---

---

---

---

---

17. Defina, por palavras suas, o que entende por vinho retinto? \*

---

---

---

---

---

18. Defina, por palavras suas, o que entende por vinho leve? \*

---

---

---

---

---

19. Defina, por palavras suas, o que entende por vinho equilibrado? \*

---

---

---

---

---

20. Defina, por palavras suas, o que entende por vinho denso? \*

---

---

---

---

---

21. Defina, por palavras suas, o que entende por vinho mineral? \*

---

---

---

---

---

22. Defina, por palavras suas, o que entende por vinho austero? \*

---

---

23. Defina, por palavras suas, o que entende por vinho encorpado? \*

---

---

---

---

---

24. Defina, por palavras suas, o que entende por vinho complexo? \*

---

---

---

---

---

25. Defina, por palavras suas, o que entende por vinho frutado? \*

---

---

---

---

---

26. Defina, por palavras suas, o que entende por vinho vivo? \*

---

---

---

---

---

27. Defina, por palavras suas, o que entende por vinho elegante? \*

---

---

---

---

---

Obrigada pela sua participação.

---

Este conteúdo não foi criado nem aprovado pelo Google.

Google Formulários

## ÍNDICE DE FIGURAS

### Índice de FIGURAS

Figura 1. Subáreas da Lexicografia (Hartmann & James, 1998:86).....	19
Figura 2. Exemplo de um guia de utilização retirado de Dicionário da Língua Portuguesa (2018). ....	43
Figura 3. Microestrutura básica essencial/mínima do artigo lexicográfico (baseado em Barbosa, 1995)..	50
Figura 4. Estrutura do texto definitório da unidade lexical especializada do artigo lexicográfico .....	60
Figura 5. Sistemas sensoriais no Córtex (retirado de Nogueira, 2011:13).....	84
Figura 6. Etapas psicológicas da percepção (retirado de Nogueira, 2011).....	86
Figura 7. Percurso do estímulo apresentado à verbalização do mesmo (adaptado de Barbosa, 2016).....	87
Figura 8. Quadro dos sentidos utilizados durante uma prova de vinho (retirado de Peynaud, 1993:17)....	93
Figura 9. Imagem frontal e transversal relativa ao olho do ser humano (retirado de Netter, 2019) .....	94
Figura 10. Roda das Cores em: <a href="http://tiny.cc/lur8gz">http://tiny.cc/lur8gz</a> (disponível e consultado em setembro de 2018) .....	96
Figura 11. Imagem lateral da cavidade nasal do ser humano (retirado de Netter, 2019).....	97
Figura 12. Roda dos Aromas adaptada para o português (retirado de INFOVINI, 2018). ....	99
Figura 13. Imagem frontal e de corte transversal da língua do ser humano (retirado de Netter, 2019)....	100
Figura 14. Roda dos Sabores em versão inglesa (retirada de Gawel, Oberholster, & Francis, 2000) .....	101
Figura 15. Avaliação do vinho (esquema retirado de Puckette & Hammack (2017)) .....	102
Figura 16. Exemplo de uma ficha descritiva de resposta aberta (retirado de Cardoso, 2007).....	103
Figura 17. Exemplo de uma ficha descritiva de resposta fechada (retirado de INFOVINI, 2018).....	103
Figura 18. Exemplo de ficha descritiva mista para 4 amostras de vinho (retirada de Wine Folly, 2018)	104
Figura 19. Exemplo de uma ficha de avaliação por pontuação (retirado de Clube de Vinhos Portugueses) .....	104
Figura 20. Triângulo de Vedel (retirado de Peynaud & Bloin (2013)).....	105
Figura 21. Exemplo de codificação de um texto recolhido para a BDEno.....	111
Figura 22. Exemplos de diferentes definições no CEnoLEX .....	126
Figura 23. Exemplos de fichas técnicas e de notas de prova.....	131
Figura 24. Excerto de concordâncias de “boca” .....	169
Figura 25. Word Sketch de “boca” .....	170
Figura 26. Diagrama Venn do número de UL de cada categoria gramatical no CEnoLEX e no CEnoTEXT .....	205
Figura 27. Exemplo da divisão em secções relativas às modalidades sensoriais.....	207
Figura 28. Imagem abstrata da rede de coocorrências retirada de Martinez e Barbosa (2018) .....	209
Figura 29. Rede de coocorrências das formas mais frequentes retirada de Martinez e Barbosa (2018)...	213
Figura 30. Rede de coocorrências de “boca” retirada de Martinez e Barbosa (2018) .....	216

