

# Construção de uma rubrica de avaliação para defesa pública: da definição de critérios ao protótipo

Simone Tomé, Nicolas Hurst, Marta Anido, Mónica Lorenzo, José Almeida, Fátima Silva, Pilar Martinez, Sónia Rodrigues, Maria Ellison & Manuela Teixeira

---

## Resumo

O artigo põe em discussão questões relacionadas com a construção de uma rubrica de avaliação para defesa pública de relatórios de mestrado em ensino de línguas. Suscitada pela necessidade de melhoria de procedimentos e harmonização de práticas e instrumentos de avaliação em diferentes cursos de mestrado em ensino de línguas, a criação de uma rubrica para avaliação da defesa pública do relatório final é o objeto desta comunicação. O texto tem a estrutura seguinte: (i) análise de documentos de referência para a defesa pública no contexto académico específico e de rubricas similares em uso noutros contextos; (ii) descrição do contexto comunicativo académico inerente ao género oral formal 'defesa pública' e das suas partes constitutivas; (iii) apresentação da estrutura da rubrica; (iv) enunciação das fases de produção do protótipo em construção.

---

## Palavras-chave:

formação inicial de professores de línguas; rubrica de avaliação; relatório de estágio; defesa pública

## Defining and defending: A criteria-based assessment rubric for the public defence

**Abstract:** The aim of this paper is to discuss issues related to the construction of an assessment rubric for the viva (public defence) of the final report resulting from teaching practice internships. Both theoretical and practical considerations which underpin the methodology adopted in this process will be discussed. The institutional context is the master's degrees in teaching languages at the Faculty of Arts and Humanities, the University of Porto (FLUP). This process was initiated in response to a jointly-perceived need to enhance quality assurance in assessment procedures and to harmonize practices and instruments used in the different masters in teaching. Thus, the subject of this paper is the construction of a criteria-based rubric to assess performance levels for the evaluation of the viva of the teaching practice final report. The process underlying its construction entailed the following phases: (i) analysis of both regulatory documents for vivas at the University of Porto and of similar rubrics in use in other university contexts; (ii) description of the communicative situation inherent in this formal, oral discourse genre in an academic context and of its constituent parts; (iii) the structure of the assessment rubric; (iv) the writing of a prototype of the rubric.

**Keywords:** Initial language teacher education; criteria-based assessment; final report; viva

## Construction d'une rubrique d'évaluation pour la soutenance: de la définition des critères au prototype

**Résumé :** L'objectif de cet article est de discuter de questions liées à la construction d'une rubrique d'évaluation pour la soutenance du rapport final résultant des stages de pratique d'enseignement. On discutera aussi bien les considérations théoriques que pratiques qui sous-tendent la méthodologie adoptée dans ce processus. Le contexte institutionnel est la maîtrise en enseignement des langues à la Faculté des Lettres de l'Université de Porto (FLUP). Ce processus a été lancé en réponse à un besoin conjointement perçu d'améliorer l'assurance de la qualité dans les procédures d'évaluation et d'harmoniser les pratiques et les outils utilisés dans les différents masters en enseignement. Ainsi, le sujet de cet article est la construction d'une grille d'évaluation fondée sur des critères pour évaluer les niveaux de performance en fonction de l'évaluation du rapport final sur la pratique d'enseignement. Le processus sous-jacent à sa construction a comporté les phases suivantes : i) l'analyse de documents réglementaires pour les soutenances à l'Université de Porto et des grilles similaires utilisées dans d'autres contextes universitaires ; ii) la description du contexte communicatif inhérent à ce genre discursif formel et oral dans un cadre académique et de ses parties constitutives ; iii) la structure de la grille d'évaluation ; iv) la rédaction d'un prototype de la grille.

**Mots-clés :** Formation initiale des professeurs de langues ; grille d'évaluation fondée sur des critères ; rapport final ; soutenance

## Construcción de una rúbrica de evaluación para la defensa pública: de la definición de criterios al prototipo

**Resumen:** El objetivo de este artículo es discutir cuestiones relacionadas con la construcción de una rúbrica de evaluación para la defensa pública del trabajo de fin de máster en los másteres en enseñanza de lenguas de la Facultad de Letras de la Universidad de Oporto (FLUP). Se presentan algunas consideraciones teóricas y prácticas subyacentes a la metodología adoptada en este proceso, que responde a una necesidad relacionada con la mejora de la calidad en los procedimientos de evaluación y con la armonización de prácticas e instrumentos de evaluación utilizados en los diferentes másteres en enseñanza. Así, el tema de este trabajo es la construcción de una rúbrica criterial para evaluar los diferentes niveles de desempeño en la defensa pública de los trabajos de fin de máster. El proceso que subyace a su construcción siguió las siguientes fases: (i) análisis tanto de los documentos de referencia para la defensa pública de la Universidad de Oporto como de rúbricas similares utilizadas en otros contextos universitarios; (ii) descripción de la situación comunicativa inherente a este género de discurso oral formal en un contexto académico y de sus partes constituyentes; (iii) estructura de la rúbrica; (iv) redacción de un prototipo de la rúbrica.

**Palabras clave:** Formación inicial de profesores de lenguas; evaluación criterial; trabajo de fin de máster; defensa pública

## Introdução

Os mestrados profissionalizantes da FLUP (ensino) concluem-se com uma defesa pública, corolário de um percurso de dois anos. Este ato público consiste na apresentação e discussão do relatório relativo a um trabalho de investigação-ação (Almeida *et al.*, 2016) realizado durante o estágio pedagógico, ao qual corresponde a conclusão da “Iniciação à Prática Profissional” (IPP). A reflexão sobre o processo de defesa pública atualmente em vigor suscita algumas questões que nos parecem dever ser objeto de discussão e de revisão, foco principal deste artigo.

A defesa pública configura um género académico formal e convencional intrinsecamente ligado ao processo de avaliação de competências específicas dos estudantes. Usualmente, contudo, as instituições não possuem ou não divulgam um quadro referencial de critérios, descritores de desempenho ou ponderações ligado a esse processo avaliativo. A FLUP, por exemplo, disponibiliza um boletim uniformizado destinado aos membros do júri, que apenas permite uma apreciação sumária e genérica da prestação do candidato.

