

Eti_ToBI:

una eina per l'etiquetatge Cat_ToBI automàtic en Praat

Wendy Elvira-García¹, Paolo Roseano^{1,2}, Ana Ma. Fernández Planas¹, Eugenio Martínez Celdrán¹

¹Universitat de Barcelona

²Universitat Pompeu Fabra



VI Workshop sobre la prosòdia del català
Barcelona, 26 de juny de 2014

Eti_ToBI en català

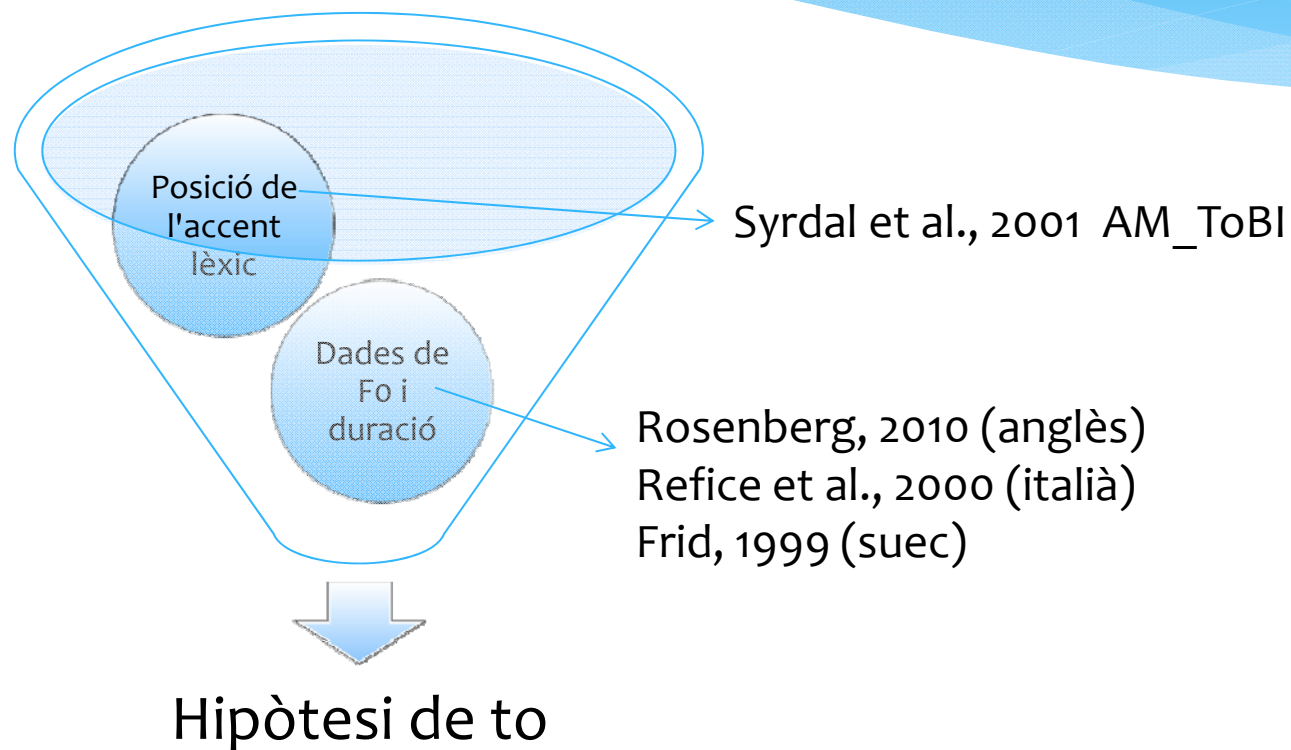


