

**TRANSCRIPCIÓ FONÈTICA I FONOLÒGICA DE L'ENTONACIÓ:
UNA PROPOSTA D'ETIQUETATGE AUTOMÀTIC**

**PHONETIC AND PHONOLOGICAL TRANSCRIPTION OF
INTONATION: A PROPOSAL OF AUTHOMATIC LABELING**

PAOLO ROSEANO

Universitat de Barcelona i Universitat Pompeu Fabra
paolo.roseano@ub.edu

ANA MA. FERNÁNDEZ PLANAS

Universitat de Barcelona
anamariafernandez@ub.edu

Artículo recibido el día: 14/05/2013

Artículo aceptado definitivamente el día: 07/06/2013

Estudios de Fonética Experimental, ISSN 1575-5533, XXII, 2013, pp. 275-332

RESUM

Aquest article té dos objectius principals. El primer és proposar un sistema de transcripció fonètic de l'entonació que sigui entenedor i que es pugui relacionar fàcilment amb un sistema de transcripció fonològica d'ús comú. El segon és presentar un conjunt d'eines informàtiques que proporcionen aquesta transcripció de l'entonació en el marc d'un projecte en concret (AMPER). La primera part de l'article conté unes reflexions prèvies sobre els sistemes de transcripció en general. La segona secció vol enfocar les característiques (així com les problemàtiques) de la transcripció fonètica i fonològica en un àmbit –el segmental– que té dos segles d'història. A la secció següent, es compara la transcripció fonètica i fonològica segmental amb aquella entonativa, que té menys tradició i, per tant, es troba en un estat menys evolucionat, malgrat l'atenció que el tema ha rebut al llarg de les últimes dècades. A la quarta part de l'article, passem a descriure les eines informàtiques que proposem per tal de proporcionar un etiquetatge prosòdic automàtic.

Paraules clau: *entonació, transcripció, anotació, etiquetatge, model mètric autosegmental, ToBI.*

ABSTRACT

This paper has two main objectives. The first is to propose a system of phonetic transcription of intonation that is understandable and that can be easily related with a widely used phonological transcription system. The second is to present a set of tools that provides this transcription of intonation in the context of a specific project (AMPER). The first part of the article contains some reflections on transcription systems in general. The second part focuses on the features (and issues) of the phonological and phonetic transcription of segments, a tradition that has been in use for over two centuries. In the following section, we compare the phonological and phonetic transcription of segments with the transcription of intonation, far less common and therefore less evolved despite the attention that it has received in recent years. In the fourth part of the article, we describe some tools that can provide the automatic transcription of the prosodic features of an utterance based on its acoustic features.

Keywords: *intonation, transcription, annotation, labeling, autosegmental-metrical model, ToBI.*

1. SISTEMES DE TRANSCRIPCIÓ EN L'ÀMBIT CIENTÍFIC

Abans de proposar un sistema de transcripció de l'entonació, és oportú intentar recordar què és un sistema de transcripció, no només en l'àmbit entonatiu o lingüístic, sinó en termes més generals en l'àmbit científic. Molt en general, un sistema de transcripció és un sistema de representació simbòlica, és a dir un conjunt de regles i de grafemes que permeten de representar amb símbols una realitat física. En aquesta definició, hi entrarien tots els sistemes que s'empren a les ciències naturals. La química, per exemple, té un sistema de representació dels fenòmens físics de què s'ocupa –encara que els químics no l'anomenin *sistema de transcripció* sinó *fòrmules*. Així, doncs, una realitat física com una molècula troba una representació simbòlica en una fórmula química. Tots els que hem cursat l'educació obligatòria sabem, per exemple, que un àtom d'hidrogen es representa amb el símbol H, que és un signe arbitrari que simbolitza una realitat física concreta: un electró donant voltes al voltant d'un protó. Per contra, el símbol He indica una realitat física incontrovertiblement diferent: dos electrons fent voltes al voltant d'un nucli format per dos protons i dos neutrons.

Una pregunta clau per entendre el sistema de transcripció de la realitat física que empren els químics és: concretament, com es passa de la realitat a la transcripció? Imaginem-nos un químic en el seu laboratori que rep un pot ple de gas amb l'encàrrec de determinar si es tracta d'hidrogen (H) o d'heli (He). Com pot decidir el químic si al pot que li han entregat ha de posar-hi una etiqueta que diu H o bé una etiqueta que diu He? Bàsicament, les estratègies per decidir quina etiqueta posar al pot són dues: una basada en la percepció individual i l'altra basada en el mètode científic. Un químic experimentat sap que si obre el pot, ensuma el gas i no passa res pot posar-hi l'etiqueta H, mentre que si després d'ensumar el gas comença a parlar amb la veu de l'Ànc Donald hi haurà de posar l'etiqueta He, perquè l'heli té aquest efecte temporal. Així, basant-se només en la seva percepció individual i la seva experiència, pot arribar a etiquetar el pot amb un grau suficient de certesa. L'altre mètode, menys divertit, més feixuc però científic, preveu una sèrie de proves amb mesuraments repetits. Per exemple el químic pot exposar el contingut del pot a un camp elèctric i mesurar les línies espectrals que l'exposició elèctrica genera a l'espectre òptic. Repeteix el procediment almenys tres vegades per assegurar-se que no hi hagin errors de mesurament i, a partir d'aquelles dades numèriques, decideix si es tracta d'hidrogen o d'heli. Per exemple, com es veu a la taula 1, els mesuraments de les freqüències de les línies espectrals de la hipotètica mostra que el nostre químic està analitzant indica que la mostra està formada per heli pur. Òbviament, a l'hora de publicar els resultats de la seva anàlisi en una revista científica, procurarà explicar amb detall quin mètode d'anàlisi ha fet servir per determinar si es tractava d'hidrogen o d'heli.

Freqüència 1	438.793 nm
Freqüència 2	443.755 nm
Freqüència 3	447.148 nm
Freqüència 4	471.314 nm
Freqüència 5	492.193 nm
Freqüència 6	501.567 nm
Freqüència 7	504.774 nm
Freqüència 8	667.815 nm

Taula 1. Resultats del mesurament de les línies espectrals de la hipotètica mostra de gas per analitzar.

El procediment que porta a etiquetar el contingut de la mostra com He es pot resumir en dues fases. La primera és el mesurament de determinades característiques físiques susceptibles de ser transformades en números. La segona és l'atribució, a partir del conjunt característic de línies espectrals, d'un símbol convencional que identifica la realitat física analitzada. Aquest procediment s'il·lustra a la figura 1.

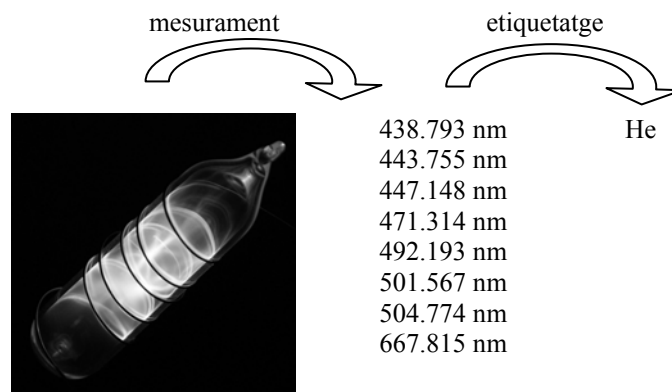


Figura 1. Sinopsi del procés de transcripció en química.

Aquest procediment que porta el químic a posar l'etiqueta H o He es recolza en una sèrie de pressuposits, entre els quals dos dels més destacats són els següents:

1. El mètode d'anàlisi és científic, és a dir que permet d'obtenir un coneixement de la realitat de manera *objectiva, fiable i verificable*.
2. El conjunt de regles i símbols que s'empren per descriure la realitat física són adequats per fer-ho. És a dir, entre d'altre coses, permeten d'obtenir una representació de la realitat que sigui suficientment *senzilla i pràctica*, però *sense perdre aspectes importants*.

En relació amb el nivell de detall de la representació simbòlica de la realitat, és imprescindible destacar com no hi ha només un sol nivell adequat. Segons les finalitats determinades, pot ser suficient una representació molt sintètica o bé es pot requerir un nivell de detall més elevat. La química ens ofereix un exemple en aquest sentit. De fet, hi ha diferents nivells de fórmules químiques, que amb un cert grau d'aproximació podem comparar amb la diferència entre una transcripció fonètica estreta i ampla. La fórmula molecular, que és la més sintètica i senzilla de totes, indica els àtoms que componen la substància i la quantitat exacta de cadascun d'ells en una molècula. A l'altre extrem de l'escala de detall tenim la fórmula desenvolupada, també anomenada fórmula estructural, que és una representació de la molècula que indica quin àtom s'uneix amb quin altre, i mitjançant quin enllaç. Per exemple, la fórmula molecular C_4H_{10} indica qualsevol molècula formada per 4 àtoms de carboni i 10 d'hidrogen, per tant es tracta de butà. Aquesta representació sintètica de la realitat química pot ser suficient a determinats nivells d'anàlisi, però no ho és en un pla més detallat. De fet, no hi ha només una molècula que es representa amb el simbolisme C_4H_{10} sinó que n'hi ha dues: el butà normal i l'isobutà, que difereixen entre elles per la disposició dels elements. La diferent disposició dels àtoms queda recollida a la fórmula desenvolupada o estructural (figura 2).

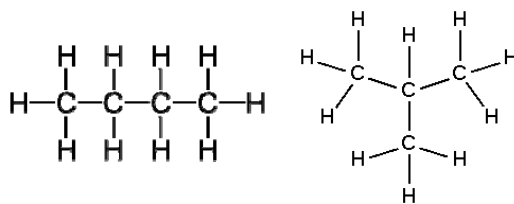


Figura 2. Fórmules desenvolupades del butà normal (a l'esquerra) i de l'isobutà (a la dreta).

Abans de passar a explorar el terreny de la transcripció en lingüística, és oportú destacar cinc característiques importants de la *transcripció* en l'àmbit de la química.

1. *Discreció*: els fets físics que s'estudien i s'etiqueten són discrets i no pas continus. Entre el fet físic H (un electró que roda al voltant d'un protó) i el fet físic He (dos electrons que giren al voltant de dos protons i dos electrons) no hi ha cap cas entremig. En la naturalesa no es dona el cas d'un electró i mig que orbiti al voltant d'un nucli de dos protons i tres quarts i un neutró coma sis.
2. *Mesurabilitat*: com que els fets físics que s'estudien són discrets, també ho són les propietats que permeten d'identificar-los. És a dir, el conjunt de freqüències de les línies espectrals que identifiquen l'heli és separat i ben diferenciat del conjunt de freqüències que identifiquen l'heli.
3. *Objectivitat*: quan un químic posa una etiqueta H o He, ho fa sobre la base del mesurament objectiu d'uns paràmetres físics (les freqüències de les línies espectrals).
4. *Universalitat*: a tot el planeta l'etiqueta H sempre indica un electró que dona voltes al voltant d'un protó. Això també significa que a tot el planeta un determinat conjunt de freqüències de les línies espectrals identifica el mateix element i no pot ser que a l'hemisferi nord un conjunt de freqüències identifiqui H i a l'hemisferi sud He. Això equival a dir que a tots als laboratoris químics del món la mateixa etiqueta descriu el mateix fet físic.
5. *Univocitat*: per indicar un electró que dona voltes al voltant d'un protó només hi ha una etiqueta, que és H. Això equival a dir que a tots als laboratoris químics del món un mateix fet físic s'etiqueta de la mateixa manera. Segons les convencions de la química, doncs, no és acceptable que el mateix àtom (identificable mitjançant un conjunt de freqüències espectrals) s'etiqueti com a H en alguns casos i com a He en altres casos.

Aquets cinc trets tenen a veure amb tres aspectes diferents del procés que porta a representar amb un símbol una realitat física. Els dos primers tenen a veure amb les característiques de l'objecte d'estudi, que són donades a priori i no es poden can-

viar. Per molt que intentem transformar un fenomen natural de discret a continu, no podem. Encara que intentem sintetitzar un àtom amb un electró i mig, serà impossible. El tercer tret és un aspecte metodològic, relacionat amb la tria del mètode (que en el cas de la química sol ser estrictament científic, mentre que en altres disciplines hi ha més tolerància per a altres mètodes). Els últims dos trets, finalment, són característiques que garanteixen la coherència del sistema de transcripció en ell mateix.

2. SISTEMES DE TRANSCRIPCIÓ DELS SEGMENTS

Ja fa alguns segles que existeix tradició de transcripcions fonètiques en el nivell segmental. Durant aquest llarg espai de temps s'han desenvolupat diferents sistemes, entre els quals en aquest article ens concentrarem només en l'IPA perquè es tracta de l'únic sistema que avui en dia es pot considerar proper a un sistema de transcripció *universal* (veg. Kemp, 1988 per a una síntesi d'altres sistemes de transcripció i la seva història). No obstant aquesta llarga tradició, la *universalitat* i *objectivitat* dels sistemes de transcripció fonètica dels segments són com a mínim problemàtiques. La mateixa International Phonetic Association recorda la dificultat de fer una transcripció molt estreta, de transcriure llengües que no són la pròpia, casos de trastorns articuladoris, o, fins fa no gaire, la dificultat de fer transcripcions informàticament, per exemple (IPA, 1999). En aquesta secció ens limitem a resumir alguns dels aspectes problemàtics en aquest sentit, i ho fem comparant el sistema de transcripció fonètica segmental amb el sistema de la química que hem descrit anteriorment i del qual hem destacat cinc trets: mesurabilitat, objectivitat, discreció, universalitat, univocitat.

Pel que fa a la *mesurabilitat*, no sembla haver-hi cap problema substancial: els fets acústics són fet físics i, per tant, es poden mesurar. El tema de l'*objectivitat* ja és més complex. Hem dit que per objectivitat entenem que quan un químic posa una etiqueta H o He, ho fa sobre la base del mesurament objectiu d'uns paràmetres físics (les freqüències de les línies espectrals). Els fonetistes fan el mateix? No tots i no sempre. Hi ha estudis acústics molt precisos que ho fan, però hi ha també estudis en què l'etiquetatge és es duu a terme més aviat d'oïda.

En relació amb la *discreció*, trobem la primera diferència important entre l'objecte d'estudi d'una disciplina com la química i el de la fonètica. Mentre que en el cas de la química els fets físics que s'estudien i s'etiqueten són discrets, en el cas de la fonètica són continus. En altres paraules, mentre en química les fronteres entre els

diferents objectes d'estudi són clares, en fonètica són difuses (Martínez Celdrán, 2002). El cas de les vocals és molt aclaridor en aquest sentit. Si acceptem que una vocal es pot definir, amb bona aproximació, a partir de la freqüència central dels seus dos primers formants, veiem que entre dues vocals, encara que siguin molt properes, sempre hi haurà infinites altres vocals possibles. Si agafem com a exemple la [e] i la [ɛ] i considerem els valors centrals típics dels seus dos primers formants en veu masculina, veiem que la [e] es caracteritza per $F1 = 410$ Hz i $F2 = 1890$ Hz, mentre [ɛ] es caracteritza per $F1 = 570$ Hz i $F2 = 1760$ (Carrera-Sabaté i Fernández-Planas, 2005). Entre aquests dos casos n'hi ha infinits d'altres, que es poden apropar més a un extrem ($F1 = 420$ Hz i $F2 = 1880$ Hz) o a l'altre ($F1 = 560$ Hz i $F2 = 1770$ Hz), o bé quedar-se a una distància intermèdia ($F1 = 490$ Hz i $F2 = 1825$ Hz). Així, mentre la química en té prou amb un conjunt limitat i finit de símbols, perquè limitat i finit és el conjunt d'elements químics, la fonètica en principi necessitaria infinits símbols, perquè infinits són els sons de la parla humana. Això és inviable i, encara que fos viable, amb tota probabilitat seria completament inútil, ja que un sistema de representació simbòlica que no simplifiqui la realitat no té cap utilitat per a la recerca. Aquesta simplificació consisteix en un procés de categorització, a partir del qual els objectes són agrupats en categories d'entitats semblants entre elles i, sobre tot, semblants amb un *prototip*. Un punt central per entendre els límits de l'etiquetatge fonètic en el nivell segmental és justament com s'arriba (o es va arribar en el seu moment) a definir aquestes categories d'objectes *semblants*, com s'arriba a definir el *prototip*.

