

## Annex A Procés de desenvolupament dels descriptors de domini de la llengua

Aquest annex analitza els aspectes tècnics de la descripció de l'assoliment dels nivells de competència lingüística. També tracta dels criteris per a la formulació de descriptors. A continuació s'enumeren les metodologies per al desenvolupament d'escala i es proporciona una bibliografia comentada.

### La formulació dels descriptors

L'experiència en la definició de nivells per a l'avaluació de llengües, la teoria sobre l'establiment de nivells en el camp més ampli de la psicologia experimental i les preferències dels docents quan se'ls ha consultat (per exemple, els esquemes de gradació d'objectius del Regne Unit o el projecte de Suïssa) suggereixen l'aplicació de les línies directrius següents per a l'elaboració de descriptors:

- **Formulació positiva.** Un tret comú de les escales de domini de la llengua centrades en l'examinador i de les escales de puntuació d'exàmens és que estan redactades de manera negativa en els nivells inferiors. Evidentment, en nivells inferiors, és més difícil formular el nivell de domini de la llengua de l'aprenent en funció d'allò que sap fer que no pas en funció del que no sap fer. Però si els nivells de domini de la llengua s'han de poder utilitzar com a objectius i no han de servir només per seleccionar els candidats, aleshores és preferible una formulació positiva. En certs casos es pot formular el mateix element de manera negativa i de manera positiva, per exemple, en relació amb el repertori de llengua.

#### Formulació positiva

- Té un repertori de llengua bàsic i d'estratègies que li permeten fer front a situacions previsible i quotidianes. (Nivell 3 d'Eurocentres: certificat).
- Té un repertori bàsic de lèxic i estratègies suficients per a la majoria de necessitats quotidianes, però generalment ha de buscar paraules i acceptar la limitació del missatge. (Nivell 3 d'Eurocentres: graella de l'examinador).
- El vocabulari gira al voltant d'àrees com objectes i llocs habituals, i els termes més comuns de relació de parentiu. (Principiant ACTFL).

#### Formulació negativa

- Té un repertori lingüístic reduït que l'obliga a utilitzar paràfrasis i a buscar paraules constantment. (Nivell 3 d'ESU).
- El domini limitat de la llengua provoca interrupcions de comunicació freqüents i malentesos en situacions poc habituals. (Finlàndia, Nivell 2).
- La comunicació es trenca perquè les limitacions lingüístiques interfereixen en el missatge. (Nivell 3 d'ESU).
- Té un vocabulari limitat. (Holanda, Nivell 1).
- La sèrie limitada de paraules i d'expressions obstaculitza la comunicació de pensaments i d'idees. (Gothenburg U)

<ul style="list-style-type: none"> <li>▪ Produeix i reconeix un conjunt de paraules i expressions breus apreses de memòria. (Trim 1978 Nivell 1).</li> </ul>	<ul style="list-style-type: none"> <li>▪ Només pot produir llistes d'expressions fetes i enumeracions de paraules. (Principiant ACTFL).</li> </ul>
<ul style="list-style-type: none"> <li>▪ Pot produir expressions quotidianes breus a fi de satisfer necessitats concretes (en l'àrea de les salutacions, la informació, etc.) (Elviri; Milà Nivell 1, 1986)</li> </ul>	<ul style="list-style-type: none"> <li>▪ Té un repertori lèxic molt bàsic i té un domini funcional de la llengua molt pobre. (Nivell 1 d'ESU).</li> </ul> <p style="text-align: right;"><b>Quadre A1</b></p>

Una complicació afegida en evitar les formulacions negatives és que certs trets del domini comunicatiu de la llengua no són acumulatius: com menys n'hi hagi, millor. L'exemple més evident el trobem en el que de vegades s'anomena independència, és a dir, el grau de dependència que l'aprenent té (*a*) del fet que l'interlocutor adapti el discurs, (*b*) de la possibilitat de demanar aclariment i (*c*) de la possibilitat de demanar ajuda per formular el que vol dir. Aquests punts sovint poden aparèixer sota la forma de condicions afegides als descriptors formulats de manera positiva, per exemple:

Pot comprendre un discurs clar i estàndard sobre temes familiars que va adreçat a ell, sempre que pugui demanar que li repeteixin o li tornin a formular de tant en tant el que li han dit.

Pot comprendre el que se li diu directament i de manera clara i lenta en una conversa quotidiana normal; pot arribar a comprendre el que s'està dient si l'interlocutor es pren la molèstia d'ajudar-lo.

o bé:

Pot interactuar amb una relativa facilitat en situacions estructurades i en converses breus sempre que l'interlocutor l'ajudi si és necessari.

