

Adam, Montserrat, Josefina Carrera, Núria Alturo i Josep Besa (2006), "Una proposta d'avaluació de l'expressió escrita per millorar la planificació docent". *Actes del 4rt Congrés Internacional de Docència Universitària i Innovació, CIDUI* (Barcelona, 5-7 July, 2006). [CD-Rom]

**I. Títol:** *Una proposta d'avaluació de l'expressió escrita per millorar la planificació docent*

**II. Autors:** Montserrat Adam, Josefina Carrera, Núria Alturo i Josep Besa

**III. Paraules clau:** *avaluació de l'expressió escrita, procés d'investigació-acció, planificació docent*

#### **IV. Resum**

Presentem aquí una proposta sobre l'avaluació de l'expressió escrita que s'emmarca en un procés d'investigació-acció (Kemmis i McTaggart 1988; Elliott 1993) orientat a millorar les habilitats de producció escrita dels estudiants de l'ensenyament de Filologia Catalana de la Universitat de Barcelona.

En el treball descrivim els paràmetres conceptuals en què es basa la correcció, per part nostra, d'un corpus de textos escrits procedents dels estudiants esmentats i presentem la combinació d'aquests paràmetres en una matriu o graella classificatòria d'errors que té una aplicació computacional. En aquesta graella es relaciona les propietats del text (adequació —que inclou la correcció—, coherència i cohesió) i les unitats susceptibles d'anàlisi que conformen el text (que estan connectades amb la morfologia, el lèxic, la sintaxi i la semàntica de l'oració i del discurs, l'ortografia, la fonologia i les convencions gràfiques). Tanca la proposta una aplicació de la «matriu classificatòria d'errors» a textos concrets que contenen errors de tipus diversos.

L'anàlisi detallada dels tipus de mancances que mostren els textos dels estudiants ha de permetre millorar l'estratègia educativa del professorat a l'hora d'afrontar l'ensenyament de les habilitats escrites.

#### **V. Desenvolupament**

##### **1. Objectius**

Els objectius del treball que presentem són els següents:

- a) Mostrar una experiència d'investigació en el marc de l'ensenyament de Filologia Catalana orientada a la millora de l'expressió escrita dels estudiants. L'experiència s'emmarca en un projecte de millora educativa basat en els pressupòsits del procés d'investigació-acció (Kemmis i McTaggart 1988; Elliott 1993). De les quatre fases que conformen aquest procés (a. *planificació*; b. *acció*; c. *observació*; d. *reflexió*), el que presentem forma part de la fase d'*observació* (c). En concret, mostrem el procediment que hem definit per tal detallar els tipus de mancances que apareixen en els textos dels estudiants, que

ens haurà de servir per poder realitzar d'una manera més precisa el procés de millora esmentat. Concretament, els resultats de l'aplicació d'aquest procediment ens hauran de ser útils en la fase següent, destinada a la *reflexió (d)* sobre l'estratègia educativa duta a terme i sobre com millorar la *planificació*.

- b) Descriure els paràmetres conceptuals en què es basa l'avaluació de l'expressió escrita en el projecte esmentat a *a*). Es tracta, bàsicament, de la interacció entre les propietats del text (adequació —que inclou la correcció—, coherència i cohesió) i les unitats susceptibles d'anàlisi que conformen el text (les quals pertanyen a la morfologia, el lèxic, la sintaxi i la semàntica de l'oració i del discurs, l'ortografia, la fonologia i les convencions gràfiques). A partir del creuament d'aquests paràmetres hem definit una «matriu o graella classificatòria d'errors» per a l'anàlisi dels textos.
- c) Mostrar l'aplicació computacional d'una «matriu classificatòria d'errors» en què es relaciona les propietats del text i les unitats susceptibles d'anàlisi a què hem fet al·lusió a *b*). El fet d'haver conceptualitzat els errors permet introduir-los com a etiqueta o marca al lloc pertinent del text que s'examini. Mitjançant programació informàtica, s'extreuen les parts concretes del text afectades per l'etiqueta i la informació s'exporta de manera que pot quedar constituïda una base de dades d'errors convenientment classificats.

## 2. Descripció del treball

### 2.1. Descripció global de l'experiència d'investigació-acció

El treball concret que aquí descrivim s'emmarca dins d'un procés global d'investigació-acció que duu a terme un equip de docents de l'assignatura de Llengua Catalana I del Departament de Filologia Catalana de la Universitat de Barcelona, el qual tracta sobre l'expressió escrita dels alumnes. Les assignatures relacionades amb aquesta matèria són les troncales Llengua Catalana I [= LC1], Llengua Catalana Estàndard [= LCE] i l'optativa Anàlisi i Producció de Textos Catalans [= APTC]. De les quatre fases d'aquest tipus de processos (segons la proposta de Kemmis i McTaggart 1988) — 1. planificació; 2. actuació; 3. observació; 4. reflexió—, el que aquí presentem correspon a la fase d'*observació*.