No sentido de aumentar o nível de transparência e diminuir o grau de subjetividade no processo, pareceu imperioso construir um documento de avaliação com critérios e descritores de desempenho relativos à defesa pública do relatório de estágio.

Neste artigo, far-se-á uma breve contextualização, o enquadramento teórico quanto ao género ‘defesa pública’ e às rubricas de avaliação, a descrição do processo de cocriação de um protótipo de rubrica, e considerações finais.

### 1. Contextualização

Com trabalho realizado desde a sua constituição, em 2015-16, o grupo intercultural e multilingue de reflexão e investigação em didática (GRIDd), de docentes e investigadores dos mestrados em ensino de línguas (português, alemão, espanhol, francês e inglês) da FLUP, tem respondido a necessidades específicas da formação de professores através de reflexão teórico-crítica e de produção de materiais. Com base numa dinâmica de “sociabilidade da investigação” (Provenzano, 2012, p. 147), são criadas soluções (Almeida *et al.*, 2016; Pazos *et al.*, 2017) para questões da formação inicial, através da discussão reflexiva mobilizadora de competências e práticas interculturais e translinguísticas.

No contexto identificado, verificou-se existir um problema análogo ao mencionado por De Miguel (2010) quando se refere ao processo administrativo de avaliação de teses de doutoramento, realçando que as universidades habitualmente não dispõem de quadros de referência claros para se elaborar relatórios normativos, sendo, por isso, necessário recorrer a dados pouco estruturados, precisos e práticos, em que os avaliadores formulam, a partir das suas crenças, juízos sobre como deve ser apresentada uma tese

de doutoramento para poder ser defendida (2010, p. 2). De Miguel explica que “tampoco existe una normativa específica sobre los criterios de evaluación que deberán utilizar los tribunales seleccionados y los procedimientos para otorgar la calificación” (2010, p.6).

De facto, o enquadramento legal dos mestrados da FLUP para o relatório de estágio e respetiva defesa pública contém também uma certa imprecisão. Quanto à avaliação do relatório, os regulamentos dos mestrados destacam: apresentação global; apresentação e qualidade da informação e referências; ortografia e outros aspetos gramaticais; estrutura e desenvolvimento lógico; contextualização teórica atual, diversificada e fundamentada; análise consistente do tema identificado, revelando originalidade, inovação e atualidade nas fontes utilizadas e clareza no resumo, na forma e no conteúdo. No entanto, não são mencionados elementos específicos que permitam avaliá-los de forma objetiva.

Do mesmo modo são referidos aspetos a que os membros do júri devem atender quanto ao discurso do candidato na defesa pública, como: a) a organização e clareza na apresentação do conteúdo; b) o profissionalismo e atitude; c) o respeito pelo tempo concedido para a apresentação; d) o grau de segurança e confiança nas respostas dadas à arguição; e) a apresentação de exemplos adicionais relevantes para as respostas (se oportuno); f) a consistência e qualidade global das respostas dadas.

Tendo em conta o exposto, percebe-se que a avaliação final e os seus fundamentos correspondem a valorações subjetivas sobre o que cada membro do júri considera “um relatório excelente/bom/aceitável/mau ou inaceitável”.

A necessidade de uma ferramenta que torne mais transparente, equitativo e formativo o processo de preparação e de avaliação da defesa pública do relatório de estágio justifica a elaboração de uma rubrica. Com esta ferramenta de avaliação, será possível dar a conhecer aos mestrandos critérios e indicadores de desempenho que orientem o processo anterior à defesa pública, harmonizar procedimentos entre júris de diferentes mestrados e procurar diminuir a subjetividade inerente ao processo de deliberação dos membros do júri (internos ou externos à instituição).

Para a criação da rubrica de avaliação da defesa pública do relatório de estágio, foi necessário aprofundar questões teóricas quanto às características da defesa pública como género textual e à elaboração de rubricas de avaliação.

## 2. Enquadramento teórico

### 2.1. O género “defesa pública”

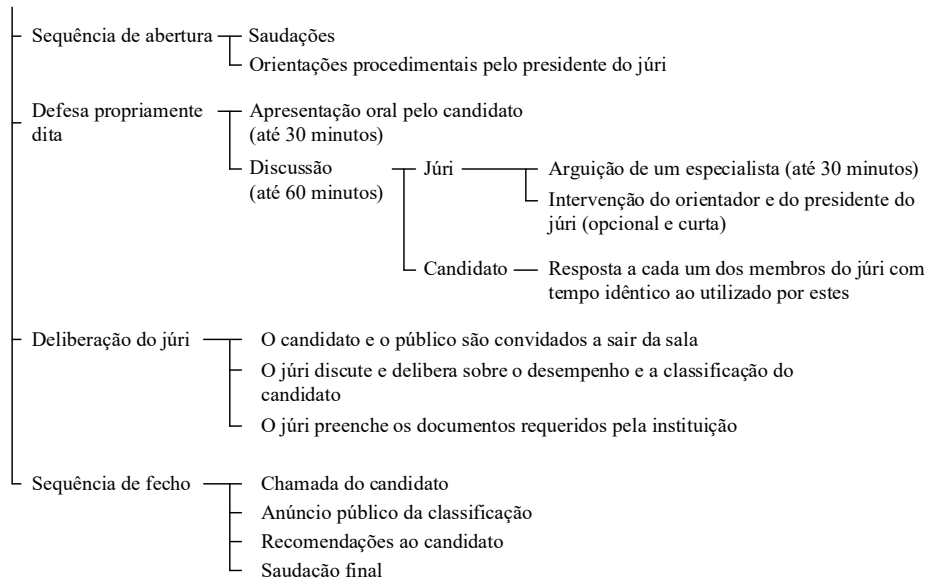
Na linha de Bakhtine (1984), vários autores (e.g. Bronckart, 1996; Coutinho, 2019; Marques, 2020) postulam que os géneros devem ser analisados na dependência das atividades sociais, pois as práticas linguísticas estão estreitamente associadas a esses contextos, sendo atualizadas em géneros orais ou escritos.

A “defesa pública” (ou *defensio*, em latim; *viva*, em inglês; *defensa oral*, em espanhol; *soutenance*, em francês) pode ser definida como um género oral formal(izado) inserido no domínio das provas académicas, ligando as áreas académica e científica.