El procés que porta a la definició dels prototips és –o hauria de ser– inductiu: s'analitza un conjunt suficientment gran de dades, es mesuren els paràmetres pertinents, s'efectua una anàlisi estadística i, a partir d'aquella, es proposen unes generalitzacions. Per exemple, per determinar les característiques acústiques de les vocals [e] i [ɛ] del català barceloní, s'enregistren diferents repeticions de paraules que contenen [e] pronunciades per diferents parlants nadius. A continuació, s'extreuen els valors centrals del primer i del segon formant de cada ítem. Aquests valors permeten de col·locar cada ítem en un diagrama de dispersió, que és un espai bidimensional en què les coordenades representen justament $F1$ i $F2$. A la figura 3 veiem com es distribueixen en aquest espai virtual les realitzacions de [e] i [ɛ] en veu masculina en barceloní segons les dades recollides per Carrera-Sabaté i Fernández-Planas (2005).

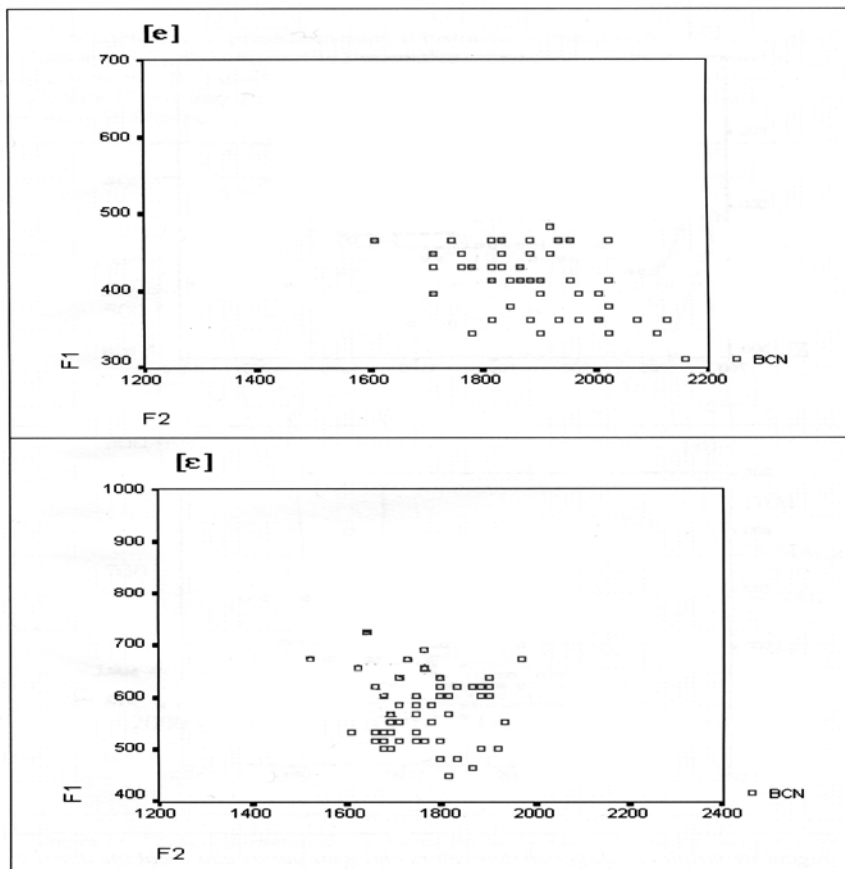


Figura 3. Distribució d'algunes realitzacions de [e] i [ɛ] en veu masculina en barceloní (gràfica estreta i adaptada de Carrera-Sabaté i Fernández-Planas, 2005: 67).

A partir de les dades que s'han il·lustrat a la figura 4, es pot proposar un prototip de [e] i [ɛ] en veu masculina en barceloní. Els dos prototips són caracteritzats per la mitjana aritmètica dels valors de F1 i F2 trobats, així que per a [e] tenim F1 = 410 Hz i F2 = 1890 Hz, mentre per a [ɛ] tenim F1 = 570 Hz i F2 = 1760 Hz (valors calculats a partir de Carrera-Sabaté i Fernández-Planas, 2005: 21-44). El prototip, doncs, es configura com a una *generalització estadística de les dades empíriques*.

A més, a partir de la distribució de les dades en el pla bidimensional, es pot delimitar l'àrea a l'interior de la qual recauen totes les realitzacions de cada vocal. Aquesta àrea, que s'anomena *camp de dispersió*, es pot delimitar amb un polígon (però també són comunes les delimitacions fetes amb una forma el·líptica). A la figura 4 proposem els camps de dispersió, en forma de polígon, de totes dues vocals.

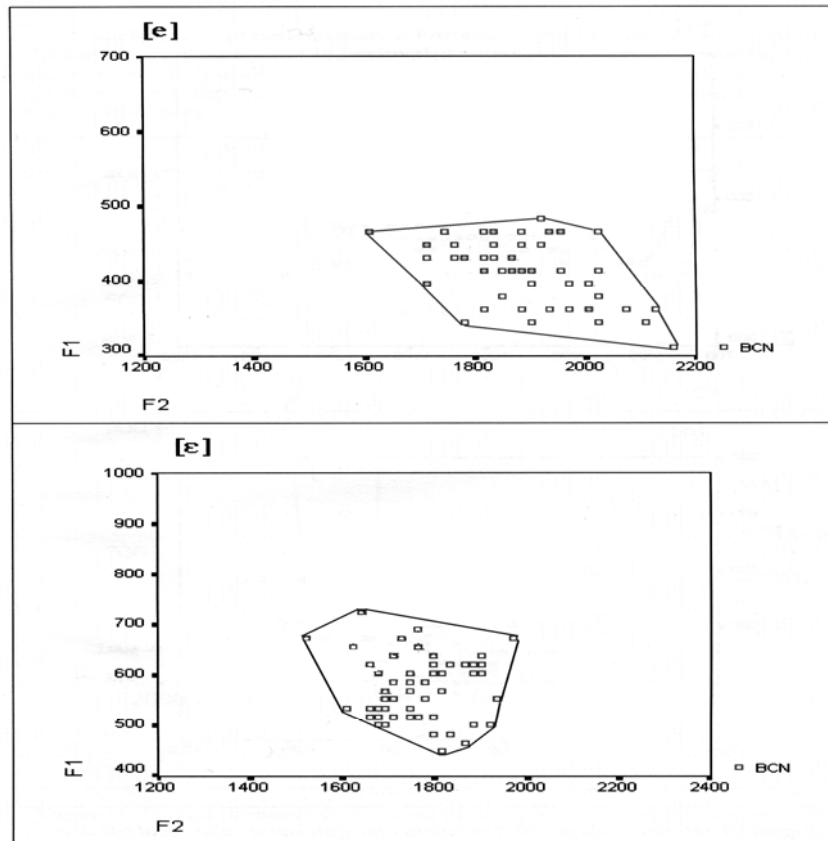


Figura 4. Camps de dispersió de les realitzacions de [e] (adalt) i [ɛ] (abaix) en veu masculina en barceloní (a partir de Carrera-Sabaté i Fernández-Planas, 2005: 67).

Arribats a aquest punt, una operació reveladora és la comparació dels camps de dispersió de dues vocals al mateix dialecte. Si sobreposem –en un pla que tingui els mateixos eixos– els camps de la figura 4, obtenim el resultat que s’il·lustra a la figura 5. El que es nota és que hi ha una intersecció dels camps de dispersió. Això significa que hi ha casos en què [e] i [ɛ], en un mateix dialecte, tenen les característiques acústiques molt semblants. Aquesta observació ens fa recordar que, quan hem descrit les característiques d’un sistema d’etiquetatge com el de la química, hem subratllat que un etiquetatge ideal és *unívoc*, en el sentit que un mateix fet físic s’etiqueta sempre de la mateixa manera. En altres paraules, hem dit que en un etiquetatge com el de la química no és acceptable que el mateix fet físic s’etiqueti com a H en alguns casos i com a He en altres casos. Per contra, la figura 5 demostra que en l’àmbit de la fonètica sí que és comunament acceptat que un mateix fet físic (en aquest cas, per exemple, una vocal que tingui F1 = 470 i F2 = 1830) es pugui etiquetar tant com a [e] que com a [ɛ]. Aquesta flexibilitat del sistema d’etiquetatge fonètic és conseqüència de la diferent naturalesa dels fenòmens observats, que són discrets en química i continus en fonètica.

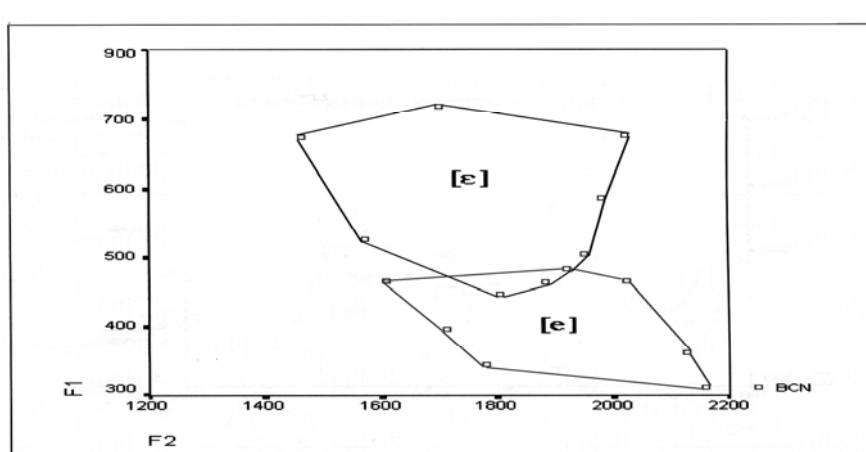


Figura 5. Superposició dels camps de dispersió de les realitzacions de [e] i [ɛ] en veu masculina en barceloní (a partir de Carrera-Sabaté i Fernández-Planas, 2005: 67).

Fins aquí hem descrit com s’arriben a definir els valors prototípics i el camp de dispersió d’una vocal *en un dialecte* en concret. Si s’amplia l’anàlisi a més d’un

dialecte, els resultats poden ser sorprenents, almenys aparentment. Per exemple, si considerem les dades de Carrera-Sabaté i Fernández-Planas (2005) per a les vocals del mallorquí i calculem els valors prototípics de les dues vocals que ens interessen, veiem que per a [e] tenim $F1 = 560$ Hz i $F2 = 1780$ Hz, mentre per a [ɛ] tenim $F1 = 690$ Hz i $F2 = 1670$ Hz (valors calculats a partir de Carrera-Sabaté i Fernández-Planas, 2005: 21-44). El fet que impressiona és que els valors prototípics de la [e] del mallorquí són pràcticament idèntics als valors de la [ɛ] del barceloní (que eren: $F1 = 570$ Hz i $F2 = 1760$ Hz). Si comparem els camps de dispersió de les vocals en els dos dialectes (figura 6), es confirma que la [e] del mallorquí i la [ɛ] del barceloní se sobreposen de manera quasi completa, alhora que els camps de dispersió de la [e] del barceloní i del mallorquí mantenen un marge de seguretat.

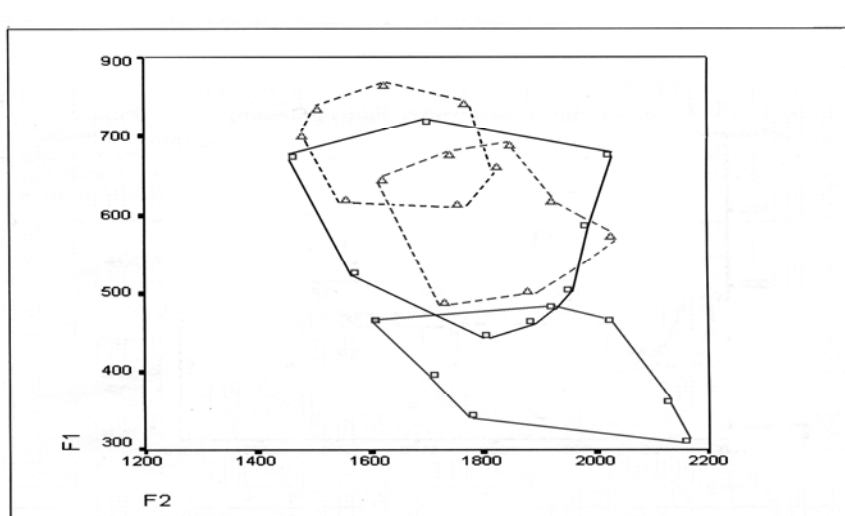


Figura 6. Superposició dels camps de dispersió de les realitzacions de [e] i [ɛ] en veu masculina en barceloní (en línies contínues superior i inferior, respectivament) i en mallorquí (en línies discontinues superior i inferior, respectivament) (a partir de Carrera-Sabaté i Fernández-Planas, 2005: 67).

Aquesta observació ens fa recordar que, quan hem descrit les característiques d'un sistema d'etiquetatge com el de la química, hem subratllat que un etiquetatge ideal

és *universal*, en el sentit que una etiqueta indica sempre una (i només una) realitat física. El cas de la comparació del vocalisme mitjà anterior barceloní i mallorquí, per contra, permet de destacar que en l'àmbit de la fonètica:

1. Una mateixa etiqueta pot indicar dos fets físics (en aquest cas dos conjunts de formants) diferents. Per exemple, com hem vist, en barceloní l'etiqueta [e] indica prototípicament una vocal que té $F1 = 410$ Hz i $F2 = 1890$ Hz, mentre en mallorquí indica un prototip ben diferent, és a dir una vocal amb $F1 = 560$ Hz i $F2 = 1780$ Hz.
2. El mateix fet físic s'etiqueta de manera diferent en dialectes diferents de la mateixa llengua. Per exemple, una vocal amb $F1 = 565$ Hz i $F2 = 1770$ Hz s'etiqueta com a [ɛ] en barceloní, però com a [e] en mallorquí.

Aquesta diferència d'ús d'una mateixa etiqueta no es troba només en l'anàlisi del català i els seus dialectes, sinó que és un fet conegut també a nivell interlingüístic. En aquest sentit es pot recordar l'estudi clàssic de Ferrari Disner (1983), que va ser un dels primers a evidenciar, mitjançant una comparació dels valors formàntics de les vocals angleses i daneses a) que el mateix símbol IPA podia indicar fets físics diferents a les dues llengües (i.e., els símbols IPA *no són unívocs*), i b) que el mateix fet fonètic pot tenir representacions simbòliques –és a dir etiquetes– diferents segons la llengua (és a dir, que els símbols IPA *no són universals*, malgrat la seva pretensió¹) (vg. figura 7).

¹ En realitat, la universalitat del sistema de transcripció IPA l'hem de trobar en la descripció de cada categoria dins la pròpia llengua. Per exemple: [i] serà sempre, en tots els sistemes, el símbol per a la vocal alta i anterior. Però la comparació entre diferents sistemes pot evidenciar que la categoria alta i anterior no correspon a les mateixes evidències físiques, com es comprova a la figura 8.

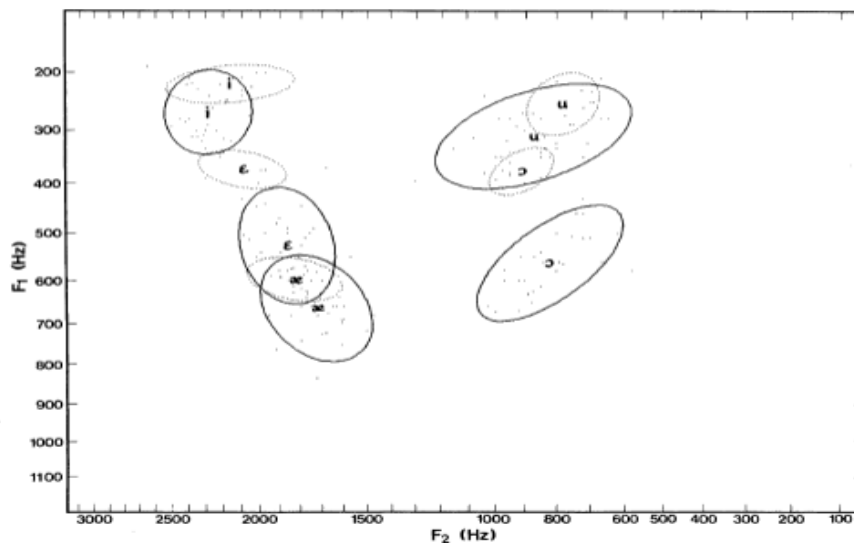


Figura 7. Superposició dels camps de dispersió de les vocals angleses i daneses (Ferrari Disner, 1983: 51).