#### 2.1.1. Les dues primeres fases de la investigació-acció

Les dues primeres fases del procés d'investigació-acció —que són la *planificació* (1) i a l'*actuació* (2)— s'havien realitzat seguint les estratègies derivades de la pràctica docent anterior del professorat.

##### a) Fase 1

En la fase 1, corresponent a la *planificació*, s'havien establert els punts bàsics sobre els quals s'havia de centrar el procés d'aprenentatge de l'alumne. Aquests van ser els relatius a les propietats del text escrit: a) adequació; b) coherència; c) cohesió; d) correcció. Igualment, vam determinar que la pràctica docent havia d'orientar-se a afavorir l'autonomia de l'estudiant pel que fa a la redacció i revisió dels seus textos.

## b) Fase 2

En la fase 2 —pròpia de l'*actuació*—, es va realitzar el material docent adient que permetés a l'alumne progressar en els paràmetres indicats a l'apartat (a).

Quant a la pràctica docent —i centrant-nos ara en les dues assignatures obligatòries (LCI i LCE)—, es van realitzar les activitats següents:

- i) anàlisi de textos per part del professor com a mitjà per descriure i valorar en cada cas l'adequació, la coherència, la cohesió i la correcció.
- ii) explicació del professor sobre els mecanismes d'elaboració de textos.
- iii) confecció de textos per part de l'alumnat. A LCI es tractava d'un mínim de tres textos escrits en diferents moments del curs —per exemple, un text argumentatiu a l'inici del curs i dos textos (un d'expositiu i un altre d'argumentatiu) al llarg del curs—; a LCE eren un mínim de tres textos argumentatius. Així mateix, els alumnes, seguint les indicacions del professorat, elaboraven una segona versió dels textos (i, si convenia, una tercera versió) per tal de millorar-los.
- iv) realització d'exercicis destinats a afavorir l'autonomia de l'estudiant (per exemple, hi ha (a) aquells en què es mostren escrits amb problemes específics relatius a aspectes textuais diversos perquè l'estudiant els identifiqui, o (b) aquells en què es demana que l'estudiant corregeixi un text breu d'un altre estudiant, o bé (c) aquells en què s'organitza discussions grupals sobre les estratègies argumentatives emprades en textos diferents que tracten sobre un mateix tema).
- v) resolució d'exercicis de llengua normativa (és aquest un recurs que aborda d'una manera específica un aspecte que, d'altra banda, inclouen també altres dels punts aquí relacionats).
- vi) aprofundiment, a partir de recursos bibliogràfics, sobre temes problemàtics o de nivell avançat de normativa i exposició d'aquests temes per part dels alumnes a l'aula.

Els cinc primers apartats (i-v) fan referència a LCI i els apartats iii, v, i vi a LCE.

### 2.1.2. La tercera fase de la investigació-acció: l'*observació*

La tercera fase del procés d'investigació-acció, que és aquella en la qual centrem el nostre treball, correspon a l'*observació*. En aquesta fase es tracta de recollir dades de l'acció docent realitzada que ens puguin servir per a la darrera fase del procés (que es pot anomenar de *reflexió*, és a dir, aquella en què s'elaboraran les conclusions del pla d'acció docent dut a terme).

És evident que, com més precisa és la informació que es recull en la fase d'*observació*, més fina és l'anàlisi que es podrà fer posteriorment en el moment de valorar els resultats que ha obtingut l'estratègia educativa desenvolupada.

#### 2.1.2.1. Un estudi longitudinal com a font d'informació

Una de les fonts d'informació que vam preveure per a aquesta fase d'*observació* va ser realitzar un estudi longitudinal que permetés comparar la progressió en la competència en expressió escrita dels estudiants de l'ensenyament de Filologia Catalana (sobre estudis longitudinals d'evolució de l'expressió escrita en universitaris, vegeu, per exemple, Haswell (1991) i Sternglass (1997)).

En el nostre cas, l'estudi longitudinal fa referència als estudiants de la Facultat de Filologia des del moment en què arriben fins quatre anys després —que és el període de durada de la llicenciatura.

La cohort escollida va ser la promoció d'estudiants que va iniciar la llicenciatura l'any 2001 i que, segons el pla d'estudis, l'any 2003 podia finalitzar el primer cicle —de dos cursos acadèmics— i l'any 2005 podia finalitzar el segon cicle —de dos cursos acadèmics. Cal dir que, per a alguns alumnes, l'any 2003 ha resultat ser simplement el final del segon any de formació i, el 2005, el final del quart any de formació (és el cas dels alumnes que necessiten més de quatre anys per finalitzar la carrera).