- **La precisió.** Els descriptors han de descriure tasques concretes i/o graus concrets d'habilitat en la realització de tasques. Aquí cal destacar dos punts. En primer lloc, cal evitar la vaguetat en el descriptor, tal com ocorre, per exemple, en "pot utilitzar una sèrie d'estratègies adequades". Què s'entén per "estratègies"? Per a què han de ser "adequades"? Com s'ha d'interpretar el terme "sèrie"? El problema dels descriptors imprecisos és que com que són molt fàcils de llegir també són molt fàcils d'acceptar, i aquesta facilitat aparent pot amagar el fet que cadascú els interpreti d'una manera diferent. En segon lloc, des dels anys 40 s'admet el principi que les distincions entre els nivells d'una escala no haurien de dependre de la substitució d'un quantitatiu com "algun" o "certs" per "nombrosos" o "la majoria", o de "bastant ampli" per "molt ampli", o de "mitjà" per "bo" en el nivell immediatament superior. Les distincions haurien de ser reals, no només terminològiques, i això pot voler dir que hi ha buits en què no es poden formular distincions concretes i significatives.
- **La claredat.** Els descriptors han de ser transparents i han d'estar redactats amb un llenguatge no especialitzat. A banda de constituir una barrera per a la compren-

sió, sovint ocorre que quan s'elimina tot l'argot d'un descriptor impressionant, resulta que no queda gran cosa. En segon lloc, els descriptors s'haurien d'escriure amb una sintaxi senzilla i haurien de tenir una estructura lògica explícita.

- **La brevetat.** Una de les escoles de pensament s'associa a les escales holístiques, sobretot les que s'utilitzen als Estats Units i Austràlia, que intenten produir un paràgraf extens que cobreixi, de manera exhaustiva, els trets que es consideren essencials. Aquestes escales aconseguen precisió mitjançant una llista exhaustiva que té per objectiu transmetre una imatge detallada d'allò que els examinadors poden reconèixer com un aprenent característic d'un nivell concret, i per aquest motiu són fonts de descripció molt valuoses. Aquest enfocament, però, té dos inconvenients. En primer lloc, cap individu és realment "característic". Els trets detallats es combinen de maneres diferents. En segon lloc, és pràcticament impossible referir-se a un descriptor de més de dues frases durant el procés d'avaluació. Sembla que els professors prefereixen lògicament els descriptors breus. En el projecte d'elaboració dels descriptors il·lustratius, els professors tenien tendència a rebutjar o retallar els descriptors de més de 25 paraules (aproximadament, dues línies).
- **La independència.** Els descriptors breus tenen dos avantatges més. En primer lloc, és probable que descriguin un comportament del qual es pugui dir: "Sí, aquesta persona pot fer això." En conseqüència, els descriptors breus i concrets es poden utilitzar com a enunciats independents de criteris en llistes de control i en qüestionaris per a l'avaluació contínua que du a terme el professor i/o per a l'autoavaluació. Aquest tipus d'integritat independent indica que els descriptor podrien servir d'objectius, ja que el seu significat no depèn només de la formulació d'altres descriptors de l'escala. Això obre una sèrie de possibilitats per utilitzar-los en diferents formes d'avaluació (veg. capítol 9).

Els usuaris del *Marc de referència* han de tenir present i, si escau, determinar:

- quins criteris de tots els presentats són més adequats, i quins altres criteris s'utilitzen implícitament o explícitament en el seu context;
- fins a quin punt és desitjable i viable que les formacions del seu sistema compleixin criteris com els que s'han enunciat.

### Les metodologies per a l'elaboració d'escales

L'existència d'una sèrie de nivells pressuposa que determinats elements es poden col·locar en un nivell i no en un altre, i que les descripcions d'un grau concret d'habilitat pertanyen a un nivell i no a un altre. Això suposa que s'aplica de manera consistent un tipus d'esglaonament. Hi ha diverses maneres d'assignar les descripcions de domini de la llengua a diferents nivells. Els mètodes disponibles es poden classificar en tres grups: mètodes intuïtius, mètodes qualitius i mètodes quantitius. La major part de les escales de domini de la llengua existents, així com altres conjunts de nivells, s'han desenvolupat mitjançant un dels tres mètodes intuïtius del primer

grup. Els millors mètodes combinen els tres enfocaments en un procés complementari i acumulatiu. Els mètodes qualitius exigeixen la preparació i la selecció intuïtiva de material, i la interpretació intuïtiva dels resultats. Els mètodes quantitius haurien de quantificar qualitativament el material comprovat prèviament, i exigeixen una interpretació intuïtiva dels resultats. Per desenvolupar els nivells comuns de referència es va utilitzar una combinació d'enfocaments intuïtius, qualitius i quantitius. Si s'utilitzen els mètodes qualitius i quantitius, hi ha dos punts de partida possibles: els descriptors o les mostres d'actuació.