VI Workshop sobre la prosòdia del català
Barcelona, 26 de juny de 2014

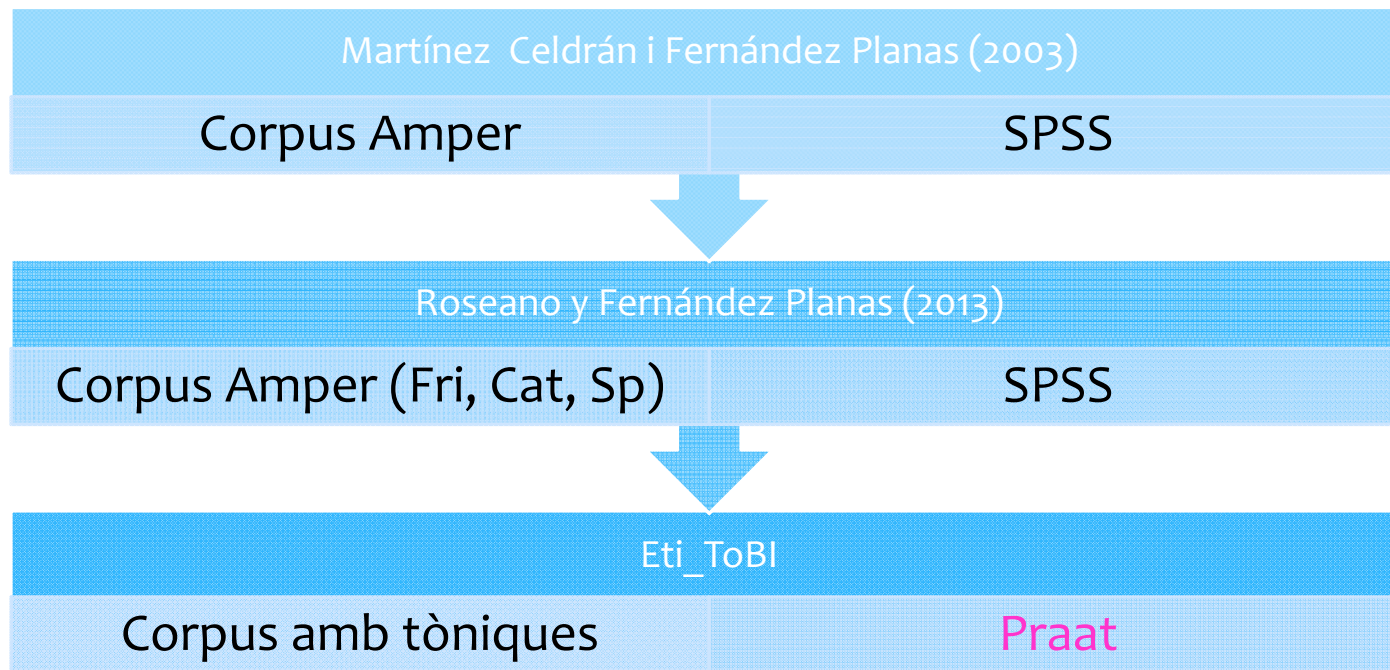
Índex

- * Introducció
 - * Els transcriptors prosòdics automàtics
 - * Antecedents de l'Eti-ToBI
- * Funcionament del sistema
 - * Passos a seguir
 - * Predicció de moviments tonals
- * Avaluació dels resultats
 - * Comprovació de detecció d'accents
 - * Test de fiabilitat