Amb l'objectiu esmentat, es van recollir textos elaborats pels estudiants en els tres punts de l'eix temporal que acabem d'esmentar: l'any 2001, l'any 2003 i l'any 2005. Els alumnes eren de Filologia Catalana i d'altres Filologies (Anglesa, Hispànica i Llatina). Cal precisar que LC1 és una assignatura troncal, obligatòria per als alumnes de Filologia Catalana, Anglesa i Hispànica i opcional per als de Filologia Clàssica (que poden triar o bé català o bé espanyol). D'altra banda, LCE i APTC són assignatures específiques de la llicenciatura de Filologia Catalana (la primera obligatòria de primer cicle i, la segona, optativa de primer cicle).

Si prenem en consideració els tres punts temporals d'elaboració de textos per part dels estudiants (2001, 2003 i 2005), el nombre total d'aquests textos que actualment conformen el nostre estudi és de 150 (disposem de textos addicionals provinents d'alumnes que només van realitzar redaccions el 2001, o bé sols el 2001 i el 2003 —vegeu *infra* i nota 1). Podem també indicar que, en una primera fase, estem comparant els textos de tipus argumentatiu (relatius al tema del terrorisme) dels anys 2001 i 2005 de catorze alumnes de Filologia Catalana —és a dir, vint-i-vuit textos.

La major part dels escrits van ser realitzats a l'aula, sense cap material de consulta i en un període d'una hora i mitja. Quant a la tipologia textual, es tracta de textos argumentatius en la seva major part en el cas dels anys 2001 i 2003; els textos del 2003 són expositius o bé redaccions corresponents a exàmens. Pel que fa als textos argumentatius, el tema suggerit girava entorn del terrorisme i les mesures de seguretat que es van començar a aplicar a partir de l'atemptat a les torres bessones de Nova York el setembre de 2001. Se'ls convidava a redactar sobre el tema esmentat mitjançant un enunciat i alguns titulars de premsa que evidenciaven posicions diferents. Detallem a continuació, com a exemple, la capçalera corresponent als textos de l'any 2005.

DNI.....

L'atemptat terrorista a Madrid de l'11 de març de l'any passat va revifar la sensació d'inseguretat que molts ciutadans van començar a tenir arran dels atemptats contra els Estats Units de l'11 de setembre de 2001.

**¿Penses que les mesures de seguretat que els governs han anat prenent per prevenir atemptats terroristes poden reduir excessivament les llibertats individuals?**

Els titulars de premsa que tens en aquest full tenen relació amb la pregunta que et fem i et poden servir, si ho creus convenient, per orientar la teva argumentació.

Extensió del text: unes 250 paraules.

*Escorcolls policials a les estacions del ferrocarril.*

*Hores de cua als aeroports.*

*Alguns països es plantegen seguir l'exemple dels Estats Units i posar un agent de seguretat armat amagat entre els passatgers dels avions.*

*Es congelen els comptes de sospitosos de terrorisme.*

*S'incrementen els intercanvis d'informació entre els països occidentals.*

Pel que fa a la mostra d'estudiants, inicialment va ser de 133. Cal dir que la mortaldat — en termes estadístics— va ser important el 2003; el 2005, en canvi, la mostra es va mantenir molt similar a la de l'any 2003.<sup>1</sup> Les causes de la baixa d'informants assenyalada van ser, o bé que aquests havien efectuat un canvi d'estudis cap a una altra llicenciatura, o bé el fet que els estudiants podien acudir o no a la sessió de redacció de textos a què se'ls convocava.

Resta afegir-hi que es va demanar a l'estudiant informació relativa a la seva situació sociocultural i acadèmica; per obtenir-la, es va utilitzar un qüestionari de noranta preguntes; la finalitat de poder disposar d'aquesta informació era la possibilitat d'establir posteriorment correlacions entre aquestes dades i els resultats corresponents al nivell lingüístic de l'estudiant i al tipus d'errors realitzats. Igualment, s'ha d'indicar que, mitjançant el que vam anomenar «enquesta de procediment», es va preguntar a l'estudiant quins recursos de tipus compostiu havia utilitzat per a l'elaboració del seu text —per tal de conèixer el grau de consciència de l'estudiant sobre l'estratègia utilitzada a l'hora d'escriure el text.