Els problemes amb la univocitat i la universalitat que hem destacat a les pàgines anteriors deriven del fet que cada etiquetatge fonètic té un biaix fonològic. De fet, quan etiquetem un corpus, ¿com decidim si una vocal que tingui $F1 = 570$ i $F2 = 1760$ s'etiqueta com a [e] o com a [ɛ]? Al cap i a la fi, dependrà del dialecte/llengua que estem etiquetant. Així, si es tracta de català mallorquí l'etiquetem com a [e], però si estem transcrivint dades de barceloní emprarem el símbol [ɛ]. Això passa perquè tant el barceloní com el mallorquí tenen un contrast fonològic entre /e/ i /ɛ/, que però es manifesta fonèticament de manera diferent a les dues varietats: si en barceloní es manifesta –pensant en termes de vocals cardinals– com a [e] vs. [ɛ], en mallorquí podríem dir que es manifesta com a [ɛ] vs. [æ]. La tendència a etiquetar fonèticament de manera diferent en els dos dialectes una vocal que tingui $F1 = 570$ i $F2 = 1760$ es deu, per tant, a la voluntat de mantenir explícit el contrast i, alhora, mantenir evident la unitat fonològica de la llengua. Si, per contra, diguéssim que en barceloní hi ha un contrast fonològic entre /e/ i /ɛ/, mentre que en mallorquí hi ha un contrast entre /ɛ/ i /æ/, estariem donant una descripció fonològica més adherent a la realitat fonètica, però perdríem de vista la unitat del sistema lingüístic.

Abans de passar als sistemes d'etiquetatge suprasegmental, és oportú resumir el que hem anat veient en relació amb l'etiquetatge segmental. Recordem que en relació amb l'etiquetatge de la química havíem destacat cinc trets: mesurabilitat i discreció dels fenòmens estudiats, objectivitat del procés d'etiquetatge, universalitat i univocitat de les etiquetes. Ara bé, en l'àmbit de la fonètica segmental els fenòmens estudiats sí són mesurables, però no són discrets. El procés d'etiquetatge pot ser objectiu (i.e. basat en la medició de fets físics), però no sempre ho és. Les etiquetes, finalment, no són ni universals ni unívokes a causa de dos fets: a) els límits de les categories són difusos, b) hi ha un biaix fonològic. Finalment, cal destacar que existeixen uns quants programes informàtics que proporcionen transcripcions fonètiques (per a un resum dels avenços en aquest sector, vg. Cucchiari i Strik, 2003), però cap d'ells està pensat per funcionar amb més d'una llengua: tots estan concebuts per transcriure'n una de sola llengua.

3. SISTEMES DE TRANSCRIPCIÓ DELS SUPRASEGMENTS

La transcripció dels aspectes suprasegmentals de la parla –i ara ens concentrarem inicialment només en l'entonació– té una tradició molt menys llarga. Segurament un dels efectes d'aquesta falta de tradició és l'existència de tota una sèrie de models de transcripció, cap dels quals ha arribat a tenir una acceptació comparable a la que té l'alfabet IPA en el nivell segmental. En aquesta secció intentarem analitzar els diferents sistemes de transcripció suprasegmental a partir dels cinc trets que hem destacat anteriorment: mesurabilitat i discreció de l'objecte d'estudi, objectivitat del mètode d'anàlisi, universalitat i univocitat del sistema d'etiquetatge.

Tal com passa a l'àmbit segmental, els fets acústics que s'estudien en l'àmbit suprasegmental –bàsicament, F0– són *mesurables* (fonamentalment en Hz i en st). En comparació amb el que hem vist per a les vocals, on el que es mesura és el valor central d'almenys dos formants, en el cas de l'entonació només considerem un valor en cada instant, la qual cosa representa un avantatge ja que, en principi, simplifica els processos de medició. Cal tenir en compte, però, que el fet que els fenòmens que s'estudien siguin *mesurables* no vol dir que siguin realment *mesurats*: durant moltes dècades, de fet, els estudis sobre l'entonació es feien d'oïda, cosa que, segons Pierrehumbert (1980: 13), és el motiu d'errors clamorosos: *transcription by ear of intonation is the source of so many gross errors in the literature that we feel it cannot be relied on.*

Pel que fa a la *discreció* de les categories, en l'àmbit de l'entonació els fets físics que s'estudien i s'etiqueten són continus i no pas discrets. Si, per exemple, volem

descriure un fet aparentment senzill com un pic d'F0, hem de fer-ho en dues dimensions: el temps (és a dir en quina posició de l'enunciat/paraula es troba el pic) i la freqüència (és a dir com d'alt és el pic, en Hz, st o altres unitats de mesura). A la figura 8 tenim tres repeticions de la mateixa frase en friülà, [la 'minima] 'la mínima', pronunciades pel mateix parlant amb la funció pragmàtica de focalització correctiva. Intuïtivament, les tres repeticions presenten el mateix patró entonatiu, caracteritzat per una pujada tonal entre la síl·laba pretònica i la tònica i una baixada entre la tònica i la posttònica. Aquesta descripció és suficient per donar una idea general, però si volem detallar més la descripció veiem que hi ha variació dels diferents paràmetres. En primer lloc, notem que hi ha variació en l'altura del pic (126 Hz, 113 Hz, 114 Hz). En segon lloc, si calculem quina variació en semitons representa el pic respecte del punt central de la vocal anterior i posterior, també notem que hi ha variació (vg. l'última tira de la figura 8). Finalment, hi ha variació també en l'alineació del pic (que indiquem amb P a la segona tira): en la primera repetició el valor màxim d'F0 a la síl·laba tònica es troba pràcticament al final de la vocal, a la segona repetició es troba després del final de la síl·laba tònica, i a la tercera repetició es troba a la primera part de la vocal tònica.

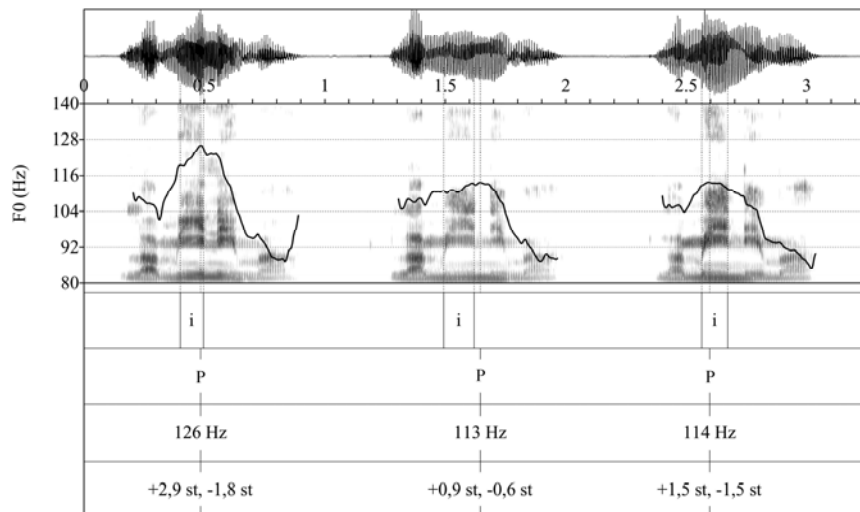


Figura 8. Oscil·lorgrama, espectrograma i corba d'F0 de tres repeticions de la frase *La minima en friülà*.

Decidir *quines* d'aquestes variacions s'han de representar en l'etiquetatge i *com* s'han de representar és una de les qüestions centrals. Una altra pregunta sobre la qual haurem de tornar és: si els fenòmens que s'estudien són continus, ¿com s'estableixen els prototips? I, un cop establerts els prototips, ¿com es decideix a quin prototip s'apropa més una corba entonativa en concret?

La tercera característica desitjable en un sistema d'etiquetatge és la seva *objectivitat*, és a dir que un etiquetatge científicament acceptable hauria de basar-se en el mesurament objectiu d'uns paràmetres físics. Els primers models d'anàlisi de l'entonació, que es van desenvolupar en una època en què les tècniques a disposició eren limitades, no poden ser considerats objectius, sinó subjectius: cada investigador escoltava una frase i en descrivia la melodia de manera intuïtiva, basant-se en la seva oïda. Així, per exemple, es van definir els models de tonema de Navarro Tomás (1944), que feia servir termes musicals com *cadència* o *anticadència* per descriure els moviments tonals que ell percebia. Amb els progressos científics, es va disposar d'eines que permetien estudiar l'entonació de manera més objectiva. És sobretot gràcies a l'anomenada escola holandesa (o IPO) que l'estudi de l'entonació va adquirir una component experimental i fonètica (Garrido Almiñana, 2003), és a dir que el sistema d'anàlisi va començar a ser *objectiu*. El model holandès va ser el primer a desenvolupar un sistema en què les corbes melòdiques dels enunciats concrets es representen amb un conjunt limitat d'etiquetes (el sistema en preveu 10) i en què aquesta transformació s'efectua per graus successius d'abstracció. Aquestes etapes són, com destaca Llisterri (2013), fonamentalment tres:

1. En la primera etapa un programa informàtic calcula la corba melòdica.
2. En la segona etapa un programa informàtic obté una corba melòdica estilitzada (és a dir simplificada) que es correspon amb la primera.
3. En la tercera etapa un programa informàtic efectua l'etiquetatge, és a dir, transforma la corba melòdica en una sèrie d'etiquetes que en representen totes les característiques lingüísticament importants.

D'una banda, a partir de l'experiència de l'IPO, s'han anat desenvolupant altres sistemes que tenen el mateix plantejament, encara que en difereixin en alguns aspectes. Entre aquests podem recordar el sistema MOMEL-INTSINT i l'IViE. Tots ells tenen en comú una sèrie de característiques, sobretot: a) que parteixen d'una anàlisi fonètica objectiva, b) que l'etiquetatge el fa un programa, la qual cosa

elimina el risc de subjectivitat de l'anàlisi, i c) que els programes en qüestió es poden aplicar a llengües diferents (Garrido Almiñana, 2003: 120). Els sistemes com l'IPO i el MOMEL-INTSINT, per tant, sembla que tenen les característiques d'*objectivitat*, *universalitat* i *univocitat*, que obtenen també gràcies a l'ús de programes informàtics.

Per altra banda, s'ha de destacar que els sistemes com el de l'IPO i el MOMEL-INTSINT no són els més utilitzats actualment. El sistema més emprat avui dia segurament és el que es recolza en el model mètric autosegmental (AM) i que es manifesta en els sistemes d'etiquetatge ToBI. La primera idea general del model AM, model bàsicament fonològic, és que a les llengües entonatives la modulació d'F0 dels enunciats constitueix un nivell independent dels altres trets fonològics. L'entonació d'una frase s'explica a partir de la successió de tons contrastius de dos nivells: H (de l'anglès *high*, alt) i L (de l'anglès *low*, baix). Aquests tons són associats a determinades síl·labes que, en principi, són mètricament prominents. Tot i això, aquesta formulació tan clara del model presenta problemes pràctics considerables en el moment en què s'aplica a l'estudi de la prosòdia d'una llengua en concret, ja que autors diferents interpreten de manera diferent les mateixes dades. Tal como observa Hualde (2003: 180), *el permitir esta multiplicidad de análisis posibles es el punto más débil del sistema métrico autosegmental en la actualidad*. En altres paraules, Hualde destaca que els sistemes d'etiquetatge basats en el model AM tenen el risc de no ser ni *objectius* ni *unívocs*.

Per reduir el risc de subjectivitat de què parla Hualde, es va crear una programa informàtic que proporciona un etiquetatge ToBI de manera automàtica (Rosenberg, 2010), és a dir un programa que opera sense la intervenció de l'investigador. Aquest programa, que s'anomena AuToBI, però, està concebut per funcionar només per transcriure l'entonació de l'anglès estàndard americà. Aquesta restricció d'ús té a veure amb un altre límit dels sistemes d'etiquetatge ToBI: encara que els fonaments del model AM siguin únics, actualment els ToBIs són molt nombrosos, pràcticament n'hi ha un per a cada llengua, i les diferències entre ells són notables. De fet, a la primera plana de la pàgina web oficial del ToBI anglès² es pot llegir que

ToBI is not [subratllat dels autors] an International Phonetic Alphabet for prosody. Because intonation and prosodic organization differ from language to language, and often from dialect to dialect within a language, there are many different ToBI systems, each one specific to a language variety and the community of researchers working on that language variety.

² <http://www.ling.ohio-state.edu/~tobi/>

Com a conseqüència d'aquesta fragmentació del model AM en diferents sistemes d'etiquetatge ToBI, hi ha una falta d'*universalitat*. Mentre en l'àmbit de la química a tot el planeta l'etiqueta H sempre indica un electró que dóna voltes al voltant d'un protó, en l'àmbit de l'entonació la mateixa etiqueta pot indicar fets acústics radicalment diferents als sistemes ToBI de diferents llengües. Un exemple d'aquest tipus d'incongruència –almenys aparent– es presenta en les figures 9 i 10. A ambdues, hi apareix la paraula [la 'li:ɹa:] pronunciada per un parlant d'espanyol argentí de Buenos Aires (figura 10) i per un parlant de friülà septentrional d'Agrons (figura 11). Les corbes entonatives presentades a les dues figures són extremadament semblants des del punt de vista acústic: presenten una síl·laba pretònica baixa, un to circumflex en la síl·laba tònica format per un ascens i un descens d'F0 i un to baix a la síl·laba posttònica. Tot i la semblança entre les dues realitzacions fonètiques, les seves representacions fonològiques són diferents. L'enunciat espanyol, de fet, constitueix un focus contrastiu, que en la varietat argentina es caracteritza per una configuració nuclear L+H*+L L% (Gabriel *et al.*, 2010). L'enunciat friülà, en canvi, és una pregunta reiterativa que en aquesta llengua té una configuració nuclear L+ɨH* L% (Roseano *et al.*, en premsa). La diferència en l'etiquetatge fonològic de dues corbes fonèticament tan semblants s'explica a partir de les diferències en les regles d'alineament i de les diferències entre els inventaris tonals d'aquestes dues llengües.

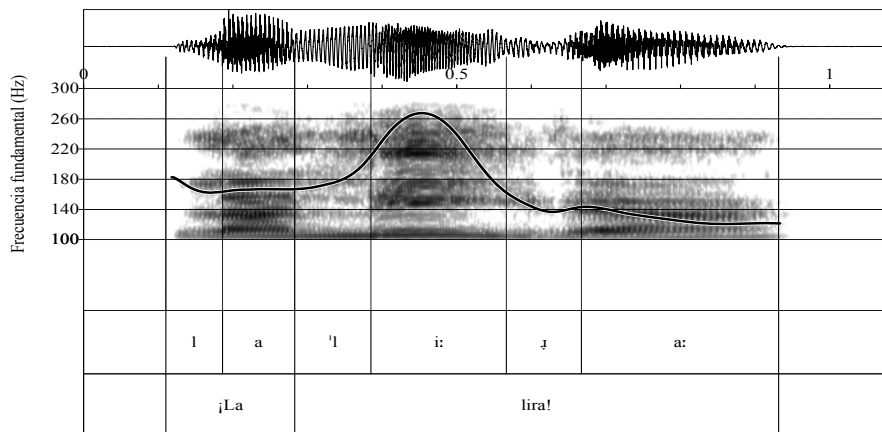


Figura 9. Oscil·lograma, espectrograma i corba entonativa de la paraula [la 'li:ɹa:] pronunciada per un parlant d'espanyol argentí de Buenos Aires.

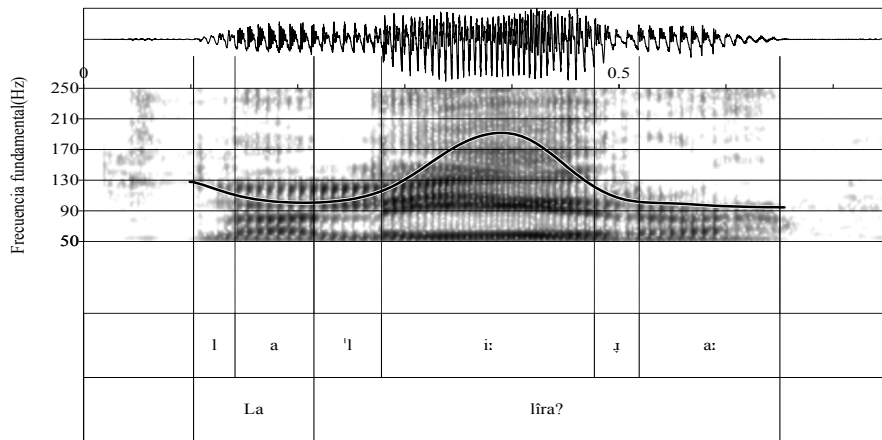


Figura 10. *Oscil·lograma, espectrograma i corba entonativa de la paraula [la 'li:ɹa:] pronunciada per un parlant de friülà septentrional d'Agrons.*

Finalment, l'últim punt problemàtic dels sistemes ToBI consisteix en com s'estableixen els prototips, i, un cop establerts, en com es decideix a quin prototip s'apropa més una corba entonativa en concret. Ambdós processos se solen basar, a la majoria dels sistemes ToBI, en el fet que els investigadors que desenvolupen un sistema ToBI són parlants nadius de la llengua que estudien (Beckman *et al.*, 2005: 41). El seu coneixement implícit de la fonologia de la llengua els permet intuir quines són les categories entonatives que contrasten i quines són les característiques que tenen, de la mateixa manera que, en el nivell segmental, el seu coneixement de la llengua els permet de saber quantes i quines són les seves vocals. En el cas dels sistemes ToBI més desenvolupats, les intuïcions fonològiques dels investigadors es comproven de manera objectiva mitjançant tests de percepció. Un cop establertes aquestes categories prototípiques, les corbes entonatives concretes s'etiqueten mitjançant mètodes d'anàlisi que, a molts escrits, no s'expliciten, però que es basen fonamentalment en la percepció per part dels investigadors i en l'anàlisi visual de la corba d'F0 extreta amb el Praat o amb altres programes. Aquesta manera d'operar fortament *language-specific* és l'avantatge i el desavantatge de cada sistema ToBI. N'és l'avantatge perquè integra al sistema de transcripció una perspectiva *fonològica* que és difícil de trobar als sistemes que presenten una aproximació més aviat fonètica. El desavantatge consisteix en la

falta d'*universalitat* de les convencions de transcripció i en la freqüent falta d'*objectivitat* del mètode d'etiquetatge.