**Partir dels descriptors.** Un punt de partida és considerar el que es vol descriure, i després escriure, reunir o corregir esborranys de descriptors de categories concretes com a informació d'entrada de la fase qualitativa. Els mètodes 4 i 9, el primer i l'últim del grup dels mètodes qualitius que es presenten més endavant, són exemples d'aquest enfocament, que és especialment adequat per desenvolupar descriptors de categories relacionades amb el currículum, com ara les activitats lingüístiques comunicatives; però també es pot utilitzar per desenvolupar descriptors d'aspectes de la competència. L'avantatge de partir de categories i descriptors és que es pot definir una base teòrica equilibrada.

**Partir de les mostres d'actuació.** L'alternativa, que només es pot utilitzar per desenvolupar descriptors per valorar actuacions, és començar amb mostres representatives d'actuacions. Es pot demanar als examinadors què veuen quan valoren actuacions d'alumnes (qualitatives). Els mètodes 5-8 són variants d'aquesta idea. De manera alternativa, es pot demanar als examinadors que avaluïn les mostres i després utilitzar una tècnica estadística per identificar les característiques bàsiques en què fonamenten les seves decisions (quantitatives). Els mètodes 10 i 11 són exemples d'aquest enfocament. L'avantatge d'analitzar les mostres d'actuacions és que es pot arribar a descripcions molt concretes basades en dades.

L'últim mètode, el 12, és l'únic que realment descriu per nivells els descriptors en un sentit matemàtic. Aquest va ser el mètode utilitzat per desenvolupar els nivells comuns de referència i els descriptors il·lustratius, després del mètode 2 (intuïtiu) i dels mètodes 8 i 9 (qualitius). No obstant això, la mateixa tècnica estadística també es pot utilitzar després del desenvolupament de l'escala per validar l'ús de l'escala en la pràctica i identificar-ne les necessitats de revisió.

### ■ Mètodes intuïtius

Aquests mètodes no exigeixen cap recull estructurat de dades, només la interpretació de l'experiència fonamentada i basada en principis.

#### Núm. 1 *Expert*

Es demana a algú que redacti l'escala, la qual cosa pot fer a partir de la consulta d'escalas existents, de documents curriculars i de qualsevol altra font de material pertinent, possiblement després d'haver analitzat les necessitats del grup objecte d'estudi en qüestió. Després pot comprovar i revisar l'escala, possiblement utilitzant informadors.

**Núm. 2 Comitè**

Com en el cas de l'expert, però s'inclou un petit equip de desenvolupament, amb un grup més gran de consultors. Els consultors comenten els esbossos i poden treballar intuïtivament sobre la base de la seva experiència i/o sobre la base de comparacions amb els aprenents o amb mostres d'actuacions. Gipps (1994) i Scarino (1996; 1997) analitzen els punts febles de les escales curriculars de l'aprenentatge de llengües modernes en la secundària, elaborades a través de comitè al Regne Unit i a Austràlia.

**Núm. 3 Experimentat**

Com el del comitè, però el procediment s'allarga considerablement en el marc d'una institució i/o d'un context d'avaluació específic abans d'arribar a un consens. Un grup de persones participa en una mateixa interpretació de criteris i nivells. L'experimentació sistemàtica i la realimentació permeten perfeccionar la formulació. Grups d'examinadors poden analitzar les actuacions dels alumnes en relació amb les definicions, i les definicions en relació amb les mostres d'actuacions. Aquesta és la manera tradicional de desenvolupar les escales de domini de la llengua (Wilds, 1975; Ingram, 1985; Liskin-Gasparro, 1984; Loewe, 1985, 1986).

---

**■ Mètodes qualitius**


---

Tots aquests mètodes suposen la participació de petits tallers de treball amb grups d'informadors i una interpretació qualitativa més que no pas estadística de les informacions recollides.

**Núm. 4 Conceptes clau: formulació**

Una vegada es disposa d'un esbós d'escala, una tècnica senzilla consisteix a fragmentar l'escala i demanar als informadors representatius dels usuaris de l'escala que *a)* col·loquin les definicions segons l'ordre que considerin convenient, *b)* expliquin perquè creuen que és així i, després d'aclarir la diferència entre l'ordre que han escollit i l'ordre que s'havia previst inicialment, *c)* identifiquin els punts clau que els han ajudat o que els han confós. De vegades, es pot introduir una millora suprimint un nivell i establint una tasca secundària per determinar si entre dos nivells hi ha un buit que indica que manca un altre nivell entre ambdós. Les escales de certificació d'Eurocentres es van desenvolupar d'aquesta manera.

**Núm. 5 Conceptes clau: actuacions**

Es comparen els descriptors amb les actuacions característiques de cada nivell a fi d'assegurar la coherència entre el que s'ha descrit i el que ha passat. Algunes guies dels exàmens de Cambridge sotmeten els professors

---

a aquest procés, i comparen la redacció de les escales amb les notes atorgades a un cert nombre d'exercicis d'expressió escrita. Els descriptors de l'IELTS (*International Language Testing System*) es van elaborar demanant a grups de correctors experimentats que identifiquessin "mostres clau d'exercicis escrits" de cada nivell i que es possessin d'acord sobre les "característiques clau" de cada exercici. Després, mitjançant debats, es van determinar els trets considerats com a característics de diferents nivells i es van incorporar als descriptors (Alderson, 1991; Shohamy *et al.*, 1992).