#### 2.1.2.2. La categorització d'errors com a eina per a l'anàlisi dels textos

A l'hora d'examinar els textos dels estudiants amb la finalitat de determinar en cada cas el tipus d'error, vam constatar que, pel que fa a tres de les quatre propietats d'avaluació textual —vegeu el punt 2.1.1 a—, era insuficient limitar-nos a assignar als errors les etiquetes genèriques, com són les de «error relatiu a l'adequació», «error relatiu a la coherència» i «error relatiu a la cohesió». D'altra banda, hi havia la particularitat que les obres de referència relatives a avaluació textual no ens aportaven una subclassificació

<sup>1</sup> El 2003 hi va haver una baixa del 67% dels informants i el 2005 la baixa va ser del 5%. És per això que la mostra d'estudiants per a l'estudi longitudinal ha resultat ser finalment d'un nombre de vint-i-nou (que són aquelles que s'ha aconseguit que l'any 2003 realitzessin encara l'exercici de redacció), amb una producció de 150 textos —xifra que abans ja hem indicat.

tipològica d'errors prou detallada —a diferència del que s'esdevé per als errors relatius al quart paràmetre (*correcció*), per al qual és habitual trobar una subclassificació força completa dels diferents tipus d'errors.

Atès que el nombre d'errors que calia classificar era certament important —podia arribar a 3000 en cas de prendre en consideració els 150 textos i, en cas de limitar-nos als vint-i-vuit textos de la primera fase de l'estudi longitudinal a què hem fet referència, hi havia una estimació prèvia de 600 errors —, ens vam proposar d'elaborar una «matriu classificatòria d'errors» (la descrivim al punt 2.2).

Cal subratllar que la confecció d'aquesta matriu és útil no sols per als estudis longitudinals que estem realitzant, sinó també per analitzar tot tipus de textos.

## 2.2. Descripció de la «matriu classificatòria d'errors» i exemplificació

### 2.2.1. Les dues primeres columnes de la matriu

A la «matriu classificatòria d'errors» es fa un creuament entre les propietats textuais (adequació —amb la correcció—, coherència i cohesió) —que ocupen la primera columna— i el que anomenem «nivell lingüístic» en què es localitza l'error, és a dir, les unitats susceptibles d'anàlisi que conformen el text (que estan connectades amb la morfologia, el lèxic, la sintaxi i la semàntica de l'oració i del discurs, l'ortografia, la fonologia i les convencions gràfiques).

#### 2.2.1.1. Les propietats textuais

##### a) Adequació

L'adequació és la propietat del text que determina les varietats lingüístiques (geogràfica, diastràtica i funcional) que cal usar segons la situació comunicativa en què s'insereix el text (cf. Gregory i Carroll 1978). En la proposta que presentem, l'adequació es distribueix en quatre apartats: convencions tipogràfiques, aspectes lèxics, aspectes oracionals i aspectes discursius o interoracionals.

##### b) Coherència

La coherència és la propietat del text per la qual aquest s'elabora (de la banda de la producció) i s'interpreta (de la banda de la recepció) com una construcció articulada d'unitats d'informació. Així doncs, dóna compte del significat global del text i explica la selecció de la informació (rellevant i irrellevant) i la seva organització en una estructura determinada (cf. Beaugrande 1991). En la proposta que presentem, la coherència es distribueix en tres apartats: aspectes morfològics i lèxics, aspectes oracionals, i aspectes discursius o interoracionals.

##### c) Cohesió

La cohesió és la propietat del text per la qual aquest està dotat de *teixit*. Reflecteix les relacions sintacticosemàntiques que s'estableixen entre els

elements del text: sintagmes, oracions i paràgrafs. Els recursos que estan al servei de la cohesió són la relació lèxica, l'anàfora, la definització i la connexió (cf. Simone 1990). En la proposta que presentem, la cohesió es distribueix en dos apartats: aspectes oracionals i aspectes discursius o interoracionals.

#### d) Correcció

Entenem per correcció d'un text el fet que aquest s'adeqüi al corpus normatiu del català aprovat com a tal. Cal assenyalar que, en sentit estricte, la correcció és un dels aspectes de l'adequació però que, per necessitats operatives i d'anàlisi (i també pel fet que té un pes específic propi), hem cregut necessari d'aïllar-la de la resta de fenòmens d'adequació. El corpus normatiu del català aprovat com a tal el conformen les obres següents:

- *Gramàtica catalana* de Pompeu Fabra (1918)
- *Diccionari de la llengua catalana* de l'Institut d'Estudis Catalans (1995)
- Documents de l'IEC declarats explícitament com a normatius.

S'ha de remarcar, tanmateix, que aquest corpus no preveu totes les possibilitats d'errors de correcció; és per això que també hem fet ús del marc de referència més ampli constituït pel que anomenem «llengua estàndard». Les formes estàndard són «les formes usuals en el llenguatge general de la comunitat acceptades majoritàriament pels professionals de la llengua (independentment que constin o no als codis normatius)» (Solà 2000: 45).

