

UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

**FACULTADE DE FILOLOXÍA
DEPARTAMENTO DE LINGUA ESPAÑOLA**

**SISTEMA DE CERTIFICACIÓN DE LINGUA GALEGA
(CELGA) NO *MARCO COMÚN EUROPEO DE REFERENCIA*
PARA AS *LINGUAS*. ESTUDO DE FIABILIDADE E VALIDEZ**

MARGARITA CHAMORRO DÍAZ

2013

Sistema de certificación de lingua galega (Celga) no
*Marco Común Europeo de Referencia para as
Linguas*. Estudo de fiabilidade e validez

Tese de doutoramento presentada no Departamento de Lingua Española, da
Facultade de Filoloxía, da Universidade de Santiago de Compostela, por
Margarita Chamorro Díaz e dirixida polo profesor *Tomás Jiménez Juliá*.

Vº e Pr. do Director

Asdo. Tomás Jiménez Juliá

Asdo. Margarita Chamorro Díaz

ÍNDICE

CAPÍTULO 1

1. Panorámica sobre a investigación en avaliación de linguas	11
1.1. Antecedentes: anos sesenta e setenta	12
1.2. Anos oitenta: inicios do método comunicativo	14
1.2.1. A avaliación e o currículo.....	15
1.2.2. Os exames a grande escala	16
1.3. A década dos noventa	18
1.3.1. Ámbitos de estudo na época	18
1.3.2. Áreas de maior impulso durante a década	20
1.3.3. Metodoloxía na investigación	21
1.4. Novo milenio, novas perspectivas	23
1.4.1. O Marco común europeo de referencia para as linguas (MECR).....	23
1.4.2. Asociacións europeas de avaliadores: ALTE, ILTE, EALTA	30
1.4.2.1. ALTE (<i>Association of Language Testers in Europe</i>).....	30
1.4.2.2. ILTA (<i>International Language Testing Association</i>)	33
1.4.2.3. EALTA (<i>European Association for Language Testing and Assessment</i>).....	35
1.4.3. Novos instrumentos para a avaliación: o Portfolio europeo das linguas	39
1.5. Liñas de investigación na área de avaliación no século XXI	41

CAPÍTULO 2

2. O Certificado de Lingua Galega (Celga)	45
2.1. Antecedentes	46
2.2. Descrición do Celga	48
2.2.1 Descrición xeral do nivel Celga	50
2.2.2. Descrición do nivel segundo as habilidades.....	51
2.2.2.1. <i>Comprensión oral</i>	51
2.2.2.2. <i>Comprensión escrita</i>	52
2.2.2.3. <i>Expresión oral</i>	53
2.2.2.4. <i>Expresión escrita</i>	55

2.2.3. Contido	55
2.2.3.1. <i>Lingua e comunicación: funcións comunicativas</i>	55
2.2.3.2. <i>Lingua e sistema: coñecemento do código</i>	57
2.3. Probas de dominio mediante tarefas	61
2.4. Definición de tarefas	61
2.5. Medidas acadadas para a implantación do sistema Celga	64
2.5.1. Medidas políticas: o Plan xeral de normalización da lingua galega	65
2.5.2 Medidas administrativas: Orde do 16 de xullo de 2007.....	70
2.5.3. Medidas metodolóxicas: formación para o profesorado, publicación de materiais didácticos e cursos preparatorios	75
2.5.3.1. <i>Materiais didácticos e cursos preparatorios</i>	76
2.5.3.2 <i>Ferramentas de axuda para os avaliadores: Manual de examinadores</i>	76
2.6. O modelo de exame	78
2.7. Sistema de corrección da proba escrita	81
2.7.1. Cualificación final da proba escrita.....	82
2.7.2. Control de calidade nas correccións dos exames	83
2.8. Criterios de avaliación da proba escrita	84
2.8.1. Adecuación	84
2.8.2. Coherencia e cohesión.....	85
2.8.3. Gramática	85
2.8.4. Léxico.....	85
3. Obxectivos do Celga	86
CAPÍTULO 3	
3.1. Estudo I	89
3.1.1. Xustificación.....	89
3.1.2. Obxectivos e hipóteses.....	91
3.1.2.1. <i>Utilidade dos criterios de corrección e avaliación</i>	93
3.1.2.2. <i>Influencia da experiencia dos avaliadores sobre a valoración dos criterios de corrección e avaliación</i>	93
3.1.2.3. <i>Influencia do nivel de lingua sobre a valoración do Celga</i>	94
3.1.2.4. <i>Equidade nas probas Celga</i>	94

3.1.3. Mostra de avaliadores do Celga	95
3.1.4. Medidas	99
3.1.4.1. <i>Variables clasificatorias</i>	99
3.1.4.2. <i>Criterios de corrección e avaliación das probas de expresión escrita</i>	101
3.1.4.3. <i>Criterios de corrección e avaliación das probas de expresión oral</i>	102
3.1.4.4. <i>Validación do construto</i>	103
3.1.4.5. <i>Equidade</i>	104
3.1.5. Instrumento	104
3.1.6. Procedemento	106
3.1.7. Análises estatísticas	107
3.1.8. Utilidade dos criterios de corrección e avaliación das probas de expresión escrita segundo os avaliadores.....	108
3.1.8.1. <i>Segundo o nivel do Celga</i>	110
3.1.8.2. <i>Segundo a experiencia dos avaliadores</i>	112
3.1.9. Utilidade dos criterios de corrección e avaliación das probas de expresión oral segundo os avaliadores	113
3.1.9.1. <i>Segundo o nivel do Celga</i>	114
3.1.9.2. <i>Segundo a experiencia profesional dos avaliadores</i>	116
3.1.10. Opinión sobre a validez dos criterios de corrección e avaliación.....	117
3.1.11. Valoración da equidade do Celga.....	127
3.1.12. Valoración da competencia comunicativa dos candidatos ao Celga ...	130
3.1.12.1. <i>Valoración da competencia comunicativa na proba escrita dos candidatos ao Celga</i>	130
3.1.12.2. <i>Valoración da competencia comunicativa na proba oral dos candidatos ao Celga</i>	132
3.2. Estudo II	135
Análise dos candidatos presentados ás probas do Celga durante os anos 2009 e 2010	135
3.2.1. Método da investigación.....	135
3.2.1.1. <i>Obxectivos</i>	135

3.2.1.2 <i>Hipóteses</i>	136
3.2.2. <i>Mostra</i>	137
3.2.3. <i>Medidas</i>	140
3.2.3.1. <i>Variables clasificatorias</i>	140
3.2.3.2. <i>Criterios de corrección na proba escrita</i>	141
3.2.3.3. <i>Criterios de corrección na proba oral</i>	143
3.2.4. <i>Procedemento</i>	144
3.2.4.1. <i>Obtención de datos para o estudo</i>	144
3.2.4.2. <i>Procedemento de avaliación das probas Celga</i>	144
3.2.5. <i>Análises estatísticas</i>	147
3.2.6. <i>Resultados segundo o perfil sociodemográfico dos candidatos</i> <i>ao Celga</i>	148
3.2.6.1. <i>Segundo a idade</i>	148
3.2.6.2. <i>Segundo o sexo</i>	152
3.2.6.3. <i>Segundo o lugar de nacemento</i>	154
3.2.6.4. <i>Segundo o lugar de residencia</i>	158
3.2.7. <i>Fiabilidade e validez nas probas Celga</i>	161
3.2.7.1. <i>Fiabilidade</i>	164
3.2.7.2. <i>Validez</i>	173
3.2.7.2. 1. <i>Validez de contido</i>	173
3.2.7.2. 2. <i>Validez de construto</i>	175
3.2.8. <i>Resultados da proba escrita</i>	178
3.2.8.1. <i>Relacións entre os criterios de corrección e avaliación</i> <i>da proba escrita</i>	179
3.2.8.2. <i>Cualificación na proba escrita dos non aptos</i>	182
3.2.8.3. <i>Cualificación na proba escrita dos aptos</i>	183
3.2.8.4. <i>Cualificación na proba escrita segundo o nivel do Celga</i>	186
3.2.8.4.1. <i>Cualificación na proba escrita dos non aptos nos niveis Celga 1</i> <i>e Celga 2</i>	187
3.2.8.4.2. <i>Cualificación na proba escrita dos non aptos nos niveis do Celga 3</i> <i>e Celga 4</i>	189
3.2.8.5. <i>Cualificación na proba escrita segundo o nivel de estudos</i>	190

3.2.8.6. Cualificación na proba escrita segundo o lugar de residencia.....	192
3.2.8.7. Cualificación na proba escrita segundo o sexo	193
3.2.8.8. Cualificacións na proba escrita segundo a idade	194
3.2.9. Resultados da proba oral.....	195
3.2.9.1. Relacións entre os criterios de corrección das probas orais.....	197
3.2.9.2. Cualificación na proba oral dos non aptos.....	199
3.2.9.3. Cualificación na proba oral dos aptos.....	200
3.2.9.4. Cualificación na proba oral segundo o nivel do Celga.....	202
3.2.9.5. Cualificación na proba oral segundo o nivel de estudos	205
3.2.9.6. Cualificación na proba oral segundo o lugar de residencia.....	206
3.2.9.7. Cualificación na proba oral segundo o sexo	207
3.2.9.8. Cualificación na proba oral segundo a idade	208
CAPÍTULO 4	
4.1. Conclusións Estudo I.....	213
4.1.1. Utilidade dos criterios de corrección e avaliación das probas de expresión escrita	213
4.1.1.1. Segundo o nivel do Celga.....	214
4.1.1.2. Segundo a experiencia dos avaliadores.....	214
4.1.2. Utilidade dos criterios de corrección e avaliación das probas de expresión oral segundo os avaliadores	214
4.1.2.1. Segundo o nivel do Celga.....	215
4.1.2.2. Segundo a experiencia profesional dos avaliadores.....	215
4.1.2.2.1. Na proba escrita	215
4.1.2.2.2. Na proba oral.....	216
4.1.3. Validez dos criterios de corrección e avaliación das probas segundo os avaliadores.....	216
4.1.3.1. Validez dos criterios de corrección e avaliación na proba escrita	217
4.1.3.2. Validez dos criterios de corrección e avaliación na proba oral	217
4.1.4. Valoración da equidade do Celga	217
4.1.5. Valoración da competencia comunicativa na proba escrita dos candidatos ao Celga	218
4.1.6. Valoración da competencia comunicativa na proba oral	

dos candidatos ao Celga	219
4.2. Conclusións Estudo II.....	221
4.2.1. Fiabilidade interxuíces	221
4.2.2. Consistencia interna.....	221
4.2.3. Validez de construto	222
4.2.4. Perfil sociodemográfico dos candidatos ao Celga en relación coa obtención da aptitude.....	223
4.2.4.1. <i>Cualificación na proba escrita segundo o nivel do Celga.....</i>	223
4.2.4.2. <i>Cualificación na proba oral segundo o nivel do Celga.....</i>	224
4.2.4.3. <i>Cualificación na proba oral dos non aptos.....</i>	224
4.2.4.4. <i>Cualificación na proba oral dos aptos.....</i>	225
4.2.5. Resultados da proba oral segundo as variables sociodemográficas	225
4.2.5.1. <i>Cualificación na proba oral segundo o nivel de estudos.....</i>	225
4.2.5.2. <i>Cualificación na proba oral segundo o lugar de residencia.....</i>	225
4.2.5.3. <i>Cualificación na proba oral segundo o sexo.....</i>	226
4.2.5.4. <i>Cualificacións na proba oral segundo a idade</i>	226
CAPÍTULO 5	
Discusións.....	229
5.1. Equilibrio no peso de todos os criterios de corrección e avaliación.....	229
5.2. Equidade das probas Celga	231
5.3. Sobre a selección dos avaliadores	233
5.4. Sobre a fiabilidade e a validez	234
5.5. Eivas detectadas polos avaliadores e propostas de mellora	236
5.6. Futuras liñas de estudo.....	238
CAPÍTULO 6	
Anexo: cuestionario	243
BIBLIOGRAFÍA.....	253

CAPÍTULO 1

Panorámica sobre a investigación en avaliación de linguas

1. Panorámica sobre a investigación en avaliación de linguas

A avaliación de linguas é un campo que experimentou un considerable desenvolvemento dende a década dos sesenta do século pasado ata os nosos días. Grazas aos diferentes estudos relacionados coa avaliación de linguas, podemos presentar un panorama histórico que aborde os distintos focos de atención sobre os que se centraron os estudosos. Así, veremos tanto os principios metodolóxicos que marcan as liñas de estudo e os avances no ensino, na aprendizaxe e na avaliación de linguas coma as propostas de orde política que transformaron a avaliación de linguas e a converteron nunha cuestión de estudo na área de lingüística aplicada.

Cando abordamos un asunto como é a avaliación de linguas, debemos delimitar o seu ámbito de acción. Neste caso, o presente traballo versará sobre a validez e a fiabilidade necesarias nun sistema de certificación oficial da competencia idiomática en galego, existente dende o ano 2006 e que se ve recollido nos certificados Celga expedidos pola Secretaría Xeral de Política Lingüística da Xunta de Galicia.

A avaliación de linguas espertou o interese da investigación en aprendizaxe de linguas por tratarse dunha cuestión que pode ser abordada dende diferentes puntos de vista, o cal permite afondar no coñecemento da adquisición dunha lingua e nos procesos aos que atinxe. Xunto a isto, cuestións coma a validez, a fiabilidade e a viabilidade das probas que permiten a certificación son asuntos que ocupan os estudosos da avaliación para conseguir unha mellora cualitativa tanto nos procesos coma no produto da avaliación de linguas. Recentemente, a incorporación das novas tecnoloxías permite facer uso de ferramentas para a investigación teórica e a súa aplicación práctica, o que revoluciona os sistemas de avaliación; concretamente a análise psicométrica permitiu a obtención de datos centrados máis no proceso de ensino ca

no resultado da instrución. Este tipo de avaliación coñécese como avaliación cualitativa e vén constituíndo nas últimas décadas unha fonte de información útil na mellora destas en aspectos como a fiabilidade e a validez, tanto para as institucións involucradas na certificación coma para os expertos en avaliación e realización de probas.

Con respecto a esta necesidade de mellora na calidade das certificación de linguas, as institucións dedicadas ao deseño e administración de exames a grande escala vense involucradas nuns procesos inxentes nos que proporcionar as ferramentas necesarias, válidas e fiables para que as súas probas sexan administradas por profesionais capacitados e formados na avaliación de linguas, co fin de que os exames sexan ferramentas de medición xustas para os candidatos que se presentan ás probas.

1.1. Antecedentes: anos sesenta e setenta

Tal e como sinala Bordón (2008), estas décadas están marcadas pola crise do modelo estruturalista, onde o que primaba era a obtención de resultados nunhas probas que revelasen o dominio do candidato sobre o código lingüístico e que non reflectían a súa habilidade para o uso da lingua. O que se procuraba nese tipo de probas pechadas, con predominio das de selección múltiple, era unha avaliación da actuación do candidato que permitise garantir a fiabilidade dos resultados cunha cualificación totalmente obxectiva grazas ás técnicas psicométricas, empregando unha metodoloxía baseada na investigación estatística para a avaliación das probas. Cabe salientar neste período os estudos que, sobre o ensino do inglés como lingua estranxeira, realizou Lado (1961), quen conseguiu que a avaliación de linguas se convertese nun asunto de interese de estudo para a lingüística aplicada, aínda que hoxe en día as súas teorías están amplamente superadas.

Un punto de inflexión neste período será, tal e como sinala Bachman (2000), a incorporación da habilidade lingüística en cada unha das catro destrezas (escoitar, falar, ler e escribir) e nunha serie de compoñentes (gramática, vocabulario e pronunciación) nos exames a grande escala e nos materiais preparatorios para probas da época. Obviamente, este tipo de exames demostra que as liñas de ensino de linguas neste período é o reflexo dunha etapa estruturalista que, a mediados dos anos setenta, xa mostraba claros sinais de crise e comezan a aparecer voces e traballos que buscan unha nova liña de ensino de linguas que desembocaría no modelo comunicativo. Neste momento publícase o *Threshold Level* (1975) e o *Nivel Umbral* (1979) na súa versión española (Slagter, 1979) e xa nestas datas o modelo de *Certificate of Proficiency in English* reflicte un novo rumbo nos tipos de probas, máis afastadas dos cortes puramente estruturalistas de probas pechadas de selección múltiple; así o novo *Certificate* constaría de cinco partes: composición, comprensión de lectura, uso do inglés, comprensión auditiva e entrevista. Weir (2002), deixando atrás un modelo que concibe a adquisición dunha lingua como unha competencia unitaria, inclínase pola habilidade lingüística do usuario, reflexo do uso das catro destrezas.

Xa nos últimos anos da década dos setenta e despois da realización do primeiro Language Testing Research Colloquium, que sitúa a área de avaliación de linguas como unha área de estudos propia da lingüística aplicada, márcanse as liñas de actuación que guiarán o ensino, a aprendizaxe e a avaliación de linguas na seguinte década, deixando atrás o modelo da habilidade lingüística como único elemento para o ensino e a avaliación das linguas. Neste momento cabe destacar os traballos de Widdowson (1978, 1979, 1983), Savignon (1972, 1983) e Canale e Swain (1981), que levarán a unha concepción do ensino de linguas centrada no seu uso e desencadeará unha revolución no ensino coñecida como método comunicativo.

1.2. Anos oitenta: inicios do método comunicativo

Nesta década, segundo sinala Bachman (2000), o modelo de ensino, aprendizaxe e avaliación de linguas estará centrado no proceso, na creación dun discurso propio, na negociación do significado e na importancia do contexto no que se desenvolve a situación comunicativa. Xunto a isto, o tipo de exames a grande escala tivo que aliñarse a un currículo de lingua para ter un modelo do cal partir. Isto provocou que se fixese unha revisión das tarefas dos exames, esixindo que estas fosen tarefas auténticas e non só actividades pechadas que avaliaban o coñecemento do código, coa finalidade de garantir a validez nos exames a grande escala que permitise facer inferencias sobre a habilidade de uso da lingua do candidato. Ademais, non se pode esquecer que o deseño dos cursos de aprendizaxe de linguas se viu influído polos exames a grande escala que os candidatos tiñan que superar. Isto provocou que tanto o ensino coma a aprendizaxe e a avaliación estaban na necesidade de abrir unha liña de traballo que permitise avanzar nunha década cara a unha nova posición, onde o candidato pasaba de ser un mero produtor de regras gramaticais e traducións a un produtor de textos orais e escritos adecuados a unha situación comunicativa.

Nesta década cabe sinalar a obra *New Directions in Language Testing* (Lee, Fok, Lord y Low, 1985), que recolle as presentacións do Internacional Symposium on Language Testing realizado en Hong Kong en 1983, onde os relatores articularon as súas propostas sobre catro áreas de interese: a avaliación e o currículo, os exames a grande escala, a avaliación da habilidade oral e, por último, a validación dos exames.

1.2.1. A avaliación e o currículo

Neste apartado son salientables as achegas de Davis (1985) e Rea (1985), ambos fan fincapé no tipo de lingua que se debe de ensinar na aula porque esta repercutirá no tipo de exame que terá que superar o candidato. Davis fai unha diferenza entre exames de aproveitamento e de dominio. A maioría dos exames seguen un modelo de ensino e aprendizaxe que levan a facer exames de aproveitamento, xa que queren comprobar se o candidato cumpriu os obxectivos do curso. Pola contra, os exames de dominio adoitan ser máis creativos, xa que os seus obxectivos e contidos poden influír no modelo de ensino ao que se verán sometidos os candidatos, xa que dan lugar a programas e cursos preparatorios para superar os exames. Para Davis, os exames de aproveitamento reflicten a concepción da lingua adoptada para o seu ensino na aula mentres que os de dominio a grande escala poden ser innovadores á hora de adoptar un modelo lingüístico que influirá no ensino.

Outro dos autores que deu luz na área da avaliación de linguas neste período foi Rea (1991), quen salientou a importancia da avaliación no proceso de ensino e aprendizaxe e as dificultades de avaliación do método comunicativo, debido á necesidade de consenso sobre a natureza da base teórica que debe sustentar os exames para garantir a validez do seu construto e do seu contido. Unha das preocupacións abordadas por Rea, propia da época, é a reflexión sobre o mantemento da dicotomía chomskiana de competencia e actuación na avaliación, e chega á conclusión de que calquera comportamento lingüístico constitúe unha mostra de actuación e, polo tanto, a competencia non pode ser avaliada directamente, senón que só a través da actuación se poden inferir niveis de competencia nos candidatos. Outros dos temas que lle preocupan a Rea son as tarefas que teñen que levar a cabo os candidatos nos exames, cando estas son actividades baseadas na análise de unidades discretas

fronte a outro tipo de exames máis integradores, para o cal fai unha proposta de dous trazos clave para a avaliación da lingua dentro do seu marco de ensino. Conclúe que a avaliación da lingua dende un método comunicativo é complexa e problemática; aínda que daquela o método comunicativo xa estaba asumido como válido para o ensino, non era así no caso da avaliación, xa que, como ben é sabido, resulta menos problemático avaliar a competencia lingüística a través de elementos discretos ca dende o propio uso da lingua.

1.2.2. Os exames a grande escala

Para relatar esta cuestión é fundamental a contribución de Swain (1985) sobre o construto dos exames a grande escala, debido á importancia e repercusión que poden ter tanto nos candidatos coma no método de ensino.

Sinala os seguintes principios que deben rexer o deseño e a construción dos exames:

a) Marco teórico: concibido como unha concepción da natureza da lingua para poder elaborar un exame de lingua, que no caso de Swain propón como modelo ideal o de competencia comunicativa, defendido xunto con Canale dende o comezo da década (Canale e Swain, 1980).

b) Concentrarse no contido: facendo referencia ao contido do material empregado, das actividades lingüísticas comunicativas e das tarefas que permitan xerar actuacións lingüísticas comunicativas. Para Swain (1985: 38), o contido debe caracterizarse por ser motivador, integrador e interactivo por natureza, para que lle resulte atractivo ao candidato.

c) Sesgar para obter unha boa finalidade: Swain propón que no exame o candidato debe sacar o mellor que ten de si para conseguir unha boa actuación. Para acadalo, suxire que o tempo sexa axeitado á tarefa, darlle

información ao candidato sobre o que se espera del e permitirlle o uso de ferramentas de axuda como dicionarios.

d) Atender á repercusión do exame (*washback*¹): isto é, estudar o efecto que o exame ten nas prácticas de ensino. Ademais, considera fundamental involucrar os profesores no proceso de deseño de exames, no protocolo de administración das probas e na súa cualificación, para que vaian en consonancia coa súa metodoloxía no ensino.

Outra cuestión que se tivo en conta nese momento e que aínda forma parte das investigacións sobre avaliación é a relativa á avaliación da expresión oral. Debido a que en todas as probas de dominio a grande escala unha parte da proba consiste na produción dun texto oral, detectouse que unha das maiores dificultades residía na avaliación das probas orais de xeito obxectivo, tamén se revelou como eiva a carencia de formación dos avaliadores para corrixir probas orais e escritas empregando os criterios establecidos. Esta cuestión continúa sendo, ata os nosos días, unha das maiores preocupacións nos sistemas de certificación a grande escala, aínda que, grazas ás achegas dos estudos en avaliación, se mellorou considerablemente.

Entre os temas recollidos no compendio de relatorios presentados no simposio de Hong Kong destaca o relativo á validez dos exames de lingua. Daquela o concepto de validez presentaba certos problemas de definición, así que Stevenson (1985) propón unha reflexión sobre o concepto de validez, sobre todo na validez aparente (*face validity*), entendida como unha proba na que unha actividade pode parecer unha tarefa da vida real, cando en realidade é unha simulación. Esas simulacións deben ser elixidas atendendo a uns

¹ Repercusión ou efecto de rebote, denomínase a influencia que poden exercer os exames no ensino, xa que, dependendo do modelo de exame empregado, a metodoloxía usada no ensino condicionará o modelo de exame. Un certificado baseado en tarefas, como é o Celga, require dun ensino da lingua baseado nun enfoque comunicativo cunha programación mediante tarefas.

principios teóricos e a súa cualificación precisa do mesmo sustento teórico para que se poidan facer inferencias válidas.

Polo tanto, podemos observar como na década dos oitenta, e segundo as presentacións recollidas na obra *New Directions in Language Testing* (Lee, Fok, Lord e Low, 1985) procedentes do Internacional Symposium on Language Testing realizado en Hong Kong, xa existía unha necesidade de establecer a competencia comunicativa como modelo de lingua e, no relativo á avaliación, as preocupacións xiraban arredor das cuestións da avaliación da expresión oral e a validez das probas, aínda latentes nos nosos días. A partir deste encontro, segundo Bachman (2000), a avaliación de linguas converteríase nun dos asuntos que cumpría estudar na área de lingüística aplicada.

1.3. A década dos noventa

1.3.1. Ámbitos de estudo na época

Esta década vén marcada por dúas liñas de traballo fundamentais nos estudos sobre adquisición de linguas; por unha banda, o método comunicativo que se adoptou como modelo para seguir no ensino de linguas e no deseño de probas de dominio e, pola outra, a competencia estratéxica ampliouse grazas aos estudos de Bachman e Palmer (1996), os cales consideran que as estratexias metacognitivas permiten establecer metas para saber que facer coa lingua, como facelo e planificar a aprendizaxe dunha lingua. Esta nova concepción influíu no construto teórico sobre ensinanza de linguas e no deseño de probas, xa que proporcionou un modelo de lingua comunicativo. Dos ámbitos que marcarían os estudos na década dos noventa destacamos o método comunicativo e a súa influencia na avaliación.

Mentres que a comezos dos anos noventa o método comunicativo xa era coñecido grazas os traballos de Canale e Swain e amplamente utilizado no ensino de linguas, non parecía que a súa influencia fose tan marcada na área da avaliación. Así, na obra editada por Charles Alderson e Brian North, *Language Testing in the 1990* (1991), que recolle os relatos expostos no simposio titulado A Avaliación da Lingua nos Anos Noventa: o Legado Comunicativo, organizado por IATEFL (*International Association of Teachers of English as a Foreign Language*), podemos atopar investigacións teóricas e prácticas sobre avaliación, ademais de modelos de avaliación que asentarían os estudos sobre este ámbito na década dos noventa.

Figuras como Carrol ou Shekan, fundamentais nos avances nesta liña investigadora e que continúan hoxe en día tendo relevancia, plasmaron as súas inquedanzas sobre temas como a resistencia aos novos modelos de exame por parte de profesores, institucións examinadoras, candidatos ou mesmo o medio académico (Carrol, 1991), que provocaron unha falta de permeabilidade para as innovacións no campo da avaliación. Para Skehan (1991) a falta de permeabilidade vén provocada polas diferentes teorías da natureza da lingua, xa que nas décadas anteriores a influencia da teoría da gramática universal de Chomsky, da gramática estrutural, ademais dos enfoques sobre os actos de fala e a concepción da lingua como discurso, provocou unha confusión naqueles que precisaban dun modelo teórico de lingua sobre o que sustentar o deseño das probas. Engade tamén que o protocolo de realización de deseño e administración de probas pode resultar moi enleado se está continuamente cambiando de modelo, xa que é necesario pilotar as probas para comprobar que se manteñen os principios de validez e fiabilidade, ademais do incremento en gastos que os ditos cambios supoñen.

Outra das cuestións tratadas no simposio e que continúa hoxe en día vixente é o uso das escalas como ferramentas para a avaliación de probas

comunicativas nas que o candidato ten que crear un texto oral ou escrito. Segundo Alderson (1991), este é un dos problemas aos que se teñen que enfrontar os avaliadores xa que, debido á necesidade de avaliar do xeito máis xusto posible unha mostra de actuación que non se pode medir tan obxectivamente coma unha proba pechada de resposta múltiple, é necesario que as escalas estean ben definidas e que sexan de axuda tanto para os avaliadores coma para os candidatos. Estas escalas serven para paliar a subxectividade inherente á cualificación porque achegan un estándar común que permite avaliar as producións máis subxectivas. Ademais, permítenlles aos candidatos saber que significa a nota que obtiveron só con ler a escala á que se corresponde a súa puntuación; por exemplo, se un candidato obtén un 3 nunha escala de 1 a 5, pode ler que é o que significa estar nese nivel no uso da lingua.

Finalmente, outros temas que foron causa de discusión no simposio foron os defendidos por Rea e Davis. Para Rea (1991) os problemas sobre avaliación neste momento residen na definición da gramática comunicativa e en se esta pode ser medida comunicativamente; defende que só pode ser medida cando o candidato pode crear significado nun contexto comunicativo que lle permita realizar un acto comunicativo. Non obstante, para Davis (1991) o foco de atención versa sobre a relación entre a avaliación e a lingüística aplicada, considera que no plano teórico debe existir maior relación entre avaliación e adquisición de linguas e no plano práctico debería existir unha correlación entre o modelo do ensino de linguas e a súa avaliación, para que vaian en sintonía e non ocorra que no ensino prime un modelo comunicativo e a avaliación presente trazos propios do modelo estruturalista.

1.3.2. Áreas de maior impulso durante a década

Bachman (2000) fai unha revisión dos ámbitos de estudo máis relevantes nesta década relacionados coa avaliación e sinala que case todos eles afondaron nas

relacións entre avaliación de linguas e outras áreas da lingüística aplicada, ademais dos avances obtidos na medición no contexto xeral da educación. Así, sinala os seguintes asuntos como aqueles que recibiron unha maior atención por parte dos estudosos: a metodoloxía para a investigación, os avances prácticos, os factores que afectan a actuación nos exames de lingua, a cualificación da actuación, a ética dos exames e a profesionalización no campo da avaliación. Trataremos cada un deles brevemente.

1.3.3. Metodoloxía na investigación

Neste ámbito destacan os avances conseguidos nas técnicas de avaliación cuantitativa. De acordo con Bachman, pasouse dunha medición con referencia á norma, na que o que se medía era a actuación dun candidato con respecto á actuación do resto dos candidatos, a unha medición da actuación realizada con referencia a un criterio, que permite interpretar os resultados dun exame en función dun dominio na habilidade lingüística (por exemplo, pódese medir a coherencia e a cohesión dun candidato nunha proba escrita). Xunto a isto, outros avances da década foron os estudos que afondaron na avaliación cuantitativa, nos que se estableceron conceptos como a teoría da xenerabilidade (*Generalizability theory*) e a teoría de resposta ao ítem (*Item response theory*), máis empregada na medición educativa, mentres que nos exames de lingua se adopta o modelo de Rasch.²

²**Teoría da xenerabilidade:** baseada na análise da varianza e pode estimar a fiabilidade dun grupo completo de avaliadores dunha vez.

Teoría da resposta ao ítem: sistema de medición que ten en conta tanto o candidato coma as características do ítem, baseada na teoría da probabilidade, mostra a probabilidade de que unha persoa acerte un ítem en particular. As notas dos candidatos e os totais de cada ítem transfórmanse nunha escala para que poidan relacionarse entre si.

Modelo de Rasch: baseado no enfoque estatístico proposto por Rasch (1906-1980), consiste nunha fórmula matemática aplicable a calquera disciplina para establecer a relación entre a probabilidade de éxito e a diferenza entre a habilidade dun individuo e a dificultade dun ítem.

Outro dos logros desta década foi a clasificación dos tipos de avaliación en función do que se quixese medir. Así, a **avaliación de probas de dominio** mide o nivel de competencia que ten un candidato nun determinado nivel de lingua –este tipo de avaliación tamén se denomina **avaliación cuantitativa**–, mentres que a **avaliación de probas de proceso** está enfocada a obter información do proceso de aprendizaxe dunha lingua, para o cal se empregan diferentes procedementos, como a análise do discurso, informes do profesor, diarios de aprendizaxe do alumno, cuestionarios ou portfolios (esta avaliación tamén se denominou **avaliación cualitativa**, avaliación auténtica, educativa, pedagóxica e alternativa). No desenvolvemento deste tipo de avaliación destacan as achegas de Brown e Hudson (1998), que prefiren empregar o termo alternativas á avaliación no canto de avaliación alternativa, para evitar as connotacións negativas que o termo puidese ter se se entende coma unha avaliación diferente do que ata agora se fixera, e que non implica o deseño de exames a toma de decisións en función dos resultados. Así, os autores consideran necesarios procedementos integrados onde se teñan en conta os datos cualitativos e cuantitativos para a avaliación de linguas.

Como reflexión sobre as melloras acadadas nesta década, Bachman (2000) destaca a aplicación da informática á avaliación, o perfeccionamento de técnicas de medición cuantitativas, a proliferación no uso de ordenadores persoais e a creación de programas de fácil accesibilidade para os usuarios.

Xa dende finais dos oitenta esta cuestión tiña defensores como Alderson (1986), que vía as posibilidades para mellorar e aumentar a interacción entre a actuación e o candidato mediante as posibles avaliacións da actuación do candidato, así como a avaliación do contido dos ítems e do exame. Deste xeito, permite coa súa análise a posibilidade de obter unha reflexión sobre o proceso de aprendizaxe e o produto obtido; ademais o uso dun programa informático aplicado á avaliación e o desenvolvemento de teorías sobre o que debe ser un

exame de diagnóstico permitirá mellorar o proceso de aprendizaxe dunha lingua. Porén, tamén advirte da limitación na obtención de datos sobre o candidato ao usaren unicamente ítems de selección múltiple.

Finalmente, Bachman tamén se refire aos factores que afectan o candidato á hora de realizar exames de lingua e revisa os aspectos sobre os que se centraron as investigacións neste ámbito. Considera que hai tres aspectos que cómpre destacar:

— As características do procedemento do exame e dos cualificadores (Chalhoub-Deville, 1996 e McNamara, 1996), que non sempre son fiables, porque se atoparon desaxustes nas cualificacións en función da lingua materna do avaliador, así como da súa experiencia e formación previa.

— Os procesos e estratexias adoptados polos candidatos ao responderen ás tarefas do exame, para o cal se empregaron cuestionarios (Púrpura, 1999) e métodos introspectivos (Wu, 1998).

— As características dos propios candidatos, prestando atención a datos censuais, como a formación académica, o sexo, a idade, a lingua materna ou a cultura, que poden repercutir na actuación do candidato no exame (Kunnan, 1995).

1.4. Novo milenio, novas perspectivas

1.4.1. O Marco común europeo de referencia para as linguas (MECR)

Sen dúbida ningunha, o punto de inflexión no ensino, aprendizaxe e avaliación de linguas no contexto europeo sería a publicación no ano 2001 do *Common European Framework of Reference for Languages: learning, teaching, assessment*, que constitúe un dos principais compoñentes da segunda fase dun proxecto do Consello de Europa de maior alcance titulado *Language Learning*

for European Citizenship (1989-1996). Foi traducido ao galego pola Xunta de Galicia no 2005 como Marco europeo común de referencia para as linguas: aprendizaxe, ensino, avaliación (en diante, MECR).

O documento foi desenvolvido a través dun proceso de investigación e de amplas consultas entre institucións e expertos. Constitúe unha ferramenta práctica para establecer obxectivos por niveis nas sucesivas fases da aprendizaxe da lingua e permite, ao tempo, a avaliación dos resultados de xeito que se poidan comparar nun contexto internacional. Asemade, constitúe unha base para o recoñecemento mutuo das certificacións de competencia lingüística para lle facilitar á cidadanía a mobilidade educativa e laboral na Unión Europea. O seu uso está estendido aos sistemas regrados de ensino e ás asociacións internacionais co obxectivo de comparar certificacións.

A vantaxe do MECR é que se trata dun documento que describe as competencias necesarias para a comunicación, os coñecementos e as destrezas relacionadas coas ditas competencias e as situacións e dominios nos que ten lugar a comunicación. Polo tanto, convértese nun pilar na construción dos currículos de linguas, dos cursos de idiomas e da maquinaria editorial que fornece materiais didácticos para a aprendizaxe de linguas, ademais de todas aqueles incluídos directamente no ensino e na avaliación, dende profesores, avaliadores e formadores de profesores ata institucións certificadoras. É, polo tanto, un marco que proporciona definicións claras sobre obxectivos e métodos de aprendizaxe e de ensino, ademais de proporcionar as ferramentas necesarias para avaliar a competencia dos candidatos.

Para Parrondo (Parrondo, 2005), o MECR ten unha influencia apreciable no xeito en que as institucións certificadoras de linguas describen os seus exames e os aspectos relacionados coa especificación dos dominios e das actividades comunicativas. A competencia comunicativa maniféstase mediante actividades lingüísticas encadradas nalgún dos catro dominios que determina o

MECR –persoal, público, laboral e educativo– e as actividades comunicativas presentan a seguinte posibilidade de variación: produtivas, receptivas, interactivas e mediadoras. Enténdense por actividades produtivas de carácter oral ou escrito tarefas variadas, que poden ir dende dirixirse a unha audiencia ata cantar en público, escribir un poema ou facer unha reclamación por escrito. As tarefas receptivas (escoitar e ler) conteñen fins específicos para esas destrezas, como ler ou escoitar, para acadar unha comprensión global, para procurar unha información específica ou para unha comprensión detallada. Por outro lado, as actividades denominadas interactivas poden ser escritas ou faladas, mentres que as tarefas de mediación consisten na tradución, interpretación, síntese ou perífrases e manifestan a capacidade de facilitar a comunicación entre as persoas.

Unha reflexión interesante sobre a importancia do MECR é a que achega Figueras (2008) facendo fincapé nos logros que levaron ao triunfo do documento como referente de uso en todo Europa. Indo máis alá da polémica suscitada polos seus detractores, para a autora o éxito do MECR radica en tres cuestións:

1. A institución que o publica é o Consello de Europa, entidade que conta coa solvencia e o rigor suficiente no ámbito europeo como para considerar o documento como o produto dun proceso de investigación serio e prolongado no tempo, que partira xa dende o ano 1971 coa publicación do *Nivel Umbral* e que nesta altura estaba dispoñible en 25 linguas europeas.

2. Ten a súa base na estratexia política educativa e didáctica empregada polo Consello de Europa para xustificar a necesidade dun marco de recoñecemento paneuropeo dos esforzos realizados nos diferentes países nos ámbitos de docencia, aprendizaxe e avaliación de linguas.

3. De tipo académico-pedagógica, está fundamentada na necesidade dun enfoque orientado á acción, tal e como recolle o MECR, considerando a lingua como instrumento de comunicación en sentido amplo. Este enfoque foi recollido e adoptado pola maioría das persoas e institucións relacionadas co ensino, aprendizaxe e avaliación de linguas en Europa.

A pesar da boa acollida e do éxito na súa implantación no ámbito do ensino e avaliación, existe unha polémica que lle afecta ao MECR en relación coa avaliación certificativa, xa que, malia todas as institucións e entes certificadores de linguas acolleren axiña o MECR como piar para o deseño das súas probas, non existen estudos de validación que os avalen como probas relacionadas co MECR. Así, no Congreso Internacional de ALTE no 2001 en Barcelona, considérase necesaria unha investigación nesta liña que atopa apoio por parte da comunidade de especialistas en avaliación e dá lugar á realización dun seminario organizado polo Ministerio de Educación finés en xullo de 2002 en Helsinqui, a partir do cal o Consello de Europa lle encarga a un equipo de especialistas en avaliación a redacción dun manual para relacionar exames cos niveis do MECR (a edición piloto de 2003, *Preliminary version of a Manual for relating examinations to the CEFR*, que sería substituída pola versión final no 2008 e que, posteriormente, se completaría coa edición de ALTE no 2011 do *Manual for Language Test Development and Examining for use with the CEFR*).

O seguinte paso para emendar esta eiva do Marco sería unha enquisa sobre o seu uso en Europa, na que se detectaron os seguintes problemas: a natureza aberta, flexible e non específica para todas as lingua provoca unha indefinición e a necesidade dun documento máis concreto que facilite os labores dos especialistas en avaliación. A isto cómpre engadirle o uso simplista que se ten feito do MECR nos materiais didácticos, deseños curriculares e probas, xa que se tomou só a dimensión vertical do documento

(descritores de dominio de lingua en seis niveis de competencia) en detrimento do uso da dimensión horizontal (que establece a descrición das categorías de aprendizaxe de idiomas e do uso da lingua). Outra das cuestións que saíron á luz coa publicación do *Manual for relating examinations* está relacionada coa case inexistente tradición avaliadora en Europa, que provoca reaccións opostas pola dificultade atopada ao empregar un documento e uns niveis de referencia pensados para a aprendizaxe de segundas linguas a adultos e a súa avaliación. Estes debates puxeron de manifesto a rixidez cara a novas vías de actuación no ámbito da avaliación e do ensino de linguas, que xa se manifestaron na dificultade de implantación de novas ferramentas na avaliación, como é o Portfolio europeo das linguas (PEL). Para facer un seguimento das liñas de actuación do Consello de Europa e da lista de actividades propostas para a mellora neste ámbito, pódense consultar na web do Consello de Europa: http://www.coe.int/t/dg4/linguistic/Conference_EN.asp

Non obstante, cómpre facer unha revisión das achegas que o MECR vén implantado nos ámbitos do ensino, aprendizaxe e avaliación, dende a súa publicación no 2001, ao longo da primeira década do novo milenio. Así, no eido da docencia, as melloras que se produciron presentan unha dobre vertente, a da formación do profesorado e a relativa ao deseño curricular e á programación.

Con respecto á formación do profesorado, o MECR ten unha función revulsiva, xa que, grazas á súa aparición, se fixo unha valoración sobre que se entende por ensinar, por lingua ou como se pode ensinar e aprender. Ademais proporciona novas propostas no ensino, xa que no seu apartado 2.1 presenta o enfoque orientado á acción que terá repercusións na organización e na programación de aula. Finalmente, o MECR converteuse no centro de atención de congresos e publicacións que permitiron mellorar a formación do profesorado nesta década.

Na planificación dos currículos e programacións de aula permitiu a profesionalización no campo da avaliación, esixindo concrecións detalladas para as distintas linguas e o deseño de probas, o que provoca problemas se estas probas deben estar aliñadas co MECR, xa que non todas as linguas conseguen especificacións para todos os seus niveis de ensino (regrado, de adultos, de estranxeiros). Pero o que é obvio é que existe unha presión para avaliar con relación ao MECR que leva a cambios estruturais tanto na aula coma nos materiais empregados para o ensino e que aínda hoxe non están totalmente equilibrados xa que non seguen a mesma liña didáctica.

Con respecto á aprendizaxe de linguas e aos beneficios que o MECR produciu neste ámbito, non hai aínda estudos que poidan achegar datos fiables sobre o impacto na aprendizaxe de linguas, pero si se percibiu unha ampla acollida por parte do alumnado e do profesorado de linguas do MECR e a súa realización máis tanxible na aula que é o Portfolio europeo das linguas (PEL).

No ámbito da avaliación, non obstante, si existen datos obxectivos sobre a ligazón entre o MECR e as probas de dominio. Xurdiron múltiples congresos, estudos e proxectos europeos que sinalan cales son as dificultades desta conxunción entre o Marco e os exames. De aquí nace a necesidade paneuropea dunha especialización profesional para as probas de dominio que abranguen todas as fases da súa elaboración, dende o deseño das probas, a súa elaboración, a administración en convocatorias de exame e a validación mediante estudos psicométricos. Isto mudaría o ámbito da avaliación e da administración de probas que ata a publicación do MECR estaba en mans dunha única asociación de examinadores de linguas europeas (ALTE). Debido á necesidade de acreditacións de coñecementos de linguas, os ministerios dos distintos países europeos concienciáronse da necesidade dunha especialización profesional no ámbito da avaliación e así creáronse novas asociacións de examinadores como EALTA (*European Association for Language Testing and*

Assessment, <http://www.ealta.eu.org>), que permite o intercambio de profesionais da avaliación de linguas europeas.

Finalmente, como sinala Figueras (2008), a materia pendente do MECR con respecto ao ensino, a aprendizaxe e a avaliación de linguas vén determinada por tres cuestións nas que aínda existen ocos de información sobre os que no futuro deben centrarse os estudos. En primeiro lugar, sinala como eiva do MECR a necesidade de distinguir entre linguas propias e linguas estranxeiras e entre as distintas etapas no desenvolvemento dos alumnos de distintas idades, así como nos distintos enfoques e categorizacións que cómpre ter presentes no deseño dos plans de estudo para o alumnado adulto. En segundo lugar, achácaselle ao MECR a centralización e favorecemento do inglés como lingua de estudo e de uso en Europa. Fronte a isto, o ideal sería o estudo e o uso doutras linguas europeas en pé de igualdade co inglés; debido a que as políticas educativas dos países europeos e as da Unión Europea parece que non lle prestan atención. No 2003 Michael Byram e Jean Claude Beacco publicaron a *Guide for the Development of Language education Policies in Europe, from languages diversity to plurilingual education*, editada polo Consello de Europa, na que se recolle a necesidade de favorecer o multilingüismo na procura de formación de cidadáns plurilingües. Xa por último, sinala Figueras, hai que resolver o equilibrio entre as políticas a longo prazo, como as do MECR, e aquelas que pretenden solucionar as necesidades máis inmediatas. O reto para a política lingüística europea está en equilibrar esas dúas vertentes, a partir dun traballo conxunto entre os ministerios de educación dos estados membros, o Consello de Europa e directrices europeas como o proxecto Surveylang.³

³ Surveylang é un consorcio europeo composto por diferentes entidades especialistas en desenvolvemento de probas e recollidas de mostras e datos, así como na psicoloxía cognitiva, a medición educativa e a elaboración de escalas e análise de datos. A Comisión Europea encomendoulle a SurveyLang a realización da enquisa europea de competencias lingüísticas.

1.4.2. Asociacións europeas de avaliadores: ALTE, ILTE, EALTA

Como xa adiantamos en seccións anteriores, o avance máis significativo no ámbito da avaliación vén dado nestas dúas últimas décadas pola profesionalización das persoas e institucións incluídas nas tarefas de avaliación en probas de dominio e a grande escala. Tres asociacións, ALTE, ILTE e EALTA, xorden no panorama europeo da avaliación; atendendo á súa orde cronolóxica de fundación, veremos cales son os seus fundamentos e a que se dedican, para o cal seguimos o estudo de Huhta (2007), no que describe a cooperación internacional na área da avaliación e das probas de segundas linguas e linguas estranxeiras en Europa, os Estados Unidos e Asia. Tales organizacións, con carácter profesional e internacional, promoven a avaliación e a certificación de linguas e afondan en cuestións relacionadas co dito ámbito para ofrecerlles ao profesorado e aos profesionais as ferramentas necesarias para avaliar competencias de lingua. As siglas ás que fan referencia os nomes das asociacións son ALTE (Asociación de Avaliadores de Lingua en Europa), ILTA (Asociación Internacional de Probas de Lingua) e EALTA (Asociación Europea de Avaliación e Probas de Lingua).

1.4.2.1. ALTE (Association of Language Testers in Europe)

É a máis antiga das tres organizacións, fundada en 1990. Diferénciase das outras en que só poden ser membros da asociación as institucións responsables da certificación das diferentes linguas de ámbito unicamente europeo; conta nas súas bases con 31 membros, representantes de 26 linguas europeas e con 40 institucións afiliadas. Dende o punto de vista xurídico, a asociación está constituída como Grupo Europeo de Interese Económico, polo que a participa-

ción nela de representantes de países que aínda non pertencen á Unión Europea debe articularse por medio da figura de membro asociado.

Unha das principais tarefas da asociación é a descrición e comparación dos diferentes exames e chegar a confeccionar un marco común de carácter referencial que permita a visualización das equivalencias.

Os principais obxectivos da asociación son:

— Establecer uns niveis comúns de competencia para promover o recoñecemento transnacional da certificación en Europa.

— Establecer uns criterios comúns para todos os estadios do proceso de avaliación de linguas, para o desenvolvemento de probas, tarefas de escritura, administración, puntuacións e cualificacións, informe de resultados, análises das probas e informes das conclusións.

— Colaborar en proxectos conxuntos e no intercambio de ideas e habilidades.

Entre os proxectos da asociación cabe destacar o proxecto de 1998 CAN-DO, que supón a validación das escalas e certificacións de ALTE co MECR nun proxecto a longo prazo, xa que aínda está vixente e orientado á validación dos descritores de capacidade lingüística dos niveis de ALTE (Acceso 1, 2, 3, 4 e 5) que se definen por medio dunha serie de declaracións de habilidade, correspondentes ás tarefas que o candidato é quen de realizar en situacións comunicativas con relación á escala de niveis. O sistema de descricións consta dunhas 400 declaracións de habilidade, subdivididas en 40 categorías que describen o que un falante típico de segunda lingua (L2) pode facer nun idioma determinado, a un determinado nivel e en cada unha das destrezas (compresión e expresión oral e escrita). Este sistema está traducido a 13 idiomas europeos e foi desenvolvido a través dun rigoroso proceso de

selección de descritores, no que participaron expertos en avaliación de diferentes países europeos e máis de 10.000 estudantes de idiomas de toda Europa. O protocolo de actuación é o seguinte: adminístraselles aos candidatos ás probas dun exame oficial un cuestionario de autoavaliación do que se obtén unha serie de datos que son cruzados cos obtidos polos candidatos nos exames oficiais, procedemento polo que se validan a correlación entre as ditas certificacións e o MERC. Ata agora, os resultados obtidos entre o cuestionario e os exames oficiais son moi estables, como se pode comprobar no seguinte sitio: <http://www.dialang.org>.

Unha das cuestións que se tiveron que solucionar ao realizar este tipo de proxectos paneuropeos foi a de acordar unha terminoloxía técnica específica de contido inequívoco que servise para facilitar os traballos conxuntos entre os membros. Así, en 1998 publicouse o glosario multilingüe de termos relacionados coa avaliación da competencia lingüística que vén empregándose como ferramenta posibilitadora de cooperación. No ano 2008 publicouno en galego a Xunta de Galicia co título *Glosario de termos para a avaliación de linguas*.

No ano 1994, a asociación ALTE decidiu adoptar un código de boa práctica para definir os niveis mínimos de calidade. Os aspectos fundamentais sobre os que incide o código e para quen os membros se comprometen a salvagardar os dereitos do usuario son:

- Desenvolvemento e confección das probas de exame.
- Interpretación dos resultados das probas.
- Compromiso de equidade e xustiza.
- Información para os usuarios.

No ano 2000, a asociación colaborou con EQUALS (*European Association of Quality Language Services*) para desenvolver o Portfolio europeo das linguas, que foi validado polo Consello de Europa no 2001 e describiremos máis adiante.

Ademais das citadas liñas de investigación e produtos obtidos pola asociación, na actualidade conta con diferentes proxectos que desenvolven grupos de traballo entre os que destacan a certificación e acreditación de profesores, a certificación da competencia lingüística con fins específicos e a colaboración co Consello de Europa no desenvolvemento de descritores do nivel A1.

As reunións de ALTE realízanse dúas veces por ano en cidades europeas onde se falan as linguas das institucións pertencentes á asociación. Nestes encontros establécense reunións dirixidas a asuntos de interese para os elaboradores de exames e obradoiros a cargo de expertos de prestixio internacional de acceso restrinxido aos membros da asociación. As reunións finalizan cunha conferencia de libre acceso na que se inclúen presentacións por parte do país anfitrión e doutras institucións membros de ALTE. Dende finais dos anos noventa tamén organiza conferencias específicas sobre probas de avaliación que teñen lugar cada tres anos.

Aínda que se trata da asociación máis restritiva para a afiliación, na web da organización están á disposición de todas aquelas persoas e institucións interesadas neste ámbito materiais de axuda que permitan mellorar as probas de dominio, como se pode ver na páxina <http://www.alte.org>

1.4.2.2. ILTA (International Language Testing Association)

ILTA é unha asociación internacional de probas de lingua que vén operando nos Estados Unidos dende mediados dos anos oitenta. Os seus obxectivos principais son:

- Estimular o desenvolvemento profesional mediante a realización de obradoiros e conferencias.
- Promover a publicación e difusión de información relacionada coa área das probas de lingua.
- Desenvolver a área das probas de lingua e formar profesionais que sexan quen de liderar a dita área.
- Proporcionarlles servizos profesionais aos seus membros.
- Incrementar o recoñecemento e o apoio á avaliación de linguas como profesión.
- Estimular o orgullo profesional entre os seus membros.
- Recoñecer a relevancia dos logros profesionais dos seus membros.
- Cooperar con outros grupos interesados nos estudos estatísticos e en lingüística aplicada.

O Language Testing Research Colloquium (en diante LRTC, que pode traducirse como o Coloquio de Investigación de Probas de Lingua) é a conferencia oficial de ILTA que centra a súa atención nos exames e nas probas de lingua a grande escala, deixando a un lado as avaliacións de aula. Isto é debido a que consideran que estas probas de dominio teñen un alto impacto, os resultados teñen consecuencias relevantes para os candidatos e, polo tanto, cómpre que sexan válidos, fiables e imparciais. Estas calidades non poden darse por supostas, senón que deben ser comprobadas mediante estudos que as verifiquen. Os resultados obtidos son máis interesantes para as entidades avaliadoras ca, por exemplo, o impacto da avaliación na aula no ensino regrado en primaria ou secundaria que, en teoría, se os resultados fosen incorrectos, poderían ser corrixidos polo profesor. Esta é a razón pola que as probas a grande escala é a área de maior investigación de LRTC, pero non a única; é importante sinalar as achegas feitas no ámbito da avaliación diagnóstica, nova

área de investigación que xurdiu a comezos do século XXI grazas ao proxecto DIALANG. Segundo Alderson (2005) e Huhta (2007), en realidade é moi pouco o que se sabe sobre como se desenvolve a competencia nunha segunda lingua e sobre que elementos lingüísticos ou de calquera outra índole son propios desa competencia nun determinado nivel. Se se soubese máis sobre ese proceso, existirían materiais didácticos, plans curriculares e probas de diagnóstico que reflectirían mellor a maneira que temos os humanos de aprender unha lingua.

Finalmente, outro dos temas das conferencias de ILTA atinxe a consideración do contexto social e das consecuencias sociais dos diferentes tipos de avaliacións. Dedicaron un simposio ás avaliacións de linguas dirixidas á obtención do status de cidadanía para os inmigrantes, facendo fincapé na importancia dos contextos socioculturais e dos aspectos cognitivos na elaboración das probas, así como da avaliación da calidade e a utilidade destas. Pódense consultar estas cuestións na páxina <http://iltaonline.com>

1.4.2.3. EALTA (European Association for Language Testing and Assessment)

Esta asociación europea de avaliación e probas de lingua foi creada en 2004 e medrou rapidamente debido á posibilidade de converter en membro asociado a todas as persoas interesadas na avaliación, fronte a outras asociacións nas que a filiación non é persoal senón institucional. A asociación xurdiu grazas a un proxecto financiado pola Unión Europea entre os anos 2002 e 2004 e aspira a achegar as probas de lingua e a avaliación ao maior número de interesados.

O principio fundacional de EALTA ten como finalidade fomentar a comprensión dos principios teóricos que rexen o desenvolvemento de probas e das tarefas de avaliación de linguas, así como mellorar e compartir as prácticas

avaliativas en toda Europa. Para isto pretende levar a cabo as seguintes accións:

- Compartir a experiencia profesional no campo da avaliación e probas de lingua.
- Incrementar a comprensión da avaliación e as probas de lingua.
- Promover o recoñecemento da avaliación e probas de lingua como disciplina profesional en Europa.
- Mellorar a práctica dos sistemas de avaliación e probas de lingua en Europa.
- Promover a formación en probas de lingua e avaliación.
- Formar expertos en probas de lingua e avaliación para a resolución dos problemas de avaliación.
- Establecer vínculos con outros grupos interesados na avaliación e nas probas de lingua.
- Participar en actividades para a mellora das probas de lingua e avaliación en Europa.

As reunións de EALTA realízanse unha vez por ano (na primavera) e, se observamos as temáticas dos relatorios, podemos ver que o foco de atención se centra no profesorado e nas avaliacións de aula máis do que nas probas de dominio de grande escala. Algúns dos títulos representativos serían “Prácticas de avaliación en primeiras e segundas linguas nas escolas”, “Avaliamos o que ensinamos, pero adquiren os aprendentes o que ensinamos?” Xunto a estes temas, tamén destaca a vinculación das probas e o MECR, tal e como vemos nos títulos dos relatorios: “O reto da construción dun sistema válido para a avaliación da expresión escrita baseada no MECR”, “Procedementos para

relacionar as probas co MECR”, “Definición de boas prácticas para relacionar os exames co MERC”.

Outro dos temas recorrentes nas conferencias das tres asociacións é o desenvolvemento de novas probas con obxectivos diferentes. Estes obradoiros son moi útiles para os deseñadores de probas e permiten partillar problemas e solucións de carácter teórico e práctico comúns aos novos instrumentos de avaliación.

Código de boas prácticas

O feito de que as tres asociacións, ILTA, ALTE e EALTA, establecesen nas súas bases fundacionais un código é un índice de madurez científica e profesional que garante uns principios básicos de calidade nos que todas aquelas persoas e institucións involucradas na avaliación de linguas e no deseño de probas debe asumir como propias para manter os estándares de calidade no seu labor. Mentres que ILTA tivo moitas diferenzas para establecer un código deontolóxico internacional, tanto ALTE coma EALTA presentan nas súas respectivas páxinas electrónicas os seus códigos de boas prácticas. Sinalamos que o código de EALTA se ofrece en 32 linguas europeas, entre elas, o galego. O documento defínese como un código tripartito de deontoloxía profesional que (...) “Como consecuencia do desexo de EALTA de ser unha asociación aberta que acolla profesionais de diversas procedencias, as súas directrices elaboráronse tendo en consideración tres tipos de posibles membros que estean implicados en: (a) a formación de profesores en avaliación, (b) a avaliación na aula, (c) o desenvolvemento de probas en organismos ou centros avaliadores públicos ou privados. Hai unha serie de principios xerais aplicables a estes tres grupos: o respecto polos estudantes/examinandos, a responsabilidade, a equidade, a fiabilidade, a validez e a colaboración coas

partes implicadas”. <http://www.ealta.eu.org/documents/archive/guidelines/-Galician.pdf>.

