

Apropament a una anàlisi de futur de la llengua i de la indústria de la llengua

Salvador Aparicio
Director de SIC, SL

D'ençà que una llengua neix no fa res més que recórrer el camí cap a la seva mort. L'únic que pot esperar és que els seus gens es perpetuïn en una nova línia evolutiva, en un nou llinatge pel qual lluitar.

Cada any moren llengües i dialectes, a tots els continents; estic content d'haver-ne salvat una. (Georges Dumézil —1898-1986—, lingüista i erudit que va documentar i confeccionar un diccionari de la llengua ubikh.)

Una jove veu l'amenaça (I)

Els avenços tecnològics dels darrers anys han donat una nova perspectiva a la visió relativista del temps. Einstein va inferir de la capacitat de modificar l'estat d'un cos i portar-lo fins a la velocitat propera a la de la llum que el temps tenia un comportament relatiu.

La tecnologia actual ens ha forçat a canviar la concepció del temps. Ara els moments també es mouen. En el context de les tecnologies de maquinari informàtic, els anys duren sis mesos; a Internet, en duren quatre.

La veu de la nova tecnologia és una amenaça per a les llengües tecnològicament dèbils, però també és el remei d'aquelles que se'n sàpiguen aprofitar: «La societat que no reïx a canalitzar la pròpia cultura dins el flux de la comunicació digitalitzada perd el tren del futur» (Miquel de Moragas).

Dues són les tecnologies que tindran una incidència cabdal en l'esdevenidor de la llengua: el reconeixement de la veu i la traducció automatitzada.

L'aplicació dels nous sistemes de reconeixement basats en la cognició dels fonemes, els micròfons intel·ligents, l'aplicació d'algoritmes de lògica difosa a patrons sonors i la suma de tots els petits i grans avenços que sorgeixen cada dia permeten, sense grans marges d'error, afirmar que en poc més de deu anys el teclat dels ordinadors serà un record anacrònic d'un passat molt proper.

I aquest serà el nostre primer gran escull. L'ortografia deixarà de ser un maldecap per a estudiants de primària. La sintaxi serà, aleshores, el seu, i el nostre, ogre particular.

Els programes de reconeixement de veu actual es troben encara lluny del que, amb un optimisme envejable, feia servir un del més coneguts (i avançats) com a motiu publicitari: *jo parlo, ell escriu*. Els productes actuals pateixen una manca de coneixement lingüístic de la parla, especialment de la parla natural, i de constituent psicològic. Ningú no parla substituint les pauses per «coma» o per «punt» ni acaba una pregunta dient «signe d'interrogació».

Però en la seva maduresa parlar a l'ordinador serà un acte natural del qual moltes llengües es veuran excloses.

Per contra, la traducció automàtica, flagell de traductors i riota de lingüistes, pot esdevenir el millor aliat de la llengua.

Les tecnologies actuals, i molt probablement les futures allà on abasta la imaginació científica, no resoldran el problema del traductor universal. Traduir és pensar, i construir màquines que pensin desperta totes les nostres pors. Però el que sí que assenyalen, com si d'una brúixola tremolosa es tractés, són les possibles dreces que ens apropen a eines bilingües prou fiables. La llengua és infinita, però alhora predictable i els recursos habituals dels seus usuaris són limitats.

Dues són les tecnologies que es fan servir per a la traducció d'informació en format electrònic: la traducció automàtica i la traducció assistida. L'única sortida raonable d'aquest coll d'ampolla on ens trobem és sumar esforços.

El cost de perfeccionament dels traductors automàtics fins a cotes de fiabilitat properes al 100% és asimptòtic amb la impossibilitat. Recursos de transferència sintàctica, de generació de metallenguatges ens acosten a la frontera, però són incapaços de travessar-la. Provar de desambiguar la llengua és lluitar contra la seva pròpia naturalesa.

Per contra, la possibilitat de comparar patrons sintàctics i semàntics preestablerts comença a ser una possibilitat econòmicament viable. Al mercat fa temps que es poden trobar productes on el flirteig entre traducció automàtica i memòria de traducció és un fet (T1, TRADOS o ATRIL en són un bon exemple).

Els resultats, els acceptables, encara són lluny.