#### 2.2.1.2. El «nivell lingüístic» de l'error

A l'inici del punt 2.2 ja ens hem referit al que anomenem «nivell lingüístic» de l'error. Aquest nivell ens indica a quina de les unitats susceptibles d'anàlisi lingüística —i, en algun cas, merament gràfica— de què es compon el text pertany l'error. Ja hem assenyalat també que aquestes unitats estan relacionades amb la morfologia, el lèxic, la sintaxi i la semàntica de l'oració i del discurs, l'ortografia, la fonologia i les convencions gràfiques.

Fruit del creuament entre les propietats textuais i el nivell lingüístic en què es realitza l'error és la segona columna de la matriu. A partir dels errors reals obtinguts en els textos examinats fins al moment, tenim que les unitats susceptibles d'anàlisi han resultat ser les següents: per a l'*adequació*, les convencions gràfiques, el lèxic, l'oració i el discurs; per a la *coherència* i la *cohesió*, l'oració i el discurs en ambdós casos; i, per a la *correcció*, l'ortografia, la morfologia, el lèxic, l'oració i el discurs.

#### 2.2. L'etiquetatge dels errors (columnes 3 i 4 de la matriu)

Les dues altres columnes de la matriu es destinen a encabir les etiquetes concretes de cadascun dels tipus d'errors que s'han pogut aïllar i conceptualitzar a partir de l'examen dels textos. A la tercera columna es troben les etiquetes corresponents als errors que, dins del text, tenen un abast local —és a dir, fan referència a un segment o fragment del

text—, i a la quarta columna es troben les etiquetes corresponents als errors de tipus global, és a dir, aquells que fan referència al text en conjunt.

La «matriu classificatòria d'errors» es troba en una fase avançada d'elaboració. Actualment conté noranta etiquetes, cadascuna de les quals ocupa una fila, sigui a la columna 3 o a la columna 4. La informació que, a les columnes 1 i 2, es correspon amb la fila de l'etiqueta de les columnes 3 o 4 ens assenyala quines propietats del text (columna 1) i quines unitats d'anàlisi (columna 2) corresponen a aquell tipus d'error.

S'assigna a cada etiqueta una sigla, que és la mateixa que s'utilitza a l'hora de fer el marcatge dels textos en el procés d'aplicació computacional de la matriu.

De les noranta etiquetes a què hem fet referència, aquí en presentem una mostra de vint-i-vuit, que són les que es troben en els dos textos transcrits.

Per a la lectura de la taula, cal dir que una casella en blanc significa que, malgrat que en els textos que aquí presentem no hi correspongui cap etiqueta, en altres textos s'han trobat errors del tipus corresponent; les caselles grises són espais de classificació per als quals els textos analitzats fins al moment no ens han indicat possibilitat d'etiquetatge.

<i>Matriu classificatòria d'errors</i>			
<b>Propietats del text</b>	<b>Elements de l'estructura lingüística (= Nivell lingüístic de l'error)</b>	<b>Trets específics</b>	<b>Trets globals</b>
Adequació	Convencions tipogràfiques	<b>APFP</b> = puntuació fi de paràgraf. <b>ATMA</b> = majúscula /minúscula.	
	Lèxic	<b>ALGR</b> = grau d'especificitat dels termes inadequat al gènere.	<b>ALRE</b> = repertori poc variat.
	Oració	<b>AOCN</b> = construcció no adequada al registre.	
	Discurs		
Coherència	Oració	<b>ROCN</b> = connector intraoracional usat en un sentit que no es correspon amb la relació semàntica entre les parts de l'oració que connecta. <b>ROTA</b> = tautologia.	



	Discurs	<p><b>RDFAS</b> = la falta d'informació afecta la interpretació del segment.</p> <p><b>RDII</b> = informació incoherent en la progressió temàtica.</p> <p><b>RDIN</b> = informació irrellevant en relació amb el conjunt.</p> <p><b>RDRE</b> = referent no identificable.</p> <p><b>ROCO</b> = col·locació semànticament anòmala.</p> <p><b>ROJC</b> = juxtaposició o coordinació incoherent.</p>	<b>RDLT</b> = la línia temàtica no és clara.
Cohesió	Oració	<p><b>SOPH</b> = puntuació intraoracional incorrecta.</p> <p><b>SOPN</b> = correlació incorrecta (excepte la temporal).</p>	
	Discurs	<p><b>SDERL</b> = element referencial llunyà.</p> <p><b>SDPI</b> = puntuació interoracional incorrecta.</p>	
Correcció	Ortografia	<p><b>CGAC</b> = accentuació.</p> <p><b>CGAP</b> = apòstrof.</p> <p><b>CGCO</b> = consonants.</p>	
	Fonologia	<b>CLCA(FON)</b> = calc (fonològic).	
	Morfologia i Lèxic	<p><b>CLCA</b> = calc.</p> <p><b>CLSI</b> = mot o mots catalans usats en un sentit incorrecte.</p> <p><b>COPI</b> = puntuació intraoracional incorrecta.</p>	
	Oració	<p><b>COCC</b> = concordança (errors d'excés i de defecte).</p> <p><b>CONI</b> = construcció incorrecta que no és concurrència.</p>	
	Discurs	<b>COPP</b> = altres problemes en l'ús de les proposicions.	