#### **Núm. 6** *Tret principal*

Els informadors classifiquen individualment les actuacions (generalment escrites) dels alumnes. Després es negocia una classificació comuna. A continuació, s'identifica i es descriu per a cada nivell el principi segons el qual s'han classificat els exercicis escrits, procurant posar en evidència els trets remarcables de cada nivell. El tret descrit (característica, constructe) determina l'ordre de la classificació (Mullis, 1980). Una variant habitual és l'organització en un cert nombre de conjunts, més que no pas en un ordre de classificació. Hi ha també una variant multidimensional interessant de l'enfocament tradicional. En aquesta versió primer es determinen els trets més significatius mitjançant la identificació de les característiques clau (el núm. 5 anterior). Després es classifiquen les mostres de cada tret de manera separada. Així, al final, es disposa d'una escala analítica dels múltiples trets més que d'una escala holística del tret principal.

#### **Núm. 7** *Decisions binàries*

Una altra variant del mètode del tret principal consisteix a classificar primer mostres representatives en conjunts de nivells. A continuació, en un debat que se centra en els límits entre els nivells, s'identifiquen les característiques clau (com al núm. 5 anterior). Després es formula la característica en qüestió com una pregunta breu de criteri que tingui per resposta Sí o No. Així es forma un arbre de tries binàries, la qual cosa ofereix a l'examinador un algoritme de decisions a seguir (Upshur i Turner, 1995).

#### **Núm. 8** *Valoracions comparatives*

Els grups examinen parelles d'actuacions i estableixen quina és la millor i per què. D'aquesta manera s'identifiquen les categories del metallenguatge utilitzades pels correctors i els trets més característics que funcionen en cada nivell. Aquests trets es poden introduir aleshores en la formulació dels descriptors (Pollit i Murray, 1996).

#### **Núm. 9** *Tasques de selecció*

Quan es disposa d'un esbós dels descriptors, es pot demanar als informadors que els classifiquin segons les categories que se suposa que descriuen i/o segons els nivells. També se'ls pot demanar que comen-

tin els descriptors, que els corregeixin o esmenin i que identifiquin els que són especialment clars, útils, adequats, etc. El banc de descriptors d'on s'han extret les escales il·lustratives es va desenvolupar i preparar d'aquesta manera (Smith i Kendall, 1963; North, 1996/2000).

### ■ Mètodes quantitius

Aquests mètodes suposen una quantitat important d'anàlisi estadística i una interpretació curiosa dels resultats.

#### **Núm. 10 *Anàlisi discriminant***

Primer, un conjunt de mostres d'actuació que ja s'han valorat (preferiblement per un equip) se sotmeten a una anàlisi detallada del discurs. Aquesta anàlisi qualitativa determina i comptabilitza la freqüència de les diferents característiques qualitatives. A continuació, s'utilitza la regressió múltiple per establir quines característiques identificades són significatives per determinar la classificació que aparentment van fer els examinadors. Aquestes característiques clau s'incorporen després a la formulació dels descriptors de cada nivell (Fulcher, 1996).

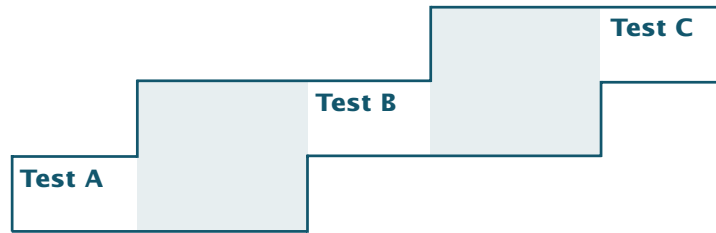
#### **Núm. 11 *Escales multidimensionals***

Tot i el nom, es tracta d'una tècnica descriptiva per identificar les característiques clau i la relació entre elles. Les actuacions es valoren mitjançant una escala analítica de diverses categories. El producte de l'anàlisi tècnica demostra quines categories són realment decisives per determinar un nivell i proporciona un diagrama que marca la proximitat o la distància entre les diferents categories. És una tècnica de recerca per identificar i validar criteris remarcables (Chaloub-Deville, 1995).

