

Informàtica Aplicada a la Traducció  
15 de gener del 2000  
Esbós de solució de la segona part

**Respostes:**

1.
  - (a) Ambigüitat estructural pura d'adjunció: el sintagma preposicional “a Martí” pot modificar a “explicarà” (oració principal) o a “ha escoltat” (oració subordinada).
  - (b) Ambigüitat estructural pura (d'adjunció): l'adverbi interrogatiu “com” pot pertànyer a l'oració principal o pot pertànyer a la subordinada i haver estat desplaçat obligatòriament (“moviment de Qu”) a la posició inicial.
  - (c) Ambigüitat lèxica del mot “llengua”.
  - (d) Ambigüitat deguda a l'el·lipsi: el subjecte del verb “vindria” pot ser “jo”, “ell”, “ella”.
  - (e) Ambigüitat extrasentencial deguda a l'anàfora: el pronom “lo” pot tenir dos antecedents diferents: “l'ordinador” i “el ratolí”.
  - (f) Ambigüitat estructural pura d'origen coordinatiu: “ibèriques” pot modificar tot el sintagma nominal coordinat o només el segon.
  - (g) Ambigüitat estructural d'origen coordinatiu, deguda a dues ambigüitats lèxiques categorials: “les” (pronom o article) i “mates” (substantiu i verb); de les quatre combinacions, dues són sintàcticament acceptables.
2. La solució semàntica és més potent i general però exigeix una anàlisi molt profunda de la frase. En alguns casos, la sintaxi podria donar pistes que permetrien una desambiguació molt aproximada. Per exemple, si *fuí* va seguit de la preposició *a*, és molt probable que es tracte del verb *ir*; d'altra banda, si va seguit d'un participi passat, és molt probable que es tracte del verb *ser*: es podria fer una categoria gramatical especial per al verb *ser* i usar tècniques de desambiguació categorial. Si trobem *podem* (*volem*) seguit d'infinitiu, és molt més probable que es tracte del verb *poder* (*voler*) que del verb *podar* (*volar*); de nou, caldria usar una categoria gramatical especial, en aquest cas per als verbs modals.
3. Es poden trobar molts exemples; per exemple, entre  $L_1 = \text{castellà}$  i  $L_2 = \text{català}$ , tenim:
  - Ayer por la mañana vino tarde  $\rightarrow$  \*Ahir pel demà vi vesprada (Ahir de matí va venir tard).
  - Río porque no llegó a la meta  $\rightarrow$  \*Riu perquè no va arribar a la fique (Ric perquè no va arribar a la meta).

4. Vegem que passa amb cada una de les oracions:

- (a) Nos ofreció un postre  $\rightarrow$  *Ens va oferir unes postres*: La traducció és correcta. Sembla que reconeix la seqüència (1) **art-subst** i propaga el nombre i el gènere del substantiu a l'adjectiu.
- (b) Nos ofreció un postre buenísimo  $\rightarrow$  *Ens va oferir unes postres boníssimes*: La traducció és correcta. Sembla que reconeix la seqüència (2) **art-subst-adj** i propaga el nombre i el gènere del substantiu tant a l'article com a l'adjectiu,
- (c) Nos ofreció un buen postre  $\rightarrow$  *\*Ens va oferir un bon postres*: No funciona. No reconeix la seqüència **art-adj-subst**, i tradueix mot per mot.
- (d) Nos ofreció un postre típico buenísimo  $\rightarrow$  *\*Ens va oferir unes postres típiques boníssim*: funciona incorrectament perquè no reconeix la seqüència completa **art-subst-adj-adj**; en canvi, sí reconeix la seqüència més curta (2) **art-subst-adj** i propaga el gènere del substantiu només a l'article i al primer adjectiu. Després, el sistema continua traduint mot per mot.
- (e) Nos ofreció un postre muy bueno  $\rightarrow$  *\*Ens va oferir unes postres molt bo*: funciona incorrectament perquè no reconeix la seqüència completa **art-subst-adv-adj**; en canvi, sí reconeix la seqüència més curta (1) **art-subst** i propaga el gènere del substantiu només a l'article. Després, el sistema continua traduint mot per mot.

El sistema només ha usat dues seqüències (1: **art-subst** i 2: **art-subst-adj**) per a intentar fer la concordança.

5. **Solució 1 ( $n$  mots)**: El traductor professional tradueix un text de  $n$  mots per  $0,05 \times n$  euros. El sistema de traducció automàtica el tradueix per  $0,03 \times n$  i corregir-lo costa  $0,10 \times (10/100) \times n$ , és a dir,  $0,01 \times n$  euros perquè només 10 de cada 100 mots són incorrectes. Per tant, el sistema semiautomàtic costa només  $(0,03+0,01) \times n = 0,04 \times n$  euros, davant dels  $0,05 \times n$  euros del traductor professional.

**Solució 2 (1000 mots)**: El traductor professional tradueix un text de 1000 mots per  $0,05 \times 1000 = 50$  euros. El sistema de traducció automàtica el tradueix per  $0,03 \times 1000 = 30$  euros, i corregir-lo costa 10 euros, perquè en 1000 paraules hi ha  $1000 \times 10/100 = 100$  mots incorrectes i corregir cada un costa 0,10 euros:  $0,10 \times 100 = 10$ . Per tant, el sistema semiautomàtic costa només 40 euros, davant dels 50 del traductor professional.