Figura 31. Projeção fatorial das formas mais usadas em cada modalidade sensorial .....	220
Figura 32. Projeção fatorial – modalidade visual .....	222
Figura 33. Projeção fatorial – modalidade olfativa.....	223
Figura 34. Projeção fatorial – modalidade gustativa .....	223
Figura 35. Projeção fatorial – verbos por modalidade sensorial.....	226
Figura 36. Esquema da segmentação da nota de prova .....	229
Figura 37. Exemplo do processo de inserção informação relativa às modalidades sensoriais no CEnoTEXT .....	230
Figura 38. Possível categorização da modalidade sensorial visual.....	231
Figura 39. Estrutura básica dos constituintes obrigatórios e opcionais do artigo lexicográfico .....	249
Figura 40. Proposta de artigo para “suor de cavalo” .....	249
Figura 41. Proposta de artigo para “opaco” .....	250
Figura 42. Proposta de artigo para “crescer” .....	250
Figura 43. Os sete fatores <i>útil/usável/achável/credível/desejável/acessível/valioso</i> que afetam o utilizador .....	252
Figura 44. Descrição do processo do UX Design (retirado de <a href="https://tinyurl.com/332mb3ad">https://tinyurl.com/332mb3ad</a> ).....	253
Figura 45. Exemplo de pesquisa de aplicações digitais relacionadas com vinho .....	259
Figura 46. Cenários e utilizadores hipotéticos.....	260
Figura 47. Crazy 8 aplicado ao DEASE .....	262
Figura 48. Evolução dos <i>wireframes</i> para construção do protótipo.....	263
Figura 49. Imagem do ecrã inicial da APP DEASE .....	264
Figura 50. Ecrã de pesquisa alfabética .....	265
Figura 51. Ecrã da pesquisa por modalidade sensorial visual .....	265
Figura 52. Ecrã do descritor “acastanhado” .....	266
Figura 53. Ecrã a que se acede a partir do botão (+) .....	267
Figura 54. Ecrã do descritor “alperce” .....	267
Figura 55. Artigo lexicográfico em versão papel .....	268
Figura 56. Artigo lexicográfico em versão digital.....	268

## ÍNDICE DE TABELAS

Tabela 1. Períodos da Lexicografia portuguesa.....	25
Tabela 2. Texto definitivo de “frutado” consoante os níveis de especialidade .....	63
Tabela 3. Períodos do século XX e respetiva descrição (adaptado de Freitas, 2008).....	80
Tabela 4. Fontes de textos recolhidos para a BDEno .....	112
Tabela 5. Modelo de ficha de avaliação das obras selecionadas .....	118
Tabela 6. Exemplo de “abafado” nas diferentes obras .....	119
Tabela 7. Exemplo da entrada para “abafado” organizado consoante a definição apresentada nas fontes .....	120
Tabela 8. Distribuição dos descritores nas obras consultadas .....	122
Tabela 9. distribuição dos descritores consoante o número de definições distintas encontradas nas obras .....	123
Tabela 10. Exemplos de definições no CEnoLEX .....	124
Tabela 11. Dimensão da definição quanto ao número de frases e palavras.....	125
Tabela 12. Exemplo de informação recolhida na ficha CEnoTEXT .....	132
Tabela 13. Exemplos de notas de prova encontrados no CEnoTEXT.....	135
Tabela 14. As 20 unidades lexicais mais frequentes do CEnoTEXT .....	140
Tabela 15. A distribuição das unidades lexicais consoante o número de ocorrências no CEnoTEXT ..	141
Tabela 16. Valores numerais após revisão manual.....	143
Tabela 17. Valores percentuais da distribuição dos lemas .....	144
Tabela 18. Excerto da listagem comparativa entre CEnoTEXT e CORLEX .....	146
Tabela 19. Listagem dos 20 verbos (lemas) mais frequentes de CEnoTEXT .....	150
Tabela 20. Listagem das 20 formas verbais mais frequentes.....	150
Tabela 21. Lista de verbos com as freq. relativas e o valor diferencial de frequência entre CEnoTEXT e CORLEX .....	153
Tabela 22. Entradas de “aparecer” e “evoluir” nos dicionários Priberam e Infopédia .....	156
Tabela 23. Detecção da existência de polissemia sensorial na língua geral ou no CEnoTEXT .....	159
Tabela 24. Listagem dos 20 nomes mais frequentes de CEnoTEXT .....	162
Tabela 25. Lista de nomes com as freq. relativas e o valor diferencial de frequência entre CEnoTEXT e CORLEX .....	167
Tabela 26. Lista de contraste dos 20 candidatos com frequência absoluta e os valores diferenciais.....	169
Tabela 27. Listagem dos 20 adjetivos mais frequentes de CEnoTEXT.....	178
Tabela 28. Lista de contraste dos 20 candidatos com frequência absoluta e os valores diferenciais.....	182
Tabela 29. Listagem de empréstimos do CEnoTEXT .....	190
Tabela 30. Definição de “refrescar” no CEnoLEX.....	195

Tabela 31. Definição de “nota” no CEnoLEX.....	197
Tabela 32. Definição de “mineral” no CEnoLEX .....	199
Tabela 33. Definição de “cheio” no CEnoLEX.....	200
Tabela 34. Definição de “vivo” no CEnoLEX .....	202
Tabela 35. Definição de “opaco” no CEnoLEX.....	203
Tabela 36. As formas mais frequentes (frequência $\geq$ 5000 ocorrências).....	210
Tabela 37. Os 10 lemas mais frequentes em cada modalidade sensorial.....	214
Tabela 38: Especificidades positivas e negativas das coocorrências de “boca” .....	215
Tabela 39. Verbos sensoriais .....	224
Tabela 40. Os 10 verbos mais frequentes em cada modalidade sensorial .....	226
Tabela 41. Distribuição de verbos por modalidades.....	227
Tabela 42. Estrutura base da ficha lexicográfica da BDEnoLex .....	242
Tabela 43. Comparação da estrutura base da ficha lexicográfica da BDEnoLex e da microestrutura do artigo lexicográfico do DEASE.....	246
Tabela 44. Elementos obrigatórios do artigo lexicográfico do DEASE .....	247
Tabela 45. Elementos opcionais do artigo lexicográfico do DEASE .....	248
Tabela 46. Lista de obras apresentadas no Inquérito .....	257

---