Não sendo objetivo uma caracterização de todas as dimensões do género, referimos essencialmente a sua estrutura. Considerando as propostas existentes sobre defesas públicas para a obtenção do grau de doutoramento (Burke, 1994; Hasan, 1994; Swales, 2004; e Mežek & Swales, 2016, e.o.), é comumente aceite tratar-se de um género académico formal e formalizado, cujas diferenças decorrem em grande medida da variação dos contextos da prova, nomeadamente em relação à cultura institucional e à sua organização. Apresentam variação terminológica, mas com significado muito próximo, apresentando tipicamente quatro movimentos.

Apesar da escassez de trabalhos sobre a defesa pública de mestrados, foi possível adaptar os modelos genéricos da defesa pública nos nossos mestrados, respeitando os parâmetros específicos (esquema 1).

#### Defesa Pública

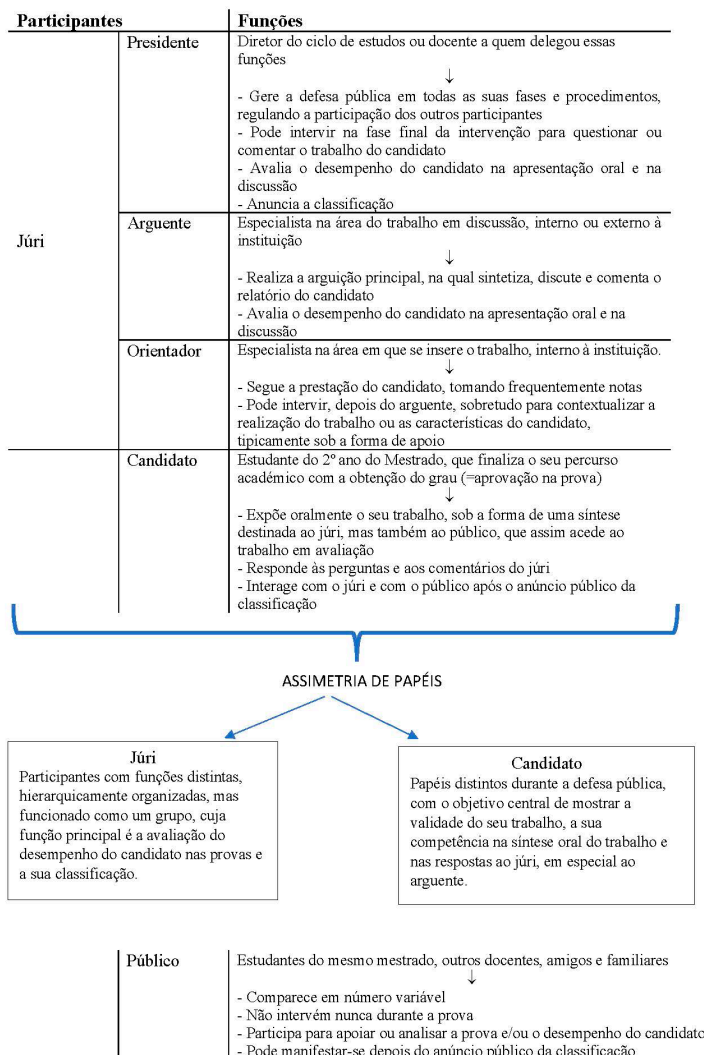


Esquema 1: Estrutura prototípica da defesa pública nos mestrados em ensino de línguas  
Fonte: Elaboração dos autores

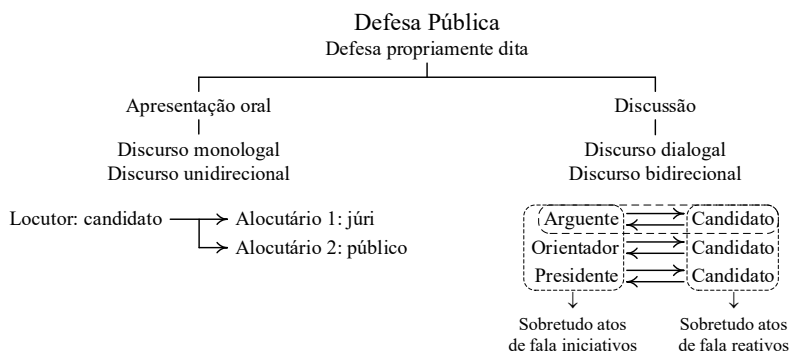
Prototipicamente, este é um evento público altamente regulado, com organização fixa. Tem sido definido por analogia ao tribunal e à atuação teatral: a primeira tem raízes nas provas académicas das universidades medievais, a segunda tem sido proposta por vários autores, na sequência do modelo Goffman (1959), entre eles van Der

Heide, Rufias & Supper (2016, p. 474), que propõem a análise da defesa pública como uma *performance* assimétrica baseada num guião predefinido.

Tendo em conta estas propostas, definimos de seguida os papéis que desempenham os participantes e as inter-relações que estabelecem entre si (esquema 2), bem como as estratégias discursivas que estruturam a fase da defesa propriamente dita (esquema 3).



Esquema 2: Papéis e funções dos participantes na defesa pública  
 Fonte: Elaboração própria dos Autores



Esquema 3: Análise discursiva global da defesa propriamente dita  
Fonte: Elaboração própria dos Autores

Intrinsecamente vinculada a este género está a avaliação (como classificação e certificação) dos estudantes. Existe necessidade de rigor na explicitação e aplicação de critérios e indicadores de desempenho, sendo capital considerar o papel das rubricas de avaliação.

## 2.2. Elaboração de rubricas de avaliação

Para Hadji (1994), a avaliação consiste numa relação entre o objeto avaliado e o modelo ideal. Na avaliação criterial, a referência é o desempenho-alvo expresso em critérios específicos que descrevem “o ser à luz de um dever-ser previamente determinado” (1994, p. 76), que ajudam a apreciar e interpretar a realidade avaliada considerando o nível de consecução desejável e não, como esclarece Fernandes (2020b, p. 3), com uma média ou um grupo de alunos.

Na defesa pública do relatório de estágio, produz-se uma avaliação sumativa. Nesta, como na avaliação formativa, para garantir uma prática transparente e rigorosa, há que definir critérios, descrever diferentes níveis de desempenho e decidir o instrumento de recolha de dados (Fernandes, 2021).