Finalmente, Huhta (2007) sinala a importancia das tres asociacións e destaca, en primeiro lugar, o espertar da conciencia do que implica a avaliación, con que instrumentos traballa e que se entende por calidade na avaliación. Considera que o profesorado de linguas, os formadores do profesorado e os deseñadores de probas deben concienciarse na importancia da avaliación e un xeito de facelo é organizando conferencias e eventos de formación, aínda que é necesario contar con diferentes métodos para influír nos encargados da toma de decisións dos órganos que expiden as certificacións de linguas en beneficio dunhas probas que sexan fiables, válidas e viables. Outro dos efectos positivos da existencia das asociacións é a influencia que poden ter nos formadores de profesorado, nos deseñadores de probas e nas xuntas de exames xa que, unha vez que toman conciencia da importancia do seu labor e da competencia necesaria para levar a cabo avaliacións de lingua, poden, pola súa vez, difundir os seus coñecementos entre os profesorado de linguas nos seus países ou rexións. Malia parecer un proceso moi lento na difusión deste coñecemento, é esencial para a constitución dun sistema de avaliación de calidade que implique a supervisión constante das distintas fases que compoñen o sistema de certificación.

Xunto a esta concienciación da importancia do sistema de certificación, as asociacións poden constituír foros para o intercambio do coñecemento sobre a área de avaliación, organización de obradoiros e cursos de formación, ademais de recursos que poden axudar a mellorar as probas e avaliacións, tales como os códigos de boas prácticas ou as listaxes para comprobar a calidade do traballo realizado que ofrece ALTE en todas as fases do uso e deseño de probas (denominadas *checklist*: http://www.alte.org/quality_assurance/code/-checklist.php).

EALTA ofrece libremente as presentacións das súas conferencias e ILTA ofrece os resúmenes das conferencias de LTRC. Son, polo tanto, ferramentas que poden axudar a todas as persoas involucradas nas tarefas de avaliación para a mellora do seu labor.

1.4.3. Novos instrumentos para a avaliación: o Portfolio europeo das linguas

Unha das propostas do Consello de Europa que maior repercusión tivo na última década no eido do ensino, aprendizaxe e avaliación foi a publicación do MECR e a segunda foi a creación do Portfolio europeo das linguas (PEL) como posta en práctica do MECR.

O PEL é un proxecto desenvolvido e pilotado polo departamento de Política Lingüística do Consello de Europa dende 1998 ata o ano 2000 e foi no 2001, Ano Europeo das Linguas, proposto de forma xeral para toda Europa. Estes dous proxectos (o PEL e o MECR) teñen tamén como obxectivos principais:

- Protexer e desenvolver a herdanza cultural e a diversidade de Europa como fonte de enriquecemento mutuo.
- Facilitar a mobilidade entre a cidadanía, así como o intercambio de ideas.
- Desenvolver un enfoque no ensino de linguas baseado en principios comúns.
- Promover o plurilingüismo.

A configuración física do PEL consta dunha carpeta dividida en tres módulos:

1. **Un pasaporte lingüístico:** é o documento onde se recollen as acreditacións e competencias lingüísticas (oficiais e non oficiais) dunha maneira recoñecida internacionalmente.

2. **Unha biografía lingüística e cultural:** na que os estudantes describen os seus coñecementos lingüísticos e experiencias no ámbito da aprendizaxe de linguas. Para iso proporciónanselles uns descritores de competencia lingüística baseados na táboa de autoavaliación do MECR.

3. **Un dossier:** no que se poden incluír mostras do progreso e desenvolvemento da aprendizaxe (certificados oficiais ou documentos de toda índole relacionados coa aprendizaxe de linguas).

Como vimos ata agora, dende a década dos noventa existe a necesidade de abrir novas liñas na investigación do ensino, aprendizaxe e avaliación de linguas. Neste eido, non cabe dúbida de que a achega do PEL é unha das que maiores resultados obtivo referida aos procedementos de avaliación da produción oral e escrita dos candidatos, xunto coas entrevistas orais. Mentres que o PEL está especializado como ferramenta para a obtención de datos cualitativos no proceso de aprendizaxe, a entrevista oral emprégase como mecanismo para a obtención de datos tamén cualitativos como un produto no momento da avaliación, en todo tipo de exames, de dominio e de aproveitamento. Polo tanto, o PEL está máis centrado no proceso de aprendizaxe e serve como ferramenta compiladora de evidencias dese proceso, mentres que a entrevista é un produto dese proceso de aprendizaxe.

A importancia do PEL vén dada pola repercusión que ten na cidadanía que decide aprender outras linguas, porque mentres que o MECR é un documento de uso case exclusivo para os profesionais do ensino e avaliación de linguas, o PEL é un documento que pertence e usa a persoa que está aprendendo unha lingua. Aínda que este documento está deseñado para ser empregado nos sistemas educativos oficiais ou para a aprendizaxe indepen-

dente, a súa versatilidade reside na capacidade de almacenamento de documentos debidamente cubertos que calquera usuario pode consultar (para traslados de expedientes académicos, solicitudes dun posto de traballo ou cursos de idiomas). Pódese comprobar na páxina <http://www.culture2.coe.int/portfolio>.

1.5. Liñas de investigación na área de avaliación no século XXI

Neste percorrido polas innovacións e estudos na área da avaliación de linguas, destacamos a visión de Bachman (2000), premonitoria a comezos de década e que continúa vixente nesta área de estudos. Considera que a avaliación ten que se converter nunha profesión que implique a formación específica dos avaliadores e a definición de estándares de práctica, ademais de sinalar a necesidade de establecer o alcance da validez e da equidade das probas, elementos necesarios para que unha proba sexa xusta.

Dentro da lingüística aplicada, a avaliación de linguas é un campo que vén desenvolvéndose nos derradeiros trinta anos, cuxas áreas de interese implican a existencia de determinados asuntos que son recorrentes e que poden ser abordados dende diferentes ángulos, tales como a necesidade de determinar a natureza da lingua e contar cun modelo para a súa avaliación, os requisitos necesarios de avaliación de validez e fiabilidade nos exames, a avaliación da expresión oral, a busca de alternativas aos exames para obter información sobre os procesos de aprendizaxe dunha segunda lingua, a aplicación do desenvolvemento da tecnoloxía informática ao deseño e cualificación de exames, o deseño de escalas de nivel de dominio ou a profesionalización dos avaliadores. Tal e como sinala Bordón, calquera reflexión teórica relacionada coa avaliación estará ligada a algún dos temas antes citados, xa que estes se relacionan entre si de tal xeito que o modelo de lingua é a substancia necesaria

para que existan a fiabilidade e a validez, e a ética é o elemento aglutinador, xa que a ética garante a equidade do instrumento avaliador, de tal xeito que, para que un exame de lingua sexa ético, precisa unha concepción da lingua, medir o que se supón que ten que medir e deben obterse resultados consistentes deste exame. De aí que o reto da avaliación neste século XXI sexa elaborar exames ou outros procedementos de avaliación que sexan xustos, viables, válidos e fiables, ademais de ser conscientes do impacto que ten a proba sobre os procesos de ensino e aprendizaxe.

CAPÍTULO 2

O Certificado de Lingua Galega (Celga)

2. O Certificado de Lingua Galega (Celga)

A realidade socioeconómica de Europa mudou nos últimos anos debido aos movementos migratorios de dentro e de fóra do continente europeo; isto deu lugar a unha sociedade europea multilingüe que precisa de linguas comúns de comunicación. Esta situación provoca que os sistemas de certificación das distintas linguas europeas acaden unha importancia cada vez maior na formación lingüística e nas necesidades de certificación da poboación. Debido a isto, o Consello de Europa no ano 2001, coa publicación do *Common European Framework for Languages: Learning, Teaching, Assessment*, marcou a actuación común nas distintas linguas europeas nas liñas de aprendizaxe, ensino e avaliación, ademais de establecer o nivel de competencia comunicativa nas distintas destrezas que ten que acadar cada usuario nos diferentes niveis establecidos.

Para favorecer o multilingüismo, existen iniciativas que parten de diferentes institucións e entidades na promoción do ensino de linguas que facilitan a mobilidade da cidadanía en Europa e o enriquecemento das persoas, na medida en que posuír unha determinada competencia comunicativa nunha lingua é un valor en alza para calquera cidadán europeo na mellora da súa profesión e mesmo no seu desenvolvemento como ser humano. Nesa liña europeísta, a Secretaría Xeral de Política Lingüística da Xunta de Galicia (en diante, SXPL) tenta favorecer a poboación con necesidades específicas na aprendizaxe e certificación de coñecemento da lingua galega mediante a realización, no mes de decembro de 2007, da primeira convocatoria dos novos certificados de lingua galega, coñecidos como Celga. Estes certificados están en paralelo cos niveis do MECRL que impulsou o Consello de Europa no 2001.

2.1. Antecedentes

Para facer unha exposición sobre os novos certificados de galego, cómpre mirar atrás e ver como se desenvolveron nos últimos anos o ensino e a certificación da lingua galega para adultos e como se está levando a cabo esa mudanza cara a un novo sistema de certificación adaptado ao MECRL.

Dende que a Lei 3/1983, do 15 de xuño, de normalización lingüística recollía as bases que asentaron o uso do galego nos ámbitos da Administración, ensino e medios de comunicación ata hoxe en día, producíronse varios cambios de normativa e diferenzas na evolución do orzamento asignado á formación de adultos en galego (Consello da Cultura Galega, 2002).

O primeiro período, que acada dende 1983 ata 1985, caracterízase por un escaso gasto e unha orientación dirixida principalmente á formación do profesorado do ensino regrado en lingua galega, o cal lles permitía impartir a materia de galego.

O segundo período, dende 1986 ata 1992, supuxo un incremento do orzamento e o asentamento das liñas de actuación en normalización lingüística da antiga Dirección Xeral de Política Lingüística. Neste momento a formación de adultos en galego estaba centrada na formación do funcionariado da Administración galega, pero a capacitación lingüística estaba dirixida á produción escrita, dado que era o ámbito público máis patente; cómpre sinalar que este corpo de funcionarios, maioritariamente, posuía unha competencia maior en comprensión e expresión oral ca en expresión escrita, polo tanto, os esforzos foron dirixidos cara á mellora na expresión escrita, facendo fincapé na norma ortográfica oficial.

Nun terceiro período, dende 1993 a 2004, increméntase o orzamento debido ao compromiso establecido entre a Xunta de Galicia e o Goberno central para compensar a débeda histórica coa normalización lingüística.

Durante todos estes anos, cabe destacar que o ensino de galego para adultos estivo marcado pola necesidade da cidadanía de conseguir un certificado requirido para o acceso á función pública ou ben como mérito nos concursos para o corpo de funcionarios. Isto provocou a necesidade de impartir cursos de formación de galego nos que primaba a asistencia como requisito para a consecución do título, e onde a didáctica destes cursos estaba baseada nun método gramatical e de tradución que pouco favoreceu o desenvolvemento da competencia comunicativa, xa que o que primaba e se valoraba era o coñecemento da gramática.

En setembro de 2004 aprobouse, por unanimidade de todas as forzas políticas con representación no Parlamento de Galicia, o Plan xeral de normalización da lingua galega (Xunta de Galicia, 2006) que establece liñas de actuación en distintos ámbitos da sociedade e que será un punto de inflexión con respecto ao antigo sistema de ensino do galego para adultos.

A partir deste momento, comeza a considerarse a necesidade de mudar o sistema de ensino de galego para adultos e a súa certificación. Así, créase un equipo de investigación sobre avaliación, formado por especialistas no ensino de linguas procedentes de distintos ámbitos da docencia que, dependendo do Centro Ramón Piñeiro para a Investigación en Humanidades e da Universidade e Santiago de Compostela, comeza a deseñar as liñas de actuación dos certificados. En primeiro lugar, faise unha tradución ao galego no 2005 do *Marco europeo de referencia para as linguas: aprendizaxe, ensino e avaliación*; manual que serve como referente para todas aquelas persoas implicadas na tarefa do ensino de linguas. O MECRL diferencia en tres niveis – A, B e C– todas aquelas habilidades (en inglés *can do*) que calquera usuario dunha lingua debe posuír para acadar a certificación dun determinado nivel.

Nunha segunda fase do proxecto, admítese o galego como lingua membro na Asociación de Examinadores de Linguas Europeas (ALTE), creada

en 1990, coa finalidade de establecer escalas e niveis comúns de competencia, constituír e implantar procedementos homologados para o desenvolvemento, confección e administración de probas de exames, de sistemas de cualificación e de comunicación de puntuacións obtidas, así como de análise e investigación de resultados. Con este motivo, os vinte e seis membros de ALTE , representantes doutras tantas linguas europeas, e da que o galego é membro de dereito dende novembro de 2005, veñen colaborando nunha serie de proxectos para o intercambio de ideas e experiencias no ámbito da avaliación de linguas. Con motivo da entrada do galego en ALTE, fíxose unha presentación dun modelo de exame para a certificación do galego dun nivel A2 do MECRL e foi admitido por decisión unánime de todos os membros de ALTE.

2.2. Descrición do Celga

Unha vez establecidos xa os parámetros de actuación do grupo de investigación en avaliación, comezou a planificación e o deseño das probas que constituirán os novos certificados de lingua galega adaptados ao MECRL. O Celga está configurado mediante distintos niveis de lingua, recolle as habilidades que reflicten a competencia comunicativa nas distintas destrezas (*can do*) e agrupa os contidos que calquera usuario de galego debe demostrar para obter este certificado nos distintos niveis de Celga, do 1 ao 5. Todos estes contidos recóllense na publicación da Xunta de Galicia do 2006 titulada *Niveis de competencia en lingua galega* que, tal e como aparece no documento, define:

(...) Os niveis de competencia en lingua galega que se presentan neste documento responden, como se indicou máis arriba, aos criterios establecidos no Marco europeo común de referencia para as linguas, coas adaptacións requiridas pola situación singular do idioma galego e a diversidade de candidatos aos que pretende atender. Pártese dunha presentación moi breve do nivel en cuestión, para continuar despois co detalle das habilidades nas que se debe reflectir esa competencia e cos contidos de distinto tipo nos que se deben apoiar. As habilidades fan referencia ás capacidades comunicativas que o aprendiz debe demostrar en cada chanzo da adquisición idiomática; é dicir, a

aquelas tarefas e interaccións sociais que debe ser quen de desenvolver a través da lingua galega para acceder á certificación correspondente. Este ‘saber facer’ remítese ás catro dimensións convencionais na aprendizaxe de calquera idioma: a comprensión oral ou auditiva, a comprensión escrita ou de lectura, a interacción e expresión oral e a interacción e expresión escrita. (Xunta de Galicia, 2007: 13)

Os contidos representan aqueles coñecementos que se consideran necesarios para poder activar de xeito satisfactorio –con corrección, adecuación e eficacia– as habilidades de comprensión e expresión ás que aludimos no parágrafo anterior. Estes coñecementos non constitúen, polo tanto, un eido de aprendizaxe con valor en si mesmo, senón que están ao servizo das capacidades comunicativas, que son as que se deben converter no eixe vertebrador da aprendizaxe do idioma e as que, en último termo, darán conta do nivel real de apropiamento realizado.

Os contidos organízanse en tres grandes bloques:

— Un de tipo comunicativo, no que se recollen de xeito máis detallado as capacidades de comprensión e expresión xa formuladas na sección de habilidades. Estas capacidades enúncianse en forma de funcións comunicativas que o usuario do idioma debe ser capaz de realizar en distintos ámbitos e situacións do seu quefacer cotián: desde os privados e informais, aos públicos e formais.

— O segundo está relacionado co código, que representa o soporte lingüístico e normativo que o aprendiz debe coñecer para facer uso do idioma neses contextos e coas finalidades ás que se alude na sección anterior. Trátase dos coñecementos fraseolóxicos, gramaticais, léxicos, fonéticos e ortográficos que se asocian máis directamente coas funcións comunicativas de cada nivel e que constitúen o requisito básico para resolvelas satisfactoriamente; un

repertorio de formas e normas coas que o usuario debe codificar e descodificar as mensaxes nas interaccións nas que participa.

— Finalmente, un terceiro bloque sobre coñecementos sociais e culturais relacionados con Galicia e cos galegos, que lle dan á aprendizaxe da lingua esa dimensión necesaria para converter o aprendiz nun usuario competente desde o punto de vista sociocultural.

Como exemplo dos contidos que conforman o documento *Niveis de competencia en lingua galega* presentamos o nivel Celga 1 e os seus contidos correspondentes:

2.2.1 Descrición xeral do nivel Celga⁴

A persoa que supere o Celga 1 posuirá un dominio da lingua *xeral que lle permita satisfacer as necesidades máis primarias de comunicación*: será capaz de comprender frases e expresións de uso frecuente, relacionadas con áreas que lle son especialmente relevantes (información elemental sobre si mesmo e a súa familia, compras, lugares de interese, ocupacións etc.); saberá relatar, describir, expoñer e demandar, en termos sinxelos, aspectos relacionados coa súa experiencia persoal, co seu contorno ou con interaccións sociais básicas; saberá comunicarse para levar a cabo tarefas cotiás que non requiran máis ca intercambios simples e directos de información sobre cuestións que lle son coñecidas ou habituais e en situacións de comunicación previsibles etc. É dicir, mostrará unhas capacidades comunicativas elementais pero suficientes para desenvolverse en contextos próximos e familiares,

⁴ Os datos que aparecen a continuación, aínda que poden ser consultados na orde que os recolle, consideramos adecuado expoñelos aquí para facilitar a comprensión dos comentarios que posteriormente se farán sobre eles.

materializadas nun repertorio lingüístico aínda esquemático e vehiculadas especialmente a través da oralidade.

2.2.2. Descrición do nivel segundo as habilidades

2.2.2.1. *Comprensión oral*

Comprenderá frases, expresións e o vocabulario máis habitual sobre temas de interese persoal e relacionado con áreas de prioridade inmediata. Interpretará o discurso de xeito eficaz para poder enfrontarse a necesidades concretas, sempre que se emita con lentitude e naturalidade, sexa breve, sinxelo, claro, redundante, práctico e ligado ao contexto. Máis en concreto, será capaz de comprender información e instrucións sinxelas, así como pasaxes curtas, mesmo gravadas, que traten sobre asuntos cotiáns e previsibles coma os seguintes:

— Intervencións simples de carácter social que permitan manter un contacto mínimo con outra persoa, poñendo en práctica habilidades sociais como presentarse ou despedirse.

— Información xeral relativa á identificación e caracterización persoal (relacións familiares, profesionais, estudos, tempo libre).

— Instrucións sobre orientación e localización no espazo, así como sobre horarios e datas.

— Información relativa á alimentación, produtos e obxectos de uso diario, prezos e formas de pagamento.

— Preguntas e instrucións sinxelas relacionadas coa saúde.

— Información básica de lugares ou oficinas públicas: correos, bancos, ambulatorios, institucións académicas etc.

— Información básica sobre feitos ou acontecementos, tanto do momento presente coma do pasado.

— Opinións sobre acontecementos de actualidade, sempre que se presenten de forma clara e breve.

— Mensaxes telefónicas sinxelas.

2.2.2.2. *Comprensión escrita*

Será capaz de ler textos moi breves e sinxelos sobre asuntos cotiáns se conteñen vocabulario moi frecuente. Saberá atopar información específica e previsible en cartas, catálogos, prospectos, menús de restaurantes, listaxes, horarios, anuncios publicitarios e artigos breves de carácter informativo nos xornais.

Comprenderá, interpretará e identificará información e instrucións sinxelas relacionadas con situacións cotiás e previsible nas que atope textos como os seguintes:

— Carteis e avisos públicos fixados en rúas, estradas e espazos de servizo público.

— Horarios e datas de viaxes, servizos de transporte, apertura e peche de establecementos.

— Información sinxela sobre a orientación e a localización no espazo.

— Normas elementais de seguridade.

— Instrucións sobre aparellos de uso frecuente (por exemplo, un teléfono público).

— Informacións e instrucións de ámbitos como a hostalería (aloxamento, menús de restaurantes etc.), a sanidade ou a vida académica.

— Indicacións básicas para o uso de medicamentos: posoloxía etc.

— Información xeral relativa á identificación e caracterización persoal: relacións familiares, profesionais, estudos, afeccións etc.

— Información sinxela relativa á alimentación e produtos de uso diario expostos nun establecemento comercial.

— Información básica de formularios e documentos normalizados das oficinas de correos, da Administración e de entidades académicas e bancarias.

— Noticias breves e titulares sinxelos dos xornais.

2.2.2.3. *Expresión oral*

Poderá comunicarse en tarefas elementais e habituais que requiran un intercambio simple e directo de información sobre actividades e asuntos cotiáns. En situacións estruturadas e previsibles da vida diaria, será capaz de participar en conversas breves con razoable comodidade, sempre que conte coa colaboración do interlocutor. Aínda que resulten moi evidentes as pausas, as dúbidas iniciais e a reformulación, farase entender en intervencións breves. Será capaz de enlazar grupos de palabras con conectores sinxelos tales como e, pero e porque. Saberá facer descricións e presentacións breves e sinxelas da familia e doutras persoas, das condicións de vida ou de traballo, das actividades diarias, dos costumes, dos gustos e preferencias, dos obxectos e posesións etc. Será capaz de facer declaracións ensaiadas de contido previsible e aprendido sobre temas importantes da vida cotiá, coa garantía de ser entendido por oíntes dispostos a concentrarse e a colaborar.

De maneira máis detallada, poderá resolver situacións e tarefas como as seguintes:

— Intervencións simples de carácter social que permitan manter un contacto mínimo con outra persoa, poñendo en práctica habilidades sociais como presentarse ou despedirse.

— Controlar a comunicación oral (solicitando repeticións, o soletreo de palabras ou a moderación na velocidade da fala).

— Dar información xeral relativa á identificación e caracterización persoal (relacións familiares, profesionais, estudos, tempo de lecer), de lugares e de obxectos.

— Solicitar e dar instrucións sinxelas sobre orientacións e localización no espazo e sobre horarios e datas.

— Reservar un cuarto nun hotel ou nun establecemento equivalente, facer preguntas sinxelas sobre o aloxamento e observacións ou reclamacións breves sobre o servizo.

— Pedir produtos expostos en establecementos comerciais e información básica sobre eles.

— Reclamar a atención do responsable dun establecemento e solicitar un servizo (restaurante, farmacia, tenda, gasoleira etc.).

— Pedir información relativa á alimentación e a produtos de uso diario, prezos e formas de pagamento.

— Darlle información básica sobre un problema de saúde a un médico ou a un farmacéutico.

— Pedir información e realizar xestións sinxelas en distintas situacións cotiás: nas oficinas de correos, bancarias, académicas ou administrativas; en estacións de transporte público, axencias de viaxes etc.

— Intervir en conversas sinxelas sobre experiencias persoais, centros de interese e acontecementos de actualidade, manifestando sentimentos, expresando acordo e desacordo e solicitando opinións.

2.2.2.4. *Expresión escrita*

Poderá escribir notas e mensaxes breves e sinxelas sobre temas relativos a áreas de necesidade inmediata e sobre aspectos cotiáns do seu contorno (persoas, lugares, condicións de vida e experiencias persoais e profesionais), empregando unha serie de frases e oracións enlazadas cos conectores máis usuais. Máis en concreto, será capaz de:

— Cubrir formularios básicos con información persoal: matrícula nun curso, apertura dunha conta nun banco etc.

— Escribir mensaxes breves de rexistro formal, por exemplo para facer unha reserva de aloxamento, pedir unha información sinxela sobre algo etc.

— Escribir unha carta ou unha postal sobre asuntos de natureza persoal.

— Escribir notas ou pequenas mensaxes a amigos sobre cuestións sinxelas.

— Escribir sobre a súa familia, lugares, condicións de vida e experiencias persoais e profesionais.

2.2.3. Contidos

2.2.3.1. *Lingua e comunicación: funcións comunicativas*

A persoa que supere este nivel será quen de dar e pedir información xeral e básica sobre persoas, lugares e obxectos; poderá expresar as súas opinións, sentimentos, desexos e preferencias, así como facer suxestións, invitacións e dar instrucións mínimas. Con respecto ao uso social da lingua, saberá saudar, presentarse a alguén, agradecer e pedir perdón. En relación coa organización do discurso, saberá dirixirse a un interlocutor, iniciar o discurso e finalizalo.

Estará en condicións de manter un control da comunicación oral que lle permita sinalar que non entende algo, solicitar unha repetición, pedir que se fale máis amodo e preguntar por unha palabra que non coñece ou esqueceu.

De xeito máis concreto, as funcións comunicativas que debe estar en condicións de desenvolver son as seguintes:

- Saudar e responder a un saúdo.
- Presentarse e presentar outra persoa.
- Concertar unha cita.
- Rematar un contacto e despedirse.
- Agradecer e desculparse.
- Felicitar e responder a unha felicitación.
- Manifestar condolencia e solidariedade.
- Invitar e responder a invitacións.
- Identificar persoas, lugares e obxectos.
- Identificar e designar as partes da casa e os obxectos domésticos de uso máis común.
 - Identificar e designar as partes do corpo humano.
 - Localizar no espazo un obxecto, persoa ou lugar.
 - Describir persoas, lugares e obxectos.
- Dar e pedir información persoal sobre: lugar de nacemento ou residencia, idade, estado civil, familia, estudos ou profesión, afeccións, estado de saúde.
 - Pedir e ofrecer información sobre situacións, tarefas e actividades cotiás: enderezos, forma de chegar a un sitio; tempo meteorolóxico; horarios e datas; prezos e fórmulas de pagamento; xestións administrativas e académicas sinxelas; acontecementos e experiencias.
 - Pedir permiso para facer algo en situacións cotiás.

- Ofrecer e pedir axuda, aceptala ou rexeitala.
- Expresar e preguntar por estados de ánimo e sentimentos: alegría, tristeza, dor, preocupación, interese etc.
- Expresar coñecemento ou descoñecemento sobre algo.
- Expresar propósitos e intencións en relación co futuro próximo.
- Mostrar aprobación ou desaprobación e acordo ou desacordo con algo.
- Solicitar información, explicacións ou axuda na aula.
- Ofrecerse para facer algo.
- Informar sobre habilidades e valoralas.
- Expresar e preguntar por gustos.
- Describir, facer valoracións e comparacións.
- Relatar algo en presente e pasado próximo.
- Datar momentos e acontecementos.
- Referirse a condicións e a circunstancias históricas.

2.2.3.2. *Lingua e sistema: coñecemento do código*

Os saberes relacionados co código lingüístico –gramaticais, fonéticos, ortográficos e léxicos– que se consideran necesarios para desenvolver de xeito satisfactorio as anteriores funcións comunicativas son os que se relacionan a continuación:

- Flexión de xénero e de número de substantivos e adxectivos; regras máis comúns de formación do plural.
- Concordancia de xénero e de número no grupo nominal.
- Diferenzas no uso de nomes comúns e nomes propios.
- Artigo determinado e artigo indeterminado: formas e usos.

- Contraccións das formas do artigo coas preposicións a, en, con, de e por.
- Casos máis usuais de presenza/omisión do artigo.
- Formas dos pronomes persoais tónicos en función de suxeito e de complemento.
- Usos de ti, vós/vostede, vostedes.
- Formas e usos dos pronomes persoais átonos; colocación nas oracións simples: afirmativas, negativas e interrogativas.
- Uso dos pronomes reflexivos.
- Demostrativos e posesivos: formas, contraccións e usos.
- Numerais e ordinais.
- Pronomes e adverbios interrogativos de uso máis común: que, quen, cal, cantos, como, onde.
- Formas, contraccións e usos dos indefinidos máis usuais: algún, ningún, alguén, ninguén etc.
- Formación, usos e valores máis comúns do indicativo (presente, pretérito, copretérito e futuro), do infinitivo, do xerundio e do participio dos verbos regulares e irregulares máis usuais (ser, estar, facer, saber, querer, ir, vir, haber, poder, ter, deber etc.).
- Uso do imperativo afirmativo.
- Perífrases verbais de uso máis habitual: estar + xer./inf. (duración), ir + inf. (futuro), haber + inf. (obriga), ter que + inf. (obriga).
- Formas e usos pronominais dos verbos máis usuais: chamarse, lavarse, vestirse etc.
- Adverbios máis comúns para a expresión de negación, afirmación, modo, tempo, lugar e dirección.
- Preposicións, conectores e enlaces de uso máis frecuente.
- Construción de frases de estrutura simple.

- Construción de oracións simples afirmativas, negativas e interrogativas.
- Correspondencia entre grafías e fonemas.
- Recoñecemento e pronuncia do sistema vocálico.
- Recoñecemento e pronuncia dos sons consonánticos.
- Entoación afirmativa, negativa e interrogativa.
- Regras básicas de acentuación.
- Nomes de países e xentilicios.
- Léxico sobre características físicas de persoas, obxectos e paisaxes: cores, tamaño, forma externa etc.
- Léxico básico sobre estados de ánimo, sentimentos ou carácter.
- Os nomes das horas, das partes do día, dos días da semana, dos meses do ano...
- Léxico relacionado coa idade, o sexo, o estado civil, o lugar de procedencia...
- Léxico para indicar parentesco.
- Léxico relacionado coas partes do corpo humano.
- Léxico relacionado coas pezas de vestir.
- Léxico relacionado coa vivenda, o fogar e o mobiliario doméstico.
- Os puntos cardinais e outro vocabulario básico referido á localización no espazo.
- Léxico relacionado con agrupamentos poboacionais e tipos de hábitat.
- Léxico relacionado con vías públicas, distribución urbana etc.
- Léxico referido a medios e sistemas de comunicación.
- Léxico relacionado co medio físico e co relevo.
- Léxico relacionado con situacións meteorolóxicas, temperatura etc.
- Léxico básico referido á fauna e á flora.

— Léxico máis habitual relacionado con alimentos, bebidas e gastronomía.

— Léxico básico relacionado cos seguintes ámbitos: a Administración; a vida académica; os servizos de telefonía, correos etc.; a saúde e a atención sanitaria; actividades de lecer e tempo libre (deporte, espectáculos etc.), actividades profesionais máis habituais etc.

— Expresións básicas de saúdo e despedida, tratamento, agradecemento e desculpa, acordo e desacordo, felicitación, condolencia etc.

A finalidade deste certificado de lingua galega é a de obter unha certificación que acredite a competencia comunicativa que un usuario ten en galego.

O Celga está dividido en cinco niveis, por orde de menor a maior dificultade: Celga 1, Celga 2, Celga 3, Celga 4 e Celga 5, aliñados ao MECRL nos seus niveis A, B e C e aos niveis de ALTE do seguinte xeito:

CELGA	MECRL	ALTE
Celga 1	A2	Nivel 1
Celga 2	B1	Nivel 2
Celga 3	B2	Nivel 3
Celga 4	C1	Nivel 4
Celga 5	C2	Nivel 5

2.3. Probas de dominio mediante tarefas

A innovación neste sistema de certificación é o tipo de exame que se deseñou para os distintos niveis do Celga, xa que é un modelo de proba de dominio⁵ baseado no ensino de linguas mediante tarefas (Gysen, 2006; Van den Braden, 2006), polo tanto, nos exames as destrezas (comprensión e expresión orais e escritas) están integradas e as tarefas que teñen que realizar os candidatos implican comunicación real e son semellantes a calquera actividade que poidan facer na súa vida cotiá na lingua da que son usuarios. Ademais de empregar a lingua para realizar tarefas significativas, a lingua tamén é significativa. Deste xeito, abandónase a tendencia de ensino en galego que se apoiaba no método tradicionalista (Sánchez Pérez, 2005), segundo o que a avaliación da lingua galega está relacionada con actividades de “encher ocos”, “corrixir erros” de gramática e ortografía ou facer traducións do español ao galego. Nesta nova etapa, preténdese que o ensino do galego para adultos se consolide ao abeiro do enfoque comunicativo (Richards e Rodgers, 1998) e de ensino de linguas mediante tarefas (Melero, 2004; Martín Peris, 2001, 2004; Zanón, 1999), o que lles permitirá aos candidatos demostrar realmente as súas habilidades e a competencia comunicativa no uso da lingua galega e non unicamente o coñecemento do código.

2.4. Definición de tarefas

Dende hai tres décadas asistimos a unha evolución do enfoque comunicativo que deu como resultado unha proposta de contidos aos que se lles engadiron os

⁵ Proba de dominio (*proficiency test*): as ditas probas non están baseadas nun programa particular, están pensadas para avaliar a habilidade lingüística de candidatos con distintas bagaxes educativas. Algunhas probas de dominio están deseñadas para mostrar se os candidatos teñen unhas habilidades necesarias para utilizar unha lingua nunha área específica, como pode ser a medicina, o turismo ou os negocios.

procesos de comunicación, orixinando así a creación de diferentes materiais de ensino de linguas, que dende os anos oitenta deron en chamarse tarefas. Para explicar que é unha tarefa, cómpre facer unha revisión sobre o dito termo partindo de Long (1985), que o acuñou por primeira vez e definiu tarefa como:

calquera actividade realizada por un mesmo ou para os demais, libremente ou con algún interese. (...) Cubrir un impreso, comprar uns zapatos, facer unha reserva de avión (...) Noutras palabras, por tarefa entendemos todo aquilo que facemos acotío na nosa vida, no traballo, para divertirnos e entre unha e outra cousa. (Long, 1985: 21)

E segundo Olga Juan (1999), nesta definición introdúcese o concepto de acción, que tamén postula o MERC, mediante o cal os estudantes teñen que “facer algo”, o obxectivo que teñen que alcanzar na súa aprendizaxe é o de cumprir cunha acción, co cal se converten en suxeitos activos no proceso de comunicación.

Segundo coas investigación sobre tarefas, Nunan (1989: 54) proporciona unha definición centrada na realidade da aula e tendo en conta a lingua como ferramenta necesaria para a realización das tarefas. Así, considéraa como:

unha unidade de traballo na aula que implique os aprendices na comprensión, manipulación, produción ou interacción na L2 mentres que a súa atención está centrada prioritariamente no significado máis ca na forma.

Posteriormente, Candlin (1990) introduce o concepto de coñecemento xa existente nos estudantes tanto no plano lingüístico coma no extralingüístico, e o de actividade xeradora de problemas que os estudantes teñen que ir resolvendo, polo tanto, para Candlin unha tarefa é “unha actividade interdependente, social e xeradora de problemas que implica a aplicación do coñecemento existente para a consecución duns obxectivos”.

Estaire e Zanón (1990) recompilan as definicións dos anteriores autores e fan unha síntese na que sinalan as características que debe presentar unha tarefa, xa que para eles cómpre que sexa:

- a) representativa de procesos de comunicación da vida real,
- b) identificable como unidade de actividade na aula,
- c) dirixida intencionalmente á aprendizaxe de linguas,
- d) deseñada cun obxectivo, unha estrutura e unha secuencia de traballo,
- e) orientada á consecución dun obxectivo de manipulación de información.

Para O. Juan (1999), os autores desta definición teñen unha proposta na que é indisoluble a aprendizaxe de linguas dunha reflexión sobre como se produce a adquisición do coñecemento lingüístico, o que dá lugar a un alumno máis activo e autónomo, xa que é o responsable da súa aprendizaxe; tamén mudará o papel do profesor, descentralízase o seu poder na aula e convértese nun asesor que actúa e intervéen segundo as necesidades do estudantado.

Segundo Martín Peris (2001, 2004) na súa reflexión sobre as implicacións do uso de tarefas na aula, para que unha actividade poida ser considerada como unha tarefa, debe realizarse en cooperación, promover a comunicación entre o alumnado e finalizar coa obtención dun resultado determinado. Máis concretamente, debe cumprir os seguintes requisitos:

1. A tarefa é un plan de traballo, nela márcase un obxectivo final e uns pasos xerais que se poden seguir para acadalo, pero non se especifica que estruturas lingüísticas deben usar os aprendices, senón que son eles os que escollen en cada momento a lingua que van empregar segundo a precisen, é dicir, as tarefas crean un marco de acción no que o alumnado escolle libremente.

2. Está centrada fundamentalmente no significado (ou sexa, a súa intención é que os aprendices se comuniquen para conseguir algo, non que utilicen simplemente algunha estrutura lingüística coma o subxuntivo). A

lingua é unha ferramenta para conseguir un obxectivo, non é o obxectivo en si mesmo.

3. Implica procesos reais de uso da lingua. As actividades poden ser semellantes ás que realizamos no mundo real (cubrir un formulario) ou non (atopar cinco diferenzas entre dúas imaxes), pero os procesos comunicativos que empregamos nelas deben ser os que empregamos no mundo real: por exemplo, nunha conversa auténtica temos que tentar explicarnos de xeito que os demais poidan comprendernos ou poder contar anécdotas da nosa vida, pero non adoitamos repetir listaxes de verbos ou recitar diálogos aprendidos de memoria. Se unha actividade ten un destes dous exemplos como obxectivos, non podemos dicir que implique procesos reais de uso da lingua.

4. Pode requirir o uso de calquera das destrezas lingüísticas, así como de varias ao tempo, o cal supón a activación da comprensión auditiva ou de lectura, da produción oral ou escrita, ou de ambas á vez.

5. Pon en marcha procesos cognitivos xa que, se os aprendices teñen que levar a cabo un obxectivo común, terán que poñerse de acordo en como o van facer, seleccionar a información que consideren relevante ou razoar sobre determinadas opinións e eleccións de lingua. Estes procesos cognitivos son moito máis complexos ca os que se activan nunha actividade na que só hai que repetir un patrón de lingua determinado.

6. Remata cun produto claramente comunicativo, que é o obxectivo mesmo da tarefa (unha tarxeta cos datos de alguén, unha listaxe de consellos para levar unha vida sa, unha carta de reclamación etc.).

2.5. Medidas acadadas para a implantación do sistema Celga

Para a implantación do sistema de certificación galego tiveron que artellarse unha serie de medidas adscritas aos ámbitos político, administrativo e metodo-

lórico, e supuxeron cambios substanciais para afianzar o prestixio do galego no ámbito da cidadanía, da Administración e do ensino para adultos que pasamos a detallar a continuación.

2.5.1. Medidas políticas: o Plan xeral de normalización da lingua galega

O Plan xeral de normalización da lingua galega aprobouse por unanimidade no Parlamento de Galicia o 21 de setembro de 2004; no documento publicado pola Xunta de Galicia no 2006 recóllense máis de catrocentas medidas concretas, elaboradas a partir da realidade lingüística de Galicia, para favorecer o uso do idioma nos distintos ámbitos da sociedade e garantir que calquera cidadán poida vivir plenamente en galego. Concibido coa participación de máis dun cento de profesionais representativos de todos os ámbitos da sociedade galega, constitúe un dos avances máis importantes en política de extensión do uso social do galego, tras a aprobación da Lei de normalización lingüística de Galicia, en xuño de 1983.

O Plan xeral de normalización da lingua galega establece obxectivos xerais e concretos, estruturados en dous grandes bloques sectoriais de actuación. O primeiro bloque integra os sectores de **carácter transversal**: dereitos lingüísticos, novas tecnoloxías e implementación do corpus. O segundo grupo correspóndese cos denominados **bloques verticais**, correspondentes aos sete ámbitos sociais nos que se pretende actuar: Administración, educación, familia e mocidade, economía, sanidade, sociedade e proxección exterior da lingua.

Tal e como aparece no Plan xeral de normalización da lingua galega (Xunta de Galicia, 2006), os obxectivos xerais son:

— Garantirle a posibilidade de vivir en galego a quen así o desexe, sabendo que conta co amparo da lei e das institucións.

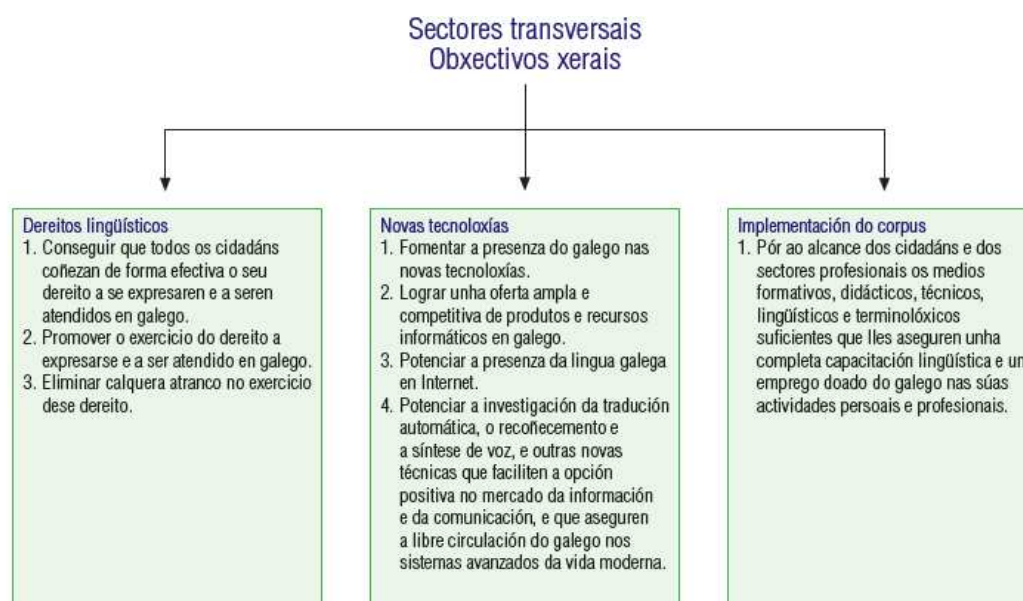
— Conseguir para a lingua galega máis funcións sociais e máis espazos de uso, priorizando a súa presenza en sectores estratéxicos.

— Introducir na sociedade a oferta positiva de atender o cidadán ou o cliente en galego como norma de cortesía dun novo espírito de convivencia lingüística.

— Promover unha visión afable, moderna e útil da lingua galega que esfarele prexuízos, reforce a súa estima e aumente a súa demanda.

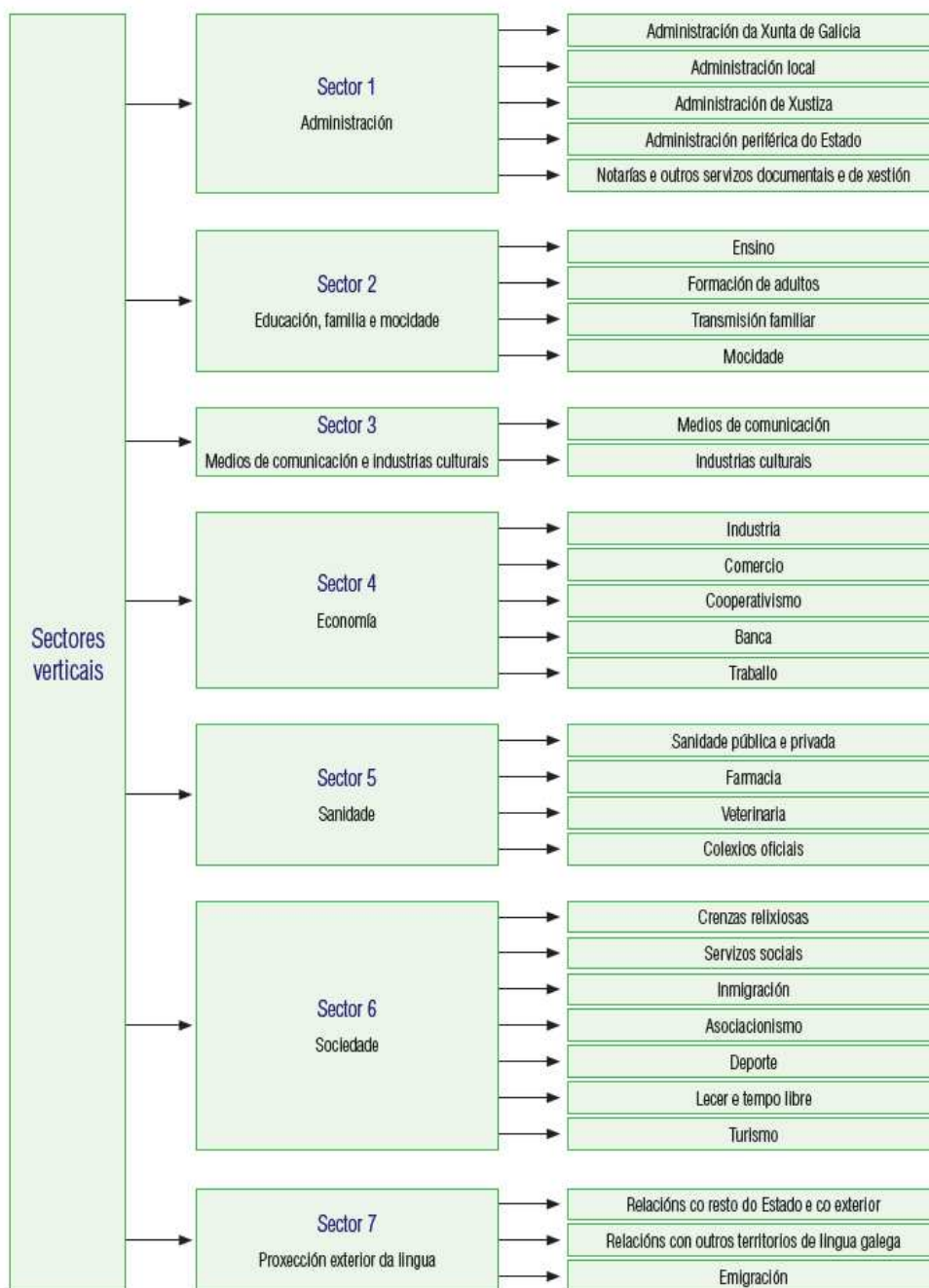
— Dotar o galego dos recursos lingüísticos e técnicos necesarios que o capaciten para vehicular a vida moderna.

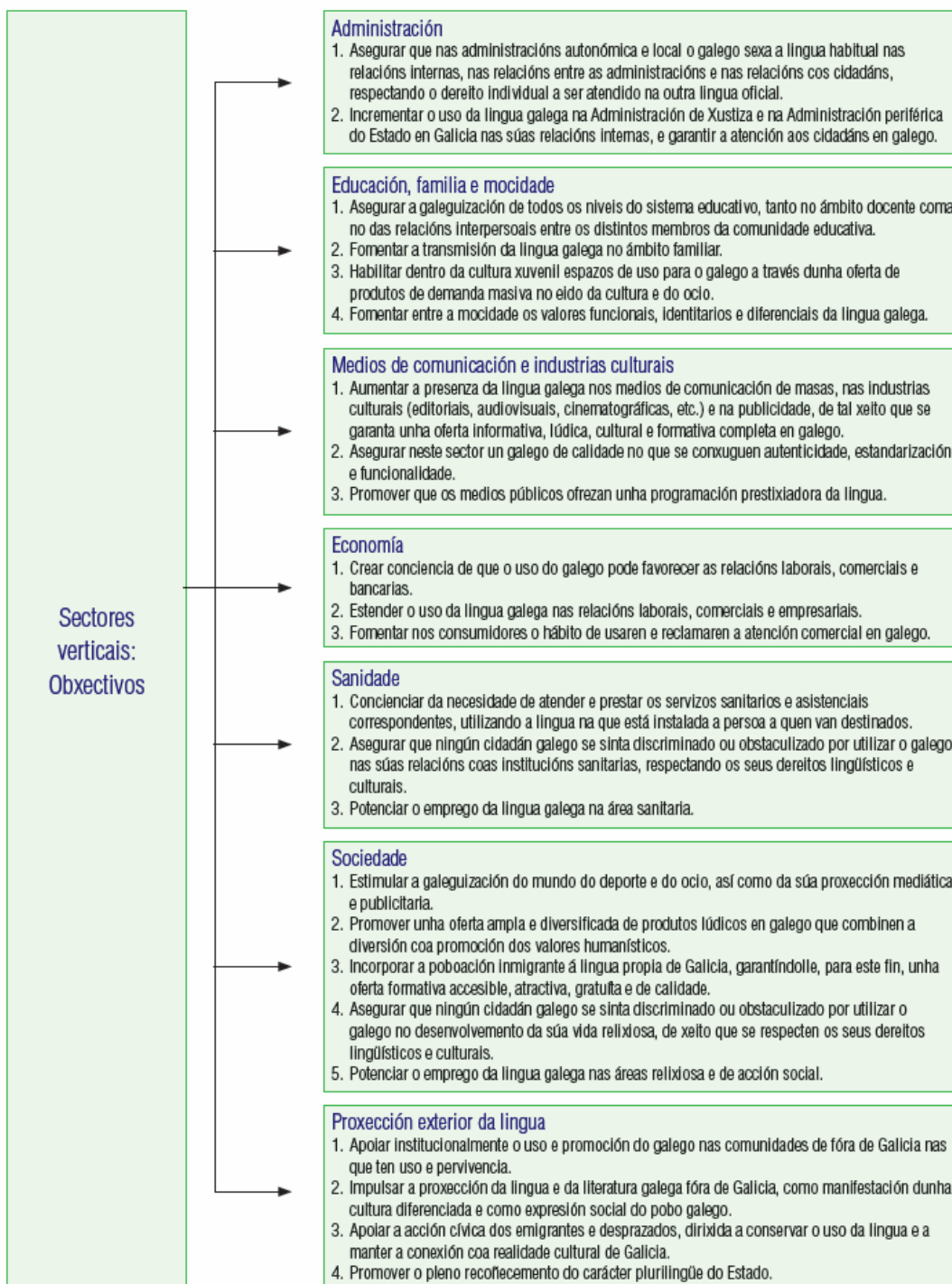
A continuación presentamos unha serie de cadros que recollen de xeito resumido os principais obxectivos do sector 2, que corresponde á familia, educación e mocidade; nos seus sectores transversais (dereitos lingüísticos, novas tecnoloxías e implantación do corpus) e verticais (Administración, educación, familia e mocidade, economía, sanidade, sociedade e proxección exterior da lingua).



Sistema de certificación de lingua galega (Celga) no *Marco Común Europeo de Referencia para as Linguas*. Estudo de fiabilidade e validez

Sectores verticais





Diagnóstico do sector

Na formación de adultos, que é o que atinxe o sistema de certificación de galego, recóllense unha serie de puntos fortes e débiles que permiten sinalar as eivas sobre as que se debe actuar. Neste caso sería o segundo punto pertencente ao cadro de puntos débiles: deficiencias no sistema de capacitación lingüística dos adultos e ausencia dun marco de referencia homologable a outras linguas.

Área 2. Formación de adultos

Puntos fortes	Puntos débiles
<ul style="list-style-type: none"> • Aumentou considerablemente a proporción de cidadanía con formación en lingua galega. • Os cursos de formación de adultos levados a cabo desde os anos oitenta tiveron resultados positivos de cara á aprendizaxe da variedade estándar. • Existe un sector do funcionamento (Administración, educación, sanidade...) que posúe unha boa competencia para desenvolver a súa actividade profesional en galego. • Existe un amplo sector da poboación, adulta non alfabetizada na lingua galega e que lle gustaría adquirir esa competencia e tamén ter un coñecemento maior de diferentes aspectos da cultura galega. 	<ul style="list-style-type: none"> • Mantéñense bolsas de persoas que non recibiron nunca formación en galego. • Obsérvanse deficiencias no sistema de capacitación lingüística dos adultos (modelo de cursos de iniciación e perfeccionamento): escasa regulamentación, mínimo control, falta de reflexión sobre obxectivos sociolingüísticos e pedagóxicos, etc. • A formación en lingua galega na educación non regrada carece dun marco de referencia homologable ao doutras linguas. • Existen aínda sectores do funcionamento que non posúen a competencia necesaria para desenvolverse en galego. • No acceso á función pública, non están singularizados e concretados os obxectivos formativos requiridos en relación co perfil de cada unha das prazas.
<ul style="list-style-type: none"> • Organizáronse actividades formativas en ámbitos temáticos sectoriais: linguaxe administrativa, etc. 	<ul style="list-style-type: none"> • Escasa adaptación da oferta formativa ás necesidades específicas e sectoriais de cada colectivo. • Insuficiencia de materiais lexicográficos (dicionarios bilingües e plurilingües, por exemplo). • Carencias a nivel de materiais terminolóxicos específicos, sobre todo en determinadas áreas de coñecemento. • Deficiencias na selección do vocabulario especializado de certas áreas e na súa divulgación entre os profesionais.
<ul style="list-style-type: none"> • Aínda que escasos, existen algúns recursos para a aprendizaxe da lingua galega en formato electrónico e a través de Internet. 	<ul style="list-style-type: none"> • Seguen sendo escasos os medios formativos que utilicen ou se aproveiten das TIC. • Non se desenvolveu a perspectiva da autoaprendizaxe da lingua, fornecendo materiais diversos en todo tipo de soportes.
<ul style="list-style-type: none"> • Nas dúas últimas décadas asentouse unha variedade normativa no que se refire á lingua escrita que facilitou o acceso a esta por parte da poboación. • Aínda que con máis dificultades, tamén se foi configurando un estándar oral que permitiu superar prexuízos dialectalizantes. 	<ul style="list-style-type: none"> • A pervivencia de certas discrepancias ortográficas puido crear desconcerto nalgunhas persoas e, sobre todo, foi utilizada de forma interesada por sectores contrarios á normalización do galego. • Con moita frecuencia, o galego utilizado nos medios de publicación ou en actos públicos caracterízase por unha total falta de autenticidade e mesmo incorrección, o que ten consecuencias nocivas tanto para a súa dignificación coma para a socialización dun modelo de lingua.

Polo tanto, unha vez establecido o sistema de certificación Celga, estaríase emendando a eiva relativa á formación de adultos en galego.

2.5.2 Medidas administrativas: Orde do 16 de xullo de 2007

A Orde do 16 de xullo de 2007, publicada no DOG número 146 do luns 30 de xullo de 2007, regula os certificados oficiais de coñecemento dos diferentes niveis de lingua galega (Celga) e as probas para obtelos. No artigo 2, niveis de coñecemento da lingua galega, establécense cinco niveis oficiais de coñecemento da lingua galega e os seus correspondentes certificados, que terán a seguinte denominación:

1. Certificado de lingua galega 1 (Celga 1): acredita unha competencia no uso da lingua que permite satisfacer as necesidades de comunicación máis básicas. Correspóndese co nivel A2 do marco europeo común de referencia para as linguas: aprendizaxe, ensino e avaliación (MECRL) e 1 da Asociación de Avaliadores de Linguas Europeas (ALTE).

2. Certificado de lingua galega 2 (Celga 2): acredita unha competencia no uso da lingua suficiente para intervir nun conxunto variado de situacións de comunicación con relativa seguridade e eficacia. Correspóndese co nivel B1 do MCERL e 2 de ALTE.

3. Certificado de lingua galega 3 (Celga 3): acredita unha competencia no uso da lingua que permite empregala cun grao importante de corrección normativa sobre temas diversos. Correspóndese co nivel B2 do MCERL e 3 de ALTE.

4. Certificado de coñecemento da lingua galega 4 (Celga 4): acredita unha competencia no uso da lingua suficiente para expresarse con fluidez e

espontaneidade na maior parte das situacións comunicativas. Correspóndese co nivel C1 do MCERL e 4 de ALTE.

5. Certificado de lingua galega 5 (Celga 5): acredita unha competencia no uso da lingua suficiente para desenvolverse sen dificultades e con corrección en calquera tipo de situación de comunicación, desde a máis informal ata a máis formal. Correspóndese co nivel C2 do MCERL e 5 de ALTE.

No artigo 3, destinatarios, establécese que terán dereito a presentarse a exame todas as persoas maiores de dezaseis anos ou aquelas que os cumpran no ano natural de convocatoria das probas, logo do pagamento dos dereitos de exame correspondentes.

No artigo 4, competencia, delimítase que a Secretaría Xeral de Política Lingüística convocará e realizará as probas. Así mesmo, expedirá os certificados citados no artigo 2 para acreditar a superación das probas.

No artigo 5, Comisión Central de Avaliación, establécese que a Secretaría Xeral de Política Lingüística nomeará unha Comisión Central de Avaliación que se responsabilizará das seguintes funcións:

- a) Programación das probas e proposta de convocatoria á Secretaría Xeral de Política Lingüística.
- b) Elaboración dos exames.
- c) Avaliación daquelas probas que lle corresponda.
- d) Elaboración de actas do resultado das avaliacións.

A Comisión Central de Avaliación das probas estará composta polos seguintes membros, nomeados pola Secretaría Xeral de Política Lingüística entre expertos na avaliación do coñecemento da lingua galega:

- a) Presidente/a.
- b) Secretario/a.
- c) Cinco vogais.

O nomeamento dos membros da Comisión Central de Avaliación terá unha vixencia anual.

Como complemento deste, preséntase no artigo 6, comisións sectoriais de avaliación, a creación de comisións sectoriais de avaliación para cada certificado; cada comisión sectorial de avaliación estará composta, cando menos, por presidente/a, secretario/a e tres vogais. Segundo o número de candidatos que opten ás probas de cada certificado, poderán designarse avaliadores auxiliares. Serán funcións de cada comisión sectorial de avaliación:

- a) A administración e avaliación da parte das probas que lle corresponda, accións para as que poderá contar coa colaboración de auxiliares nomeados pola Secretaría Xeral de Política Lingüística.

- b) A confección das correspondentes actas das probas que lle correspondan e a súa remisión á Comisión Central de Avaliación.

- c) A remisión dos exames á Comisión Central de Avaliación.

No artigo 7, convocatoria das probas, establécese que a Secretaría Xeral de Política Lingüística, logo da proposta da Comisión Central de Avaliación, convocará as probas para a obtención dos certificados de coñecemento da lingua galega por medio dunha resolución que se publicará no *Diario Oficial de Galicia*. A dita resolución indicará:

- a) Os certificados obxecto de convocatoria.

- b) As datas e lugares de realización das probas. As datas de realización das probas dos distintos certificados non poderán coincidir temporalmente entre elas.

- c) A contía das taxas de exame e o sistema de liquidación.
- d) Os lugares e prazos de inscrición e de publicación de resultados.
- e) Os procedementos de reclamación ou de revisión de probas.