#### **Núm. 12 *Teoria de la resposta a l'ítem (TRI) o anàlisi dels trets latents***

La TRI ofereix un conjunt de models d'elaboració d'escales de mesura. El més simple i el més sòlid és el model de *Rasch*, que pren el nom de George Rasch, matemàtic danès. La TRI és un desenvolupament de la teoria de probabilitats i s'utilitza principalment per determinar la dificultat dels ítems individuals, procedents d'un banc d'ítems, que apareixen en una prova. Si s'està en un nivell avançat, les probabilitats de respondre una pregunta de nivell elemental són molt altes; si s'està en un nivell elemental, les probabilitats de respondre un ítem de nivell avançat són molt baixes. El model de Rasch ha permès convertir aquest fet tan simple en una metodologia de descripció de nivells que es pot utilitzar per calibrar els ítems d'una mateixa escala. Un desenvolupament d'aquest enfocament permet utilitzar-lo per distribuir en nivells tant els descriptors de domini comunicatiu de la llengua com els ítems de les proves.

En l'anàlisi de Rasch es pot plantejar un encadenament de diferents proves o qüestionaris mitjançant la utilització d'"ítems àncora", que són comuns a les proves adjacents. En el diagrama següent els ítems àncora estan ombrejats. D'aquesta manera, parts de la prova es poden dissenyar per a grups d'aprenents concrets, però es poden enllaçar en una escala comuna. S'ha de ser prudent, no obstant això, en aquest procés, ja que el model distorsiona els resultats de les puntuacions altes i de les puntuacions baixes de cada prova.



L'avantatge del model de Rasch és que pot proporcionar una mesura independent de mostres i d'escala; és a dir, una descripció de nivells independent de les mostres o de les proves i qüestionaris utilitzats en l'anàlisi. Es proporcionen els valors de l'escala que romanen constants per a tots els grups futurs a condició que aquests subjectes futurs es puguin considerar grups nous en el marc de la mateixa població estadística. Els canvis sistemàtics de valors que tenen lloc amb el temps (per exemple, a causa d'un canvi del currículum o en la formació dels examinadors) es poden quantificar i adaptar. Així mateix, es poden quantificar i adaptar les variacions sistemàtiques entre els tipus d'aprenents o d'examinadors (Wright i Masters, 1982; Linacre, 1989).

Les anàlisis de Rasch es poden utilitzar de diferents maneres per situar descriptors en una escala:

- a)** Amb el model de Rasch, les dades de les tècniques qualitatives núm. 6, 7 o 8 es poden traslladar en una escala aritmètica.
- b)** Es poden elaborar proves curosament per fer operatius els descriptors de domini de la llengua, especialment en alguns ítems concrets de proves. Aleshores aquests ítems es poden descriure per nivells amb l'anàlisi de Rasch i els seus valors es poden utilitzar en l'escala per indicar la dificultat relativa dels descriptors (Brown *et al.*, 1992; Carroll, 1993; Masters, 1994; Kirsch, 1995; Kirsch i Mosenthal, 1995).
- b)** Es poden utilitzar els descriptors com a preguntes d'un qüestionari per a l'avaluació que el docent fa als aprenents ("Pot fer això?"). D'aquesta manera, els descriptors es poden calibrar directament en una escala aritmètica, de la mateixa manera que els ítems de la prova es descriuen per nivells en els bancs d'ítems.
- d)** Les escales de descriptors incloses en els capítols 3, 4 i 5 es van desenvolupar d'aquesta manera. En els tres projectes descrits en els annexos B, C i D s'ha utilitzat la metodologia de Rasch per descriure per nivells els descriptors i per comparar les escales de descriptors resultants.



A més de ser útil per al desenvolupament d'una escala, el model de Rasch també es pot utilitzar per analitzar la manera com s'utilitzen realment els nivells en una escala d'avaluació. Això pot contribuir a destacar la redacció imprecisa, la utilització errònia o la sobrevaloració d'un nivell, i donar-ne a conèixer la necessitat de revisió (Davidson, 1992; Milanovic *et al.* 1996; Stansfield i Kenyon, 1996; Tyndall i Kenyon, 1996).

Els usuaris del *Marc de referència* han de tenir present i, si escau, determinar:

- fins a quin punt les notes que s'atribueixen en el seu sistema d'avaluació tenen el mateix sentit i unes definicions comunes;
- quins dels mètodes descrits anteriorment –o quins altres mètodes– s'utilitzen per desenvolupar aquestes definicions.

## Bibliografia seleccionada i comentada: Escala de domini de la llengua

ALDERSON, J.C. «Bands and scores». A: ALDERSON, J.C.; NORTH, B. [ed.], *Language testing in the 1990s. Developments in ELT*. Londres: British Council/Macmillan, 1991, p. 71-86

Analitza els problemes causats per la confusió de finalitats i orientació, i el desenvolupament de les escales d'expressió oral d'IELTS.

BRINDLEY, G. «Defining language ability: the criteria for criteria». A: ANIVAN, S. [ed.] *Current developments in language testing*. Singapur: Regional Language Centre, 1991.

Crítica raonada de la reivindicació de les escales de domini de la llengua per representar l'avaluació amb referència a un criteri.