Per començar els algorismes de la llengua són encara força desconeguts, i encara ho és més la manera de reduir-los a operacions matemàtiques coherents. Però els treballs en aquesta línia són constants. CREL, CLIC, IULA... i molts altres són referències fàcils i encoratjadores.

Però el coneixement cal aplicar-lo, i per fer-ho cal disposar d'uns corpus de text paral·lel de magnituds ciclòpies i perfectament acotats. L'emmagatzemament de memòries de traducció és gairebé una fal·lera entre els traductors *tecnològicament avançats*. Institucions, empreses i professionals liberals de la llengua recopilen informació, malauradament sense cap criteri estandaritzat. TMX o Martif, si bé solucions parcials i incompletes, són totalment desconeguts per un nombre molt elevat d'usuaris de les tecnologies de traducció assistida.

Coneixement i dades són un tamboret de dues potes al qual cal un gestor de dades, un ordinador, de potència i capacitat de càlcul inabastable ara per ara.

La nanotecnologia i l'aplicació dels principis de la mecànica quàntica al desenvolupament de productes informàtics ens obre la porta a especulacions poc sibil·lines. En poc més de vint-i-cinc anys, nanoconductors de carboni, commutadors moleculars, emmagatzemament de dades en estructures cristal·lines, permetran capacitats de càlcul i de gestió de dades que requeriran noves unitats de mesura.

Aquesta terra de promissió es dibuixa en un horitzó de poc més d'un quart de segle. La veritable era digital encara ha d'arribar. Utopies de somiadors podran fer-se realitat: ordinadors de la mida d'un gra de sorra, previsió climàtica, anàlisi genètica, i qui sap què, a més de traductors universals.

Però no tothom hi serà present. La major part de les llengües no hi seran a temps.

A més, tampoc no podem oblidar el component sociolingüístic del problema.

En un entorn socioeconòmic on preval la *ratio* cost/benefici, el manteniment de símbols d'identitat es fa a costa de voluntats personals i polítiques. La pèrdua de valors és un fet constatat. Abandonar una reivindicació activa, fins i tot militant, de la defensa de la llengua i adoptar una postura conformista és un perill que una llengua com la catalana no pot permetre's.

En un entorn social que avança cap a la unió paneuropea podem observar senyals prou evidents d'evolució lingüística i cultural. De les onze llengües oficials reconegudes per als 377 milions d'habitants de la Unió Europea, el 95% de la documentació generada es troba disponible, només, en anglès o en francès. Llengües tan significatives i potents com l'espanyola o l'alemanya han de compartir amb altres set idiomes un misèrrim 5% de quota de mercat. Quan es trigarà a adoptar una llengua franca? Només és qüestió de temps.

Cal, doncs, esmerçar recursos i esforços per a provar d'evitar l'inevitable? Provem-ho.

Una jove veu l'amenaça (II)

Les noves generacions, nascudes i formades en les lleis de la competitivitat i la productivitat poden arribar a qüestionar la validesa de determinats conceptes inamovibles per a nosaltres. Les veus del trilingüisme a Catalunya són ja un fet. I volen saber què els en compensarà l'esforç.

Les empreses estrangeres manifesten «una preocupació latent» per l'existència del català com a element desincentivador per impulsar les seves activitats a Catalunya. Les *Majors* fa temps que ens etziben bastonades, ara ja sense cap mena de pudor i els mercats laborals evolucionen cap a un mestissatge on prevalen molts altres valors aliens a la llengua.

En una societat on el temps i els costos són els valors més preuats la millor moneda de canvi és l'estalvi de temps i de diners.

Però, en canvi, el mercat de la llengua a Catalunya es troba marcat per un fort sentiment català. Pel nostre pes demogràfic podem considerar-nos una veritable potència en coneixement, investigació i desenvolupament de productes relacionats amb la nostra llengua i amb la llengua en general. Ara bé, tota aquesta potencialitat pateix, en general, una manca de planificació global, especialment des de les diverses plataformes públiques.