O artigo 8, programa e contido das probas, establece os niveis de coñecementos e os contidos esixidos para a superación das probas que outorguen dereito á obtención dos diferentes certificados. As probas correspondentes a cada un dos certificados centraranse na avaliación das catro destrezas lingüísticas: expresión oral, comprensión oral, expresión escrita e comprensión escrita.

O artigo 9, expedición de certificados, establece que o/a secretario/a xeral de Política Lingüística lles expedirá os certificados que regula esta orde ás persoas que superasen as probas correspondentes e obteñan a cualificación de aptitude. A Secretaría Xeral de Política Lingüística levará un rexistro administrativo de todos os certificados expedidos.

Finalmente, o artigo 10, validacións, establece que serán validables polo Certificado de lingua galega 1 (Celga 1): o certificado do curso de nivel elemental de lingua e cultura galegas para estranxeiros impartido polas universidades de Galicia ata a entrada en vigor desta orde; o certificado do curso de lingua galega para non galegofalantes impartidos polas universidades galegas ata a entrada en vigor desta orde e o certificado do nivel básico dos estudos de galego da UNED ata a entrada en vigor desta orde.

Serán validables polo Certificado de lingua galega 2 (Celga 2): o título de graduado escolar, sempre que se estudase toda a EXB ou primaria en Galicia e se cursase de maneira oficial a materia de lingua galega en todos os cursos dos ditos estudos; o certificado do curso de nivel medio de lingua e cultura galegas para estranxeiros impartido polas universidades de Galicia ata a entrada en vigor desta orde; o certificado do nivel medio dos estudos de galego

da UNED ata a entrada en vigor desta orde e o certificado de ter cursadas e aprobadas as materias de ciencias sociais e as materias optativas de galego da ESO nos centros do Bierzo e Sanabria acollidos ao programa de formación previsto no convenio para a promoción do idioma galego nos territorios limítrofes das comunidades autónomas de Galicia e Castela e León, subscrito entre ambas as dúas comunidades.

Serán validables polo Certificado de lingua galega 3 (Celga 3): o título de graduado en educación secundaria obrigatoria (ESO) ou técnico auxiliar (FP1), sempre que se estudase en Galicia na súa totalidade e se cursase de maneira oficial a materia de lingua galega en todos os cursos dos ditos estudos; o certificado do ciclo inicial de galego expedido polas escolas oficiais de idiomas; o certificado de ter cursadas e aprobadas as materias de ciencias sociais e as materias optativas de galego na ESO e no bacharelato nos centros do Bierzo e Sanabria acollidos ao programa de formación previsto no convenio para a promoción do idioma galego nos territorios limítrofes das comunidades autónomas de Galicia e Castela e León, subscrito entre ambas as dúas comunidades.

Serán validables polo Certificado de lingua galega 4 (Celga 4): o título de bacharelato (LOXSE ou BUP) ou técnico de FP2, sempre que se estudase en Galicia na súa totalidade e se cursase de maneira oficial a materia de lingua galega en todos os cursos dos ditos estudos; a certificación do curso de especialización en lingua galega para profesores de EXB.

Serán validables polo Certificado de lingua galega 5 (Celga 5): o título de licenciado en Filoloxía Galega ou Filoloxía Hispánica (subsección de galego-portugués); calquera outro título de licenciatura, diplomatura ou equivalente, sempre que se acompañe da certificación de que se cursaron, polo menos, 18 créditos de lingua galega, e se estea en posesión do Celga 4; o certificado do ciclo superior de galego, expedido polas escolas oficiais de

idiomas conforme o programa previsto no Real decreto 47/1992, do 24 de xaneiro, polo que se regulan os contidos mínimos correspondentes ao ensino especializado das linguas españolas impartidas nas escolas oficiais de idiomas.

2.5.3. Medidas metodolóxicas: formación para o profesorado, publicación de materiais didácticos e cursos preparatorios

Outra das medidas levadas a cabo pola SXPL para desenvolver o cambio no sistema de ensino e certificación da lingua galega consistiu na formación do profesorado que imparte os cursos preparatorios dos Celga. Estes profesores asistiron a cursos de didáctica de linguas impartidos por Cursos Internacionais da USC, coa finalidade de solucionar a eiva na formación en didáctica que ata agora tiñan os profesores de galego, facendo que se pasase dunha metodoloxía gramatical a outra comunicativa. A estes cursos de formación didáctica sumáronselle as Xornadas de Formación para Avaliadores dos Celga, que permiten crear un corpo de examinadores-avaliadores que dean apoio durante a realización do exame e no proceso de avaliación das probas, mantendo criterios homoxéneos que garantan a fiabilidade e a validez⁶ na avaliación (Alderson *et al.*, 1998: 175-279). Todos estes avaliadores reciben unha formación específica que os capacita para seguir o protocolo da realización da proba oral e a súa avaliación, ademais de permitirlles avaliar as probas escritas.⁷

⁶ Fiabilidade (*reliability*): consistencia ou estabilidade das medidas dunha proba. Unha proba é máis fiable cantos menos erros aleatorios conteña. Unha proba que conteña erros sistemáticos, coma un prexuízo contra un grupo determinado, pode ser fiable pero non válida.

Validez (*validity*): grao no que as puntuacións dunha proba permiten extraer conclusións adecuadas, significativas e útiles, de acordo cos propósitos dunha proba. Existen varios aspectos de validez, como a validez do contido, de criterio e de construto. Estes aspectos proporcionan diferentes tipos de evidencia para xulgar a validez total dunha proba cunha finalidade concreta.

⁷ A descrición do curso de formación preséntase máis detallada no capítulo 3, apartado 1.2. Mostra.

2.5.3.1. *Materiais didácticos e cursos preparatorios*

Para a preparación dos Celga existen dúas posibilidades: a primeira consiste en preparalo de maneira autodidacta e, para iso, cóntase coa axuda de materiais de apoio e un curso en liña con tutoriais que se poden atopar na web da SXPL (www.xunta.es/linguagalega/celga); unha segunda possibilidade consiste en asistir a un curso preparatorio impartido pola SXPL do nivel de Celga ao que decida presentarse. Estes cursos tamén están concibidos nunha liña comunicativa e de ensino de linguas mediante tarefas (Nunan, 1996), para o cal se crearon materiais de clase para os profesores e alumnos –*Aula de Galego*–, dos que existen catro manuais que se corresponden cos niveis de Celga 1, 2, 3 e 4.

Os cursos impartidos en todo o territorio galego e no exterior substitúen os anteriores cursos de iniciación e perfeccionamento. A realización deses cursos preparatorios non implica a obtención do Celga; para iso, o candidato terá que presentarse ás probas, que serán unificadas e avaliadas por unha comisión central avaliadora. Ata o ano 2010 estes certificados de lingua galega realizáronse en dúas convocatorias anuais en Galicia, España e Europa, unha en maio e outra en novembro; e unha en outubro en Sudamérica e África. Actualmente non se realizan probas fóra do territorio español e só existe unha convocatoria anual do Celga.

2.5.3.2 *Ferramentas de axuda para os avaliadores: Manual de examinadores*

Como ferramenta empregada para levar a cabo o protocolo de administración das probas Celga, existe un documento de consulta interno da SXPL –*Manual dos examinadores*– que se actualiza en cada convocatoria das probas Celga e no que os examinadores e avaliadores poden consultar a seguinte información:

- convocatoria, datas e lugares de realización das probas,
- material e documentación que debe portar o candidato,
- horario e locais de realización das probas,
- normas sobre seguridade, vixilancia e civismo,
- casos de discapacidade aos que se poden acoller os candidatos,
- tipos de adaptación das probas dependendo das discapacidades,
- formato e duración das probas,
- persoal que participa nas probas e as súas funcións,
- control e seguridade do material do exame,
- xuntanza previa para a coordinación das comisións sectoriais de avaliación,
- organización nos centros examinadores: preparación do espazo, dos medios técnicos, distribución de funcións, organización das listaxes e chamamentos,
- información sobre a distribución da proba escrita: entrada, acomodación e acollida dos candidatos, desenvolvemento da proba, recollida do material por parte dos examinadores que atenden cada aula de exame, entrega do material na sala da comisión,
- sistema de corrección da proba escrita (será tratado en profundidade no seguinte capítulo)
- introdución das cualificacións da proba escrita na aplicación informática,
- reunión previa á proba oral no centro de exame,
- administración da proba oral: espazo e preparación da proba, o equipo avaliador, deberes dos examinadores, deberes dos candidatos, tratamento de incidencias, mecánica da entrevista ao candidato,
- sistema de avaliación da proba oral: cubrir as grellas de avaliación, exemplos de proba oral con documentación gráfica,

- instrucións e cronograma do día do exame especificado por niveis do Celga 1 ao Celga 4,
- listaxe de material para o día do exame,
- criterios de corrección e avaliación da expresión escrita nos niveis do Celga 1 ao Celga 4,
- criterios de corrección e avaliación da expresión oral nos niveis do Celga 1 ao Celga 4,
- modelos de grellas de avaliación da expresión oral (folla do avaliador e do entrevistador),
- documentos administrativos: enquisa de calidade, folla de recollida de datos bancarios para o pagamento dos avaliadores, folla de reclamación, certificación de asistencia, informe sobre o desenvolvemento das probas Celga no exterior de Galicia.

2.6. O modelo de exame

Tal e como avanzamos neste capítulo, o máis significativo dos exames que conforman as probas Celga é o seu deseño, xa que os candidatos deben resolver tarefas significativas para que se poida certificar a súa competencia comunicativa⁸ en galego nos distintos niveis do Celga. Nestas probas non existe unha proba específica de gramática para demostrar o coñecemento do código, senón que as tarefas presentan as destrezas⁹ integradas e a gramática

⁸ Competencia comunicativa (*communicative competence*): capacidade para utilizar adecuadamente unha lingua en diversas situacións. Bachman (1990: 16-68).

⁹ Destreza (*skill*): capacidade para facer algo. Nas probas de lingua distínguense catro destrezas: comprensión escrita, expresión escrita, comprensión oral e expresión oral. Cando se fala das destrezas integradas, significa que nunha mesma proba poden darse dúas ou máis destrezas xuntas que deben ser empregadas para a resolución da tarefa. Por exemplo, a proba A1 do Celga implica a comprensión oral e a expresión escrita. *Glosario de termos para a avaliación de linguas* (2008: 33).

convértese nun parámetro máis de avaliación xunto con outros que analizaremos máis adiante.

Os exames Celga engloban dúas probas no seu formato, unha escrita e outra oral. Cada unha delas inclúe diferentes tarefas que o candidato deberá resolver; a primeira delas, a proba escrita de xeito individual, e a segunda, a proba oral en colaboración co examinador que dirixirá a entrevista. Para obter a cualificación de ‘apto’, requírese superar as dúas probas na mesma convocatoria de exame. A proba escrita realizarana todos os candidatos ao mesmo tempo, na data e no horario fixados. A proba oral será individual e terá lugar, como norma xeral, o mesmo día da proba escrita; tamén se poderá realizar nos días seguintes se o número de candidatos o fixese necesario.

O modelo de exame está dividido en dúas partes:

Proba escrita: que consta, pola súa vez, de dúas probas:

— Proba A1. As destrezas implicadas son a comprensión oral e a expresión escrita)

— Proba A2. Desenvólvese unha tarefa de comprensión escrita e expresión escrita.

O candidato realizará a proba escrita de maneira conxunta co resto dos candidatos nas aulas habilitadas para tal fin. Nesta parte da proba, solicítaselle ao candidato unha tarefa na que as destrezas están integradas e se avalían tamén de maneira conxunta, xa que, para a realización da produción escrita en ambas as partes (A1 e A2), o candidato ten que comprender o texto oral na proba A1 e o texto escrito na proba A2 que lle serven como *input* para a súa produción escrita. A duración das dúas probas escritas será de entre dúas e tres horas, en función do nivel de Celga; os cinco primeiros minutos estarán dedicados á lectura do *input* para resolver a tarefa, logo escoitarán dúas veces

seguidas un audio e o tempo restante, ata completar as dúas/tres horas, servirá para realizar as dúas probas escritas (A1 e A2), que o candidato organizará como considere oportuno.

Proba oral: consta dunha única proba na que como *input* pode activarse calquera destreza comprensiva (comprensión oral ou comprensión escrita) e a produción ten que implicar a expresión oral. Nesta proba establécese unha aula de audio onde os candidatos, en grupos de non máis de 15 persoas, poderán escoitar un audio ou ler un texto (se na proba oral o *input* que se propón é un audio, este escoitarase dúas veces seguidas e o candidato poderá tomar as notas que considere oportunas para logo consultalas durante a entrevista co examinador). Despois, os candidatos desprazaranse ás aulas onde os recibirán o avaliador e o entrevistador, que dirixirá a entrevista; o examinador proporalle, a partir do texto oral ou escrito que acaba de escoitar ou ler, a resolución dunha tarefa, ao tempo que un avaliador estará valorando a súa produción oral.

A continuación mostramos unha serie de cadros que resumen os contidos da proba e a súa duración en cada un dos niveis Celga.

CELGA 1		TEXTO	TAREFA	DURACIÓN	ESTRUTURA DA PROBA (tempos aproximados)	
PROBA ESCRITA	PROBA A1	Comprensión oral	Expresión escrita	2 horas (9.15-11.15 h)	Lectura da tarefa	5 min
					Audición 2 veces do audio da proba A1	5 min
	PROBA A2	Comprensión escrita	Expresión escrita		Realización das tarefas	1.50 h
Ao remate da proba escrita publicaranse os chamamentos para a proba oral, indicando o día e hora aproximada para a súa realización.						
PROBA ORAL	PROBA B	Comprensión escrita ou oral	Expresión oral	Audición ou lectura noutra sala	Benvida	1 min
				10 min entrevista	Preguntas persoais	3 min
				5 minutos avaliación	Resumo	2 min
					Interacción	4 min

Sistema de certificación de lingua galega (Celga) no *Marco Común Europeo de Referencia para as Linguas*. Estudo de fiabilidade e validez

CELGA 2		TEXTO	TAREFA	DURACIÓN	ESTRUTURA DA PROBA (tempos aproximados)	
PROBA ESCRITA	PROBA A1	Comprensión oral	Expresión escrita	2 horas (9.15-11.15 h)	Lectura da tarefa	5 min
					Audición 2 veces do audio da proba A1	5 min
	PROBA A2	Comprensión escrita	Expresión escrita		Realización das tarefas	1.50 h
Ao remate da proba escrita publicaranse os chamamentos para a proba oral, indicando o día e hora aproximada para a súa realización.						
PROBA ORAL	PROBA B	Comprensión escrita ou oral	Expresión oral	Audición ou lectura noutra sala	Benvida e contacto persoal	1 min
					Resumo	2 min
				10 min entrevista	Relato (o candidato ten que contar unha anécdota ou episodio en pasado)	3 min
				5 minutos avaliación	Interacción	4 min

CELGA 3		TEXTO	TAREFA	DURACIÓN	ESTRUTURA DA PROBA (tempos aproximados)	
PROBA ESCRITA	PROBA A1	Comprensión oral	Expresión escrita	2.30 horas (9.15-11.45 h)	Lectura da tarefa	5 min
					Audición 2 veces do audio da proba A1	15 min
	PROBA A2	Comprensión escrita	Expresión escrita		Realización das tarefas	2.10 h
Ao remate da proba escrita publicaranse os chamamentos para a proba oral, indicando o día e hora aproximada para a súa realización.						
PROBA ORAL	PROBA B	Comprensión escrita ou oral	Expresión oral	Audición ou lectura noutra sala	Benvida e contacto persoal	1 min
					Resumo	3 min
				15 min entrevista	Exposición	4 min
				5 minutos avaliación	Interacción (pode haber debate, pero non hai que buscar explicitamente os argumentos contrarios do candidato)	4 min

CELGA 4		TEXTO	TAREFA	DURACIÓN	ESTRUTURA DA PROBA (tempos aproximados)	
PROBA ESCRITA	PROBA A1	Comprensión oral	Expresión escrita	3.00 horas (9.15-12.15 h)	Lectura da tarefa	5 min
					Audición 2 veces do audio da proba A1	10 min
	PROBA A2	Comprensión escrita	Expresión escrita		Realización das tarefas	2.45 h
Ao remate da proba escrita publicaranse os chamamentos para a proba oral, indicando o día e hora aproximada para a súa realización.						
PROBA ORAL	PROBA B	Comprensión escrita ou oral	Expresión oral	Audición ou lectura noutra sala	Benvida e contacto persoal	1 min
					Resumo	3 min
				15 min entrevista	Exposición	4 min
				5 minutos avaliación	Interacción (nalgún momento o entrevistador ten que buscar a contraargumentación)	4 min

2.7. Sistema de corrección da proba escrita

A corrección dos exames escritos lévaa a cabo o profesorado colaborador da SXPL escollido para cada convocatoria que recibise a formación específica para tal fin. Realizarase tendo en conta as seguintes cuestións:

— Non se deben facer marcas nos exames dos candidatos. Para a puntuación do exame os avaliadores dispoñen da grella de avaliación, onde marcarán as cualificacións correspondentes a cada un dos parámetros de avaliación (adequación, cohesión e coherencia, gramática e léxico).

— En cada un dos parámetros as puntuacións que o candidato pode obter sitúanse nunha escala do 1 ao 4. As dúas partes da proba (A1 e a A2) avalíanse por separado; polo tanto, cada corrector fornece oito notas en total para cada candidato; a folla de avaliación dos correctores inclúe a seguinte grella:

	Adecuación				Cohesión e coherencia				Gramática				Léxico				Cualificación	
A1	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	Apto	Non apto
A2	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4		

2.7.1. Cualificación final da proba escrita

Os criterios mínimos empregados para establecer a aptitude ou non dun candidato son os seguintes:

— Se o candidato obtén un 1 nalgún dos oito parámetros, considérase ‘non apto’.

— Se o candidato obtén máis dun 2 nos oito parámetros, considérase ‘non apto’.

— Se o candidato obtén *só un 2 nalgún dos oito parámetros*, considérase ‘apto’.

— Se o candidato *non obtén ningún 1 nin ningún 2*, considérase ‘apto’.

2.7.2. Control de calidade nas correccións dos exames

Coa finalidade de asegurar a fiabilidade das probas, estableceuse un sistema de correccións múltiple mediante o cal os exames poden chegar a ter ata catro revisións. O protocolo de corrección é o seguinte:

— *Primeira corrección*

O corrector da proba escrita terá que consignar as notas do exame escrito na folla de avaliación da proba e na aplicación informática da SXPL. Para que un candidato sexa considerado apto, como máximo poderá haber un único 2 nunha das oito notas da grella de corrección.

— *Segunda corrección*

Lévaa a cabo un supervisor de probas designado pola SXPL que descoñece a puntuación que se lle outorgou ao exame. Esta segunda corrección realízase sobre exames non aptos e sobre unha selección aleatoria de exames aptos e as notas consígnanse na folla de supervisión da proba. Se o avaliador e o supervisor concordan na nota, esta mantense; en caso de discordancia, pasarase o exame a unha terceira corrección.

— *Terceira corrección*

Se un exame é apto para un dos correctores e non apto para o outro, será sometido a unha terceira corrección, realizada por outro supervisor que descoñece a puntuación outorgada polos dous anteriores, a cal determinará a nota definitiva. Esta nota consígnase na folla de revisión da proba.

— *Cuarta corrección*

Realízase no caso de que o candidato presente unha reclamación contra a cualificación final xa publicada. Levarana a cabo os membros da Comisión Central Avaliadora.

2.8. Criterios de avaliación da proba escrita

Os criterios de avaliación da proba escrita son aqueles que se recollen no *Manual de examinadores* e que, como veremos a continuación, nalgún caso só se citan de xeito xeral e sen especificacións do que se debe de ter en conta en cada un dos parámetros. A razón vén dada porque, supostamente, estes parámetros son tratados en profundidade nos cursos de formación específica aos que os avaliadores deben asistir e que os capacita para realizar as funcións de avaliar e examinar nas probas Celga. O mesmo acontece cos parámetros de avaliación da proba oral, que non presentaremos neste apartado porque os trataremos en profundidade no 3.1.4.3.

2.8.1. Adecuación

Neste parámetro valórase que o candidato se axuste á tarefa solicitada e cumpra os seguintes requisitos:

— *Extensión*: cada nivel de Celga ten un número de palabras asignado en cada tarefa da proba escrita (Celga 1, 150 palabras; Celga 2, 200 palabras; Celga 3, 300 palabras e Celga 4, 400 palabras) e sobre este permítese un 20 % máis ou menos de palabras. A superación deste límite, por moi ben que cumpra os demais requisitos, indicará que non houbo unha resposta adecuada á tarefa e que non ofrece elementos suficientes para a súa avaliación. De calquera xeito, será necesario que este aspecto quede ben especificado ao comezo da proba nas indicacións para os candidatos.

— *Tratamento dos puntos da tarefa*: en canto ao tratamento dos puntos da tarefa, só contarán aqueles que se traten de forma explícita e con información pertinente.

— *Formato textual*: o escrito debe axustarse ao formato textual apropiado, aínda que poden permitirse leves alteracións nas convencións que configuran a macroestrutura dos textos.

— *Rexistro*: debe utilizarse, en función do nivel, un rexistro adecuado á tarefa e un modo adecuado de dirixirse ao interlocutor.

2.8.2. Coherencia e cohesión

Neste parámetro valórase nos textos:

- puntuación,
- estruturación do texto,
- presentación clara das ideas.

2.8.3. Gramática

Neste parámetro valórase a corrección ortográfica, morfolóxica e sintáctica:

- ortografía,
- acentuación,
- estruturas morfosintácticas,
- coñecemento da norma.

2.8.4. Léxico

Neste parámetro valórase a corrección, precisión e riqueza léxica:

- incorreccións (castelanismos...),
- precisión,
- riqueza léxica.

3. Obxectivos do Celga

Os obxectivos xerais do novo sistema de certificación de lingua galega tentan asentar as bases dos Celga como referente na avaliación da lingua galega, coa finalidade de conseguir:

- Consolidarse como sistema de certificación de lingua galega adaptado ao MECRL.
- Permitir a convivencia do Celga xunto con outros certificados de linguas europeas.
- Garantir a fiabilidade, validez e viabilidade na realización das probas.
- Solucionar as eivas existentes na formación e certificación de galego para adultos; para o cal a SXPL pon á disposición da poboación recursos humanos e materiais que permitan o éxito nestes campos.

Coa creación do sistema de certificación galega preténdese tomar medidas homoxéneas no ensino, aprendizaxe e avaliación da lingua galega para adultos asentada nos piares do MECRL e no Plan xeral de normalización da lingua galega. Así, o pilar do MECRL, xunto co enfoque comunicativo, fundamentarán a liña de formación do profesorado, o ensino da lingua, a creación dos materiais de aula e o deseño das probas Celga. Xunto a isto, o Plan xeral de normalización da lingua galega será a guía que marcará a actuación cara a solucionar as eivas existentes en relación coa formación de adultos en lingua galega.

CAPÍTULO 3

Estudios

3.1. Estudo I

3.1.1. Xustificación

O presente estudo ten como obxecto de investigación o sistema de certificación de lingua galega denominado Celga e a análise abórdase desde dous enfoques, a opinión dos avaliadores sobre este sistema de certificación e os resultados dos candidatos nas probas Celga. Esta investigación pretende fornecer datos que permitan optimizar o sistema Celga, aínda que cómpre salientar que, desde a súa creación, tanto a SXPL coma o equipo de investigación, formado polo grupo de especialistas en aprendizaxe, ensino e avaliación de linguas que leva a cabo o deseño das probas, mostran unha actitude de reflexión constante a prol da excelencia no ecosistema Celga; deste xeito procúrase unha reflexión que favoreza a calidade das probas e o labor dos avaliadores.

Alén desta primeira motivación na realización deste traballo, cómpre sinalar a necesidade de investigación en avaliación e certificación de galego para adultos, xa que non existen precedentes de análises sobre lingua galega nesta área, coa finalidade de solucionar unha eiva no desenvolvemento de estudos sobre avaliación da lingua galega fóra do ámbito dos estudos regrados. Pola contra, si existen estudos sobre a competencia lingüística en galego e español que adquiren os alumnos ao final do ensino regrado (i.e.: Loredó, Fernández, Suárez, 2005; Silva, Barros, Loredó, Núñez, Núñez, Sanz, Castro, Davila, 2010).

O obxectivo xeral desta tese é analizar a validez e fiabilidade do instrumento, que neste caso son as probas empregadas para a avaliación; tanto a validez coma a fiabilidade son dous requisitos básicos que debe ter calquera proba.

Deste xeito, a proba e o procedemento de avaliación teñen validez na medida en que se poida demostrar que o que se avalía (o construto) é o que se debería avaliar realmente e a información obtida é unha representación exacta do dominio lingüístico que posúen os candidatos que realizan o exame. Por outra banda, a fiabilidade é o grao de precisión dunha medida.

Con tal premisa preséntanse os resultados de dous estudos empíricos. O primeiro, “Opinión dos avaliadores sobre os criterios de corrección do Celga”, presenta os resultados das opinións que teñen os avaliadores sobre o propio sistema Celga, centrándonos nos criterios de corrección e avaliación que utiliza este sistema para extraer o construto no que se basean os avaliadores para cualificar os candidatos, é dicir, o interese está centrado na consideración que teñen sobre as ferramentas empregadas para avaliar as probas e da que se pode extraer un construto, un concepto teórico que condiciona a actuación dos avaliadores e polo tanto o funcionamento do sistema de avaliación; na obtención destes datos empregouse como instrumento un cuestionario administrado grupalmente aos avaliadores (ver anexo 1).

O segundo estudo empírico, “Análise das cualificacións dos candidatos das probas do Celga nos anos 2009-2010”, presenta as análises das puntuacións obtidas polos candidatos nas convocatorias do Celga dos anos 2009 e 2010, datos que nos permiten obter unha visión xeral sobre os resultados acadados polos candidatos e estudar como funcionan os criterios de avaliación na poboación obxectivo, así como presentar os resultados das análises da fiabilidade interxuíces e dos coeficientes de consistencia interna e da influencia das variables clasificatorias máis importantes.

Estes estudos permitirán tirar conclusións sobre o sistema de avaliación e focalizar as eivas que poida presentar o sistema, así como presentar as propostas de mellora que se poidan extraer deles.

3.1.2. Obxectivos e hipóteses

O obxectivo xeral desta investigación é dar a coñecer as opinións dos avaliadores do Celga sobre os criterios de avaliación empregados como ferramentas no proceso avaliativo. Para acadar este obxectivo participaron neste estudo un total de 152 examinadores aos que se lles pediu que cubriran un cuestionario (véxase anexo 1).

Para a análise de datos, a maiores dos resultados descritivos, realizáronse diferentes tipos de análises bivariantes, é dicir, observouse simultaneamente o efecto dunha variable sobre outra, neste caso para coñecer a influencia do nivel do Celga e experiencia do avaliador sobre os criterios de avaliación das probas de expresión oral e escrita.

Tamén se empregaron estatísticas multivariantes, método estatístico mediante o cal se pode determinar a contribución de varios factores a un resultado, para desvelar a estrutura subxacente dos criterios de avaliación (construto), a través do cal os avaliadores consideran os criterios de avaliación como ferramentas útiles, tendo en conta o nivel do Celga.

Os obxectivos específicos dos que partimos na investigación do presente traballo están relacionados coas hipóteses que fixemos, de tal xeito que pretendemos que a suma dos obxectivos específicos sexa o resultado do obxectivo xeral: coñecer as opinións dos avaliadores sobre o sistema Celga. Nos seguintes apartados presentamos os obxectivos específicos:

- Valorar a utilidade dos criterios de corrección e avaliación.
- Valorar a utilidade dos criterios de corrección e avaliación das probas de expresión escrita atendendo a dous parámetros: segundo o nivel do Celga e segundo a experiencia profesional dos avaliadores. Coñecer a opinión dos avaliadores sobre o construto, isto é, sobre a validez dos criterios de corrección e avaliación.

— Valorar a equidade das probas Celga atendendo ás opinións dos avaliadores.

— Analizar a estrutura que subxace na avaliación da competencia comunicativa oral e escrita mediante a representación da percepción de utilidade dos criterios de avaliación.

As hipóteses das que partimos están relacionadas co uso que os avaliadores poden facer dos criterios de corrección e avaliación, é dicir, se os empregan dun xeito adecuado, outorgándolles a cada un o valor que lle corresponde, ou se hai algún deles que poida ter un maior peso á hora de decidir que un candidato sexa apto ou non. Nesta liña as hipóteses que tentamos demostrar son as seguintes:

— O criterio máis útil á hora de decidir se un candidato é apto ou non na proba escrita é o de gramática, debido ao peso do tradicional método de ensino de linguas centrado no coñecemento do código e non na adquisición da competencia comunicativa.

— A valoración da utilidade dos criterios de corrección e avaliación varía en función da experiencia dos avaliadores. A dita utilidade é directamente proporcional á experiencia como avaliador no sistema Celga, canta máis experiencia máis positiva será a opinión sobre a utilidade de todos os criterios.

— A valoración da utilidade dos criterios de corrección e avaliación varía en función do nivel de Celga avaliado, de tal xeito que, a medida que os niveis de lingua son superiores, os criterios de adecuación á tarefa e gramática serán cada vez máis relevantes ca o resto de criterios.

— As probas Celga están deseñadas para evitar prexuízos culturais, isto é, que un candidato sexa quen de resolver con éxito as tarefas que se lle solicitan depende da competencia comunicativa que posúa e non de seu perfil sociolóxico. Polo tanto, a equidade é un dos trazos que caracterizan as probas Celga.

3.1.2.1. Utilidade dos criterios de corrección e avaliación

Unha das maiores preocupacións do grupo de investigación que diseña as probas Celga é que as ferramentas que se lles ofrecen aos avaliadores sexan útiles ao longo de todo o proceso de avaliación, tanto no día da convocatoria coma nos posteriores, cando avalían os exames escritos. Tendo en conta que os criterios de corrección e avaliación están deseñados para seren empregados como elementos preceptivos, estes deberían estar ben interiorizados por parte dos avaliadores e todos os criterios deberían ter o mesmo peso á hora de seren empregados; non obstante, cabe a posibilidade de que non todos teñan a mesma relevancia para os avaliadores á hora de decidir se un candidato é considerado apto ou non.

3.1.2.2. Influencia da experiencia dos avaliadores sobre a valoración dos criterios de corrección e avaliación

Dada a tradición do noso país no ensino, aprendizaxe e avaliación de linguas, caracterizada por unha formación do profesorado centrada na gramática e nas actividades de corte máis estruturalista, o punto de inflexión que se produce coa aparición do sistema de certificación Celga, no que prima o ensino e a avaliación de linguas baseados en tarefas, pode dar lugar a un paradoxo nos avaliadores, xa que recibiron formación no ensino de linguas dende unha perspectiva comunicativa, seguindo as pautas ditadas polo MCRL e, aínda así, o peso da tradición formativa pode ter unha influencia excesiva na importancia de determinados criterios de corrección e avaliación, como pode ser a gramática. Consideramos que esta posibilidade pode ser superada na medida en que o avaliador vai adquirindo máis experiencia na participación dos labores de avaliación e se vai familiarizando co método comunicativo e co ensino e avaliación de linguas mediante tarefas.

3.1.2.3. *Influencia do nivel de lingua sobre a valoración do Celga*

Outra das hipóteses que tentamos resolver está relacionada coa valoración do Celga en función do nivel de lingua que se avalía. Pode ocorrer que, dependendo do nivel de Celga, o peso dun criterio de corrección e avaliación poida ser máis significativo para un avaliador. Isto é, na proba escrita, a gramática pode ter menos peso no nivel Celga 1 ca no nivel Celga 4 ou, na proba oral, o criterio de pronuncia e entoación pode ter menos peso nun nivel baixo ca nun alto; isto pode deberse a que consideran que algúns criterios de corrección e avaliación non delimitan os distintos niveis de lingua, ou incluso poden crer que os criterios teñen distinto peso na suma total das puntuacións dependendo de se avalían unha proba escrita ou unha oral.

3.1.2.4. *Equidade nas probas Celga*

Con respecto á equidade (do inglés *fairness*) no sistema de certificación Celga, enténdese esta como a posibilidade de xustiza na avaliación de calquera candidato sen prexuízos ou favoritismos derivados do seu perfil sociolóxico, ou sexa, que o sistema Celga avalía a competencia comunicativa dos candidatos en función dos niveis de lingua e non en función de cuestións como o nivel cultural, a idade, a procedencia etc.

Trátase de demostrar que nas probas Celga non existen prexuízos culturais –termo procedente do inglés *cultural bias*, que consiste na distribución sistemática dun erro de medida causada pola orixe cultural dos candidatos–. Considérase que unha proba é parcial se un sector concreto da poboación de candidatos resulta privilexiado ou desfavorecido por algunha característica da proba, unha proba non imparcial culturalmente pode favorecer ou prexudicar os candidatos de determinadas orixes culturais (Bachman, 1990).

3.1.3. Mostra de avaliadores do Celga

A mostra está composta por 152 avaliadores do sistema de certificación de lingua galega, aínda que nas bases de datos da SXPL a poboación total de avaliadores é de 280 profesores colaboradores, é dicir, que a mostra utilizada correspóndese co 54,29 % do total (Figura 1). A dita mostra caracterízase por non utilizar un método de selección aleatoria. As características específicas da poboación, coa súa dispersión e falta de estabilidade á hora de participar como avaliadores, fan moi difícil utilizar os métodos máis fiables de mostraxe. Por todo isto, utilizamos un curso de formación ao que os avaliadores tiveron que asistir de xeito obrigatorio antes da convocatoria das probas Celga de 2012 para aplicar o cuestionario á maior poboación posible; deste xeito chegouse a unha cifra que consideramos importante para ter en consideración as conclusións que se poidan extraer deste estudo.

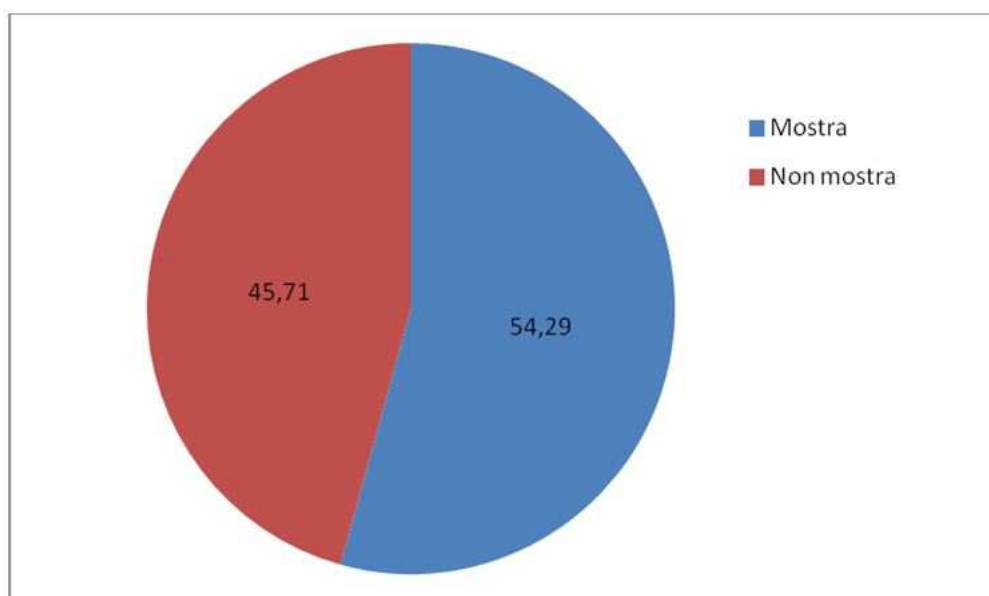


Figura 1. Porcentaxe da poboación que conforma a mostra de avaliadores do Celga

Unha das variables que se analizou nos examinadores foi a súa experiencia profesional no proceso de avaliación das competencias escritas e orais do galego no sistema Celga. Os avaliadores tiñan que explicar a súa experiencia tanto na competencia escrita coma na oral contestando unha pregunta aberta que despois se dividiu en categorías a partir das súas respostas.

Con respecto á avaliación da competencia escrita a maioría dos avaliadores tiña unha ampla experiencia con máis de tres convocatorias (39,3 %), un 18 % declarou que participara en tres convocatorias e un 23,3 % tiña menos de dúas convocatorias de experiencia fronte a só un 19,3 % que declarou que non tiña experiencia na avaliación da competencia escrita (Figura 2).

Na avaliación da competencia oral todos os suxeitos tiñan algún tipo de experiencia, un 40,7 % tiña entre catro e sete convocatorias e un 44,7 % tiña máis de sete convocatorias de experiencia (Figura 3).

En xeral, os avaliadores tiñan máis experiencia no proceso de avaliación oral ca no de avaliación escrita. Isto débese a cuestións de organización da SXPL, xa que procura que, en cada convocatoria de probas, participen todos como avaliadores da proba oral e algúns deles como avaliadores da proba escrita, en función das necesidades do servizo. De feito, dáse o caso de avaliadores con moita experiencia nas probas orais e pouca nas probas escritas.

Sistema de certificación de lingua galega (Celga) no *Marco Común Europeo de Referencia para as Linguas*. Estudo de fiabilidade e validez

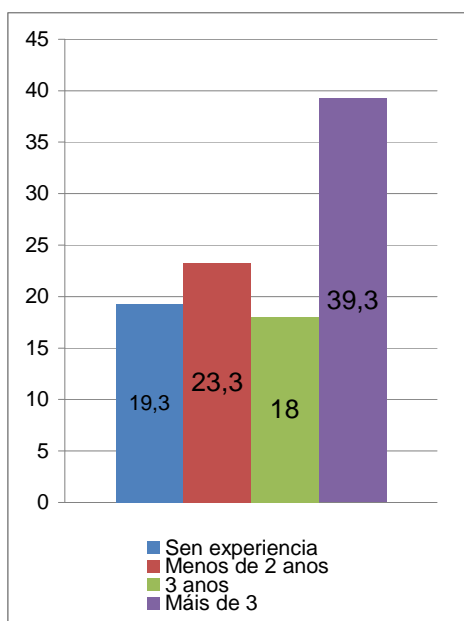


Figura 2. Experiencia dos avaliadores na avaliación da competencia escrita no Celga

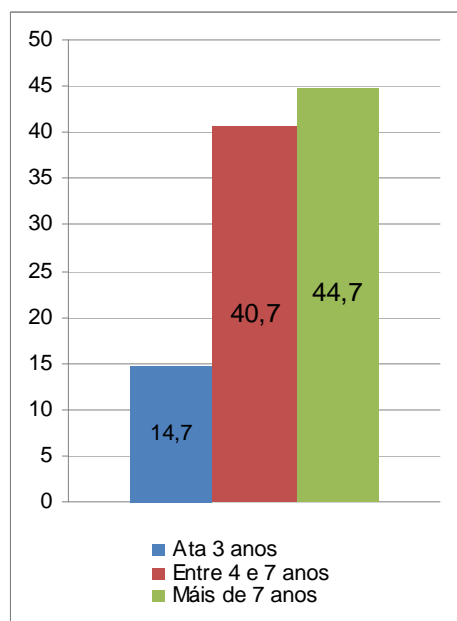


Figura 3. Experiencia dos avaliadores na avaliación da competencia oral no Celga

Como se pode ver na táboa seguinte (táboa I), a maioría da mostra ten unha ampla traxectoria nos dous campos de avaliación das probas orais e escritas. O 20,8 % da mostra participou en máis de catro convocatorias na avaliación da proba escrita no sistema Celga e na proba oral entre catro e sete convocatorias.

			Valoración da expresión oral			Total
			Ata 3	Entre 4 e 7	Máis de 7	
Valoración da expresión escrita	Sen experiencia	Count	10	13	6	29
		% of Total	6,7 %	8,7 %	4,0 %	19,5 %
	Menos de 2 anos	Count	7	14	14	35
		% of Total	4,7 %	9,4 %	9,4 %	23,5 %
	3 anos	Count	0	12	15	27
		% of Total	0 %	8,1 %	10,1 %	18,1 %
	Máis de 4	Count	5	22	31	58
		% of Total	3,4 %	14,8 %	20,8 %	38,9 %
Total	Count	22	61	66	149	
	% of Total	14,8 %	40,9 %	44,3 %	100,0 %	

Táboa I. Experiencia dos avaliadores na valoración da expresión escrita segundo a súa experiencia na avaliación da expresión oral.

A poboación de examinadores do Celga presenta o seguinte perfil de formación:

a. *Formación xeral*: licenciados en Filoloxía Galega, licenciados en Tradución e Interpretación ou noutra licenciatura de filoloxía, sempre e cando cursasen materias de Lingua Galega.

b. *Formación específica* para impartir os cursos preparatorios destinados aos candidatos que se presentarán ás probas do Celga, ademais de habilitalos como avaliadores das probas escritas e orais. A formación recibida consiste nun seminario de 32 horas lectivas cuxos obxectivos son:

— Dotar os asistentes das ferramentas necesarias para a planificación e a avaliación no proceso de ensino e aprendizaxe da lingua galega, sempre dentro dun enfoque comunicativo, permitindo a presenza nas aulas dunha diversidade de actividades e oportunidades de aprendizaxe dirixidas ao desenvolvemento da competencia comunicativa do alumnado.

— Responder ás esixencias do contexto de ensino, aprendizaxe e avaliación do galego que precisa a sociedade actual.

— Reflexionar sobre o propio labor docente, como elemento dinamizador da formación do profesorado de galego.

— Contribuír ao coñecemento do sistema funcional e formal do galego e dotar o alumnado dun tratamento didáctico axeitado aos principios comunicativos do ensino de linguas.

Os contidos que se abordan no seminario de formación consisten en:

— Introducción teórica ao enfoque comunicativo. Presentación do Marco europeo común de referencia para as linguas, aprendizaxe, ensino e avaliación. Os novos certificados de lingua galega: Celga.

— Conceptos básicos de sociolingüística. Implicacións didácticas.

— Didáctica das destrezas: comprensión e expresión orais e escritas. Tratamento de actividades na aula para favorecer a competencia comunicativa en cada unha delas. Ensino por tarefas.

— Gramática pedagóxica. Estudo sobre casos conflictivos e o seu tratamento nas aulas.

— Didáctica do vocabulario.

— A avaliación e a autoavaliación na aprendizaxe.

3.1.4. Medidas

Enténdense como medidas aquelas variables que teremos en conta á hora de analizar o sistema Celga. Por unha banda, están as variables clasificatorias, que atinxen os diferentes niveis de lingua que forman o sistema de certificación, dende o Celga 1 ao Celga 4, e a experiencia dos avaliadores como membros dos equipos participantes nas convocatorias do Celga; por outra, aténdense os criterios de corrección e avaliación empregados polos avaliadores nas probas orais e escritas definidas a continuación.

3.1.4.1. Variables clasificatorias

a). Niveis do Celga

Tal e como avanzamos no capítulo 2, onde se presenta unha extensa explicación sobre os diferentes niveis de lingua, lembramos que o sistema de certificación Celga está dividido en catro niveis que se corresponden cos establecidos polo MCRL do seguinte xeito:

Celga 1	A2
Celga 2	B1
Celga 3	B2
Celga 4	C1

b). Experiencia dos avaliadores

O grupo de avaliadores que colabora coa SXPL é un colectivo que recibiu unha formación específica para participar nas convocatorias de exame do Celga, tal como avanzamos no apartado 3.1.1. (Mostra de avaliadores do Celga).

Para determinar a experiencia dos avaliadores empregouse o cuestionario administrado (anexo 1), no que tiñan que contestar á seguinte pregunta:

*1. Sinale o número de convocatorias do Celga nas que participou
-como avaliador da proba escrita:
-como avaliador da proba oral:*

Obtéñense así os datos sobre a súa participación na proba oral e escrita, xa que un avaliador que participe na proba oral non sempre participa como avaliador na proba escrita, de aí que moitos avaliadores teñan máis experiencia nas probas orais e menos nas escritas.

Á hora de seleccionar os avaliadores da proba oral, preténdese que todos os compoñentes da listaxe participen como avaliadores dalgún dos catro niveis de Celga, de tal xeito que interveñan polo menos nun dos días da convocatoria e como máximo en dous días. Lembremos que a convocatoria das probas Celga se divide en dúas fins de semana, é dicir, catro días nos que en cada un se avalía un nivel de Celga.

Para a asignación dos avaliadores da proba escrita, o criterio establecido é o de participar como tal nas últimas tres convocatorias. Normalmente cada avaliador debe corrixir unha media de trinta exames escritos, aínda que o número de avaliadores que participan na proba escrita tamén varía en función do número de candidatos que se presentan en cada convocatoria.

3.1.4.2. Criterios de corrección e avaliación das probas de expresión escrita

a). *Adecuación á tarefa*: mediante este parámetro valórase que o candidato se axuste ao tipo de tarefa que se lle esixe para completar a proba, é dicir, se se lle solicita que redacte un texto dunha determinada modalidade (expositivo, argumentativo ou narrativo) e cun determinado número de palabras (que varía segundo o nivel de Celga), que o cumpra.

b). *Coherencia e cohesión*: mediante este parámetro avalíase que o texto producido polo candidato presente unha estrutura adecuada na presentación de ideas, que use conectores que axuden a progresar a información e que empregue os mecanismos gramaticais e léxicos necesarios para o mantemento do referente. Consideramos:

A coherencia coma unha noción que inclúe as relacións pragmáticas ademais das relacións semánticas intratextuais. É un concepto que se refire ao significado do texto na súa totalidade, abarcando tanto as relacións das palabras co contexto coma as relacións entre as palabras no interior do mesmo texto. Alude á estabilidade e á consistencia temática subxacente, asociada á macroestrutura (contido), á superestrutura (esquema de organización) do texto, á súa ancoraxe enunciativa (protagonistas, tempo e espazo) e ás inferencias que activan os falantes para interpretalos a partir de coñecementos previos. A cohesión é un concepto que se refire a un dos fenómenos propios da coherencia, o das relacións particulares e locais que se dan entre os elementos lingüísticos, tanto os que remiten uns aos outros como os que teñen a función de conectar e organizar. Así, a coherencia, que inclúe a cohesión, pode considerarse tanto unha propiedade fundamental que dá cobertura ao conxunto de significacións dun texto coma un principio necesario para que o texto sexa definido como tal (Casalmiglia e Tusón, 2002: 221).

- a). *Gramática*: mediante este parámetro valórase a corrección morfolóxica e sintáctica no discurso do candidato e a súa adaptación á norma estándar do galego.
- b). *Léxico*: neste parámetro tense en conta a corrección, precisión e riqueza léxica, de acordo coas esixencias de cada nivel.

3.1.4.3. *Criterios de corrección e avaliación das probas de expresión oral*

- a). *Adecuación e interacción*: mediante este parámetro valórase que o candidato se axuste ao tipo de tarefa que se lle esixe para completar a proba, é dicir, se se lle solicita que produza un texto dunha determinada modalidade (expositivo, argumentativo ou narrativo).
- b). *Gramática*: mediante este parámetro valórase a corrección morfolóxica e sintáctica no discurso do candidato e a súa adaptación á norma estándar do galego.
- c). *Léxico*: neste parámetro valórase a corrección, precisión e riqueza léxica, de acordo coas esixencias de cada nivel.
- d). *Cohesión e fluidez*: neste parámetro valórase que o candidato estructure correctamente as súas ideas e que teña unha fluidez e un ritmo de fala acorde co nivel correspondente.
- e). *Pronuncia e entoación*: neste parámetro valórase que o usuario utilice na oralidade o sistema fonolóxico do galego.
- f). *Interacción*: neste parámetro valórase que o candidato colabore co interlocutor e manteña vivo o diálogo nas partes da proba que requiren interacción.

3.1.4.4. Validación do construto

Segundo Alderson *et al.* (1998: 21), cada proba ten unha teoría que a sustenta, isto é, unha crenza abstracta sobre que é a lingua e en que consiste ter competencia comunicativa nesa lingua, que implica aprender esa lingua e que poden facer os falantes con esa lingua. Todos eses conceptos conforman o construto, ou conceptos psicolóxicos, que son os principais compoñentes dunha teoría e entre eles establecen unha serie de relacións. Por exemplo, algunhas teorías sobre a lectura afirman que esta inclúe distintos construtos (lectura globalizadora ou *skimming*, lectura selectiva ou *scanning*).

A validación do construto supón examinar ata que punto unha proba avalía os seus construtos. Para poder levar a cabo a validación do construto é fundamental que un sistema de certificación desenvolva nas súas especificacións os fundamentos teóricos sobre os que se sustenta, que no caso do Celga veñen marcados polas pautas do MECRL e do sistema comunicativo no ensino, aprendizaxe e avaliación de linguas, ademais da relación entre a teoría e o obxectivo co cal se elaborou a proba.

No contido das especificacións do exame debe constar a información sobre que habilidades se avalían, sobre as interrelacións nesas habilidades, sobre que métodos de avaliación se empregarán e sobre que criterios se utilizarán para a avaliación. No caso do sistema de certificación Celga, por tratarse de probas baseadas en tarefas, as habilidades veñen determinadas pola integración das destrezas de comprensión e expresión oral e escrita que os candidatos deben manexar para resolver as tarefas solicitadas na proba. A partir disto, os avaliadores empregarán unha serie de ferramentas coñecidas como criterios de corrección e avaliación que no presente traballo analizaremos para medir a validez do seu construto.

3.1.4.5. Equidade

Tal e como avanzamos no apartado 3.1.4. (Medidas) deste capítulo, a equidade ou *fairness* esixe que unha proba sexa o máis xusta posible para os candidatos, polo que cómpre evitar os prexuízos culturais. Así, existen códigos de boas prácticas como o establecido por ALTE na versión en galego no que se recolle:

Esfuerzo pola imparcialidade

Os membros de ALTE tentan facer os exames coa maior imparcialidade posible para todos os candidatos de diferentes procedencias (por exemplo, raza, xénero, orixe étnica, discapacidade física ou mental, etc.). Na práctica, isto implica que os membros de ALTE se comprometen a facer o seguinte:

- Analizar e revisar as tarefas do exame e os materiais relacionados, para evitar contido ou linguaxe politicamente incorrecta.
- Establecer procedementos que axuden a asegurar que as diferenzas na actuación de candidatos están relacionadas principalmente coas destrezas que se están a avaliar e non con factores irrelevantes, como a raza, o xénero ou a orixe étnica.
- Sempre que sexa factible, facer os cambios pertinentes nos exames ou nos procedementos administrativos para lles facilitar a situación aos candidatos con algún tipo de discapacidade física ou mental. (ALTE, 2007: 2)

3.1.5. Instrumento

O instrumento elaborado *ad hoc* para valorar as opinións dos avaliadores sobre os criterios de corrección das probas do Celga é un cuestionario que consta de 16 ítems e está dividido en tres partes.

As dúas primeiras partes teñen como obxectivo avaliar a utilidade dos criterios de corrección dos niveis Celga para as probas de expresión escrita na primeira parte e para as probas de expresión oral na segunda parte. Os avaliadores podían escoller nunha escala de intervalo de 1 (a puntuación máis baixa) ao 10 (a máis alta).

A última parte está destinada a analizar a relevancia e a forma de corrección dos criterios utilizados para avaliar os candidatos tanto na proba

escrita coma na oral e nos diferentes niveis do Celga. Complementariamente, examinouse a opinión dos avaliadores sobre se faltaba ou sobraba algún criterio, así como a súa consideración xeral do Celga como sistema de avaliación. Para acadar estes obxectivos, utilizáronse once preguntas nominais de alternativa múltiple, unha pregunta filtro e dúas preguntas abertas (anexo I).

3.1.6. Procedemento

Para a creación do cuestionario utilizouse un comité de expertos composto por catro persoas. Tres delas teñen formación específica no ensino, aprendizaxe e avaliación de linguas, ademais dunha ampla experiencia neste eido; o cuarto membro é un psicólogo con formación en estatística e elaboración de instrumentos de medición cuantitativos e cualitativos, así como cunha grande experiencia en proxectos e publicacións de avaliación das competencias lingüísticas.

Inicialmente, elaboráronse a primeira e a segunda parte do cuestionario para, posteriormente, complementalo coa terceira. O método de traballo empregado consistiu no uso dun servidor privado ao que tiñan acceso os membros do comité e ao que achegaban as súas suxestións para a mellora do cuestionario de consulta para os avaliadores. O proceso finalizou cando ningún dos catro membros tiña ningunha cousa que engadir.

Para comprobar o proceso de administración da proba e o instrumento, e dicir, o cuestionario, realizáronse dúas probas piloto. Na primeira, utilizáronse tres avaliadores con grande experiencia aos que se lles solicitou que contestasen e, no caso de atoparen algún problema de comprensión ou calquera outra dificultade, que o reflectisen no corpo da mensaxe. O resultado desta aplicación do instrumento foi a modificación do enunciado de dous ítems para facelos máis comprensibles.

A segunda proba piloto do cuestionario consistiu na aplicación nunha mostra máis grande de avaliadores. O procedemento utilizado para a aplicación foi o envío masivo por parte da SXPL do cuestionario a todos os membros coñecidos da poboación. Os membros da mostra recibían, mediante correo electrónico, un documento adxunto en formato .doc que tiñan que completar e devolver ao mesmo enderezo de onde procedía o envío. Do total da poboación formada por 280 avaliadores, responderon 85, número suficiente para os nosos obxectivos, de aí que non se realizase un segundo envío recordatorio.

Ao analizar os cuestionarios devoltos, houbo que eliminar tres por ter unha porcentaxe moi alta de non resposta. Estes 82 avaliadores que quedaron eran o 29,29 % do total da poboación de avaliadores (Figura 4). A análise das respostas fixo modificar algunha das alternativas da última parte do cuestionario, debido aos problemas que presenta o envío dos cuestionarios con este tipo de preguntas, e procurouse un tipo de administración alternativa.

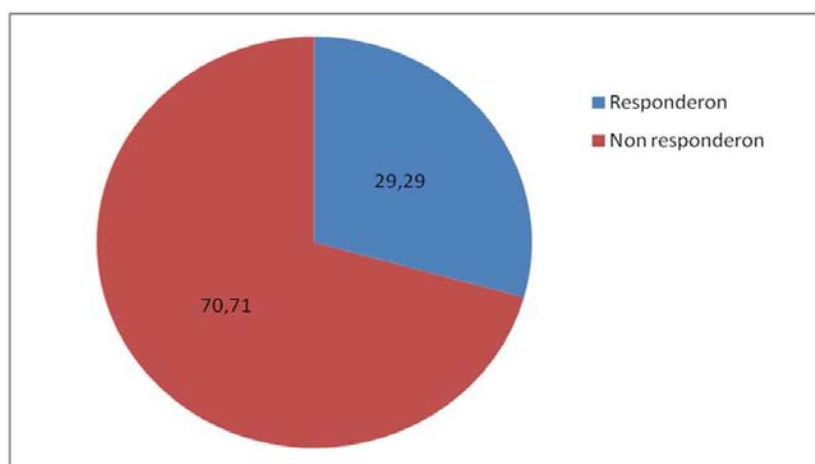


Figura 4. Mostra piloto de avaliadores do Celga

Unha vez definido o cuestionario, aproveitando unhas xornadas de formación organizadas pola SXPL para os avaliadores previas á convocatoria de xuño de 2012, administróuselles aos avaliadores o cuestionario e estes responderon de xeito anónimo. As xornadas de formación leváronse a cabo no

mes de xuño do 2012 en Santiago de Compostela e Vigo. O feito de elaborar un cuestionario presenta vantaxes sobre outros sistemas, xa que se fai de xeito anónimo e axiliza a administración, recollida e análise dos datos.

3.1.7. Análises estatísticas

O enfoque analítico polo que se optou para o tratamento das opinións dos avaliadores sobre os criterios do Celga parte dun uso complementario de técnicas de tipo descritivo inferencial segundo os obxectivos de investigación que se pretendía acadar. A seguir, detallamos as análises estatísticas utilizadas e explícase en que momento se utilizaron, así como os obxectivos que xustificaron a súa utilización. Cómpre salientar que a descrición do proceso de análise estatístico empregado non é de doada comprensión para os alleos á psicometría, polo tanto agardamos que a lectura deste apartado sexa entendida como un intento de clarificar un procedemento altamente técnico que require dun coñecemento sobre psicometría que non se espera do lector do presente traballo.

Na primeira parte do estudo utilizáronse estatísticos descritivos (medias e frecuencias) para coñecer as opinións dos avaliadores sobre a utilidade e validez dos criterios de avaliación, tanto da expresión oral coma da escrita, así como da equidade do procedemento de avaliación do Celga. Para sondar a relación entre as variables dependentes de tipo de intervalo e variable nivel do Celga con máis de dous niveis de resposta, utilizouse a técnica paramétrica de análise de varianza dun factor, xa que se cumprían os requisitos de normalidade e homoxeneidade da varianza das submostras cos contrastes *a posteriori* Tukey b.

Para afondar na relación entre a experiencia profesional e a utilidade dos criterios de avaliación, utilizouse unha proba non paramétrica paralela á

análise de varianza, xa que estes non cumprían os requisitos de normalidade e homoxeneidade. A proba escollida aplicada foi a H de Krustal-Wallis. Para afondar nas relacións entre os diferentes niveis de resposta, adminístrase a proba U de Mann Whitney, acompañada das correccións conservadoras de Bonferroni segundo a fórmula: $p' = 1 - (1 - p)^c$.

Escolleuse a proba non paramétrica coeficiente de incerteza para as variables de medidas mediante escalas de tipo nominal (equidade). Trátase dunha medida baseada na redución da proporción do erro que se amosa máis robusta ca outros contrastes de uso habitual, como o lambda, xa que reduce a posibilidade de efectuar prognósticos erróneos ao considerar non só as probabilidades de que os casos pertencen á categoría máis probable senón a todas as categorías dispoñibles.