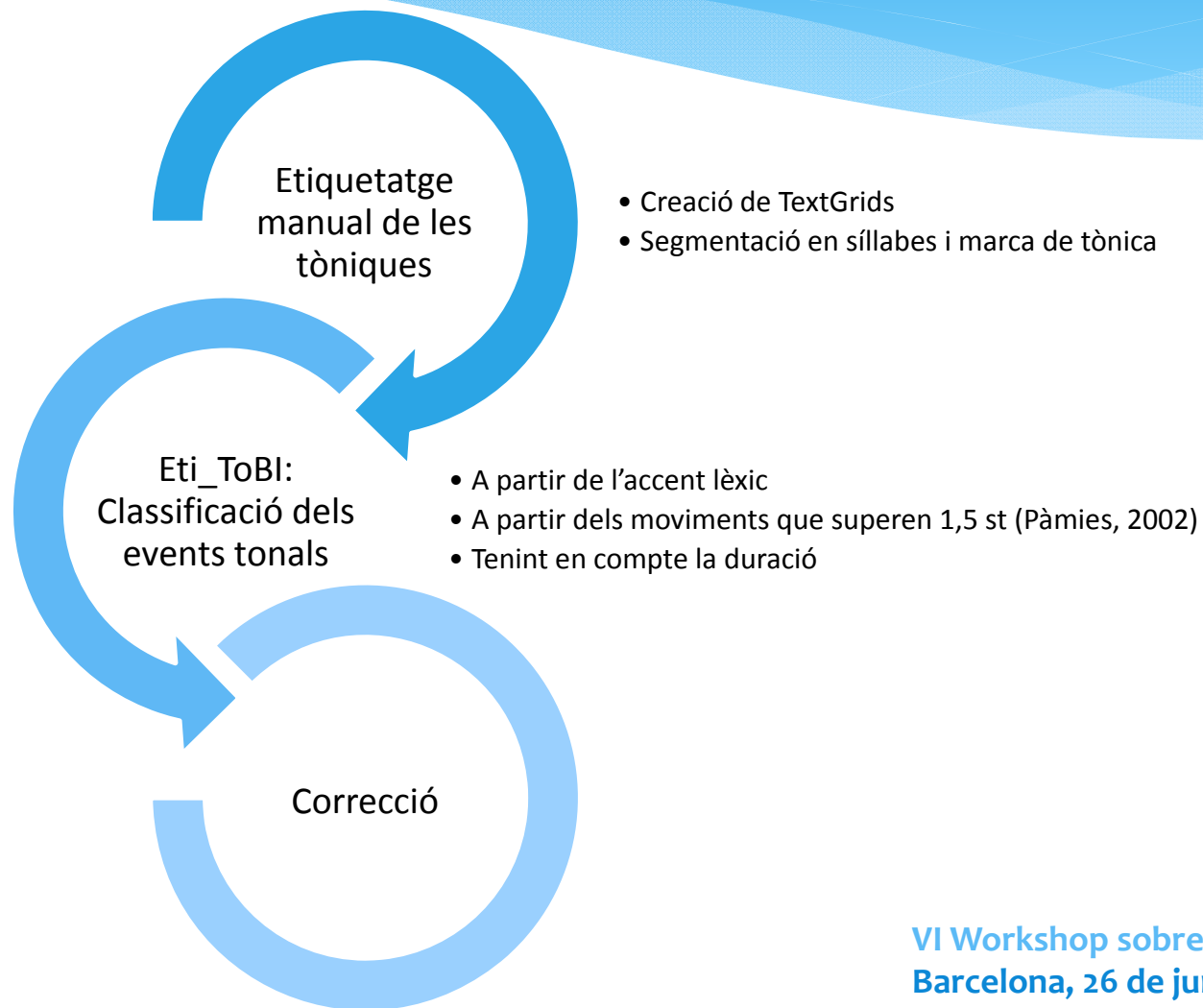
Introducció I: altres transcriptors ToBI automàtics



Introducció II: Antecedents



Com funciona?