Al llarg de les properes pàgines explorarem la possibilitat de proposar un sistema d'etiquetatge entonatiu automàtic que intenti combinar els avantatges dels sistemes fonètics com l'IPO (universalitat, univocitat, objectivitat) amb els avantatges dels sistemes de transcripció basats en el model mètric autosegmental (perspectiva fonològica, àmplia acceptació a la comunitat científica).

4. UN SISTEMA D'ANOTACIÓ PROSÒDICA (EN EL MARC DEL PROJECTE AMPER I EN EL SI DEL LABORATORI DE FONÈTICA DE LA UB)

El projecte de l'Atlas Multimèdia de Prosòdia de l'Espai Romànic (AMPER) preveu que es recullin dades prosòdiques de les varietats romàniques (Contini, 1992). Entre les dades en qüestió, destaca la presència d'un corpus anomenat fix, que inclou frases de dues modalitats (declaratives neutres i interrogatives absolutes informatives) formades per un subjecte, un verb i un complement. Subjecte i verb poden estar formats per una sola paraula fonològica (p. ex., en castellà, *La guitarra*) o per dos (p. ex. *La guitarra magrebí*). Gràcies a la combinació de paraules agudes, planes i esdrúixoles en cada posició sintàctica, en la majoria de llengües romàniques el corpus fix inclou 63 frases de cada modalitat oracional. L'anàlisi acústica d'aquestes frases es duu a terme actualment amb el programa Amper2006, elaborat per l'equip d'AMPER-Astur (López Bobo *et al.*, 2007), a partir d'una versió anterior elaborada per Antonio Romano (Romano, 1999). Amper2006 extreu tres valors d'F0 de cada vocal (F0 inicial, F0 central, F0 final), un valor d'intensitat de cada vocal i la durada de cada vocal. Aquests valors es guarden en un arxiu en format txt, com el que apareix a la figura 11. Aquest tipus de corpus facilita la tasca de comparació entre mostres sintàcticament iguals o equivalents de diferents varietats romàniques.

Les dades acústiques emmagatzemades als arxius .txt com el de la figura 12 són el punt de partida per al procés d'etiquetatge que es presenta en aquest treball i que correspon al que es porta a terme en tres fases mitjançant tres aplicacions informàtiques: AMPEReno, AMPERExtra i AMERETi. Aquest sistema d'etiquetatge automàtic, desenvolupat al Laboratori de Fonètica de la UB a partir d'una versió inicial (Martínez Celdrán i Fernández Planas, 2003), es va aplicar per primera vegada en la seva versió actual a la descripció del friülà (Roseano, 2012).

8012twta0.txt - WordPad

Archivo Edición Ver Insertar Formato Ayuda

C:\Amper\ficheros0txt1puntoHz\8012005p0p0a0u0e03x6x&x11a0.txt size: 29068
08-Feb-2011

zona	duration [ms]	energy [dB]	fo1	fo2	fo3 [Hz]
1	38	98	184	184	182
2	35	96	179	179	180
3	102	102	185	191	209
4	49	103	236	251	262
5	46	96	243	233	224
6	50	98	219	219	219
7	39	99	224	231	238
8	46	97	243	243	241
9	55	102	234	230	226
10	106	90	161	154	151
11	122	89	168	186	193

values at:
1329 1636 1944 3007 3288 3569 5392 6206 7021 7575 7963 8352 10099 10466 10834
12389 12786 13183 13936 14247 14559 15948 16312 16676 17805 18244 18682 21137
21983 22828 25286 26262 27238

Para obtener Ayuda, presione F1

Figura 11. Arxiu en format .txt generat per Amper2006 que conté els valors d' F_0 , intensitat i durada de cada vocal d'una frase del corpus fix d'AMPER.

4.1. AMPERReno

Per tal de poder analitzar fonèticament i estadísticament les dades acústiques contingudes als arxius .txt, aquestes han de ser posades en una matriu de dades que permeti, en primer lloc, de calcular les estructures prosòdiques i, en segon lloc, d'analitzar-ne la freqüència estadística. Si el procés d'inserció de les dades acústiques a la matriu es fes manualment, requeriria un temps molt llarg i segurament es donarien errors. Per aquest motiu es va decidir crear un programa que efectués de manera automàtica el transferiment de dades dels arxius .txt a la matriu de dades, com es veurà a l'apartat següent; però abans, calia canviar el nom dels arxius, perquè el primer problema de programació que es va haver de solucionar va ser el de trobar el sistema per donar ordres a l'aplicació. De fet, cap programa és capaç de saber quines dades s'han de transferir a la matriu i quines no, ja que als .txt les síl·labes tòniques no estan marcades de cap manera. Per aquest motiu es va

pensar d'utilitzar el nom de l'arxiu .txt com a *string* per donar aquestes informacions a l'aplicació.

El nom dels arxius .txt havia de contenir una sèrie d'informacions que el programa de transferència de dades necessitava per individualitzar les dades que havia d'exportar a la matriu i que no es podien deduir a partir dels codis de les frases que s'utilitzen habitualment en el marc AMPER:

1. Localitat.
2. Informant.
3. Número correlatiu de la frase.
4. Modalitat oracional.
5. Estructura sintàctica de la frase.
6. Posicions accentuals de les paraules (és a dir, si cada paraula era aguda, plana o esdrúixola).
7. Número total de síl·labes de l'oració.
8. Posició des la síl·labes accentuades en l'oració.
9. Timbre de les vocals tòniques.

Totes aquestes informacions van ser codificades en el nom de l'arxiu txt. En aquest context seria inoportú explicar de manera detallada els criteris de classificació, que de tota manera queden aclarits exhaustivament a Roseano (2012: 880-892). Un cop establerts els criteris de codificació, quedava el problema de canviar els noms dels arxius txt que havien estat posats segons la codificació d'AMPER pels noms nous, que contenen les instruccions per al programa d'extracció de dades. Aquesta operació de canvi de noms s'efectua amb un programa que ha estat creat específicament i que s'ha anomenat AMPERRenó. Mitjançant aquesta aplicació, per exemple, el nom inicial de l'arxiu wp12bwky0.txt passa a ser el nom següent wp12120p0a0eao0o0ii3X8XAF15y0.txt.

4.2. AMPERExtra

Un cop els arxius txt tenien els noms nous, s'havien d'extreure les dades importants i exportar-les a una matriu de dades en un format que en permetés l'elaboració i l'anàlisi estadística, concretament en una base de dades d'Excel.

D'aquesta operació, se n'encarrega un altre programa: AMPERExtra. Gràcies a les instruccions codificades als noms dels arxius txt, AMPERExtra exporta a un full en format .xls les dades següents:

1. Llengua.
2. Dialecte.
3. Localitat.
4. Informant.
5. Número de frase.
6. Frase amb vocals llargues o breus.
7. Tipus de paraula (aguda, plana, esdrúixola o cap) a cada accent lèxic.
8. F0 inicial, central i final de la vocal pretònica, tònica i posttònica de cada accent lèxic.
9. F0 inicial, central i final de les vocals adjacents a les fronteres.
10. Intensitat de la vocal pretònica, tònica i posttònica de cada accent lèxic.
11. Durada de la vocal pretònica, tònica i posttònica de cada accent lèxic.
12. Timbre de la vocal tònica de cada accent lèxic.
13. Modalitat oracional.
14. Número de síl·labes de l'oració.

4.3. Metodologia per a l'anàlisi entonativa de les dades (AMPEREti)

Un cop les dades eren al full de càlcul en format .xls, es podien importar a una matriu de dades d'SPSS17, un paquet estadístic que permet d'efectuar dues operacions: calcular les estructures prosòdiques a partir de les dades acústiques i analitzar la seva distribució estadística.

Pel que fa al càlcul de les estructures, s'ha creat expressament un conjunt de 856 fórmules que permeten calcular-les. Constitueixen una rutina que hem anomenat AMPEREti. Aquestes fórmules, inspirades en una primera versió d'algorismes d'etiquetatge entonatiu creats al Laboratori de Fonètica de la Universitat de Barcelona (Martínez Celdrán i Fernández Planas, 2003), determinen les següents unitats prosòdiques:

1. Les estructures entonatives superficials i profundes de cada accent lèxic (a partir de la diferència en semitons entre la freqüència de la vocal tònica i cada una de les vocals adjacents, tenint en compte el llindar psicoacústic d'1,5 st).
2. El to de frontera inicial de la frase entonativa.
3. El to de frontera final de la frase entonativa.
4. El to de frontera final de frase intermèdia, si n'hi ha.
5. Les estructures duratives de cada accent lèxic (a partir del llindar psicoacústic d'1/3 en més o en menys respecte de la durada de la vocal tònica).
6. Les estructures d'intensitat de cada accent lèxic (a partir del llindar psicoacústic del 15% en més o en menys respecte de la intensitat de la vocal tònica).

La finalitat d'AMPEREti és facilitar als investigadors una *proposta* d'etiquetatge automàtic de tipus Mètric Autosegmental tant en àmbit entonatiu com pel que fa als paràmetres de durada i d'intensitat. L'objectiu és que l'etiquetatge proporcionat per AMPEREti sigui el més semblant possible a un etiquetatge realitzat per etiquetadors humans.

4.3.1. Etiquetatge entonatiu

La primera idea general del model Mètric Autosegmental és que l'entonació d'una frase s'explica a partir de la successió de tons de dos nivells: H (de l'anglès *high*, alt) i L (de l'anglès *low*, baix). Aquests tons són associats a determinades síl·labes que, en principi, són mètricament prominents. Per tal d'evitar que l'atribució de les etiquetes H i L (en les seves combinacions que s'explicaran a continuació) no sigui un procés subjectiu, proposem aplicar un mètode d'etiquetatge basat en criteris numèrics i *objectius* perquè es pretén arribar a les estructures fonològiques a partir d'unes anàlisis fonètiques rigoroses, tal com es proposa a Martínez Celdrán i Fernández Planas (2003) i Fernández Planas i Martínez Celdrán (2003). Per etiquetar entonativament un mot, el primer pas és *mesurar* el valor de l'F0 (en Hz) de tres vocals adjacents (pretònica, tònica i posttònica). A continuació, es calculen les diferències entre les vocals en semitons i es considera significativa cada diferència superior al llindar psicoacústic establert en 1,5 semitons³ (Rietveld i Gussen-

³ L'ús dels semitons i del llindar psicoacústic és important perquè normalitza les diferències interparlants i intraparlants.

hoven, 1985; Pamies *et al.*, 2002; Fernández Planas *et al.*, 2002; Martínez Celdrán i Fernández Planas, 2003; Fernández Planas i Martínez Celdrán, 2003). Per tant, si entre dues vocals hi ha una diferència de to significativa, es poden donar dues situacions: o bé hi ha un descens de la primera a la segona (indicat convencionalment amb els símbols H+L), o bé hi ha un ascens (L+H). Si no hi ha cap diferència significativa entre elles, es tractarà o d'un to alt (H) o d'un to baix (L), segons el rang en què es realitza.

El segon postulat del model AM, que el distingeix d'altres models, és que la síl·laba tònica serveix de *punt d'ancoratge* dels fenòmens tonals (Hualde, 2003: 159). Això significa que la successió de tons que formen un contorn es recolzen en una síl·laba tònica. Tot i això, a les llengües en què la melodia tonal no té valor lèxic, no hi ha relació fixa entre la síl·laba tònica i un tipus concret de contorn tonal: a una síl·laba tònica poden trobar-se associats diferents contorns tonals, segons el tipus d'enunciat, la modalitat, la posició a l'interior de la frase, la rellevància de la paraula a l'interior de l'enunciat i, òbviament, la llengua.

En el moment en què es vol investigar com es relaciona el to amb l'accent tònic, a aquesta cadena de tons s'afegeix la indicació de la tonicitat de la vocal (marcada convencionalment amb un asterisc). D'aquesta manera les situacions possibles són les que s'indiquen a la taula 2.

Estructura tonal	Accent tònic en la vocal	
	1a	2a
H+L	H*+L	H+L*
L+H	L*+H	L+H*
L	L*	
H	H*	

Taula 2. Estructures tonals i accent tònic.

La primera d'aquestes estructures és H*+L i es pot descriure com a una *baixada a partir de la tònica*. La segona, H+L*, es pot parafrasejar com a una *baixada amb un pic a la pretònica*. La tercera, L*+H, pot ser enunciativa com a una *pujada amb un pic després de la tònica*. La quarta, L+H*, correspon a una *pujada amb un pic a*

la tònica. Finalment, L* i H* corresponen amb vocals realitzades amb un to baix o alt, respectivament, sense diferències significatives a les vocals adjacents.

El tercer concepte rellevant del mateix model és que existeixen dos nivells d'estructures entonatives: el superficial i el profund (Ladd, 1996). Les estructures profundes o subjacents, que són les representades a la taula 1, constitueixen models abstractes que troben les seves manifestacions empíriques als actes de parla en forma d'estructures superficials. El nombre d'estructures superficials és major que el nombre d'estructures profundes, ja que algunes de les superficials són només variants d'una mateixa estructura subjacent (Ladd, 1996).

És important remarcar que considerem que les estructures tritonals⁴ es consideren al·lòtons d'estructures subjacents bitonals (Martínez Celadrán i Fernández Planas, 2003; Fernández Planas i Martínez Celadrán, 2003). Com que mesurem l'F0 de ternes de vocals (la tònica, la pretònica i la posttònica de cada accent lèxic de les frases), és força probable que es trobin estructures tritonals. Les estructures tritonals possibles són dotze⁵ (vegeu la taula 3), només sis de les quals són pertinents. Les altres sis no ho són ja que es consideren només les ternes vocàliques l'element tònic de les quals és el central. La taula 3 recull les dotze estructures tritonals possibles. D'acord amb la proposta d'etiquetatge de Martínez Celadrán i Fernández Planas (Martínez Celadrán i Fernández Planas, 2003; Fernández Planas i Martínez Celadrán, 2003), els parèntesis indiquen quin dels dos intervals tonals és *menys* significatiu en semitons. Per tant, tot i que ambdues anotacions (L+H*)+L i L+(H*+L) indiquin un moviment tonal circumflex ascendent-descendent amb el pic a la vocal tònica, difereixen entre elles per la menor prominència acústica i perceptiva del moviment ascendent en el primer cas i del moviment descendent en el segon (vegeu també la taula 4 per a una exemplificació gràfica i numèrica).





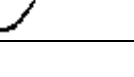
Per tant, a un conjunt de dades empíriques, podrem trobar dotze estructures superficials, que poden ser relacionades amb les quatre subjacents de la manera indicada a la taula 3, que representa l'actualització de la proposta original de Fernández Planas i Martínez Celadrán (2003: 173-174).

⁴ Les estructures tritonals s'obtenen, evidentment, quan es mesura la freqüència de tres vocals consecutives en lloc de dues i totes dues relacions sobrepassen el llindar establert.




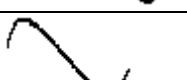
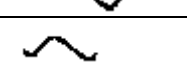
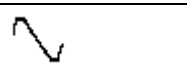

⁵ S'obtenen eixamplant una estructura profunda bitonal a l'esquerra o a la dreta.

Estructura profunda	Eixample a l'esquerra		Eixample a la dreta	
	H	L	H	L
(L+H*)	H+(L+H*)	il·lògica	(L+H*)+H	(L+H*)+L
(L*+H)	H+(L*+H)	il·lògica	(L*+H)+H	(L*+H)+L
(H+L*)	il·lògica	L+(H+L*)	(H+L*)+H	(H+L*)+L
(H*+L)	il·lògica	L+(H*+L)	(H*+L)+H	(H*+L)+L

Taula 3. Estructures tritonals (en negreta les pertinents en aquesta recerca).