Escolleuse a técnica multivariante Individual Differences Scalling (INDSCAL) para sondar a estrutura que subxace á valoración dos correctores das persoas que se presentan para obter o certificado de lingua galega atendendo ao nivel do Celga. Esta técnica proporciónanos unha representación gráfica dimensional (mapa perceptual) dos criterios de avaliación, tanto oral como escrita, que consistentemente se utilizan para a valoración dos participantes (para ver unha descrición máis exhaustiva da técnica, Valderrey, 2010).

O programa utilizado para a codificación dos datos e a súa posterior análise foi o SPSS 13.0.

3.1.8. Utilidade dos criterios de corrección e avaliación das probas de expresión escrita segundo os avaliadores

A continuación preséntanse (Figura 5) os resultados dos niveis de utilidade adxudicados polos avaliadores aos criterios de corrección e avaliación

propostos polo grupo de investigación do Celga, utilizando as puntuacións medias como puntuación descritiva. En xeral, obsérvase que os avaliadores consideran que estes criterios son altamente útiles para a función que desempeñan, xa que non hai ningún criterio que teña unha media de utilidade inferior a 7 e o intervalo de puntuacións oscila entre 3 e 10 en todos os criterios, excepto o de gramática que vai desde 2 a 10. A porcentaxe de avaliadores que lles adxudican un suspenso¹⁰ aos criterios de utilidade non son moitos, xa que non chegan ao 9% en adecuación á tarefa en utilidade (8,8% e 8,6%, respectivamente) e ao 13,5% en gramática e léxico (12,3% e 13,1%, respectivamente).

Atendendo ás puntuacións medias en utilidade adxudicadas polos avaliadores, pódense dividir os criterios en dous grupos: un primeiro considerado máis útil e composto por adecuación á tarefa e coherencia e cohesión, e outro que, aínda que se considera que ten moita utilidade, ten algunha menos ca o anterior e está composto por gramática e léxico.

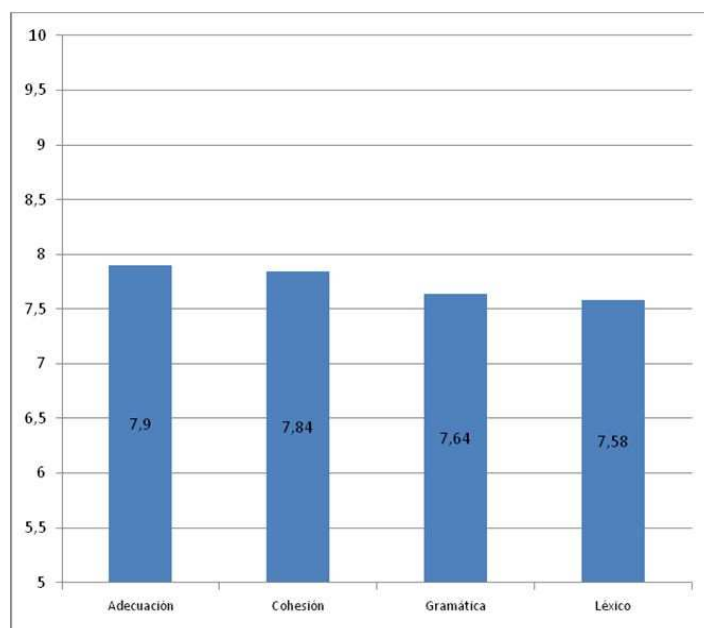


Figura 5. Nivel de utilidade dos criterios de corrección e avaliación para as probas de expresión escrita do CELGA.

¹⁰ Considera 'suspender' adxudicar unha puntuación inferior ao 5,5 na escala de utilidade.

3.1.8.1. Segundo o nivel do Celga

Seguindo a liña dos resultados anteriores, obsérvase que as puntuacións medias de utilidade adxudicadas atendendo ao nivel do Celga varían substancialmente. Os resultados amosan que a utilidade aumenta a medida que o nivel do Celga aumenta, é salientable a importancia que se adxudica aos criterios no nivel Celga 4 fronte ao resto. Cómpre sinalar que os criterios de avaliación se consideran útiles en todos os niveis da variable (Figura 6).

As puntuacións oscilan entre a máis baixa, adxudicada ao criterio léxico no nivel Celga 1 (7,15), e a máis alta, dos criterios de coherencia e cohesión e adecuación á tarefa do Celga 4 (8,32).

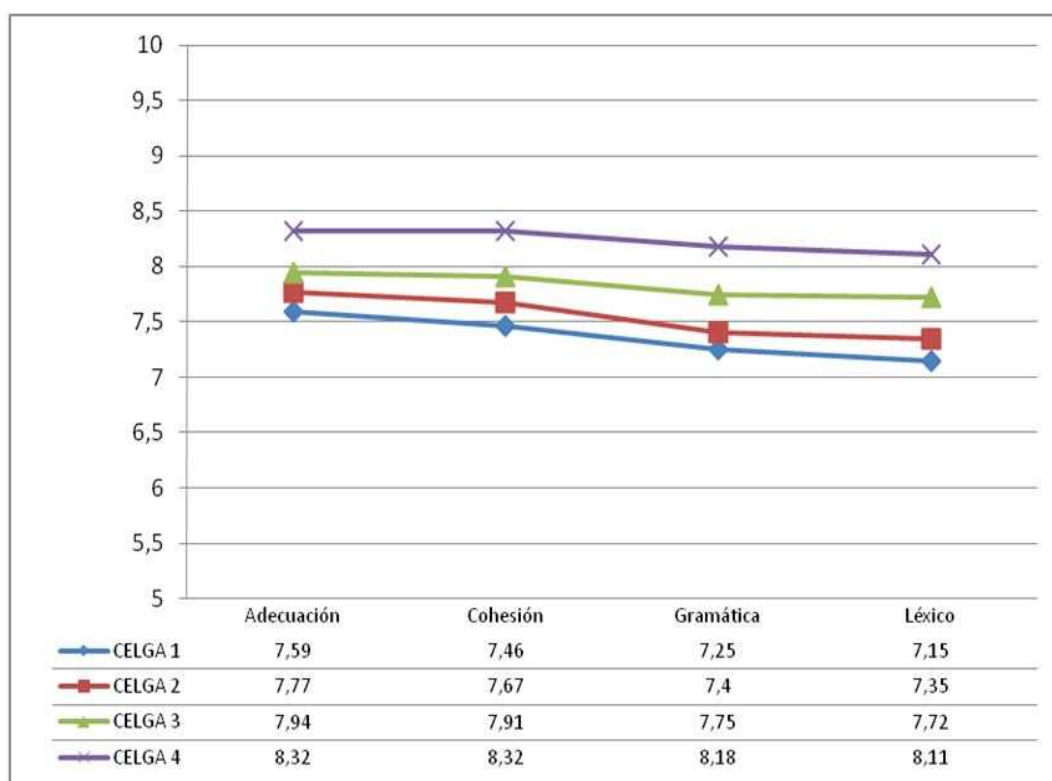


Figura 6. Nivel de utilidade dos criterios de corrección e avaliación para as probas de expresión escrita segundo os niveis dos Celga.

Por outra banda, as valoracións sobre a utilidade dos criterios de corrección e avaliación varían dun xeito relevante atendendo aos distintos niveis do Celga. Ao analizar cada un deles criterios individualmente e os niveis diferentes do Celga, obsérvase a existencia de diferenzas significativas que se detallan na táboa 2.

Os resultados que reflicte a análise individual dos criterios de avaliación segundo o nivel do Celga ao que se presente o avaliado, utilizando a análise de varianza con contraste *a posteriori* Tukey b, indican que:

— A adecuación á tarefa se considera máis útil para o Celga 4 ca para o Celga 1 e Celga 2.

— Os criterios de coherencia e cohesión, gramática e léxico se consideran máis útiles para o nivel Celga 4 ca para os primeiros –Celga 1 e Celga 2–. Así mesmo, estes criterios son considerados máis útiles no Celga 2 ca no Celga 1.

Agrupando os resultados individuais, obtemos un panorama xeral que indica a existencia de dous perfís claros, un deles composto polo Celga 4, no que todos os criterios para a valoración dos suxeitos se consideran máis útiles ca nos niveis máis baixos (Celga 1 e Celga 2). Por outra banda, teríamos un perfil intermedio (Celga 3) no que as valoracións son semellantes ás do Celga 2 e ás do Celga 4, aínda que se diferenciaría claramente do Celga 1 en todos os criterios, agás no de adecuación á tarefa.

	F	Graos de liberdade	P	<i>Post hoc</i> Tukey b
Adecuación	5,958	552	p = 0,001	Celga 1*Celga 4 Celga 2*Celga 4
Cohesión	8,547	552	p < 0,001	Celga 1*Celga 3 Celga 1*Celga 4 Celga 2*Celga 4
Gramática	9,706	552	p < 0,001	Celga 1*Celga 3 Celga 1*Celga 4 Celga 2*Celga 4
Léxico	10,414	552	p < 0,001	Celga 1*Celga 3 Celga 1*Celga 4 Celga 2*Celga 4

Táboa 2. Diferenzas nas valoracións de utilidade dos criterios de corrección escrita atendendo ao nivel do Celga. Resultados da aplicación da proba Anova Oneway e contraste *a posteriori* Tukey b.

3.1.8.2. Segundo a experiencia dos avaliadores

Para analizar se a experiencia dos avaliadores influía no proceso de avaliación aplicouse a proba non paramétrica chamada H Krustal Wallis, da que se deduce que esta variable non inflúe significativamente sobre a percepción de utilidade de ningún dos criterios de corrección e avaliación: adecuación ($K = 7,551$; $p = 0,056$), coherencia e cohesión ($K = 5,543$; $p = 0,136$), gramática ($K = 0,673$; $p = 0,880$) e léxico ($K = 0,106$; $p = 0,991$).¹¹

¹¹ Non se aplicou a proba *a posteriori* U de Mann Whitney para comprobar as diferenzas entre os diferentes niveis da variable experiencia.

3.1.9. Utilidade dos criterios de corrección e avaliación das probas de expresión oral segundo os avaliadores

Da mesma forma que no apartado anterior se presentan os resultados dos niveis de utilidade adxudicados polos avaliadores aos criterios de corrección e avaliación, neste apartado presentamos os resultados sobre a utilidade para as probas de expresión oral. Mentres que para a expresión escrita se utilizaban catro criterios, para as probas de expresión oral engádense dous criterios pola especificidade da competencia oral (pronuncia e interacción).

A figura seguinte (Figura 7) presenta os resultados dos niveis de utilidade adxudicados a cada criterio sen ter en conta o nivel do Celga ao que se presenta o candidato. De igual forma que para a expresión escrita os criterios se consideran útiles porque a media non baixa en ningún dos casos de 7, os intervalos de puntuación oscilan entre 2 e 10 no caso de adecuación e pronuncia e 3 e 10 no resto dos criterios.

En xeral, obsérvase que os correctores consideran que estes criterios son útiles para a función que desempeñan, xa que non existe ningún criterio que teña unha media de utilidade inferior a sete; o intervalo de puntuacións oscila entre 3 e 10 en todos os criterios, agás o de gramática que vai desde 2 a 10. A porcentaxe de avaliadores que lles adxudican un suspenso¹² aos criterios de utilidade non son moitos: adecuación á tarefa (11,7 %), coherencia e cohesión (13,6 %), léxico e á gramática (15,2 % e 15,1 %, respectivamente), pronuncia (18,3 %).

Atendendo ás puntuacións medias en utilidade adxudicadas polos avaliadores, obsérvase que os criterios que obteñen unha maior puntuación son interacción e coherencia e cohesión (7,67 e 7,6, respectivamente) fronte a

¹² Véxase nota 10.

pronuncia, que acada a puntuación máis baixa (7,17), o resto obtén puntuacións moi semellantes.

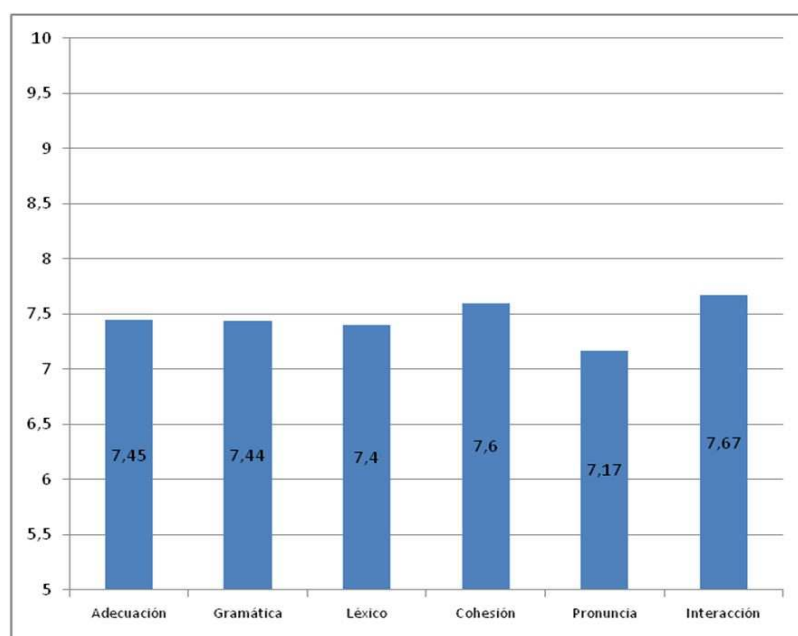


Figura 7. Nivel de utilidade dos criterios de corrección e avaliación para as probas de expresión oral do Celga.

3.1.9.1. Segundo o nivel do Celga

Ao analizar os resultados de utilidade dos criterios de corrección para expresión oral en función do nivel do Celga, obsérvase que se produce o mesmo fenómeno ca na expresión escrita: a medida que aumenta o nivel do Celga aumenta a importancia que lle adxudican os avaliadores a estes (Figura 8). A puntuación máis baixa é inferior ao seis e medio (6,45) e correspóndese coa pronuncia no Celga 1, é dicir, os avaliadores non o perciben como excesivamente útil para o primeiro nivel, circunstancia que se repetirá no Celga 2, e que despois se atenuará segundo aumenta o nivel do Celga. Por outra banda, os avaliadores consideran que os criterios máis útiles son os de adecuación á tarefa, interacción e coherencia e cohesión (8,27, 8,26, 8,22,

respectivamente), que se corresponden co Celga 4. Cómpre sinalar que en niveis máis baixos as puntuacións máis altas por nivel se refiren ao criterio de adecuación á tarefa.

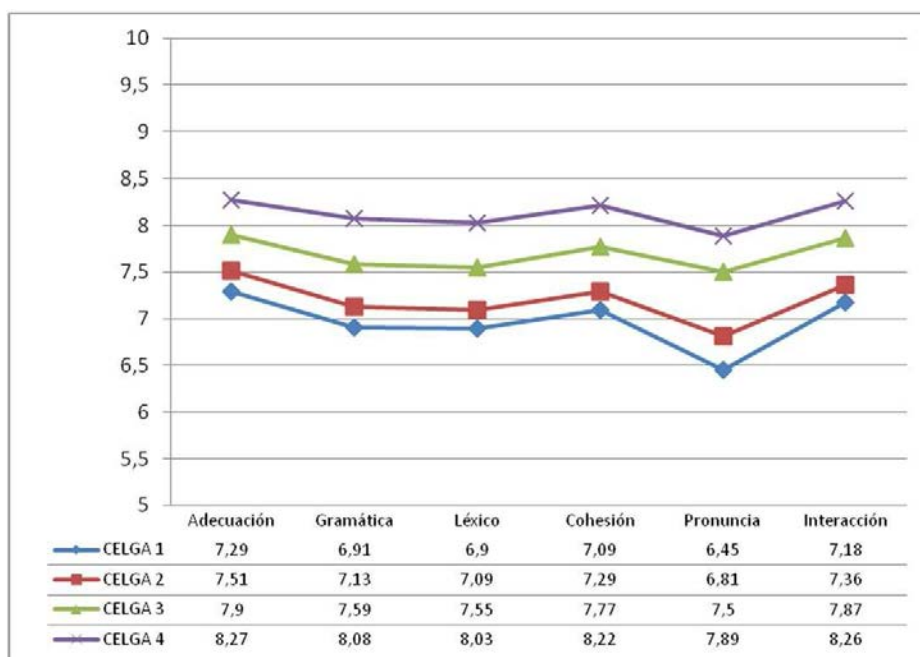


Figura 8. Nivel de utilidade dos criterios de corrección e avaliación para as probas de expresión oral dos distintos niveis Celga.

Os resultados que reflicte a análise individual dos criterios de avaliación segundo o nivel do Celga ao que se presente o candidato, utilizando a análise de varianza con contraste *a posteriori* Tukey b, indican que:

— A adecuación á tarefa se considera máis útil para o Celga 4 ca para o Celga 1 e 2.

— Os criterios de coherencia e cohesión, gramática e léxico se consideran máis útiles para Celga 4 ca para os primeiros (Celga 1 e Celga 2) e todos eses criterios son máis útiles para o Celga 3 ca para o Celga 1.

	F	Graos de liberdade	P	Post hoc Tukey b
Adecuación	10,938	577	$p < 0,001$	Celga 4* Celga 1 Celga 4* Celga 2 Celga 3* Celga 1
Gramática	14,902	577	$p < 0,001$	Celga 4* Celga 3 Celga 4* Celga 2 Celga 4* Celga 1 Celga 3* Celga 2 Celga 3* Celga 1
Léxico	13,948	576	$p < 0,001$	Celga 4* Celga 3 Celga 4* Celga 2 Celga 4* Celga 1 Celga 3* Celga 2 Celga 3* Celga 1
Cohesión	15,894	575	$p < 0,001$	Celga 4* Celga 3 Celga 4* Celga 2 Celga 4* Celga 1 Celga 3* Celga 2 Celga 3* Celga 1
Pronuncia	20,741	575	$p < 0,001$	Celga 4* Celga 1 Celga 4* Celga 2 Celga 3* Celga 2 Celga 3* Celga 1
Interacción	13,058	578	$p < 0,001$	Celga 4* Celga 1 Celga 4* Celga 2 Celga 3* Celga 2 Celga 3* Celga 1

Táboa 3. Diferenzas nas valoracións de utilidade dos criterios de corrección oral atendendo ao nivel do Celga. Resultados da aplicación da proba Anova Oneway e contraste *a posteriori* Tukey b.

3.1.9.2. Segundo a experiencia profesional dos avaliadores

A experiencia dos avaliadores nas probas orais inflúen significativamente na percepción de utilidade en cinco dos seis criterios de avaliación utilizados nas probas orais: adecuación ($K = 20,352$; $p < 0,001$), gramática ($K = 14,015$; $p = 0,001$), coherencia ($K = 26,839$; $p < 0,001$), pronuncia ($K = 14,378$; $p = 0,001$) e interacción ($K = 14,087$; $p = 0,001$); non inflúe no criterio de léxico ($K = 2,250$; $p = 0,284$).

A continuación acudiuse á proba U-Mann-Whitney e ao cálculo manual da proba de Bonferroni para perfilar o patrón de diferenzas considerando as variables dúas a dúas. Os resultados indicaron que:

a) medida que aumentaba a experiencia dos avaliadores, estes percibían que a coherencia e a cohesión tiñan máis utilidade. Existen diferenzas entre os que teñen menos de tres anos de experiencia fronte aos que teñen entre catro e sete ($p' = 0,004$) e os que teñen máis de sete ($p' < 0,001$) e entre os que teñen entre catro e sete fronte aos que teñen máis de sete ($p' = 0,006$).

b) Nos criterios de adecuación á tarefa, gramática e interacción había diferenzas significativas entre os correctores con menos de tres anos de experiencia e máis de sete ($p' < 0,001$; $p' = 0,008$; $p' = 0,002$) e entre os que tiñan entre catro e sete anos de experiencia e máis de sete anos de experiencia ($p' = 0,002$; $p' = 0,002$; $p' = 0,006$). Todos ían na mesma dirección, a medida que tiñan máis experiencia percibían que estes criterios tiñan máis importancia.

c) No criterio de interacción había diferenzas significativas entre os avaliadores con menos de tres anos de experiencia e os de máis de sete anos de experiencia ($p' < 0,001$).

3.1.10. Opinión sobre a validez dos criterios de corrección e avaliación

Na segunda parte do cuestionario, pedíáselles aos avaliadores que escollesen entre os criterios de corrección e avaliación que considerasen menos válidos á hora de determinar a competencia comunicativa tanto escrita coma oral (preguntas 10 e 11 do anexo 1).

Como resultados sobre os criterios menos válidos na proba escrita, máis da metade (59,6 %) dos avaliadores optaba por escoller a alternativa “e” que consideraba a todas as alternativas como válidas (Todos son igual de válidos).

Non obstante, hai unha porcentaxe de avaliadores que considera o criterio adecuación á tarefa como o menos importante (21,9 %), o resto de alternativas obtiveron puntuacións inferiores a 9. Hai que considerar que só cinco avaliadores escolleron o criterio de coherencia e cohesión como o menos válido (Figura 9).

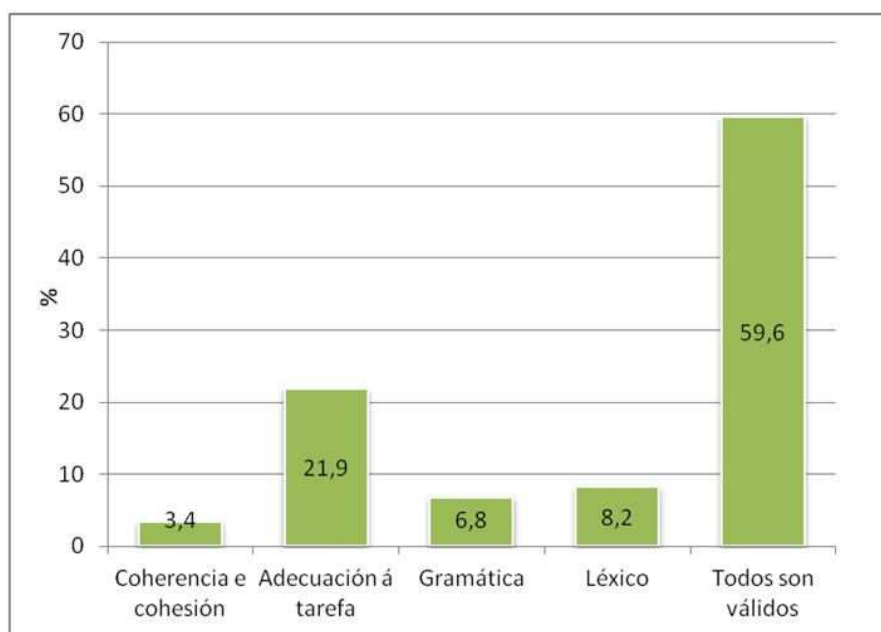


Figura 9. Criterio de corrección e avaliación que lles parecía menos válido á hora de determinar a competencia comunicativa na proba escrita.

Existe una relación significativa entre a experiencia profesional dos avaliadores e a elección dos criterios que consideran como válidos ou non válidos á hora de determinar a competencia comunicativa ($I = 0,025$; $p = 0,018$). En xeral, pódese inferir que os avaliadores cunha maior experiencia (isto é, máis de catro convocatorias) son os que consideran en maior medida que todos os criterios son válidos (63,2 %) fronte aos correctores con experiencia en tres convocatorias ou menos de dúas (55,6 % e 55,9 %, respectivamente). Hai un grupo que non concorda coa tendencia, que é dos avaliadores sen experiencia, o cal, curiosamente, coincide nas respostas cos

avaliadores máis experimentados (61,5 % os avaliadores sen experiencia e 63,2 % os avaliadores con máis de catro convocatorias). A causa desta disonancia é que este grupo está composto fundamentalmente por persoas con moita experiencia na avaliación da proba oral, de aí que poidan compartir os mesmos criterios ca as persoas con experiencia da proba escrita.

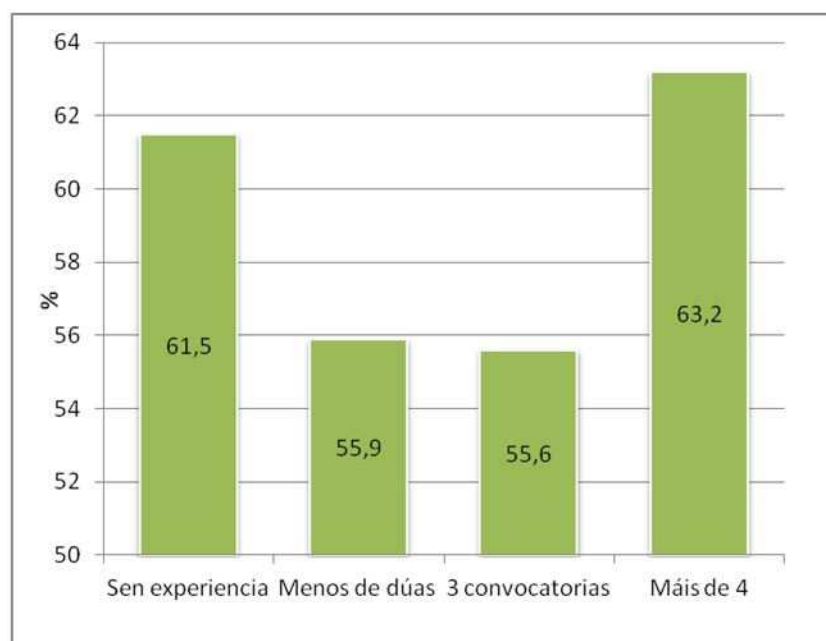


Figura 10. Avaliadores que elixiron a alternativa “todos son válidos” á hora de determinar a competencia comunicativa escrita segundo o nivel de experiencia.

Cando se lles solicitou aos avaliadores unha valoración sobre os criterios de corrección menos válidos para avaliar a competencia oral (na pregunta 11 do anexo 1), a metade (51,4 %) escollía as alternativa “todos son igual de válidos”. O criterio escollido por unha porcentaxe maior de avaliadores como menos válido foi o de pronuncia e entoación (17,6 %), mentres os criterios que se elixiron un menor número de veces foron cohesión e fluidez e gramática (6,8 % nos dous casos).

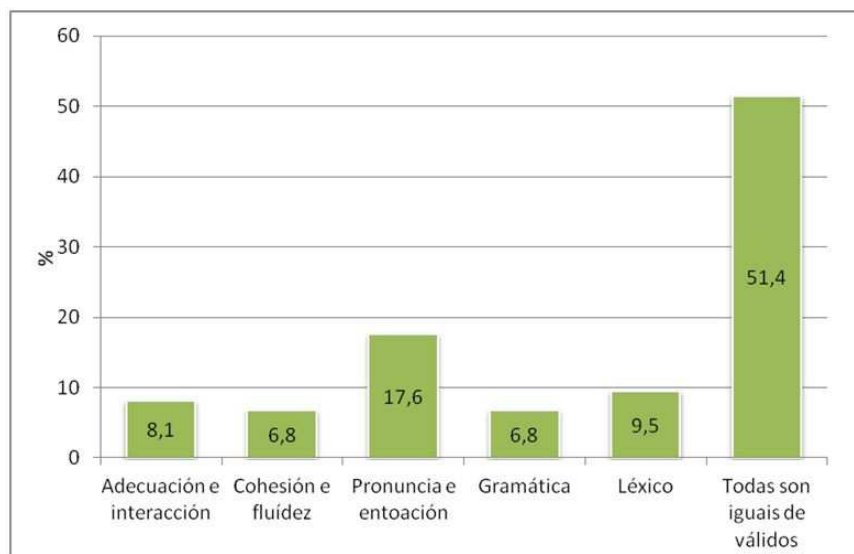


Figura 11. Criterio de corrección e avaliación que lles parecía menos válido á hora de determinar a competencia comunicativa na proba oral.

A experiencia profesional ten unha relación directa na determinación dos criterios máis válidos á hora de avaliar a competencia comunicativa na proba oral ($I = 0,017$; $p = 0,012$). Como se pode ver na Figura 11, a medida que aumenta a experiencia na avaliación da competencia oral, diminúe a consideración sobre que todos os criterios son válidos, pasando do 61,9 % nos avaliadores con ata tres convocatorias ao 45,5 % nos que teñen máis de sete convocatorias.

Centrándonos nos avaliadores que conforman o grupo de máis experiencia, estes concédennlle máis importancia ao criterio de cohesión e fluidez, soamente un 3 % non o consideran válido; pola contra, o criterio de pronuncia e entoación segue a ser o que se considera menos válido, cunha porcentaxe notable (19,7 %).

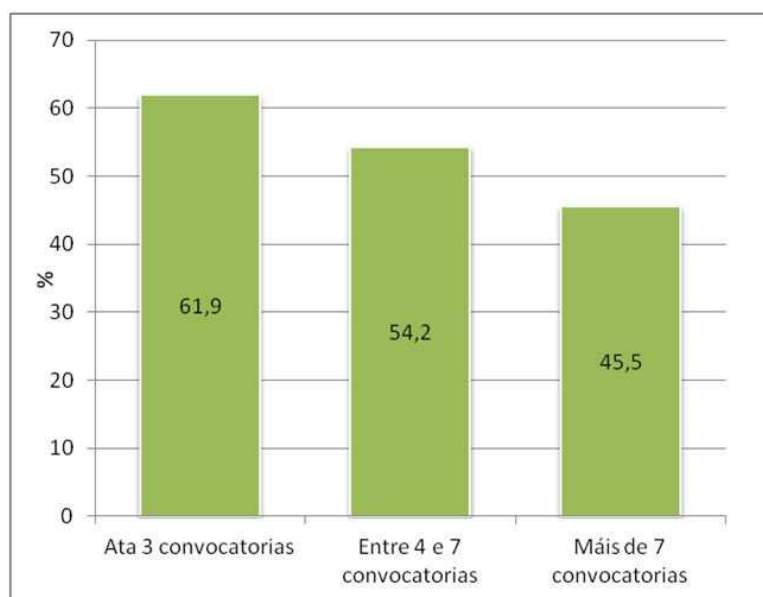


Figura 12. Avaliadores que elixiron a alternativa “todos son válidos” á hora de determinar a competencia comunicativa oral segundo o nivel de experiencia

É salientable que, aínda que os criterios de corrección teñen unha consideración de maior ou menor utilidade, case a totalidade dos avaliadores (94,4 %) considera todos os criterios de corrección e avaliación importantes, xa que os enquisados opinan que non se debería eliminar ningún (pregunta 13 do anexo 1) (Figura 13).

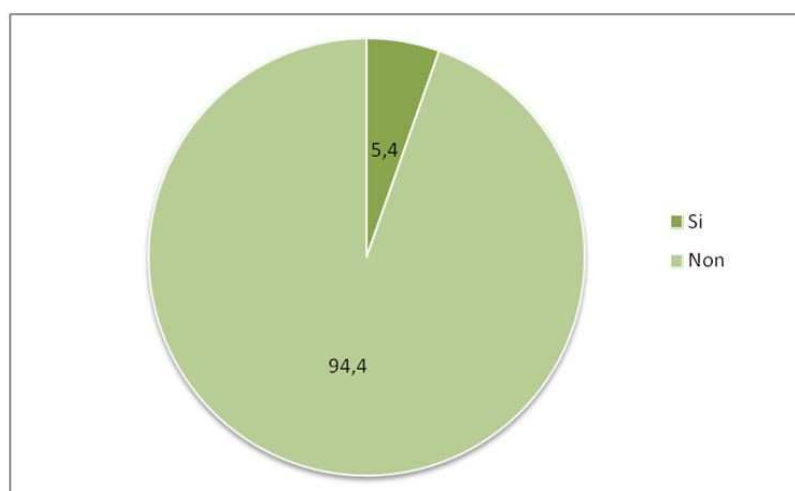


Figura 13. Considera que se debería eliminar algún dos criterios de corrección e avaliación?

Existe unha relación significativa entre a opinión sobre a eliminación de criterios na avaliación da proba escrita e a experiencia dos avaliadores ($I = 0,072$; $p = 0,032$). Mentres que a totalidade dos avaliadores sen experiencia consideran que non se debería eliminar ningún, un 9,1 % opina que si se debería eliminar algún (Figura 13).

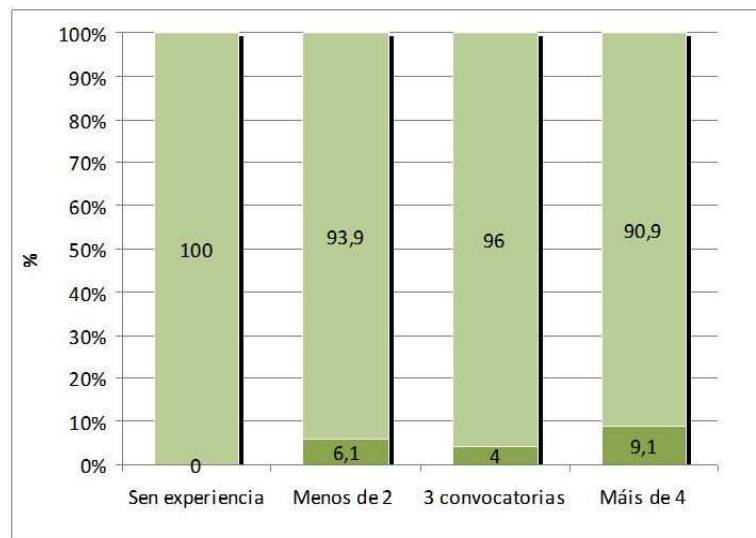


Figura 14. Eliminación dalgún criterio de avaliación da competencia escrita atendendo á experiencia.

Entre a opinión sobre a eliminación de criterios na avaliación da proba oral (pregunta 14 do anexo 1) e a experiencia dos avaliadores tamén existe unha relación significativa aínda que máis forte ca no caso da proba escrita ($I = 0,182$; $p = 0,044$). Ao contrario do que sucedía na proba escrita, a medida que aumenta a experiencia dos avaliadores, a opinión é máis contraria á eliminación dalgún criterio de avaliación (Figura 14). Un 15 % deles con menos experiencia opina que se debería eliminar algún criterio, mentres que a totalidade dos avaliadores con máis experiencia pensan que non se debería eliminar ningún.

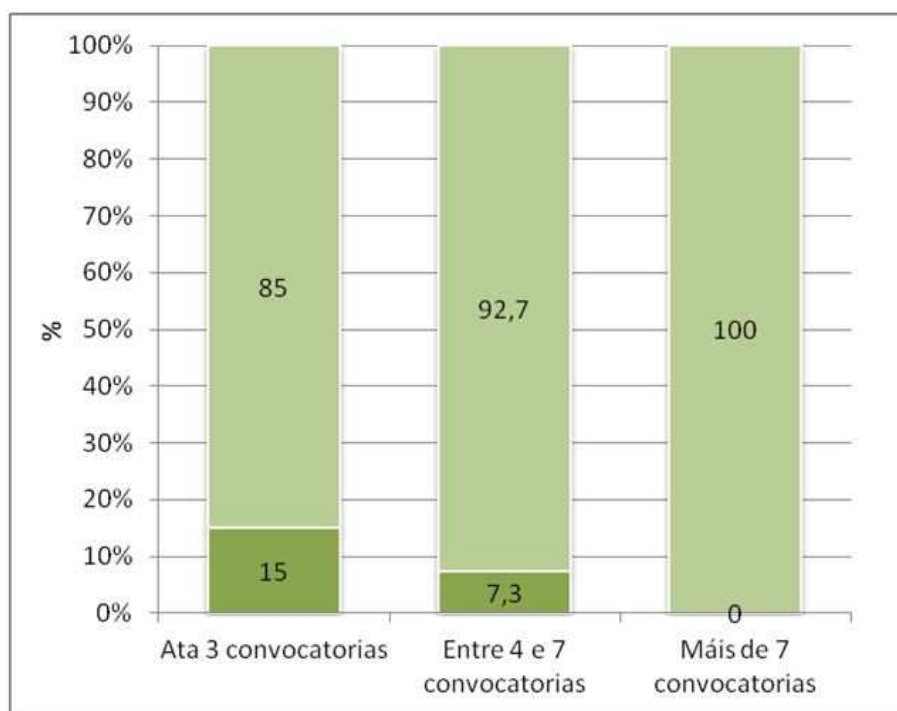


Figura 15. Eliminación dalgún criterio de avaliación da competencia oral atendendo á experiencia.

Para aclarar que criterios son susceptibles de eliminación segundo a opinión dos avaliadores, indicóuselles que, se consideraban que debían descartar algún criterio, especificasen cal sería tanto para a proba oral coma para a escrita (pregunta 14 do anexo 1).

Como se pode ver na Figura 16, soamente cinco persoas de 152, é dicir, un 0,69 % da mostra, sinalaron que debería eliminarse algún criterio da avaliación da proba escrita. Estes cinco avaliadores decantáronse por dous criterios: adecuación á tarefa e coherencia e cohesión.

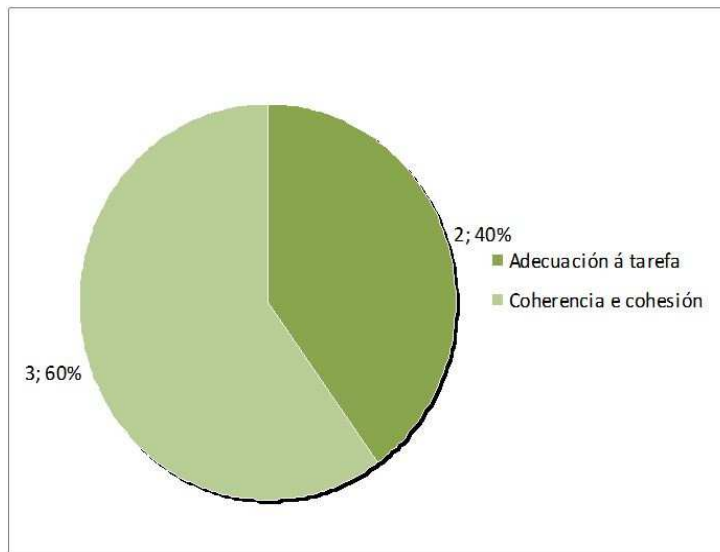


Figura 16. Criterio de corrección e avaliación que se debería eliminar na proba escrita. Frecuencias e porcentaxes.

Por outra banda, nove avaliadores escollen un criterio de corrección e avaliación para eliminar na proba oral (Figura 17), algo máis ca no caso da proba escrita, pero que segue a ser unha porcentaxe da mostra pequena 5,92 % (n = 9). Case a metade destas persoas (n = 4) escolle eliminar adecuación e interacción.

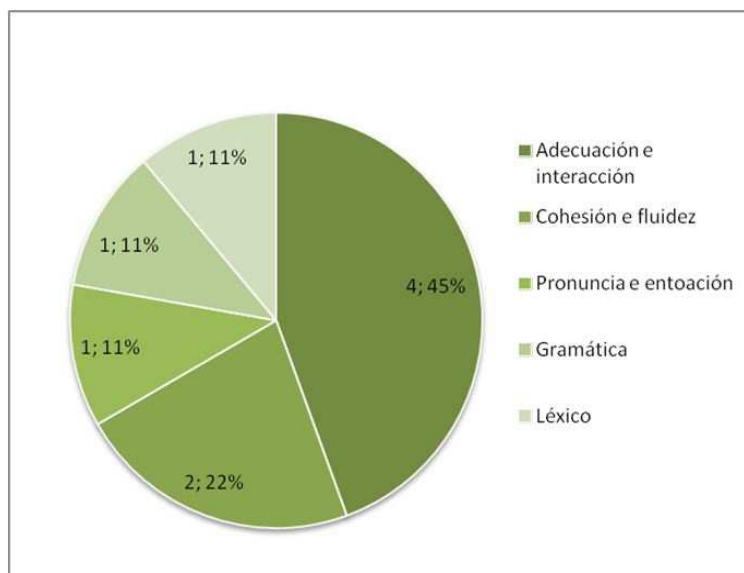


Figura 17. Criterio de corrección e avaliación que se debería eliminar na proba oral. Frecuencias e porcentaxes

Sistema de certificación de lingua galega (Celga) no *Marco Común Europeo de Referencia para as Linguas*. Estudo de fiabilidade e validez

Para complementar as opinións sobre a validez dos criterios de corrección e avaliación, preguntábaselles aos correctores se habería que incluír algún criterio máis (pregunta 12 do anexo 1). Un 12,2 % ofreceu unha alternativa a maiores das xa incluídas no proceso de avaliación e un 2,6%, dúas alternativas.

A continuación detállanse as alternativas que os avaliadores consideraron como criterios de corrección e avaliación que se deben incluír:

p12a

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 88	135	88,8	88,8	88,8
a adecuación sociolingüística	1	,7	,7	89,5
actitude cara ao galego	1	,7	,7	90,1
adequación ao rexistro	1	,7	,7	90,8
coherencia asociada á cohesión	1	,7	,7	91,4
falante en galego	1	,7	,7	92,1
formación da persoa	1	,7	,7	92,8
galegofalantes e non galegofalantes	1	,7	,7	93,4
nivel formación	1	,7	,7	94,1
ponderación dalgúns criterios	1	,7	,7	94,7
pronuncia/entoación	1	,7	,7	95,4
reducir marxe do número de palabras a un 10 %	1	,7	,7	96,1
riqueza	1	,7	,7	96,7
ser galegofalante	1	,7	,7	97,4
ser galego funcional	1	,7	,7	98,0
ser galegofalante	1	,7	,7	98,7
situación sociolingüística	1	,7	,7	99,3
uso habitual do galego	1	,7	,7	100,0
Total	152	100,0	100,0	

p12 b

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 88	148	97,4	97,4	97,4
galego falantes nativos	1	,7	,7	98,0
nivel cultural	1	,7	,7	98,7
non recoñecer nivel cultural	1	,7	,7	99,3
variacións dialectais	1	,7	,7	100,0
Total	152	100,0	100,0	

Figura 18a. Considera que hai algún/s criterio/s de corrección e avaliación que se debería/n engadir e que non se ten/teñen en conta á hora de avaliar os exames? Resposta directa

		Frecuencia	Porcentaxe
Non engade resposta			
Si engade resposta	Categoría I Sociolingüística	10	
	Categoría II Sociolóxica	4	
	Categoría III Textual	7	
Total	152 avaliadores	Frecuencia	Porcentaxe

Figura 18. Pregunta 12: considera que hai algún/s criterio/s de corrección e avaliación que se debería/n engadir e que non se ten/teñen en conta á hora de avaliar os exames?

Como se pode observar nas táboas anteriores (correspondentes á pregunta 12 do anexo 1), os avaliadores á hora de engadir outros criterios de corrección e avaliación presentan unha serie de posibilidades que están lonxe das características que teñen os actuais criterios, centrados en aspectos que conforman un texto, fronte a cuestións que poden agruparse nas seguintes categorías:

— De orde sociolingüística, como ser galegofalante, ser falante nativo, situación sociolingüística, uso habitual do galego, ser galego funcional, non ser galegofalante, adecuación sociolingüística, actitude cara ao galego .

— De carácter sociolóxico, como nivel de formación da persoa.

— De orde textual, como adecuación ao rexistro, coherencia asociada á cohesión, pronuncia e entoación, reducir a marxe de palabras a un 10 %, riqueza.

Pola contra, na táboa p12b (obtida da pregunta 12 do anexo 1) obsérvase que as alternativas que ofrecen son, por unha banda, de carácter sociolingüístico, como galegofalantes nativos, variacións dialectais, e pola outra de carácter sociolóxico, como nivel cultural, non recoñecer nivel cultural.

Polo tanto, enténdese que á hora de establecer novos criterios de corrección e avaliación, existe unha certa confusión sobre cuestións de carácter meramente textual e doutra orde diferente, como a sociolóxica ou a sociolingüística, que, aínda que tamén forman parte do ecosistema da competencia comunicativa, pode que non sexan tan útiles como ferramentas á hora de avaliar a competencia comunicativa nunha produción textual oral ou escrita.

3.1.11. Valoración da equidade do Celga

Ao preguntarles aos colaboradores se pensan que as probas Celga teñen un sistema xusto de avaliación, a maior parte deles considera que é equitativa

(74,5%). A pesar desta maioría que cre que o procedemento de avaliación é xusto, existe unha porcentaxe significativa de avaliadores que non opina o mesmo (25,5 %).

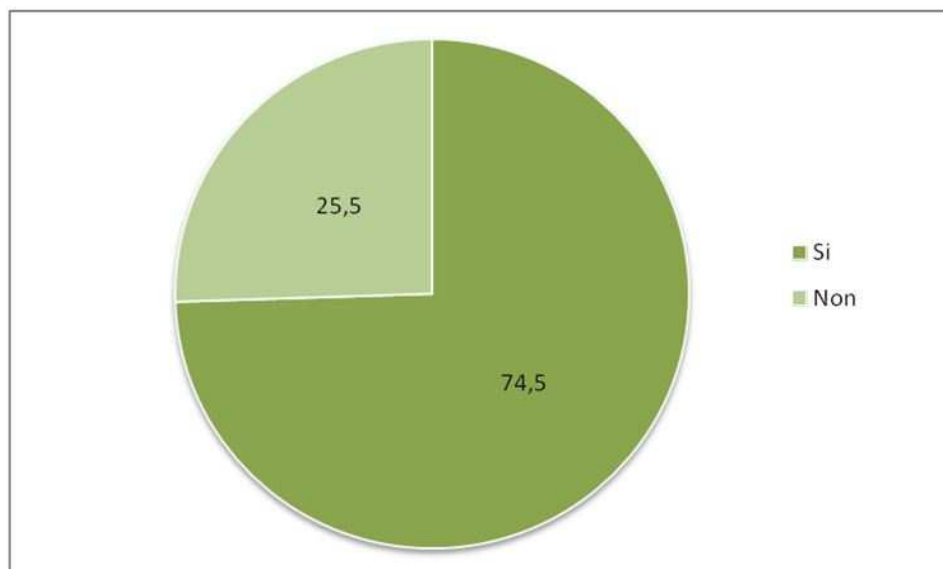


Figura 19. Opinión sobre a equidade do Celga

Ao analizar a relación entre a experiencia dos avaliadores na proba escrita e a opinión sobre a equidade do sistema, observouse unha variación significativa ($I = 0,015$; $p = 0,019$). Sen considerar os avaliadores sen experiencia, pode observarse que, a medida que aumenta a experiencia como avaliadores na proba escrita, incrementase a percepción de equidade do sistema, pasando do 35,3 % daqueles con menos de dúas convocatorias ao 23,2 % dos avaliadores con máis de catro convocatorias. Semella que os incluídos na categoría sen experiencia son suxeitos con moita experiencia na avaliación oral e ningunha na proba escrita, de aí as cifras tan disonantes (Figura 19).

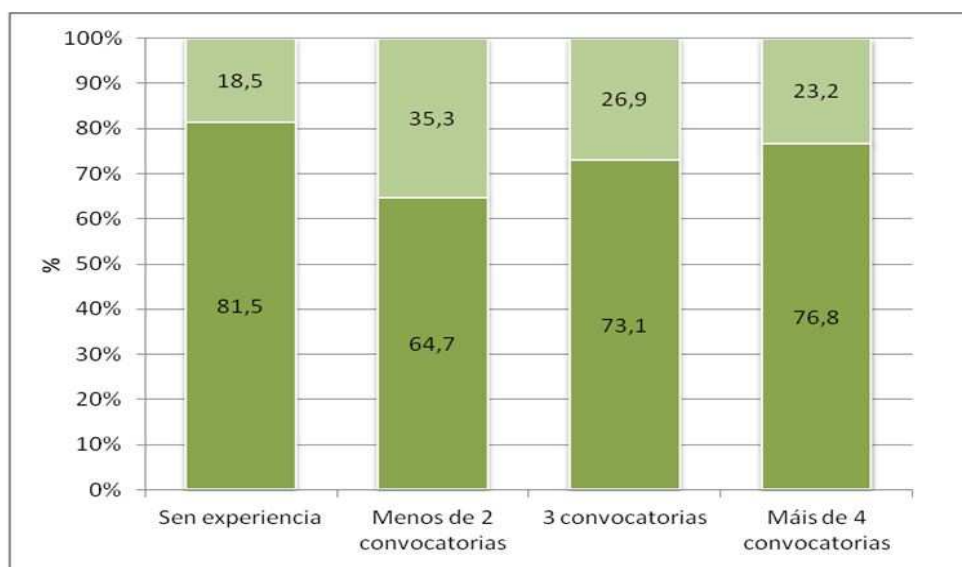


Figura 20. Relación entre a equidade do sistema de avaliación e a experiencia como avaliadores na proba escrita

Ao examinar a relación entre a experiencia dos avaliadores na proba oral e a opinión sobre a equidade do sistema, observouse unha variación significativa ($I = 0,006$; $p = 0,012$). Os avaliadores cunha menor experiencia na proba oral (isto é, ata tres convocatorias) son os que opinan en maior medida (33,3 %) que o sistema é menos xusto.

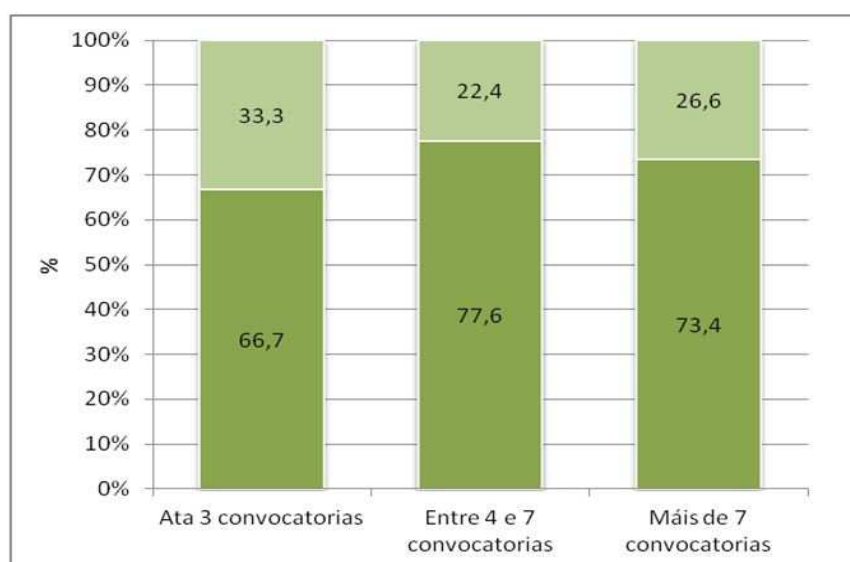


Figura 21. Relación entre a equidade do sistema de avaliación e a experiencia como avaliadores na proba oral

Porén, podemos observar como aqueles avaliadores con entre catro e sete convocatorias ou máis de sete convocatorias de experiencia na proba oral obteñen resultados moi similares cando consideran que o Celga é un sistema xusto de avaliación. Polo tanto, podemos inferir que a consideración da equidade no sistema de avaliación é proporcional á experiencia dos avaliadores nas distintas probas orais escritas.

3.1.12. Valoración da competencia comunicativa dos candidatos ao Celga

Este apartado ten como finalidade amosar o construto empregado polos avaliadores no proceso de avaliación das probas escritas e orais do Celga. Con este obxectivo en mente e tendo presentes os resultados xerais e a influencia do nivel do Celga sobre a percepción da utilidade dos avaliadores, aplicamos a análise INDSCAL.

3.1.12.1. Valoración da competencia comunicativa na proba escrita dos candidatos ao Celga

A análise multidimensional empregada para desentrañar o proceso de valoración realizada polos avaliadores sobre a competencia comunicativa dos candidatos na proba escrita, mediante a utilización dos criterios de corrección e avaliación, proporciónanos un modelo que explica case a totalidade da varianza (99 %), con niveis de Stress de Krustal aceptables. O valor mínimo aconsellable queda por baixo de 0,10 (Stress: 0,01), o cal indica un nivel moi alto de axuste entre o modelo e os datos. Acadar unha varianza explicada alta cun nivel de stress tan baixo significa que se obtivo un modelo correcto segundo os datos avaliados.

Cómpre sinalar que a primeira dimensión (dimensión I) é a que realmente soporta o resultado do modelo de avaliación na proba escrita, xa que explica un 98,5 % da varianza mentres a segunda explica un 1,5 %.

Este nivel tan alto de varianza explicada e tan baixo de stress na primeira dimensión indica a existencia dunha única dimensión para a avaliación da proba escrita en todos os niveis do Celga. Partindo da observación anterior, analizouse polo miúdo o modelo cunha solución onde todos os niveis do Celga teñen un peso superior a 0,90 na dimensión I, que é a que explica case a totalidade da varianza (táboa 22). A dimensión II explica un 2,5 % da varianza e o peso máis grande é no Celga 1 ($w = 0,22$).

	Weirdness	Dimensión I	Dimensión II
CELGA 1	0,56	0,97	0,22
CELGA 2	0,86	1	0,01
CELGA 3	0,09	0,99	0,09
CELGA 4	1	1	0
Varianza explicada		98,5 %	2,5 %

Figura 22. Peso dos niveis na avaliación da proba escrita do Celga en cada dimensión e valores da varianza explicada

Os criterios de corrección e avaliación que teñen un peso maior na dimensión I son (*vid.* táboa 22) adecuación á tarefa ($w = -1,25$), léxico ($w = -0,98$) e gramática ($w = -0,98$); mentres na dimensión II son os criterios de adecuación á tarefa ($w = 1,39$) e gramática ($w = -1,29$).

Criterios de avaliación	Dimensión I	Dimensión II
Adecuación á tarefa	1,25	1,39
Coherencia e cohesión	0,71	0,49
Gramática	-0,98	0,40
Léxico	-0,98	-1,29

Figura 23. Peso dos criterios de corrección e avaliación na proba escrita en cada dimensión

A dimensión I, despois de revisar os datos (*vid.* Figura 23) e as diferenzas por nivel do Celga (Figura 24), amosa que os avaliadores consideran os criterios de gramática e léxico cun menor peso e son máis similares entre eles fronte ao criterio de adecuación á tarefa na avaliación da proba escrita do Celga e independentemente do nivel do certificado. Complementariamente, a dimensión II sinala que para o Celga I o criterio cun peso máis alto é o de adecuación á tarefa fronte ao de léxico.

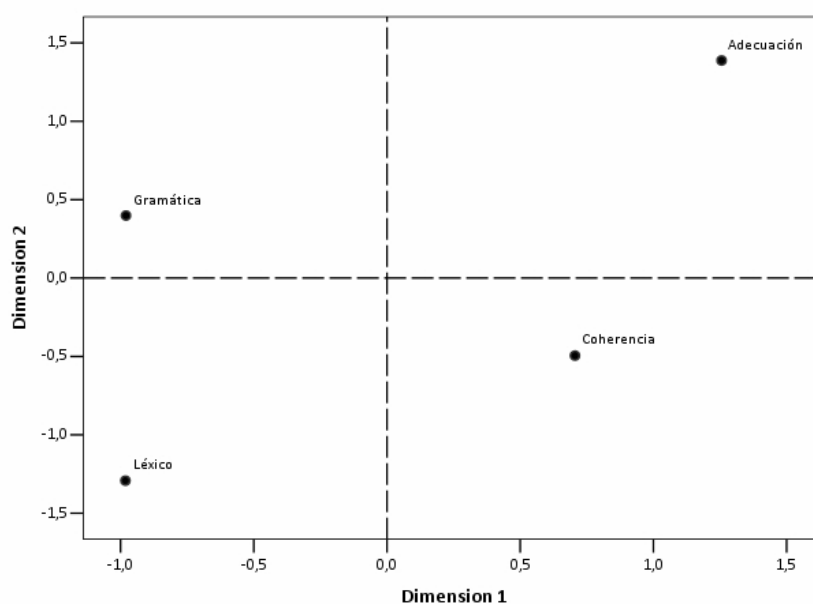


Figura 24. Mapa perceptual dos criterios de avaliación escrita do Celga.

3.1.12.2. Valoración da competencia comunicativa na proba oral dos candidatos ao Celga

O procedemento utilizado para analizar o construto que utilizan os avaliadores para valorar a competencia oral foi igual ao da escrita. Neste caso o modelo multidimensional obtido mediante a análise INDSCAL explica o 99 %

da varianza e un valor de Stress de 0,03. Estes resultados indican un alto nivel de axuste entre o modelo e os datos.

O modelo no que se fundamenta a valoración da competencia comunicativa dos candidatos na proba oral componse de dúas dimensións: a primeira explica un 59 % da varianza e a segunda explica algo menos, un 41 %. Na dimensión I é onde os niveis do Celga 1 e do Celga 2 teñen o peso máis alto ($w = 0,94$ e $w = 0,92$, respectivamente), mentres na dimensión II son os niveis de Celga 4 e Celga 3.

	Weirdness	Dimensión I	Dimensión II
CELGA 1	0,48	0,94	0,33
CELGA 2	0,37	0,92	0,4
CELGA 3	0,36	0,56	0,82
CELGA 4	0,38	0,55	0,83
Varianza explicada		59 %	41 %

Figura 25. Peso dos niveis do Celga na proba oral en cada dimensión e valores da varianza explicada

Os criterios que máis determinan a avaliación da competencia oral na dimensión I serían (Figura 25) a adecuación ($w = -1,76$) fronte á pronuncia e entoación ($w = -1,69$), mentres na segunda dimensión os criterios que máis o determinan son a interacción ($w = -1,46$) e a adecuación ($w = -1,04$) fronte ao léxico ($w = 1,09$), á gramática ($w = 1,06$) e, en menor medida, pronuncia e entoación ($w = -0,64$).

Criterios de avaliación	Dimensión I	Dimensión II
Adecuación	-1,76	-1,04
Gramática	-0,07	1,06
Léxico	-0,05	1,09
Cohesión e fluidez	0,03	-0,29
Pronunciación e entoación	1,69	0,64
Interacción	0,15	-1,46

Figura 26. Pesos dos criterios de avaliación da competencia oral en cada dimensión.