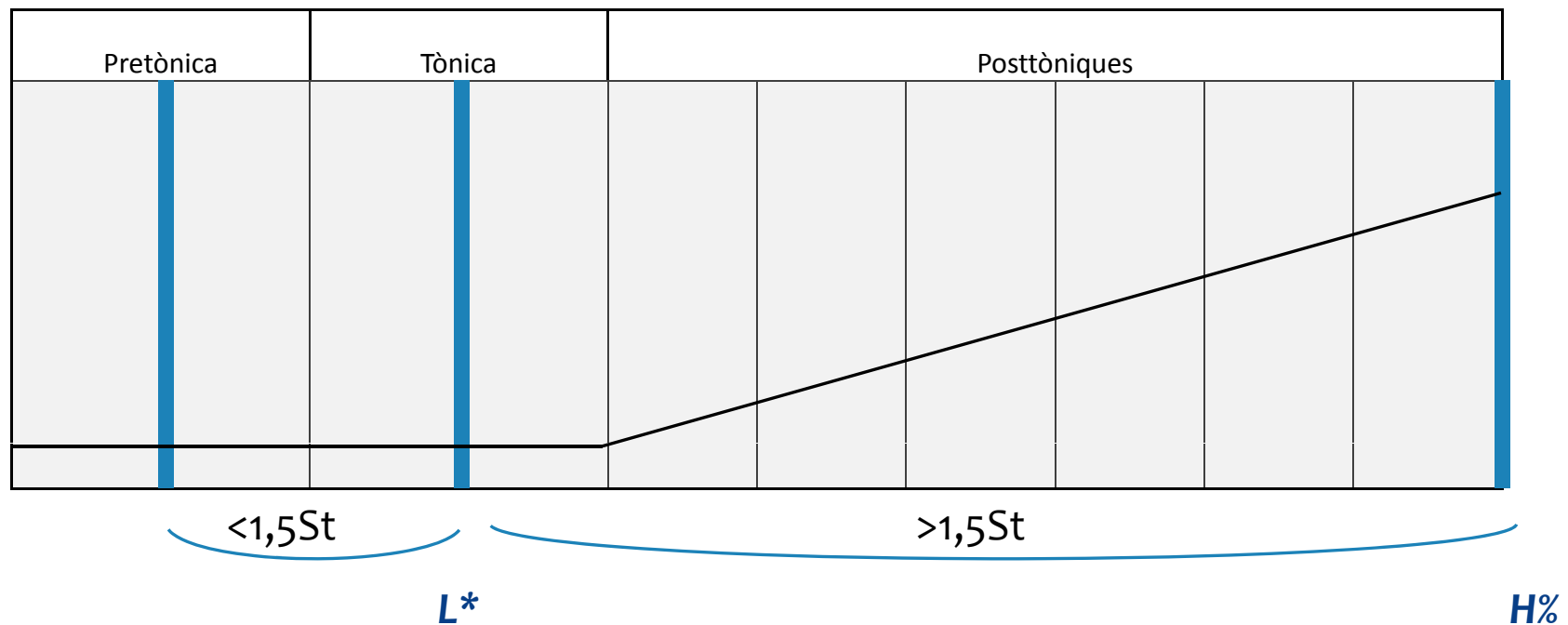


Quines dades serveixen per calcular el to?

Pretònica: Valors inicial, central i final

Tònica: Valors inicial, central i final.

Posttòniques: 6 valors¹



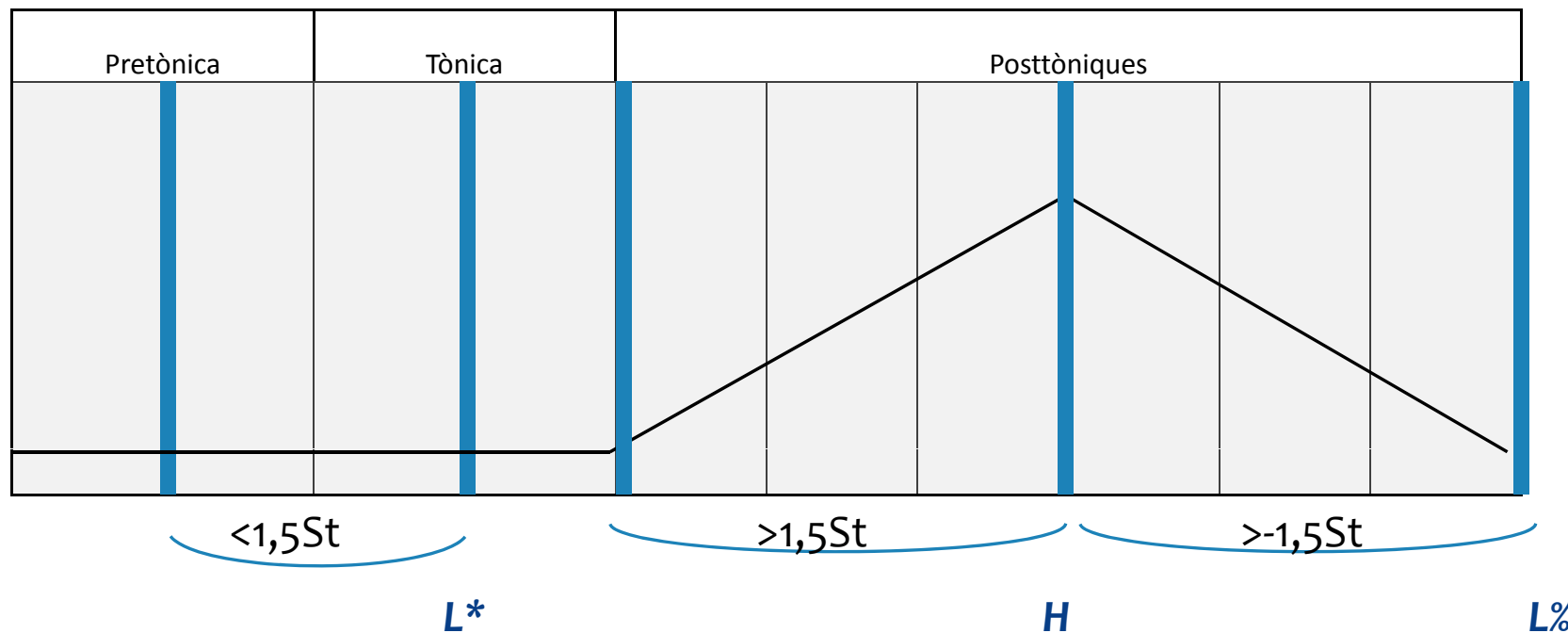
¹ En el cas de les paraules oxítones hi ha dotze valors a la tònica

Quines dades serveixen per calcular el to?