## Mostra de l'etiquetatge de textos assistit per ordinador

### Text 1 (any 2001)

```
<?xml version="1.0" encoding="xxx-xxx-x"?>
<!DOCTYPE document SYSTEM "jerarquia.dtd">
<document>
  <capçalera>
    <idpersona>xxxxxxxx</idpersona>
    <subcorpus>cc</subcorpus>
    <any>2001</any>
    <lloc>aula</lloc>
    <tema>terrorisme</tema>
  </capçalera>
  <text alre="2">
  <text rdlt>
```

<cgac>Després</cgac> dels atemptats terroristes que van tenir lloc als Estats Units el dia 11 de <atma>Setembre</atma>, <cgap>l'</cgap>inseguretat de la població està en creixement</coni> <rocn>envers</rocn> la seguretat als aeroports. A partir d'aquest succés les mesures que <clca>garantitzen</clca> <coni>la confiança</coni> dels passatgers s'estan modificant i millorant<sdpi>,</sdpi> <coni>trobem per exemple la presència</coni> d'un agent de seguretat armat entre els passatgers com a mesura més <clsi>precària</clsi> i <copi>fins i tot,</copi> els satèl·lits que <rdin>Rússia</rdin> llançarà per <rota>espiar i vigilar</rota> l'Afganistan com a <cgac>mètode</cgac> més dràstic.

<coni>Podem comentar</coni> que encara que hi <cooc>hagin</cooc> tants mètodes i mesures de seguretat un atemptat tan preparat <cgac>seria</cgac> molt difícil de <clsi>detectar</clsi>. S'ha dit que <rdre>aquestes persones</rdre> no portaven ni tan sols armes de foc <sdpi>,</sdpi> les eines que van utilitzar <copi>doncs</copi> són estris que pot portar una persona "normal" a la seva bossa de viatge.

D'altra banda, <cgap>sí</cgap> <cgac>hauria</cgac> de mirar la documentació dels passatgers<sopi>, cosa que no es fa de vegades als aeroports,</sopi> ja que pot ser que el <cgco>bitllet</cgco> hagi estat robat.

<sderl>Aquest atemptat</sderl> ha estat una obra preparada durant molt de temps, anys i tot, i la persona que ha estat capaç de fer <algr>això</algr> ha de ser intel·ligent i perversa, podem dir que <clca>gairebé segur</clca> no va passar cap detall per alt.

Per concloure podem dir que és una pena que ja al segle XXI <cooc>hi hagin</cooc> tantes rivalitats entre nacions, ja que <sderl>aquest atemptat</sderl> no <cgac>hauria</cgac> estat possible en un món de comprensió i respecte <copp>envers al</copp> <clca>prójim</clca>.

</document>

## Mostra de l'etiquetatge de textos assistit per ordinador (*cont.*)

### Text 2 (any 2001)

```
<?xml version="1.0" encoding="xxx-xxx-x"?>
<!DOCTYPE document SYSTEM "jerarquia.dtd">
<document>
  <capçalera>
    <idpersona>xxxxxxxx</idpersona>
    <subcorpus>cc</subcorpus>
    <any>2001</any>
    <lloc>aula</lloc>
    <tema>terrorisme</tema>
  </capçalera>
  <text alre="0">
```

És cert que els atemptats terroristes contra els Estats Units han fet créixer la sensació d'inseguretat als països occidentals. Penso que <coni>la raó principal</coni> és la magnitud d'aquests atemptats<sopi></sopi> perquè és gairebé inimaginable. El fet que <rdfas>una sola persona</rdfas> hagi pogut <cgco>movilitzar-ne</cgco> tantes i aconseguir fer tant de mal desborda qualsevol ment humana. Per tant, s'ha demostrat que qualsevol persona pot ser capaç de fer trontollar l'equilibri mundial si compta amb els recursos necessaris.<sderl> Aquesta inseguretat</sderl> derivada dels atemptats, ha fet prendre mesures molt estrictes<sopi>, al meu parer, fins i tot massa</sopi>.

