

El projecte de constitució de recursos lingüístics per a les escoles. Plantejaments inicials

Toni Badia

Professor de la Universitat Pompeu Fabra i membre del CREL

En aquesta ocasió em pertoca presentar-vos el projecte *Constitució de recursos lingüístics per a la xarxa telemàtica educativa*, que l'Institut d'Estudis Catalans ha posat en marxa dins del projecte més ampli *El català en la societat de la informació*, dirigit per M. Teresa Cabré. El projecte es troba actualment en els seus estadis inicials, de manera que en aquesta presentació no podré sinó referir-me als seus plantejaments inicials.

El projecte té com a objectiu fonamental crear una xarxa lingüística per al sistema educatiu català que es vehicularà a través d'Internet, aprofitant la infraestructura proporcionada pel projecte Educàlia. Aquesta xarxa lingüística ha de fer possible en primer lloc que es posin a disposició de les escoles i els instituts de Catalunya els materials i recursos de què disposa la Secció Filològica de l'IEC, de manera que els mestres i professors els puguin usar directament, sigui per a les seves classes, sigui per a la seva formació i preparació. A la vegada, la xarxa ha de permetre que els mestres i professors de Catalunya interaccionin amb l'IEC, per tal de disposar d'informació permanentment actualitzada sobre la normativa, de poder fer consultes lingüístiques i de plantejar les seves necessitats i neguits a l'IEC. En el seu conjunt, doncs, aquest projecte pretén reforçar l'ensenyament de la llengua catalana i en llengua catalana, i en general la labor educativa dels mestres. A la vegada, i d'una manera indirecta, el projecte contribuirà a reconèixer el valor social de l'ensenyament, en fer palès que els recursos i materials de la Secció Filològica, convenientment adaptats, són també útils per a finalitats didàctiques, per a les quals, és clar, no havien estat pensats inicialment. Finalment, aquesta xarxa lingüística contribuirà també a prestigiar la llengua catalana davant dels nens i nenes del nostre país: l'accés a informació a través d'Internet, així com els exercicis i materials de consulta respecte de la llengua catalana que s'hi poden trobar, faran que l'aurèola de prestigi que tenen per als nens les noves tecnologies cobreixi també una mica la llengua catalana.

S'ha previst, naturalment, que el projecte es desenvolupi en diverses fases, que resumim a continuació. En primer lloc, es du a terme un estudi previ sobre les necessitats reals de les escoles catalanes i sobre la manera com aquestes necessitats poden ser cobertes per recursos existents a l'IEC. A continuació caldrà fer el disseny tècnic dels mòduls de què es compondrà la

xarxa: a partir de les necessitats detectades i dels recursos assenyalats s'haurà de preveure l'estructura tècnica dels diversos mòduls i de l'enllaç entre les bases de dades de la Secció Filològica i la interfície amb els usuaris escolars. La tercera fase consistirà en la realització efectiva dels mòduls; serà fonamentalment, per tant, una etapa de programació informàtica. El resultat d'aquesta fase de programació haurà de ser posat a prova i avaluat, abans de generalitzar-ne la implantació en les escoles del país. És clar que les tres darreres etapes (realització dels mòduls, experimentació i avaluació i implantació general) es podran dur a terme de manera esglaonada: no tots els mòduls hauran d'estar acabats al mateix temps, ni hauran de ser avaluats a l'hora. Tenint en compte, doncs, que l'estudi previ estarà acabat a finals de 1999, hom pot preveure que les primeres proves experimentals es deuran poder fer durant el curs 2000-2001.

En el que resta d'espai per a aquesta presentació em centraré en l'estudi previ que estem realitzant actualment. Per a aquest estudi calia comptar amb mestres que tinguessin un coneixement complet de la realitat de les escoles catalanes. Així doncs, coordinat per qui signa aquest escrit, l'equip que col·labora en aquesta etapa està format per quatre mestres:

- una mestra d'una escola del Raval (Barcelona), especialista en educació infantil (de tres a sis anys) i cicle inicial de primària (de sis a vuit anys);

- una mestra de l'escola de Santpedor (Bages), especialista en cicle mitjà de primària (de vuit a deu anys);

- una mestra d'un institut de Manresa (Bages), especialista en cicle superior de primària (de deu a dotze anys) i educació secundària obligatòria (de dotze a setze anys);

- una mestra de suport del Servei d'Ensenyament del Català que treballa a diverses escoles de Barcelona.