Pretònica: Valors inicial, central i final

Tònica: Valors inicial, central i final.

Posttòniques: 6 valors¹

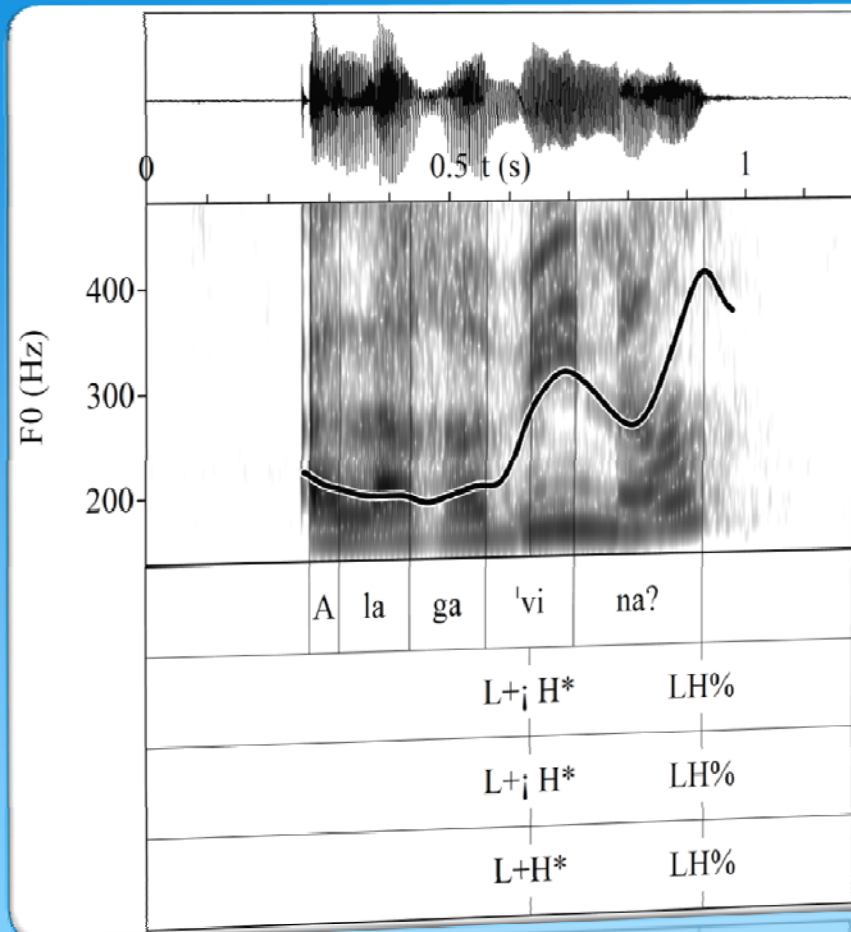


¹ En el cas de les paraules oxítones hi ha dotze valors a la tònica

55 fórmules i 26 subfórmules

```
# H+L* i desaccentuació si és una baixada des d'un punt alt que pertany a un to anterior L*+H
#Calcul·la si hi ha declinació entre l'últim target i la pretònica de la tònica actual. Si n'hi ha significa que no hi ha un target alt a la pretònica i, per tant, no es H+L*
if difpreton < 'umbralnegativo'
  etiquetatono$ = "H+L*"
  etiquetaprofunda$ = "H+L*"
  tonicaH=0
  if lengua = 3 or lengua =2
    select TextGrid 'base$'
    numeropuntosahora = Get number of points: 'tier_Tones'
    if numeropuntosahora >=1
      labeltonicaanterior$ = Get label of point: tier_Tones, numeropuntosahora
      if numeropuntosahora >1
        tpuntoanterior = Get time of point: tier_Tones, numeropuntosahora-1
        intervaloulitimotono = Get interval at time: tier_tonicidad, tpuntoanterior
        labeltonoanterior$ = Get label of point: tier_Tones, numeropuntosahora-1
        if (labeltonoanterior$ = "L*+H") or (labeltonoanterior$ = "H+(L*+H)") or (labeltonoanterior$ = "(H+L*)+H")
          inicio_target = Get start point: tier_tonicidad, intervaloulitimotono+1
          fin_target = Get end point: tier_tonicidad, intervaloulitimotono+1
          select Pitch 'base$'
          target_anterior = Get maximum: inicio_target, fin_target, "Hertz", "Parabolic"
