

Experiències

## Sobre la I Jornada del Processament Computacional del Català

### Autors

Gemma Boleda  
*Universitat Politècnica de Catalunya*

Montse Cuadros  
*Universitat Politècnica de Catalunya*

Cristina España-Bonet  
*Universitat Politècnica de Catalunya*

Maite Melero  
*Universitat Pompeu Fabra*

Lluís Padró  
*Universitat Politècnica de Catalunya*

Martí Quixal  
*Universitat Pompeu Fabra*

Carlos Rodríguez  
*Universitat Pompeu Fabra*

**L'article fa balanç de la Jornada del Processament Computacional del Català. S'hi exposen els objectius, es planteja una visió del processament des dels àmbits de recursos i recerca i el perfil dels participants, es plantegen els principals fils d'argumentació del debat que s'hi va portar a terme i se n'extreuen unes conclusions.**

El processament computacional de la llengua abraça qualsevol activitat relacionada amb la creació, gestió i utilització de tecnologia i recursos lingüístics. En el pla científic, aquesta activitat és central en disciplines com ara la lingüística de corpus, l'enginyeria lingüística o el processament del llenguatge natural escrit o parlat. En el pla quotidià, el processament s'inclou en un ampli ventall d'aplicacions cada cop més habituals: sistemes automàtics d'atenció telefònica, traducció automàtica, etc.

La gran majoria d'aquestes aplicacions requereixen eines i recursos lingüístics específics per a cada llengua. Per a llengües amb un mercat ampli, com l'anglès o el castellà, l'oferta de productes i serveis basats en tecnologia lingüística és variada i habitual. Per al cas de llengües com el català, és més difícil trobar productes i serveis que s'ofereixin ja "de fàbrica" amb aquesta tecnologia.

Per tal de reflectir l'estat actual de les tecnologies de la llengua aplicades al català, de posar en contacte els membres d'aquesta comunitat, i d'impulsar iniciatives que les potenciïn, al març del 2009 es va celebrar al Palau Robert de Barcelona la primera Jornada del Processament Computacional del Català (JPCC).<sup>1</sup> La Jornada tenia l'objectiu d'esdevenir un punt de trobada i alhora un aparador per als grups de recerca de l'àrea, i encetar el debat sobre com articular la comunitat per tal de potenciar l'ús i el desenvolupament del català tant en la tecnologia lingüística com en els productes i serveis que en depenen. Aquest article presenta un resum del contingut i les conclusions de la Jornada.

### Introducció

Cada cop són més habituals els productes tecnològics que incorporen cert grau de processament computacional de la llengua. Des del senzill predictor de paraules del telèfon mòbil, fins a un traductor automàtic integrat al navegador o al correu electrònic, les interfícies tecnològiques intenten apropar-se cada cop més a la forma d'interacció més natural per a les persones: la llengua.

1. <http://sites.google.com/site/jornadacatala/>

El desenvolupament d'aquestes tecnologies és sovint costós, especialment pel que fa a les eines i recursos que són específics per a cada llengua. Per tant, cada llengua que s'afegeix als productes suposa una inversió que no sempre resulta rendible per a les empreses que els desenvolupen. En el cas de llengües com el català, és clau el paper que poden tenir les administracions, universitats i centres de recerca a l'hora de facilitar l'accés a aquestes eines i recursos per tal que desenvolupar aquests productes en català suposi una rendibilitat comparable a fer-ho en una altra llengua amb un mercat més ampli.

La primera Jornada del Processament Computacional del Català (JPCC) va tenir lloc al març del 2009 amb els objectius següents: 1) millorar la comunicació i la col·laboració entre els diferents grups de recerca, empreses i institucions que desenvolupen eines i recursos computacionals per al català; 2) trobar maneres d'aprofitar de forma eficient els recursos existents, i 3) donar visibilitat a la recerca en el tractament computacional del català. Així, doncs, la Jornada anava dirigida principalment al col·lectiu més vinculat a la recerca, tot i que, com veurem, també va despertar l'interès de persones d'altres àmbits socials, sobretot empresa i administració.