Entre os possíveis instrumentos para uma avaliação criterial destaca-se a rubrica (Alsina, 2013; Cano, 2015; Fernandes, 2020b), um conceito atualmente popular na literatura sobre a avaliação de competências (Cano, 2015). Para esta autora, a rubrica é “un registro evaluativo que posee ciertos criterios o dimensiones a evaluar y lo hace siguiendo unos niveles o gradaciones de calidad y tipificando los estándares de desempeño” (2015, p. 267). Trata-se dum instrumento para medir o desempenho na execução duma tarefa concreta, em formato de grelha, a partir de escalas quantitativas e/ou qualitativas associadas a critérios pré-estabelecidos organizados hierarquicamente em distintos níveis de desempenho, do mais para o menos exigente (Alsina, 2013; Fernandes, 2020b; Goodrich, 1996; Martínez-Rojas, 2008).

As rubricas também podem ser usadas na avaliação sumativa “para mobilizar informação para efeitos da atribuição de classificações”, devendo além da descrição da tarefa, dos critérios e dos níveis de desempenho relativos a cada critério, incluir a definição de uma escala que atribui a cada nível de desempenho uma dada classificação ou pontuação (Fernandes, 2020b, p. 6).

Convém clarificar desde já dois conceitos. Por um lado, por critérios (de avaliação) entendem-se as “afirmações [...] que identificam o que se consideram ser as características ou os atributos que o desempenho dos alunos deve ter quando estão a trabalhar numa dada tarefa de avaliação”, a redigir de forma simples e breve (2020b, p. 4). Tal como adverte Fernandes, esses critérios não devem ser confundidos com os de classificação, definidos como “meras técnicas, mais ou menos arbitrárias, mais ou menos intuitivas, que se utilizam para atribuir classificações.” (2020a, pp. 4-5). Por outro lado, as descrições dos níveis de desempenho são “descrições de graus de consecução possíveis ou diferentes níveis de aprendizagem para um dado critério” (2020b, p. 5). Quando associados a uma dada classificação ou pontuação, sobretudo em práticas de avaliação sumativa que requerem uma classificação, são designados por alguns especialistas como *standards* (2020a, p. 5).

Existem diferentes tipos de rubricas em função do objetivo e das suas componentes. A maioria dos especialistas distingue entre rubricas holísticas e analíticas (Del Pozo, 2013; Fernandes, 2021; Martínez-Rojas, 2008). As primeiras apresentam uma descrição dos níveis de desempenho globais e fornecem uma valoração integrada dos distintos critérios (Fernandes, 2021, p. 13; Martínez-Rojas, 2008), normalmente associadas a uma avaliação mais sumativa (Del Pozo, 2013; Fernandes, 2021, p. 13), pois tornam mais fácil o processo de atribuição duma classificação, embora sem informação concreta sobre os diversos critérios (Fernandes, 2021, p. 13). As segundas são mais abrangentes: ao contrário das anteriores, definem os critérios que vão ser avaliados, a sua ponderação quanto à nota global e os diferentes níveis de consecução para cada descritor separadamente. Neste caso, o processo de avaliação, apesar de mais lento, fornece um *feedback* muito mais preciso, sendo estas as rubricas mais indicadas para uma avaliação formativa (Del Pozo, 2013; Martínez-Rojas, 2008). Contudo, nada impede que sejam usadas na avaliação sumativa (Fernandes 2021, p. 12), pois as rubricas analíticas permitem “classificar melhor, com mais transparência e mais justiça, do que qualquer outro procedimento mais ou menos intuitivo ou mais ou menos baseado em algoritmos” (2021, p. 15), se se ultrapassar a dificuldade que colocam a uma descrição completa de um determinado desempenho e a uma perceção clara da interação entre os distintos critérios que a integram (2021, p. 15).

A elaboração de rubricas implica (Del Pozo, 2013; Martínez-Rojas, 2008) indicar o objeto a avaliar, os critérios e, se oportuno, a ponderação de cada um na classificação final; determinar o tipo de rubrica a criar e decidir os níveis de valoração de cada um



dos critérios. Del Pozo (2013) recomenda: começar com uma rubrica de três níveis (máximo, mínimo e intermédio), podendo proceder-se ou não à sua ampliação; redigir os descritores com níveis de desempenho específicos; construir uma grelha com todas estas informações e implementar este instrumento para verificar se cumpre o objetivo, sendo aconselhada a sua revisão por outros especialistas. Também Helvoort (2010, p. 27) apresenta uma proposta para a construção faseada duma rubrica, com três momentos: identificação dos critérios por *brainstorming*; descrição do nível de desempenho ‘excelente’ e ‘insuficiente’ de cada critério; elaboração, que envolve a decisão do número de níveis de desempenho a integrar e a descrição dos níveis intermédios para cada critério; testagem, que implica a utilização da rubrica a partir de exemplos de trabalhos de estudantes de vários níveis e correções necessárias.

Como referido, quando bem elaborada e compreendida pelos utilizadores, a rubrica oferece vantagens a todos os agentes envolvidos no processo avaliativo (Alsina, 2013; Cano, 2015; Fernandes, 2020b; Helvoort, 2010), contribuindo para “una evaluación fácilmente interpretable, justa y transparente para profesores y estudiantes” (García-Sanz, 2014, p. 92). Serve, como afirma Alsina, de “guía u hoja de ruta”, uma vez que “muestra las expectativas que alumnado y profesorado tienen y comparten sobre una actividad o varias actividades, organizadas en diferentes niveles de cumplimiento” (2013, p. 8).

Assinalamos, contudo, algumas dificuldades ligadas à construção e ao uso de rubricas. Primeiro, as competências não podem ser diretamente avaliadas e o processo de transformação em resultados de aprendizagem e em condutas observáveis pode ser perigoso, havendo o risco duma excessiva desfragmentação da competência ou do esquecimento de aspetos invisíveis da competência que são difíceis de avaliar (Cano, 2015). Assim, elaborar uma boa rubrica que desafie esses perigos, que descreva desempenhos com rigor para reduzir o nível de subjetividade dos avaliadores, seguindo uma gradação de qualidade bem articulada e uso duma linguagem clara, é uma tarefa difícil e demorada (Cano, 2015; Helvoort, 2010; Goodrich, 1996).