        else
          select Pitch 'base$'
          target_anterior = Get maximum: tpuntoanterior-0.05, tpuntoanterior+0.05, "Hertz", "Parabolic"
        endif
      endif
    endif
    if numeropuntosahora <= 1
      inicio_frase = Get start point: tier_tonicidad, 2
      select PitchTier 'base$'
      target_anterior = Get value at time: inicio_frase
    endif
    #aquí si pongo con el pto 3 de la pretónica deja de ver los H+L*
    difconlaanterior = (12 / log10 (2)) * log10 ('f01pre' / 'target_anterior')
    if numeropuntosahora <= 1 and difconlaanterior < 'umbralnegativo'
      etiquetatono$ = "H+L*"
      etiquetaprofunda$ = "H+L*"
    endif
    if (numeropuntosahora > 1) and (labeltonicaanterior$ <> "H+L*") and (labeltonicaanterior$ <> "H+(L*+H)") and (difconlaanterior < 'umbralnegativo')
      etiquetatono$ = "H+L*"
      etiquetaprofunda$ = "L*+H"
    endif
  endif
endif
endif
endif
endif
```

Els etiquetatges



1
2
3

1. Etiquetatge superficial.
Etiqueta tots els moviments >1,5 i només aquests.

2. Etiquetatge profund: Hi ha una primera interpretació, es tenen en compte aspectes com la duració i els efectes de la declinació a la corva de Fo.

3. Etiquetatge normalitzat:
Substitueix etiquetes no recollides al Cat_ToBI com L+;H ;HL% per d'altres que ho són.

Comprovació de resultats

Test 1:

- * Funcionament de les fórmules
- * Reconeixement de moviments en frases reals.

Test 2:

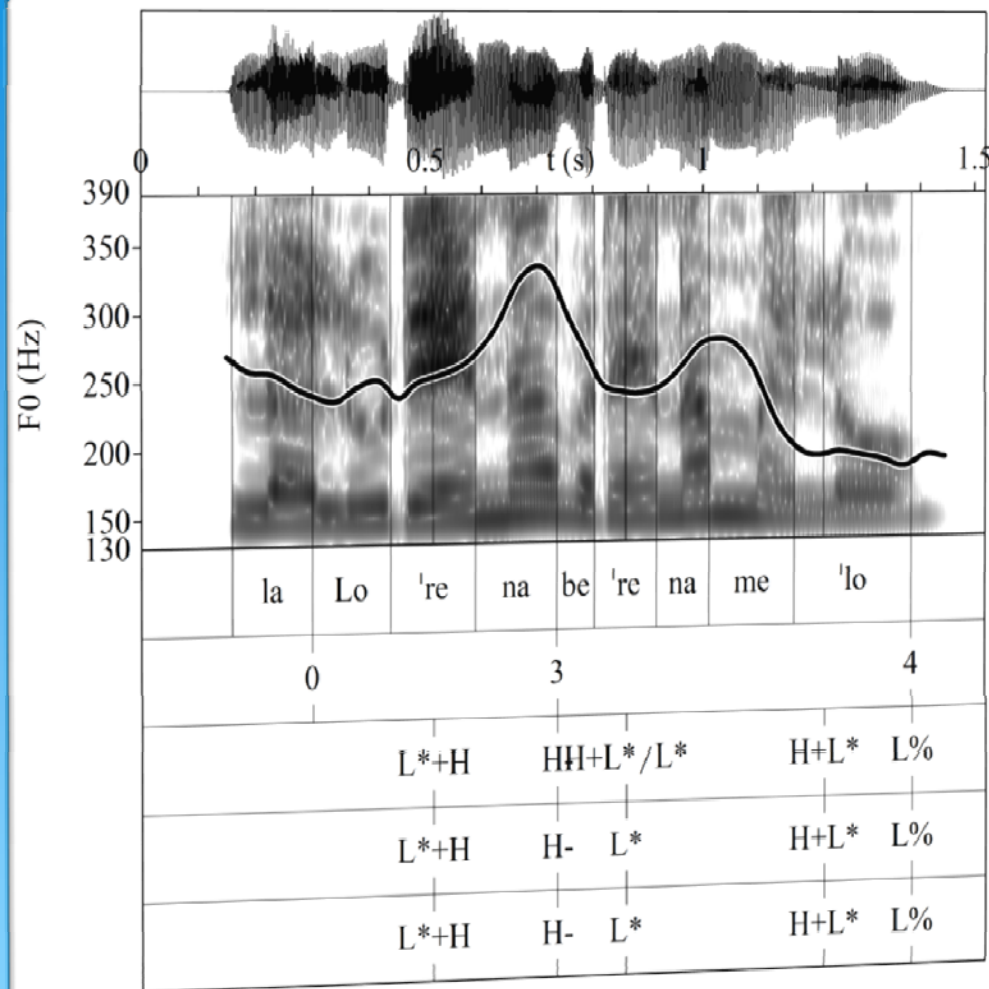
- * Test de fiabilitat de resultats (kappa)