BRINDLEY, G. «Outcomes-based assessment and reporting in language learning programmes, a review of the issues». *Language Testing* [s. l.] 15, 1 (1998), p. 45-85.

Crítica l'èmfasi en els resultats en funció d'allò que els aprenents poden fer més que en els aspectes de la competència emergent.

BROWN, A.; ELDER, C.; LUMLEY, T; McNAMARA, T.; McQUEEN, J. «Mapping abilities and skill levels using Rasch techniques.» Article presentat al catorzè *Language Testing Research Colloquium*, Vancouver. Reimprès a: *Papers in Applied Linguistics*, [Melbourne] 1/1 (1992), p. 37-69.

Ús tradicional de la gradació d'ítems de proves amb el model de Rasch per tal de produir una escala de domini de la llengua a partir de tasques de comprensió lectora provades en ítems diferents.

CARROLL, J.B. «Test theory and behavioural scaling of test performance». A: FREDERIKSEN, N.; MISSLEVY, R.J.; BEJAR, I. [ed.] *Test theory for a new generation of tests*. Hillsdale, N.J.: Lawrence Erlbaum Associates, 1993, p. 297-323.

Article fonamental que recomana l'ús del model de Rasch per esglaonar els ítems d'una prova i elaborar així una escala de domini de la llengua.

CHALOUB-DEVILLE, M. «Deriving oral assessment scales across different tests and rater groups.» *Language Testing* [s. l.], 12, 1 (1995), p. 16-33.

Estudi que revela els criteris que utilitzen els parlants nadius d'àrab per valorar els aprenents. És pràcticament l'única aplicació de la descripció per nivells multidimensional en proves de llengua.

DAVIDSON, F. «Statistical support for training in ESL composition rating». A: HAMP-LYONS [ed.] *Assessing second language writing in academic contexts*. Norwood, N.J.: Ablex, 1992, p. 155-166.

Un relat molt clar de la manera de validar una escala d'avaluació en un procés cíclic amb l'anàlisi de Rasch. Defensa un enfocament "semàntic" de la descripció per nivells més que no pas un enfocament "concret" recollit, per exemple, en els descriptors il·lustratius.

FULCHER. «Does thick description lead to smart tests? A data-based approach to rating scale construction». *Language Testing* [s. l.], 13, 2 (1996), p. 208-38.

Enfocament sistemàtic sobre el desenvolupament dels descriptors i de l'escala, començant per una anàlisi detallada del que succeeix realment en l'actuació. Un mètode que requereix molt de temps.

GIPPS, C. *Beyond testing*. Londres: Falmer Press, 1994.

A favor de l'avaluació que du a terme el professor i que se centra en un nivell d'exigència amb relació a punts comuns de referència establerts mitjançant el treball en grup. Anàlisi dels problemes provocats per descriptors imprecisos en el currículum nacional anglès. Multidisciplinar.

KIRSCH, I.S. «Literacy performance on three scales: definitions and results». A: *Literacy, economy and society: Results of the first international literacy survey*. París: Organisation for Economic Cooperation and development (OECD), 1995, p. 27-53.

Informe senzill i no tècnic sobre un ús sofisticat del model de Rasch per elaborar una escala de nivells a partir d'unes dades d'exàmens. Mètode desenvolupat per predir i explicar la dificultat dels ítems d'una prova partint de les tasques i de les competències implicades, és a dir, amb relació a un marc.

KIRSCH, I.S.; MOSENTHAL, P.B. «Interpreting the IEA reading literacy scales». A: BINKLEY, M.; RUST, K.; WINGLEE, M. [ed.] *Methodological issues in comparative educational studies: The case of IEA reading literacy study*. Washington DC.: US Department of Education, National Center for Education Statistics, 1995, p. 135-192.

Versió més detallada i tècnica de l'informe anterior que traça el desenvolupament del mètode a través de tres projectes relacionats.

LINACRE, J.M. *Multi-faceted Measurement*. Chicago: MESA Press, 1989.

Punta de llança en estadística que permet tenir en compte el rigor dels examinadors a l'hora d'informar dels resultats d'una avaluació. Aplicat al projecte per desenvolupar els descriptors il·lustratius a fi de comprovar la relació entre els nivells i els cursos escolars.