Estructura superficial				Estructura profunda
Cadena	Representació esquemàtica	Exemple (valors en Hz i diferència en st)		
L*+H		117-120-140	<1,5 >1,5	L*+H
(L+H*)+H		100-110-120	>1,5 >1,5	
(H+L*)+H		120-110-130	>1,5 >1,5	
H+(L*+H) ⁶		130-110-120	>1,5 >1,5	
L+H*		105-120-117	>1,5 <1,5	L+H*

⁶ D'aquesta estructura també s'ha considerat la variant H+(L*+!H), que es caracteritza per una pujada d'F0 d'una amplada entre 1 i 1,5 st. Considerem que aquesta pujada, tot i no passar el llindar psicoacústic d'1,5 st, podria tenir una importància lingüística ja que és el resultat d'un moviment articulari voluntari contrari a la declinació que es donaria naturalment.

L+(H*+L)		100-120-100	>1,5 >1,5	
H*+L		117-120-105	<1,5 >1,5	H*+L
(L+H*)+L		107-120-100	>1,5 >1,5	
H+L*		120-105-110	>1,5 <1,5	H+L*
H*		257-260-255	<1,5 <1,5	H*
L*		110-105-100	<1,5 <1,5	L*
(H+L*)+L		130-119-100	>1,5 >1,5	desaccentuat

Taula 4. *Relació entre estructures tonals superficials i profundes amb exemples estilitzats.*

A tall d'exemple considerem la figura 12 en la qual s'il·lustren diferents tipologies tonals d'estructures superficials i profundes, així com el mètode per determinar-les a partir del llindar psicoacústic d'1,5 semitons. L'exemple correspon a la frase *Il sarcandul al mangjava la verdura* 'El pidolaire menjava la verdura' pronunciada per una parlant de friülà septentrional.

La figura 13 presenta, de manera esquemàtica, el càlcul de les estructures entonatives superficials possibles a partir de les dades acústiques d'F0 a cada accent lèxic i la següent transformació en estructures entonatives profundes per part d'AMPEREti.

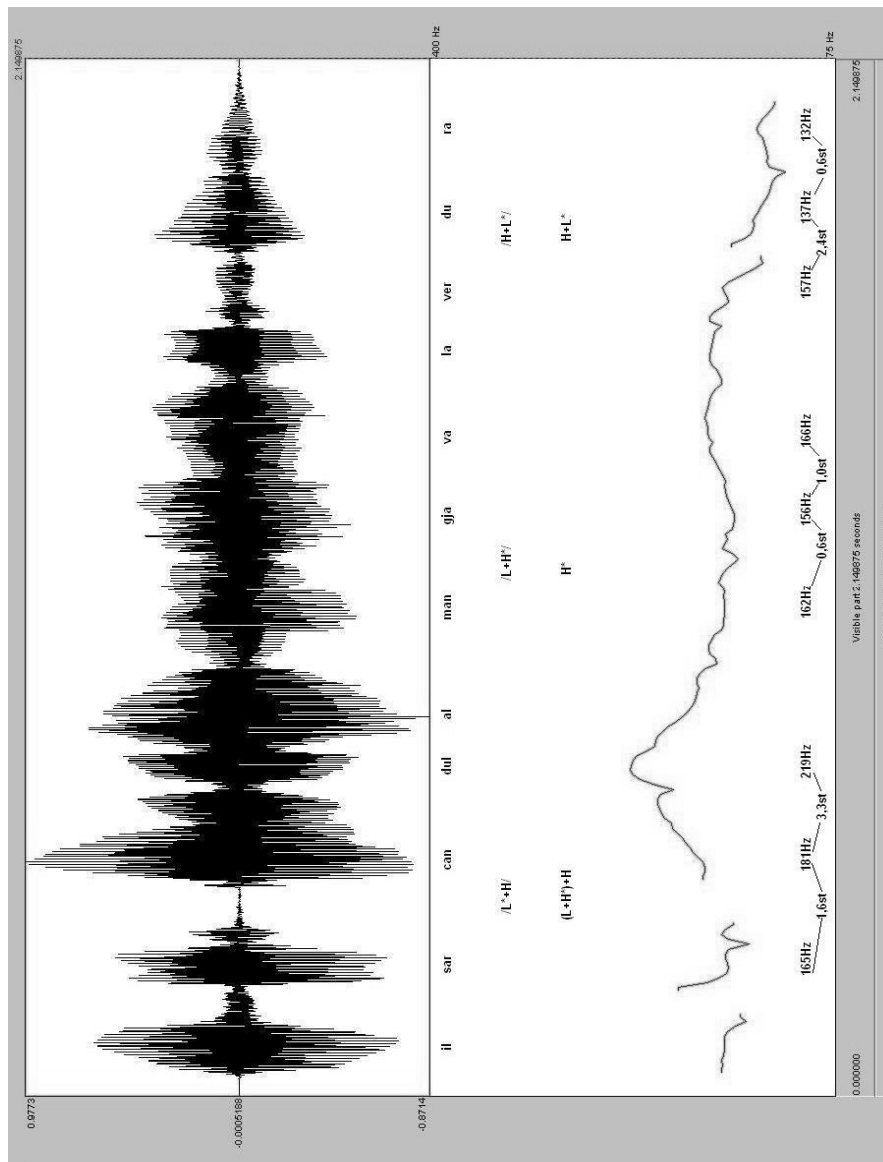


Figura 12. Exemple del mètode de determinació de les estructures tonals.

		$\Delta_1 = F0 \text{ pretònica} - F0 \text{ tònica}$					
		$ \Delta_1 < 1,5 \text{ st}$	$ \Delta_1 > 1,5 \text{ st} \wedge \Delta_1 < \Delta_2 $	$ \Delta_1 > 1,5 \text{ st} \wedge \Delta_1 > \Delta_2 $			
$\Delta_2 < 0$	$\Delta_1 < 0$						
	$\Delta_1 > 0$						
$\Delta_2 > 0$	$\Delta_1 < 0$						
$\Delta_2 > 0$	$\Delta_1 > 0$						
$ \Delta_2 < 1,5 \text{ st}$	$\Delta_2 < 0$						
	$\Delta_2 > 0$						
$ \Delta_2 > 1,5 \text{ st}$ \vee $ \Delta_2 < \Delta_1 $	$\Delta_2 < 0$						
	$\Delta_2 > 0$						
$ \Delta_2 > 1,5 \text{ st}$ \vee $ \Delta_2 > \Delta_1 $	$\Delta_2 < 0$						
	$\Delta_2 > 0$						

Figura 13. Representació esquemàtica del càlcul de les estructures entonatives superficials a partir de les dades acústiques d'F0 i transformació en estructures entonatives profundes per part d'AMPEREti.

El quart i últim aspecte rellevant del model AM és representat pels tons de frontera. Els tons no estan associats només amb síl·labes amb accent lèxic, sinó també amb el final de les frases. En aquest últim cas es parla de *tons de frontera* o de *tons de juntura*, que es tractaran amb més detall a continuació. Els tons de frontera final de frase en el model clàssic poden ser només dos: un to alt (indicat amb H%) i un to baix (indicat amb L%). Tot i això, aplicacions successives del model AM a llengües diferents de l'anglès va portar a ampliar l'inventari de tons de frontera, que pot incloure també un to mitjà (normalment etiquetat com a !H%) i tons de frontera complexos, és a dir formats per dos tons o fins i tot més de dos.

L'última idea fonamental del model AM, per tant, és que, a més d'haver-hi accents tonals, també hi ha tons de frontera. Mentre que un accent tonal és un to (o una seqüència de tons) associat amb una síl·laba accentuada, un to de frontera s'associa amb el límit de la frase prosòdica (Hualde, 2003: 157, 167). El model de Pierrehumbert (1980) proposa dos tipus de frases prosòdiques⁷: la frase intermèdia i la frase entonativa (que anomena *intermediate phrase* i *intonational phrase*, respectivament). La frase entonativa és la unitat tonal de nivell més elevat, que pot incloure al seu interior unes frases intermèdies. La frase entonativa es caracteritza per un to de frontera inicial i un to de frontera final, indicats amb %T i T%, respectivament. De la frase intermèdia se sol considerar només el to final, indicat amb T-.

Aquesta distinció entre accents tonals i tons de frontera es fonamenta en la idea que l'entonació té més d'una funció. De fet, l'F0 no realitza només prosòdicament els accents lèxics (en aquest cas es parla d'accents tonals), sinó que també organitza el discurs en unitats prosòdicament autònomes que ajuden l'oient a agrupar i segmentar adequadament la cadena sonora (Prieto, 1999). Per exemple, en friülà oriental la frase *La vecja lanza la menaza* és sintàcticament ambigua i pot tenir dues interpretacions. Si el subjecte és *la vecja* i el verb és *lanza*, l'oient entén que hi ha una dona vella que està amenaçant algú (figura 14). Per contra, si el subjecte és *la vecja lanza* i el verb és *menaza*, l'oient entén que una dona és amenaçada per una vella lança (figura 15). Aquesta diferència, que no es reflecteix a la llengua escrita, es realitza a l'acte de parla mitjançant l'entonació. A la primera frase, de fet, hi ha un to alt de frontera intermèdia (H-) entre els mots *vecja* i *lanza*, mentre que a la segona frase el mateix to H- es troba entre *lanza* i *la* (Hualde, 2003: 168-169; Prieto, 1999).

⁷ Convé assenyalar que la teoria de la fonologia prosòdica distingeix un conjunt més ampli de constituents prosòdics, que reben denominacions diferents segons l'autor (p. ex. Selkirk, 1986; Nespor i Vogel, 1986).

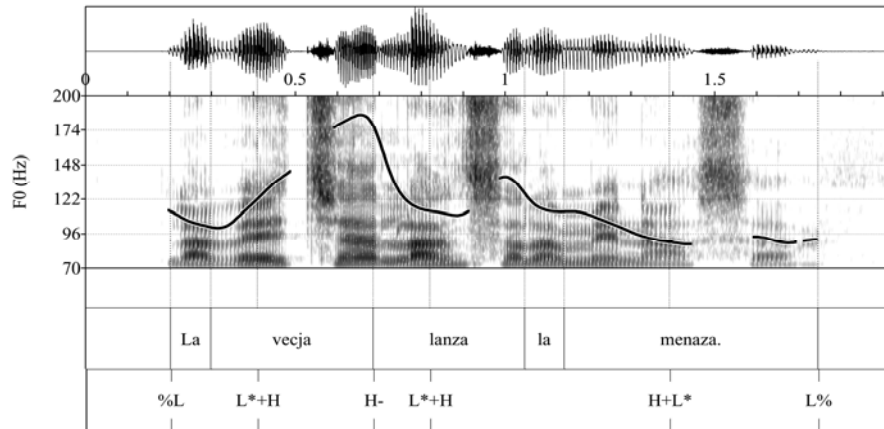


Figura 14. Oscil·lograma, espectrograma, corba d'F0 i estructura entonativa de la frase [La vecja]_{SUBJ} [lanza]_{VERB} [la menaza]_{OBJ} pronunciada per un parlant de friülà oriental.

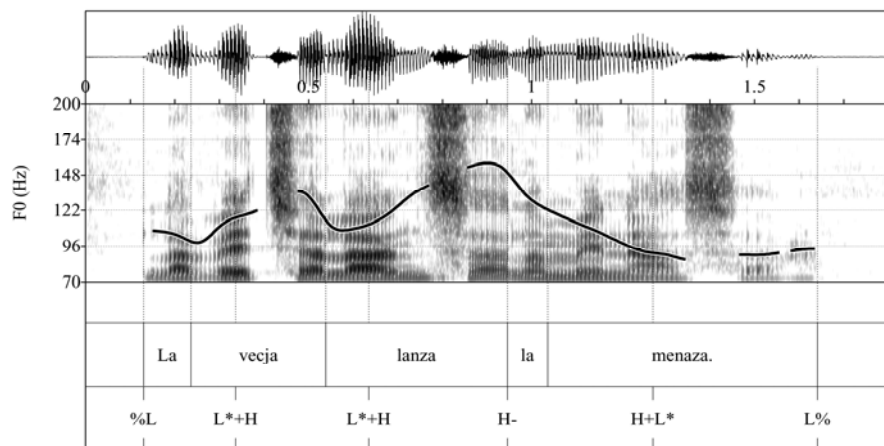


Figura 15. Oscil·lograma, espectrograma, corba d'F0 i estructura entonativa de la frase [La vecja lanza]_{SUBJ} [la]_{OBJ} [menaza]_{VERB} pronunciada per un parlant de friülà oriental.

A més, a la paraula friülana *elefant*, la síl·laba *-fant* és tònica i es percep com a més prominent que les altres del mateix mot, de la mateixa manera que la síl·laba *-can* de la paraula *african* és tònica i més prominent. A una frase com *L'elefant african!* (figura 16), però, la síl·laba tònica de l'última paraula es percep com a més prominent que la de la primera. Per això es diu que la segona paraula porta accent nuclear (Hualde, 2003: 157). És evident, per tant, que l'F0 pot estar relacionada, entre altres factors⁸, amb l'estructura sintàctica de la frase.

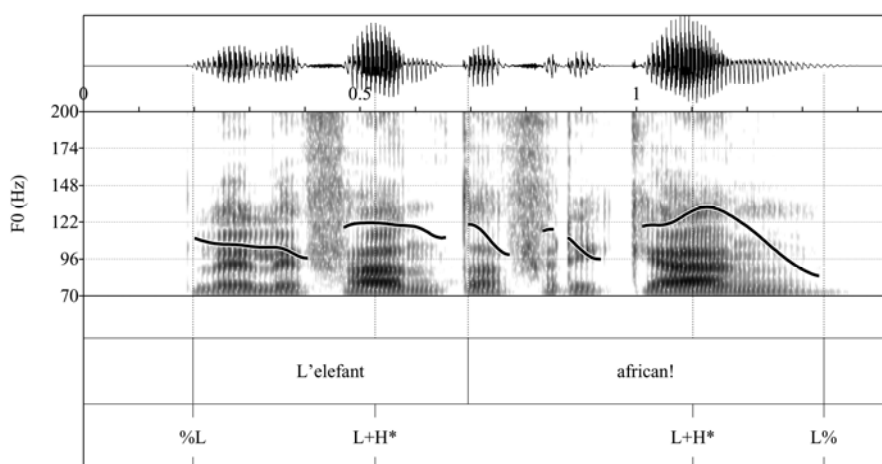


Figura 16. *Oscil·lograma, espectrograma, corba d'F0 i estructura entonativa de la frase exclamativa L'elefant african! pronunciada per un parlant de friülà septentrional.*

Entre els tons de frontera de frase entonativa, els que solen ser importants són els finals, mentre que els inicials juguen un paper molt més limitat (Hualde, 2003: 169; Prieto *et al.*, 2009: 299). Per aquest motiu, ens concentrarem sobretot en els tons

⁸ Tal com remarca Prieto (1999), l'agrupació prosòdica dels enunciat depèn de diferents factors que són: a) la velocitat d'elocució; b) la sintaxi de l'enunciat; c) la longitud dels constituents (com més llargs són els constituents, més probable és que estiguin separats per una frontera prosòdica); d) l'estructura informativa de l'enunciat (les fronteres prosòdiques se solen utilitzar per marcar la separació entre informació coneguda i informació nova); e) la configuració accentual de l'enunciat (xocs accentuals etc.).

finals de frontera, i postularem que només n'hi ha de dos nivells: alt (H%) i baix (L%). Per a les frases intermèdies considerarem només el to final, al qual atribuirem els mateixos dos nivells, indicats en aquest cas amb H- i L-, respectivament.

En analitzar les dades, considerem els valors d'F0 pròxims a les fronteres intermèdies i finals. En concret, agafem una finestra de dues vocals: la vocal immediatament següent a la frontera i la vocal immediatament anterior a ella. Per exemple, si tenim una frase com a *La fantate si nete la barete*, primer de tot establím on es col·loquen les fronteres finals de mot prosòdic, que a la figura 17 indiquem amb dues barres verticals.

La fantate || si nete || la barete. ||

Figura 17. *Fronteres intermèdies i frontera final.*

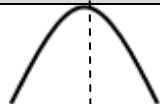
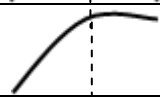
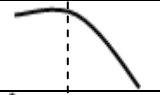
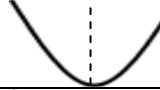
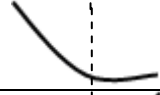
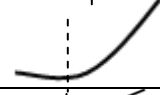

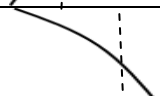
En un segon pas, agafem els valors d'F0 pròxims a cada frontera. Per exemple, si considerem el primer sintagma de la mateixa frase que s'ha fet servir a la figura anterior (és a dir la paraula esdrúixola *La fantate*), extraíem el valor inicial i final d'F0 de les vocals que s'han emfasitzat i subratllat a la figura 18.