Com es pot veure, s'ha procurat que l'equip constituït cobreixi tant la diversitat sociològica de les escoles de Catalunya com els diferents nivells de l'ensenyament obligatori. Així doncs, entre les quatre, aquestes mestres aporten una visió força representativa de les diverses situacions en què es desenvolupa l'ensenyament obligatori en el nostre país.

Els objectius de l'estudi previ són fonamentalment dos: analitzar les necessitats bàsiques de les escoles respecte de la llengua catalana i estudiar els projectes de la Secció Filològica per veure com poden cobrir, parcialment, és clar, aquestes necessitats. Es tracta, per tant, de saber exactament què convé fer per tal que la xarxa resultant del projecte sigui realment útil per a les escoles. Noteu que amb l'anàlisi de necessitats no n'hi hauria prou. Cal, també, analitzar els materials de la Secció Filològica; la raó és molt senzilla: aquests materials i recursos no han estat pensats amb cap finalitat didàctica, sinó estrictament per a la recerca. No és evident, doncs, que puguin servir per a cobrir necessitats de les escoles de primària i secundària; cal veure-ho, cal analitzar què pot servir i què no. Així doncs, en acabar l'estudi previ, disposarem d'una anàlisi detallada de cada un dels grans projectes de la Secció Filològica que comporten la constitució de recursos lingüístics: el *Diccionari de la llengua catalana* i altres materials de les Oficines Lexicogràfiques, el Corpus Textual Informatitzat de la Llengua Catalana, l'Atlas Lingüístic del Domini Català, l'Oficina d'Onomàstica i la Gramàtica Catalana. D'aquests projectes, però, els que hauran

estat analitzats més a fons seran els dos primers (els materials de les Oficines Lexicogràfiques i el Corpus Textual), perquè són els que aporten més materials útils per a finalitats didàctiques. Els altres, sigui per la limitació dels materials que recopilen, sigui per l'estat incipient del projecte, no proporcionen gaires recursos que puguin ser usats ara mateix a les escoles.

En aquests moments, estem en disposició de presentar un avanç dels resultats, sobretot pel que fa a l'anàlisi de necessitats i als recursos que es poden derivar dels dos projectes majors que hem mencionat. L'informe sobre les necessitats de les escoles catalanes detecta que la llengua catalana es troba en una situació relativament precària, a pesar de la generalització del català com a llengua vehicular de l'ensenyament. Hem constatat tres factors bàsics per a aquesta relativa precarietat: la situació sociolingüística de les escoles, la formació dels mestres i els materials didàctics.

La realitat sociolingüística de les escoles és, naturalment, el reflex de la situació general de la societat; i el que apareix amb molta claredat és que hi ha, cada cop més, una gran varietat lingüística, que afecta, cada vegada més, més escoles. La llengua familiar dels nens d'una escola no és, en general, uniforme, i el grau de coneixement de la llengua catalana varia molt. A més, en anar-se estenent la immigració procedent de fora de l'Estat, es produeixen situacions noves: nens que ni tenen el castellà com a llengua materna o que, fins i tot, desconeixen totalment tant el català com el castellà.

En general, no és fàcil per als mestres adequar els coneixements adquirits en la seva formació a les necessitats específiques de cada escola i de cada col·lectiu de nens. La formació inicial, bàsicament teòrica, no és suficient per a respondre a les noves necessitats. A més, des de l'aplicació de la reforma, a les escoles ja no hi ha mestres especialistes en llengua (i, per tant, en llengua catalana): actualment, els únics mestres especialistes ho són en gimnàstica, música o idioma estranger, mentre que la resta són tots «generalistes», és a dir, mestres preparats per a cobrir totes les facetes de l'ensenyament primari. Això comporta, necessàriament, una dificultat a l'hora de respondre a noves situacions lingüístiques. La resposta general que s'està donant a aquests problemes consisteix a promoure la formació a l'escola, de manera que resulti adequada a les necessitats reals de cada lloc. Això, però, té el problema que alenteix el procés de formació, i que normalment hi ha assumptes més urgents que passen davant de les qüestions plantejades sobre la llengua.