- LISKIN-GASPARRO, J.E. «The ACTFL proficiency guidelines: Gateway to testing and curriculum». A: *Foreign Language Annals* [s. l.], 17/5 (1984), p. 475-489.  
Esbós dels propòsits i desenvolupaments de l'escala americana d'ACTFL, a partir de la de l'Institut de Serveis Estrangers (FSI).
- LOWE, P. «The ILR proficiency scale as a synthesising research principle: the view from the mountain». A: JAMES, C.J. [ed.] *Foreign Language Proficiency in the Classroom and Beyond*. Lincolnwood (Ill): National Textbook Company, 1985.  
Descripció detallada del desenvolupament de l'escala de la Interagency Language Roundtable (ILR) dels Estats Units, a partir de la seva predecessora, l'escala de l'FSI. Funcions de l'escala.
- LOWE, P. «Proficiency: panacea, framework, process? A Reply to Kramsch, Schulz, and particularly, to Bachman and Savignon». A: *Modern Language Journal* [s. l.] 70/4 (1986), p. 391-397.  
Defensa d'un sistema que va funcionar bé –en un context específic– malgrat la crítica acadèmica que va suscitar la difusió d'aquesta escala i de la seva metodologia basada en entrevistes per a l'educació (amb ACTFL).
- MASTERS, G. «Profiles and assessment». *Curriculum Perspectives* [s. l.], 14, 1 (1994), p. 48-52.  
Breu informe de com s'ha utilitzat el model de Rasch per descriure per nivells els resultats d'exàmens i les avaluacions que fa el professor amb la finalitat de crear un sistema curricular de perfils a Austràlia.
- MILANOVIC, M.; SAVILLE, N.; POLLITT, A.; COOK, A. «Developing rating scales for CASE: Theoretical concerns and analyses.» A: CUMMING, A.; BERWICK, R. *Validation in language testing*. Clevedon, Avon: Multimedia Matters, 1996, p. 15-38.  
Explicació tradicional de l'ús del model de Rasch per perfeccionar una escala de puntuació utilitzada en una prova d'expressió oral, reduint els nivells de l'escala a un nombre que els examinadors puguin utilitzar de manera eficaç.
- MULLIS, I.V.S. *Using the primary trait system for evaluating writing*. Manuscrit núm. 10-W-51. Princeton, NJ: Educational Testing Service, 1981.  
Explicació tradicional del mètode del tret principal en l'expressió escrita en llengua materna per desenvolupar una escala d'avaluació.
- NORTH, B. *The development of descriptors on scale of proficiency: perspectives, problems, and a possible methodology*. Article ocasional NFLC (National Foreign Language Center). Washington DC: [s. n.], abril 1993.  
Crítica del contingut i del desenvolupament de la metodologia de les escales tradicionals de domini de la llengua. Proposta d'un projecte per desenvolupar els descriptors il·lustratius amb professors i esglaonar-los amb el model de Rasch, partint de l'avaluació que fa el docent.
- NORTH, B. *Scales of language proficiency: a survey of some existing systems*. CC-LANG (94) 24. Estrasburg: Consell d'Europa, 1994.  
Estudi exhaustiu d'escales curriculars i d'escales de puntuació, analitzades i uti-

litzades posteriorment com a punt de partida del projecte per desenvolupar descriptors il·lustratius.

NORTH, B. *The development of a common framework scale of language proficiency*. Tesi doctoral (1996/2000). Thames Valley University. Reimprès a Nova York: Peter Lang, 2000.

Anàlisi de les escales de domini de la llengua, de com es relacionen els models de competència i l'ús de la llengua amb les escales. Informe detallat de les fases de desenvolupament del projecte d'elaboració dels descriptors il·lustratius, dels problemes sorgits i de les solucions adoptades.

NORTH, B. *Scales for rating language performance in language tests: descriptive models, formulation styles and presentation formats*. Article d'investigació de TOEFL. Princeton, NJ: Educational Testing Service. [En premsa.]

Anàlisi detallada i estudi històric dels tipus d'escales de valoració utilitzades en proves d'expressió oral i d'expressió escrita: avantatges, inconvenients, dificultats, etc.

NORTH, B.; SCHNEIDER, G. «Scaling descriptors for language proficiency scales.» *Language Testing* [s. l.], 15/2 (1998), p. 217-262.

Visió global del projecte d'elaboració dels descriptors il·lustratius. Examina els resultats i l'estabilitat de les escales. En un annex s'ofereixen exemples d'instruments i de productes.

POLLITT, A.; MURRAY, N.L. «What raters really pay attention to.» A: MILANOVIC, M.; SAVILLE, N. [ed.] *Performance testing, cognition and assessment*. Studies in Language Testing 3. Articles seleccionats del quinzè *Language Testing Research Colloquium*, que va tenir lloc a Cambridge i a Arnhem del 2 al 4 d'agost de 1993. Cambridge: University of Cambridge Local Examinations Syndicate, 1996, p. 74-91.

Article metodològic interessant que vincula l'anàlisi de graelles de repertori utilitzades en psicologia amb una tècnica senzilla de descripció de nivells, per determinar en què se centren els examinadors en els diferents nivells de domini de la llengua.

SCARINO, A. «Issues in planning, describing and monitoring long-term progress in language learning», a les actes del desè Congrés de Llengües Nacionals d'AFMLTA, 1996, p. 65-75.