Les següents seccions presenten el contingut de la JPCC: primer es resumeixen les presentacions de grups de recerca de l'àrea, i es posa de manifest la varietat i l'abundància de productes tecnològics existents per al català. Així mateix, es mostren les dades de participació. Seguidament, es resumeixen les aportacions principals del debat celebrat durant la Jornada. Finalment, es presenten les conclusions dels organitzadors sobre el contingut de la trobada. Remetem el lector al lloc web de la JPCC per accedir a tota la documentació generada al voltant de la Jornada (debat, presentacions, demostracions i pòsters).<sup>2</sup>

### **Visió general del processament computacional del català**

Els grups de recerca que treballen en el processament computacional del català presenten un ampli ventall d'interessos i perfils. La major part d'aquests grups té un vincle amb universitats públiques o privades. En aquesta secció fem un breu resum dels grups i de la recerca presentada.

#### ***Recursos i recerca***

La primera JPCC va reunir 14 grups o centres de recerca provinents de Catalunya i del País Valencià en els quals es treballa per al català, però també per a d'altres llengües. Així mateix, es van presentar 27 eines, recursos i aplicacions existents per al català.

Quant als grups, se'n van presentar de força veterans tant pel que fa al vessant lingüístic com al computacional, però també hi va haver grups que s'han format en els darrers anys. Dels 14 grups, 11 són creadors i usuaris d'eines i recursos basats en tecnologia lingüística en català, i els altres tres en són principalment usuaris. Tots treballen o bé en la recerca teòrica o en l'aplicada, i un d'ells específicament en la transferència de tecnologia, tot i que de diverses maneres tots els grups s'involucren en aquesta tasca. El finançament dels projectes és tant autonòmic com estatal i europeu i, a més dels concursos específics per a la recerca, la majoria dels grups tenen col·laboracions amb empreses i l'Administració pública.

---

2. <http://sites.google.com/site/jornadacatala/>

Pel que fa a les línies de treball i eines representades, podríem dir que hi eren gairebé totes les àrees del processament del llenguatge natural i de la parla: des del processament lingüístic de textos més bàsic (segmentació de paraules i oracions, anotació automàtica de categories gramaticals, etc.) fins al d'interacció amb sistemes de teleoperadors automàtics (reconeixement i síntesi de la parla), passant per sistemes de traducció i correcció automàtica o assistida, reconeixement automàtic de textos manuscrits, eines d'aplicació a la lingüística forense, d'ajut a la creació i manteniment de bases de dades lèxiques i terminològiques, subtítolació automàtica, etc. Pel que fa a recursos, s'hi van presentar corpus de textos escrits i orals de naturalesa diversa, amb anotacions automàtiques i manuals, i una bona varietat de recursos lexicogràfics i terminològics de diverses àrees temàtiques i des de diferents perspectives teòriques.

A continuació presentem una taula que resumeix les línies de treball dels grups que es van presentar a la Jornada.

<p>Grup de Veu i Llenguatge (Barcelona Media Centre d'Innovació) i GLiCom (Universitat Pompeu Fabra)</p>
<p><i>Objectiu:</i> el desenvolupament de nous productes i serveis, i la potenciació de la recerca i el desenvolupament (R+D). <i>Línies de treball:</i> extracció d'informació, correcció i revisió de textos, traducció automàtica, síntesi de veu, llengua de signes i tecnologies per a la docència de la traducció i la lingüística. <i>URL:</i> <a href="http://www.barcelonamedia.org">http://www.barcelonamedia.org</a> i <a href="http://www.glicom.upf.edu">http://www.glicom.upf.edu</a></p>
<p>Grup de Processament del Llenguatge Natural, Centre de Tecnologies i Aplicacions del Llenguatge i la Parla (Universitat Politècnica de Catalunya)</p>
<p><i>Objectiu:</i> la recerca en el camp del processament computacional del llenguatge natural, així com la creació de recursos lingüístics. <i>Línies de treball:</i> recerca en processadors lingüístics, adquisició, integració i explotació de coneixement lexicosemàntic, anàlisi semàntica des de diferents punts de vista (WSD, SRL, NLU), aprenentatge automàtic aplicat al PLN, resum automàtic i traducció automàtica. <i>URL:</i> <a href="http://www.lsi.upc.edu/~nlp">http://www.lsi.upc.edu/~nlp</a></p>
<p>Grup de Tractament Automàtic del Llenguatge Natural, Departament de Tecnologies de la Informació i les Comunicacions (Universitat Pompeu Fabra i Barcelona Media Centre d'Innovació)</p>
<p><i>Objectiu:</i> fer recerca i aplicacions en alguns camps específics del processament automàtic del llenguatge natural, en el marc teòric de la Teoria Sentit-Text (TST). <i>Línies de treball:</i> generació multilingüe de llenguatge natural i altres continguts, resum automàtic, paràfrasi, traducció, lexicologia computacional i aprenentatge automàtic orientat a l'adquisició de recursos lingüístics. <i>URL:</i> <a href="http://www.recerca.upf.edu/taln">http://www.recerca.upf.edu/taln</a></p>
<p>Servei Lingüístic (Universitat Oberta de Catalunya)</p>
<p><i>Objectiu:</i> fer un ús intensiu de la tecnologia disponible per tal d'optimitzar els processos productius (rendibilitzar costos i aportar qualitat i homogeneïtat). <i>Línies de treball:</i> l'explotació i la millora dels sistemes de traducció automàtica, i la creació d'eines i recursos de suport als processos de correcció i traducció. <i>URL:</i> <a href="http://www.uoc.edu/serveilinguistic">http://www.uoc.edu/serveilinguistic</a></p>