Els materials didàctics, finalment, de què es disposa a l'escola (tant els llibres com els materials complementaris en format electrònic), no solen tenir en compte la diversitat lingüística que pot haver-hi a les aules. Per una banda, no estan pensats per a ser usats en entorns amb una varietat gran de nivells lingüístics; i, per l'altra, no contempen situacions extremes, però cada cop més usuals, que comentàvem abans.

En detallar les necessitats actuals de les escoles, ens hem fixat en els aspectes que ens han semblat més significatius:

- els alumnes, que mostren la varietat de situacions lingüístiques que acabem de comentar;
- la llengua oral, que és (ha de ser) el vehicle de comunicació a l'escola, funció que, a vegades, es veu dificultada per la manca de fluïdesa i naturalitat del mestre;

— la formació del professorat, que massa vegades no és prou adequada perquè pugui donar resposta a les situacions noves amb què es troba;

— l'autoaprenentatge, per tal com és una de les respostes didàctiques més adequades a la varietat de situacions d'aprenentatge en una mateixa aula i, a més, es veu clarament afavorit amb les noves tecnologies, i

— la relació de la llengua amb les altres matèries, ja que només en la mesura que el català és la llengua vehicular de tot l'ensenyament es pot garantir l'adopció d'hàbits lingüístics que suposin algun ús del català.

Quant als projectes de l'IEC, actualment hem analitzat ja amb un cert detall els dos que preveiem que oferiran més possibilitats amb vista al nostre projecte: els materials de les Oficines Lexicogràfiques i el Corpus Textual Informatitzat. En tots dos casos, la metodologia que hem emprat ha estat la mateixa. En primer lloc, hem analitzat la informació continguda a les bases de dades de l'IEC, per tal de saber quina informació ens podien proporcionar que resultés útil per a les escoles. En aquesta observació no ens hem centrat només en el format aparent de la informació, sinó que hem tingut en compte l'estructura i el contingut interns de la base de dades. La diferència entre les dues aproximacions és molt evident en el cas del diccionari. La base de dades amb què treballen a les Oficines Lexicogràfiques té molta més informació de la que surt publicada en el diccionari imprès; en ser el nostre un projecte que ha d'aprofitar els recursos en format electrònic, la informació realment rellevant és la que apareix a la base de dades, independentment de si apareix també en el diccionari imprès.

En segon lloc, i a partir d'aquesta informació, hem estudiat en quines activitats escolars podria ser usada. Així, hem anat examinant per a tots els nivells de l'ensenyament obligatori quins dels aspectes de la formació lingüística que s'hi preveuen podrien ser reforçats amb aquest tipus d'informació, i hem proposat models d'exercicis que la podrien aprofitar. El resultat d'aquesta anàlisi és un llistat de classes d'exercicis segons la temàtica en què es basen; aquest llistat, a més, està classificat per nivells d'ensenyament, segons la gradació següent:

— exercicis que poden ser usats a partir de l'educació infantil i el cicle inicial de primària;

— exercicis que poden ser usats a partir del cicle mitjà de primària;

— exercicis que poden ser usats a partir del cicle superior de primària;

— exercicis que poden ser usats a l'ensenyament secundari obligatori i pels mestres per a la seva formació o per a la preparació de les classes.

Finalment, hem identificat a les bases de dades de l'IEC la informació que pot ser útil en cada cas. D'aquesta manera, disposem d'una primera aproximació de quines coses poden ser aprofitades de cada base de dades i amb quina finalitat.

Respecte dels materials de les Oficines Lexicogràfiques, hem observat que podien ser útils en dues operacions bàsiques: la classificació i la cerca de mots. En molts casos, no obstant això, el corpus lexicogràfic a partir del qual s'hauria de fer la tasca de classificació o de cerca no pot ser el de tota la base de dades lèxica, sinó el d'un subconjunt d'aquesta. Hi ha una diferència notable entre el vocabulari rellevant per a les edats escolars i el vocabulari general recollit en aques-

ta base de dades; per això, si volem que el recurs sigui realment útil, no podem usar la base de dades general, sinó que s'haurà de partir d'un subconjunt d'aquesta, escollit amb criteris didàctics. Amb la finalitat de mostrar efectivament el tipus de vocabulari que és convenient de tenir en compte, hem elaborat també unes taules de mots classificats per temes, d'acord amb el que és habitual en els exercicis de vocabulari de cada nivell de l'ensenyament.