És <rdii>excessiu i perillós</rdii> portar un agent de seguretat armat dins d'un avió, <copi>sobretot perquè veient el que ja han aconseguit fer</copi>, aquest agent es podria tornar un perill per la força que dóna tenir una arma entre les mans. Si s'ha de portar un agent camuflat entre els passatgers, <aocn>que vagi sense armes</aocn> (si més no, armes de foc), perquè hi ha altres maneres d'aturar segrestadors i terroristes, <rdii>sobretot perquè tampoc pot ser que ells vagin armats</rdii>.

La mesura que trobo encertada i confortant és la <sopn>d'augmentar els controls dels aeroports, encara que s'hagi de fer hores de cua</sopn>. Si aquestes hores fan que l'avió pugui ser un transport més segur i eficaç, és <clsi>un bon mètode que cal dur a terme</clsi><apfp></apfp>

Per altra banda, es parla de les fonts monetàries que han servit als terroristes per a la seva formació de pilots i per realitzar aquests atemptats. No serveix de gaire que ara s'investiguin els moviments de les empreses de la borsa<sopi></sopi> perquè el mal ja està fet. En aquest punt es mostra la falta de vigilància i de seguretat dels nostres diners i <rojcd>de</rojcd> les <roco>repercussions que això pot comportar</roco> <copp>a</copp> les vides humanes.

Per últim, penso que els Estats Units vol retornar la plena seguretat als seus ciutadans i això fa que totes les mesures siguin poques per a ells. La realitat és que mai s'estarà segur del tot en aquest món. </text>  
</document>

### 2.3. Implementació computacional

Volíem que l'aplicació de la «matriu classificatòria d'errors» als textos pogués fer-se de manera tan automatitzada com fos possible. En aquest punt, hem tingut el suport del Servei de Tecnologia Lingüística de la Facultat de Filologia de la Universitat de Barcelona.

Amb l'objectiu esmentat, els textos —que inicialment van ser escrits a mà, per la raó que es tractava d'un exercici a l'aula, i que conservem en arxius d'imatge— van ser introduïts en un processador de textos convencional i, posteriorment, captats dins del que es coneix com a *editor de text*.

Un *editor de text* és un programa que permet escriure i modificar arxius digitals que només contenen text; es coneixen també com a arxius de text pla. El tret diferencial entre l'editor de text i el processador de text és que el primer s'usa només per escriure text sense format (en alguns editors, els formats es poden marcar amb codis especials de

text). D'editors de text n'hi ha de moltes menes (d'ús general, dissenyats per escriure o programar en un llenguatge, senzills, amb una gran quantitat de funcions, etc.). Exemples d'editors de textos són: *Bloc de notes* (editor integrat a Windows, que també es coneix com a *Notepad*), *Emacs* (editor molt comú a Unix), *jEdit* (editor multiplataforma) *Vi* (editor d'ús comú a Unix).

L'editor de text Emacs <http://www.gnu.org/software/emacs/emacs.html#Whatis#Whatis> <http://es.gnu.org/cgi-bin/emacs-es.pl> <http://es.wikipedia.org/wiki/Emacs>) és un editor de text que va ser creat per Richard Stallman i que està distribuït sota la llicència lliure GPL; actualment el manté la Free Software Foundation i forma part del projecte GNU. Les funcions principals de l'editor Emacs són: 1) escriure text pla (text conegut com a txt, MS2 o ASCII –American Standard Code for Information Interchange- de 7 bits), 2) programar (determinar un conjunt concret d'instruccions que pot executar una computadora) o 3) escriure scripts (guió computacional). Aquestes funcions es desenvolupen per mitjà del llenguatge elisp (Emacs Lisp). Una versió millorada de l'editor Emacs destinada a treballar de manera integrada a un sistema operatiu basat en un entorn de finestres (tipus Windows) és XEmacs.

Per editar/corregir un text en XEmacs cal obrir un fitxer amb l'extensió \*.txt que estigui prèviament guardat dins d'una carpeta en què hi ha d'haver també un altre fitxer amb l'extensió \*.dtd —aquest darrer fitxer conté la informació de les marques jerarquitzaades que s'han d'introduir ( propietats del text, abast de l'error i nivell de l'error).

Posteriorment, el fitxer amb extensió \*.txt s'ha de gravar com a \*.xml. Un cop activat el programa de marcatge, s'introdueixen a la capçalera les marques que permeten identificar el text i qui l'ha escrit. Seguidament, un menú desplegable que, a cada part concreta del text, conté en forma de codi les etiquetes d'error que en aquell punt poden necessitar-se, permet que aquestes s'introdueixin de manera semiautomatitzada davant i darrere de l'error després que l'usuari hagi seleccionat el fragment en què aquest es troba i l'etiqueta d'error que li correspon.