Podem ser un dels pobles amb més pàgines web per habitant, però amaguem el cap sota l'ala al fet que generen un volum de negoci molt inferior a les seves homòlogues en castellà o anglès. Les nostres universitats virtuals són de les millors del món, però la seva supervivència depèn de la seva expansió en mercats estrangers. Desenvolupem productes d'alt nivell tecnològic, però els hem de posar una interfície en anglès si volem projectar-los com a productes rendibles. Disposem d'un bon grapat d'opcions adaptades a la nostra llengua, però, o han estat mal gestionades (el cas del Windows 98 en català), o ens hem conformat amb un breu maquillatge facial (la data automàtica en català al Microsoft Word pren la ridícula forma de «diumenge, 29 / abril / 2001» quan el mateix programa compta amb recursos suficients per controlar el problema de l'apostrofació de la preposició).

Tot aquest ventall de queixes no ens portarà enlloc.

En futurologia, i no parlem de cap altra cosa, hi ha tres temps ben definits: és massa tard, aquest és l'últim tren i encara tenim temps per fer proves. Estem a les acaballes d'aquest darrer estadi, encara som a temps de provar, d'equivocar-nos i de rectificar, encara som a temps d'escollir entre més d'un camí.

Ara, la pregunta obligada sorgeix sense esforç: quins són aquests camins? I la resposta és clara: els que la imaginació i el coneixement ens marquin.

Calen moltes més inversions en el reconeixement de la parla: en investigació, en tecnologia del so i en desenvolupament de productes. L'actual oferta de mercat es redueix a un producte desenvolupat per Philips, el FreeSpeech, que és poc més que inoperant.

Calen moltes més inversions en lingüística aplicada i en lingüística computacional. El mercat de la indústria de la llengua necessita nous coneixements en què fonamentar el desenvolupament de nous productes. Sense una empenta seriosa en aquest camp no podrem resoldre els tres reptes que molt encertadament defineix en la seva presentació el CREL: el repte econòmic de la concurrència en el mercat lingüístic; el repte cultural de la supervivència de les llengües com a llengües vehiculars aptes per a qualsevol tipus de comunicació i tractament; i el repte polític d'adaptació a la modernitat i d'inserció en el multilingüisme.

Calen nous plantejaments de política de la llengua que en fomentin l'ús i abarateixin els costos de conviure en una societat plurilingüe: disponibilitat real dels fons de dades públiques i la seva adaptació a formats estàndards que permetin la seva reutilització en qualsevol àmbit, públic o privat.

Cal un nou plantejament, més coordinat, a la societat electrònica: recopilar i classificar tot el que hi ha a la xarxa que faci referència a la llengua catalana en un únic portal independent, és a hores d'ara, un projecte inajornable.

Calen nous plantejaments educatius. Els professionals i els estudiosos de la llengua, en una gran majoria, presenten grans mancances formatives en tecnologies que els haurien de ser molt properes. I els nous professionals que surten de les nostres universitats no es troben en condicions gaire més bones pel que fa a aquest apartat.

Cal un nou plantejament polític més bel·ligerant, però alhora més conseqüent. Les justifica-

cions i les solucions donades, per exemple, en el cas de la traducció de la darrera fornada d'aplicacions Microsoft, amaguen un desconeixement esfereïdor del tema.

Molts més són els camins que hauríem de fer, però parafrasejant un americà famós no hem de preguntar-nos què pot fer la tecnologia per nosaltres sinó què podem fer nosaltres per la llengua i adonar-nos que ja que Mahoma s'apropa molt lentament a la muntanya, potser fóra bo que la muntanya s'atansés una mica cap a Mahoma.

Amb la recança que sempre fa obrir la capsa dels trons, el marc d'aquestes Jornades és el més adient per llançar la pregunta a l'aire. És possible interactuar sobre la llengua per tal de facilitar l'entesa home-màquina? És possible dissenyar un nou estàndard, un llenguatge controlat per a registres determinats? Com s'hauria de fer, a què es podria renunciar? N'ignoro les respostes, però crec saber el temps de què disposem per obtenir-les.

Ens apropem al centenari d'una època de discussions, pactes i resolucions que van elevar el català al rang de llengua moderna. Potser és l'hora de reprendre de nou aquest tarannà.

Al cap i a la fi, fa temps que se sap, que va ser en fer-se un xic ençà que va néixer l'Empordà.