A cocriação de uma rubrica para a avaliação da defesa pública do relatório de estágio teve em conta as características da defesa pública e das rubricas de avaliação. Descrevemos agora esse processo.

### 3. Processo

O processo de cocriação do protótipo teve 3 etapas: recolha e análise documental, análise comparativa e construção do protótipo.

#### 3.1. *Recolha e análise documental*

Iniciou-se o processo com a recolha de documentos sobre defesas públicas de trabalhos de final de ciclo em instituições nacionais e internacionais. O material reunido

não foi abundante, mas permitiu um *corpus* de rubricas de avaliação (cf. anexo I) com nove documentos de cinco universidades, de cursos ministrados em alemão, espanhol, francês e inglês.

Os documentos continham informações sobre critérios de avaliação, embora apresentassem configurações, formatos e estruturas muito diferentes entre si, quanto a: (i) título ou designação do documento (manual, rubricas, orientações...); (ii) objeto de avaliação (trabalho escrito em combinação com a prova oral); (iii) estrutura (grelha para registar dados resultantes da observação; texto com apresentação de critérios; tabela de dupla entrada com critérios, descritores e escala com correspondência ou não a uma ponderação quantitativa); (iv) escalas (com intervalos de 3, 4, 7 ou 8 níveis qualitativos); (v) formulação dos critérios (frases afirmativas; frases interrogativas; nominalizações).

### 3.2 *Análise comparativa*

A análise, baseada em Bardin (2014), foi realizada em duas fases: (i) identificação dos critérios de avaliação através de análise categorial; (ii) sistematização das características do desempenho dos estudantes através de uma análise da expressão.

No primeiro procedimento foi utilizada a análise temática, com um sistema de categorias resultante de dois processos: umas definidas a partir de referências teóricas sobre rubricas de avaliação de aprendizagens, tais como “objeto de avaliação”, “critério” e “ponderação” (cf. capítulo 2.2.); outras emergentes da leitura do *corpus*. Esta análise incluiu 112 unidades de registo. Embora o número limitado da amostra não permita dados quantitativos robustos, a contagem destas unidades por (sub)categoria permitiu captar certas tendências de uma rubrica de avaliação de defesa pública.

Para o segundo procedimento, a análise da expressão baseou-se em indicadores lexicais, a partir da identificação de palavras valorativas associadas a características do desempenho dos estudantes em aspetos como exposição, discurso, material. O foco incidiu sobre as categorias sintáticas “adjetivo”, “nome” e “advérbio”, associadas a critérios, descrições de desempenho por níveis e escalas.

O conjunto dos documentos exhibe elementos fulcrais numa rubrica e todos indicam a tarefa que é objeto de avaliação de desempenho.

A componente mais referida é a ‘Exposição oral’ (ocorre em 75% dos casos, contra 13% de resposta a questões e 12 % da competência comunicativa), no âmbito da qual se consideram seis critérios (estrutura, conteúdo, discurso, suporte visual, linguagem corporal e voz, gestão do tempo), cuja distribuição se apresenta no gráfico 1:



Gráfico 1: distribuição dos critérios relativos a 'Exposição oral'  
Fonte: Elaboração dos Autores

Três (RE03, RE04, RUK01) documentos (33%) apresentam descrições de níveis de desempenho. Em 78% dos documentos, são usadas escalas com linhas horizontais, onde o avaliador pode marcar a sua avaliação, existindo variação quanto ao número de níveis, entre 3 ([RE05]) e 8 ([RUK01]), e quanto à expressão designativa de cada nível, numéricas em dois casos e verbais nos restantes cinco.

Os diferentes níveis de desempenho são designados através de expressões diversas. Considerando apenas os dois polos opostos, as expressões são: *very good - fail*; *completamente adecuado - inadecuado*; *excelente - insuficiente*; *adecuado - insuficiente*; *présente une qualité exceptionnelle - très insatisfaisant*; *exemplary first class honours - fail*. Quatro escalas começam com o intervalo de avaliação máxima e duas com a mínima. [RE03] apresenta uma escala numérica, sendo 1 o intervalo de avaliação mínimo e 4 o intervalo de avaliação máxima. Em três documentos ([RA01], [RE01], [RUK01]), a escala qualitativa está associada à correspondente escala quantitativa. Um documento apresenta a ponderação para a atribuição da classificação final.

Os dados analisados constituíram matéria de reflexão para a segunda fase do trabalho, a construção da rubrica de avaliação.

### 3.3 Início da construção do protótipo

A construção teve início com decisões quanto à finalidade, ao tipo e aos elementos constitutivos, tendo em conta o contexto institucional específico de formação inicial de professores de línguas (cf. capítulo 1). Foi ainda decidido que o debate decorreria em português, mas o protótipo seria redigido em inglês. O trabalho realizado simultaneamente em duas línguas (inglês e português) constituiu uma estratégia intencional para aprofundamento da reflexão. O trabalho colaborativo teve como método o debate em torno de questões conceptuais e terminológicas.

Assumiu-se, assim, a construção de uma rubrica de avaliação que:

- a. se dirigisse não só a professores (orientadores e arguentes), como instrumento de avaliação sumativa (classificação), mas sobretudo aos estudantes, como guia para a preparação da defesa pública, fornecendo-lhes informação específica sobre as competências a desenvolver e a preparação a fazer para atingir o nível de desempenho mais elevado possível;
- b. possibilitasse a conversão da informação relativa à escala qualitativa em quantitativa através de uma fórmula matemática para a classificação final;
- c. contivesse descrições o mais informativas possível quanto aos níveis de desempenho (ou descritores) para cada um dos critérios de avaliação, contribuindo para a transparência no processo de avaliação.

O propósito do instrumento é a sua utilidade não só para orientadores e arguentes, mas também para os próprios estudantes. Por isso, considerou-se pertinente conceber um formato com parâmetros e descritores para cada nível de desempenho.

#### *a) Critérios*

Na definição de critérios, o debate centrou-se em conceitos usualmente expressos por diferentes designações. Por exemplo, o termo inglês “delivery” é unívoco, mas em português ocorre com designações diversas (“discurso oral”, “exposição oral”, “apresentação oral”): há diferenças entre elas ou apenas sinonímia? Após reflexão conjunta, optou-se por: **A. Content**; **B. Structure**; **C. Delivery**; **D. Discussion** (cf. anexo II).