La fantate_u || si_i nete || la barete. ||

Figura 18. *Terna de vocals per a cada frontera de frase intermèdia.*

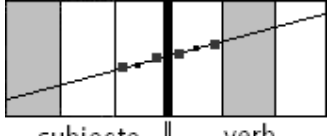
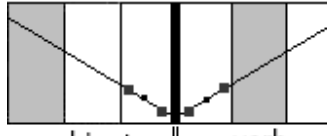
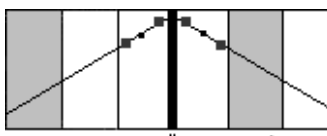
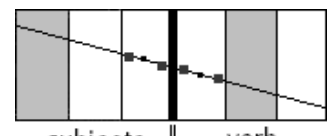
Aquests quatre valors són suficients per indicar, de manera esquemàtica, el comportament d'F0 en proximitat a les fronteres. Considerarem que hi ha una frontera entonativa intermèdia baixa (L-) si en correspondència amb la frontera final d'un sintagma es pot identificar el punt d'inflexió d'un moviment tonal còncau, és a dir si l'F0 baixa (o es manté baixa) a l'última síl·laba del subjecte i puja a la següent. Al contrari, considerem que al final del sintagma hi ha una frontera intermèdia alta (H-) si en correspondència amb la frontera final del sintagma mateix es pot identificar el vèrtex d'una inflexió tonal còncaua, és a dir si l'F0 puja (o es manté alta) a l'última síl·laba del subjecte i baixa a la següent. En tots els altres casos, és a dir quan no és possible identificar un punt d'inflexió tonal a la frontera entre els sintagmes, considerem que no hi ha cap esdeveniment tonal i, per tant, que la frontera sintagmàtica no està marcada entonativament.

Les situacions que es poden verificar en relació amb les fronteres de frase intermèdia són les que s'il·lustren a la taula 5. A la primera columna de la taula apareix un exemple de corba, amb una línia discontinua vertical que indica la posició de la frontera de mot. A la segona columna consten els quatre valors d'F0 que es consideren (és a dir, els valors inicial i final de les dues vocals adjacents a la frontera sintagmàtica). Finalment, a la columna dreta, apareix el tipus de to de frontera que correspon a cada tipus de corba.

Representació esquemàtica	Exemple (valors en Hz)	To de frontera
	100-120-120-100	H-
	105-120-119-117	
	117-120-110-90	
	120-110-110-130	L-
	120-105-108-110	
	123-120-130-140	
	100-110-120-130	Cap frontera
	130-119-100-90	

Taula 5. Tons de frontera intermèdia.

AMPEREti també proposa un etiquetatge de la frontera interna que apareix, a diferents llengües romàniques, entre el subjecte i el verb, tant als enunciats declaratius com als interrogatius. El programa considera que hi ha una frontera intermèdia baixa (L-) si, en correspondència amb la frontera final del sintagma que fa de subjecte, es pot identificar el vèrtex d'una inflexió tonal còncava; és a dir, si F0 baixa (o es manté baix) a l'última síl·laba del subjecte i puja a la següent. Al contrari, AMPEREti considera que al final del subjecte hi ha una frontera intermèdia alta (H-) si, en correspondència amb la frontera final del sintagma que fa de subjecte, es pot identificar el vèrtex d'una inflexió tonal còncava, és a dir si F0 puja (o es manté alt) a l'última síl·laba del subjecte i baixa a la següent. En tots els altres casos, és a dir quan no és possible d'identificar un vèrtex tonal a la frontera entre subjecte i verb, el programa considera que no hi ha cap esdeveniment tonal i, per tant, no proposa cap etiqueta. La taula 6 presenta, de manera esquemàtica, el càlcul dels tons de frontera interna final del subjecte a partir dels valors inicial i final d'F0 de les vocals adjacents a la frontera entre subjecte i verb.

		$\Delta_1 = F0 \text{ final última síl·laba del subjecte} - F0 \text{ final última síl·laba del subjecte}$	
		$\Delta_1 > 0$	$\Delta_1 < 0$
$\Delta_2 = F0 \text{ final primera síl·laba del verb} - F0 \text{ final primera síl·laba del verb}$	$\Delta_2 > 0$	 <p>subjecte verb Cap frontera</p>	 <p>subjecte verb L-</p>
	$\Delta_2 < 0$	 <p>subjecte verb H-</p>	 <p>subjecte verb Cap frontera</p>

Taula 6. Representació esquemàtica del càlcul del to de frontera intermèdia per part d'AMPEREti.

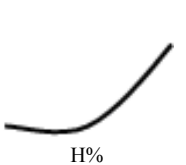

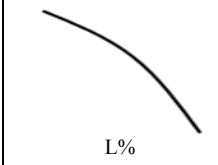
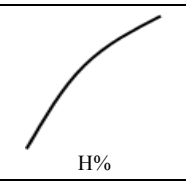
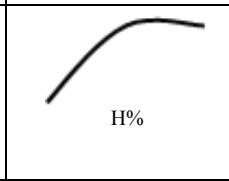
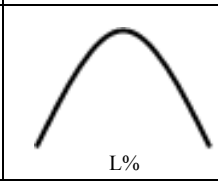
Pel que fa a les fronteres finals, disposem de dos valors d'F0: el valor central de l'última vocal tònica i el valor final de l'última vocal de l'enunciat (veg. figura 19).

La zovine || si nete || la barete. ||

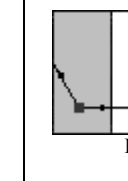
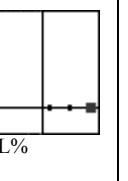
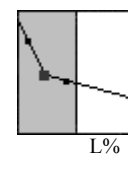
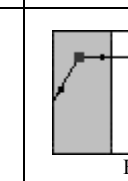
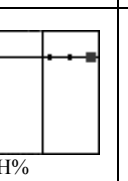
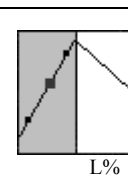
Figura 19. *Vocals analitzades per al càlcul de la frontera final de frase entonativa.*

Aquests dos valors són suficients per indicar, de manera esquemàtica, el comportament d'F0 en proximitat a la frontera final. Les situacions que es poden verificar en relació amb la frontera final de frase entonativa són les que s'il·lustren a la taula 7. En relació amb aquesta taula, cal destacar sobretot que, si no hi ha cap diferència psicoacústicament significativa entre els dos valors d'F0 que es consideren, es postula que el to de frontera és del mateix nivell del to final de l'últim accent tonal superficial de la frase. Per exemple, si l'últim accent tonal de la frase és H+L*, a falta de diferències psicoacústiques significatives, s'ha d'inferir que el to final de frontera de frase entonativa ha de ser L%.

AMPEREti analitza també els esdeveniments tonals que tenen lloc en correspondència amb les fronteres de l'enunciat. Pel que fa a la frontera final, AMPEREti calcula en primer lloc la diferència en semitons entre el valor central d'F0 de l'última síl·laba tònica de la frase i el valor final d'F0 de l'enunciat (en el cas de mots aguts són dos valors a la mateixa síl·laba). Si aquest valor és superior a 1,5 semitons (és a dir si hi ha un ascens final psicoacústicament significatiu), el programa considera que el to final és alt (H%). Si el valor en qüestió és superior a -1,5 st (és a dir si hi ha un descens psicoacústicament significatiu), AMPEREti considera que el to final és baix (L%). En els casos en què no hi ha cap diferència perceptible (és a dir que el valor està comprès entre -1,5 i 1,5 st), l'aplicació pren en consideració l'accent tonal anterior: si aquest és alt o ascendent, considera que el to final és alt, també. Si és baix o descendent, considera que el to final és igualment baix. La taula 8 presenta, de manera esquemàtica, el càlcul dels tons de frontera final a partir de les dades acústiques d'F0 i dels tipus d'accents tonals de la síl·laba nuclear. A la figura en qüestió el rectangle gris representa la síl·laba tònica nuclear i els rectangles blancs les síl·labes posttòniques. Cal destacar que, encara que a l'esquematització que es proposa s'hagin representat dues síl·labes posttòniques, el mateix càlcul es pot aplicar també si n'hi ha només una (és a dir, si el mot és pla) o cap (és a dir, si la paraula és aguda).



		Diferència d'F0 entre l'última vocal tònica i el final de l'última vocal de l'enunciat		
		Pujada psicoacústicament significativa	Cap diferència psicoacústicament significativa	Baixada psicoacústicament significativa
Accent nuclear	Baix o descendent			
	Alt o ascendent			

Taula 7. Tons de frontera final de frase entonativa.

		$\Delta = F0 \text{ última tònica} - F0 \text{ final última vocal}$		
		$\Delta > 1,5 \text{ st}$	$ \Delta < 1,5 \text{ st}$	$\Delta < -1,5 \text{ st}$
Accent nuclear	Baix o descendent			
	Alt o ascendent			

Taula 8. Representació esquemàtica del càlcul del to de frontera final per part d'AMPEREti.

A més a més, AMPERETi també proposa un etiquetatge del to inicial de l'enunciat, a partir de la comparació entre el valor inicial d'F0 de la primera vocal de la frase i el valor mitjà d'F0 de l'enunciat. Si el valor d'F0 inicial és més alt que el valor mitjà, el programa considera que el to és alt (%H) i si és inferior que és baix (%L). La taula 9 presenta, de manera esquemàtica, el càlcul dels tons de frontera inicial a partir de les dades acústiques d'F0.

$\Delta = F0 \text{ inicial primer vocal} - F0 \text{ mitjà}$	
$\Delta < 0$	$\Delta > 0$
 <p style="text-align: center;">%L</p>	 <p style="text-align: center;">%H</p>

Taula 9. Representació esquemàtica del càlcul del to de frontera inicial per part d'AMPERETi.

Com a exemple considerem la figura 20 en la qual s'il·lustren diferents tipologies de tons de frontera, així com el mètode per determinar-les a partir del llindar psicoacústic d'1,5 semitons. L'exemple correspon a la frase *Il sarcandul al mangjava la verdura* 'El pidolaire menjava la verdura' pronunciada per una parlant de friülà septentrional.

L'etiquetatge *superficial* és successivament transformat en una proposta d'etiquetatge *profund*, que representa una aproximació a l'etiquetatge fonològic i constitueix la base per a l'última anàlisi duta a terme pels investigadors. La transformació de les etiquetes *superficials* en etiquetes *profundes* es fonamenta en les regles d'implementació fonètica dels tons que són pròpies de cada llengua. De fet, la transformació del fonètic en fonològic, tant en l'àmbit segmental com en l'àmbit suprasegmental, ha de tenir en compte les especificitats de cada llengua. En l'àmbit segmental, per exemple, la realització fonètica [ˈwɪk] té representacions fonològiques diferents segons es tracti de la paraula polonesa <łyk> /'wɪk/ o de la paraula <vuic> /'wɪk/ del friülà centreoriental, ja que en la primera llengua [ɪ] és la realització del fonema /i/ mentre que en la segona és la realització al·lofònica del fonema /i/ en posició tònica (figura 21).

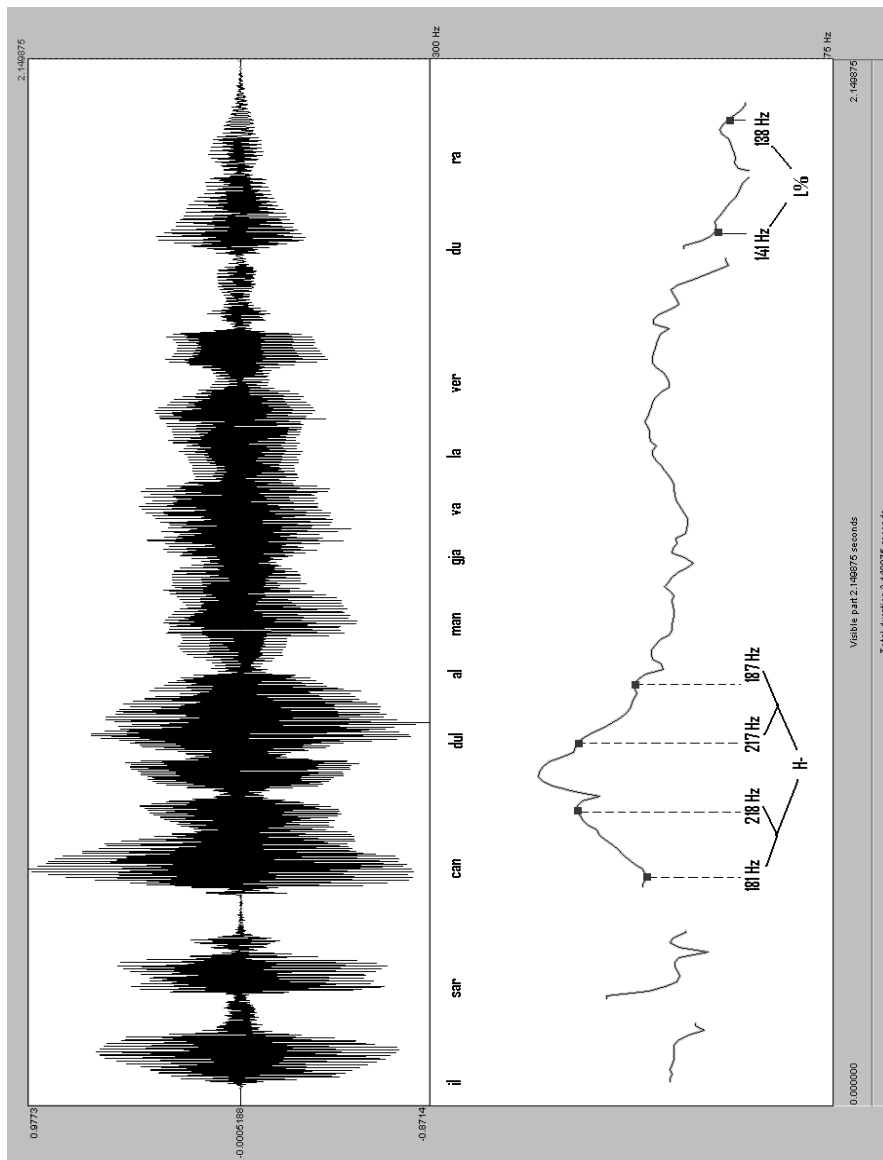


Figura 20. Exemple del mètode de determinació dels tons de frontera intermèdia i final.

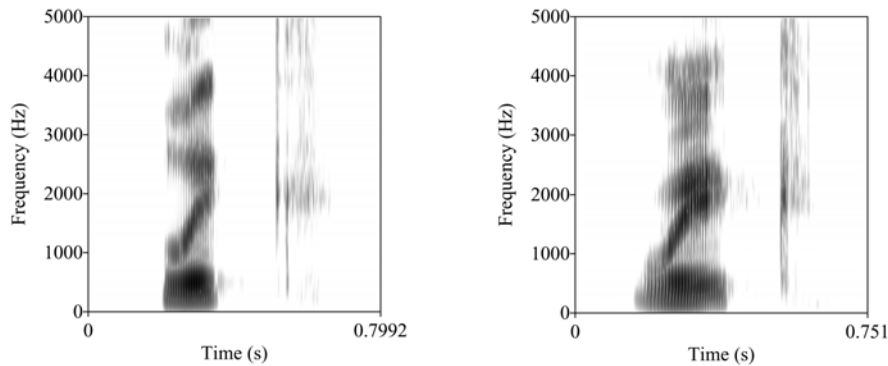


Figura 21. Espectrograma de la paraula ['wik] pronunciada per un parlant de polonès (panell de l'esquerra) i per un parlant de friülà central (panell de la dreta).

Com en l'àmbit segmental, els mateixos fets acústics poden tenir representacions fonològiques diferents segons la llengua també en l'àmbit suprasegmental, ja que la mateixa corba entonativa pot correspondre a representacions fonològiques diferents en funció de la llengua. Un exemple d'aquest tipus es presenta a les figures 10 i 11 que hem vist anteriorment. La versió actual de la part d'AMPEREti dedicada a l'etiquetatge entonatiu ha estat concebuda per al friülà, però s'està ampliant de manera que pugui ésser aplicada amb èxit a d'altres llengües romàniques on la fonologia entonativa és, lògicament, diferent.