<p>FlexSem, Departament de Filologia Francesa i Romànica (Universitat Autònoma de Barcelona)</p>
<p><i>Objectiu:</i> la recerca aplicada al processament automàtic del llenguatge natural des d'una perspectiva de formalització lingüística i la creació de recursos lingüístics. <i>Línies de treball:</i> prosòdia, patologies de la parla i intercomprensió entre llengües romàniques; les vinculacions i vehiculacions semàntiques del lèxic amb aplicabilitat en la creació de diccionaris electrònics i mòduls de processament lingüístic. <i>URL:</i> <a href="https://masters.uab.es/flexsem">https://masters.uab.es/flexsem</a></p>
<p>Centre de Llenguatge i Computació (Universitat de Barcelona)</p>
<p><i>Objectiu:</i> desenvolupar recursos de tecnologia lingüística (corpus anotats, lexicons, analitzadors) que són a la base de les aplicacions basades en processament del llenguatge, amb cura dels fonaments lingüístics i metodològics. <i>Línies de treball:</i> processament massiu de textos, anàlisi morfològica i sintàctica, corpus anotats, analitzadors basats en ML, resolució de la coreferència, identificació de paràfrasis. <i>URL:</i> <a href="http://clic.ub.edu">http://clic.ub.edu</a></p>
<p>Institut Universitari de Lingüística Aplicada (Universitat Pompeu Fabra )</p>
<p><i>Objectiu:</i> la creació de recursos lingüístics i la recerca basada en la lingüística de corpus i les tecnologies lingüístiques. <i>Línies de treball:</i> demolingüística, discurs especialitzat, documentació digital, enginyeria lingüística, extracció d'informació, fonologia aplicada, lexicografia, lexicologia, lingüística forense, neologia, sociolingüística, terminologia i variació lingüística. <i>URL:</i> <a href="http://www.iula.upf.edu">http://www.iula.upf.edu</a></p>
<p>Grup de Recerca en Tecnologies Mèdia Enginyeria i Arquitectura la Salle (Universitat Ramon Llull)</p>
<p><i>Objectiu:</i> treballar la multidisciplinarietat centrada en la innovació en noves tecnologies multimèdia i multimodals. <i>Línies de treball:</i> síntesi de la parla expressiva, creació d'avatars (cares parlants i LSC), reconeixement de la parla, reconeixement audiovisual d'emocions i, finalment, disseny i enregistrament de corpus de veu. <i>URL:</i> <a href="http://www.salle.url.edu/~iriondo">http://www.salle.url.edu/~iriondo</a></p>
<p>Grup de Recerca Interuniversitari en Aplicacions Lingüístiques, Departament de Lingüística General (Universitat Autònoma de Barcelona)</p>
<p><i>Objectiu:</i> desenvolupar, avaluar i aplicar recursos lingüístics de gran cobertura, creant i fent ús de gramàtiques, lèxics, corpus i memòries de traducció. <i>Línies de treball:</i> gramàtiques (dependències, HPSG), recursos lèxics, corpus (anotació i explotació), adquisició d'informació (esquemes de subcategorització i restriccions de selecció) i memòries de traducció. <i>URL:</i> <a href="http://grial.uab.es">http://grial.uab.es</a></p>

