

La traducció automàtica, encara lluny

## 'Ordinadore, traduttore'

La traducció automàtica continua sent un somni. Després de gairebé quaranta anys de recerca els especialistes coincideixen en l'opinió que la màquina de traducció universal és un projecte sense concreció en un futur pròxim.

**T**ot i així s'ha avançat molt des de les primeres experiències realitzades als anys 50, amb èxits parcials en camps temàtics limitats, i avui existeixen diversos sistemes de traducció automàtica al mercat com l'Atlas de Fujitsu que tradueix de l'anglès al japonès; el Translation Manager/2 llançat per IBM pel juny de l'any passat, capaç de tractar 19 llengües europees (gairebé totes les oficials al continent); o el més reeixit, segons els especialistes, el sistema METAL (Machine Translation & Analysis of Natural Language) de Siemens i Nixdorf, adaptat a la traducció de manuals d'ús, contractes i altres documents administratius de l'alemany a l'anglès, el francès, l'italià i l'espanyol, i també del neerlandès al francès. Igualment, es troben en desenvolupament diversos programes de traducció automàtica, alguns amb noms tan suggerents com Rosetta (el nom de la pedra trilingüe que va servir a Champollion per a desxifrar els jeroglífics de l'egipci antic), Philips o Gral (el màgic calze cercat per Perceval) de Fiat i Aerospaziale, sigles de "Grammars Re-used to Automatically Analyse Languages", el MU-2 del govern japonès o el programa de l'Europa Comunitària, Eurotra, que justament enguany comença el seu desenvolupament industrial, transformat en el programa ALEP.

Eurotra ha estat un dels projectes de traducció automàtica més importants del món, tant pel nombre de llengües implicades (totes les comunitàries amb excepció de l'irlandès i el luxemburguès), com pel nombre de científics que hi treballen (uns 200) i el volum pressupostari, estimat en 37,5 milions d'ecus aportats per la CE, més una quantitat gairebé idèntica aportada pels països participants en el projecte.

Aquest programa de traducció multilingüe va començar de manera efectiva al 83 amb la participació de les quatre llengües comunitàries germàniques: alemany, danès,



L'objectiu final del programa de traducció automàtica Eurotra és de resoldre els problemes econòmics, po-

holandès i anglès; el grec i dues de romàniques: francès i italià. Posteriorment, amb la incorporació a la CE de l'estat espanyol i de Portugal, s'hi van afegir l'espanyol i el portuguès, igualment romàniques.

Totes les llengües han estat analitzades a quatre nivells –de l'estructura morfològica a l'anàlisi sintàctica i semàntica– i sintetitzades, "generades" en un altra llengua, co-

sa que suma un conjunt de 72 mòduls (cada una de les nou llengües per les vuit restants). A l'estat espanyol aquest treball s'ha fet des d'aquí, sota la direcció del catedràtic de Lingüística General de la Universitat de Barcelona, Ramon Cerdà. Des de Barcelona, a instàncies del CIRIT es va fer també una proposta d'incorporar el català a Eurotra a través del castellà i

l'anglès –un grup de la Universitat d'Essex (Regne Unit) hi estava interessat– “però no va prosperar tot i que era relativament barat –explica Ramon Cerdà– perquè al final no es va arribar a un acord sobre qui havia de pagar aquests dos mòduls”, si Generalitat o govern espanyol.

L'objectiu final d'Eurotra quan va començar, al 1982, era de resoldre els problemes econòmics, polítics i socials que origina la coexistència de tantes llengües diverses, especialment amb vista al mercat únic i, en primer terme, reduir les despeses en traducció de documentació interna i externa, una de les partides més oneroses del

i tipus de tests limitats, però per a donar suport a la sofisticació d'un sistema Eurotra operatiu manca encara un progrés tècnic substancial en els aspectes lingüístic i informàtic en el terreny del tractament del llenguatge natural.