4.3.2. Etiquetatge de la durada

La durada vocàlica ha estat objecte d'anàlisi des dels començaments de la fonètica experimental. Els autors que s'han ocupat d'aquest tema s'han enfrontat a dues preguntes:

1. La durada vocàlica és fonològica?
2. La durada vocàlica està relacionada amb l'accent?

Les respostes que han donat són diferents d'una llengua a l'altra. De fet, és sabut que hi ha llengües que tenen una diferència fonològica entre vocals llargues i breus, com ara l'àrab estàndard, el txec (Pamies, 1993: 35-37) i el friülà. Hi ha altres llengües, per contra, en les quals aquesta diferència –quan existeix– només és fonètica i depèn de factors microprosòdics, com ara el grau d'obertura de les vocals, l'estructura sil·làbica i la llargada de la paraula. De fet sembla que hi ha regles universals que determinen microvariacions de la durada (Pamies, 1993: 39; Carton, 1974: 104-108; Malmberg, 1966: 85-86; Recasens, 1991: 52). En general, es veu que:

1. Com més oberta és una vocal, més llarga resulta.
2. Com més llarga és una paraula, més breus són les seves vocals.
3. La mateixa vocal és més llarga si va seguida per una consonant sonora.

Pel que fa a la relació entre tonicitat i durada vocàlica, hi ha diferents opinions, tant entre autors que s'han ocupat de llengües diferents, com entre autors que han estudiat la mateixa llengua (Pamies, 1993: 35-50). D'una banda, de fet, hi ha autors que argumenten que les vocals tòniques són més llargues de les àtones mentre que, de l'altra, n'hi ha que intenten demostrar que no hi ha una relació clara entre tonicitat i durada vocàlica. Si agafem l'italià veiem que es tracta d'un cas força afortunat, ja que hi ha un cert acord entre els autors sobre la idea que la durada vocàlica està relacionada amb l'accent tònic. En particular, sembla que les vocals tòniques en sil·laba oberta són més llargues que les àtones en la mateixa posició (Canepari, 1979: 75). Altres autors (Bertinetto, 1981) arriben a dir que, en italià, la durada vocàlica és sempre el paràmetre que indica la tonicitat.

Encara que molts autors hagin destacat el paper prosòdic de la durada, cap dels models que s'han il·lustrat als paràgrafs anteriors li dedica una atenció comparable a la que ha rebut l'F0. També l'aplicació clàssica del model AM a la prosòdia es limita només a un dels seus paràmetres responsables: la modulació de la freqüència fonamental. Uns estudis més recents de Fernández Planas i Martínez Celdrán (2003) proposen un sistema anàleg al de l'anàlisi de l'F0 per a l'estudi de la durada. Emulant els postulats del model AM, els autors en qüestió (2003: 168) assumeixen que en les frases, les sil·labes, especialment les vocals, se succeeixen en el temps i la seva anàlisi s'ha de dur a terme en funció de la successió de dues dimensions duratives: G (gran) i P (petita). Per tal d'etiquetar una vocal amb G o P, després de mesurar el valor de la seva durada (en ms), es calcula la diferència entre

les vocals i es considera significativa cada diferència superior al llindar d' $1/3$ de la durada en més o en menys de la primera vocal⁹.

En el moment en què es vol investigar com la durada es relaciona amb la tonicitat lèxica, a aquesta cadena de durades s'afegeix la indicació convencional de la tonicitat de la vocal, mitjançant un asterisc (*). Com que al llarg d'aquest treball mesurarem la durada de ternes de vocals, al conjunt de dades que analitzarem podrem trobar onze estructures duratives superficials, que apareixen a la primera columna de la taula 10. Les estructures duratives superficials es poden transformar en estructures profundes que són bàsicament dues: la primera d'elles (G*) indica que la vocal tònica és significativament més llarga que les adjacents, mentre la segona (P*) indica que la vocal tònica és significativament més petita que les vocals adjacents. També es contempla la possibilitat que no hi hagi cap diferència significativa de durada a la terna de vocals considerades, o bé que les dues diferències significatives siguin en la mateixa direcció, cas en el qual es considera que no hi ha cap estructura durativa relacionada amb la tonicitat.

Pel que fa a la durada, AMPERETi compara, fonamentant-se en el llindar psicoacústic de perceptibilitat d' $1/3$ de la durada (Pamies *et al.*, 2002), els valors de la vocal pretònica, tònica i posttònica de cada accent lèxic i obté un etiquetatge coherent amb el marc teòric descrit anteriorment. La figura 22 presenta, de manera esquemàtica, el càlcul de les estructures de durada a partir de les dades acústiques per part d'AMPERETi.

Com a exemple considerem la figura 23 en la qual s'il·lustren diferents tipologies d'estructures duratives superficials i profundes, així com el mètode per determinar-les a partir del llindar psicoacústic d' $1/3$ de la durada. L'exemple correspon a la frase *Il sarcandul al mangjava la verdura* 'El pidolaire menjava la verdura' pronunciada per una parlant de friülà septentrional.

4.3.3. Etiquetatge de la intensitat

La intensitat és un paràmetre que va tenir molta fama durant el període en què no hi havia mitjans tècnics per poder-lo analitzar empíricament. De fet, els primers autors de l'escola britànica associaven l'accent lèxic amb la intensitat (García-

⁹ També l'ús d'aquest llindar normalitza les diferències interparlants i intraparlants. El llindar d' $1/3$ de la durada ha estat establert per al castellà (Pamies *et al.*, 2002) i es farà servir també per al friülà.

Lecumberri, 2003: 38). També els primers fonetistes espanyols pensaven que l'accent lèxic es realitzava mitjançant la intensitat (Navarro Tomás, 1944; Alarcos Llorach, 1950; Gili Gaya, 1950). Quan es va fer comú l'ús de l'oscil·loscopi, es va començar a veure que la intensitat no estava relacionada de manera tan clara amb la tonicitat, ja que la vocal tònica era la més intensa només en una minoria de casos a la majoria de les llengües estudiades (Pamies, 1993: 50-60), en parla continua, no en llistes de paraules aïllades. Tot i això, s'ha de destacar l'existència d'estudis molt recents que arriben a la conclusió que, fins i tot en una llengua com l'anglès, per a la qual normalment es dona per descomptat que l'accent lèxic es realitza com a accent melòdic, la preeminència d'una síl·laba es deu, acústicament, en primer lloc a la intensitat, en segon lloc a la durada i, només en tercer lloc, a la freqüència fonamental (Kochanski *et al.*, 2005; Mo, 2008). Altres autors, al contrari, subratllen el paper primari de la durada, al qual s'afegeix el de la intensitat, que té una importància més limitada (Turk *et al.*, 1996).

Estructura superficial				Estructura profunda
Cadena	Representació esquemàtica	Exemple (valors en ms i diferència)		
P+G*	___ ___	44-90-75	>1/3 <1/3	G*
G*+P	_____	79-93-50	<1/3 >1/3	
(P+G*)+P	___ _____	60-155-56	>1/3 >1/3	
P+(G*+P)	___ _____	20-76-27	>1/3 >1/3	
G+P*	_____	141-62-49	>1/3 <1/3	P*
P*+G	___ ___	53-54-77	<1/3 >1/3	
(G+P*)+G	_____	58-40-59	>1/3 >1/3	
G+(P*+G)	_____	100-50-95	>1/3 >1/3	
-	_____	65-79-65	<1/3 <1/3	cap
G+P*+P	_____	129-95-48	>1/3 >1/3	
P+G*+G	_____	42-79-179	>1/3 >1/3	

Taula 10. Estructures duratives superficials i profundes.

		$\Delta_1 = (Durada\ tònica - Durada\ pretònica) / Durada\ pretònica$					
		$ \Delta_1 < 1/3$	$\Delta_1 > 0$	$ \Delta_1 > 1/3 \wedge \Delta_1 < \Delta_2 $	$ \Delta_1 > 1/3 \wedge \Delta_1 > \Delta_2 $		
$\Delta_2 < 0$	$ \Delta_2 < 1/3$						
	$ \Delta_2 > 1/3$						
$\Delta_2 > 0$	$ \Delta_2 < 1/3$						
	$ \Delta_2 > 1/3$						
$\Delta_2 < 0$	$ \Delta_2 < 1/3$						
	$ \Delta_2 > 1/3$						
$\Delta_2 > 0$	$ \Delta_2 < 1/3$						
	$ \Delta_2 > 1/3$						

Figura 22. Representació esquemàtica del càlcul de les estructures superficials de durada a partir de les dades acústiques i transformació en estructures profundes per part d'AMPEREti.

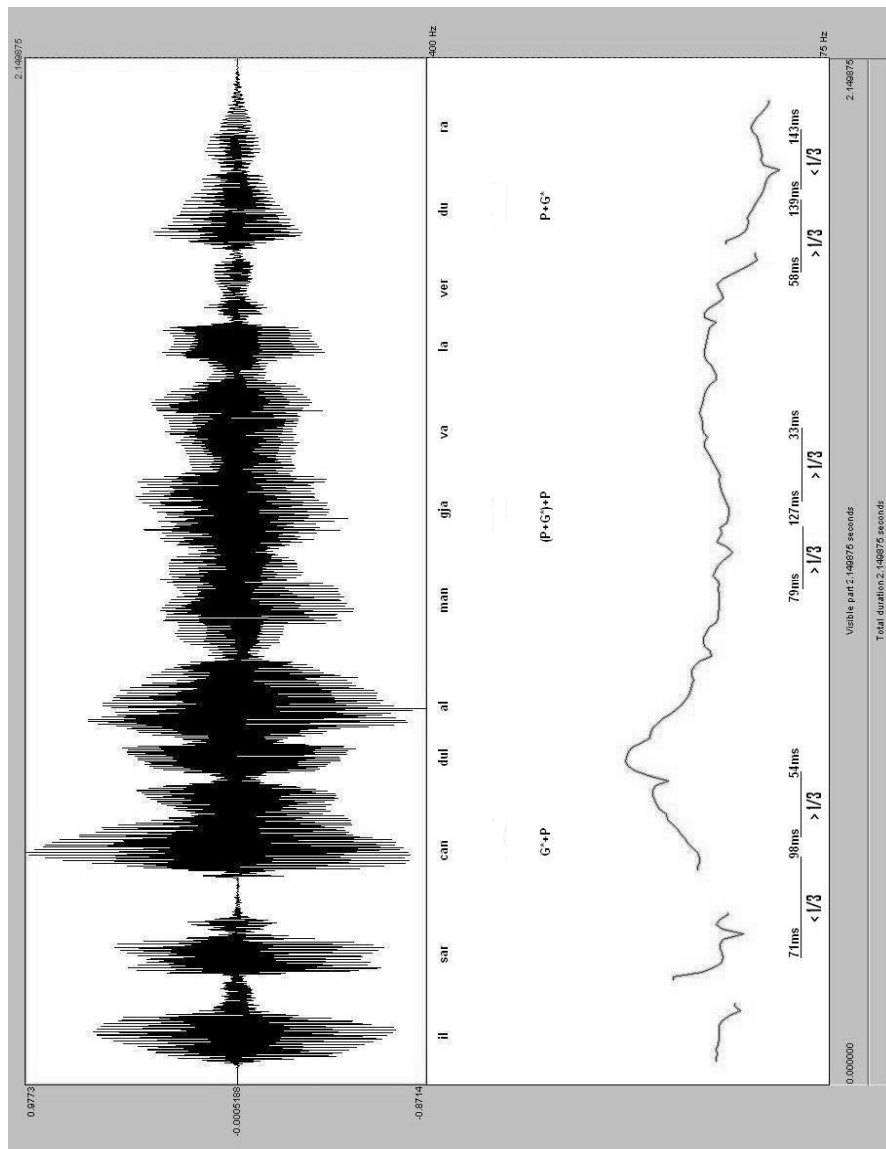


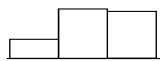
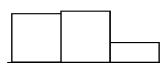
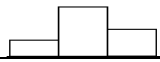
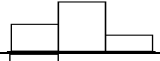
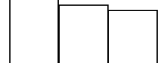



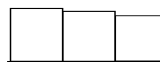

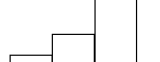
Figura 23. Exemple del mètode de determinació de les estructures duratives.

La dificultat tècnica de mesurar de manera precisa i unívoca la intensitat, juntament amb la impossibilitat de trobar una relació evident de la intensitat amb l'accent i el fet que l'anàlisi d'F0 ha capturat pràcticament tota l'atenció dels fonetistes i els fonòlegs al llarg dels últims anys, juntament amb la dificultat per establir l·lindars psicoacústics semblants als que s'ha proposat per a l'F0 i la durada, fan que no hi hagi a disposició cap model fonològic de tipus Mètric Autosegmental que intenti descriure i explicar el paper de la intensitat mitjançant estructures anàlogues a les estructures tonals i duratives que s'han il·lustrat als epígrafs anteriors.

Amb la intenció d'omplir aquest buit metodològic, proposem per a l'estudi de la intensitat un sistema anàleg al que s'ha descrit als paràgrafs anteriors per a l'F0 i la durada. Emulant els postulats del model AM, assumim que en les frases, les síl·labes, especialment les vocals, se succeeixen com una concatenació d'intensitats diferents i la seva anàlisi es pot dur a terme en funció de la successió de dos nivells: F (fort) i D (dèbil). Per tal d'etiquetar una vocal amb F o D, després de mesurar el valor de la seva intensitat (en dB), es calcula la diferència entre les vocals i es considera significativa cada diferència superior al l·lindar del 15% d'intensitat en més o en menys respecte de la primera vocal¹⁰. Com que al llarg d'aquest treball mesurarem la intensitat de ternes de vocals, al conjunt de dades que analitzarem podrem trobar onze estructures superficials, que apareixen a la primera columna de la taula 11. Les estructures superficials es poden transformar en estructures profundes, que són bàsicament dues: la primera d'elles (F*) indica que la vocal tònica és significativament més intensa que les adjacents, mentre la segona (D*) indica que la vocal tònica és significativament menys intensa que les vocals adjacents. També es contempla la possibilitat que no hi hagi cap diferència significativa d'intensitat a la terna de vocals considerades, o bé que les dues diferències significatives vagin en la mateixa direcció, situació en la qual es considera que no hi ha cap estructura d'intensitat relacionada amb la tonicitat.

Pel que fa a la intensitat, AMPERETi compara, fonamentant-se en el l·lindar psicoacústic de perceptibilitat del 15% de diferència d'intensitat (Riesz, 1928; Llisterri, 2012), els seus valors a la vocal pretònica, tònica i posttònica de cada accent lèxic i obté un etiquetatge coherent amb el marc teòric descrit anteriorment. La figura 24 presenta, de manera esquemàtica, el càlcul de les estructures d'intensitat a partir de les dades acústiques per part d'AMPERETi.

¹⁰ També l'ús d'aquest l·lindar normalitza les diferències interparlants i intraparlants. El l·lindar del 15% de la intensitat ha estat establert per Riesz (Riesz, 1928; Llisterri, 2012).

Estructura superficial				Estructura profunda
Cadena	Representació esquemàtica	Exemple (valors en dB i diferència)		
D+F*		40 - 70 - 68	>15% <15%	F*
F**+D		68 - 70 - 40	<15% >15%	
(D+F*)+D		20 - 70 - 34	>15% >15%	
D+(F**+D)		34 - 70 - 20	>15% >15%	
F+D*		120 - 70 - 68	>15% <15%	D*
D**+F		72 - 70 - 120	<15% >15%	
(F+D**)+F		120 - 70 - 100	>15% >15%	
F+(D**+F)		100 - 70 - 120	>15% >15%	
-		72 - 70 - 68	<15% <15%	cap
F+D**+D		130 - 70 - 35	>15% >15%	
D+F**+F		40 - 70 - 130	>15% >15%	

Taula 11. Estructures superficials i profundes d'intensitat.

		$\Delta_1 = (\text{Intensitat tònica} - \text{Intensitat pretònica}) / \text{Intensitat pretònica}$			
		$ \Delta_1 < 15/100 * \Delta_2 $ $\Delta_1 < 0$	$ \Delta_1 > 15/100 * \Delta_2 $ $\Delta_1 > 0$	$ \Delta_1 < 15/100 * \Delta_2 $ $\Delta_1 < 0$	$ \Delta_1 > 15/100 * \Delta_2 $ $\Delta_1 > 0$
$\Delta_2 = (\text{Intensitat posttònica} - \text{Intensitat tònica}) / \text{Intensitat tònica}$	$ \Delta_1 < 15/100 * \Delta_2 $				
	$ \Delta_1 > 15/100 * \Delta_2 $				
$ \Delta_2 < 0$	$ \Delta_1 < 15/100 * \Delta_2 $				
	$ \Delta_1 > 15/100 * \Delta_2 $				
$ \Delta_2 > 0$	$ \Delta_1 < 15/100 * \Delta_2 $				
	$ \Delta_1 > 15/100 * \Delta_2 $				

Figura 24. Representació esquemàtica del càlcul de les estructures superficials d'intensitat a partir de les dades acústiques i transformació en estructures profundes per part d'AMPEREti.