<p>Pattern Recognition and Human Language Technology, Institut Tecnològic d'Informàtica (Universitat Politècnica de València)</p>
<p><i>Objectiu:</i> el reconeixement de patrons, l'aprenentatge automàtic i la interacció multimodal. <i>Línies de treball:</i> traducció de textos i veu (automàtica o assistida), reconeixement de la parla, reconeixement de textos manuscrits (automàtic o assistit), biometria i recuperació d'imatges interactiva. <i>URL:</i> <a href="http://prhlt.iti.es">http://prhlt.iti.es</a></p>
<p>Grup de Tractament de la Parla, Centre de Tecnologies i Aplicacions del Llenguatge i la Parla (Universitat Politècnica de Catalunya)</p>
<p><i>Objectiu:</i> la recerca en el camp del processament computacional de la llengua, així com en la creació de recursos lingüístics, amb un interès especial cap a la traducció automàtica. <i>Línies de treball:</i> reconeixement automàtic de la parla, conversió de text a veu, traducció automàtica oral i textual, creació de recursos lingüístics (orals i textuals) i el processat de la parla i àudio en interfícies multimodals. <i>URL:</i> <a href="http://gps-tsc.upc.es/veu">http://gps-tsc.upc.es/veu</a></p>
<p>Grup Transmedia Catalonia, Departament de Traducció i d'Interpretació (Universitat Autònoma de Barcelona )</p>
<p><i>Objectiu:</i> l'experimentació mitjançant tecnologies com l'eye-tracking o el reconeixement de parla aplicades a la traducció audiovisual i a l'accessibilitat. <i>Línies de treball:</i> traducció audiovisual (doblatge, subtitulació, veus superposades), accessibilitat als mitjans (audiodescripció, subtitulació per a sords i audiosubtitulació) i accessibilitat en la docència. <i>URL:</i> <a href="http://www.fti.uab.cat/transmediacatalonia">http://www.fti.uab.cat/transmediacatalonia</a></p>
<p>Projecte institucional d'investigació IVITRA (Universitat d'Alacant)</p>
<p><i>Objectiu:</i> la recerca d'eines relacionades amb la traducció automàtica i d'aplicacions per al tractament de corpus multilingües i paral·lels. <i>Línies de treball:</i> elaboració de corpus del català antic, textos paral·lels (traduïts) i totes les eines relacionades amb el seu tractament. <i>URL:</i> <a href="http://www.ivitra.ua.es">http://www.ivitra.ua.es</a></p>
<p>Transducens, Departament de Llenguatges i Sistemes Informàtics (Universitat d'Alacant)</p>
<p><i>Objectiu:</i> la recerca i la creació d'aplicacions que emprin tecnologies de processament del llenguatge natural. <i>Línies de treball:</i> l'educació assistida per ordinador, la indexació i llenguatges de marcat en biblioteques digitals, i la inferència gramatical a partir de mostres estocàstiques i les seves aplicacions; traducció automàtica, anotació fonètica de textos per a la lectura en veu alta, i eines per al processament de textos paral·lels i generació de memòries de traducció. <i>URL:</i> <a href="http://transducens.dlsi.ua.es">http://transducens.dlsi.ua.es</a></p>