Les activitats de classificació i cerca previstes hauran de poder-se fer a partir de diversos criteris, que contemplen aspectes gramaticals, de forma i de significat. Així, s'hauran de poder classificar i cercar mots a partir de:

- les seves propietats ortogràfiques i de partició sil·làbica;
- la seva categoria gramatical;
- les seves propietats de flexió;
- els processos de formació de mots a què estan sotmesos;
- la seva definició;
- les relacions semàntiques codificades (sinonímia i antonímia).

Respecte del Corpus Textual Informatitzat de la Llengua Catalana, hem vist que hi havia dos tipus d'informació que podien ser útils per a les escoles: la selecció de textos segons algunes de les seves característiques i l'observació de mots en el seu context. Aquí també hem vist que habitualment convindria partir d'un subconjunt dels materials continguts en el Corpus. Per una banda, els textos inclosos a la base de dades textuals són massa nombrosos per a poder-hi fer cerques amb finalitats didàctiques; i, per l'altra, hi ha textos d'èpoques massa allunyades perquè puguin representar un model de llengua en les activitats escolars. Així doncs, creiem que caldrà partir d'una selecció prèvia dels materials del Corpus, efectuada amb criteris didàctics.

Per a l'obtenció de textos representatius, preveiem que es podran utilitzar diversos criteris, com ara la distinció (bàsica en l'organització del Corpus) entre textos literaris i no literaris, la classificació en subgèneres (de cada un dels dos grans blocs que acabem de mencionar), la classificació per l'època de producció del text, o fins i tot el nom de l'autor del text. En aquest ús del Corpus Textual, la finalitat bàsica serà l'obtenció de textos per a l'observació i l'anàlisi (i la posterior producció de textos similars).

La segona utilització dels materials del Corpus Textual serà l'observació de les paraules en el seu context. Així, si amb el diccionari es poden extreure les propietats intrínseques de les paraules, a partir del Corpus Textual es pot observar com aquestes propietats repercuteixen en el seu ús. Com a conseqüència d'això, preveiem que els criteris de cerca seran pràcticament els mateixos que per al diccionari, en la mesura que formin part del marcatge del Corpus Textual, és a dir, seqüències de formes, lemes i etiquetes gramaticals per a observar:

- la capacitat de combinació de cada una de les categories bàsiques;
- l'estructura dels diversos sintagmes;
- l'estructura de l'oració simple;
- les oracions compostes;
- l'estructura discursiva del text.

Ara mateix estem acabant aquesta anàlisi prèvia amb relació als altres projectes de la Secció Filològica, cosa que està prevista que estigui llesta aquest any 1999. Després convindrà fer una anàlisi dels resultats des del punt de vista tècnic, amb l'objectiu de preparar la primera implementació dels mòduls. Aquesta anàlisi tècnica consistirà bàsicament en la classificació de les operacions d'extracció de la informació que hi ha al darrere de les propostes d'activitats didàctiques previstes en l'estudi previ. En aquest moment, sembla raonable preveure que aquesta classificació s'haurà de fer atenent, almenys, els criteris següents:

- la base de dades de partida;
- el tipus de programes que s'hagin d'usar;
- la interfície amb l'usuari que convingui preveure.

Amb aquesta anàlisi s'acabaran les etapes prèvies a la implementació i ja es podrà començar a treballar en la construcció efectiva dels primers mòduls del projecte.

D'aquests primers passos del projecte, ja en podem treure alguna conclusió, que resumeixo breument aquí. En primer lloc, es constata clarament que les escoles catalanes tenen necessitats clares amb relació a la llengua catalana, fruit en gran part de la situació sociolingüística del nostre país. En segon lloc, s'observa, ben clarament també, que algunes d'aquestes necessitats poden ser cobertes adequadament amb informació que és present a les bases de dades que la Secció Filològica ha anat constituint en els seus projectes dels darrers anys. Finalment, es comprova que aquesta informació, convenientment classificada segons els criteris rellevants, pot ser tractada adequadament en cada cas.