Critica l'ús de formulacions imprecises i la manca d'informació sobre la qualitat de les actuacions dels aprenents d'acord amb els enunciats descriptius característics del currículum al Regne Unit i a Austràlia per a l'avaluació que fa el professor.

SCARINO, A. «Analysing the language of frameworks of outcomes for foreign language learning», a les actes de l'onzè Congrés de Llengües Nacionals d'AFMLTA, 1997, p. 241-258.

Com l'anterior.

SCHNEIDER, G.; NORTH, B. "In anderen Sprachen Kann ich". *Skalen zur Beschreibung, Beurteilung und Selbsteinschätzung der fremdsprachlichen Kommunikations-*

*fähigkeit*, NFP 33/SKBF (Umsetzungsbericht), Berna/Aarau: [s. n.], 1999.

Breu informe sobre el projecte d'elaboració de les escales il·lustratives. També presenta la versió suïssa del Carnet (40 pàgines A5).

SCHNEIDER, G.; NORTH, B. "Dans d'autres langues, je suis capable de..." *Echelles pour la description, l'évaluation et l'auto-évaluation des compétences en langues étrangères*. PNR 33/CSRE (informe de valoració), Berna/Aarau: [s. n.], 2000.  
Com l'anterior.

SCHNEIDER, G.; NORTH, B. *Fremdsprachen können was heisst das? Skalen zur Beschreibung, Beurteilung und Selbsteinschätzung der fremdsprachlichen Kommunikationsfähigkeit*. Chur/Zurich: Verlag Rüegger AG, 2000.  
Informe complet del projecte d'elaboració de les escales il·lustratives. Capítol senzill sobre la descripció de nivells en anglès. També introdueix la versió suïssa del Carnet.

SKEHAN, P. «Issues in the testing of English for specific purposes». A: *Language Testing* [s. l.] 1/2, 1984, p. 202-220.  
Critica les referències a la norma i la formulació de les escales d'ELTS

SHOHAMY, E.; GORDON, C.M.; KRAEMER, R. «The effect of rater's background and training on the reliability of direct writing tests». *Modern Language Journal* [s. l.], 76 (1992), p. 27-33.  
Explicació senzilla del mètode qualitatiu per al desenvolupament d'una escala analítica per a l'expressió escrita. Va proporcionar una fiabilitat sorprenent entre examinadors no professionals.

SMITH, P.C.; KENDALL, J.M. «Retranslation of expectations: an approach to the construction of unambiguous anchors for rating scales». A: *Journal of Applied Psychology* [s. l.], 47/2 (1963).  
El primer enfocament de l'esglaonament de descriptors i no tan sols de la redacció d'escales. De gran influència. Molt difícil de llegir.

STANSFIELD, C.W.; KENYON, D.M. «Comparing the scaling of speaking tasks by language teachers and the ACTFL guidelines.» A: CUMMING, A.; BERWICK, R. *Validation in language testing*. Clevedon, Avon: Multimedia Matters, 1996, p. 124-153.  
Utilització del model d'esglaonament de Rasch per confirmar l'ordre de classificació de les tasques que apareixen a les guies d'ACTFL. Estudi metodològic interessant que va inspirar l'enfocament recollit en el projecte d'elaboració dels descriptors il·lustratius.

TAKALA, S.; KAFTANDJIEVA, F. «Council of Europe scales of language proficiency: A validation study». A: ALDERSON, J.C. [ed.] *Case studies of the use of the Common European Framework*. Consell d'Europa. [En premsa.]  
Informe sobre la utilització d'un model de Rasch més desenvolupat a l'hora de graduar per nivells l'autoavaluació en relació amb adaptacions dels descriptors il·lustratius. Context: El projecte DIALANG: proves per al finès.

TYNDALL, B.; KENYON, D. «Validation of a new holistic rating scale using Rasch multifaceted analysis». A: CUMMING, A.; BERWICK, R. *Validation in language testing*. Clevedon, Avon: Multimedia Matters, 1996, p. 9-57.

Explicació senzilla de la validació d'una escala per validar entrevistes en anglès com a segona llengua per a la prova d'ESL d'accés a la universitat. Ús tradicional del model plurifacètic de Rasch per identificar les necessitats de formació.

UPSHUR, J.; TURNER, C. «Constructing rating scales for second language tests.» *English Language Teaching Journal* [s. l.], 49, 1 (1995), p. 3-12.

Desenvolupament més sofisticat de la tècnica del tret principal per elaborar gràfics de decisions binàries. Molt rellevant per al sector escolar.

WILDS, C.P. «The oral interview test». A: SPOLSKY, B.; JONES, R. [ed.] *Testing Language Proficiency*. Washington DC: Center for Applied Linguistics, 1975, p. 29-44.

La primera versió de l'escala original de valoració del domini de la llengua. Mereix una lectura atenta per observar matisos que s'han perdut en la major part dels enfocaments d'entrevista des d'aleshores.