Els problemes amb què ha topat Eurotra són els mateixos amb què es troben els sistemes de traducció automàtics, és a dir: que funcionen molt bé dins el camp dels llenguatges especialitzats que cobreixen un domini precís, com per exemple la medicina, d'on ha estat bandejada tota ambigüitat, però en altres textos queden atrapats en les trames de la sintaxi i fracassa-

destinació” van pensar els primers informàtics que als anys 50, especialment als Estats Units, van començar les primeres experiències de traducció automàtica. El sistema de traducció paraula per paraula va fracassar estrepitosament, “sovint no es podia ni tan sols endevinar de què tractaven els textos ‘traduïts’ –explica el director d'Eurotra– si giraven entorn de la paleontologia o de l'art culinària...”, cosa que demostra allò que tots els traductors professionals saben: que traduir no és senzillament fer un trasllat literal paraula per paraula, sinó interpretar un text en una llengua i traspassar aquesta interpretació en una altra.

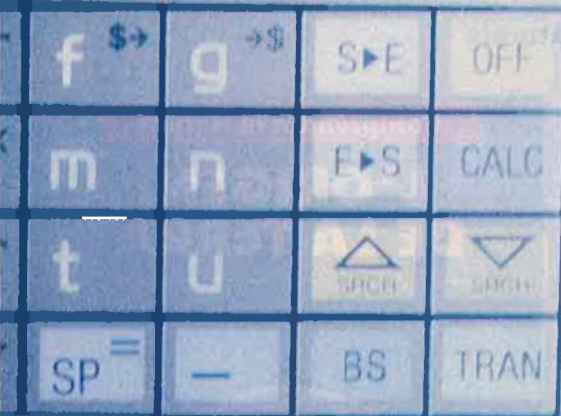
Durant les dècades dels 60 i els 70 van sorgir dos canvis substancials: d'una banda, els primers grans llenguatges de programació, com el COBOL o el BASIC, i d'una altra, la teoria generativo-transformacional, que Chomsky va desenvolupar en el camp de la lingüística teòrica, d'un formalisme tan rigorós que s'acostava a un llenguatge de programació. De fet, la gramàtica generativo-transformacional, el famós “arbre” chomskià, es troba en els fonaments de tots els sistemes de traducció actualment operatius, i ajudant a convertir el llenguatge natural, ambigu, en una operació matemàtica vist que, com assenyalava la cap del programa Eurotra a Barcelona, Núria Bel, “els ordinadors necessiten una informació no ambigua perquè treballen amb processos algorítmics”.

Però si la gramàtica generativa, és a dir, l'anàlisi sintàctica en profunditat, pot revelar-se útil entre llengües d'una mateixa família, esdevé inservible entre llengües del tot dispars com ara el català i el japonès. Per a això caldria emprar un model d'interlingua que significa passar per un estadi interpretatiu de representacions semàntiques abstractes, és a dir, formulades al marge de tota llengua natural, i per tant l'ordinador hauria de passar a formalitzar una modelització del món, val a dir: a entendre el significat de conceptes com “calor”, “fred”, “amor” o “vida” i pràcticament a “pensar” com un ésser humà.

Actualment hi ha obert un gran projecte de coneixement informàtic del món a Carnegie Mellon (EUA), encara en una fase molt primitiva. Sembla que falta molt de temps per a aconseguir que un ordinador tingui un coneixement efectiu de la realitat. Tant, si més no, com el que manca perquè arribi a ésser capaç de traduir tan bé com un home immers en el món que coneix.

**Magda Bertran**

## SPANISH → ENGLISH TRANSLATOR



lítics i socials que origina la coexistència de tantes llengües diverses.

CARME PUERTOLAS

pressupost comunitari, tant com d'agilitar aquesta tasca, perquè malgrat l'existència d'un veritable exercit de traductors i intèrprets a Brussel·les, mai no aconsegueixen de cobrir totalment les fites de traducció assenyalades. Finalitzat el programa, s'ha aconseguit l'objectiu tècnic, és a dir, crear un prototip de sistema de MT (Machine Translation) per a un camp temàtic limitat

sen com a traductors d'una llengua en conjunt.

Segons Ramon Cerdà, “cada vegada es veu que és menys senzill el processament del llenguatge natural per mitjans informàtics, cosa que d'entrada semblava molt fàcil”. “Res de més senzill que traduir amb ordinador assignant a cada paraula el seu equivalent en la llengua de