Diferentemente da tendência verificada na análise documental e comparativa apresentada (cf. 4.1. e 4.2.), em que os critérios “Content” e “Structure” estavam dentro do parâmetro “Delivery”, optou-se por assumi-los como autónomos, pois o estudante revela competências científicas (ligadas ao conteúdo e à estrutura) não só na exposição, mas também na discussão.

No critério “Content” agregou-se: tema, problemática, quadro teórico e revisão da literatura, termos técnicos e linguagem científica, metodologia, validade e fiabilidade dos dados (recolha, interpretação e apresentação). Incluiu-se o conhecimento científico específico sobre metodologia de investigação qualitativa em didática de línguas por este estar no cerne do estágio pedagógico.

Relativamente ao critério “Structure”, incluiu-se uma frase introdutória, com informação sobre os aspetos valorizados na apresentação oral: justificação do estudo, questões de investigação, fundamentação teórica, técnicas da recolha de dados, apresentação e discussão dos dados, conclusões (tendo em conta objetivos da pesquisa e o desenvolvimento pessoal e profissional do candidato), limitações e recomendações e bibliografia principal. A informação complementar refere a obrigatoriedade de adequação às normas de um trabalho científico, coerência entre diversos elementos (objetivos e metodologia, técnicas de análise de dados, a clareza da formulação dos

objetivos), conclusões e resultados, relevância da fundamentação teórica, atualidade e relevância da bibliografia e organização lógica dos diversos elementos.

O critério “Delivery” salienta as características do género académico oral formal e incorpora aspetos como volume de voz, fluência do discurso oral, segurança na postura, legibilidade e clareza dos suportes visuais, gestão de tempo.

O critério “Discussion” diz respeito ao desempenho do estudante após a tomada de palavra do arguente. Prende-se com a capacidade de responder às questões suscitadas pelo arguente com profundidade, segurança, clareza, coerência e espírito crítico.

### ***b) Descritores***

De seguida, produziu-se uma descrição o mais completa possível para cada um dos critérios, tendo em conta o nível máximo de desempenho. Neste processo, foram partilhadas as expectativas máximas que cada elemento do grupo de trabalho considerava quanto às aprendizagens a desenvolver nos estudantes. Tratou-se de um exercício de explicitação de um perfil de estudante que fosse comum aos diferentes cursos de mestrado em ensino de línguas.

A estratégia de reflexão utilizada consistiu em comparar o conceito, a sua designação e significado nas várias línguas. Seguindo essa estratégia, descreveu-se o nível de desempenho de “excelente” em inglês. De seguida, procedeu-se à tradução do texto de inglês para português e novamente para inglês. O processo de criação de textos equivalentes em duas línguas adicionou uma nova dinâmica à definição dos descritores: a transformação ou transcodificação teve de respeitar o conteúdo da “mensagem”, bem como características formais e funcionais. A evolução recíproca de ambos os textos por meio de rascunhos sucessivos e traduções resultou numa explicitação concetual significativa.

A definição de critérios não é tarefa fácil. A sua explicitação obrigou a sucessivas retextualizações devido a questões linguísticas e lexicais. Por exemplo, um dos aspetos alterados da primeira para a versão final foi o tempo verbal. (1) e (2) mostram a modificação realizada depois de concluirmos que os descritores deviam estar no presente do indicativo, dado o seu valor genérico.

1. The candidate *addressed* all major aspects of the research project. (Version 1)
2. The candidate *fully addresses the main issues* raised by the discussant/jury, *always at appropriate depth and detail*. (Final version)

### ***c) Distinções de nível***

Embora o protótipo construído corresponda, por enquanto, ao “excelente”, a fixação dos descritores de desempenho neste nível específico levou a um debate sobre expressões linguísticas para distinguir graus da escala. Esse exercício consistiu em

explicitar a intensidade e grau (desempenhos qualitativos) e a frequência (desempenhos quantitativos). Para esta graduação, foram cruciais adjetivos, advérbios e expressões linguísticas. Por exemplo, o descritor “The candidate addresses the main issues raised by the discussant/jury.” terá cinco formas, cada uma correspondendo a um determinado nível de desempenho. Neste caso, o escopo da variação é o modo como a abordagem das questões suscitadas pelo arguente é realizada: as palavras e/ou expressões linguísticas que marcam o nível “excelente” são constituintes do predicado, desempenhando a função de modificadores: “completamente” (“plenamente”, “totalmente”) e “com profundidade e detalhes apropriados”. Existe ainda o valor de tempo marcado pelo advérbio “sempre”. Em suma, o enunciado que expressa o desempenho em causa no nível “excelente” é “The candidate fully addresses the main issues raised by the discussant/jury, always at appropriate depth and detail.”

Neste nível (“Excelente”), os elementos associados à qualidade foram sobretudo adjetivos: *secure, deep, up-to-date, wide-ranging, rigorous, full, clear, coherent, appropriate, correct, very positive, systematic ... excellent*. A frequência foi definida a partir de advérbios como *always, continuously, consistently, invariably, unfailingly, regularly, repeatedly*.

O processo de redação de enunciados que traduzissem com clareza o desempenho associado a um determinado critério no nível específico foi predominantemente linguístico, com aprofundamento devido à comparação entre línguas. De entre as muitas questões em debate, destaca-se a relacionada com o advérbio “sempre”. O seu uso em certos contextos é inviável, dado que há comportamentos que não se conseguem contabilizar durante uma interação verbal como a defesa pública como, por exemplo, “The candidate maintains good eye contact when communicating with jury members (...)” vs. “The candidate always maintains contact when communicating with jury members (...)”.

#### **d) Ponderação**

Tendo criado uma rubrica analítica que compreende as modalidades de avaliação formativa e sumativa, procedeu-se à elaboração de uma fórmula de cálculo para a atribuição de uma classificação final, conduzindo, assim, a um juízo globalizante sobre a defesa do relatório. Para a conversão da informação relativa à escala qualitativa em quantitativa, foi discutido o peso relativo de cada critério. A distribuição não foi equitativa, uma vez que se considerou que o “Content” e a “Discussion” são as componentes fundamentais de uma defesa pública: Content (40 pontos); Structure (20); Delivery (20); Discussion: (40). Estabeleceu-se os 120 pontos como pontuação máxima por comodidade de quantificação, procedendo depois à conversão para a escala de 0 a 20 valores. Para os restantes níveis de execução (muito bom, bom e satisfatório), os pontos foram distribuídos proporcionalmente.