Reconeixement configuració nuclear

Grau de reconeixement segons
sexe de l'informant

| | % encert |
|---------------|----------|
| Veü femenina | 97% |
| Veü masculina | 75% |
| | 86% |

- * Corpus ad hoc
 - * 15 configuracions nuclears
 - * Paraules oxítones i no oxítones
 - * Sonores
- * 2 informants (masculí i femení)
 - * N final: 60
- * Factors del no reconeixement:
 - * Errors a la corva de *pitch* Praat
 - * $L+H^* L!H\% \rightarrow L+H^* LH\%$



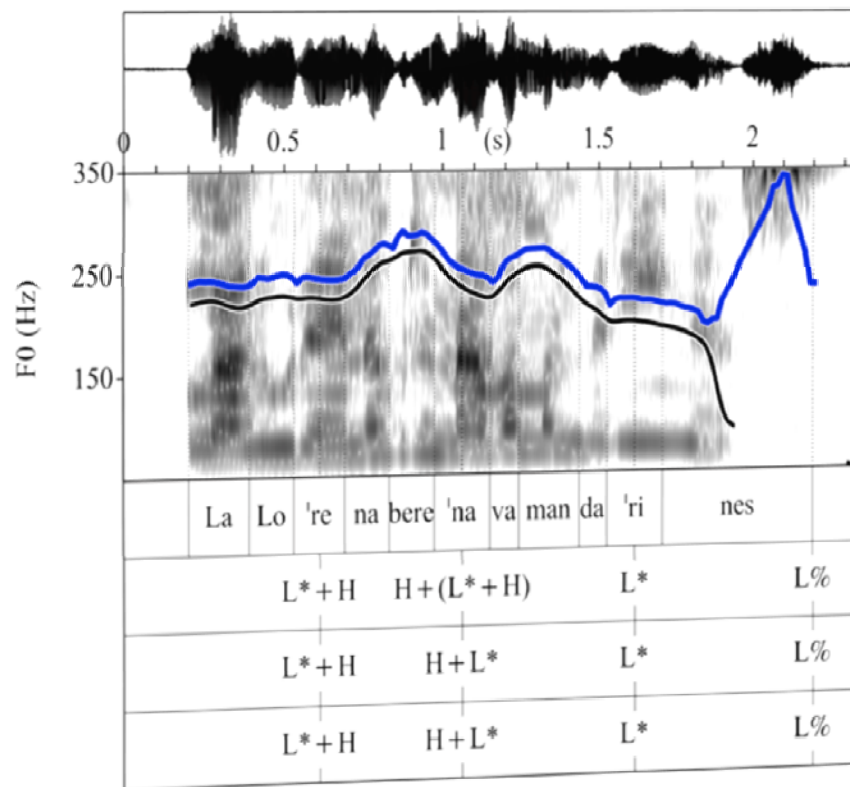
Tipus d'error 1:

manca de dades per passar dels fonètics superficials als profunds

Última paraula de IP **aguda**.