Ao analizar o peso e direccións das escalas en cada unha das dimensións, pódese dicir que a valoración da competencia comunicativa dos candidatos na proba oral nos niveis iniciais (Celga 1 e Celga 2) se fundamenta especialmente no criterio de adecuación e interacción fronte ao de pronuncia e entoación (dimensión I, Figura 26). Por outra banda, nos niveis máis altos do Celga (Celga 3 e Celga 4), a valoración fundaméntase en máis parámetros (dimensión II, Figura 27). Os avaliadores céntranse nos criterios de adecuación e interacción fronte aos de gramática, léxico e pronuncia e entoación.

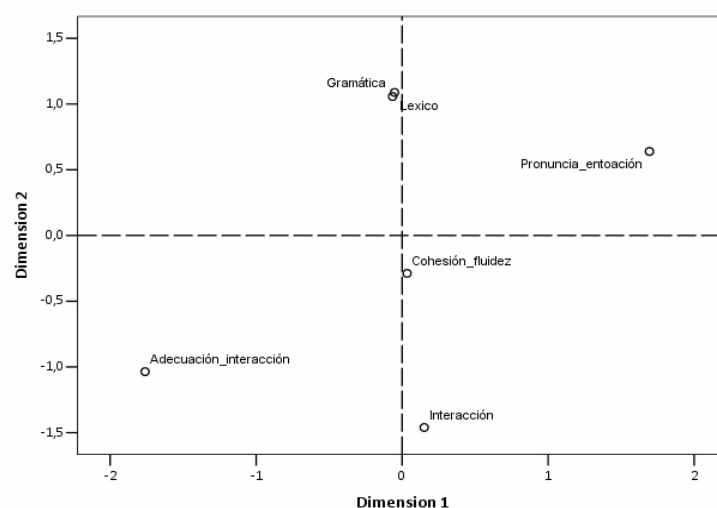


Figura 27. Mapa perceptual dos criterios de avaliación oral do Celga.

3.2. Estudo II: Análise dos candidatos presentados ás probas do Celga durante os anos 2009 e 2010

A continuación presentamos a análise efectuada sobre os datos obtidos dos candidatos ao Celga presentados nas catro convocatorias correspondentes aos dous anos 2009 e 2010 que, xunto cos datos do estudo I, conforman o total da investigación sobre o sistema Celga.

3.2.1. Método da investigación

Mediante un estudo realizado aos candidatos que se presentaron ao Celga durante os anos 2009 e 2010 en dúas convocatorias anuais –maio e decembro–, conseguimos datos coa finalidade de analizar a fiabilidade e validez das ditas probas. O número de candidatos é de 8016, cun perfil sociodemográfico heteroxéneo, xa que se consideran variables como idade, sexo, nivel de estudos, lugar de nacemento e lugar de residencia. A motivación principal para presentarse ás probas vén dada pola necesidade dos candidatos de obter unha certificación que lles permita optar a algún posto dos distintos niveis da Administración pública galega.

Os candidatos facilitáronlle os seus datos á SXPL o mesmo día da proba e acordouse que só se faría uso deles con fins de mellora do sistema. Os datos foron tomados en conxunto, sen diferenciación por anos, xa que non é relevante analizalos por separado para demostrar a fiabilidade e a validez das probas.

3.2.1.1. Obxectivos

Do mesmo xeito que avanzamos no apartado que recolle os obxectivos e hipóteses do estudo I, os obxectivos específicos dos que partimos na

investigación deste estudo están relacionados coas hipóteses que elaboramos, de tal xeito que pretendemos que a suma dos obxectivos específicos sexa o resultado do obxectivo xeral, neste caso comprobar a fiabilidade e validez das probas Celga. Con tal finalidade, consideramos como obxectivos específicos os seguintes:

— Comprobar a fiabilidade interxuíces nas probas Celga, é dicir, entre os avaliadores e os supervisores, coa finalidade de corroborar se os avaliadores e supervisores avalían do mesmo xeito as mesmas probas.

— Comprobar a consistencia interna, que é a que se obtén da relación que entre si presentan os parámetros de corrección e avaliación empregados nas probas Celga.

— Comprobar a validez de construto, referida ao grao no que un procedemento de medida considera de forma axeitada o construto teórico que pretende medir.

— Analizar o perfil sociodemográfico dos candidatos ao Celga en relación coa obtención da aptitude.

— Analizar os resultados das proba escrita e oral segundo as variables sociodemográficas dos candidatos.

3.2.1.2 Hipóteses

As hipóteses que presentamos están relacionadas cos obxectivos xerais e específicos que conforman o apartado anterior, articuladas do seguinte xeito:

— **Con respecto á fiabilidade interxuíces:** os resultados da primeira rolda de avaliación e da segunda, levada a cabo polos supervisores, terán un alto nivel de correlación á hora de considerar a un candidato como apto ou non apto.

— **Con respecto á consistencia interna:** esperamos que sexa alta, é dicir, que os parámetros de corrección e avaliación das probas teñan un alto coeficiente de consistencia.

— **Con respecto á validez de construto:** esperamos que tanto na proba escrita coma na oral os parámetros de corrección e avaliación que remiten ao coñecemento do código teñan máis peso na avaliación da competencia comunicativa.

— **Con respecto ás variables clasificatorias:** as variables como sexo, lugar de nacemento e lugar de residencia non terán relevancia á hora de acadar a aptitude do sistema, mentres que o nivel de estudos e a idade serán determinantes para que un candidato sexa apto ou non. Do mesmo xeito, consideramos que o nivel de Celga ao que se presente o candidato influirá significativamente para acadar a súa aptitude.

3.2.2. Mostra

A mostra deste estudo está composta polos candidatos das probas do Celga durante os anos 2009 e 2010 nas dúas convocatorias de maio e de novembro. Os datos foron proporcionados pola Secretaría Xeral de Política Lingüística da Xunta de Galicia.

Os candidatos ao Celga conforman unha poboación heteroxénea tanto na súa procedencia coma na formación, xa que se poden atopar cidadáns galegos, doutras comunidades españolas e estranxeiros. O perfil dos candidatos varía en idade, dende os 18 anos en adiante, e en niveis de estudo que poden ser básicos, medios ou universitarios.

A totalidade da mostra está composta por un total de 8024 participantes, dos cales a maior parte (36,3 %) se corresponde coas persoas que se presentaron á proba do Celga 4, fronte a un 11,7 % que se presentaron ao Celga

1, mentres que o Celga 2 se compón dun 24,1 % e o Celga 3 dun 27,9 % do total dos candidatos (Figura 28).

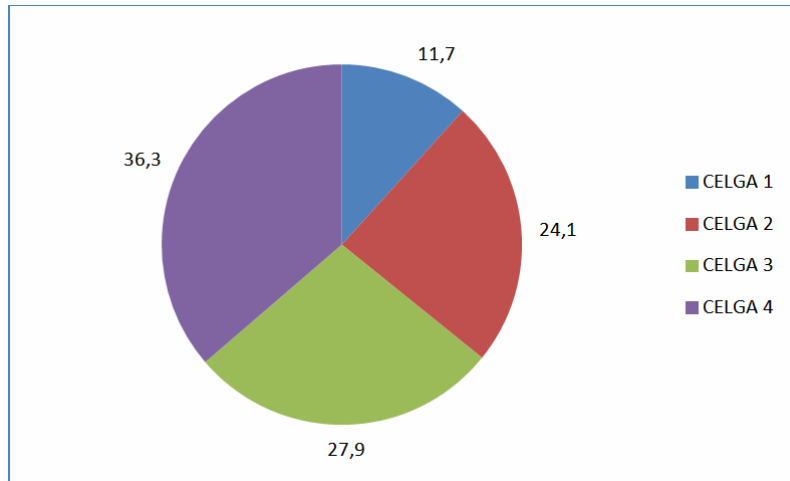


Figura 28. Distribución porcentual da mostra por Celga

Ao comparar o número de persoas que se presentaron nas convocatorias de 2009 co das que se presentaron en 2010, obsérvase que as cifras son moi semellantes, aínda que son algo superiores as de 2009 (Figura 28).

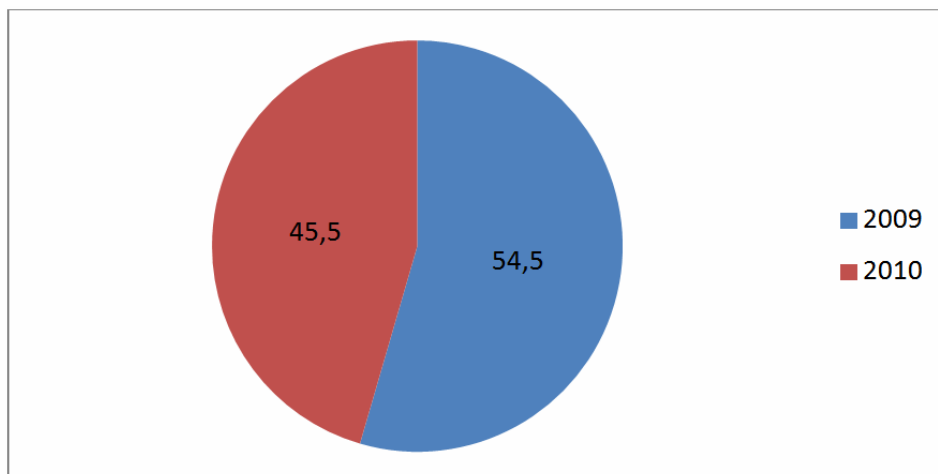


Figura 29. Distribución porcentual da mostra por anos

Podemos observar na Figura 30 que no ano 2009 houbo un número maior de candidatos presentados a todos os niveis do Celga (54,5 %), fronte ao 2010 no que a porcentaxe total de candidatos que se presentaron aos distintos niveis do Celga foi dun 45,5 %. Isto pode deberse a que, no inicio da existencia do certificado de lingua galega, a cidadanía que xa posuía os certificados Celga obtidos polas validacións que decretou a SXPL consideraba que se conseguían a certificación mediante o exame podería obter dobre puntuación para determinados postos de traballo ou nos concursos de traslados da función pública galega.

Ao analizar a evolución temporal do número de persoas que se presentaron ás probas do Celga obsérvase que (Figura 30):

- a. Hai unha certa diminución das persoas que se presentan ás convocatorias en todos os niveis.
- b. O número de persoas que se presentan ao Celga 1 diminúe progresivamente.
- c. A última convocatoria para o Celga 2 tivo un número substantivamente menor de participantes.
- d. A falta de datos de máis convocatorias, semella que hai unha estabilización no número de persoas que se presentan á proba do Celga 3.
- e. Na proba para acadar o Celga 4, despois de ter unha diminución progresiva, volve aumentar o número de persoas que se presentan para acadar este certificado.

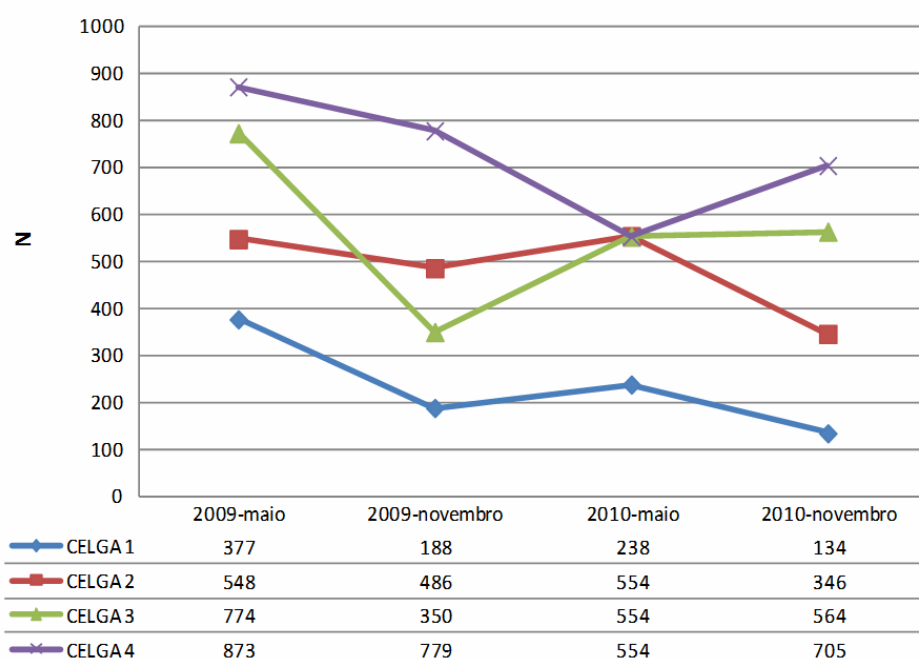


Figura 30. Distribución da mostra por Celga, ano e convocatoria

3.2.3. Medidas

Neste apartado analizamos as variables que afectan os candidatos do Celga, dividímolas en variables clasificatorias e criterios de corrección e avaliación empregados polos avaliadores tanto na proba oral coma na escrita.

3.2.3.1. Variables clasificatorias

a. **Ano da convocatoria:** os niveis da variable son os seguintes: 2009 e 2010 e convocatoria de maio e convocatoria de decembro.

b. **Nivel do Celga ao que se presenta:** os niveis son os seguintes: Celga 1, Celga 2, Celga 3 ou Celga 4.

c. **Nivel de estudos:** os niveis da variable son os seguintes: sen estudos, estudos primarios, secundario, estudos universitarios.

d. **Lugar de residencia:** os niveis da variable son os seguintes: aldea, vila ou cidade.

e. **Lugar de nacemento:** os niveis da variable son os seguintes: Galicia, outras rexións de España, Europa, América, África, Asia, Oceanía.

f. **Sexo:** os niveis son varón e muller.

e. **Idade:** as idades dos candidatos foron clasificadas nos seguintes intervalos: menores de 30 anos, entre 30 e 40 anos, entre 40 e 50 anos, entre 50 e 60 e máis de 60 anos.

3.2.3.2. *Criterios de corrección na proba escrita*

A proba escrita, denominada proba A, consta, pola súa vez, de dúas probas, unha primeira chamada A1, na que se avalía a comprensión oral e a expresión escrita, e unha segunda proba A2, na que se avalía a comprensión escrita e a expresión escrita.

As variables empregadas como ferramenta de corrección e avaliación da proba escrita polos avaliadores constan dos parámetros que se detallan a continuación:

Adecuación á tarefa: mediante este parámetro valórase que o candidato se axuste ao tipo de tarefa que se lle esixe para completar a proba, é dicir, se se lle pide que redacte un texto dunha determinada modalidade (expositivo, argumentativo ou narrativo) e cun determinado número de palabras (que varía segundo o nivel de Celga), espérase que o cumpra.

Coherencia e cohesión: mediante este parámetro valórase que o texto producido polo candidato amose unha estrutura adecuada na presentación de ideas, empregue conectores que axuden a progresar a información e utilice os

mecanismos gramaticais e léxicos necesarios para o mantemento do referente.

Consideramos que

la coherencia es una noción más extensa que incluye las relaciones pragmáticas, además de las relaciones semánticas intratextuales. Es un concepto que se refiere al significado del texto en su totalidad, abarcando tanto las relaciones de las palabras con el contexto como las relaciones entre palabras en el interior del mismo texto. Alude a la estabilidad y a la consistencia temática subyacente, asociada a la macroestructura (contenido), a la superestructura (esquema de organización) del texto, a su anclaje enunciativo (protagonistas, tiempo y espacio) y a las inferencias que activan los hablantes para interpretarlo a partir de conocimientos previos. La cohesión es un concepto que se refiere a uno de los fenómenos propios de la coherencia, el de las relaciones particulares y locales que se dan entre elementos lingüísticos, tanto los que remiten unos a otros como los que tienen la función de conectar y organizar. Así, la coherencia, que incluye la cohesión, se puede considerar tanto una propiedad fundamental que da cobertura al conjunto de significaciones del texto como un principio necesario para que un texto sea definido como tal. (Casalmiglia e Tusón, 2002: 221-222)

Gramática: mediante este parámetro valórase a corrección morfolóxica e sintáctica no discurso do candidato e a súa adaptación á norma estándar do galego.

Léxico: neste parámetro valórase a corrección, precisión e riqueza léxica, de acordo coas esixencias de cada nivel.

A competencia comunicativa na proba escrita está composta por dous niveis (apto/non apto). Para considerar se un candidato é apto ou non acudiuse ás puntuacións obtidas nos parámetros anteriores que seguen unha escala do 1 ao 4: 1 e 2 son non apto e 3 e 4 son apto.

Competencia comunicativa escrita: os niveis da proba escrita serán apto e non apto. Para esta codificación consideraranse os criterios do Celga e, por tratarse en realidade de dúas probas (A1 e A2), permítese unha puntuación menor a 3 nalgún dos parámetros e só nunha das probas (na A1 ou na A2) e a

proba considerárase apta. En caso de aparecer máis dunha puntuación menor a 3, a proba considerárase non apta.

3.2.3.3. *Criterios de corrección na proba oral*

Na proba oral, denominada proba B, os parámetros de corrección e avaliación que teñen que empregar os avaliadores como ferramenta no seu labor son os seguintes:

- Adecuación
- Gramática
- Léxico
- Cohesión e fluidez
- Pronuncia e entoación
- Interacción

Como podemos observar, os únicos que varían con respecto aos parámetros da proba escrita son os seguintes, que pasamos a detallar:

— **Pronuncia e entoación:** neste parámetro valórase que o usuario utilice na oralidade o sistema fonolóxico do galego.

— **Interacción:** neste parámetro valórase que o candidato colabore co interlocutor e manteña vivo o diálogo nas partes da proba que requiren interacción.

— **Cohesión e fluidez:** neste parámetro valórase que o candidato estructure correctamente as súas ideas e que teña unha fluidez e un ritmo de fala acorde co nivel correspondente.

A competencia comunicativa na proba oral está composta por dous niveis (apto ou non apto). Para considerar se un candidato é apto ou non apto acudiuse ás puntuacións obtidas nos parámetros anteriores que seguen unha escala do 1 ao 4, de 1 a 2 son non apto e de 3 a 4 apto.

3.2.4. Procedemento

3.2.4.1. Obtención de datos para o estudo

Os datos empregados no presente apartado proceden das convocatorias do Celga dos anos 2009 e 2010 sometidos a unha análise psicométrica. Compiláronse mediante unha serie de táboas Excel onde se recollen os datos referidos aos candidatos, as puntuacións que obtiveron nas probas que realizaron e os códigos das persoas que avaliaron as ditas probas. Posteriormente, estes datos trasladáronse a unha matriz de datos de SPSS para a súa análise.

Todos os datos obtidos para o presente traballo fóronlles solicitados aos candidatos o día da proba, aos que se lles explicou que só se faría uso deles para investigación a prol da mellora do sistema Celga.

3.2.4.2. Procedemento de avaliación das probas Celga

Os candidatos que optan á consecución dun certificado Celga terán que superar dúas probas, unha escrita e outra oral, que pasamos a detallar a continuación.

PROBA ESCRITA

Proba A1

Este proba A1 correspóndese co primeiro exercicio que o candidato ten que resolver no exame. A partir dun texto oral que escoita, debe redactar un texto

escrito segundo unha serie de pautas que se lle indican e que lle permitirán adecuarse á tarefa solicitada, ben sexa creando un texto expositivo, narrativo ou argumentativo. Deste xeito, o candidato demostrará **a súa competencia comunicativa tanto na comprensión oral coma na expresión escrita** dos distintos niveis de Celga, ao ser quen de construír un texto entendido como “una unidad comunicativa de orden distinta a la oracional; una unidad intencional y de interacción” (Casalmiglia e Tusón, 2002: 219).

Proba A2

Nesta segunda parte da proba escrita, denominada A2, a partir do *input* proporcionado por un texto escrito, o candidato debe redactar outro segundo unha serie de pautas que lle permitirán adecuarse á tarefa solicitada, creando un texto de modalidade expositiva, narrativa, ou argumentativa. Deste xeito, o candidato demostrará **a súa competencia comunicativa na comprensión e a expresión escrita** dos distintos niveis de Celga.

PROTOCOLO DE AVALIACIÓN DA PROBA ESCRITA

Todos os parámetros que conforman a grella de avaliación son empregados polos avaliadores colaboradores da SXPL que recibiron unha formación específica para desempeñar o dito labor. Para manter o anonimato e a fiabilidade na avaliación das probas escritas, o sistema establece que existan varias figuras que corrixan as probas; todas as persoas implicadas no proceso de corrección descoñecen a puntuación outorgada previamente polos seus compañeiros, así como a identidade do candidato –a proba non presenta ningún risco nin marca ou sinal que poida permitir a identificación do candidato–. Cada proba pode pasar ata por catro avaliadores diferentes que descoñecen a puntuación que obtivo esa proba anteriormente. Así, o sistema de avaliación do Celga presenta as seguintes figuras:

— Un **corrector**: é a primeira persoa que avalía a proba segundo os parámetros establecidos e decide se esta é apta ou non apta.

— Un **supervisor**: que descoñece a puntuación outorgada anteriormente polo corrector á proba e avalía de novo, xa que todas as probas non aptas e, aleatoriamente, algunhas aptas son corrixidas unha segunda vez por un supervisor.

— Un **revisor**: no caso de que as notas do corrector e do supervisor sexan discordantes, esa proba pasa a unha terceira corrección na figura do revisor, que tamén descoñece a puntuación outorgada á proba polos anteriores corrector e supervisor.

— Un **membro da Comisión Central Avaliadora**: no caso de reclamación por parte do candidato, que en última instancia será quen decide se a proba é apta ou non apta.

Os datos empregados no presente traballo son propiedade da SXPL e foron xenerosamente cedidos para seren analizados e permitir unha optimización e mellora do sistema de certificación Celga.

PROBA ORAL

Logo de realizar a proba escrita, os candidatos deberán demostrar a súa competencia comunicativa en galego na proba oral. Para isto, a partir dun *input* que pode ser unha comprensión oral ou unha comprensión escrita, deberán realizar unha entrevista na que estarán presentes o examinador e o avaliador e que será rexistrada para conservar a evidencia da actuación.

O examinador levará o peso da entrevista co candidato, mentres o avaliador, sen participar na conversa, tomará nota da actuación do candidato para despois cubrir as grellas de avaliación coa puntuación que considere oportuna.

Cando remata a entrevista e sen mediar palabra co avaliador, o entrevistador tamén cubrirá as grellas de avaliación e outorgaralle ao candidato a puntuación que considere axeitada. Unha vez cubertas as grellas por parte do avaliador e do examinador, exporán se consideran o candidato apto ou non, sen detallar a puntuación que cada un lle asignou aos parámetros de avaliación; en caso de non chegar a un acordo, deberán debater e, de non poñerse de acordo, anotarán na grella como incidencia a imposibilidade de chegar a un acordo na aptitude ou non do candidato, de tal xeito que pasaría a ser avaliado por un superviso.

3.2.5. Análises estatísticas

Para analizar os resultados dos candidatos nas probas do Celga nas convocatorias do ano 2009 e 2010, utilizáronse complementariamente diferente tipos de técnicas. Para comprobar a fiabilidade do instrumento de avaliación do Celga, empregouse a proba Alpha de Crobach e así obter o coeficiente de consistencia, e para analizar a fiabilidade interxuíces escolleuse o coeficiente de kappa complementado co estatístico de correlación de Pearson.

No estudo da validez de construto da proba aplicouse unha análise de compoñentes principais con rotación varimax, acompañado das probas de Kaiser-Meyer-Olkin e de Barlett que indican se se cumpren os requisitos para aplicar esta proba.

Para afondar na relación entre as variables clasificatorias recollidas (nivel de Celga, nivel de estudos, idade, sexo...) e os resultados obtidos polos candidatos no Celga na proba oral e na proba escrita (apto/non apto), realizáronse contrastes de tipo non paramétrico que permiten establecer o nivel de significatividade das relacións.

Escolleuse entre as probas non paramétricas dispoñibles a do coeficiente de incerteza, técnica fundamentada na redución da proporción de erro que se amosa máis robusta ca outras técnicas como a lamba, xa que reduce a posibilidade de realizar prognósticos erróneos ao considerar non só as posibilidades de que os casos pertencen á categoría máis probable senón a todas as categorías dispoñibles.

Tamén se utilizou a proba Anova One Way con contrastes *a posteriori* Tukey b para analizar a relación de determinadas variables clasificatorias sobre os criterios de corrección e avaliación das probas de competencia comunicativa escrita e oral.

Os datos foron exportados dunha base Excel a unha matriz de datos de SPSS para a súa posterior recodificación e análise.

3.2.6. Resultados segundo o perfil sociodemográfico dos candidatos ao Celga

A continuación expóñense as características xerais dos candidatos ao Celga durante os anos 2009 e 2010. Para iso iremos analizando variables sociodemográficas tales como a idade, o sexo, o nivel de estudos, o lugar de nacemento e o lugar de residencia dos candidatos –datos recollidos o día da proba e concedida a autorización á SXPL para facer uso deles con fins de mellora no sistema Celga–.

3.2.6.1. Segundo a idade

Unha das variables que caracterizan os datos sociais dos candidatos ao Celga é a idade. A continuación preséntase a distribución porcentual dos participantes nas probas para a obtención do Celga nos anos 2009 e 2010 por tramos de idade. Cómpre sinalar que soamente se obtiveron as idades de 2487 persoas, é dicir, o 55,6 % da mostra, polo tanto, este dato débese ter en conta con certa precaución. Despois de obter as súas idades, agrupáronse estas en cinco tramos,

situados desde menos de 30 anos ata máis de 60 anos. Ao realizar as análises descritivas directas, obtivemos que a mediana é de 41 anos –o participante de menor idade tiña 17 anos e o de maior 76–, mentres a media é 46 anos (Figura 31).

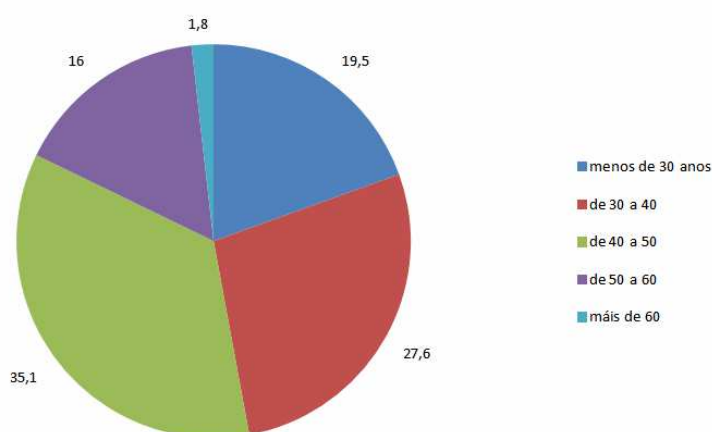


Figura 31. Distribución dos candidatos por idade

Como podemos observar na Figura 32, na que se recollen as distribucións dos participantes por idade e ano de convocatoria, no ano 2009 preséntase en cada un dos tramos de idade unha proporción maior ca na convocatoria do ano 2010, aínda que as diferenzas non son moi relevantes, pois non superan máis de tres puntos por tramo de idade.

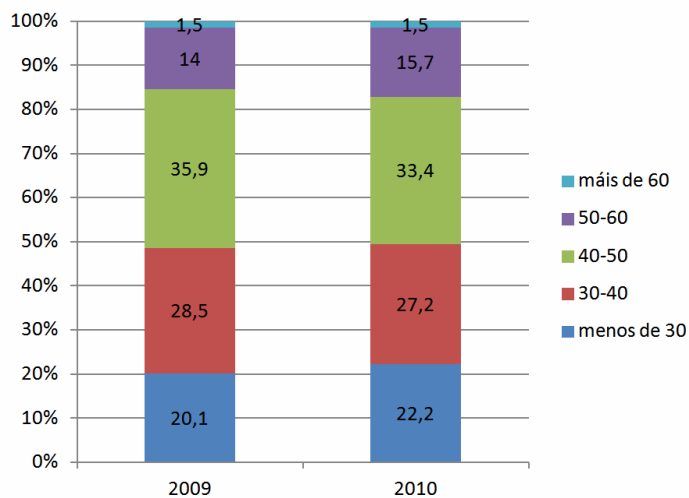


Figura 32. Distribución dos candidatos por idade e ano de convocatoria

Pódese observar na Figura 33 como nas quendas de maio e decembro o número de candidatos é case igual e a diferenza por tramos de idade é pouco relevante. En casos como o dos candidatos cunha idade entre 40 e 50 anos presentan unhas porcentaxes do 35,1 % no mes de maio fronte ao 35,3 % do mes de decembro, isto é, que a poboación candidata ao Celga é moi similar e non varía dunha convocatoria a outra.

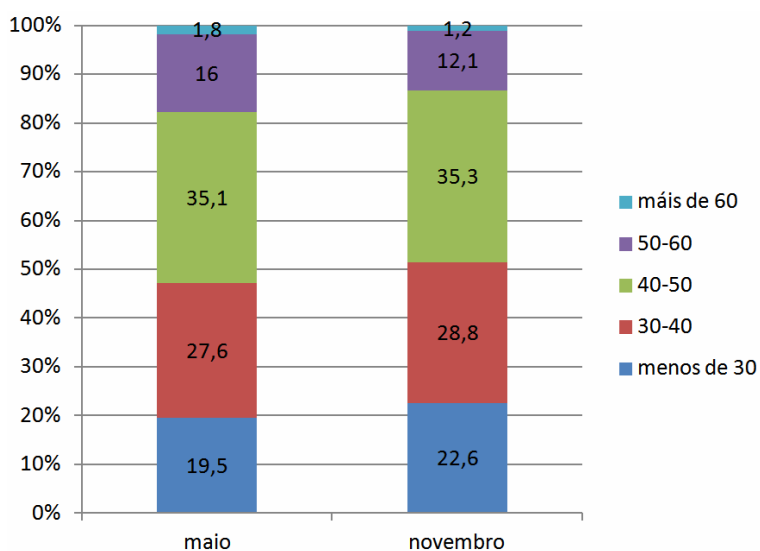


Figura 33. Distribución dos candidatos por idade e convocatoria

Como podemos observar na Figura 34, na que se plasma a distribución dos participantes segundo a idade e a proba á que se presentan, o tramo de idade que vai dos 40 aos 50 anos é o maioritario nos niveis de Celga 2, Celga 3 e Celga 4, mentres que no Celga 1 o tramo de idade no que se presentan máis candidatos é o que vai dos 30 aos 40 anos. Isto débese a que neste tramo de idade a maioría dos candidatos que se presentan de entre 30 e 40 anos non estiveron escolarizados en galego e adoitan ser estranxeiros e españois que se presentan a este nivel e que pretenden optar aos postos máis baixos da Administración galega (grupos C, D e E).

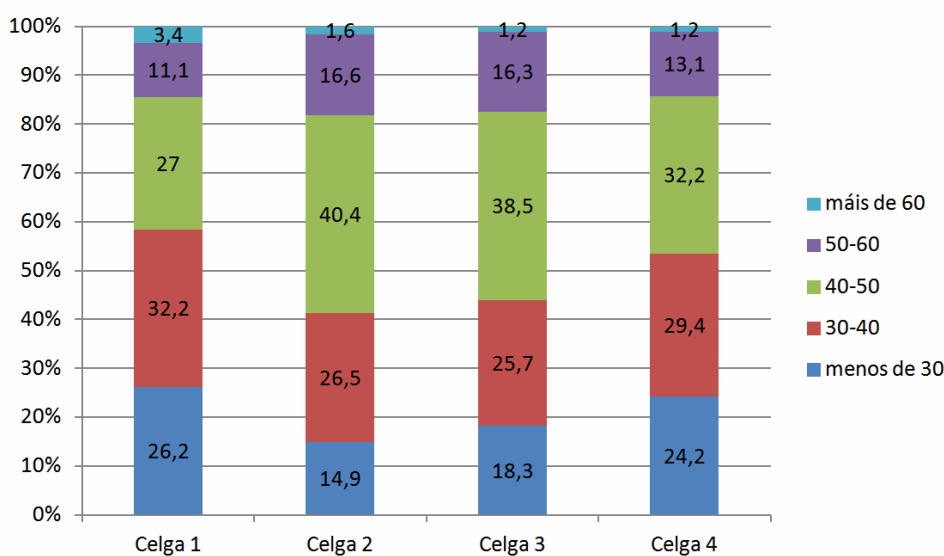


Figura 34. Distribución dos candidatos por idade e proba á que se presentan

Obsérvase que aqueles candidatos cunha idade entre 40 e 50 anos, que non estiveron escolarizados en galego e que maioritariamente son galegos, se presentan aos grupos Celga 2 e Celga 3. Este colectivo merecerá unha especial atención, xa que moitos deles non conseguirán a aptitude nas probas aínda que adoitan ser galegofalantes, cuestión que será analizada con máis detalle nas conclusións.

3.2.6.2. Segundo o sexo

Con respecto a esta variable, obtense da análise psicométrica un dato interesante, xa que, tendo en conta que a poboación xeral está distribuída case por igual entre homes e mulleres, as tres cuartas partes dos participantes nas probas Celga son mulleres (70,6 %) (Figura 35).

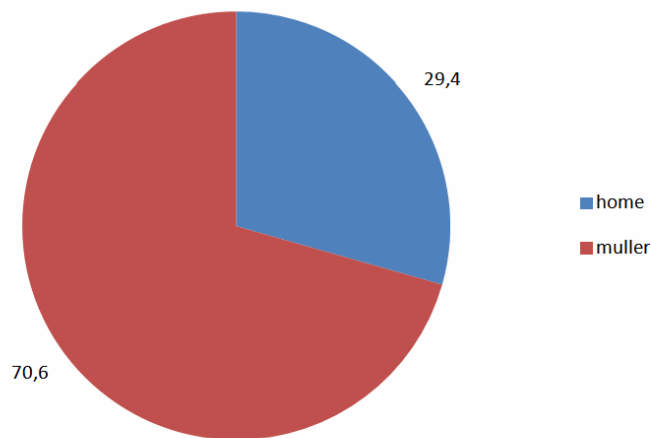


Figura 35. Sexo dos participantes

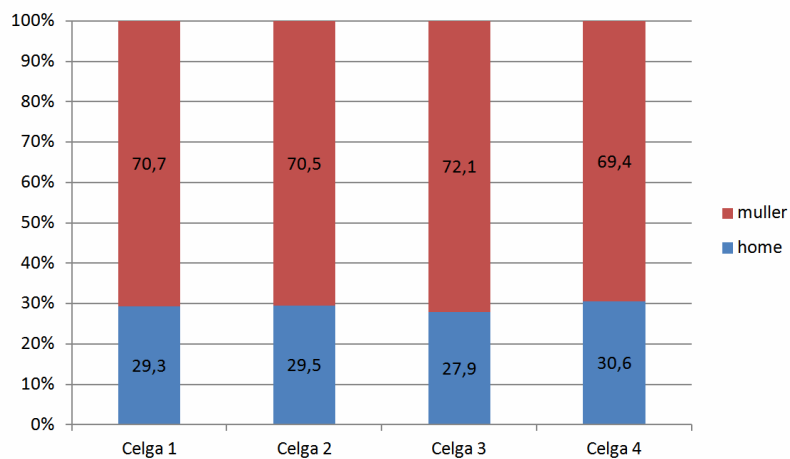


Figura 36. Distribución do sexo dos candidatos segundo a proba á que se presentan

Como podemos observar na Figura 36, a distribución do sexo dos participantes vén marcada por unha maioría de mulleres que ocupan o 70 % de participación en todos os niveis do Celga.

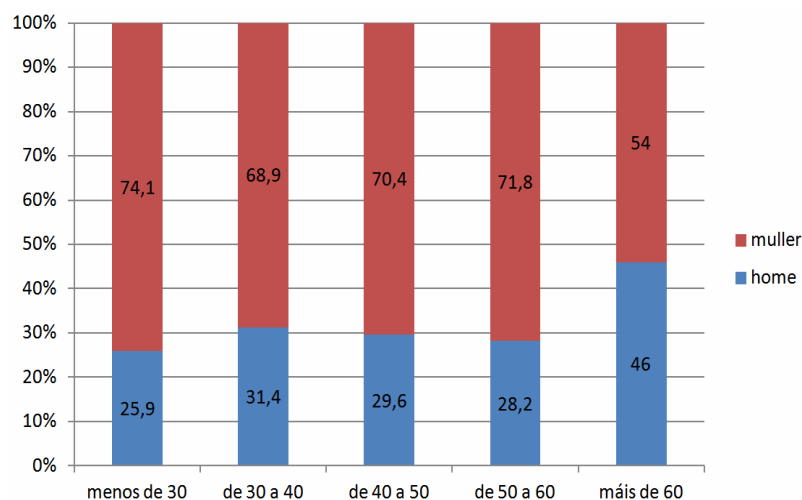


Figura 37. Distribución do sexo dos candidatos segundo a idade

Nesta outra Figura 37 podemos apreciar como, sendo as mulleres o colectivo maioritario nos catro niveis de Celga, as menores de 30 anos presentan un maior número de candidatas (74,1 %), seguidas polas mulleres entre 50 e 60 anos que tamén acadan cotas moi altas de participación (71,8 %) nas probas para obtención do Celga.

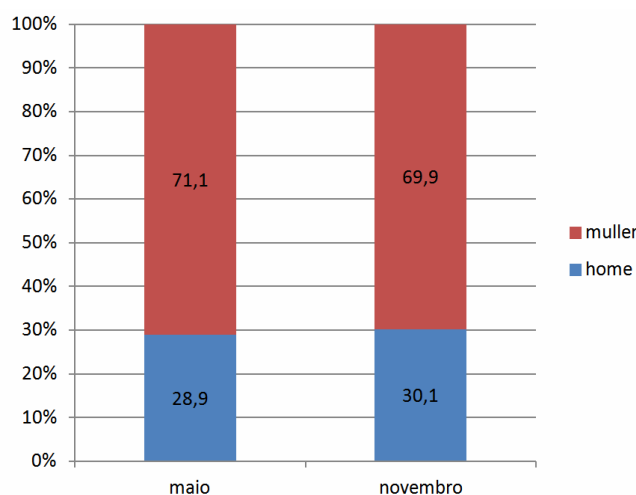


Figura 38. Distribución do sexo dos candidatos segundo a convocatoria

Observamos na Figura 38 que continúan sendo as mulleres as candidatas maioritarias nas probas Celga. Non é relevante a época do ano na que se realiza a convocatoria para que o número de candidatas varíe, xa que son mínimas as diferenzas entre unha convocatoria en maio (71,1 % de mulleres presentadas) ou en decembro (69,9 % de mulleres presentadas).

3.2.6.3. Segundo o lugar de nacemento

Ao analizar o lugar de nacemento das persoas participantes, apréciase que algo máis da metade dos avaliados naceu en Galicia (55,3 %), mentres o 29,6 % pertence ao resto de comunidades autónomas do Estado español. Existe unha porcentaxe mínima de persoas que naceu no continente americano (8,3 %) e no resto de Europa (6,5 %). Cómpre sinalar que a esta pregunta soamente contestou menos da metade dos avaliados (46,5 %), polo tanto, estes datos deben ser tomados con precaución.

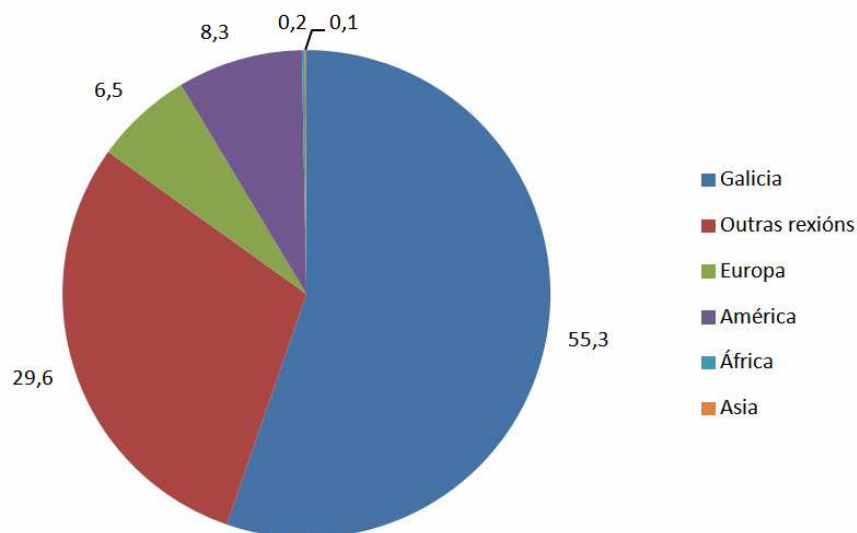


Figura 39. Lugar de nacemento dos candidatos

Como xa adiantamos, a maioría dos candidatos presentados ao Celga son galegos, un 55,3 %, seguidos por un 29,6 % de españois e o resto confórmano candidatos nados en América, Europa. A presenza de candidatos nados en África e Asia é case anecdótica.

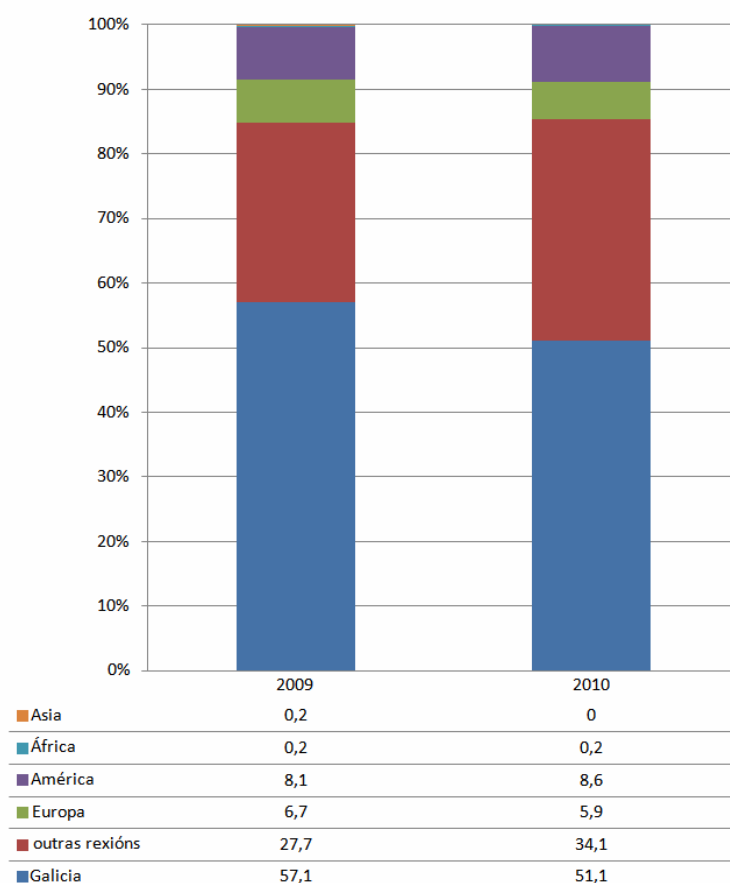


Figura 40. Lugar de nacemento dos candidatos segundo o ano da convocatoria

Podemos observar na Figura 40 que, dependendo do ano da convocatoria, o lugar de nacemento dos participantes non presenta unha variación digna de mención. Deste xeito, obsérvase que no ano 2009 o número de candidatos galegos é aproximadamente dun 57,1 %, fronte ao 51,1 % do ano 2010; a categoría de candidatos que si experimentou un crecemento maior foi a

que atinxe os nacidos noutras rexións españolas diferentes á galega (un 27,7 % no ano 2009 e un 34,1 % no ano 2010). Con respecto a aqueles candidatos nacidos en África, América, Asia ou Europa non presentan unha diferenza salientable máis alá dunhas décimas.

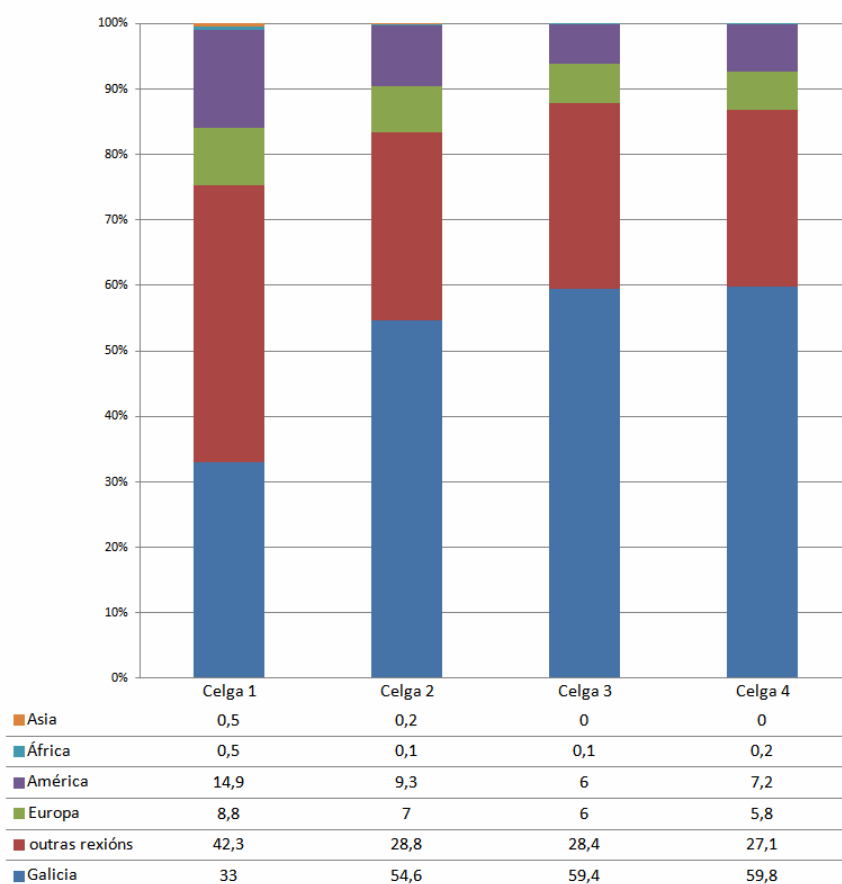


Figura 41. Lugar de nacemento dos candidatos segundo o nivel do CELGA ao que se presentaban

Con respecto ao lugar de nacemento dos candidatos segundo o nivel de Celga ao que se presentan (Figura 41), podemos observar como os galegos son a poboación maioritaria que participa nas probas e os que optan polos niveis máis altos, Celga 3 (59,4 %) e Celga 4 (59,8 %) son os máis numerosos. Isto

pode deberse a que os postos da Administración galega máis demandados son aqueles nos que se esixen os niveis de competencia máis altos en lingua galega.

Como segundo grupo maioritario preséntanse aqueles candidatos nados noutras áreas xeográficas españolas; non obstante, é salientable que este colectivo se presenta maioritariamente ao Celga 1 (42,3 %) e nos outros tres niveis de Celga nunha proporción similar: Celga 2 (28,8 %), Celga 3 (28,4 %) e Celga 4 (27,1 %).

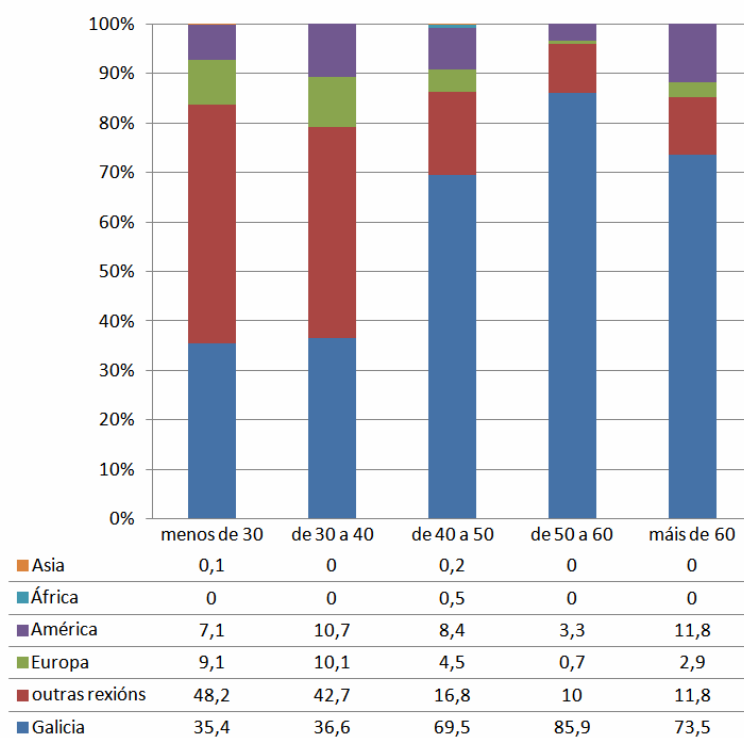


Figura 42. Lugar de nacemento dos candidatos segundo a idade

Segundo a idade dos candidatos (Figura 42), observamos que o perfil maioritario dos candidatos presentados ás probas é o de persoa de procedencia galega e entre 50 e 60 anos (85,9 %), moi seguido polos tamén galegos e maiores de 60 anos (73,5 %). Isto débese a que nesta franxa de idade a cidadanía non estivo escolarizada en galego e precisa dunha certificación que acredite a súa competencia lingüística.

Pola contra, só unha porcentaxe dun 35,4 % dos candidatos galegos menores de 30 anos se presentan ás probas, xa que eles foron escolarizados en Galicia e en galego e por medio das validacións poden obter a certificación que precisen. Contrariamente, dos candidatos con orixe noutras áreas xeográficas españolas preséntanse ás probas os menores de 30 anos (48,2 %) e de entre 30 e 40 anos (42,7 %). Isto pode deberse a que nestas franxas de idade os candidatos decidan presentarse aos postos laborais da Administración galega ou ben queiran posuír unha certificación noutra lingua diferente á do seu lugar de orixe.

3.2.6.4. Segundo o lugar de residencia

Os candidatos presentados ás probas Celga nos anos 2009 e 2010 vivían maioritariamente en cidades (56,7 %), seguidos daqueles que residen nas vilas (37 %) e, por último, unha porcentaxe moi pequena de persoas que viven na aldea (6,3 %).

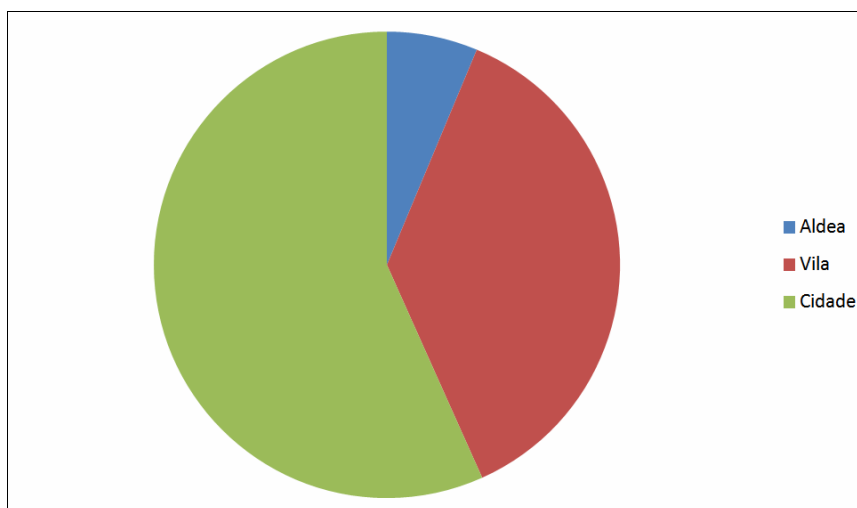


Figura 43. Lugar de residencia dos candidatos

Como se pode comprobar no gráfico (Figura 43), o perfil sociolóxico dos candidatos con respecto ao lugar de residencia é marcadamente urbano, xa

que un 56,7 % do total dos participante vive nalgunha cidade galega. Cómpre sinalar que un 60,6 % das persoas non proporcionou a información solicitada sobre o lugar de residencia no exame, polo tanto, estes datos deben ser tomados con precaución.

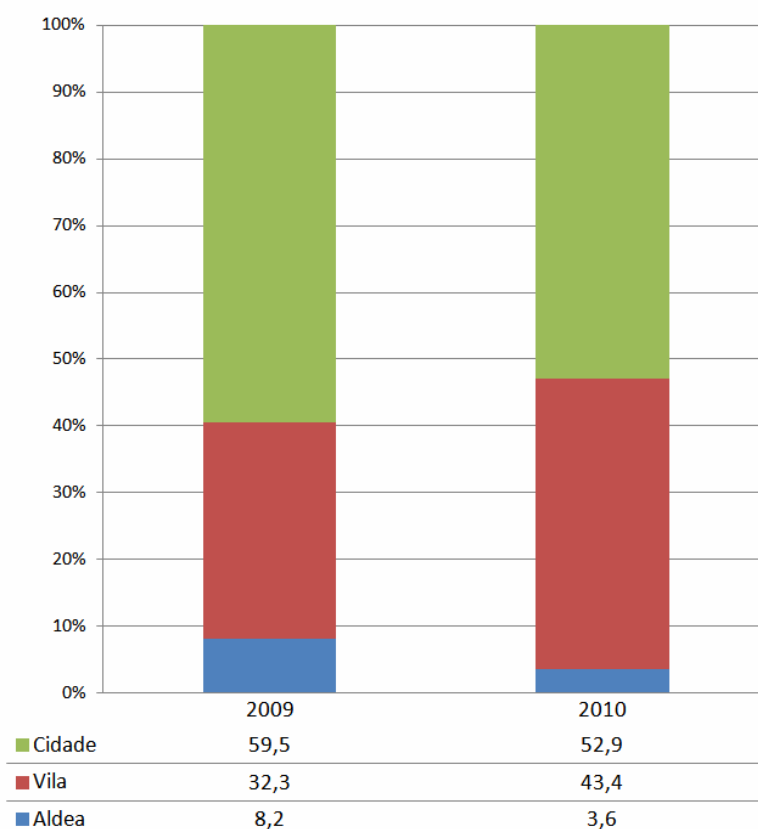


Figura 44. Lugar de residencia dos candidatos segundo o ano da convocatoria

Do mesmo xeito ca na anterior figura (Figura 44), podemos apreciar que, dependendo do ano da convocatoria, non presenta gran variación o lugar de residencia e son as cidades as que maior número de candidatos achegan, tanto no 2009 (59,5 %) coma no 2010 (52,9 %).

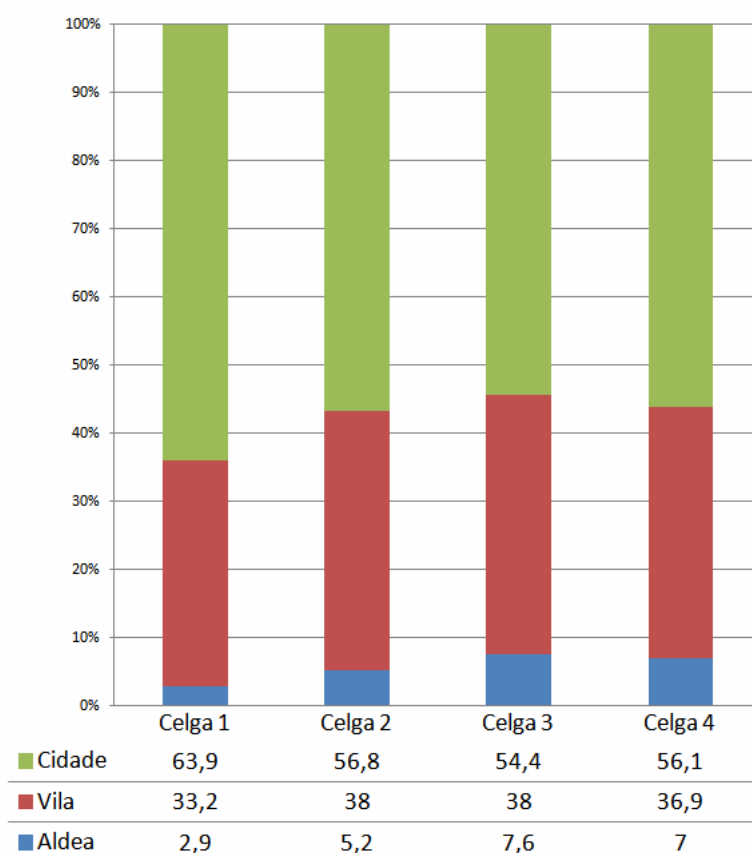


Figura 45. Lugar de residencia dos candidatos segundo o nivel do Celga ao que se presentaban

Con respecto á variable do nivel de Celga ao que se presentan os candidatos e a relación co lugar de residencia (Figura 45), non se observan diferenzas salientables. En calquera nivel do Celga, dende o máis baixo ao máis alto, as porcentaxes de candidatos repártense nunhas proporcións semellantes; sinalamos como diferenza salientable aqueles que moran na cidade, cun 56,8 % do Celga 2, un 54,4 % no Celga 3 e un 36,9 % no Celga 4, fronte aos candidatos do Celga 1 moradores na cidade que presentan unha porcentaxe maior, cun 63,9 %.

Sistema de certificación de lingua galega (Celga) no *Marco Común Europeo de Referencia para as Linguas*. Estudo de fiabilidade e validez

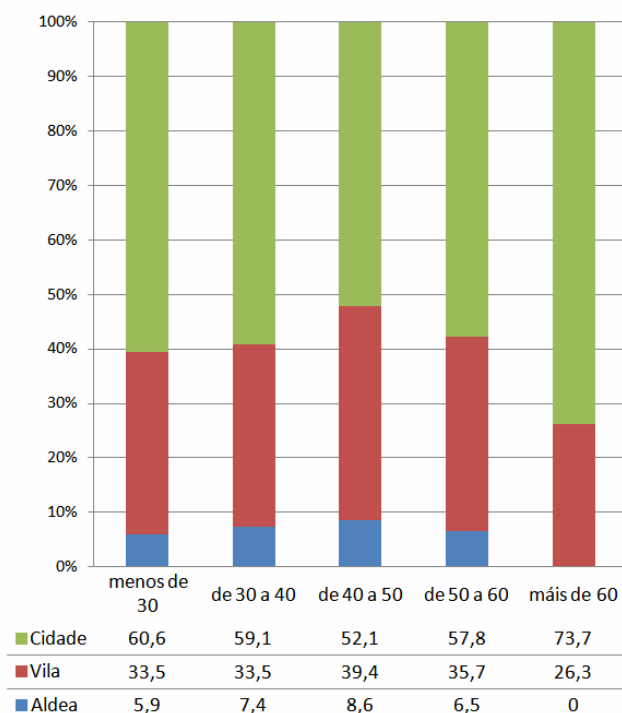


Figura 46. Lugar de residencia dos candidatos segundo a idade

Podemos ver na figura anterior (Figura 46) que o lugar de residencia e a idade nos candidatos só é significativa naqueles que teñen máis de 60 anos, xa que neste caso unicamente se presentan ás probas os que viven nas cidades (un 73,7 %) e nas vilas (un 26,3 %). Os habitantes das aldeas con máis de 60 anos son un perfil de candidato que non existe no sistema Celga; pola contra, obsérvase que os candidatos que viven nas cidades entre menos de 30 e menos de 60 anos se sitúan nunha porcentaxe que supera o 40 %.

