

Informàtica Aplicada a la Traducció  
15 de gener del 2000  
Segona part (1 hora i 40 minuts)

**Instruccions:**

- A l'eixida us donaré una solució possible de l'examen perquè us pugueu autoavaluar.
- En aquesta part de l'examen podeu consultar qualsevol material que hàgeu portat però no podeu compartir material ni parlar amb ningú.
- Si teniu algun dubte, alceu la mà i aniré a atendre-us.

**Qüestions:**

1. Indiqueu quina classe d'ambigüitat presenten aquestes frases (justifiqueu breument la vostra resposta):
  - (a) "Aureli explicarà les històries que ha escoltat a Martí" (1: li les explicarà a Martí; 2: li les ha escoltades a Martí);
  - (b) "Com dius que l'has matat?" (1: Com l'has matat?; 2: Com dius això?)
  - (c) "Un problema de llengua li impedeix parlar bé" (1: el problema és físic, anatòmic; 2: el problema és lingüístic, psicològic)
  - (d) "Li he dit que vindria més tard" (qui?)
  - (e) "Connecteu el ratolí a l'ordinador i activeu-lo" (1: activeu el ratolí; 2: activeu l'ordinador).
  - (f) "Han trobat gerres i àmfores ibèriques" (les gerres també són ibèriques?).
  - (g) "Les gallines han destrossat el sembrat, però no les mates" (1: no les sacrifiqueu; 2: no les plantes).
2. Moltes vegades, l'ambigüitat lèxica no és ni polisèmia (*estació, bomba*) ni ambigüitat lèxica amb canvi de categoria gramatical (*sobre* [preposició, substantiu i, en valencià, verb], *riu* [substantiu i verb]) sinó que succeeix perquè dues formes *de la mateixa categoria lèxica* són homògrafes: *volem* pot ser una forma del verb *volar* i del verb *voler*; *podeu* pot ser una forma del verb *poder* i del verb *podar*; en castellà, *creo* és una forma de *crear* o de *crear*, *fui* és una forma de *ir* o de *ser*, etc. Per a resoldre l'ambigüitat d'un mot polisèmic s'ha d'usar informació semàntica; per a resoldre l'ambigüitat lèxica categorial sol ser suficient usar informació sintàctica (per exemple, la categoria gramatical del mots anterior i posterior); però, és possible resoldre l'ambigüitat

deguda a l'homografia de mots de la mateixa categoria usant només la sintaxi o és necessari l'ús d'informació semàntica?

3. Els sistemes de traducció mot per mot poden cometre, per exemple, errors deguts a l'elecció incorrecta de la categoria gramatical d'un mot lèxicament ambigu. Elegiu dues llengües,  $L_1$  i  $L_2$  i poseu dos exemples de traduccions errònies de  $L_1$  a  $L_2$ , indicant la frase original, la frase mal traduïda i la frase correcta.
  
4. Els sistemes de traducció automàtica entre dues llengües amb sintaxi similar no necessiten fer massa reordenaments perquè l'ordre dels mots no varia massa d'una llengua a altra. A pesar d'això, la traducció mot per mot no és practicable perquè el gènere i el nombre gramatical d'alguns substantius varia i els adjectius, articles, etc., que l'acompanyen no concorden correctament: cast. *una señal muy clara* → cat. *\*una senyal molt clara* (correcte: *un senyal molt clar*); cast. *me gusta la leche fría* → ital. *\*mi piace la latte fredda* (correcte: *mi piace il latte freddo*). Una manera d'identificar zones on s'ha d'establir la concordança és detectar seqüències de mots, de manera similar a com es fa en els sistemes de transferència morfològica, però sense reordenar-les. Per exemple, detectar la seqüència **art-subst** pot servir per propagar el gènere i el nombre del substantiu a l'article. Fixeu-vos en les frases castellanes següents i les traduccions al català fetes per un sistema que usa aquesta estratègia i deduiu quines són les seqüències que detecta i quines no. Justifiqueu la vostra resposta.
  - (a) Nos ofreció un postre → *Ens va oferir unes postres*
  - (b) Nos ofreció un postre buenísimo → *Ens va oferir unes postres boníssimes*
  - (c) Nos ofreció un buen postre → *\*Ens va oferir un bon postres*
  - (d) Nos ofreció un postre típico buenísimo → *\*Ens va oferir unes postres típiques boníssim*
  - (e) Nos ofreció un postre muy bueno → *\*Ens va oferir unes postres molt bo*
  
5. Imagineu que un traductor professional cobra 0,05 euros per mot de text traduït i que un corrector de textos cobra 0,10 euros per mot de text corregit. Imagineu que tenim un sistema de traducció automàtica que ens costa uns 0,03 euros per mot traduït i que produeix un 10% de mots incorrectes en les traduccions. Convé adoptar-lo i contractar el corrector o és millor contractar el traductor professional? (si no sabeu calcular-ho en general, feu els càlculs amb un text de, per exemple, 1000 mots).