Es pot indicar que el format de dades de text pla sobre el qual s'introdueixen els errors en XEmacs és XML, que conté marques estructurades, però, a diferència de l'HTML, no té un conjunt predefinit de marques sinó que cal assignar-les-hi. Algunes característiques formals de l'XML són: 1) les marques són diferents si s'escriuen en majúscules o minúscules; 2) totes les marques han d'aparèixer per parelles i tots els elements han de tenir una marca d'obertura (<element>) i una marca de tancament (</element>); vegeu, per a més detall, [http://tools.arlut.utexas.edu/gash2/doc/xml/introt\\_oxml.html](http://tools.arlut.utexas.edu/gash2/doc/xml/introt_oxml.html).

Cal assenyalar, finalment, que s'ha realitzat també la programació adequada per tal que tota la informació relativa a cada error pugui ser emmagatzemada en una base de dades convencional, externa al programa XEmacs.

### 3. Resultats i conclusions

a) L'estudi longitudinal (2001-2003-2005) que vam preveure realitzar per tal d'avaluar la millora de l'expressió escrita dels estudiants de Filologia ens permet disposar de 150 textos escrits, en la seva major part argumentatius i que poden encadenar-se en el temps.

b) Per tal de poder extreure una informació precisa del tipus d'errors dels textos i detallar més del que se sol fer quan s'avaluen, concretament, tres de les propietats textuais (adequació, coherència i cohesió), s'ha elaborat una matriu classificatòria d'errors. Actualment aquesta matriu es compon de noranta etiquetes.

La posició que ocupen les etiquetes dins la graella es correspon amb la propietat textual a què l'error fa referència (adequació, coherència, cohesió i correcció), l'àmbit d'anàlisi (global o local) i el nivell d'anàlisi (ens referim a les unitats lingüístiques relacionades amb la morfologia, el lèxic, la sintaxi i la semàntica de l'oració i del discurs, l'ortografia i la fonologia, així com a les convencions gràfiques).

c) Hem realitzat, d'altra banda, una aplicació de la graella als textos mitjançant un procediment computacional que se serveix de les possibilitats de l'editor X-Emacs.

El nostre projecte ha servit, així mateix, per desenvolupar una eina d'etiquetatge semiautomatitzat que implementa l'editor XEmacs, així com un programa que permet emmagatzemar de manera automatitzada la informació corresponent a cada error (el seu entorn oracional i, mitjançant l'etiqueta, l'especificació de les propietats textuais, àmbit d'anàlisi i nivell d'anàlisi) en una base de dades externa apta per a estudis de tipus divers.

d) Indiquem també que la matriu classificatòria d'errors que hem elaborat és una eina que permet sistematitzar la feina del corrector. A més, pot ser d'ús general per a tot tipus de textos.

e) Hem d'assenyalar igualment que la confecció de la matriu esmentada demana un treball considerable per tal de clarificar i delimitar en cada cas conceptes de tipus textual o de nivell lingüístic. Un cop elaborada, és una mitjà útil per a l'observació detallada dels problemes relatius a la competència en l'expressió escrita —que és el nostre objectiu immediat. Una altra possibilitat seria utilitzar-la com a instrument quotidià d'avaluació de l'aprenentatge i d'interacció amb l'alumne. De moment no hem explotat aquesta darrera via; l'avantatge d'emprar-la amb aquest objectiu seria el detall en la concreció de l'error; la dificultat d'aplicació estaria relacionada sobretot amb el fet que alentiria el treball de correcció de textos i que els alumnes haurien de tenir alguns coneixements previs per poder-la interpretar.

#### 4. Bibliografia

BEAUGRANDE, R. de (1991) «Semantics and text meaning: retrospects and prospects», *Journal of Semantics* 5, p. 89-121.

ELLIOTT, J. (1993) *El cambio educativo desde la investigación*, Madrid, Morata.

GREGORY, M., i CARROLL, S. (1978) *Language and situation. Language varieties and their social contexts*, Londres, Routledge & Kegan Paul Ltd.

HASWELL, R. H. (1991) «Documenting Improvement in College Writing. A Longitudinal Approach», *Written Communication* 17, 3, p. 307-352.

KEMMIS, S., i MCTAGGART, R. (1988) *Cómo planificar la investigación-acción*. Barcelona, Laertes.

MILIAN, M. (1997) «Avaluar els textos escrits», dins Teresa Ribas (coord.) *L'avaluació formativa en l'àrea de llengua*, Barcelona, Graó, p. 109-129.

STERNGLASS, M. S. (1997), *Time to Know Them. A Longitudinal Study of Writing and Learning at the College Level*. Laurence Erlbaum, 1997. Cap. 2 i 3.

SOLÀ, J. (2000) «Reflexió breu sobre el concepte de normativa», dins Macià, J., i Solà, J. (ed.) *La terminologia lingüística en l'ensenyament secundari*, Barcelona, Graó, p. 39-46.