3.2.7. Fiabilidade e validez nas probas Celga

Neste apartado analizaremos a fiabilidade e a validez da probas Celga que miden a competencia comunicativa escrita e oral dos candidatos. Fiabilidade e validez son conceptos fundamentais para calquera instrumento de medición ou

diagnóstico; por unha banda, debe medirse con precisión, estabilidade e coherencia (fiabilidade), achegando medidas axeitadas segundo diversas facetas destas que posteriormente abordamos, por outra, debemos obter probas de que o que estamos a medir é o que é necesario medir e non outras cousas (validez).

Como sinala Kerlinguer (1975), en termos absolutos a fiabilidade pode definirse como a proporción da varianza total das puntuacións obtidas cun procedemento que é a varianza verdadeira, isto é, da mesma forma que se descompón a puntuación observada en puntuación verdadeira e erro aleatorio de medida ($X=V+e$), onde a correlación é nula, tamén poden descompoñerse as súas varianzas ($S_t^2= S_v^2+ S_e^2$). Como dende o ámbito das ciencias sociais o concepto teórico da puntuación verdadeira separado do erro de medida non está dispoñible, trabállase desde enfoques relativos de fiabilidade.

No campo da educación nada é constante nin os fenómenos se repiten máis ca aparentemente, polo tanto, os metodólogos desenvolveron un conxunto de estratexias para analizar a fiabilidade relativa da medida desde diversos planos. Estas estratexias que analizan a fiabilidade dos instrumentos son (García, 2003):

1. A estabilidade ou constancia das puntuacións obtidas polos mesmos suxeitos nunha mesma proba aplicada en dúas ocasións, o que dá lugar ao procedemento *test-retest*, no que a correlación estima a fiabilidade. Isto coñécese como coeficiente de estabilidade, o cal pretende estimar o erro de aplicación do instrumento. No caso do Celga, o posible erro de medida é o relacionado coa corrección da proba por parte dos avaliadores, de aí que nos centraremos na fiabilidade do control deste erro.

2. A equivalencia das puntuacións ou resultados obtidos polos mesmos suxeitos sobre a base de dúas probas paralelas ou instrumentos considerados equivalentes.

3. A consistencia interna ou coherencia interna das puntuacións obtidas no marco dun mesmo procedemento de medida parte do suposto de que todos os elementos do procedemento (neste caso criterios de corrección e avaliación) conducen á medida dun mesmo trazo: a competencia comunicativa.

Como esbozabamos ao inicio, o concepto de validez pódese definir dunha forma moi sinxela como o grao no que un procedemento de medida recolle o que pretende medir e non outras cousas (Garrett, 1937). Na actualidade, están consolidados tres enfoques que resumen as achegas básicas sobre o problema da validez dos procedementos de medida: de contido, de criterio (concorrente e predictiva) e de construto.

1. A **validez de contido** na práctica consiste no exame sistemático do contido dun instrumento para determinar se é unha mostra representativa e relevante do dominio (Martínez, 1996). Este proceso alude de forma necesaria á participación de expertos na definición rigorosa dos dominios e compara a valoración da relevancia e representatividade dos elementos considerados no procedemento de medida. Polo tanto, a discusión deste tipo de validez ten compoñentes basicamente conceptuais e teóricos máis ca empíricos.

2. A **validez criterial**, referida á correlación entre as puntuacións de medida obtidas co procedemento que se quere validar e un criterio externo a el. Esta correlación constitúe o **coeficiente empírico de validez criterial**. Existen dous tipos de validez criterial; considérase validez concorrente a correlación entre dúas puntuacións de dous procedementos que se aplican conxuntamente e

validez predictiva cando os datos do criterio se recollen *a posteriori* á medida (García, 2003).

3. A **validez de construto**, referida ao grao no que un procedemento de medida considera de forma axeitada o construto teórico que pretende medir e en que medida as hipóteses derivadas del se confirman empiricamente mediante o citado procedemento (García, 1986, 2003).

3.2.7.1. *Fiabilidade*

Para comprobar a fiabilidade do procedemento de avaliación da competencia comunicativa, analizouse, en primeiro lugar, a súa consistencia interna utilizando o coeficiente de correlación de alpha de Cronbach. Para esta análise, por unha banda, agrupáronse todos os parámetros de avaliación da competencia comunicativa na proba escrita e na oral e, pola outra, utilizáronse todos os candidatos dos anos 2009 e 2010 despois de comprobar que non existían diferenzas na consistencia interna da proba na súa aplicación nas diferentes convocatorias. Tamén se comprobou a consistencia sen o parámetro de avaliación na competencia oral de interacción e con el, xa que a súa valoración se aplicaba consistentemente ata os niveis Celga 1 e Celga 2 e non nos niveis de Celga 3 e Celga 4. En definitiva, a proba alpha de Cronbach deu un coeficiente estandarizado de 0,863 ($F(13,4087) = 654,468, p < 0,001$), resultado que indica un grao de homoxeneidade entre os compoñentes bastante alto. A media de correlacións interítems é de 0,311, cun máximo de 0,746 e un mínimo de -0,023, a desviación típica é de 0,20. Este coeficiente aumenta ata 0,875 ($F(12,7958) = 654,468, p < 0,001$) cando non se considera na análise o

criterio de interacción, subindo o número de casos incluídos na análise a 7.959 fronte aos 4.088 no caso de ter en conta o criterio.

Analizando a relación de cada parámetro co total, obsérvase que a eliminación dalgún dos criterios baixaría o coeficiente de correlación, excepto a pronuncia, que o aumentaría lixeiramente (0,869). Non obstante, se se analiza soamente a competencia comunicativa na proba oral, obsérvase que o coeficiente de correlación é alto (0,809) e a eliminación de calquera dos parámetros faríao diminuír, incluso o criterio pronuncia, aínda que en menor medida ca os demais.

Neste tipo de probas, para estimar empiricamente os estatísticos de fiabilidade (erro típico de medida e coeficiente de fiabilidade) utilízanse métodos para medir a **consistencia das puntuacións dos distintos cualificadores** (Prieto e Delgado, 2010). Para minimizar este erro de medida no procedemento de avaliación do Celga, utilizáronse cualificadores de control (véxase apartado de procedemento). Non obstante, o procedemento de control céntrase nos falsos negativos (isto é, non aptos) e non hai un control tan exhaustivo nos falsos positivos (ou sexa, aptos).

Se analizamos o coeficiente de correlación dos parámetros de avaliación da competencia comunicativa na escrita tanto para a proba A1 como para a A2 (ambas as probas, A1 e A2, forman o total da proba escrita) entre os primeiros avaliadores e os avaliadores supervisores, obsérvase a existencia dunhas correlacións significativas en todos os parámetros; porén, estas correlacións non son especialmente altas sobre todo nalgún parámetro de corrección e avaliación como léxico 2 (o aplicado á proba A2). Chama a atención que na maioría dos casos as correlacións entre criterios diferentes

acadan correlacións maiores ca os mesmos criterios avaliados por avaliadores diferentes (véxanse as táboas 4 e 5).

Criteria	N	Alpha de Pearson	P
Proba A1: Adecuación	3138	0,454**	p < 0,001**
Proba A1: Cohesión	3138	0,392**	p < 0,001**
Proba A1: Gramática	3138	0,400**	p < 0,001**
Proba A1: Léxico	3138	0,289**	p < 0,001**
Proba A2: Adecuación	3138	0,365**	p < 0,001**
Proba A2: Cohesión	3138	0,346**	p < 0,001**
Proba A2: Gramática	3138	0,350**	p < 0,001**
Proba A2: Léxico	3138	0,264**	p < 0,001**

Táboa 4. Coeficientes de correlación nas cualificacións dos avaliadores, coordinadores e supervisores na proba de competencia comunicativa escrita

Para afondar na consistencia dos diferentes avaliadores da proba escrita, dividíronse as puntuacións dos parámetros de avaliación en 1 e 2 para os non aptos e 3 e 4 para os aptos, malia esta non ser a forma exacta para considerar un apto ou un non apto no Celga, si se pode considerar que o punto de corte da escala sería acadar máis dun 2.

Sistema de certificación de lingua galega (Celga) no *Marco Común Europeo de Referencia para as Linguas*. Estudo de fiabilidade e validez

Adecuación proba A1		<i>Cualificación do avaliador_1</i>		Total
		Non apto coinciden avaliador e supervisor	Apto coinciden avaliador e supervisor	
<i>Cualificación do supervisor_2</i>	Non apto	63,2 %	23,1 %	43,4 %
	Apto	36,8 %	76,9 %	56,6 %
Total (70,05 %)		100 (50,5 %)	100 (49,5 %)	100 %
Cohesión proba A1		<i>Cualificación do avaliador_1</i>		Total
		Non apto	Apto	
<i>Cualificación do supervisor_2</i>	Non apto	51,4 %	23,3 %	34,9 %
	Apto	48,6 %	76,7 %	65,1 %
Total (64,05 %)		100 (41,1 %)	100 (58,9 %)	100 %
Gramática proba A1		<i>Cualificación do avaliador_1</i>		Total
		Non apto	Apto	
<i>Cualificación do supervisor_2</i>	Non apto	67,9 %	34,5 %	57,2 %
	Apto	32,1 %	65,5 %	42,8 %
Total (66,7 %)		100 (68 %)	100 (32 %)	100 %
Léxico proba A1		<i>Cualificación do avaliador_1</i>		Total
		Non apto	Apto	
<i>Cualificación do supervisor_2</i>	Non apto	39,8 %	19,4 %	23,8 %
	Apto	60,2 %	80,6 %	76,2 %
Total (48,8 %)		100 % (21,8 %)	100 % (78,2 %)	100 %
Adecuación proba A 2		<i>Cualificación do avaliador_1</i>		Total
		Non apto	Apto	
<i>Cualificación do supervisor_2</i>	Non apto	59,7 %	25,8 %	42,1 %
	Apto	40,3 %	74,2 %	57,9 %
Total (66,95 %)		100% (48,2 %)	100% (51,8 %)	100 %
Cohesión proba A2		<i>Cualificación do avaliador_1</i>		Total
		Non apto	Apto	
<i>Cualificación do supervisor_2</i>	Non apto	47,4 %	24 %	33,8 %
	Apto	52,6 %	76 %	66,2 %
Total (61,7 %)		100% (42,1%)	100% (57,9%)	
Gramática proba A2		<i>Cualificación do avaliador_1</i>		Total
		Non apto	Apto	
<i>Cualificación do supervisor_2</i>	Non apto	64,6 %	35,3 %	55,1 %
	Apto	35,4 %	64,7 %	44,9 %
Total (64,65 %)		100 (67,7 %)	100 (32,3 %)	100 %
Léxico proba A2		<i>Cualificación do avaliador_1</i>		Total
		Non apto	Apto	
<i>Cualificación do supervisor_2</i>	Non apto	38,3 %	19,8 %	23,5 %
	Apto	61,7 %	80,2 %	76,5 %
Total (59,25 %)		100 (20 %)	100 (80 %)	

Táboa 5. Relacións porcentuais entre as cualificacións dos avaliadores e supervisores na proba de competencia comunicativa escrita despois da transformación dos criterios en apto e non apto.

En xeral, a táboa 5 que relaciona as puntuacións de apto e non apto dos criterios de competencia comunicativa escrita nunha mostra de exercicios avaliada polos avaliadores coordinadores (primeiros avaliadores) e avaliadores supervisores (segundos avaliadores) indica que o grao de concordancia¹³ xeral está entre o 70,05 % da adecuación da proba A1 e o 48,8 % da proba A1 de léxico. O grao de concordancia medio en todos os parámetros de avaliación da competencia escrita é dun 62,77 %. En liñas xerais, advírtese que hai unha diferenza na rigorosidade entre os coordinadores e supervisores, os primeiros son máis rigorosos ca os segundos, de aí que exista unha tendencia en todos os parámetros, excepto léxico, a considerar menos non aptos a máis xente na primeira avaliación ca na segunda. Este é un erro que afecta seriamente á fiabilidade na consistencia dos avaliadores que comprobaremos a través dos índices kappa.

A continuación aplicouse o índice kappa, que ofrece unha estimación do grao de acordo non debido ao azar a partir da proporción de acordo observado e a proporción de acordo esperado. Este proba aplícase para analizar o grao de concordancia cando se parte de variables dicotómicas, a proba kappa é a máis empregada cando se avalía a fiabilidade interxuíces ou interobservadores. Na proba kappa todos os resultados inferiores a 0,20 se consideran malos, nunha escala que vai de 0 a 1; neste caso os resultados expostos na táboa 6 indican que o grao de concordancia entre os avaliadores nos parámetros de léxico-proba 1 (0,20) e coherencia e cohesión-proba 1 (0,17) son **malos** (< 0,20), mentres que o de adecuación-proba 1 (0,40) se achega a **moderado** (moderado 0,41-0,60). No resto de parámetros, o grao de concordancia é **baixo** (0,21-0,40).

¹³ O grao de concordancia obtense de sumar e dividir entre dous a porcentaxe acordo entre os supervisores e avaliadores para os aptos e para os non aptos.

Parámetros de avaliación	Coefficiente kappa
Adecuación-proba A1	0,40
Coherencia Cohesión-proba A1	0,17
Gramática-proba A1	0,31
Léxico-proba A1	0,20
Adecuación-proba A 2	0,34
Coherencia Cohesión-proba A 2	0,24
Gramática-proba A2	0,27
Léxico-proba A2	0,17

Táboa 6. Coeficiente kappa para analizar a consistencia entre os avaliadores e os supervisores nas probas de competencia comunicativa escrita

Fiabilidade da proba oral

Logo de analizar a fiabilidade interxuíces na proba de competencia comunicativa escrita, estudouse este mesmo parámetro na proba de competencia comunicativa oral. Para estudar a consistencia da avaliación dos avaliadores utilizáronse as correccións do entrevistador e do avaliador na primeira fase do procedemento, antes de que puidesen chegar a calquera acordo sobre as cualificacións. Non utilizamos os datos da segunda fase, onde se teñen que poñer de acordo en caso de disonancia, nin as cualificacións da comisión avaliadora en caso de desacordo entre entrevistador e avaliador porque non serven para analizar a consistencia interxuíces para a fiabilidade da proba, xa que soamente serve como elemento de control da proba.

A relación entre cualificacións xerais na proba de competencia comunicativa oral dos avaliadores e entrevistadores presentámolas na táboa 7.

Os resultados indican a existencia dun alto grao de acordo na adxudicación do apto (91,5 %), porcentaxe que baixa cos non apto (83,4 %); o grao de concordancia xeral é do 87,45 %. O coeficiente kappa xeral é 0,741 ($p < 0,001$), é dicir, un coeficiente **bo** (0,61-0,81), máis cerca de **excelente** (0,81-1,00) ca de **moderado** (0,41-0,60).

Despois de analizar as cualificacións xerais na proba oral do Celga, estudamos a fiabilidade destas observando a relación entre a cualificación outorgada polo entrevistador e polo avaliador (*vid.* táboa 7). Os resultados indican a existencia dun alto grao de acordo na adxudicación do apto (91,5 %), porcentaxe que baixa lixeiramente cos non apto (83,4 %).

ORAL		Cualificación do entrevistador		Total
		Non apto	Apto	
Cualificación do avaliador	Non apto	83,4 %	16,6 %	100 %
	Apto	8,5 %	91,5 %	100 %
Total		31 %	69 %	100 %

Táboa 7. Relación entre a cualificación da proba oral do avaliador e do entrevistador

Analizando máis detidamente a relación entre as puntuacións dos avaliadores e dos entrevistadores, estudáronse as puntuacións adxudicadas aos mesmos criterios de corrección e avaliación na proba oral utilizando o estatístico de correlación de Pearson. Os resultados amosaron unha alta e positiva correlación en todos os criterios, superior en todos os casos ao $r = 0,600$ (*vid.* táboa 8). A correlación máis alta foi no caso do criterio de gramática ($r = 0,665$; $p < 0,001$) e a máis baixa a de interacción ($r = 0,639$; $p < 0,001$), é dicir, no caso da gramática o avaliador e o entrevistador comparten o 47,2 % da varianza e un 41 % no caso da interacción. Ao contrastar os resultados desta táboa (*vid.* táboa 8) coa anterior (*vid.* táboa 7), pode observarse

que a varianza está centrada nas puntuacións extremas, non nas puntuacións de corte que diferencian o non apto do apto.

	N	Correlación de Pearson	Sig.
Adecuación	7967	0,665	p < 0,001
Gramática	7967	0,687	p < 0,001
Léxico	7967	0,656	p < 0,001
Cohesión	7967	0,652	p < 0,001
Pronuncia	7967	0,663	p < 0,001
Interacción	7967	0,639	p < 0,001

Táboa 8. Correlación entre os mesmos criterios de cualificación oral dos avaliadores e entrevistadores

Os **graos de concordancia** obtidos de cruzar as cualificacións dos entrevistadores e dos avaliadores en cada un dos criterios de avaliación da competencia comunicativa oral son altas, xa que superan o 80 % –oscilan entre o 82,55 % do parámetro cohesión ata o 86,95 % do parámetro gramática (táboa 9)–. Non obstante, ao examinar máis polo miúdo os datos presentados na táboa 9, pode observarse que as diferenzas na adxudicación do apto ou do non apto en cada criterio se centran sistematicamente no non apto. O entrevistador considera aptos os que o avaliador considera non aptos (entre un 21,2 % do criterio gramática ata un 31 % do criterio de cohesión). Apenas hai diferenzas na consideración de aptos, o entrevistador soamente considera non aptos os que o entrevistador considera aptos (entre un 3,6 % do criterio interacción e un 5,1 % do criterio léxico). Parece que hai diferentes criterios para adxudicar as puntuacións máis baixas da escala polos diferentes avaliadores, circunstancia que se repetía na avaliación da competencia comunicativa escrita e que baixa a fiabilidade da escala.

Adecuación		<i>Cualificación do entrevistador</i>		Total
		Non apto	Apto	
<i>Cualificación do avaliador</i>	Non apto	72,1 %	27,9 %	100 %
	Apto	3,8 %	96,2 %	100 %
Total (84,15 %)		11,6 %	88,4 %	100 %
Gramática		<i>Cualificación do entrevistador</i>		Total
		Non apto	Apto	
<i>Cualificación do avaliador</i>	Non apto	78,8 %	21,2 %	100 %
	Apto	4,9 %	95,1 %	100%
Total (86,95%)		18 %	82 %	100 %
Léxico		<i>Cualificación do entrevistador</i>		Total
		Non apto	Apto	
<i>Cualificación do avaliador</i>	Non apto	75,3 %	24,7 %	100 %
	Apto	5,1 %	94,9 %	100 %
Total (85,1 %)		16,1 %	83,9 %	100 %
Cohesión		<i>Cualificación do entrevistador</i>		Total
		Non apto	Apto	
<i>Cualificación do avaliador</i>	Non apto	69 %	31 %	100 %
	Apto	3,9 %	96,1 %	100 %
Total (82,55 %)		11,1 %	88,9 %	100 %
Pronuncia		<i>Cualificación do entrevistador</i>		Total
		Non apto	Apto	
<i>Cualificación do avaliador</i>	Non apto	73,6 %	26,4 %	100 %
	Apto	3,7 %	96,3 %	100 %
Total (84,95 %)		10,2 %	89,8 %	100 %
Interacción		<i>Cualificación do entrevistador</i>		Total
		Non apto	Apto	
<i>Cualificación do avaliador</i>	Non apto	74 %	26 %	100 %
	Apto	3,6 %	96,4 %	100 %
Total (85,2 %)		9,5 %	90,5 %	100 %

Táboa 9. Relacións porcentuais entre as cualificacións dos avaliadores e dos entrevistadores na proba de competencia comunicativa oral despois da transformacións dos criterios en apto e non apto

O coeficiente kappa aplicado a cada un dos parámetros (táboa 9) indica o significativo nivel de acordo en todos os criterios de avaliación da competencia comunicativa oral, xa que este coeficiente oscila 0,65 de cohesión e o 0,73 de gramática. É dicir, pódese considerar que o nivel de acordo é bo en todos os criterios, especialmente no parámetro de gramática.

Parámetros de avaliación	Coeficiente kappa
Adecuación	0,68**
Gramática	0,73**
Léxico	0,69**
Cohesión	0,65**
Pronuncia	0,67**
Interacción	0,66**

** $p < 0,001$

Táboa 10. Coeficiente kappa para analizar a consistencia entre os avaliadores e os entrevistadores nas probas de competencia comunicativa oral

3.2.7.2. Validez

3.2.7.2. 1. Validez de contido

A **validez de contido** (*content validity*) depende dunha análise sistemática do contido da proba coa finalidade de demostrar que esta contén unha mostra representativa das destrezas lingüísticas relevantes. A dita validez precisa da recollida das opinións de expertos, persoas nas que se está disposto a confiar aínda que se discrepe coa propia opinión (Alderson, 1998). Isto implica que os expertos emitan xuízos de forma sistemática, analizando o contido dunha proba

e comparándoo coa definición do que debería ser o seu contido que aparece plasmado nas especificacións desa proba.

No caso do Celga, os criterios de corrección e avaliación empregados para a proba oral e escrita aparecen definidos nas especificacións da proba e son froito dunha fonda reflexión por parte do grupo de expertos durante meses. O mesmo ocorre co propio deseño das tarefas que os candidatos deben realizar; o feito de seren probas nas que as destrezas aparecen integradas e así deben ser avaliadas supón un coidado maior á hora de cumprir as especificacións. Así, por cada proba deseñada realízase unha serie de análises e pilotaxes que tentan demostrar que realmente están a avaliar o que din que avalían. O paradoxo desta situación na que se utilizan expertos é que se considera que as opinións dos expertos son decisivas sempre e cando se escollan as persoas adecuadas para emitir os xuízos, pero desgraciadamente isto non sempre é así, e debe terse en conta que a validez dunha proba é máis relativa ca absoluta, se os problemas de validez de contido son evidentes, deben recollese outros tipos de evidencia como a validez de construto ou a validez aparente ou de resposta. É necesario aceptar que na avaliación de linguas aínda existe moito terreo descoñecido e hai que continuar á busca da validez ou modificar a proba ou as especificacións que a definen sempre en prol dunha mellora do sistema de avaliación.

Con respecto aos criterios de corrección e avaliación empregados no Celga, xorde sempre a necesidade por parte do grupo de investigación de melloralos, e con ese espírito de crítico organízanse xornadas de reflexión sobre estes, con axuda de expertos externos ao sistema que ofrecen unha visión nova e permiten planificar cambios que redunden en beneficio da validez. Lembremos que no estudo I os avaliadores confirmaron a validez de contido da proba outorgándolles aos criterios de corrección e avaliación uns valores altos na utilidade para o seu labor, ademais de considerar que todos eles eran necesarios e non solicitaron que se eliminase ningún deles.

3.2.7.2. 2. Validez de construto

Para contrastar a validez de contido do instrumento de medida da competencia comunicativa escrita e oral xustificada polo grupo de expertos, aplicouse unha análise factorial exploratoria (ver unha descrición máis exhaustiva da técnica en Hair, Anderson, Tatham e Black, 1999; García, Gil, e Rodríguez, 2000) aos resultados nos criterios de avaliación da competencia comunicativa escrita, por unha banda, e aos resultados nos criterios de avaliación da competencia comunicativa oral, pola outra.

Para analizar a validez de construto do instrumento de medida da avaliación da competencia comunicativa escrita composto polos mesmos catro criterios da proba A1 e da A2 (adecuación á tarefa, coherencia e cohesión, léxico e gramática), comprobouse a pertinencia de aplicar a análise factorial exploratorio mediante a proba de adecuación mostral de Kaiser-Meyer-Olkin (KMO), con valores que flutúan entre 0 e 1. Convén lembrar que, para a utilización da análise factorial exploratoria, se aconsella acadar valores por riba de 0,5, o que indica que as correlacións entre os pares poden explicarse polo efecto doutras variables. A maiores tamén se aplicou a proba de esfericidade de Barlett, que informa da ausencia de correlacións significativas entre as variables, circunstancia que, de darse, desaconsellaría a realización de análises deste tipo. O resultado na proba KMO foi superior a $> .80$ e a proba de esfericidade de Barlett foi significativa. Estes valores indican que é pertinente realizar a análise factorial da matriz de correlacións (Tabachinnick e Fidell, 1989: 604).

A través da análise de compoñentes principais (método de rotación ortogonal varimax), obtivéronse dous factores, con *eigenvalue* con parada en

0,95,¹⁴ que explicaron un 73,37 % da varianza total. O primeiro factor explica un 46,25 % da varianza total e o segundo un 27,11 %. Na Figura 47 aparecen os pesos ou cargas relevantes de cada criterio en cada factor. Considéranse relevantes para cada factor aqueles criterios que amosan unha carga superior a 0,40, criterio máis restritivo ca o utilizado por outros autores que empregan 0,30.¹⁵

O primeiro factor está composto polos criterios de gramática, léxico e coherencia e cohesión das probas A1 e A2, a porcentaxe de varianza que explica é dun 46,25 %. Pódese denominar coñecemento do código.

O segundo factor está composto polos criterios de adecuación á tarefa na proba A1 e A2 e polos criterios de coherencia e cohesión na proba A1 e A2 compartidos co factor primeiro. Este factor pódese denominar de estrutura textual.

	F ₁	F ₂
<i>Gramática_A1</i>	,843	
<i>Gramática_A2</i>	,835	
<i>Léxico_A1</i>	,822	
<i>Léxico_A2</i>	,818	
<i>Cohesión_A2</i>	,657	,535
<i>Cohesión_A1</i>	,645	,512
<i>Adecuación_A1</i>		,847
<i>Adecuación_A2</i>		,828
% varianza	46,25 %	27,11 %

Figura 47. Análise factorial das competencias comunicativas escritas (rotación ortogonal varimax)

¹⁴ Para establecer o punto de parada non se utilizou o método de Guttman-Kaiser que o establece en 1, senón que se analizou o gráfico de sedimentación.

¹⁵ Para máis información, en DiStefano, Zhu e Mindrila (2009).

Os resultados previos para comprobar a idoneidade da análise factorial exploratoria realizada sobre os criterios para avaliar a competencia comunicativa oral foron favorables. O valor do índice KMO foi de 0,793. Pola súa parte, o test de Barlett rexeitou que a matriz fose de identidade ($\chi^2 = 12.420,671$ g.l = 15, $p < 0,001$), polo que se considera que a factorización é unha técnica adecuada para analizar os criterios de avaliación da competencia comunicativa oral.

A través da análise de compoñentes principais (método de extracción rotación ortogonal varimax), obtivéronse dous factores que explicaron un 77,86 % da varianza total. Na Figura 48 aparecen as cargas relevantes de cada criterio introducido na análise en cada factor. Como sinalabamos anteriormente, considéranse relevante para cada factor aqueles criterios que superan o peso de 0,40 (véxase Figura 48).

	F ₁	F ₂
<i>Gramática</i>	,891	
<i>Pronuncia</i>	,836	
<i>Léxico</i>	,809	
<i>Interacción</i>		,902
<i>Adecuación</i>		,871
<i>Cohesión</i>		,799
% varianza	39,03 %	38,84 %

Figura 48. Análise factorial das competencias comunicativas orais (rotación ortogonal varimax)

O primeiro factor está composto polos criterios de gramática, pronuncia e léxico e explica un 39,03 % da varianza total. O factor pódese denominar de coñecemento do código.

O segundo factor está composto polos criterios de interacción, adecuación e cohesión e explica un 38,84 %, case o mesmo ca o primeiro factor; pódese denominar de estrutura textual.

3.2.8. Resultados da proba escrita

A continuación presentamos unha análise polo miúdo dos datos obtidos na proba escrita en todos os niveis de Celga, unha mostra que atinxe dende os datos máis xerais, aqueles que nos permiten coñecer a porcentaxe dos candidatos aptos e non aptos, ata unha descrición detallada na que, atendendo a variables sociodemográficas como o sexo, o lugar de residencia, idade e o nivel de estudos, poderemos comprobar se son relevantes para acadar a aptitude no certificado de lingua.

Así, como vemos na Figura 49, os resultados indican que existe unha porcentaxe do 54,8 % de candidatos aptos fronte ao 41,6 % de candidatos que non superou a proba escrita.

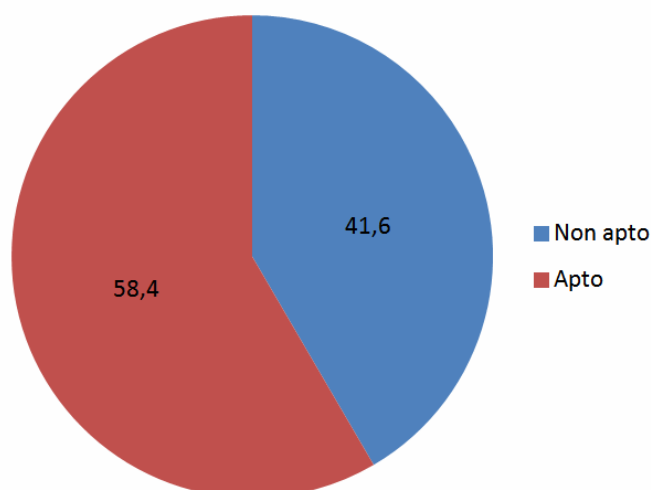


Figura 49. Cualificación na proba escrita

Lembramos que a proba escrita se divide en dúas partes: a proba A1, na que os candidatos a partir dun audio teñen que construír un texto activando as destrezas de comprensión oral e expresión escrita; e unha segunda proba denominada A2, na que partindo como *input* dun texto escrito os candidatos deben redactar outro, demostrando a súa competencia comunicativa nas destrezas de comprensión escrita e expresión escrita.

En primeiro lugar, presentamos unha análise dos datos obtidos grazas ás correlacións das puntuacións acadadas polos candidatos na proba escrita, tanto na proba A1 coma na proba A2 e en todos os niveis de Celga.

Como xa adiantamos no apartado 2.2.a, a proba escrita foi avaliada segundo os parámetros de adecuación á tarefa, coherencia e cohesión, gramática e léxico. As puntuacións que se poden conseguir en cada un dos parámetros seguen unha escala do 1 ao 4. No caso da proba escrita, por tratarse en realidade de dúas probas (A1 e A2), permitírase a presenza dun 2 só nun dos parámetros e só nunha das probas (na A1 ou na A2) e a proba considerárase apta. En caso de aparecer máis dun 2, a proba considerárase non apta.

3.2.8.1. Relacións entre os criterios de corrección e avaliación da proba escrita

Podemos observar na táboa seguinte (táboa 11) o nivel de relación entre as puntuacións obtidas nos diferentes parámetros de avaliación tanto da proba A1 coma da A2 polos 8.016 candidatos que realizaron as probas Celga nos anos 2009 e 2010. Na táboa 11, aparecen na horizontal e na vertical todas as variables que forman parte dos criterios de avaliación da proba escrita (tanto da proba A1 coma da A2).

Por exemplo, cada un dos cadros da liña horizontal fai referencia a unha variable e presenta unha serie de números que é a puntuación obtida en

correlación con outra variable na liña vertical; no caso de cohesión 1 en relación con adecuación 1, presenta a seguinte numeración: 0,526: esta cifra é unha puntuación que vai do 1 ao -1 e indica a correlación entre a variable da fila e da columna, se ten un signo positivo implica que nas variables a medida que sobe a puntuación nunha tamén sobe noutra, é correlativo. A medida que o número se achega máis a 1, a correlación entre ambas as variables será maior e, a medida que o número se acerque máis a 0, a correlación será menor. O seguinte número do cadro indica o nivel de significativade da correlación, se é inferior a $p < 0,001$, a relación é significativa. Por último, o último número correspóndese coa mostra coa que se está a realizar a análise, neste caso $N = 8.016$.

	Adecuación _1	Coherencia Cohesión_1	Gramática _1	Léxico_1	Adecuación _2	Coherencia Cohesión_2	Gramática_2	Léxico_2
Adecuación_1	1 ,000 8016	0,526 ,000 8016	0,387 ,000 8016	0,420 ,000 8016	0,574 ,000 8016	0,497 ,000 8016	0,395 ,000 8016	0,404 ,000 8016
Cohesión_1	0,526 ,000 8016	1 ,000 8016	0,609 ,000 8016	0,604 ,000 8016	0,451 ,000 8016	0,767 ,000 8016	0,575 ,000 8016	0,568 ,000 8016
Gramática_1	0,387 ,000 8016	0,609 ,000 8016	1 ,000 8016	0,671 ,000 8016	0,382 ,000 8016	0,591 ,000 8016	0,742 ,000 8016	0,624 ,000 8016
Léxico_1	0,420 ,000 8016	0,604 ,000 8016	0,671 ,000 8016	1 ,000 8016	0,401 ,000 8016	0,578 ,000 8016	0,633 ,000 8016	0,747 ,000 8016
Adecuación_2	0,574 ,000 8016	0,451 ,000 8016	0,382 ,000 8016	0,401 ,000 8016	1 ,000 8016	0,534 ,000 8016	0,422 ,000 8016	0,443 ,000 8016
Cohesión_2	0,497 ,000 8016	0,767 ,000 8016	0,591 ,000 8016	0,578 ,000 8016	0,534 ,000 8016	1 ,000 8016	0,631 ,000 8016	0,643 ,000 8016
Gramática_2	0,395 ,000 8016	0,575 ,000 8016	0,742 ,000 8016	0,633 ,000 8016	0,422 ,000 8016	0,631 ,000 8016	1 ,000 8016	0,687 ,000 8016
Léxico_2	0,404 ,000 8016	0,568 ,000 8016	0,624 ,000 8016	0,747 ,000 8016	0,443 ,000 8016	0,643 ,000 8016	0,687 ,000 8016	1 ,000 8016

Táboa 11. Correlacións entre as puntuacións obtidas polos candidatos na proba escrita

Da análise dos datos podemos extraer a seguinte información:

— **En xeral, existen altas correlacións entre todos os parámetros:** partimos de que todas as variables dependentes están relacionadas porque todas elas presentan puntuacións por riba de 1, algunhas presentan unha puntuación maior e outras menor (por exemplo, unha puntuación de 0,380 é normal pero unha de 0,742 é alta), e como podemos observar na táboa en xeral existen altas correlacións entre todos os parámetros ou variables dependentes. Apréciase na táboa, se facemos unha lectura en diagonal, que aparecen cadros onde coinciden a mesma variable na fila e na columna (adecuación 1 e adecuación 1); nese caso a puntuación obtida é 1 porque non se relaciona o parámetro consigo mesmo.

— **As maiores correlacións obtivéronse entre as probas A1 e A2 excepto no criterio de adecuación á tarefa:** isto pode deberse a que o criterio de adecuación é o máis significativo á hora de decidir se unha proba é apta ou non, xa que para cumprir con este parámetro deben crear un texto cunha modalidade solicitada e, ademais, cumprir cun número determinado de palabras, no caso contrario, ben sexa porque se supera ou non se chega nun 20% ao número de palabras estipuladas, a proba considérase non apta.

— **As correlacións máis baixas son as existentes entre os criterios de adecuación á tarefa e gramática tanto na proba A1 como na proba A2:** obsérvase que na correlación entre as variables dependentes de adecuación á tarefa e gramática tanto na proba A1 como na proba A2 se obteñen os seguintes resultados:

— Adequación 1 e a gramática 1: 0,387

— Adequación 1 e gramática 2: 0,395

— Adequación 2 e gramática 1: 0,382

— Adequación 2 e gramática 2: 0,387

Son resultados baixos comparados co resto das correlacións entre as outras variables dependentes. Isto pode deberse a que sexan os criterios que máis peso teñan para os avaliadores á hora de delimitar se un candidato é apto ou non apto.

3.2.8.2. Cualificación na proba escrita dos non aptos

Por outra banda, segundo a cualificación dos candidatos non aptos, podemos observar que existen diferenzas se se atende ás variables dependentes ou parámetros. Como se pode apreciar na Figura 50, onde na lenda que está sinalada cos números do 1 ao 4 se fai referencia á puntuación obtida en cada un dos parámetros, lembremos que a proba escrita se considera apta sempre que se obtivesen puntuacións de 3 e 4 en todos os parámetros, agás nun. Vemos que aparecen por duplicado as lendas referidas aos parámetros, xa que se refiren ás dúas probas escritas, por exemplo, a adecuación I fai referencia á da proba A1 e a adecuación II correspóndese coa da proba A2.

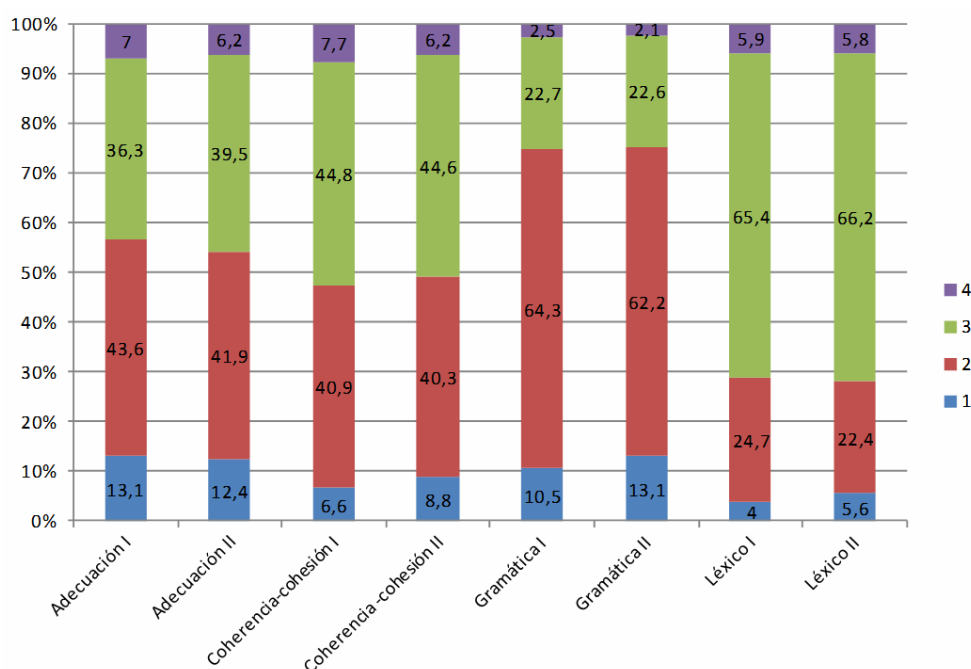


Figura 50. Cualificación na proba escrita dos non aptos

1 punto: nos parámetros de adecuación á tarefa nas probas A1 e A2, observamos as porcentaxes máis altas (13,1% e 12,4 %, respectivamente) ao lado das puntuacións de gramática I (10,5 %) e gramática II (13,1 %); fronte ás máis baixas de léxico I (4 %) e léxico II (5,6 %). Situado nun plano intermedio, atopamos a coherencia e cohesión (6,6 % na proba A1 e 8,8 % na proba A2).

2 puntos: neste caso o parámetro que destaca é o de gramática, que obtén un 64,3 % na proba A1 e un 62,2 % na proba A2, e esta variable dependente é a que parece máis significativa á hora de marcar a diferenza entre os candidatos non aptos; seguida da adecuación á tarefa e coherencia e cohesión, que parecen ser os parámetros que na media entre 40 % e 43,6 % van delimitar a aptitude. Pola contra, o léxico presenta os resultados máis baixos, cun 24,7 % (proba A1) e 22,4 % (proba A2).

3 puntos: a diferenza coas puntuacións anteriores nos que primaban a gramática e a adecuación, xorde o léxico cunha puntuación elevada, 65,4 % na proba A1 e 66,2 % na proba A2.

4 puntos: é salientable como neste caso os parámetros máis significativos son a adecuación e coherencia e cohesión, aínda que con puntuacións baixas –entre 6 % e 7 %– nas probas A1 e A2 e, pola contra, é a gramática a variable que obtén a puntuación máis baixa 2,5 % e 2,1 %.

3.2.8.3. Cualificación na proba escrita dos aptos

A continuación presentamos unha análise dos datos obtidos nas probas dos candidatos aptos, onde observaremos diferenzas nos parámetros que delimitan a aptitude dos candidatos e analizaremos aqueles parámetros que obtiveron puntuacións que van do 2 ao 4 –lembremos que na proba escrita só pode obterse un 2 nalgún dos parámetros das probas A1 e A2, o resto teñen que ser 3 e 4–:

2 puntos: a adecuación á tarefa (10,2 %) seguida da gramática son as variables cunha maior puntuación, que oscila entre o 8,8 % e 8,9 % da gramática nas probas A1 e A2. O parámetro do léxico, cun 0,7 %, é o que obtén o resultado máis baixo.

3 puntos: é a gramática o parámetro que máis puntuación obtén (64,1 % na proba A1 e 64,4 % na proba A2) seguida da variable do léxico (54,1 % na proba A1 e 55 na proba A2).

4 puntos: neste caso é a variable da coherencia e cohesión a que obtén un maior resultado (47,3 % na proba A1 e 46,2 % na proba A2), seguida do léxico (44,9 % na proba A1 e 44,3 % na proba A2).

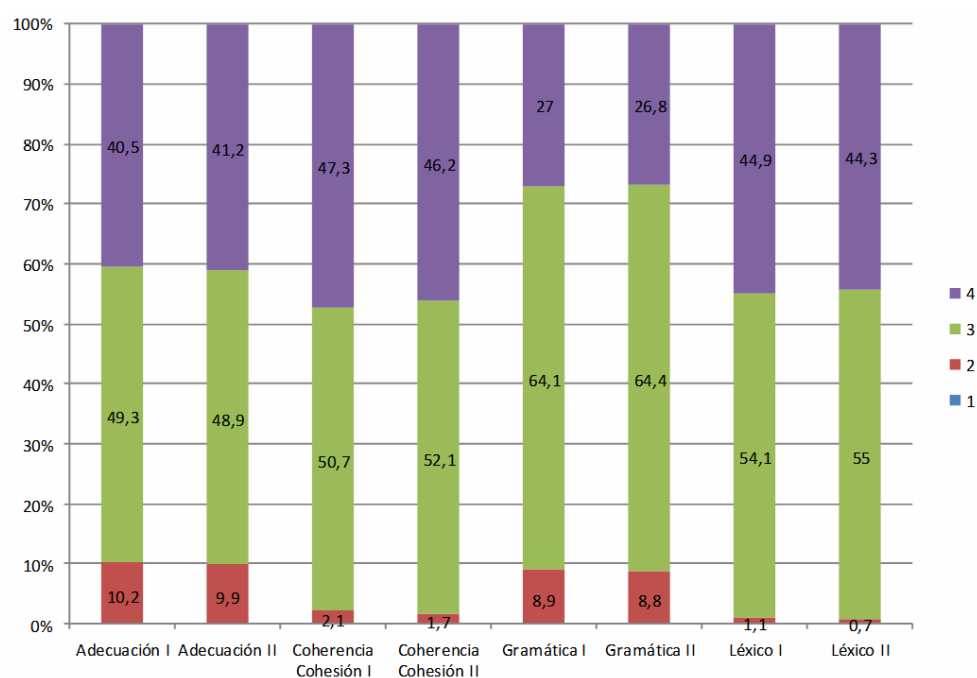


Figura 51. Cualificación na proba escrita dos aptos

No seguinte cadro comparativo entre os resultados dos candidatos aptos e non aptos, observamos como, dependendo de cada nivel, cambia a variable significativa que obtén unha puntuación maior e menor.

Sistema de certificación de lingua galega (Celga) no *Marco Común Europeo de Referencia para as Linguas*. Estudo de fiabilidade e validez

CANDIDATOS NON APTOS		CANDIDATOS APTOS	
Maior puntuación	Menor puntuación	Maior puntuación	Menor puntuación
1 punto Adecuación: 13,1 %	1 punto Léxico: 4 %	Non consta	Non consta
2 puntos Gramática: 64,3 %	2 puntos Léxico: 22,4 %	2 puntos Adecuación: 10,2 %	2 puntos Léxico: 0,7 %
3 puntos Léxico: 66,2 %	3 puntos Gramática: 22,6 %	3 puntos Gramática 64,4 %	3 puntos Adecuación: 48,9 %
4 puntos Coherencia e cohesión: 7,7 %	4 puntos Gramática: 2,1 %	4 puntos Coherencia e cohesión: 47,3 %	4 puntos Gramática: 26,8 %

Así, observamos que coinciden como variables dependentes significativas tanto nos candidatos non aptos coma nos aptos, con 4 puntos, a coherencia e a cohesión; este parámetro obtén a puntuación maior fronte á gramática, que obtén a puntuación menor. Pode ser significativo que a gramática é o parámetro relevante que obtén menos puntuación, pero que está presente tamén no apartado dos 3 puntos, onde o seu papel é diferenciado entre os candidatos aptos e os non aptos. Isto pode deberse a que nos niveis altos de lingua, no Celga 3 e Celga 4, a gramática é unha variable cun peso relevante, tanto nos candidatos aptos coma nos non aptos; vemos que a gramática é o parámetro que obtén menor puntuación nos candidatos non aptos fronte aos candidatos aptos onde obtén os resultados máis altos, é dicir, o parámetro da gramática pode ser decisivo á hora de delimitar que unha proba sexa apta ou non.

Non obstante, para os candidatos non aptos, o léxico é a variable dependente que obtén os resultados máis baixos, tanto en 1 punto coma en 2 puntos e, pola contra, a adecuación á tarefa é a que obtén a maior puntuación, con 1 punto dos candidatos non aptos, e a gramática é a que obtén a puntuación maior no apartado de 2 puntos. De novo a gramática volve ser o parámetro que obtén a maior puntuación con 2 puntos. Isto tamén pode ser que sexa o parámetro que delimita que unha proba sexa apta ou non.

3.2.8.4. Cualificación na proba escrita segundo o nivel do Celga

O nivel de Celga ao que concorran os candidatos ten unha influencia significativa nas cualificacións obtidas ($I = 0,026$; $p < 0,001$), como se pode ver na figura seguinte (52).

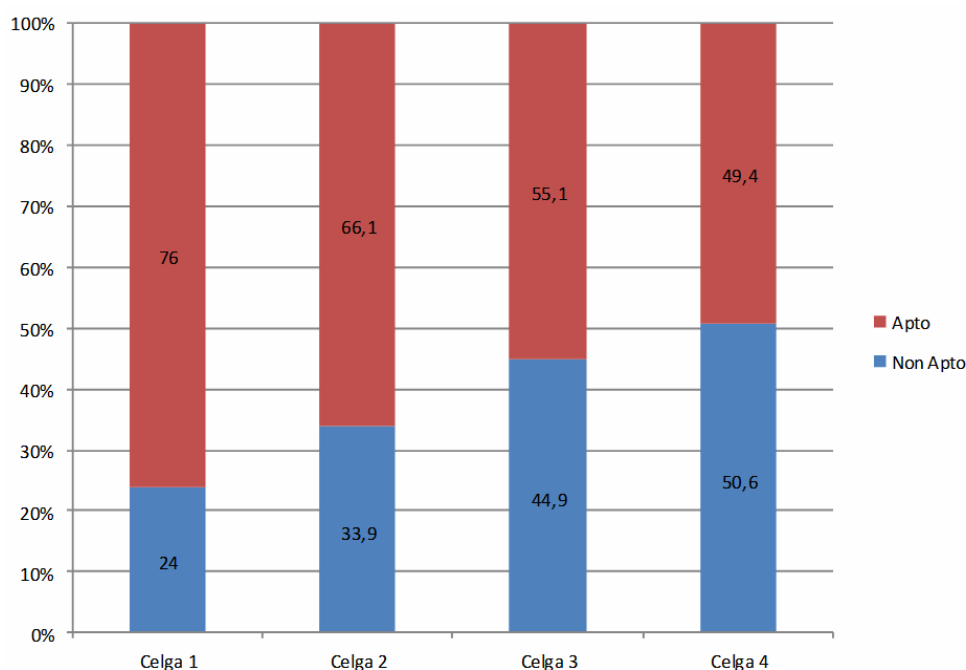


Figura 52. Cualificación na proba escrita segundo o nivel do Celga

Podemos comprobar que o sistema de certificación de lingua galega presenta unha porcentaxe de candidatos aptos na proba escrita proporcional ao nivel de lingua. Nos niveis máis baixos, Celga 1 e Celga 2, o índice de aprobados é superior (75 % e 66,1 %, respectivamente), fronte aos niveis superiores, Celga 3 –o índice de candidatos aptos é dun 55,1 %– e Celga 4 –presenta un 49,4 % de candidatos aptos–. Isto é totalmente normal, xa que, a medida que o nivel de esixencia na competencia lingüística é superior, decrece

o número de candidatos que cumpren os requisitos para demostrar a súa aptitude.

Aínda así, cómpre salientar que no nivel Celga 4, que é o máis alto e complexo á hora de obter a aptitude, debemos ter en conta que o perfil sociolóxico dos candidatos de Celga é maioritariamente de galegos. Ao analizar o lugar de nacemento dos participantes, obsérvase que algo máis da metade dos candidatos naceron en Galicia (55,3 %), o 29,6 % pertence ao resto de rexións do Estado español e hai unha porcentaxe de persoas que naceron no continente americano (8,3 %) e no resto de Europa (6,5 %). Isto condiciona que a proximidade das linguas dos candidatos, español e galego, permita unha porcentaxe tan elevada nos niveis de Celga 3 e Celga 4.

3.2.8.4.1. Cualificación na proba escrita dos non aptos nos niveis Celga 1 e Celga 2

A continuación presentamos os resultados das puntuacións obtidas polos candidatos non aptos nos niveis máis baixos do Celga, que é onde o número de candidatos aptos na proba escrita adoita ser máis elevado, tal e como se presentou no apartado 3.4.4 (Celga 1: 76 % e Celga 2: 66,1 %).

Nótese que na Figura 53 aparecen os parámetros de corrección e avaliación das dúas probas que conforman a proba escrita (A1 e A2) sinalados como adecuación I que pertence á proba A1 e adecuación II pertence á proba A2.

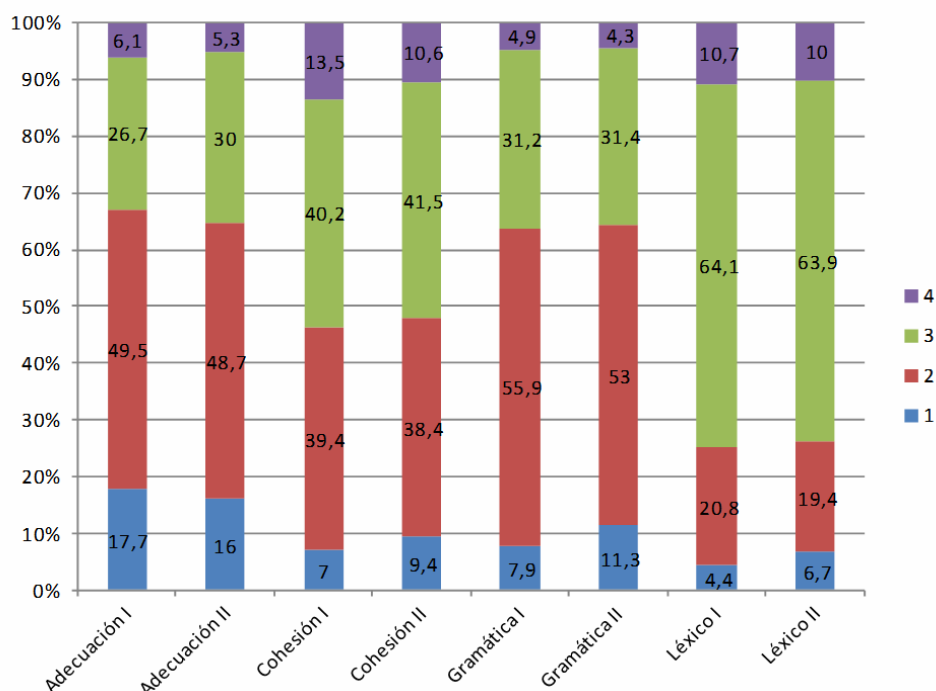


Figura 53. Cualificación na proba escrita dos non aptos nos niveis do Celga 1 e Celga 2

Neste caso, por tratarse de candidatos non aptos, atopamos nas puntuacións máis baixas unha porcentaxe elevada:

1 punto: é o parámetro de adecuación á tarefa o que obtén os resultados máis elevados: 17,7 en adecuación á tarefa I e 16 en adecuación á tarefa II. Isto pode deberse a que os candidatos incumpren o requisito do número de palabras estipulado para a redacción dos textos. Non obstante, é significativa a diferenza entre este criterio e os outros, xa que en adecuación á tarefa I e II só hai un 1,6 % de diferenza, mentres que en cohesión I e II hai un 2,4 % de diferenza, en gramática I e II hai 3,4 % de diferenza e en léxico I e II hai 2,3 % de diferenza.

2 puntos: obsérvase que é unha das puntuacións máis empregadas polos avaliadores e é significativo en parámetros como adecuación á tarefa e gramática; por unha banda, en gramática I obtén o 55,9 % da puntuación e en

gramática II obtén o 53 % e, por outra, en adecuación á tarefa I rexístrase un 49,5 % e en adecuación á tarefa II un 48,7 % da puntuación.

3 puntos: desta puntuación destaca o parámetro de léxico I, que obtén un 64,1 % e léxico II, un 63,9 %.

4 puntos: como era de esperar neste caso, non aparece unha gran porcentaxe nesta puntuación por tratarse de candidatos non aptos. Daquela os parámetros de cohesión e léxico son os que acadan as maiores porcentaxes: 13,5 % en coherencia e cohesión I e 10,6 % en coherencia e cohesión II, mentres que o léxico rolda o 10 % da puntuación.

3.2.8.4.2. Cualificación na proba escrita dos non aptos nos niveis do Celga 3 e Celga 4

Do mesmo xeito ca no apartado anterior, presentamos as puntuacións obtidas nos distintos parámetros de avaliación nos candidatos non aptos nos niveis máis altos, Celga 3 e Celga 4.

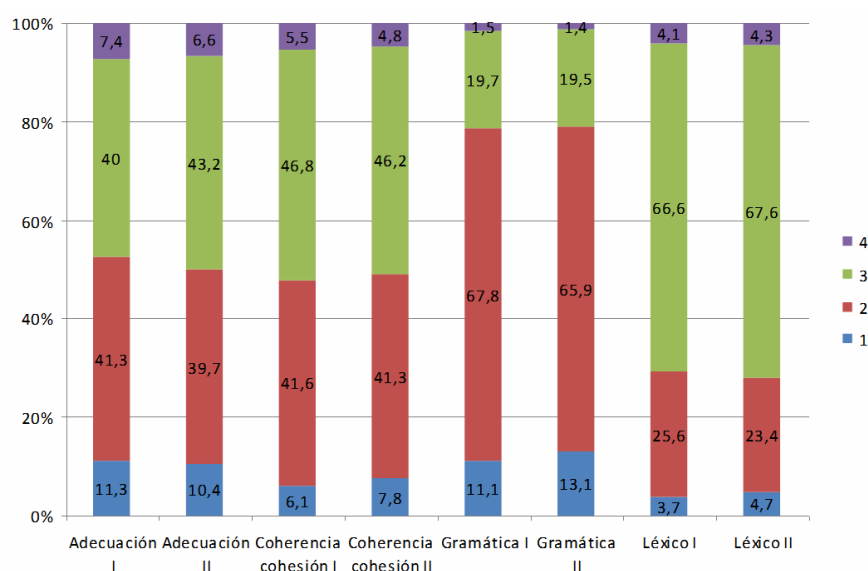


Figura 54. Cualificación na proba escrita dos non aptos nos niveis do Celga 3 e Celga 4

1 punto: nesta categoría a puntuación máis baixa acadaa o léxico I (3,7 %) seguido do léxico II (4,7 %) e, pola contra, a gramática I obtén 11,1 % e a gramática II, 13,1 %.

2 puntos: é significativo que a gramática I acade un 67,8 % e a gramática II un 65,9 %, seguida dos parámetros de adecuación á tarefa e coherencia e cohesión, arredor do 41 %, e, por último, os que acadan as puntuación máis baixas son o léxico I (25,6 %) e o léxico II (23,4 %).

3 puntos: é no léxico I (66,6 %) e léxico II (67,6 %) onde se obteñen as maiores puntuacións, seguidos de cohesión e adecuación á tarefa, que varían dende o 40 % ao 46 % da puntuación e, por último, a gramática acada as porcentaxes máis baixas (19 %).

4 puntos: como cabe esperar por tratarse de probas non aptas, non haberá puntuacións moi elevadas nesta categoría: a mínima é 1,4 % en gramática II e 1,5 % en gramática I; a máxima, 7,4 % en adecuación á tarefa I, seguida de adecuación á tarefa II, 6,6%; os parámetros de léxico e coherencia e cohesión roldan o 5%.

3.2.8.5. Cualificación na proba escrita segundo o nivel de estudos

O nivel de estudos ten unha influencia significativa nas cualificacións obtidas ($I = 0,040$; $p < 0,001$), como se pode ver na figura seguinte (Figura 55), aínda que cómpre sinalar que non se inclúen as persoas que carecen de estudos porque non son suficientes para extraer ningunha conclusión.