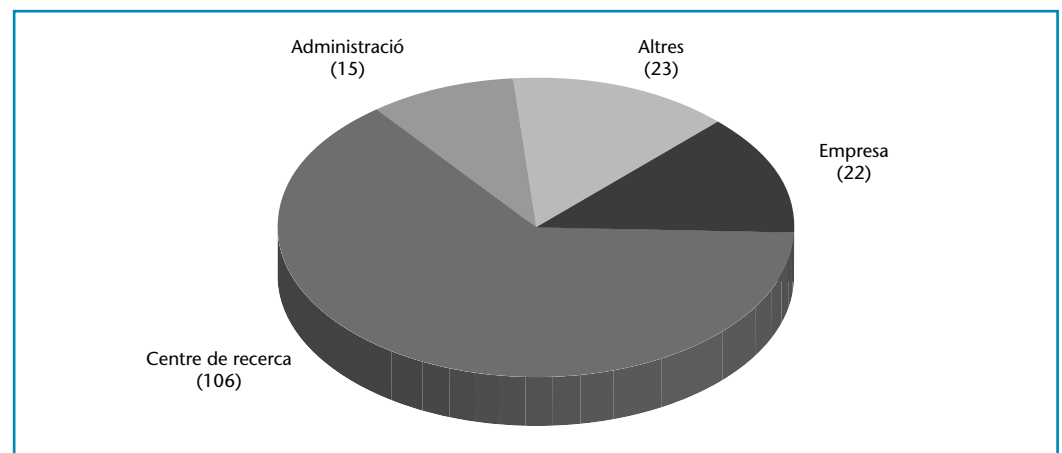
### Perfil dels participants

El processament computacional del català involucra de manera directa o indirecta persones de diferents àmbits de la societat, i això es va veure reflectit en la procedència dels assistents a la Jornada. Amb 166 persones inscrites provinents principalment de Catalunya i del País Valencià, la Jornada va ser tot un èxit de participació.

Un dels objectius principals de la trobada era donar a conèixer a la comunitat la recerca que s'està duent a terme en l'àrea, i això va fer que gran part dels assistents, 109, fossin representants d'universitats o centres de recerca. Com ja s'ha dit, 14 grups de recerca van presentar les seves línies d'investigació relacionades amb el català.

Els altres elements importants, l'empresa privada i l'Administració pública, també es van veure representats a la Jornada, més o menys a parts iguals. La distribució es pot veure a la figura 1, on es mostra el nombre de participants classificats segons la seva pertinença a un dels quatre grups d'afiliacions laborals distingits: centre de recerca, Administració, empresa privada i altres. Sota la denominació "Altres" s'inclouen des de representants d'institucions com l'Institut d'Estudis Catalans i el Termcat, fins a representants d'entitats del sector de la comunicació i la tecnologia en català com ara la Corporació Catalana de Mitjans Audiovisuals, Softcatalà o VilaWeb.

Gràfic 1. Distribució dels 166 inscrits segons l'àmbit de procedència. Alguns dels assistents pertanyen a dos dels àmbits considerats.



### Debat

El debat va ser moderat per Vicent Partal (Vilaweb) i els ponents foren Toni Hermoso (SoftCatalà), M. Antònia Martí (U. Barcelona, Thera), Mikel Forcada (U. Alacant, Transducens) i Joan Soler (IEC). El debat es va estructurar al voltant de tres preguntes generals, i el resum que en presentem l'articularem també en funció d'aquestes qüestions.

**Com es podria fer per ampliar la presència de la llengua catalana en els serveis i productes tecnològics? I per augmentar la transferència de la tecnologia que es desenvolupa als centres de recerca?**

Els ponents van estar d'acord que la salut computacional del català és força bona; de fet, els recursos i eines disponibles són molts més que els que li correspondrien per demografia o poder polític, i la seva situació de gairebé normalitat en aquest sentit és excepcional per a una llengua sense estat.

Això no obstant, es van identificar una sèrie de problemes que dificulten la difusió de les eines computacionals del català tant per a objectius de recerca com per a aplicacions d'usuari final, i que expliquem a continuació.

En primer lloc, *la baixa demanda social de recursos tecnològics del català* per part del món de l'empresa i els serveis públics (es van esmentar per exemple sectors com ara la banca o la sanitat) en què el català no és llengua vehicular real. Això fa que el desenvolupament privat de recursos en català sovint es faci per "voluntarisme". Així i tot, es va suggerir que les mateixes tecnologies lingüístiques podrien fer el paper d'actor en la recuperació sociolingüística del català. Si es fan eines en català prou útils i atractives, l'usuari les escollirà per aquestes raons.

També es va apuntar *la desconexió entre el desenvolupador i l'usuari final*. En aquest sentit, Mikel Forcada va argumentar que les aplicacions distribuïdes amb llicències de codi obert (p. e. Apertium)<sup>3</sup> faciliten aquesta connexió gràcies a les comunitats que s'estructuren al voltant d'aquest tipus de codi. Aquestes comunitats comparteixen i amplien les aplicacions segons els seus interessos i necessitats, i poden aglutinar tant grups de recerca com usuaris o empreses orientades a serveis.