4.3.4. Prova de fiabilitat

S'ha testat la primera versió d'AMPEREti amb un corpus que conté 162 frases de tres tipus (declarativa neutra, interrogativa absoluta neutra i -només en català-interrogativa absoluta neutra introduïda per *que*), de quatre llengües (català, espanyol, friülà i italià) i deu punts d'enquesta (Barcelona, Girona, Lleida, Castelló de la Plana, Bullas, Arrecife, Siena, Agrons, Beivars i Gradisca). Totes les frases d'aquest corpus han estat etiquetades prèviament pels autors d'aquest treball fins arribar a un acord entre ells. Els resultats de l'etiquetatge coincideixen amb els que proposen els sistemes ToBI per a cada llengua a la literatura (entre d'altres: Dorta i Hernández, 2005; Martínez Celdrán, Fernández Planas i Roseano, 2008; Estebas i Prieto, 2009; Prieto, en premsa; Roseano, Vanrell i Prieto, en premsa).

Les 972 etiquetes *profundes* proposades per AMPEREti s'han comparat sistemàticament amb els etiquetatges *humans* per valorar el grau de fiabilitat de les fórmules a partir del nivell de concordances observades. Mitjançant l'Online Kappa Calculator¹¹, s'han calculat els índexs de fiabilitat. Per als tons de frontera, el percentatge de coincidència global entre l'etiquetatge *humà* i el que proporciona AMPEREti és de 0,96 (essent el màxim 1,00), mentre que el Free-marginal kappa és 0,92593. Per als accents tonals, el percentatge de coincidència global entre l'etiquetatge *humà* i el que proporciona AMPEREti és de 0,76, mentre el Free-marginal kappa és 0,72. Si recordem que el valor màxim de kappa és 1,00 i que valors de kappa superiors a 0,70 se solen considerar indicadors d'un bon nivell de concordança entre els etiquetadors, podem concloure que les anotacions entonatives proposades per AMPEREti tenen un bon grau de fiabilitat.

5. CONCLUSIONS

Les rutines que s'han presentat compleixen amb una sèrie d'objectius de tipus pràctic i metodològic. Des del punt de vista pràctic, constitueixen una ajuda eficaç per a l'anàlisi entonativa dels corpus fixos en el marc del projecte AMPER. Des del punt de vista metodològic analitzen les dades d'una manera *objectiva* (és a dir basada en la mesura dels fets acústics) i mitjançant una rutina informàtica, la qual cosa redueix el risc de subjectivitat, és a dir de diferències d'anàlisi entre investigadors. A més, creiem que l'aplicació de les mateixes fórmules a quatre llengües diferents pot ser un pas cap a l'elaboració d'un sistema interlingüístic d'anotació entonativa AM, un sistema que alhora sigui capaç de recollir els

¹¹ <http://justusrandolph.net/kappa/>

contrastos entonatius fonològics de cada llengua i de representar-los d'una manera que sigui transparent, és a dir fonamentada en la fonètica i semblant entre llengües.

La feina que queda per fer en aquest sentit és encara molta. En un futur pròxim, en primer lloc seria oportú donar flexibilitat a les aplicacions que s'han creat per etiquetar les frases del corpus fix d'AMPER, de manera que es puguin emprar per etiquetar enunciats amb un nombre de sintagmes i de síl·labes no predeterminat. Això permetria estendre'n l'ús a l'anàlisi de textos de parla espontània. En segon lloc, s'hauran de concebre fórmules addicionals que tinguin la capacitat de reconèixer i etiquetar correctament algunes estructures que són fonològicament distintives en d'altres varietats romàniques o en d'altres modalitats oracionals i que no estan recollides a l'inventari relativament senzill d'estructures profundes que hem emprat. Per exemple, caldrà tenir en compte que diverses llengües romàniques disposen d'un to de frontera final mitjà –que s'empra típicament al vocatiu, on sol anar precedit per un accent tonal ascendent– que la versió actual d'AMPEREti no diferenciaria. Més en general, les fórmules d'AMPEREti, ja que estan basades en la versió clàssica del model AM que preveu l'existència de només dos nivells tonals (alt o H i baix o L), no són aptes per detectar contrastos entre tres o quatre nivells tonals que es descriuen per a algunes llengües romàniques (p.e. Estebas i Prieto, 2009 per al castellà; Prieto *et al.*, 2009 per al català). Una altra millora futura del conjunt de fórmules d'etiquetatge haurà de consistir a incloure-hi algorismes que siguin capaços d'etiquetar els moviments tonals complexos que, en algunes llengües com el català i el castellà, s'associen a la frontera final d'oracions no neutres. La versió actual d'AMPEREti, de fet, només reconeix els tons de frontera final monotonals, que solen caracteritzar les modalitats neutres a les llengües romàniques.

En tercer i últim lloc cal recordar que l'extensió de l'ús de l'aplicatiu AMPEREti a altres llengües requereix unes reflexions sobre la relació entre etiquetatge fonètic i etiquetatge fonològic. El debat sobre aquest tema –delicat i crucial– neix de la necessitat de conciliar exigències diferents. D'una banda hi ha la necessitat de fonamentar fonèticament l'etiquetatge fonològic, cosa que demanaria que una mateixa corba entonativa s'etiquetés de la mateixa manera a totes les llengües. D'altra banda, hi ha l'exigència de motivar un etiquetatge fonològic a partir de la seva distintivitat i contrastivitat en un sistema lingüístic determinat, cosa que pot anar en contra d'un etiquetatge estrictament fonètic. Per exemple, s'ha demostrat (Roseano, 2012: 248-250) que en friülà la pujada tonal que es troba al final de les declaratives i que no arriba a un nivell tan alt com la de les interrogatives (i, per tant, correspon amb un to final mitjà i s'hauria d'etiquetar !H%) en realitat és una realització al·lotònica d'un to final baix, i per això fonològicament s'etiqueta com a

L%. En català, per contra, la mateixa pujada final fins a un nivell mitjà és fonològica ja que vehicula un sentit pragmàtic diferent (concretament un matís d'obvietat, veg. Prieto *et al.*, 2009). A causa de la seva contrastivitat, en el sistema Cat_ToBI aquesta pujada final rep l'etiquetatge fonològic !H%, que en marca la diferència en comparació amb els tons finals baix (L%) i alt (H%). Ens trobem, per tant, en la situació en què el mateix fenomen acústic, és a dir una lleugera pujada final d'F0, s'etiqueta fonològicament com a L% en friülà i com a !H% en català. És evident que això genera dues discrepàncies: la primera és entre llengües, en el sentit que el mateix contorn s'etiqueta de manera diferent a llengües de la mateixa família (tal com subratllen també Prieto i Hualde, 2012), i la segona és entre etiquetatge fonètic i fonològic, en el sentit que en friülà un to final fonèticament mitjà (!H%) es considera fonològicament baix (L%). Creiem que al llarg dels propers anys les dificultats generades per aquests dos tipus de discrepàncies –que segons la nostra manera de veure constitueixen un dels aspectes més problemàtics però més interessants dels etiquetatges entonatius– requeriran una reflexió aprofundida per part de la comunitat científica, a la qual esperem que aquest article pugui contribuir.

6. REFERÈNCIES BIBLIOGRÀFIQUES

- ALARCOS LLORACH, E. (1950): *Fonología española*, Madrid, Gredos.
- BECKMAN, M. E.; J. HIRSCHBERG i S. SHATTUCK-HUFNAGEL (2005). «The original ToBI system and the evolution of the ToBI framework», dins S. A. Jun (ed.): *Prosodic Typology. The Phonology of Intonation and Phrasing*, Oxford, Oxford University Press, pp. 9-54.
- BERTINETTO, P. M. (1981): *Strutture prosodiche dell'italiano*, Florència, Accademia della Crusca.
- CANEPARI, L. (1979): *Introduzione alla fonetica*, Torí, Einaudi.
- CARRERA SABATÉ, J. i A. M. FERNÁNDEZ PLANAS (2005):
- CARTON, F. (1974): *Introduction à la phonétique du français*, París, Bordas.
- CONTINI, M. (1992): «Vers une géoprosodie romane», dins G. Aurrekoetxea i X. Bidegain (ed.): *Actas del Nazioarteko Dialektologia Biltzarra Agiriak*, Bilbao, Publicaciones de la Real Academia de la Lengua Vasca, pp. 83-109.

-
- CUCCHIARINI, C. i H. STRIK (2003): «Automatic phonetic transcription: An overview», dins M. J. Solé, D. Recasens i J. Romero (eds): *Proceedings of International Congress of Phonetic Sciences*, Barcelona, pp. 347-350.
- DORTA, J. i B. HERNÁNDEZ (2005): «Acento y entonación: interrogativas vs declarativas SVO sin expansión en Canarias», *Revista Internacional de Lingüística Iberoamericana (RILI)*, 3, 2(6), pp. 85-108.
- ESTEBAS VILAPLANA, E. i P. PRIETO (2009): «La notación prosódica en español. Una revisión del Sp_ToBI», *Estudios de Fonética Experimental*, XVIII, pp. 263-283.
- FERNÁNDEZ PLANAS, A. M.; E. MARTÍNEZ CELDRÁN; V. SALCIOLI; G. TOLEDO i J. CASTELLVÍ (2002): «Taxonomía autosegmental en la entonación del español peninsular», dins J. Díaz García (ed): *Actas del II Congreso de Fonética Experimental*, Sevilla, Universidad de Sevilla, pp. 180-186.
- FERNÁNDEZ PLANAS A. M. i E. MARTÍNEZ CELDRÁN (2003): «El tono fundamental y la duración: Dos aspectos de la taxonomía prosódica en dos modalidades de habla (enunciativa e interrogativa) del español», *Estudios de Fonética Experimental*, 12, pp. 165-200.
- FERRARI DISNER S. (1983). *Vowel quality: the relation between universal and language specific factors*, UCLA Working Papers in Phonetics, 58.
- GABRIEL, C.; I. FELDHAUSEN; A. PEŠKOVÁ, L. COLANTONI; S.A. LEE; V. ARANA I L. LABASTÍA (2010): «Argentinian Spanish Intonation», dins P. Prieto i P. Roseano (eds.): *Transcription of Intonation of the Spanish Language*, Munic, Lincom Europa, pp. 285-317.
- GARCÍA-LECUMBERRI, M. L. (2003): «Análisis por configuraciones: la escuela britànica», dins P. Prieto (ed.): *Teorías de la entonación*, Barcelona, Ariel, pp. 35-62.
- GARRIDO ALMIÑANA, J. M. (2003): «La escuela holandesa: el modelo IPO», dins P. Prieto (ed.): *Teorías de la entonación*, Barcelona, Ariel, pp. 97-122.
- GILI GAYA, S. (1950): *Elementos de fonética general*, Madrid, Gredos.
- HUALDE, J.I. (2003): «El modelo métrico y autosegmental», dins P. Prieto (ed.): *Teorías de la entonación*, Barcelona, Ariel, pp. 155-184.
-

-
- INTERNATIONAL PHONETIC ASSOCIATION (1999): *Handbook of the International Phonetic Association*, Cambridge, Cambridge University Press.
- KEMP, J. A. (1988): «Universal phonetic alphabets: an historical perspective», dins C. Bamroongraks *et al.* (ed.): *The International Symposium on Language and Linguistics*, Bangkok, Thailandia, pp. 1-6.
- KOCHANSKI G.; E. GRABE; J. COLEMAN i B. ROSNER (2005): «Loudness predicts prominence: fundamental frequency lends little», *Journal of the Acoustical Society of America*, 118(2), pp. 1038-1054.
- LLISTERRI, J. (2013): *La representación fonética suprasegmental de corpus orales*. http://liceu.uab.cat/~joaquim/language_resources/spoken_res/Repres_fon_supraseg.html [14.05.2013].
- LLISTERRI, J. (2012): «Umbrals diferencials auditius», dins Llisterra, J., *Nociones básicas de psicoacústica*. http://liceu.uab.es/~joaquim/phonetics/fon_percept/psicoacustica/psicoacustica.html#umbrals_audicion/ [23.09.2012].
- LÓPEZ BOBO, M. J.; C. MUÑIZ CACHÓN; L. DÍAZ GÓMEZ; N. CORRAL BLANCO; D. BREZMES ALONSO i M. ALVARELLOS PEDRERO (2007): «Análisis y representación de la entonación. Replanteamiento metodológico en el marco del proyecto AMPER», dins J. Dorta (ed.): *La prosodia en el ámbito lingüístico románico*, Santa Cruz de Tenerife, La Página, pp. 17-34.
- MALMBERG, B. (1966): *La phonétique*, París, PUF.
- MARTÍNEZ CELDRÁN, E.; A. M. FERNÁNDEZ PLANAS i P. ROSEANO (2008): «Aproximación al estudio de la entonación de la región de Murcia: Caravaca de la Cruz y Bullas», dins A. Turculeț (ed.): *La variation diatopique de l'intonation dans le domaine roumain et roman*, Iași, Editura Universității «Alexandru Ioan Cuza », pp. 75-92.
- MARTÍNEZ CELDRÁN, E. i A. M. FERNÁNDEZ PLANAS (2003): «Taxonomía de las estructuras entonativas de las modalidades declarativa y 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*, México, El Colegio de México, pp. 267-294.

-
- MARTÍNEZ CELDRÁN, E. (2002): «Aproximación a una fonética no discreta», dins J. Díaz García (ed): *Actas del II Congreso de Fonética Experimental*, Sevilla, Universidad de Sevilla, pp. 35-48.
- MO, Y. (2008): «Duration and intensity as perceptual cues for naïve listeners' prominence and boundary perception», dins P. A. Barbosa, S. Madureira i C. Reis, (eds.): *Speech Prosody 2008 Proceedings*, Campinas, Brasil, pp. 739-742.
- NAVARRO TOMÁS, T. (1944): *Manual de entonación española*, Madrid, Guadarrama.
- PAMIES BERTRÁN, A.; A. M. FERNÁNDEZ PLANAS; E. MARTÍNEZ CELDRÁN; A. ORTEGA ESCANDELL i M. C. AMORÓS CÉSPEDES (2002): «Umbrals tonals en el espanyol peninsular», dins J. Díaz García (ed): *Actas del II Congreso de Fonética Experimental*, Sevilla, Universidad de Sevilla, pp. 272-278.
- PAMIES BERTRÁN, A. (1993): *Acento, ritmo y lenguaje (Elementos de prosodia general y comparada)*, tesi doctoral inèdita, Universidad de Granada.
- PIERREHUMBERT, J. (1980): *The Phonology and Phonetics of English Intonation*, tesi doctoral, Massachusetts Institute of Technology.
- PRIETO, P. (en premsa): «The Intonational Phonology of Catalan», dins S-A. Jun (ed.): *Prosodic Typology 2. The Phonology of Intonation and Phrasing*, Oxford, Oxford University Press.
- PRIETO, P.; L. AGUILAR; I. MASCARÓ; F. TORRES-TAMARIT i M. M. VANRELL (2009): «L'etiquetatge prosòdic Cat_ToBI», *Estudios de Fonética Experimental*, XVIII, pp. 287-309.
- PRIETO, P. i J. I. HUALDE (2012): «Towards developing a standard for prosodic annotation», presentació al *Workshop "Advancing Prosodic Transcription for Spoken Language Science and Technology"*, Stuttgart.
- PRIETO, P. (1999): *Entonació. Mètodes, teoria, models*, Barcelona, Ariel.
- RECASENS VIVES, D. (1991): *Fonètica descriptiva del català*, Barcelona, Institut d'Estudis Catalans.
- RIESZ, R. R. (1928): «Differential sensitivity of the ear for pure tones», *Physical Review*, 31(5): pp. 867-875.
-

- RIETVELD, A. C. i C. GUSSENHOVEN (1985): «On the relation between pitch excursion, size and prominence», *Journal of Phonetics*, 13, pp. 299-308.
- ROMANO, A. (1999): *Analyse des structures prosodiques des dialectes et de l'italien régional parlés dans le Salento (Italie): approche linguistique et instrumentale*, tesi doctoral, Université Stendhal Grenoble 3.
- ROSEANO, P.; M. M. VANRELL i P. PRIETO (en premsa): «The intonational phonology of Friulian and its dialects», dins S. Frota i P. Prieto (eds.) *Intonational variation in Romance*, Cambridge, Cambridge University Press.
- ROSEANO, P. (2012): *La prosòdia del friülà 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», *Proceedings of Interspeech 2010*, Makuhari, Chiba (Japó), pp.146-149.
- TURK, A. E. i J. R. SAWUSCH (1996): «The processing of duration and intensity cues to prominence», *Journal of the acoustical society of America*, 99(6), pp. 3782-3790.