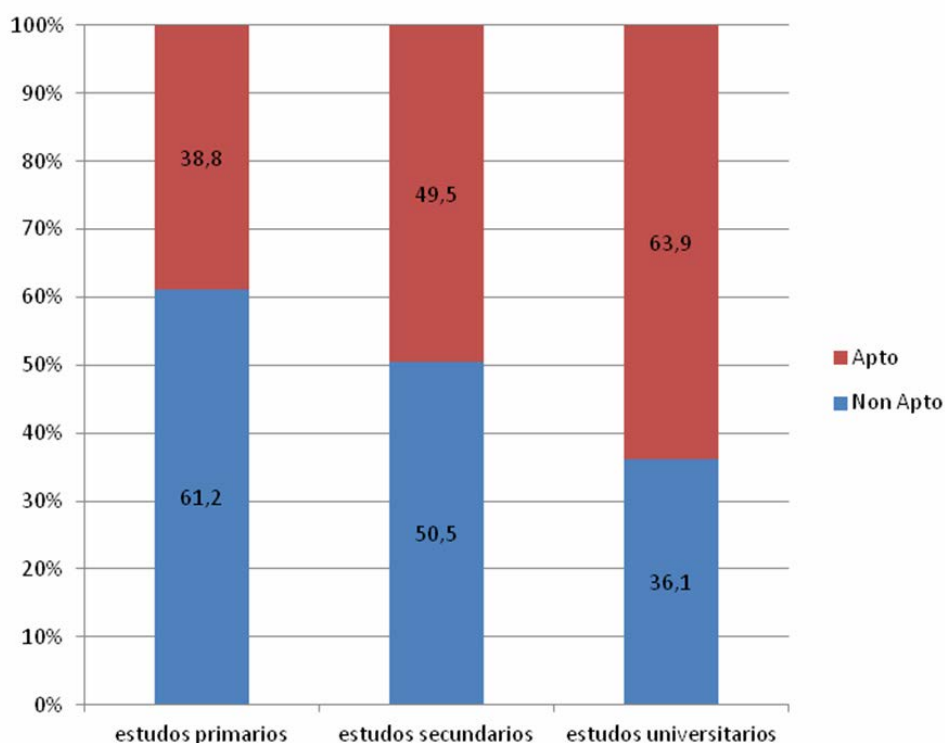


Figura 55. Cualificación na proba escrita segundo o nivel de estudos

Dividimos os candidatos en tres grandes grupos segundo o nivel de estudos: aqueles que teñen estudos primarios, secundarios e universitarios. Comprobamos que a variable nivel de estudos está relacionada coa consecución do éxito na proba, pois dos tres grupos o que presenta unha porcentaxe maior de candidatos aptos é aquel que posúe estudos universitarios (63,9 %), seguido do grupo de candidatos con estudos secundarios (49,5 %) e, por último, o grupo co nivel de estudos máis baixo, os estudos primarios, presenta unha porcentaxe dun 38,8 %. Estes datos non presentan ningunha sorpresa, xa que normalmente ás persoas que teñen máis estudos –e que probablemente xa foron escolarizadas en Galicia e en galego– resúltalles máis doado enfrentarse á aprendizaxe e avaliación dunha lingua ca ás persoas que non tiveron ocasión de formarse no ensino regrado.

3.2.8.6. Cualificación na proba escrita segundo o lugar de residencia

O lugar de residencia ten unha influencia significativa nas cualificacións obtidas ($I = 0,01$; $p < 0,001$) como se pode ver na figura seguinte (Figura 56).

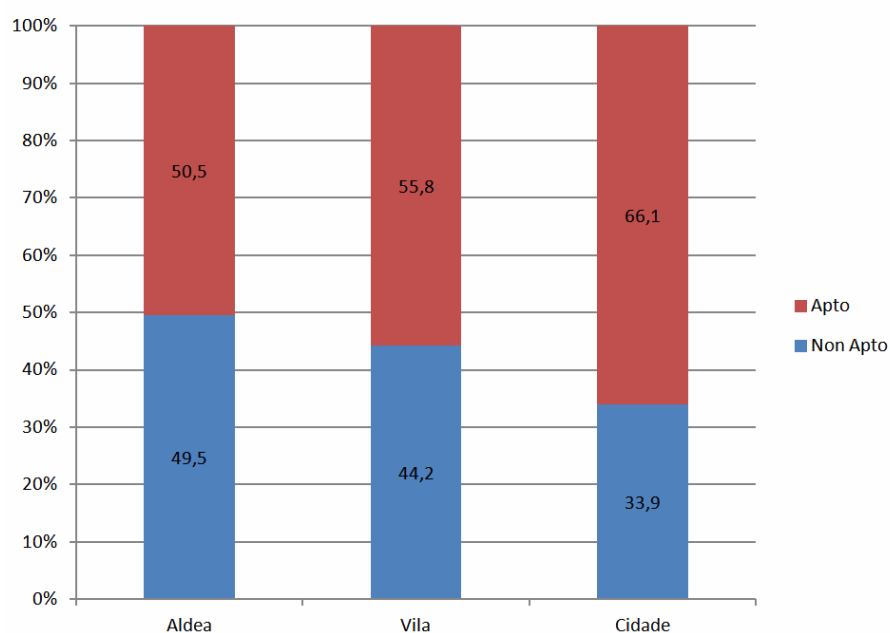


Figura 56. Cualificación na proba escrita segundo o lugar de residencia

Tal e como podemos comprobar na Figura 56, o lugar de residencia dos candidatos –aldea, vila ou cidade– é unha variable clasificatoria significativa. Aínda que a diferenza de porcentaxe non é extremadamente alta, si podemos comprobar que os candidatos que habitan nas cidades presentan unha porcentaxe maior de aptos (66,1 %), seguidos dos habitantes das vilas (55,8 %) e, finalmente, os candidatos residentes nas aldeas presentan unha porcentaxe de aptos menor (50,5 %). Isto pode ser debido a que esta variable clasificatoria do lugar de residencia tamén está asociada a outra variable como é o nivel de estudos que xa analizamos na Figura 55, onde comprobamos que os candidatos que realizaron estudos universitarios presentan unha porcentaxe maior de aptos.

3.2.8.7. Cualificación na proba escrita segundo o sexo

O sexo ten unha influencia significativa nas cualificacións obtidas ($I = 0,001$; $p < 0,001$), como se pode ver na figura seguinte (Figura 57).

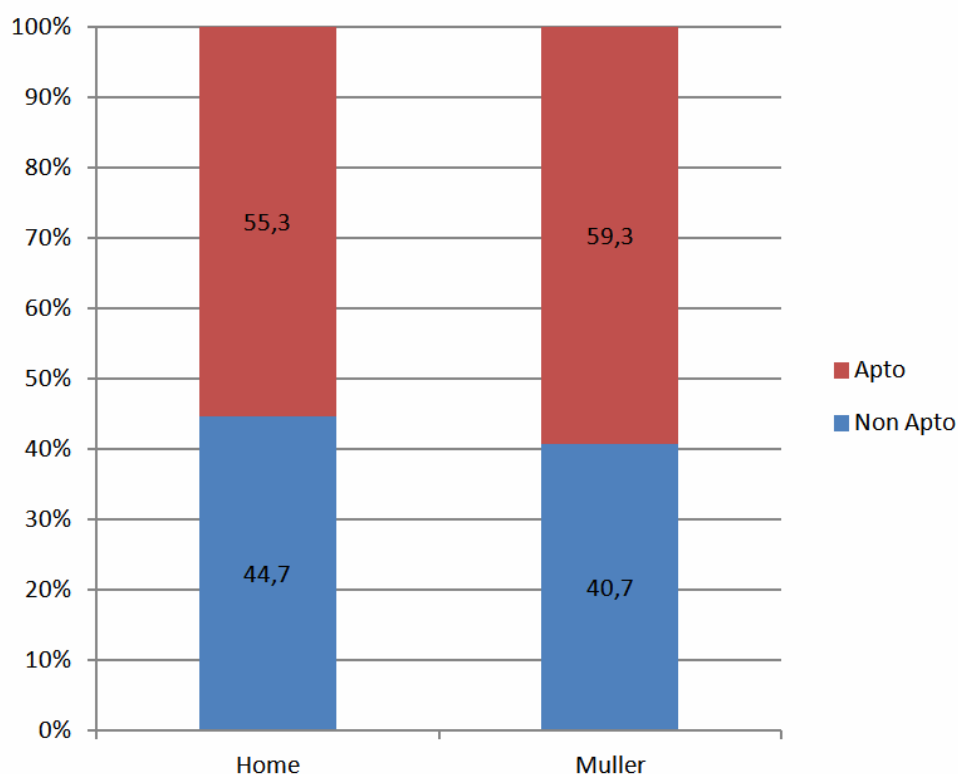


Figura 57. Cualificación na proba escrita segundo o sexo

Divididos os candidatos en dous grupos segundo o sexo –homes e mulleres–, observamos que son as mulleres as que presentan unha porcentaxe maior de aptos (59,3 %) fronte ao 55,3 % dos candidatos homes. Isto pode deberse a que a poboación xeral participante das convocatorias de probas está distribuída case por igual entre homes e mulleres, pero as tres cuartas partes das persoas participantes nas probas Celga son mulleres (70,6 %), tal como vimos no apartado 3.3.

3.2.8.8. Cualificacións na proba escrita segundo a idade

A idade ten unha influencia significativa nas cualificacións da proba escrita obtidas ($I = 0,014$; $p < 0,001$), como se pode ver na figura seguinte (Figura 58). Os resultados indican que, a medida que aumenta a idade, a porcentaxe de persoas aptas na proba escrita diminúe.

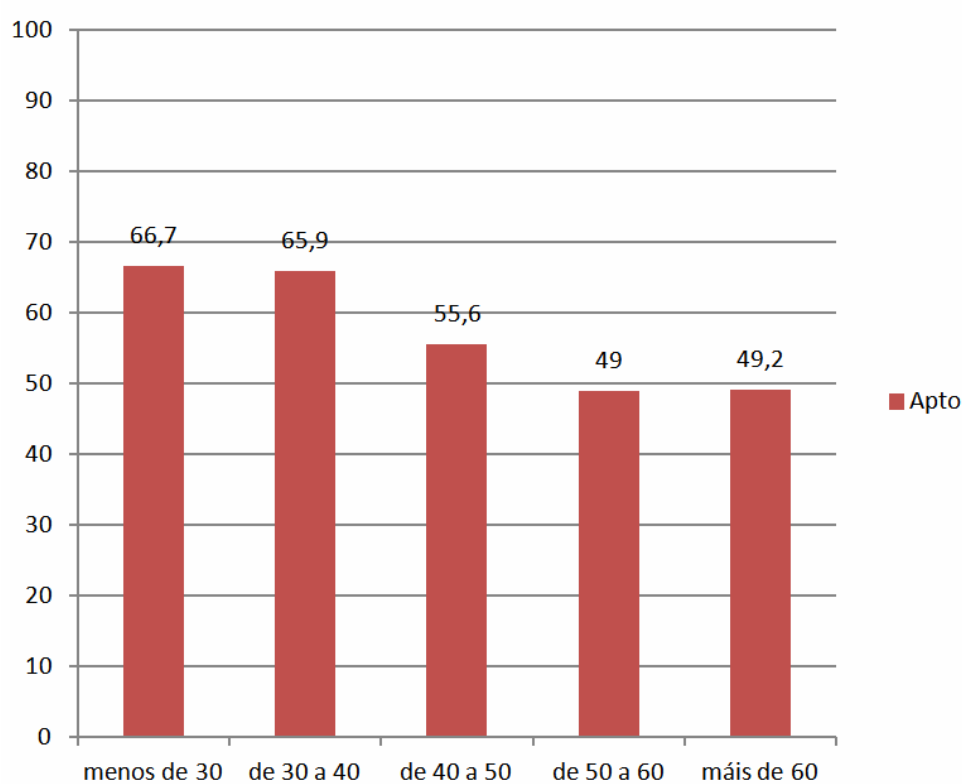


Figura 58. Aptos na proba escrita segundo a idade

A proba ANOVA que se administrou atendendo ás puntuación obtidas en cada un dos criterios de avaliación da proba escrita segundo a idade dos candidatos (Figura 59) indica que existen diferenzas significativas en todos os criterios, confirmando os resultados obtidos na cualificación global da proba escrita. Aplicando a proba *post hoc* de Tukey b, que nos permite analizar as diferenzas entre os diferentes tramos da variable idade, observouse que as

diferenzas se centran nas persoas de menos de 40 anos fronte ao resto, aínda que estas diferenzas se dilúen nos parámetros de léxico e se acentúan nos de adecuación.

Criterios de avaliación	<i>gl</i>	F	<i>p</i>
Adecuación 1	4	27.545	$p < 0,001$
Cohesión 1	4	29.553	$p < 0,001$
Gramática 1	4	22.684	$p < 0,001$
Léxico 1	4	15.977	$p < 0,001$
Adecuación 2	4	21.399	$p < 0,001$
Cohesión 2	4	25.755	$p < 0,001$
Gramática 2	4	14.780	$p < 0,001$
Léxico 2	4	11.193	$p < 0,001$

Figura 59. ANOVA das puntuacións dos candidatos nos parámetros de avaliación escrita segundo a súa idade

Debemos ter en conta que, debido á idiosincrasia galega, se presenta o caso nos candidatos ao Celga maiores de 40 anos que, sen recibiren formación en galego no ensino regrado, adoitan ter maior dificultade para conseguir a aptitude nas probas.

3.2.9. Resultados da proba oral

A continuación discútiense os datos tirados da análise psicométrica sobre a proba oral. Cómpre lembrar que esta é avaliada por dúas persoas, un avaliador e un entrevistador, este último guía a conversa co candidato, pero avaliador e entrevistador teñen que estar de acordo para lle outorgar a aptitude ao

candidato, sempre que acade unha nota de 3 ou 4 puntos en todos os parámetros de avaliación: cohesión e fluidez, adecuación, gramática, léxico, pronuncia e interacción.

O desenvolvemento da proba consiste nunha conversa entre o entrevistador e o candidato, mentres que o avaliador puntúa os parámetros antes citados. Unha vez rematada a participación do candidato, o entrevistador tamén puntúa de xeito individual o candidato sen comentar as notas co avaliador; acabada a puntuación, o entrevistador e o avaliador dan a súa valoración sobre o candidato, isto é, se se considera apto ou non; en caso de non chegar a un acordo, poden debater as razóns pero non as puntuacións que lle outorgaron cada un deles nos parámetro de avaliación; se aínda así non se chegase a unha nota de consenso, pasaríase a unha segunda fase de avaliación na que un supervisor decidiría a nota do candidato.

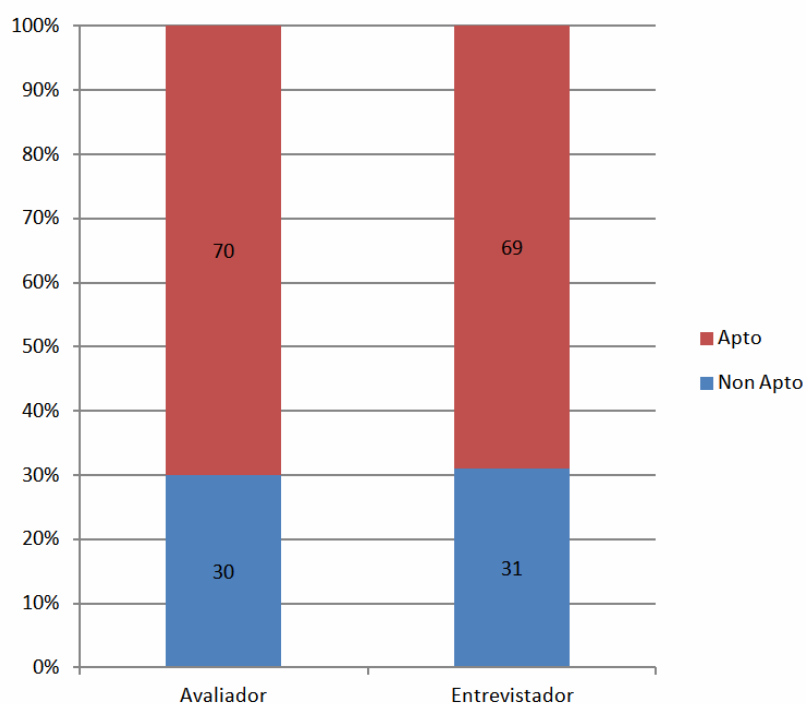


Figura 60. Cualificación na proba oral segundo a valoración do entrevistador e do avaliador

Vemos que tanto o avaliador coma o entrevistador presentan, en xeral, puntuacións moi semellantes, a diferenza é do 1 % entre a porcentaxe de non aptos para o avaliador (30%) e a do mesmo colectivo para o entrevistador (31%). Aínda que neste caso non está diferenciado por niveis do Celga, observamos que, no conxunto do sistema de certificación, a homoxeneidade que presentan os avaliadores á hora de puntuar e considerar a aptitude ou non dos avaliadores na proba oral demostra que os parámetros de avaliación son efectivos e provocan poucos conflitos entre os avaliadores.

3.2.9.1. Relacións entre os criterios de corrección das probas orais

Na táboa seguinte (táboa 12) preséntase o nivel de relación nas puntuacións obtidas polos candidatos ao Celga nos diferentes parámetros de avaliación empregados para a corrección da proba oral. Ao comparar as valoracións dos criterios entre si, obsérvase que a pronuncia e entoación é o criterio que está menos correlacionado. Por outra banda, as valoracións do resto de criterios son bastante semellantes, aínda que se pode observar que a cohesión e fluidez se considera lixeiramente máis útil ca o resto.

	Adecuación	Gramática	Léxico	Cohesión	Pronuncia	Interacción
Adecuación	1 ,000 7967	0,304 ,000 7967	0,381 ,000 7967	0,662 ,000 7967	0,242 ,000 7967	0,677 ,000 4093
Gramática	0,304 ,000 7967	1 ,000 7967	0,698 ,000 7967	0,460 ,000 7967	0,648 ,000 7967	0,251 ,000 4093
Léxico	0,381 ,000 7967	0,698 ,000 7967	1 ,000 7967	0,495 ,000 7967	0,543 ,000 7967	0,312 ,000 4093
Cohesión	0,662 ,000 7967	0,460 ,000 7967	0,495 ,000 7967	1 ,000 7967	0,441 ,000 7967	0,695 ,000 4093
Pronuncia	0,242 ,000 7967	0,648 ,000 7967	,543 ,000 7967	0,441 ,000 7967	1 ,000 7967	0,238 ,000 4093
Interacción	0,677 ,000 4093	0,251 ,000 4093	0,312 ,000 4093	0,695 ,000 4093	0,238 ,000 4093	1 ,000 4093

Táboa 12. Correlacións entre as puntuacións obtidas polos candidatos na proba oral

Polo tanto, da análise dos datos obtidos na anterior táboa podemos tirar as seguintes conclusións:

— Existen correlacións significativas entre todos os criterios de corrección, do mesmo xeito que ocorría na proba escrita; partimos de que todas as variables dependentes están relacionadas porque todas elas presentan puntuacións por riba de 1, algunhas presentan unha puntuación maior e outra menor, (ex: unha puntuación de 0,238 é baixa pero unha de 0,698 é alta) e, como podemos observar na táboa, en xeral existen altas correlacións entre todos os parámetros ou variables dependentes. Apréciase na táboa, se facemos unha lectura en diagonal, que aparecen cadros onde coinciden a mesma variable na fila e na columna (adecuación 1 e adecuación 1); nese caso a puntuación obtida é 1, porque non se relaciona o parámetro consigo mesmo.

— As correlacións máis altas céntranse nas relacións entre interacción e cohesión (0,695), gramática e léxico (0,698), interacción e adecuación (0,677), cohesión e adecuación á tarefa (0,662) e pronuncia e gramática (0,648). Como vemos, todos os parámetros obteñen unha puntuación máis ou menos igual, o que indica que para os avaliadores todos os parámetros de corrección teñen o mesmo peso e validez á hora de empregalos, en contraposición á proba escrita, onde a adecuación á tarefa presentaba unha correlación menor.

— As correlacións máis baixas –pronuncia e adecuación (0,242), interacción e gramática (0,251), interacción e pronuncia (0,238)– obteñen puntuacións moi semellantes e baixas con respecto aos outros parámetros; iso pode sinalar que eses parámetros poden ter menos valor para os avaliadores á hora de considerar a un candidato apto ou non apto.

— Dos datos extraídos, podemos concluír finalmente que se producen agrupacións entre gramática, léxico e pronuncia por unha banda e adecuación á

arefa, coherencia e cohesión e interacción pola outra, en función do valor outorgado aos ditos parámetros por parte do avaliadores para considerar que un candidato sexa apto ou non na proba oral.

3.2.9.2. Cualificación na proba oral dos non aptos

Se observamos polo miúdo como os avaliadores fan uso das ferramentas de avaliación de que dispoñen, é dicir, dos criterios de corrección e avaliación, pódese apreciar que no caso dos candidatos non aptos vai variando a relevancia de cada criterio, que obterá distinta puntuación. Cómpre recordar que na proba oral para que un candidato sexa considerado apto debe obter máis de 2 en todos os parámetros, ao contrario da proba escrita onde poderían ter un 2 só nun parámetro e no resto 3 ou 4. Polo tanto, tal e como aparece na táboa seguinte (Figura 61) e dado que son candidatos non aptos, o parámetro que aparece con máis porcentaxe de puntuación por debaixo do considerado apto, isto é, 1 ou 2, é o de gramática, que presenta un 54,5 %, seguido do léxico, cun 48,3 %, o cal pode apuntar a unha tendencia nos avaliadores a que sexan estes os parámetros que teñen máis peso á hora de decidir se un candidato é apto ou non.

Se observamos os criterios de corrección e avaliación que obtiveron unha porcentaxe de puntuacións altas, 3 e 4, vemos que en interacción aparece un 57 %, seguido de cohesión e fluidez e pronuncia, cun 51,5 % cada un deles.

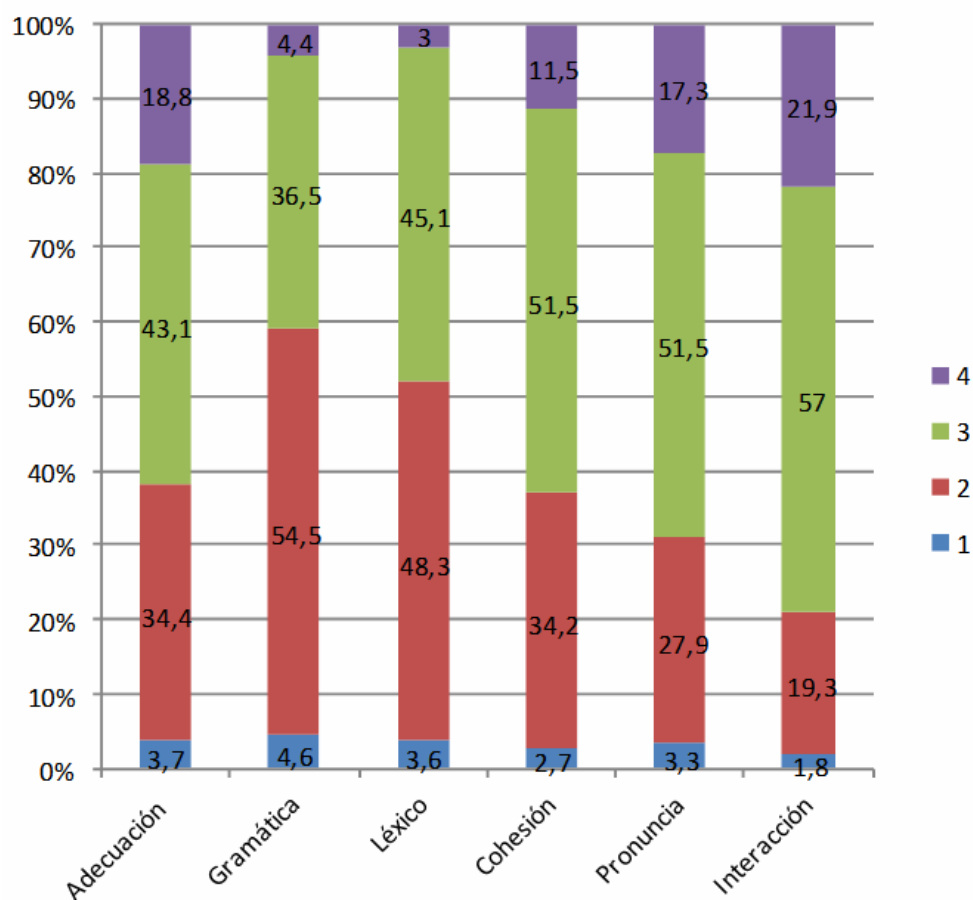


Figura 61. Cualificación na proba oral dos non aptos segundo os avaliadores

3.2.9.3. Cualificación na proba oral dos aptos

Se observamos a análise dos candidatos aptos na proba oral (Figura 62), é dicir, aqueles que obtiveron en todos os parámetros de corrección e avaliación unha puntuación de 3 ou 4, é salientable que o léxico é o parámetro que obtén unha porcentaxe maior (66,8 %), seguido polo de gramática (65,1 %), ambos puntuados con 3.

Os parámetros que obtiveron a puntuación máis elevada, un 4, son os de interacción (58 %), adecuación (54,5 %) e cohesión e fluidez (54,3 %).

Polo tanto, podemos deducir que existe un agrupamento de criterios, do mesmo xeito que ocurría nas puntuacións dos criterios das probas orais presentado no apartado 3.2.9.2., no que, por unha banda, están interacción, adecuación e cohesión e fluidez (con puntuacións sobre 4) e, pola outra, o resto dos criterios, gramática e léxico (puntuados sobre 3).

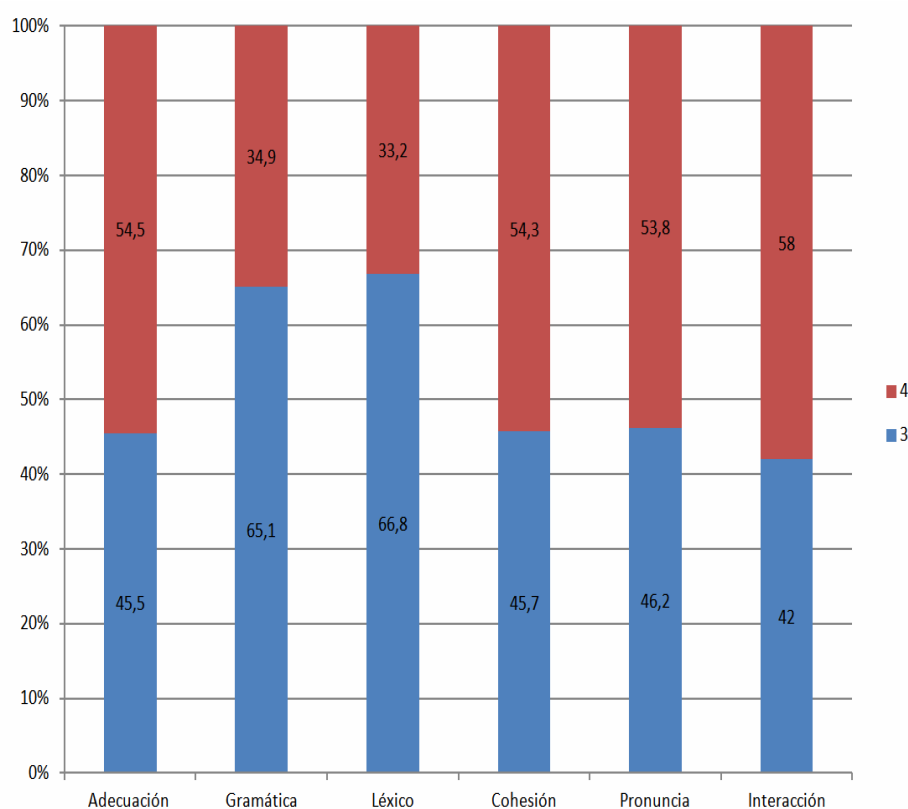


Figura 62. Cualificación na proba oral dos aptos segundo os avaliadores

3.2.9.4. Cualificación na proba oral segundo o nivel do Celga

O nivel do Celga ten unha influencia significativa nas cualificacións obtidas ($I = 0,022$; $p < 0,001$) como se pode ver na figura seguinte (63).

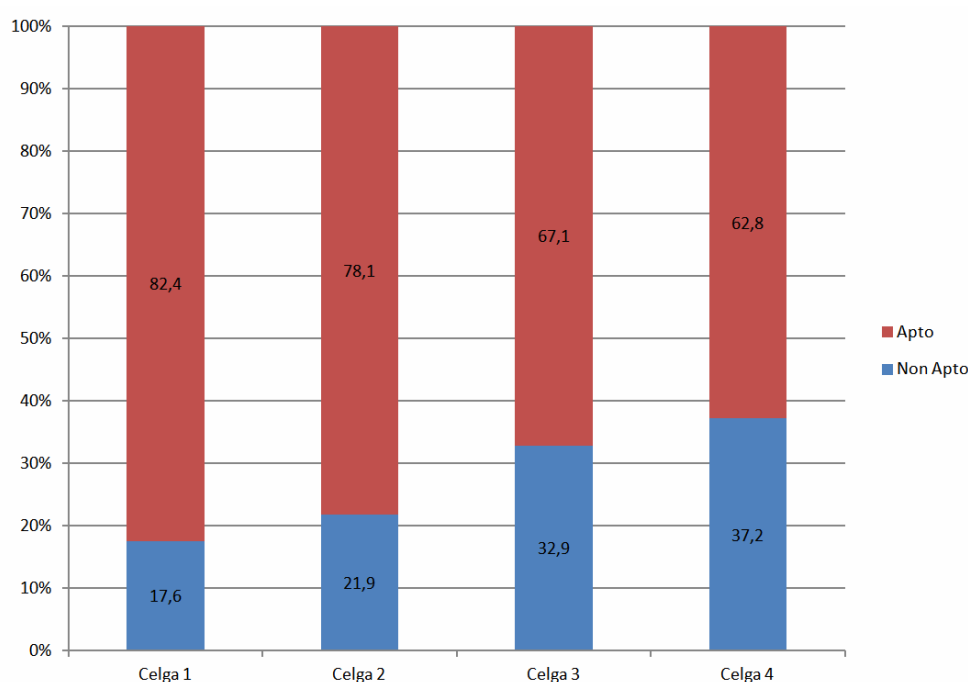


Figura 63. Cualificación na proba oral segundo o nivel do Celga

Así, nos niveis máis baixos aparece unha cantidade maior de candidatos aprobados (82,4 % no Celga 1 e 72,1 % no Celga). Isto é o habitual neste tipo de probas, nas que é directamente proporcional a dificultade no nivel de lingua á diminución do número de candidatos aptos (67,1 % no Celga 3 e 62,8 % no Celga 4). Polo tanto, os niveis de Celga 1 e Celga 2 son os que acadan o maior número de candidatos aptos na proba oral.

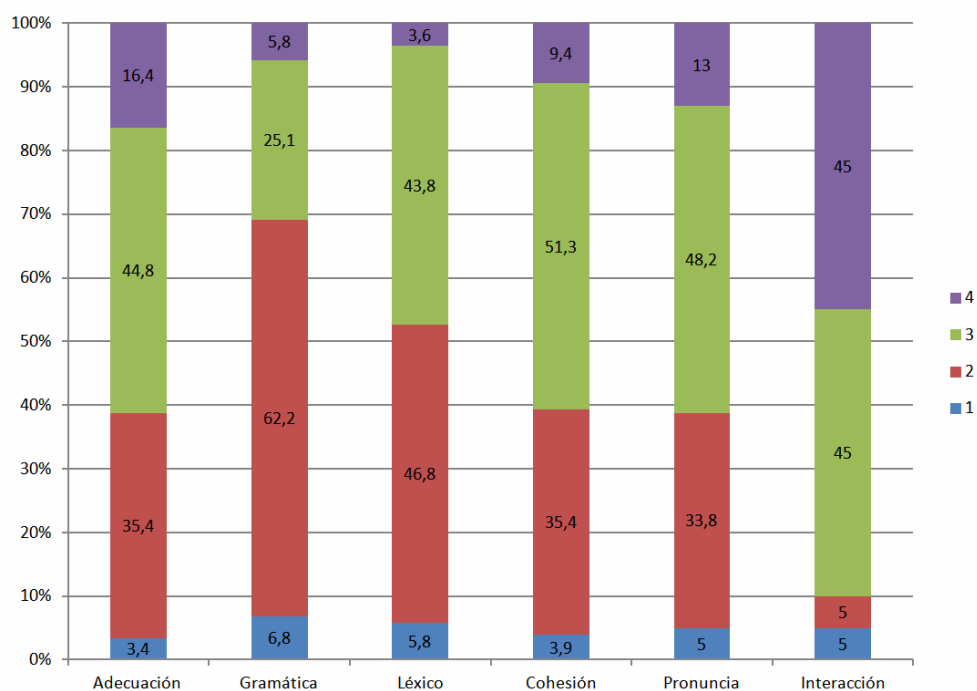


Figura 64. Cualificación na proba oral dos candidatos non aptos nos niveis do Celga 1 e do Celga 2

Nos niveis máis baixos de lingua, Celga 1 e Celga 2, onde xa vimos que a porcentaxe de candidatos que supera a proba oral é maior ca nos niveis máis altos, observamos (Figura 64) que os aqueles candidatos que non superaron a proba presentan unha porcentaxe de puntuación de 2 e 3 maioritariamente. Presentan unha porcentaxe de 2 puntos en gramática (62,2 %) e en léxico (46,8 %), mentres que obtiveron unha porcentaxe de 3 puntos os parámetros de cohesión e fluidez (54,3 %) e adecuación (44,8 %).

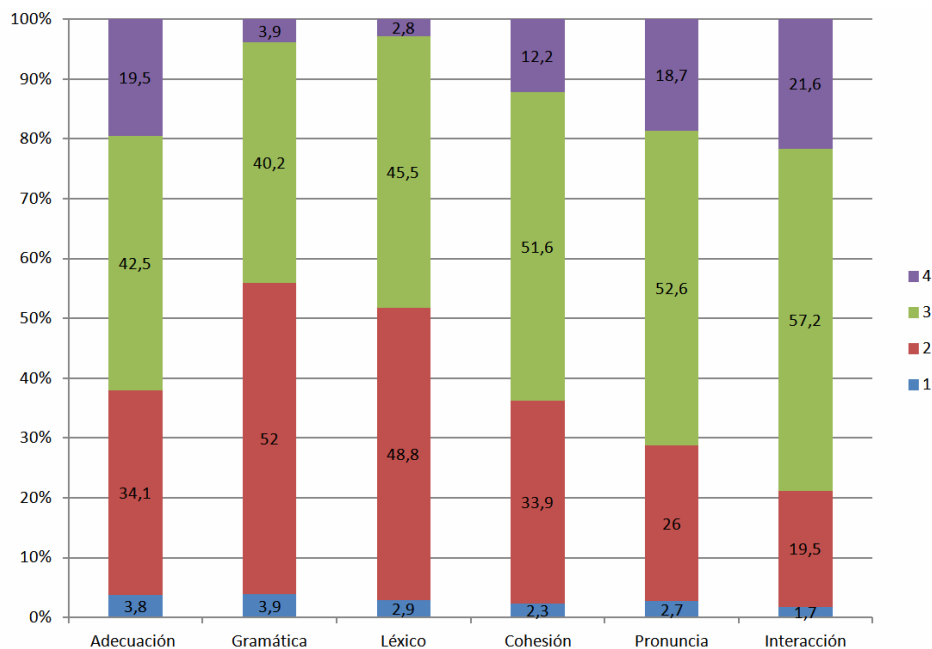


Figura 65. Cualificación na proba oral dos non aptos nos niveis do Celga 3 e do Celga 4

No caso da proba oral os candidatos non aptos presentados aos niveis máis altos, Celga 3 e Celga 4 (Figura 65), obtiveron unha puntuación maioritaria entre o 2 e o 3; así, candidatos que obtiveron como puntuación un 2, resultou o parámetro de gramática o maioritario cun 52% do total, mentres que cun 2 como puntuación o parámetro que obtivo a menor porcentaxe de puntos foi o de interacción cun 19,5 %. Pola contra, para os candidatos que obtiveron un 3 como puntuación o parámetro con maior porcentaxe foi o de interacción cun 57,2% fronte a gramática cun 40,2 %.

3.2.9.5. Cualificación na proba oral segundo o nivel de estudos

O nivel de estudos ten unha influencia significativa nas cualificacións obtidas ($I = 0,044$; $p < 0,001$) como se pode ver na figura seguinte (Figura 66).

Hai que sinalar que non se inclúe as persoas que carecen de estudos porque non son suficientes para extraer ningunha conclusión (véxase a Figura 66, onde están clasificadas as persoas sen estudos).

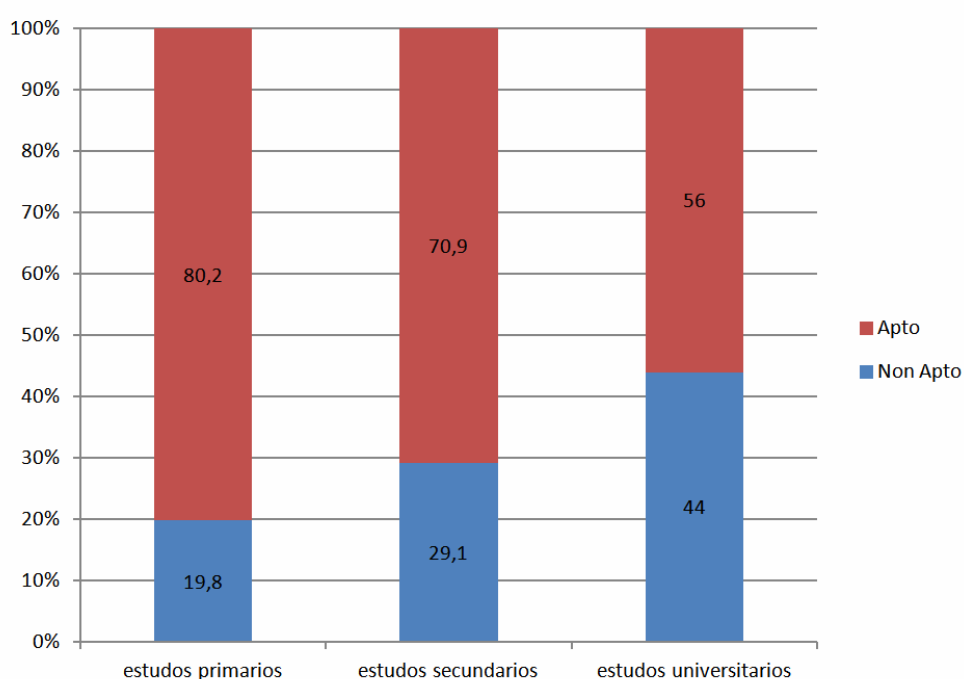


Figura 66. Cualificación na proba oral segundo o nivel de estudos

Como podemos observar, a relación entre os candidatos aptos no Celga é directamente proporcional ao nivel de estudos que posúan. Un 56 % dos candidatos con estudos universitarios acada o apto, un 70,09 % dos candidatos con estudos medios consegue a aptitude nas probas e, finalmente, un 80,2 % dos candidatos con estudos primarios tamén son aptos; esta correlación na porcentaxe de aptos naqueles candidatos que teñen estudos primarios ten que ver co nivel de Celga ao que se presentan, xa que no Celga 1 é onde máis

candidatos con niveis primarios se presentan fronte aos niveis máis altos, sobre todo o Celga 4, no que se presentan candidatos con estudos superiores.

3.2.9.6. Cualificación na proba oral segundo o lugar de residencia

O lugar de residencia apenas inflúe na cualificación na proba oral ($I = 0,002$; $p = 0,01$), como se pode ver na figura seguinte.

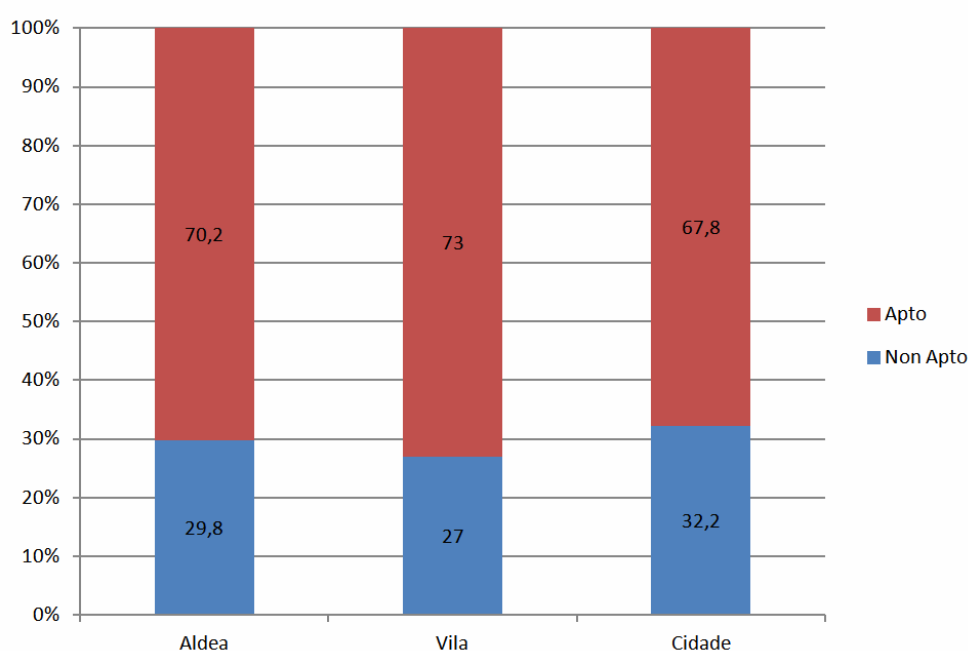


Figura 67. Cualificación na proba oral segundo o lugar de residencia

O lugar de residencia dos candidatos –aldea, vila ou cidade– non é unha variable clasificatoria significativa. Aínda que a diferenza de porcentaxe non é extremadamente superior, si podemos comprobar que os candidatos que habitan nas cidades presentan unha porcentaxe menor de aptos (67,8 %), seguidos dos habitantes das aldeas, que presentan unha porcentaxe de aptos maior, un 70,2 %. Finalmente, os candidatos residentes nas vilas presentan un 73,8 %. Isto pode ser debido a que a variable clasificatoria do lugar de residencia está asociada á variable nivel de estudos que xa analizamos na

Figura 67, onde comprobamos que os candidatos que teñen estudos universitarios presentan unha porcentaxe maior de candidatos aptos.

3.2.9.7. Cualificación na proba oral segundo o sexo

O sexo ten unha influencia significativa nas cualificacións obtidas ($I = 0,002$; $p < 0,001$), como se pode ver na figura seguinte (Figura 68).

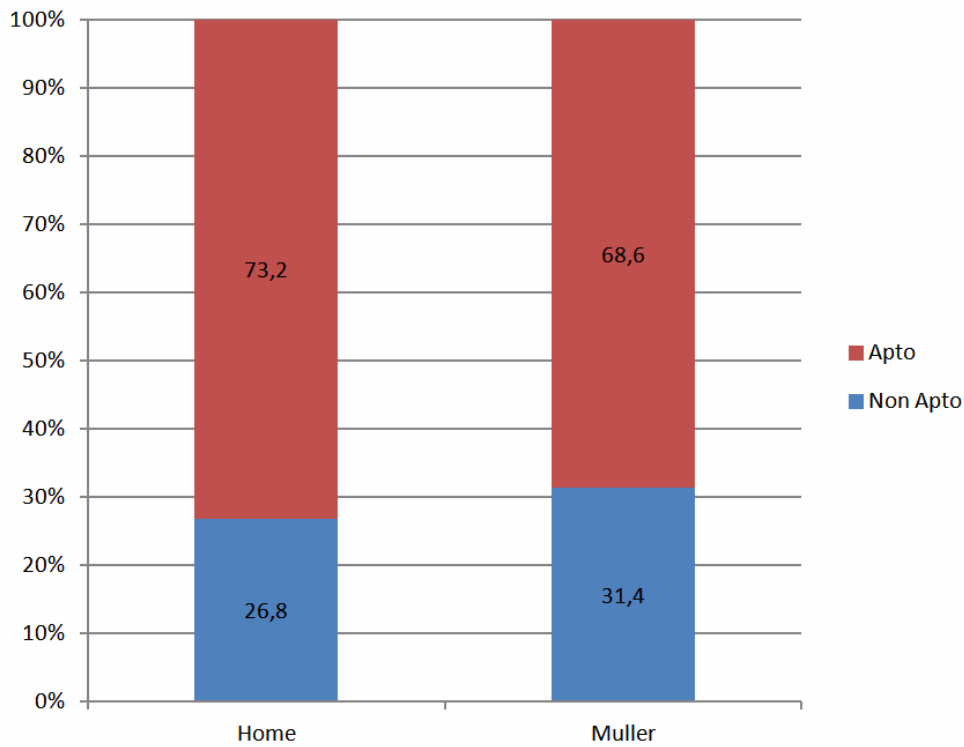


Figura 68. Cualificación na proba oral segundo o sexo

Divididos os candidatos en dous grupos segundo o sexo, homes e mulleres, vemos que son os homes os que presentan unha porcentaxe maior de aptos (73,2 % fronte ao 68,6 % das mulleres). Ao contrario do que acontecía na proba escrita, tendo en conta, ademais, que a poboación xeral participante nas convocatorias de probas está distribuída case por igual entre homes e mulleres,

as tres cuartas partes das persoas participantes nas probas Celga son mulleres (70,6 %), tal e como analizamos no apartado 3.3.

3.2.9.8. Cualificación na proba oral segundo a idade

Como se pode ver na Figura 69, a idade inflúe en certa medida sobre a cualificación da proba oral, aínda que este efecto é menor e de signo contrario ao caso da cualificación escrita ($I = 0,003$; $p = 0,007$).

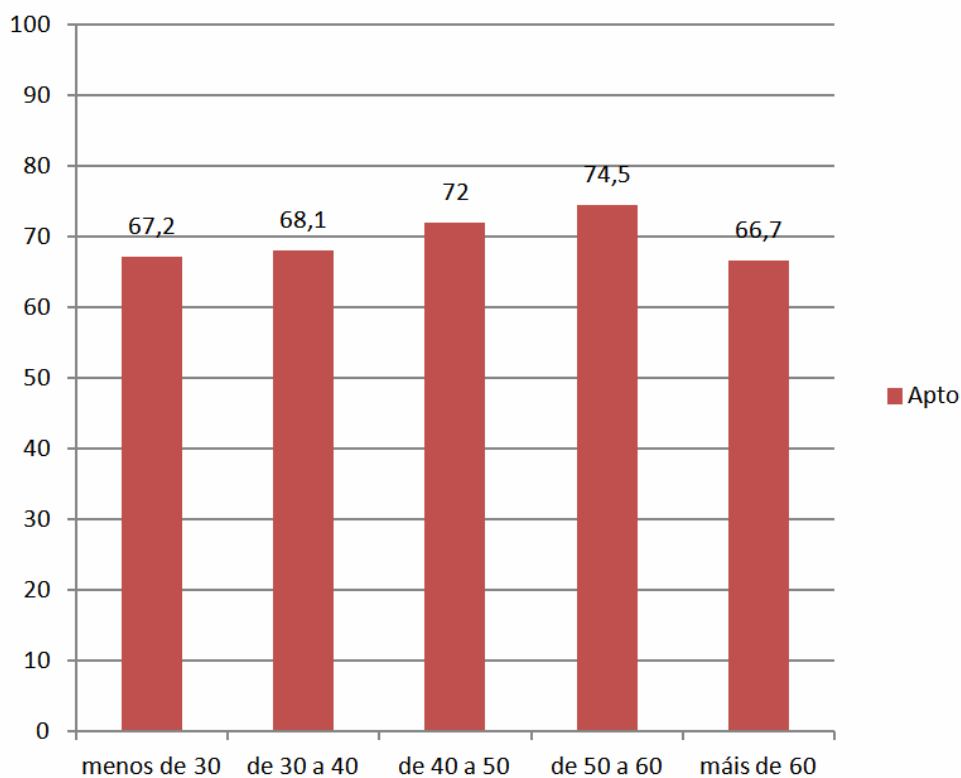


Figura 69. Aptos na proba oral segundo a idade

Analizando as puntuacións obtidas en cada un dos criterios de avaliación (Figura 70), os resultados da ANOVA revelan que a idade non ten capacidade de discriminación co parámetro interacción; pola contra, o criterio pronuncia ten unha forte variación segundo a idade do candidato. Observamos

que entre os candidatos de máis idade –maiores de 60 anos (66,7 %)- e os de menos de trinta anos (67,25%) existe só unha diferenza de pouco máis de 1 punto; e as ratios de idade que obteñen unha maior porcentaxe de candidatos aptos son as das persoas entre 50 e 60 anos (74,5 %) e os de 40 a 50 anos (72 %). Hai que lembrar que deste grupo poboacional non todos recibiron unha alfabetización en galego durante a súa aprendizaxe regrada; non obstante, son maioritariamente galegofalantes e teñen menos dificultade na proba oral ca na escrita.

Criterios de avaliación	<i>GI</i>	F	<i>p</i>
Adecuación	4	7.126	p < 0,001
Gramática	4	13.463	p < 0,001
Léxico	4	3.892	p = 0,004
Cohesión	4	3.124	p = 0,014
Pronuncia	4	41.455	p < 0,001
Interacción	4	1.867	p = 0,114

Figura 70. ANOVA das puntuacións dos candidatos nos parámetros de avaliación escrita segundo a súa idade

Aplicando a proba *post hoc* de Tukey b e como se pode ver na Figura 71, na que se representan as puntuacións medias nos diferentes criterios segundo a idade dos candidatos, as persoas de máis de 40 anos acadan unha maior puntuación no criterio de pronuncia ca as de menor idade. Este mesmo fenómeno, aínda que con menos forza, prodúcese no parámetro de gramática. Pola contra, no resto de criterios existe un grupo que se diferencia claramente do resto, os maiores de 60 anos, persoas que obteñen puntuacións significativamente máis baixas ca o resto en adecuación, cohesión e léxico; presentan deficiencias na riqueza de léxico, dificultade á hora de cohesionar

textos e de adecuarse á tarefa solicitada, pero non na pronuncia nin na gramática.

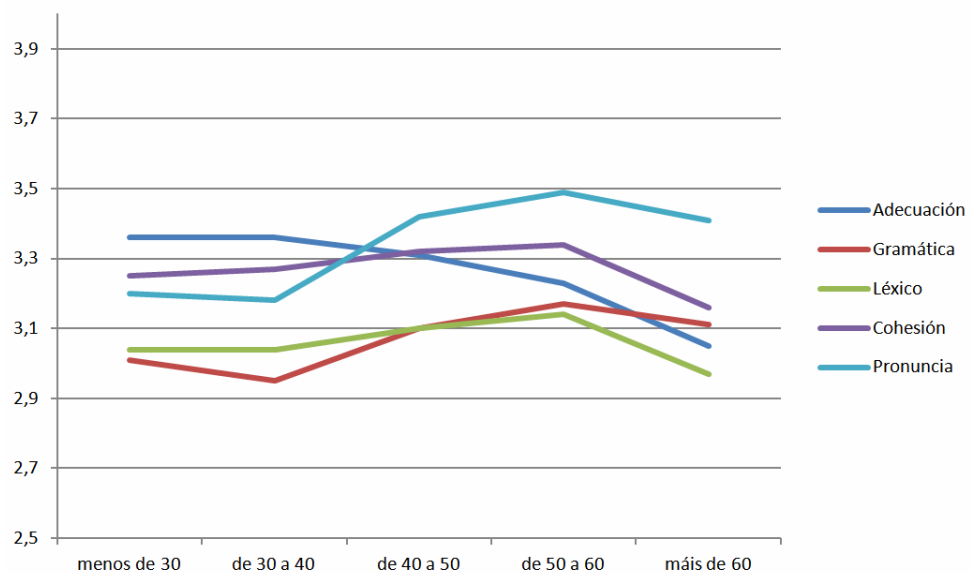


Figura 71. Medias das puntuacións dos candidatos nos parámetros de avaliación oral segundo a súa idade

CAPÍTULO 4

Conclusiones

4.1. Conclusións Estudo I

A continuación presentamos as conclusións ás que chegamos logo de examinar os datos procedentes do presente estudo. Para facilitar o seu seguimento, agrupámolas segundo os apartados correspondentes ao capítulo 3 de análise de datos que, pola súa vez, se corresponden cos obxectivos e hipóteses presentadas no apartado 1.1 deste capítulo.

4.1.1. Utilidade dos criterios de corrección e avaliación das probas de expresión escrita

Neste apartado o que se presenta é unha compilación das opinións que os avaliadores teñen sobre a utilidade dos parámetros de corrección e avaliación como ferramentas necesarias para o seu labor nas probas escritas; tamén presentaremos os datos tirados sobre a utilidade dos parámetros atendendo ao nivel de Celga e a experiencia dos avaliadores, así podemos observar que:

— Os correctores consideran que estes criterios **son altamente útiles** para a función que desempeñan, xa que ningún deles ten unha media de utilidade inferior a 7 e o intervalo de puntuacións oscila entre 3 e 10 en todos os criterios, excepto o de gramática que vai desde 2 a 10.

— Atendendo ás puntuacións medias en utilidade adxudicadas polos avaliadores, pódense dividir os criterios en dous grupos:

— Un deles considerado **máis útil, composto por adecuación á tarefa e coherencia e cohesión.**

— Outro que, aínda que se considera que ten utilidade, ten algunha **menos ca o anterior, composto por gramática e léxico.**

4.1.1.1. Segundo o nivel do Celga

— A medida que o nivel de Celga é superior, tamén aumenta a utilidade adxudicada a todos os criterios de corrección e avaliación. Non obstante, existen diferenzas relevantes sobre cada un dos criterios:

— A **adecuación á tarefa** considérase máis útil para o Celga 4 e o Celga 3 ca para o Celga 1 e o Celga 2.

— Os criterios de coherencia e cohesión, gramática e léxico considéranse máis útiles para o nivel Celga 4 ca para o Celga 1 e o Celga 2, mentres que estes criterios son considerados máis útiles no Celga 2 ca no Celga 1.

4.1.1.2. Segundo a experiencia dos avaliadores

— A experiencia dos avaliadores no proceso de avaliación escrita do Celga non inflúe significativamente sobre a súa percepción da utilidade de ningún dos criterios de corrección e avaliación. Porén e tal e como avanzamos:

— Podemos afirmar que nas probas escritas a **adecuación á tarefa** é o criterio considerado máis útil.

— As medias de valoración dos criterios de corrección e avaliación son altas; esta está (establecida de 1 a 10) en 6, o cal indica que os avaliadores consideran útiles os criterios como ferramentas para o seu traballo.

4.1.2. Utilidade dos criterios de corrección e avaliación das probas de expresión oral segundo os avaliadores

— Obsérvase que os correctores consideran que estes criterios **son útiles** para a función que desempeñan, xa que non hai ningún criterio que teña

unha media de utilidade inferior a sete –o intervalo de puntuacións oscila entre 3 e 10 en todos os criterios, agás o de gramática, que vai desde 2 a 10–.

— Os criterios que obteñen unha maior puntuación son **interacción e coherencia e cohesión** (7,67 e 7,6, respectivamente) fronte á **pronuncia**, coa puntuación máis baixa (7,17). O resto acada puntuacións moi semellantes.

4.1.2.1. Segundo o nivel do Celga

Prodúcese o mesmo fenómeno ca na expresión escrita, a medida que aumenta o nivel do Celga aumenta a importancia que lles adxudican os avaliadores aos criterios de corrección e avaliación.

— Os avaliadores consideran que os criterios **máis útiles son os de adecuación á tarefa, interacción e coherencia e cohesión** (8,27; 8,26; 8,22, respectivamente), sobre todo no Celga 4.

— A **adecuación á tarefa** considérase **máis útil** para o Celga 4 ca para o Celga 1 e 2.

— Os avaliadores perciben como **menos útil a pronuncia** para o Celga 1 e Celga 2, diferenzas que se irán atenuando a medida que aumenta o nivel de Celga.

— Os criterios de coherencia e cohesión, gramática e léxico considéranse máis útiles para o Celga 4 ca para os primeiros (Celga 1 e Celga 2) e todos eses criterios son máis útiles para o Celga 3 ca para o Celga 1.

4.1.2.2. Segundo a experiencia profesional dos avaliadores

4.1.2.2.1. Na proba escrita

— Aplicouse a proba non paramétrica chamada H Krustal Wallis, da que se deduce que a experiencia profesional non inflúe significativamente sobre a percepción de utilidade de ningún dos criterios de corrección e

avaliación: adecuación ($K = 7,551$; $p = 0,056$), coherencia e cohesión ($K = 5,543$; $p = 0,136$), gramática ($K = 0,673$; $p = 0,880$) e léxico ($K = 0,106$; $p = 0,991$).¹⁶

4.1.2.2.2. Na proba oral

— A medida que aumentaba a experiencia dos avaliadores, estes percibían que a coherencia e cohesión tiña máis utilidade. Existen diferenzas entre os que teñen menos de tres anos de experiencia fronte aos que teñen entre catro e sete anos de experiencia ($p' = 0,004$) e os que teñen mais de sete anos ($p' < 0,001$) e entre os que teñen entre catro e sete fronte aos que teñen máis de sete ($p' = 0,006$).

— Nos criterios de adecuación á tarefa, gramática e interacción había diferenzas significativas entre os correctores con menos de tres anos de experiencia e máis de sete anos ($p' < 0,001$; $p' = 0,008$; $p' = 0,002$) e entre os que tiñan entre catro e sete anos de experiencia e máis de sete anos de experiencia ($p' = 0,002$; $p' = 0,002$; $p' = 0,006$).