Un punt crucial és també *la dificultat en l'accés a certes tecnologies i recursos per problemes de lliçència*. Aquí, de nou Forcada va subratllar la importància que augmenti el nombre de recursos de lliure distribució. En el cas que un recurs no es pugui fer de lliure distribució per problemes de propietat o *copyright*, com ara un corpus amb textos que siguin propietat d'autors o editorials, caldria facilitar com a mínim el seu ús per a objectius de recerca. En aquest cas, cal que les condicions d'ús estiguin ben especificades i que l'accés als recursos sigui àgil i sense obstacles burocràtics. Actualment, es dona una situació molt complicada, amb molta diversitat de llicències i condicions d'ús. Així, no és estrany el cas que per accedir a un recurs que és lliure per a recerca calgui esperar gairebé un any perquè cal signar un conveni entre les institucions implicades.

Especialment nocives són *la duplicació i dispersió d'esforços* en la construcció de corpus, eines, etc. El català és una llengua relativament petita, els recursos lingüístics són costosos i el finançament escàs. La impressió general és que no ens podem permetre com a comunitat malbaratar esforços d'aquesta manera. Cal doncs millorar la coordinació i fomentar la cooperació entre grups de recerca i també amb les empreses.

Es va esmentar, així mateix, la relativament *poca interacció entre lingüistes i informàtics* a l'hora de crear i compartir aplicacions i recursos. Cal més interdisciplinarietat, i això es veu dificultat per les rígides estructures en l'organització docent de la universitat, així com en la manca de centres de recerca híbrids.

Finalment, M. Antònia Martí també va posar sobre la taula una certa indiferència per part dels investigadors en tecnologies del llenguatge per la qüestió de la llengua. Això fa que llengües com

---

3. <http://xixona.dlsi.ua.es/apertium>

l'anglès, que són les que compten amb més recursos, actuen com a atractors dels esforços investigadors, desequilibrant cada vegada més la proporció de recursos entre les llengües. Per combatre aquest fenomen, cal afavorir la presència de recursos textuais per al català, com ara corpus anotats, en competicions internacionals per a tasques computacionals.

***Quines estratègies de finançament són necessàries? Quina relació hi ha d'haver entre l'origen del finançament i les condicions d'accessibilitat i distribució dels recursos?***

M. Antònia Martí va apuntar que l'empresari català tendeix més cap a la seguretat que cap al risc i la innovació. Per tant, calen polítiques d'incentivació a les empreses. Les polítiques actuals, però, estan orientades a projectes massa grossos només assumibles per grans empreses, com Microsoft. Les accions de l'Administració podrien ser més efectives si es fessin polítiques orientades a PIMES.

Així mateix, per tal d'afavorir el codi obert, es va proposar que en les convocatòries de projectes competitius els grups hagin d'especificar les condicions d'ús dels recursos que es generaran, per tal que els avaluadors ho puguin tenir en compte a l'hora de concedir el projecte i de controlar-ne el progrés. Des del públic es va suggerir que aquesta demanda es podria fer en l'àmbit estatal a fi que s'incorpori a la nova Llei de la ciència.

Un problema de la lliure distribució dels recursos generats pels projectes de recerca és que cal tenir en compte els costos de l'empaquetament final de l'aplicació. Aquests costos, així com la preparació de llicències i el mecanisme de distribució, s'haurien de preveure en la definició inicial dels projectes.

***Com hem d'articular la comunitat que formem les persones vinculades d'una manera o altra a l'àmbit del processament del llenguatge natural i de la parla de la llengua catalana?***

La major part dels ponents es va mostrar favorable a l'existència d'un portal web de les tecnologies lingüístiques del català que centralitzés tota la informació referent als recursos, les eines, els grups, etc. Es va posar com a model la Linguateca<sup>4</sup> de la llengua portuguesa. Respecte al suport necessari per crear el portal i sobretot mantenir-lo, es van plantejar diferents alternatives. La més repetida va ser que tocava als poders públics finançar aquest manteniment, possiblement aprofitant el marc d'algun projecte ja existent, ja sigui en l'àmbit de l'Administració local o estatal (p. e. programa AVANZA). Per la seva banda, Mikel Forcada va apuntar la possibilitat de crear-lo col·lectivament en forma de *wiki* cooperativa. Des del públic, Lluís de Yzaguirre va recordar l'existència del projecte europeu CLARIN, que ja inclou el català, i que pretén facilitar l'explotació de recursos lingüístics per part dels investigadors en ciències socials.