	Excellent (20-19)	Very good (18-17)	Good (16-15-14)	Satisfactory (13-12-11-10)
A. Content (40)	40-39-38-37	36-35-34-33	32-31-30-29-28-27	26-25-24-23-22-21-20
B. Structure (20)	20-19	18-17	16-15-14	13-12-11-10
C. Delivery (20)	20-19	18-17	16-15-14	13-12-11-10
D. Discussion (40)	40-39-38-37	36-35-34-33	32-31-30-29-28-27	26-25-24-23-22-21-20
Total points scored: /120 points				Final mark: /20

Tabela 1: fórmula para classificação final

## Conclusão

É consabido o potencial das rubricas na transparência dos processos de avaliação formativa e sumativa. Foi justamente a procura de transparência e equidade que motivou a realização do trabalho apresentado, de natureza colaborativa interdisciplinar.

O debate como técnica de cocriação, pela sua natureza interrogativa, revelou-se ideal por abrir espaço por confronto de perspetivas, aprofundamento de reflexão crítica, partilha de saberes, busca de precisão concetual e terminológica.

Apesar de ser usual fazer a apresentação de rubricas de avaliação com foco no produto, optamos por descrever o processo de construção de uma rubrica para destacar questões e reflexões, nem sempre tão conhecidas, que emergem ao longo do percurso. Há ainda questões por responder: será que a nossa grelha apenas traz mais burocracia para o processo de avaliação no ensino superior?

O desenvolvimento do protótipo prosseguirá até à implementação para testagem, que permitirá uma avaliação de eficácia com base em dados de estudantes e de membros de júris de atos de defesa pública.

## Referências

- Almeida, J. et al. (2016). *Projetos de investigação-ação. Orientações gerais para a elaboração do relatório de estágio em ensino de Português e de línguas estrangeiras*. (pp.1-25).Edições FLUP.
- Alsina Masmitjà, J. (Coord.). (2013). *Rúbricas para la evaluación de competencias*. ICE y Ediciones Octaedro.
- Bakhtine, M. (2004/1984). *Esthétique de la Création Verbale*. Gallimard.
- Bardin, L. (2014). *Análise de conteúdo*. Edições 70.
- Burke, P. J. (1994). Segmentation and control of a dissertation defense. *Advances in Discourse Processes*, 43, 95-95.

- Cano García, E. (2015). Las rúbricas como instrumento de evaluación de competencias en educación superior: ¿uso o abuso? *Profesorado. Revista de currículum y formación del profesorado*, 19 (2), 265-280. Recuperado de <https://bit.ly/2Ds8zzy>
- Coutinho, A. 2019. *Texto e(m) linguística: teorias, cruzamentos, aplicações*. Edições Colibri.
- De Miguel, M. (2010). La evaluación de tesis doctorales. Propuesta de un modelo. *RELIEVE: Revista Electrónica de Investigación y Evaluación Educativa*, 16(1), 1-18. Recuperado de <https://bit.ly/3AT4S00>
- Del Pozo, J. A. (2013). *Competencias profesionales. Herramientas de evaluación: el portafolios, la rúbrica y las pruebas situacionales*. Narcea Ediciones.
- Fernandes, D. (2020a). Folha Critérios de Avaliação. Projeto de monitorização, acompanhamento e investigação em avaliação pedagógica (Projeto MAIA). Recuperado de <https://bit.ly/3kL7rMe>
- Fernandes, D. (2020b). Folha Rúbricas de Avaliação. Projeto de monitorização, acompanhamento e investigação em avaliação pedagógica (Projeto MAIA). Recuperado de <https://bit.ly/3mah7iE>
- Fernandes, D. (2021). Folha Para uma Iniciação às Práticas de Classificação Através de Rubricas. *Projeto de monitorização, acompanhamento e investigação em avaliação pedagógica (Projeto MAIA)*. Recuperado de <https://bit.ly/3zPI28q>
- García-Sanz, M. P. (2014). La evaluación de competencias en Educación Superior mediante rúbricas: un caso práctico. *Revista Electrónica Interuniversitaria de Formación del Profesorado*, 17(1), 87-106. Recuperado de <https://bit.ly/2BgVhDk>
- Goffman, E. (1959). *The Presentation of Self in Everyday Life*. Doubleday.
- Goodrich, H. (1996). Understanding Rubrics. *Educational Leadership: Teaching for Authentic Student Performance*, 54(4). Recuperado de <https://bit.ly/2MwkihU>
- Hadji, C. (1994). *A Avaliação, regras do jogo: das intenções aos instrumentos*. Porto Editora.
- Hasan, R. (1994). Situation and the definition of genres. In A. D. Grimshaw (Ed.), *What's going on here: Complementary studies of professional talk* (pp. 127-167). Ablex Publishing Corporation.
- Helvoort J. van (2010). A scoring rubric for performance assessment of information literacy in *Dutch Higher Education*. *Journal of Information Literacy*, 4(1), 22-39 <http://dx.doi.org/10.11645/4.1.1256>
- Marques, M. A. (2020). Os géneros das provas académicas: entre discurso académico e discurso científico. *Matraga. Estudos Linguísticos & Literários*, 27 (50), 469-486. DOI: <https://doi.org/10.12957/matraga.2020.50983>
- Martínez-Rojas, J. G. (2008). Las rúbricas en la evaluación escolar: su construcción y su uso. *Avances en Medición*, 6, 129-138. Recuperado de <https://bit.ly/2FcoALC>
- Mežek, Š & Swales, John M. (2016). PhD defences and vivas. In K. Hyland & P. Shaw. *The Routledge Handbook of English for Academic Purposes*. (pp. 127-167) Routledge.
- Pazos, M. et al. (2017). Construção de um referencial de avaliação de desempenho do estudante de línguas em estágio pedagógico: problemas e desafios. *Sensos-e*, Vol. IV, n.º 1, 3-11.
- Provenzano, F. (2012). Francophonie. Idéologie, variation, canon : modèles québécois pour la francophonie littéraire. *Tangence*, 100, 133-152.