— Polo tanto, podemos deducir que a medida que adquiren máis experiencia consideran que estes criterios teñen máis importancia.

4.1.3. Validez dos criterios de corrección e avaliación das probas segundo os avaliadores

— Case a totalidade dos avaliadores (94,4 %) considera importantes todos os criterios de corrección e avaliación, xa que os enquisados opinan que non se debería eliminar ningún (pregunta 13 do cuestionario, *vid.* anexo 1).

¹⁶ Non se aplicou a proba *a posteriori* U de Mann Whitney para comprobar as diferenzas entre os diferentes niveis da variable experiencia.

4.1.3.1. Validez dos criterios de corrección e avaliación na proba escrita

— Logo de lles preguntar aos avaliadores sobre os criterios de corrección e avaliación considerados menos válidos á hora de determinar a competencia comunicativa na proba escrita (pregunta 10 do cuestionario, *vid.* anexo 1), máis da metade (59,6 %) opina que todos os criterios de corrección e avaliación son igual de válidos.

— Os avaliadores sen experiencia, curiosamente, coinciden en que todos os criterios son válidos cos avaliadores máis experimentados (61,5 % os avaliadores sen experiencia e 63,2 os avaliadores con máis de catro convocatorias).

— A totalidade dos avaliadores sen experiencia cre que non se debería eliminar ningún criterio na proba escrita, aínda que un 9,1 % con máis experiencia opina que si se debería eliminar algún.

4.1.3.2. Validez dos criterios de corrección e avaliación na proba oral

— Cando se lles pregunta aos avaliadores sobre os criterios de corrección menos válidos para avaliar a competencia oral (na pregunta 11 do cuestionario, *vid.* anexo 1), a metade (51,4 %) escollía a alternativa “todos son igual de válidos”.

— O criterio elixido por unha porcentaxe maior de avaliadores como menos válido foi pronuncia e entoación (17,6 %), mentres os criterios que se elixiron un menor número de veces foron cohesión e fluidez e gramática (6,8 % nos dous casos).

4.1.4. Valoración da equidade do Celga

— Ao lles preguntar aos avaliadores se consideran que as probas Celga teñen un sistema xusto de avaliación, a maior parte deles considera que é

equitativa (74,5 %), aínda que existe unha porcentaxe moi significativa de avaliadores que non o considera xusto (25,5 %).

— A medida que aumenta a experiencia como avaliadores na proba escrita increméntase a percepción de equidade do sistema.

— Aqueles avaliadores con entre catro e sete convocatorias ou máis de sete convocatorias de experiencia na proba oral obteñen resultados moi similares cando consideran que o Celga é un sistema xusto de avaliación (77,6 % e 77,3 %). Pola contra, os avaliadores cunha menor experiencia na proba oral (isto é, ata 3 convocatorias) opinan nun 33,3% que o sistema é menos xusto.

— Polo tanto, podemos inferir que a consideración da equidade no sistema de avaliación é proporcional á experiencia dos avaliadores nas probas orais escritas.

4.1.5. Valoración da competencia comunicativa na proba escrita dos candidatos ao Celga

— A primeira dimensión (dimensión I) é a que realmente soporta o resultado do modelo de avaliación na proba escrita, xa que explica un 98,5 % da varianza, mentres a segunda soamente explica un 1,5 %. Este nivel tan alto de varianza explicada e tan baixo de stress indica a alta relación entre todos os criterios de corrección e avaliación en todos os niveis do Celga. Esta dimensión estaría centrada especialmente no criterio de adecuación á tarefa, seguido, en menor medida, polos criterios de léxico e gramática. O cal provoca na avaliación da proba escrita a existencia dun supracriterio empregado polos avaliadores que ten máis peso que o resto de criterios de corrección e avaliación á hora de decidir que un candidato sexa apto ou non apto.

4.1.6. Valoración da competencia comunicativa na proba oral dos candidatos ao Celga

Neste apartado preséntase a consideración que os avaliadores teñen sobre a competencia comunicativa que presentan os candidatos nas probas orais, atendendo sempre aos parámetros de avaliación empregados para cualificar dita competencia e ao nivel do Celga, porén obsérvase que:

— Os criterios que máis determinan a avaliación da competencia comunicativa na proba oral na dimensión I serían adecuación á tarefa e interacción ($w = -1,76$), fronte a pronuncia e entoación ($w = -1,69$). Mentres na segunda dimensión os criterios que máis o determinan son a interacción ($w = -1,46$) e a adecuación á tarefa e cohesión ($w = -1,04$).

— A valoración da competencia comunicativa dos candidatos na proba oral nos niveis iniciais (Celga 1 e Celga 2) fundaméntase especialmente no criterio de adecuación á tarefa e interacción fronte ao de pronuncia e entoación.

— Por outra banda, nos niveis máis altos do Celga (Celga 3 e Celga 4), a valoración fundaméntase nos criterios de adecuación á tarefa e interacción fronte aos de gramática, léxico e pronuncia e entoación.

— Xa que logo, podemos asegurar que, na valoración da competencia comunicativa dos candidatos na proba escrita, os avaliadores consideran que existe un supracriterio que reúne a todos os demais, o de adecuación á tarefa; mentres que na valoración da proba oral non se presenta un criterio que teña máis peso ca os outros, senón que existe un equilibrio na importancia de todos os criterios. Isto provoca que a proba escrita presente unha diferenza na avaliación con respecto á proba oral, xa que a adecuación á tarefa da proba escrita vén limitada polo número de palabras que, segundo o nivel, os candidatos deben producir, requisito que na proba oral non se considera.

4.2. Conclusións Estudo II

A continuación presentamos as conclusión do estudo II, no que se trará cuestións coma a fiabilidade entre os avaliadores e supervisores, a validez de construto das probas Celga e por último, o perfil sociodemográfico dos candidatos ás probas.

4.2.1. Fiabilidade interxuíces

Con respecto á relación entre os primeiros avaliadores e os segundos, os resultados indican que a fiabilidade non é tan alta como cabía esperar. Este feito pode ser causado polo protocolo de actuación á hora de corrixir as probas, xa que, despois de avaliar a primeira vez a todos os candidatos, se fai unha segunda volta de correccións da proba escrita por parte dos supervisores na que se avalían todos os exames non aptos e, aleatoriamente, unha cantidade de exames aptos. É dicir, céntrase o control nos non aptos e non nunha mostra aleatoria dos exames, aptos e non aptos, que sería o ideal; ao facer isto dáse por suposto que existe un problema de fiabilidade, o cal provoca que mentres nos aptos hai un alto índice de relación nas puntuación dos avaliadores, nos non aptos este índice baixa perigosamente.

4.2.2. Consistencia interna

En xeral, hai unha alta consistencia interna que indica que todos os criterios de avaliación están medindo o mesmo construto. O cal significa que o construto do Celga está moi interiorizado e fortemente asentado nos avaliadores.

4.2.3. Validez de construto

Existen altas correlacións entre todos os parámetros de corrección e avaliación: partimos de que todas as variables dependentes están relacionadas porque todas elas presentan puntuacións por riba de 1. As maiores correlacións obtivéronse entre as probas A1 e A2, excepto no criterio de adecuación á tarefa, e isto pode deberse a que o dito criterio é o máis significativo á hora de decidir se unha proba é apta ou non, xa que para cumprir con este parámetro os candidatos deben crear un texto cunha modalidade específica e, ademais, cumprir cun número determinado de palabras, en caso contrario –ben sexa porque se supera ou non se chega nun 20% ao número de palabras estipuladas– a proba considérase non apta.

Por outra banda, as correlacións máis baixas son as existentes entre os criterios de adecuación á tarefa e gramática tanto na proba A1 como na proba A2. Seguramente se deba isto a que son os criterios que máis peso teñen para os avaliadores á hora de delimitar se un candidato é apto ou non apto.

No caso da proba escrita, existe un dobre agrupamento dos criterios de corrección e avaliación en función do uso que os avaliadores fan de ditos criterios. O primeiro factor, denominado de coñecemento do código, está composto polos criterios de gramática, léxico e coherencia e cohesión das probas A1 e A2 e a porcentaxe de varianza que explica é dun 46,25 %. O segundo factor pódese denominar de estrutura textual e está composto polos criterios de adecuación á tarefa na proba A1 e A2 e polos criterios de coherencia e cohesión na proba A1 e A2, compartidos co factor primeiro. O cal significa que o primeiro factor ‘coñecemento do código’ no caso da proba escrita, ao contrario das opinións expresadas polos avaliadores é a que explica unha porcentaxe de varianza maior. Esta circunstancia significa que aspectos formais de código teñen unha maior importancia á hora de decidir se un

candidato é APTO ou NON. No caso da proba oral, a importancia dos factores iguálase. A varianza do factor ‘coñecemento do código’ e do factor ‘estrutura textual’ é moi semellante.

Así, na proba oral, o factor de coñecemento do código, que agrupa os criterios de gramática, pronuncia e léxico, explica un 39,03 % da varianza total. O segundo factor –de estrutura textual– está composto polos criterios de interacción, adecuación e cohesión e explica un 38,84 %, case o mesmo ca o primeiro factor. Estes datos sinalan que ao contrario que na proba escrita, na proba oral non existe unha diferenza tan marcada entre o factor que atinxe á estrutura textual e o factor de coñecemento do código (38,84% e 39,03% respectivamente).

4.2.4. Perfil sociodemográfico dos candidatos ao Celga en relación coa obtención da aptitude

4.2.4.1. Cualificación na proba escrita segundo o nivel do Celga

Podemos comprobar que o sistema de certificación de lingua galega presenta unha porcentaxe de candidatos aptos na proba escrita proporcional ao nivel de lingua. Nos niveis máis baixos, Celga 1 e Celga 2, o índice de aprobados é superior (75 % e 66,1 %, respectivamente) ao dos niveis superiores (55,1 % no Celga 3 e 49,4 % no Celga 4).

Aínda así, cómpre salientar que no nivel Celga 4 o perfil sociolóxico dos candidatos é maioritariamente de galegos [ao analizar o lugar de nacemento dos participantes obsérvase que algo máis da metade dos candidatos naceron en Galicia (55,3%)], o 29,6 % pertence ao resto de rexións do Estado

español e hai unha porcentaxe de persoas que naceron no continente americano (8,3 %) e no resto de Europa (6,5%). Isto condiciona que a proximidade das linguas dos candidatos, español e galego, permita unha porcentaxe tan elevada nos niveis de Celga 3 e Celga 4, malia seren os máis altos e complexos á hora de obter a aptitude.

4.2.4.2. Cualificación na proba oral segundo o nivel do Celga

Os niveis de Celga 1 e Celga 2 son os que acadan o maior número de candidatos aptos na proba oral (82,4 % no Celga 1 e 72,1 % no Celga 2), mentres no Celga 3 e no Celga 4 esta porcentaxe descende (67,1 % e 62,8 %, respectivamente). O habitual neste tipo de probas é que sexa directamente proporcional a dificultade no nivel de lingua coa diminución do número de candidatos aptos, xa que a medida que o nivel de lingua aumenta a dificultade para acadar a competencia comunicativa necesaria é maior.

4.2.4.3. Cualificación na proba oral dos non aptos

No caso dos exames avaliados como non aptos, os parámetros que aparecen con máis porcentaxe de puntuación por debaixo do considerado apto, isto é, 1 ou 2, son os de gramática (54,5 %) e léxico (48,3 %), o cal pode apuntar a unha tendencia nos avaliadores a que sexan estes os parámetros que teñen máis peso á hora de decidir se un candidato é apto ou non.

Se observamos os criterios de corrección e avaliación que obtiveron unha porcentaxe de puntuación alta (3 e 4), aínda que os exames resultasen finalmente considerados non aptos, vemos que en interacción aparece un 57 %,

seguido de cohesión e fluidez e pronuncia, cun 51,5 % cada un deles. Isto pode apuntar a unha tendencia nos avaliadores a considerar os criterios de interacción, cohesión e fluidez e pronuncia como secundarios á hora de considerar aos candidatos como aptos.

4.2.4.4. Cualificación na proba oral dos aptos

Existe un agrupamento de criterios no que, por unha banda, están interacción, adecuación e cohesión e fluidez (con puntuacións sobre 4) e, pola outra, o resto dos criterios, gramática e léxico (puntuados sobre 3). Isto pode sinalar unha tendencia nos avaliadores a ser máis esixentes en criterios relacionados co coñecemento do código que os criterios relacionados coa estrutura textual, por tratarse estes de criterios cos que están menos familiarizados.

4.2.5. Resultados da proba oral segundo as variables sociodemográficas

4.2.5.1. Cualificación na proba oral segundo o nivel de estudos

A relación dos candidatos aptos no Celga é directamente proporcional ao nivel de estudos que posúan. Os que teñen estudos universitarios presentan unha porcentaxe maior de aptitude no Celga (63,9 %).

4.2.5.2. Cualificación na proba oral segundo o lugar de residencia

Non é significativa, xa que os candidatos que habitan nas cidades presentan unha porcentaxe menor de aptos (67,8 %), seguidos dos habitantes das aldeas que presentan unha porcentaxe de aptos maior (70,2 %). Finalmente, os candidatos residentes nas vilas presentan un 73,8 %.

4.2.5.3. Cualificación na proba oral segundo o sexo

Trátase dunha variable significativa. Son os homes os que presentan unha porcentaxe maior de aptos (73,2 % fronte ao 68,6 % das mulleres). Ao contrario do que acontecía na proba escrita, tendo en conta, ademais, que a poboación xeral participante nas convocatorias de probas está distribuída case por igual entre homes e mulleres, as tres cuartas partes dos participantes nas probas Celga son mulleres (70,6 %).

4.2.5.4. Cualificacións na proba oral segundo a idade

Aínda que non é unha variable significativa, observamos diferenzas entre o parámetro de interacción, que non ten capacidade de discriminación, e o criterio de pronuncia, que ten unha forte variación segundo a idade do candidato.

Os maiores de 60 anos obteñen puntuacións significativamente máis baixas ca o resto en adecuación, cohesión e léxico. Este colectivo, como veremos nas discusións, require dunha especial atención por tratarse de galego falantes sen formación regrada en galego, cun baixo nivel de estudos e con necesidades de adquisición dun certificado oficial que acredite a súa competencia comunicativa en galego.

CAPÍTULO 5

Discusións

Discusións

Como cabe esperar, neste apartado dedicado ás discusións presentamos unha serie de cuestións que consideramos de interese sobre a utilidade dos criterios de corrección e avaliación do Celga, xa que, derivado da toma de decisións sobre os problemas expostos, podería mellorarse a calidade do sistema de certificación de lingua galega. Polo tanto, opinamos que a implementación do sistema debería ter en conta os temas que expoñemos a continuación.

Cómpre salientar que aínda que o sistema de certificación presenta unha serie de fallos que serán analizados e aos que lles daremos unha proposta de solución, debemos manter a percepción de que o Celga é un sistema de certificación efectivo en cuestións de constructo que o soporta, organización administrativa das probas, deseño das mesmas, protocolo de actuación durante a realización dos exames, custodia dos mesmos, transmisión de información dos resultados aos candidatos e impacto na cidadanía, como así demostra a rápida acollida e asentamento que tivo desde a súa aparición na sociedade galega.

5.1. Equilibrio no peso de todos os criterios de corrección e avaliación

Para os avaliadores na proba escrita, cando os resultados na escrita sinalaban que o peso do código era maior, o criterio de adecuación á tarefa converteuse nun supracriterio que agrupa todas as variables nunha única. Isto provoca que en ocasións se resolva a aptitude dun candidato en función dese único criterio que, en realidade, está fundamentado no cómputo de número de palabras permitidas na redacción do texto. Se analizamos como se define o criterio no manual de avaliadores:

Adecuación á tarefa na proba escrita

Neste parámetro valórase que o candidato se axuste á tarefa

— Extensión: A extensión permitida será dun 20 % como máximo, xa que a superación deste límite, por moi ben que cumpra os demais requisitos, indicará que non houbo unha resposta adecuada á tarefa e que non ofrece elementos suficientes para a súa avaliación. De todos os modos será necesario que este aspecto quede ben especificado ao comezo da proba nas indicacións para os candidatos.

— Tratamento dos puntos da tarefa: En canto ao tratamento dos puntos da tarefa, só contarán aqueles que se traten de forma explícita e con información pertinente.

— Formato textual: O texto debe axustarse ao formato textual pertinente, se ben poden permitirse leves alteracións nas convencións que configuran a macroestrutura dos textos.

— Rexistro: Debe utilizarse, en función do nivel, un rexistro adecuado á tarefa e un modo adecuado de dirixirse ao interlocutor.

Dos catro aspectos que engloban este criterio –extensión, tratamento dos puntos da tarefa, formato textual e rexistro–, o primeiro deles, a extensión do texto, é o que pode provocar que os avaliadores atendan a este parámetro como a unha mera acción mecánica de contar palabras no texto producido polo candidato e que o penalicen se presenta unha extensión un 20 % inferior ou superior ao número de palabras esixido segundo o nivel de lingua. Desta acción deriva a creación dun supracriterio que domina non só os tres aspectos restantes do criterio de adecuación á tarefa, senón que determina que un candidato sexa apto ou non.

Como resultado da situación na que un candidato pode considerarse non apto porque se desvía en máis dun 20% do número de palabras esixidas na proba escrita aínda que obtivese as puntuación necesarias para ser apto,

deberíamos reflexionar sobre a validez do criterio de adecuación á tarefa neste subapartado de extensión, xa que, tal e como defende Henning (1991, 285) no seu traballo *How valid is valid?* atendendo á importancia que se lles atribúe ás decisións que se toman a partir das puntuacións obtidas nunha proba, debemos ter en conta que no caso do Celga o feito de obter ou non a aptitude pode supoñer para un candidato optar a un traballo na Administración galega, e o impacto que pode ter no candidato fai que afecte a súa vida persoal e laboral de non ser o esperado. Porén, xa neste momento estase a levar a cabo unha revisión dos descritores que definen cada un dos criterios de corrección e avaliación, coa finalidade de mellorar os parámetros e dotar así aos avaliadores dunhas ferramentas máis fiables e xustas para desenvolver o seu labor, solventando así esta eiva e implementando o sistema de certificación Celga.

5.2. Equidade das probas Celga

Tal e como avanzamos no apartado de 3.1.2. (Hipóteses e obxectivos) deste capítulo, a equidade nas probas é unha das constantes que debe primar en calquera sistema de avaliación, de tal xeito que non exista ningún colectivo que se vexa prexudicado polo tipo de probas ou polo seu contido.

A pilotaxe do cuestionario realizado aos avaliadores amosou que unha das eivas que consideraban máis flagrante, repetida con insistencia, é que o sistema de certificación favorece a aquelas persoas que foron escolarizadas en galego e dificulta a posibilidade de aprobar dos galegofalantes maiores de 50 anos que carecen de formación académica en galego e non son quen de superar as probas escritas. Esta situación dáse en candidatos que se presentan ás probas dos niveis do Celga 1 e Celga 2, solicitados para os postos máis baixos da escala laboral da Administración galega. Ao non poder superar as probas, prodúcese nos candidatos un sentimento de rexeitamento cara ao sistema de

certificación de lingua proposto pola Administración galega; asemade, moitos avaliadores consideran que esta situación non dignifica o galego e predispón os candidatos contra a súa propia lingua.

Con respecto a esta eiva, cómpre analizar o perfil dos candidatos, persoas maiores de 50 anos e sen formación académica en galego que, aínda que sexan galegofalantes, non significa que teñan a suficiente competencia comunicativa como para poder superar as probas de certificación, nas que se esixe un nivel de lingua formal, lonxe do nivel de lingua coloquial que probablemente sexa o que empregan os ditos candidatos en calquera ámbito e situación comunicativa. De aí pode xurdir a confusión dos avaliadores á hora de considerar inxusto o Celga para persoas non escolarizadas en galego, tratándose unicamente dunha cuestión de nivel de lingua e non de equidade nas probas, xa que do deseño das probas, dos temas que hai que tratar ou das tarefas que se solicitan non se desprende que existan colectivos ou grupos de candidatos que se vexan favorecidos ou prexudicados á hora de poder completalas.

Polo tanto, e para evitar esa sensación nos candidatos maiores de 50 anos, galegofalantes e sen formación académica en galego cuxo obxectivo é obter un certificado de lingua que lles permita acceder aos postos máis baixos da categoría laboral da escala administrativa, cómpre un estudo empírico de perfís competenciais da función pública no que se especifique en que tipo de destreza debe ter competencia comunicativa o candidato que queira optar a determinados traballos, o cal permitirá a certificación de competencia comunicativa por destrezas (comprensión e expresión orais ou escritas) para os niveis máis baixos da escala laboral da Administración galega.

5.3. Sobre a selección dos avaliadores

Con respecto a este apartado centrarémonos nunha actuación cuestionable de carácter administrativo que implica a selección dos avaliadores de exames orais ou escritos. Segundo os criterios da SXPL, en cada convocatoria de probas realízase un chamamento para que, da bolsa de profesores colaboradores, todos eles poidan participar polo menos nunha xornada como avaliadores das probas orais en calquera das comisións repartidas nas diferentes cidades galegas que acollen as probas, de tal xeito que se garanta que todos os avaliadores poidan participar nas probas; con respecto á selección dos avaliadores das probas escritas, o criterio que prima é o da experiencia, pero non na proba oral, senón na proba escrita, co cal se dá o paradoxo de ter avaliadores cunha ampla experiencia nas probas orais (máis de catro convocatorias) e ningunha como avaliadores da proba escrita. Iso provoca que a valoración nos criterios de corrección da proba escrita e oral sexa diferente, tal como se demostrou no capítulo anterior, e provoquen que a curva de interiorización dos criterios de corrección e avaliación sexa significativo, tal como se viu no capítulo 3., apartado 1.10., na Figura 10 (Avaliadores que elixiron a alternativa todos son válidos á hora de determinar a competencia comunicativa escrita segundo a nivel de experiencia), onde podemos observar a existencia dunha curva que vai dende os avaliadores sen experiencia aos que teñen máis experiencia na avaliación da proba escrita.

Outra das cuestións que se debería ter en conta é que as xornadas de formación dos avaliadores están baseadas na análise de probas escritas. Como consecuencia, a todos aqueles avaliadores sen experiencia nas probas escritas pode non resultarlle tan útil xa que descoñecen a dita actividade. Aínda que os criterios de corrección son semellantes, sabemos que na avaliación da proba

escrita existe un supraconstruto marcado polo peso que ten o criterio de adecuación á tarefa, que non ocorre na avaliación da proba oral.

Polo tanto, recoméndase que todos os profesores colaboradores teñan a posibilidade de se converter tamén en avaliadores das probas escritas para que adquiran a perspectiva necesaria sobre a totalidade do exercicio de avaliación das probas escritas e orais, o cal lles permitirá interiorizar os criterios de corrección e isto reverterá nunha mellora no sistema de certificación, permitindo unha validez de construto firme e equilibrando a equidade na distribución dos roles entre os avaliadores, do cal resultarán avaliadores con máis experiencia, tanto nas probas orais coma nas escritas.

5.4. Sobre a fiabilidade e a validez

— **Sobre a fiabilidade interxuíces:** tal como avanzamos no apartado das conclusións, a fiabilidade interxuíces é baixa porque está centrada nos falsos positivos da proba escrita (todos os exames non aptos que corrixen os supervisores por segunda vez). Para solucionar esta eiva, cómpre que a fiabilidade se considere como un problema que hai que resolver *a priori*, é dicir, que antes de cada convocatoria todos os avaliadores e supervisores deberían recibir unha formación específica para sinalar cal é o nivel de corte dos aptos. Con isto evitariáse que os supervisores resultasen máis transixentes á hora de decidir a aptitude ou non dun candidato porque, na situación actual, os avaliadores pode que non teñan claro cal é o corte que separa o apto do non apto e aplican os criterios de xeito moi restrictivo, mentres que os supervisores son máis permisivos á hora de considerar apto un candidato.

— O sistema de control establecido como unha segunda volta de corrección de todos os exames non aptos e só algúns aptos provoca un erro na fiabilidade porque, se o nivel de corte estivese ben establecido, non sería

precisa esta segunda volta. Así mesmo, é importante sinalar que os candidatos non aptos poderían ter un sistema de control sempre que reclamasen unha segunda revisión e evitaríase unha solución *a posteriori* como a que existe actualmente que, ademais, non é fiable.

— **Sobre a validez de criterio:** o sistema Celga carece de validez de criterio porque non existe ningunha proba ou mecanismo que mida o mesmo ca a proba, isto é, non se lle aplica unha segunda proba de control aos candidatos ou un instrumento de competencia autopercebida que indique o nivel de competencia comunicativa que cada candidato cre que ten en galego. Para solucionar este defecto pódese aplicar un test de competencia autopercebida que, mediante unha escala de 1 ao 5, estableza a consideración que o candidato ten sobre a súa competencia comunicativa en galego. Outro xeito de medir a validez de criterio sería contrastando as puntuacións obtidas coas notas obtidas no sistema regrado ou noutras probas.

— **Sobre a validez de construto:** tal e como establecemos nas conclusións, a validez de construto é alta, xa que todos os criterios de corrección e avaliación están correlacionados e ademais agrupados en dous factores, un que atinxe os parámetros relacionados co coñecemento do código e outro factor que fai referencia aos parámetros da estrutura textual. Podería ser interesante en futuras investigacións observar se estes dous factores varían segundo o nivel de Celga ou se se manteñen en todos os niveis por igual.

— **Sobre a validez de contido:** con respecto aos criterios de corrección e avaliación empregados no Celga, xorde sempre a necesidade por parte do grupo de investigación de melloralos e, con ese espírito crítico, organízanse as xornadas de reflexión sobre estes, con axuda de expertos externos ao sistema que ofrecen unha visión nova e permiten planificar cambios que redunden en beneficio da validez. Lembremos que no estudo I os avaliadores confirmaron a validez de contido da proba outorgándolles aos criterios de corrección e

avaliación uns valores altos na utilidade para o seu labor, ademais de considerar que todos eles eran necesarios e non solicitar que se eliminase ningún deles.

— **Sobre o deseño das probas:** para comprobar a validez e fiabilidade da proba cómpre analizar as tarefas, é dicir, se todos os ítems que os candidatos teñen que completar se avalían un por un, aplicando os parámetros de corrección e avaliación.

— A validez diverxente debería medirse en futuras investigacións, é dicir, comprobar mediante outra proba que se está a medir o que se di que se está medir. Nestes momentos isto está comprobado mediante un comité de expertos e mediante a opinión dos avaliadores no estudo I.

5.5. Eivas detectadas polos avaliadores e propostas de mellora

Unha vez analizados todos os cuestionarios cubertos polos avaliadores, detectáronse unha serie de coincidencias nas opinións sobre o sistema de certificación. A maioría dos enquisados considera que o Celga é un sistema xusto e equitativo (74,5 % fronte o 25,5 %), mais isto non é óbice para sinalar unha serie de eivas que perciben dado o seu labor en todo o proceso, dende impartir cursos preparatorios para as probas ata avaliar estas. A continuación preséntanse as impresións dos avaliadores sobre o sistema e as propostas de mellora que se desprenden da lectura dos cuestionarios:

— Unha das eivas significativas que detectan os avaliadores está relacionada cos criterios de corrección e avaliación que deben empregar como referencias para desempeñar o seu labor. Consideran que poden resultar confusos e dar lugar a interpretacións subxectivas por parte de cada avaliador, dado que non se lles fornece un número aproximado de erros que poden ser permitidos en criterios como léxico ou ortografía. Polo tanto, solicitan

exemplos claros en cada parámetro como referencia de modelo óptimo de avaliación. Un dos criterios relevantes é o de adecuación á tarefa, xa que consideran que o número de palabras esixidas nas probas escritas non deberían ser un elemento discriminatorio á hora de decidir se un candidato pode ou non ser apto. Tamén consideran que algúns criterios non delimitan os distintos niveis de lingua e na proba oral os criterios de pronuncia e entoación deberían ter menos peso na suma total dos puntos.

— Opinan que o procedemento administrativo das probas pode ser mellorado, xa que atopan non equitativas as validacións establecidas pola SXPL dos sistemas de certificación antigos e os obtidos das validacións co ensino regrado, establecidos na Orde do 16 de xullo de 2007 (publicada no DOG número 146, do luns 30 de xullo de 2007), que regula os certificados oficiais de coñecemento dos diferentes niveis de lingua galega (Celga) e as probas para obtelos, así como as validacións no artigo 10 (*vid.* capítulo 2).

Por outra banda, da lectura dos cuestionarios despréndese un interese xeral dos avaliadores por mellorar o sistema achegando propostas e ideas, o que suxire a existencia dunha gran responsabilidade co seu labor e unha fonda preocupación por optimizar tanto os resultados dos candidatos coma a vinculación emocional que poden sentir coa lingua grazas ao sistema de certificación actual. Así, consideran positivo:

— O sistema de certificación Celga está aliñado co Marco europeo común de referencia para as linguas, o que coloca o galego en pé de igualdade con calquera lingua europea, dotándoa das ferramentas de ensino, aprendizaxe e avaliación promovidos polo Consello de Europa dende 2002, e permitindo a saída do localismo que viña sufrindo o galego e a vinculación cun proxecto europeísta.

— O tipo de probas que conforman o sistema de certificación atinxe as diferentes destrezas e permite calibrar o nivel de competencia comunicativa en cada unha delas dun xeito integrado, moi próximo ao que un usuario de lingua ten que realizar no seu quefacer cotián para comunicarse.

— Que, como sistema de certificación baseado en tarefas, permite que o avaliador obteña unha evidencia da competencia comunicativa do candidato, tanto na creación de textos orais coma escritos, afastando o Celga do movemento estrutural e gramatical de ensino, aprendizaxe e avaliación de linguas.

Polo tanto, cómpre ter en conta que os avaliadores son membros imprescindibles para o perfecto engraxe e funcionamento do sistema de certificación e, xa que logo, debería aumentar o número de horas de formación que reciben antes de cada convocatoria anual para garantir homoxeneidade e equidade á hora de empregar os criterios de corrección e avaliación.

5.6. Futuras liñas de estudo

Unha vez analizado o Celga observamos que se trata dun sistema cunha fonda consistencia no constructo teórico que o sustenta, dando como resultado unha asimilación moi natural na cidadanía galega. Debido a esta situación, estamos en condicións de sinalar que o Celga pode ser fonte de futuras liñas de estudo que permitan avances e melloras tanto no propio sistema como na investigación sobre a avaliación de linguas.

Non debemos esquecer que o Celga, xunto co sistema de certificación neerlandés, son os dous únicos certificados de lingua europeos baseados en tarefas, e polo tanto sería interesante para a comunidade de investigadores sobre avaliación de linguas un futuro estudo sobre as vantaxes e dificultades que supón este cambio didáctico no ensino, na aprendizaxe e, sobre todo, na

avaliación, xa que marcou un punto de inflexión neste século na didáctica e metodoloxía empregada, afectando a todos os ámbitos e persoas implicadas no proceso de certificación dunha lingua (Van de Branden, K., 2006).

CAPÍTULO 6

Anexo: cuestionario

VALORACIÓN DOS CRITERIOS DE CORRECCIÓN E AVALIACIÓN DOS Celgas

A continuación presentámoslle un cuestionario sobre o Celga. Agradecemoslle a súa colaboración neste estudo que pretende mellorar a calidade e efectividade do sistema de certificación, ademais de ter en conta as súas suxestións e opinións como avaliador e membro imprescindible no proceso de certificación de lingua galega. Todas as respostas son totalmente confidenciais e só serán empregadas coa finalidade de optimizar os resultados do Celga. Grazas pola súa colaboración.

1. Sinale o número de convocatorias do Celga nas que participou

-como avaliador da proba escrita:

-como avaliador da proba oral:

EXPRESIÓN ESCRITA

2. Valore o nivel de UTILIDADE dos criterios de corrección e avaliación do CELGA 1 para as probas de EXPRESIÓN ESCRITA cunha puntuación de 1 a 10, tendo en conta que o 1 é a puntuación máis baixa e o 10 a máis alta (marque cunha X a alternativa elixida):

	1	2	3	4	5	6	7	8	9	10
Adecuación										
Cohesión										
Gramática										
Léxico										

3. Valore o nivel de UTILIDADE dos criterios de corrección e avaliación do CELGA 2 para as probas de EXPRESIÓN ESCRITA cunha puntuación de 1 a 10, tendo en conta que o 1 é a puntuación máis baixa e o 10 a máis alta (marque cunha X a alternativa elixida):

	1	2	3	4	5	6	7	8	9	10
Adecuación										
Cohesión										
Gramática										
Léxico										

4. Valore o nivel de UTILIDADE dos criterios de corrección e avaliación do CELGA 3 para as probas de EXPRESIÓN ESCRITA cunha puntuación de 1 a 10, tendo en conta que o 1 é a puntuación máis baixa e o 10 a máis alta (marque cunha X a alternativa elixida):

	1	2	3	4	5	6	7	8	9	10
Adecuación										
Cohesión										
Gramática										
Léxico										

5. Valore o nivel de UTILIDADE dos criterios de corrección e avaliación do CELGA 4 para as probas de EXPRESIÓN ESCRITA cunha puntuación de 1 a 10, tendo en conta que o 1 é a puntuación máis baixa e o 10 a máis alta (marque cunha X a alternativa elixida):

	1	2	3	4	5	6	7	8	9	10
Adecuación										
Cohesión										
Gramática										
Léxico										

EXPRESIÓN ORAL

6. Valore o nivel de UTILIDADE dos criterios de corrección e avaliación do CELGA 1 para as probas de EXPRESIÓN ORAL cunha puntuación de 1 a 10, tendo en conta que o 1 é a puntuación máis baixa e o 10 a máis alta (marque cunha X a alternativa elixida):

	1	2	3	4	5	6	7	8	9	10
Adecuación										
Gramática										
Léxico										
Cohesión										
Pronuncia										
Interacción										

7. Valore o nivel de UTILIDADE dos criterios de corrección e avaliación do CELGA 2 para as probas de EXPRESIÓN ORAL cunha puntuación de 1 a 10, tendo en conta que o 1 é a puntuación máis baixa e o 10 a máis alta (marque cunha X a alternativa elixida):

	1	2	3	4	5	6	7	8	9	10
Adecuación										
Gramática										
Léxico										
Cohesión										
Pronuncia										
Interacción										

8. Valore o nivel de UTILIDADE dos criterios de corrección e avaliación do CELGA 3 para as probas de EXPRESIÓN ORAL cunha puntuación de 1 a 10, tendo en conta que o 1 é a puntuación máis baixa e o 10 a máis alta (marque cunha X a alternativa elixida):

	1	2	3	4	5	6	7	8	9	10
Adecuación										
Gramática										
Léxico										
Cohesión										
Pronuncia										
Interacción										

9. Valore o nivel de UTILIDADE dos criterios de corrección e avaliación do CELGA 4 para as probas de EXPRESIÓN ORAL cunha puntuación de 1 a 10, tendo en conta que o 1 é a puntuación máis baixa e o 10 a máis alta (marque cunha X a alternativa elixida):

	1	2	3	4	5	6	7	8	9	10
Adecuación										
Gramática										
Léxico										
Cohesión										
Pronuncia										
Interacción										

10. Que criterio de corrección e avaliación lle parece menos válido á hora de determinar a competencia comunicativa na PROBA ESCRITA? Sinale só unha opción.

- a) Coherencia e cohesión
- b) Adecuación á tarefa
- c) Gramática
- d) Léxico
- e) Todos son igual de válidos

11. Que criterio de corrección e avaliación lle parece menos válido á hora de determinar a competencia comunicativa na PROBA ORAL? Sinale só unha opción.

- a) Adecuación e interacción
- b) Cohesión e fluidez

- c) Pronuncia e entoación
- d) Gramática
- e) Léxico
- d) Todos son igual de válidos

12. Considera que hai algún/s criterio/s de corrección e avaliación que se debería/n engadir e que non se teñen en conta á hora de avaliar os exames?

- a)
- b)
- c)

13. Considera que se debería eliminar algún dos criterios de corrección e avaliación?

- a) Si (pase á pregunta 15)
- b) Non (pase á pregunta 16)

14. En caso de responder si, cal deles?

Na proba oral:

- a) Adecuación e interacción
- b) Cohesión e fluidez
- c) Pronuncia e entoación
- d) Gramática
- e) Léxico

Na proba escrita:

- a) Coherencia e cohesión
- b) Adecuación á tarefa

c) Gramática

d) Léxico

15. Considera que o Celga é un sistema xusto de avaliación?

a) Si

b) Non

16. Explique brevemente as razóns da súa resposta anterior.

BIBLIOGRAFÍA

BIBLIOGRAFÍA CAPÍTULO 1

- ALTE (1994). *The ALTE code of practics*. Cambridge: ALTE.
- ALTE (2003). *Preliminary version of a manual for relating examinations to the CEFR*. Cambridge: ALTE.
- ALTE (2008). *Manual for relating examinations to the CEFR*. Cambridge: ALTE.
- ALTE (2011). *Manual for language test development and examining for use with the CEFR*. Cambridge: ALTE.
- Alderson, Ch. (1986). "Computers in language testing". En G. Leech e Ch. Candlin (eds.), *Computers in English language learning and research*. Harlow: Longman.
- Alderson, Ch. (1991). "Bands and scores". En Ch. Alderson e B. North (eds.), *Language testing in the 1990s*. Londres: Modern English Publications e The British Council.
- Alderson, Ch. e Brian, N. (1991). *North language testing in the 1990s*. Londres: Modern English Publications e The British Council.
- Alderson, J. Charles (2005). *Diagnosing foreign language proficiency: the interface between learning and assessment*. Londres: Continuum.
- Alderson, J. C. et al. (2006). "Analysing test of reading and listening in relation of the common european framework of reference: the experience of the dutch CEFR construct project". *Language Assessment Quarterly*, 3 (1), 3-30. Laurence Erlbaum Associates, Inc.
- Bachman, L. F. e Palmer, A. (1996). *Language testing in practice*. Oxford: Oxford University Press.
- Bachman, L. F. (2000). "Modern language testing at the turn of the century: assuring that what count counts". *Language Testing*, 17, 1.

- Byram, M. e Beacco, J. C. (2003). *Guide for the development of language education policies in Europe, from languages diversity to plurilingual education*. Council of Europa.
- Bordón, T. (2004). "Panorama histórico del desarrollo de algunas de las cuestiones fundamentales en la evaluación de segundas lenguas". *Carabela*, 55, 5-29. Madrid.
- Brindley, G. (1994). "Task-centred assessment in language learning: the promise and the challenge". En N. Bird et al. (eds.) (1994), *Language and learning*. Hong Kong: Hong Kong Education Department.
- Brown, J. D. e Hudson, T. (1998). "The alternatives in language assessment". *TESOL Quarterly*, 32, 4, 653-675.
- Brown, J. D. (2002). *An investigation of second language task-based performance assessments*. Honolulu: University of Hawaii Press.
- Bygate, M. (ed.) (2001). *Researching pedagogic tasks: second language learning, teaching and testing*. New York: Pearson Education.
- Canale, M. e Swain, M. (1980). "Theoretical bases of communicative approaches to second language teaching and testing". *Applied Linguistics*, 1, 1, 1-47.
- Canale, M. (1983). "On some dimensions of language proficiency". En Oller, J. W. (ed.) (1983), *Issues in language testing research*. Rowley: Newbury House.
- Carrol, B. (1991). "Resistance to change". En Ch. Alderson e B. North (eds.), *Language testing in the 1990s*, 22-27.
- Chalhoub-Deville, M. (1996). "Performance assessment and the components of the oral construct across different tests and rater groups". En M. Milanovic e N. Saville (eds.), *Performance testing, cognition and assessment* (55-73). Cambridge: University of Cambridge, Local Examinations Syndicate e Cambridge University Press.

- Chalhoub-Deville, M. (2001). "Task-based assessments: characteristics and validity evidence". En M. Bygate et al. (eds.) (2001). *Researching pedagogic tasks: second language learning, teaching and testing*. New York: Pearson Education.
- Cassany, D. (2007). "Del portfolio al e-PEL", *Carabela*, 60 (*El portfolio europeo de las lenguas (o PEL) y la enseñanza de lenguas extranjeras/segundas lenguas*). Madrid: SGEL.
- Council of Europe (2001). *Common european framework for languages: learning, teaching, assessment*.
- Council of Europe (1989). *Language learning for European citizenship (1989-1996)*.
- Council of Europe (2003). *Manual for relating examinations to the common European framework of reference for languages. Preliminary pilot version*.
- Council of Europe (2001). *Portfolio europeo das linguas*.
- Davis, A. (1985). "Follow my leader: is that what language tests do?". En Y. P. Lee (ed.), *New directions in language testing (3-13)*. Oxford: Pergamon.
- Davis, A. (1991). "Language testing in the 1990s". En Ch. Alderson e B. North (eds.), *Language testing in the 1990s*, 136-149.
- Ellis, R. (2003). *Task-based language learning and teaching*. Oxford: Oxford University Press.
- Figueras, N. (2008). "El MCER, más allá de la polémica". *MarcoELE*, 7, 26-36. Madrid.
- Fulcher, G. (1996). "Testing tasks: issues in task design and the oral group". *Language Testing*, 13, 23-51.
- Huhta, A. (2008). "Asociaciones internacionales de evaluación y pruebas de lengua. ¿Qué son y qué hacen?". *MarcoELE*, 7, 42-52. Madrid.

- Huhta, A. (2007). "Diagnostic and formative assessment". En Spolsky, B. e Hult, F. (eds.), *Handbook of educational linguistics* (469-482). Blackwell.
- Lado, R. (1961). *Language testing: the construction and use of foreign language tests*. Londres: Longman.
- Lee, Fok, Lord e Low (1985). *New directions in language testing*. Oxford: Pergamon.
- Lynch, B.; Hamp-Lyons, L. (1999). "Perspectives on research paradigms and validity: tales from the language testing research colloquium". *Language Testing*, 8, 57-93. Melbourne.
- Long, M. H.; Norris, J. M. (2000). "Task-based language teaching and assessment". En Bayram, M. (ed.) (2000), *Encyclopedia of language teaching*. Londres: Routledge.
- McNamara, T. (1996). *Language testing*. Oxford: Oxford University Press.
- McNamara, T. (1996). *Measuring second language performance: a new era in language testing*. New York: Longman.
- Parrondo, J. R. (2004). "Modelos, tipos y escalas de evaluación". En Sánchez Lobato, J.; Santos Gargallo, I. (dirs.), *Vademécum para la formación de profesores* (967-982). Madrid: SGEL Ross.
- Púrpura, J. E. (1999). *Modeling the relationship between test takers' reported cognitive and metacognitive strategy use and performance on language tests*. Cambridge: University of Cambridge Local Examinations Syndicate e Cambridge University Press.
- Rea, P. (1985). "Language testing and the communicative language teaching curriculum". En Y. P. Lee (ed.), *New directions in language testing* (15-32). Oxford: Pergamon.
- Rea, P. (1991). "What makes a grammar test communicative?". En Ch. Alderson e B. North (eds.), *Language testing in the 1990s*, 112-131.

- Skehan, P. (1991): "Progress in language testing: the 1990s". En Ch. Alderson e B. North (eds.), *Language testing in the 1990s*, 3-6.
- Slageter, P. (1979). *Nivel umbral*. Estrasburgo: Consello de Europa.
- Stevenson, P. (1985). "Pop validity and performance testing". En Y. P. Lee e outros (eds.), *New directions in language testing* (111-118). Oxford: Pergamon.
- Swain, M. (1985). "Large-scale communicative language testing: a case study". En Y.P. Lee (ed.), *New directions in language testing* (35-46). Oxford, Pergamon.
- Trim, J. (2007). *Modern languages in the Council of Europe 1954-1997*. Estrasburgo: Language Policy Division.
- Van Ek, J. A. e Trim, J. (1980). *Threshold level English*. Oxford: Pergamon.
- Van Ek, J. A. e Trim, J. (1991). *Threshold level 1990* (edición revisada e corrixida). Cambridge: CUP.
- Xunta de Galicia. Presidencia. Secretaría Xeral de Política Lingüística (2005). *Marco europeo común de referencia para as linguas: aprendizaxe, ensino, avaliación*. Santiago de Compostela: Xunta de Galicia.
- Xunta de Galicia. Presidencia. Secretaría Xeral de Política Lingüística (2008). *Glosario de termos para a avaliación de linguas*. Santiago de Compostela: Xunta de Galicia.
- Wall, D. (1997). "Impact and washback in language testing". En C. Clapham e D. Corson (eds.), *Encyclopaedia of language and education*, 7 (*Language Testing and Assessment*). Dordrecht: Kluwer Academic.
- Weir, C. e Milanovic, M. (eds.) (2002). *Innovation and continuity: revising the Cambridge Proficiency Examination*. Cambridge: Cambridge University Press.

Wu, Y. (1998). "What do tests of listening comprehension test? A retrospection study of EFL test-takers performing a multiple-choice task". *Language Testing*, 15, 1, 21-44.

Sitios web:

ALTE: www.alte.org

ALTE checklist: http://www.alte.org/quality_assurance/code/checklist.php

Consello de Europa: http://www.coe.int/t/dg4/linguistic/Conference_EN.asp

Dialang: <http://www.dialang.org>

EALTA: <http://www.ealta.eu.org>

EALTA: Código de boas prácticas: <http://www.ealta.eu.org/documents/-archive/guidelines/Galician.pdf>

ILTA: <http://iltaonline.com>

Manual para relacionar exámenes con los niveles del Marco común europeo de referencia: <http://www.coe.int/T/DG4/Portfolio/documents/Overview.-doc>

Portfolio das linguas: <http://www.culture2.coe.int/portfolio>

SURVEYLANG: <http://www.surveylang.org/es/surveylang.html>)

Xunta de Galicia.Secretaría Xeral de Política Lingüística: www.xunta.es/-linguagalega

BIBLIOGRAFÍA CAPÍTULO 2

- Alderson, C.; Clapham, C.; Wall, D. (1998). *Exámenes de idiomas. Elaboración y evaluación*. Cambridge: Cambridge University Press.
- Candlin, C. (1990). *Hacia la enseñanza del lenguaje basada en tareas. Comunicación, lenguaje y educación*, 7-8.
- Council of Europe (2001). *Common european framework for languages: learning, teaching, assessment*.
- Estaire e Zanón (1990). *El diseño de unidades didácticas en L2 mediante tareas: principios y desarrollo*. Instituto Cervantes.
- Gysen, S. (2006). "Developing an introducing task-based language test". En *Task-based language education: from theory to practice*. Cambridge Applied Linguistics Series.
- Hoz, C. *Experiencias en el aula mediante tareas*. Segundo encuentro práctico de profesores de español en Alemania. International House Barcelona-Editorial Difusión. <http://www.encuentro-practico.com>.
- Juan Lázaro, O. (1999). "La enseñanza mediante tareas". *Actas do XXXIII Congreso Internacional da Asociación Europea de Profesores de Español*. Soria.
- Long, M. (1990). *Studies in second language acquisition*. Cambridge Univ Press.
- Marín Peris, M. (2001). "¿Qué significa trabajar en clase con tareas comunicativas?". En *Llengua i Ús*, 21. Barcelona: Direcció General de Política Llingüística, Generalitat de Catalunya.
- Melero Abadía, P. (2004). "De los programas nocional-funcionales a la enseñanza comunicativa". En *Vademécum para la formación de profesores* (2004) (689-714). Madrid: SGEL.

- Monteagudo, H.; Bouzada, X. M. (coords.) (2002). *O proceso de normalización do idioma galego (1980-2000). Volume I. Política Lingüística: análise e perspectivas*. Santiago de Compostela: Consello da Cultura Galega.
- Nunan, D. (1996). *El diseño de tareas para la clase comunicativa*. Madrid: Cambridge University Press.
- Richards, J. C. e Rodgers, T. S. (1998, 2.^a ed. 2001). *Enfoques y métodos en la enseñanza de idiomas*, Madrid: Cambridge University Press.
- Sánchez Pérez, A. (2004). “Metodología: conceptos y fundamentos”. En *Vademécum para la formación de profesores (2004)* (665-688). Madrid: SGEL.
- Van den Branden, K. (ed.) (2006). *Task- based language education: from theory to practice*. Cambridge Applied Linguistics Series.
- Xunta de Galicia. Presidencia. Secretaría Xeral de Política Lingüística (2005). *Marco europeo común de referencia para as linguas: aprendizaxe, ensino, avaliación*. Santiago de Compostela: Xunta de Galicia.
- Xunta de Galicia. Presidencia. Secretaría Xeral de Política Lingüística (2006). *Niveis de competencia en lingua galega*. Santiago de Compostela: Xunta de Galicia.
- Xunta de Galicia. Presidencia. Secretaría Xeral de Política Lingüística (2006). *Plan xeral de normalización da lingua galega*. Santiago de Compostela: Xunta de Galicia.
- Zanón, J. (1990). “Los enfoques por tareas para la enseñanza de las lenguas extranjeras”. *Cable*, 5.
- Zanón, J. (coord.) (1999). *La enseñanza del español mediante tareas* (53-57). Madrid, Edinumm.

BIBLIOGRAFÍA CAPÍTULO 3

- Alderson, C.; Clapham, C.; Wall, D. (1998). *Exámenes de idiomas. Elaboración y evaluación*. Cambridge: Cambridge University Press.
- ALTE (1994). *The ALTE code of practice*. Cambridge: ALTE.
- Bachman, L. F. (1990). *Fundamental considerations in language testing* (271-272). Oxford: Oxford University Press.
- Casalmiglia, H. e Tusón, A. (2002). *Las cosas del decir*. Barcelona: Ariel Lingüística.
- DiStefano, C.; Zhu, M. e Mindrila, D. (2009). “Understanding and using factor scores: considerations for the applied researcher”. *Practical Assessment, Research & Evaluating*, 14 (20). Disponible en liña: <http://pareonline.net/getvn.asp?v=14&n=20>
- García, E.; Gil, J. e Rodríguez, G. (2000). *Análisis factorial*. Madrid: La Muralla.
- García, R. (2003). “La medida en educación: concepto e implicaciones en las actividades diagnósticas. En R. García (ed.), *Técnicas e instrumentos de diagnóstico. Proyecto docente del departamento MIDE de la Universidad de Sevilla* (131-211). Sevilla: Editorial IETE.
- Garrett, H. E. (1937). *Statistics in psychology and education* (2.^a ed.). Oxford: Longmans, Green.
- Hair, Jr.; Anderson, R.; Tatham, R. L. e Black, W. C. (1999). *Análisis multivariante* (5.^a ed.). Madrid: Prentice Hall.
- Henning, G. (1991). “How valid is valid?”. *Language Learning*, 41 (*Communicative writing profiles: an investigation of the transferability of a multiple-trait scoring instrument across ESL writing assessment contexts*).

- Kerlinguer, F. N. (1975). *Investigación del comportamiento: técnicas y metodología*. México: Nueva Editorial Interamericana.
- Loredo, X.; Fernández, A.; Suárez, I. (2005). "Análise da expresión escrita en galego e castelán ao final do ensino obrigatorio: resultados dun estudo realizado na comarca de Santiago de Compostela". *Cadernos de lingua*, 27, 109–127.
- Martínez, R. (1996). *Psicometría: teoría de los test psicológicos y educativos*. Madrid: Síntesis.
- Pérez, C. (2004). *Técnicas de análisis multivariante de datos*. Madrid: Prentice Hall.
- Prieto, G. e Delgado, A. R. (2010). "Fiabilidad y validez". *Papeles del Psicólogo: Revista del Colegio Oficial de Psicólogos*, 31 (1), 67-74.
- Real, J. E. (2001) *Escalamiento multidimensional*. Madrid: Muralla.
- Sánchez, P.; Aguilar de Armas, I.; Fuentelsaz, C.; Casbas, M.; Moreno, T. e Hidalgo, R. (2005). "Fiabilidad de los instrumentos de medición en ciencias de la salud." *Enfermería Clínica*, 15, 4 , 227-236.
- Silva, B.; Barros, R.; Lorenzo, M.; Loredo, X.; Núñez, X.; Núñez, M.; Sans, D.; Castro, A.; Dávila, S. (2010). *Avaliación da competencia do alumnado de 4º da ESO nos idiomas galego e castelán* (216). Santiago de Compostela: Universidade de Santiago de Compostela, Investigación Educativa.
- Tabachnick, B. G. e Fidell, L. S. (1989). *Using multivariate statistics*. Nova York: Harper Collins.
- Valderrey, P. (2010). *SPSS 17. Extracción del conocimiento a partir del análisis de datos*. Paracuellos del Jarama: Rama.
- Van de Branden, K. (2006). *Task-based language education. Form theory to practice*. Cambridge: Cambridge University Press.

BIBLIOGRAFÍA XERAL

- Bachman, L. F. e Palmer, A. S. (1996). *Language testing in practice: designing and developing useful language tests*. Oxford: Oxford University Press.
- Bachman, L. F. (2002). “Some reflections on task-based language performance assessment”. *Language Testing* 19, 453-476.
- Brindley, G. (1994). “Task-centred assessment in language learning: the promise and the challenge”. En N. Bird et al. (eds.) (1994), *Language and learning*. Hong Kong: Hong Kong Education Department.
- Brown, J. D. et al. (2002). *An investigation of second language task-based performance assessments*. Honolulu: University of Hawaii Press.
- Bygate, M. (ed.) (2001). *Researching pedagogic tasks: second language learning, teaching and testing*. New York: Pearson Education.
- Canale, M. (1983). “On some dimensions of language proficiency”. En Oller, J. W. (ed. 1983). *Issues in language testing research*. Rowley: Newbury House.
- Canale, M. e M. Swain (1980). “Theoretical basis of communicative approaches to second language teaching and testing”. *Applied Linguistics*, 1, 1-47.
- Canale, M. (1983). “From communicative competence to communicative language pedagogy”. En Richards, J. C. e Schmidt, R. (eds.), *Language and communication*. Londres: Longman.
- Chalhoub-Deville, M. (2001). “Task-based assessments: characteristics and validity

- evidence”. En M. Bygate et al. (eds.) (2001), *Researching pedagogic tasks: second language learning, teaching and testing*. New York: Pearson Education.
- Ellis, R. (2003). *Task-based language learning and teaching*. Oxford: Oxford University Press.
- Ellis, R. (2005). “*La adquisición de segundas lenguas en un contexto de enseñanza*”. En Biblioteca Virtual redELE 2006. Disponible en línea: <http://www.sgci.mec.es/redele/biblioteca>
- Estaire, S. (1990). “La programación de unidades didácticas a través de tareas”. *Cable*, 5.
- Estaire, S. (1999a). “Tareas para el desarrollo de un aprendizaje autónomo y participativo”. En Zanón, J. (coord.), 1999.
- Estaire, S. (1999b). *Tareas para hacer cosas en español: principios y práctica de la enseñanza de lenguas extranjeras mediante tareas*. Colección Aula de Español, Universidad Antonio de Nebrija.
- Estaire, S. (2000). *Un procedimiento para diseñar unidades didácticas mediante tareas (I e II) y un instrumento para planificar clases*. DidactiRed del Centro Virtual del Instituto Cervantes ([//cvc.cervantes.es/aula/didactired](http://cvc.cervantes.es/aula/didactired))
- Fernández, S. (2001). *Tareas y proyectos en clase*. Madrid: Edinumen.
- Fernández, S. (2004). *Propuesta curricular y marco europeo de referencia: desarrollo por tareas*. Madrid: Edinumen.
- Fulcher, G. (1996). “Testing tasks: issues in task design and the oral group”. *Language Testing*, 13, 23-51.
- Gysen, S. e Van Avermaet, P. (2005). “Issues in functional language performance assessment: the case of the certificate dutch as a foreign language”. *Language Assessment Quarterly* 2, 1, 51-68.

- Long, M. H. & Norris, J. M. (2000). "Task-based language teaching and assessment". En Bayram, M. (ed.) (2000). *Encyclopedia of language teaching*. London: Routledge.
- Martín Peris, E. (1998). "Gramática y enseñanza de segundas lenguas". *Carabela*, 43.
- Martín Peris, E. (1999). "Libros de texto y tareas". En Zanón, J. (coord.), 1999.
- McNamara, T. (1996). *Measuring second language performance: a new era in language testing* (Monográficos marcoELE 2008, 7. Evaluación 121). Nova York: Longman.
- McNamara, T. (1997). "Performance assessment". En Clapham, C. e Corson, D. (eds.) (1997), *Encyclopedia of Language and Education*, 7, (*Language Testing and Assessment*). Dordrecht: Kluwer.
- Messick, S. (1994). "The interplay of evidence and consequences in the validation of performance assessments". *Educational Researcher*, 23, 13-23.
- Norris, J. M. et al. (1998). *Designing second language performance assessments*. Honolulu: University of Hawaii Press.
- Nunan, D. (2004). *Task-based language teaching*. Cambridge: Cambridge University Press.
- Skehan, P. (1998). *A cognitive approach to language learning*. Oxford: Oxford University Press.
- Soler Canela, O. (2007). *Diferentes enfoques en la evaluación de lenguas mediante tareas* (Monográficos marcoELE 2007, 8. Evaluación).
- Parrott, M. (1993). *Tasks for language teachers. A resource book for training and development*. Cambridge: Cambridge University Press.

- Skehan, P. (1998). *A cognitive approach to language learning*. Oxford: Oxford University Press.
- Skehan, P. (2001). "Tasks and language performance assessment". En M. Bygate et al. (eds.) (2001), *Researching pedagogic tasks: second language learning, teaching and testing*. New York: Pearson Education.
- Shekan, P. (2003). *Task based instruction*. *Language teaching* , 36, 1, xaneiro 2003, 1-14.
- Willis, J. (1996). *A framework for task-based Learning*. Harlow: Longman.
- Zanón, J. (1990). "Los enfoques por tareas". *Cable*, 5.
- Zanón, J. (coord.) (1999). *La enseñanza del español mediante tareas*. Madrid: Edinumen.