To teòric: L*

To posat per l'script: H+L*



Tipus d'error 2:

Errors de l'algoritme de detecció de Fo de Praat

- Creaky voice
- Barres d'explosió
- Sibilants

Mesures presses per minimitzar l'efecte dels errors a la corba de Fo

“Praat may then hallucinate pitches in other places that you would prefer to consider voiceless” (Boersma, 2006)

- * Mesures contra la detecció d'Fo en *creaky voice*
 - * Mesures de quantils per calcular el rang (Hirst, 2007)
- * Mesures per evitar la detecció d'Fo en fricatives i barres d'explosió
 - * Filter (**stop** Hann band): 900, 20000

Test de fiabilitat

| | Nº | % coincidència | Kappa (de Cohen) | Valoració | Kappa ponderado | Valoració |
|----------------|----|----------------|------------------|-----------|-----------------|-----------|
| Transcriptor 1 | 98 | 77.45% | 0.66 | good | 0.85 | very good |
| Transcriptor 2 | 98 | 82% | 0.752 | good | 0.779 | good |
| Transcriptor 3 | 98 | 75% | 0.661 | good | 0.756 | good |
| Transcriptor 4 | 98 | 65.45% | 0.549 | moderate | 0.57 | moderate |
| mitjana | | 74.98% | | | | |

| A la bibliografia | Nº | % coincidència |
|----------------------------------|-----|----------------|
| Escudero <i>et al.</i> (2012) | 264 | 73.635%* |
| Escudero i Aguilar (2010) | 133 | 74.49% |

* Mitjana del resultat de *pitch accent* i *boundary tone*, que apareix a l'article

Conclusions i línies de investigació futures

- * Conclusions
- * L'Eti_ToBi és el primer intent d'aplicar el sistema Cat_ToBi de manera automàtica al català que funciona sense un corpus específic.
- * Ha demostrat tenir una eficàcia igual als transcriptors humans.
- * Passes següents
- * Aplicació a altres llengües.
- * Incorporació de millores als algoritmes que permetin reduir els casos ambigus.
- * Tests d'accent prenuclear

Referències

- * Beckman, M. E., i Hirschberg, J. (1994). The ToBI annotation conventions. Ohio State University. Unpublished manuscript, Ohio State University and AT&T Bell Telephone Laboratories.
- * Boersma, Paul i Weenink, David (2014). Praat: doing phonetics by computer [Computer program]. Versió 5.3.70, extret el 02 d'abril de 2014 de <http://www.praat.org/>
- * Cohen, J. A coefficient of agreement for nominal scales. Educational and Psychological Measurement, 1960, 20, 37-46.
- * Frid, J. (1999). «An environment for testing prosodic and phonetic transcriptions», dins *Proceedings of ICPhS 99* (pp. 2319-2322). University of California: Berkeley.
- * Hirst, D. (2007) Detect_pitch.praat [praat script]
- * Martínez Celadrán, E. i Fernández Planas, A.M. (2003): «Taxonomía de las estructuras entonativas de las modalidades declarativa e interrogativa del español peninsular según el modelo AM en habla de laboratorio», dins E. Herrera i P. Martín (eds.). *La tonía: dimensiones fonéticas y fonológicas*, El Colegio de México: Mèxic, pp. 267-294.
- * Pamies Bertrán, A; A. M. Fernández Planas; E. Martínez Celadrán; A. Ortega Escandell i M. C. Amorós Céspedes (2002): «Umbrals tonals en espanyol peninsular» dins *Actas del II Congreso de Fonética Experimental: Sevilla*, pp. 272-278.
- * Prieto, P. (2014). «The Intonational Phonology of Catalan», dins Sun-Ah Jun (ed.). *Prosodic Typology 2. The Phonology of Intonation and Phrasing*. Oxford University Press: Oxford, pp. 43-80.
- * Roseano, P. (2012). *La prosòdia del friül. en el marc de l'Atles Multimèdia de Prosòdia de l'Espai Romànic*, tesi doctoral, Universitat de Barcelona.
- * Rosenberg, A. (2010). «AutoBI-a tool for automatic ToBI annotation», dins *Actes INTERSPEECH 2010, 11th Annual Conference of the International Speech Communication Association*, Makuhari, Chiba: Japan, pp. 146-149.
- * Syrdal, A. K.; Hirschberg, J.; McGory, J.; Beckman, M. (2001). «Automatic ToBI prediction and alignment to speed manual labeling of prosody». *Speech Communication* 33 (2001) 135-151.

Agraïments

- * Al projecte de recerca [FFI2012-35998](#) fons públics, obtinguts en convocatòria competitiva.
- * A l'Ajut al Personal Investigador en Formació de la Universitat de Barcelona.

Gràcies

Mila esker

Gracias

Podeu trobar més scripts a:

<http://stel.ub.edu/labfon/ca/scripts-de-praat-praat-scripts>

O podeu sol·licitar un script personalitzat al Laboratori de Fonètica UB