Així mateix, Joan Soler va suggerir la necessitat d'una associació que donés estructura a la comunitat vinculada amb l'àmbit de les tecnologies del català i que possibilités iniciatives com ara el portal web. Aquesta proposta va ser acollida favorablement per altres participants. M. Antònia Martí va apuntar la possibilitat de reconvertir la SEPLN<sup>5</sup> en una federació d'associacions ibèriques organitzades per llengua. De Yzaguirre, per la seva banda, va proposar la figura de l'institut interuniversitari, dependent de l'Administració catalana però vinculat a les universitats. Altres assistents van estar d'acord amb aquestes propostes però van insistir en la importància de mantenir també un vincle amb les empreses. Va haver-hi consens que una associació permetria donar veu al col·lectiu de

4. <http://www.linguateca.pt/>

5. Sociedad Española para el Procesamiento del Lenguaje Natural (<http://www.sepln.org/>)



recerca i alhora fer de nexa tant amb les empreses com amb l'Administració i els usuaris finals. De l'Administració es va demanar sobretot suport a la difusió i ajuda en el tema de les llicències.

Finalment, també es va proposar donar periodicitat a la Jornada i fer-la itinerant a altres poblacions dels Països Catalans. També es va suggerir la creació d'una llista de distribució inicial a partir de la llista dels assistents a la Jornada per tal de difondre les notícies relacionades amb la tecnologia lingüística del català.

### Conclusions

El primer objectiu de millorar la comunicació entre els membres de la comunitat es va aconseguir amb rotunditat, ja que, a més de congregar 14 grups de recerca provinents de diversos punts geogràfics en què es parla o es treballa amb el català, també vam tenir la sort de comptar amb 27 presentacions d'eines, recursos i aplicacions existents per al català. Creiem que l'interès despertat per la Jornada, així com la quantitat i qualitat de grups, eines, recursos i àmbits d'aplicació que s'hi van presentar, són una prova de la vitalitat tecnològica del català.

A partir d'aquí, i recollint un sentiment expressat repetidament pels participants, cal articular iniciatives per intensificar la comunicació i la col·laboració entre els diferents agents de la comunitat, tant entre els mateixos grups de recerca com amb l'empresa, ja sigui desenvolupadora o usuària. A la Jornada van sorgir dues propostes en aquest sentit: la creació d'un portal web que aglutini les eines i recursos computacionals del català (la Llenguateca?) i la constitució d'una associació.

L'objectiu de donar una major visibilitat a la recerca que es fa en el tractament computacional del català també es va aconseguir, ja que aproximadament un terç dels assistents no provenia del món de la recerca. En futures edicions de la JPCC seria bo ampliar la presència d'empreses, que en aquesta edició només eren convidades com a assistents, no com a participants.

Pel que fa als punts relatius a la col·laboració i l'aprofitament dels recursos, tal com hem vist, un bon nombre de membres de la comunitat opinen que hi ha maneres d'aconseguir que es faciliti l'accés a les eines i recursos, així com la reutilització, i que per a això cal una acció comuna i activa per part de la comunitat, però també suport institucional. En particular, per establir una política més oberta de distribució i una major facilitat d'ús de la tecnologia lingüística cal un suport decidit dels poders públics, tant autonòmics com estatals i europeus, ja que són ells els que estableixen les condicions de finançament dels projectes de recerca i de transferència de tecnologia.

En conjunt, la Jornada del Processament Computacional del Català va presentar l'estat de les tecnologies lingüístiques en català i va fer d'altaveu de les inquietuds i necessitats de diferents agents de la comunitat. Esperem que les iniciatives sorgides arran de la Jornada fructifiquin per donar resposta a les necessitats tecnològiques de la societat de parla catalana en l'era de la societat de la informació.

### Agraïments

Els organitzadors agraeixen a la Secretaria de Política Lingüística i a la Secretaria de Telecomunicacions i Societat de la Informació de la Generalitat de Catalunya el seu interès, suport i els recursos esmerçats en la Jornada. També agraeixen la presència dels ponents del debat Toni Hermoso, Mikel Forcada, M. Antònia Martí i Joan Soler, i especialment la de Vicent Partal, que va fer de moderador.

Igualment, també volen agrair la participació a tots els grups de recerca i persones que van presentar eines i recursos, i l'interès i participació de tots els assistents.