Swales, John M. (2004). Research Genres. *Explorations and Applications*. CUP.

Van der Heide, A., Rufas, A. & Supper, A. (2016). Doctoral Dissertation Defenses: Performing Ambiguity Between Ceremony and Assessment. *Science as Culture*. 25 (4), p. 473– 495. DOI: <http://dx.doi.org/10.1080/09505431.2016.1183610>

## Anexo I

Código	Designação	Instituição	Grau
RA01	Präsentation Bachelor-Arbeit	Universität Hohenheim Agrartechnik (Alemanha)	Bachelor/ Master
RA02	Defensio: Richtlinien	Universität Graz (Áustria) Translationswissenschaft (Tradução)	
RE01	Plantilla de evaluación del Trabajo Fin de Máster	Universidad Complutense de Madrid (Espanha)	Máster
RE02	Trabajo de Fin de Máster UC3M: Defender el TFM. Criterios	Universidad Carlos III Madrid (Espanha)	Máster
RE03	Rúbrica UCM. Matriz de criterios para la presentación y defensa oral	Universidad Carlos III Madrid (Espanha)	Máster
RE04	Rubrica para evaluar la defensa oral de un Trabajo Final de Grado en GIOI	Universitat Politècnica de València/Escuela Técnica Superior de Ingenieros Industriales (Espanha)	
RE05	Rúbrica para la evaluación de la defensa oral del trabajo fin de grado	Universidad del País Vasco Facultad de relaciones laborales y trabajo social (Espanha)	
RF01	Évaluation TFS		
RUK01	Law School Undergraduate Grading Descriptors for Oral Presentations. University of Hull Grading Descriptors UGMA	University of Hull Law School (Reino Unido)	Undergraduate Grading

## Anexo II

	<b>Teacher Education: Viva Assessment Criteria: Public Defence of Final Report (Action Research Project).</b> <b>Grade: Excellent</b>
<b>A. Content. (/40)</b> The topic of the research is contextualised and pedagogically justified. The background to and relevance of the research project are perfectly defined. The candidate demonstrates a solid knowledge of the field (state of the art). Key terms are well-defined in specialized language appropriate for the academic context. The candidate demonstrates a sound, ample knowledge of the specific literature of the chosen areas of study. Information included is highly relevant (selected and synthesised) and completely aligned with the stated objectives/research questions of the AR project. The research methodology (data collection instruments, analysis and presentation) is appropriate and clearly described and rigorously justified. The results/findings are clearly presented, described and analysed. The interpretation of data is fully informed by the literature, taking specific account of the objectives and research questions of the project. The contribution of the study is recognised and its relative importance fully understood. The project is both innovative and original in authorship.	
<b>B. Structure. (/20)</b> <i>The presentation must include justification of the research, objectives, research question(s), theoretical overview, data collection techniques, presentation and interpretation of data, conclusions (taking into account the research objectives and the personal and professional development of the candidate), limitations and recommendations and main bibliographical references.</i> The type of presentation is entirely in keeping with the norms associated with academic/scientific presentations. Each section is completely coherent with the stages and purpose of the research, presenting information in a clear, logical sequence. The presentation reflects the development of the action research project and includes both matchless theoretical and empirical evidence. A sufficient number of relevant, up-to-date references are cited. There is a robust summary of the main findings of the research project within a conclusion which is entirely coherent with the research itself. Recommendations are clear and relate specifically to the research project. Limitations of the project are fully acknowledged and explained.	
<b>C. Delivery. (/20)</b> The candidate uses a clear, audible voice throughout. The presentation is delivered in a highly fluent and accurate manner while maintaining a confident posture. The tone of the presentation is suitably formal and academic. Content is well communicated through the spoken channel, without any reliance on a written script. There are few hesitations and limited use of fillers. Visual aids (PowerPoint/Prezi) are very well prepared, in terms of quantity and density of information displayed. All slides are easily legible. The candidate maintains good eye contact when communicating with jury members and is appropriately dynamic (e.g. gestures, body posture, etc.). The length of presentation strictly adheres to the assigned time limit.	
<b>D. Discussion. (/40)</b> The candidate fully addresses the main issues raised by the discussant/jury, always at appropriate depth and detail. The candidate has anticipated potential questions and prepared responses. The candidate demonstrates ability to answer thoroughly and articulately a range of questions posed, to do with the content, structure and challenges of the research project. The candidate's answers/clarifications are consistently clear and coherent and indicative of the ability to think critically and utilize higher order thinking skills.	

Simone Tomé

Faculdade de Letras da Universidade do Porto

Email: stome@letras.up.pt

Nicolas Hurst

Faculdade de Letras da Universidade do Porto

Email: nrhurst@letras.up.pt

Orcid : <http://orcid.org/0000-0001-7079-369X>

Marta Anido

Faculdade de Letras da Universidade do Porto

Email: manido@letras.up.pt

Mónica Lorenzo

Faculdade de Letras da Universidade do Porto

Email: mlorenzo@letras.up.pt

**José Almeida**

Faculdade de Letras da Universidade do Porto  
Email: jalmeida@letras.up.pt  
Orcid : <http://orcid.org/0000-0002-4564-2766>

**Fátima Silva**

Faculdade de Letras da Universidade do Porto  
Email: mhenri@letras.up.pt  
Orcid : <http://orcid.org/0000-0003-2360-5136>

**Pilar Martinez**

Faculdade de Letras da Universidade do Porto  
Email: mmartinez@letras.up.pt  
<https://www.cienciavita.pt/B316-C925-2079>

**Sónia Rodrigues**

Faculdade de Letras da Universidade do Porto  
Email: srodrigues@letras.up.pt  
<https://www.cienciavita.pt/4F10-1BE5-36A2>

**Maria Ellison**

Faculdade de Letras da Universidade do Porto  
Email: mellison@letras.up.pt  
Orcid : <http://orcid.org/0000-0001-9060-4147>

**Manuela Teixeira**

Faculdade de Letras da Universidade do Porto  
Email: dteixeira@letras.up.pt

**Correspondência**

José Domingues de Almeida  
Faculdade de Letras da Universidade do Porto  
Via Panorâmica, s/n  
4150-564 Porto – Portugal

Data de submissão: Abril 2021´

Data de avaliação: Maio 2021

Data de publicação: Novembro 2021