

## Capítol 2. Aspectes fonètics de l'entonació

---

### 2.1. La fisiologia de l'entonació

El procés de producció dels sons sol dividir-se en tres períodes successius: el període respiratori, el període fonatori i el període articulatori. Primer de tot, perquè es produeixi qualsevol so s'ha de crear un corrent d'aire. L'emissió d'aquest corrent s'anomena iniciació aerodinàmica i s'engega des dels pulmons. Després, el període fonatori es produeix en la cavitat laríngia, on se situen les cordes vocals, les quals són les principals responsables de les variacions melòdiques de la parla. Finalment, el període articulatori es produeix en les cavitats supraglotal: en aquest període, la posició dels articuladors actius (llavis, llengua, maxil.lar inferior, vel del paladar) i dels passius (maxil.lar superior, paladar dur) és la principal responsable de la major part de variacions de qualitat del so.

El correlat acústic principal del to subjectiu és la freqüència fonamental ( $F_0$ ), que articulàtorialment es correspon amb la freqüència vibratòria de les cordes vocals produïda durant el període fonatori i es mesura en cicles per segon o hertzs (1 Hz = 1 cicle per segon, 1 cps). Com més ràpida és la freqüència de vibració de les cordes més agut és el to, i a l'inrevés. Ara bé, ¿com es desencadena fisiològicament la vibració de les cordes vocals?; i, ¿com controla el parlant l'elevació i el descens del to? Des de les investigacions de van den Berg la teoria coneguda amb el nom de **teoria mioelàstica** —o també "aerodinàmico-mioelàstica"— de la fonació constitueix el model més acceptat sobre el funcionament del procés fonatori. Aquesta teoria manté que la vibració de les cordes vocals es produeix per l'acció complementària de dos factors: les propietats mioelàstiques de la musculatura de la laringe i les relacions aerodinàmiques que s'estableixen entre les cavitats glotal i supraglotal. D'una banda, les forces musculars i moviments mecànics dels cartílags (tiroide, aritenoide i cricoide) i dels músculs de la laringe (sobretot del múscul cricotiroïdal) apliquen una tensió voluntària sobre les cordes que fan modificar la seva configuració espacial —fan que s'apropin o que s'allunyin— i incrementar el seu grau de tensió. D'altra banda, la diferència de pressió entre l'aire dels pulmons i el que es troba en les cavitats supraglotal origina una pressió subglotal positiva que fa iniciar la circulació d'un flux d'aire cap a les cavitats supraglotal. En aquest moment, si les cordes vocals es troben suficientment aprop, el flux fa que els plec musculars se separin i deixin passar una petita part de l'aire obstaculitzat; tot seguit els propis plec fan una força de succió (anomenada efecte Bernoulli) que els torna a unir, fet que completa un cicle de vibració. En resum, per tal que les cordes vocals puguin vibrar s'han de donar dues condicions principals: 1) aquestes han d'estar suficientment juntes (**condició mioelàstica**); i 2) hi ha de passar una quantitat d'aire suficient creada per les diferències de pressió supraglotal i subglotal (**condició aerodinàmica**). Finalment, si no hi ha cap modificació en l'ajustament de les cordes vocals, la vibració pot començar a produir-se de forma periòdica, tot iniciant el procés fonatori.<sup>1</sup>

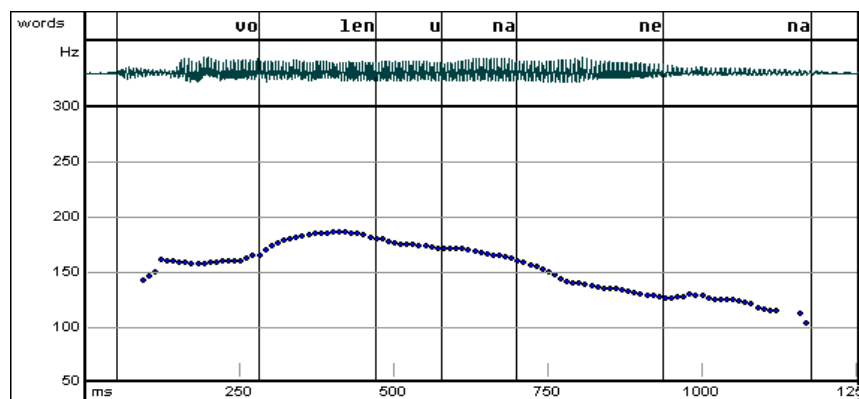
El parlant pot alterar voluntàriament la velocitat de vibració de les cordes, i, per tant, controlar el to de veu resultant. Així, en català, els parlants poden articular els enunciats amb una cadència descendent final (i produir un contorn declaratiu) o amb un ascens melòdic força marcat (i produir un contorn interrogatiu). La velocitat de les cordes vocals es regula per mitjà del reajustament dels músculs laringis, i molt

---

<sup>1</sup> Per a més referències sobre la producció entonativa, remetem el lector a les aportacions de Hirose (1997) i Stevens (1998).

especialment del múscul cricotiroide, que es considera com el principal responsable de la regulació del to. En llengües com l'anglès, el japonès i el xinès, hom ha constatat que l'activitat del múscul cricotiroide augmenta just abans de l'elevació del to, i a l'inrevés, la seva activitat minva abans de la davallada tonal. La contracció d'aquest múscul aplica una tensió sobre les cordes vocals que produeix un efecte d'allargament i de disminució de la seva massa vibratòria, la qual cosa fa accelerar la velocitat de vibració. El mecanisme actiu que produeix l'abaixament del to, en canvi, no ha estat tan ben descrit com el mecanisme d'elevació i encara està essent examinat (vg. Hirose 1997:133).

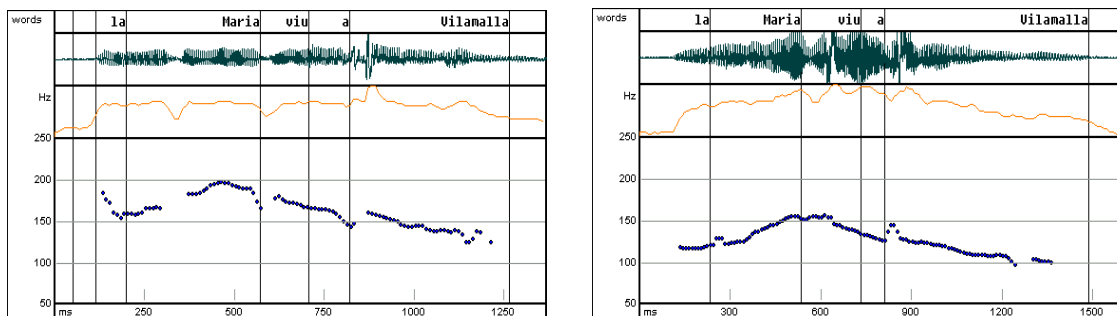
Actualment es considera que l'augment del flux de l'aire que passa per la glotis és relativament independent de la velocitat de vibració de les cordes. Així, encara que l'emissió d'una síl·laba tònica —produïda amb un volum superior d'aire que una d'àtona— s'ha associat típicament amb un increment dels valors d' $F_0$ , aquesta tendència també es pot capgirar segons el contorn melòdic en què s'insereixi aquesta síl·laba. En català, per exemple, la darrera síl·laba tònica d'un enunciat declaratiu (percebuda com la síl·laba portadora del màxim de prominència prosòdica de l'enunciat) és 'sorprenentment' una de les síl·labes amb els valors d' $F_0$  més baixos de la seqüència, tal com mostra l'esquema següent.



Un factor que resta fora del control del parlant i que determina parcialment la velocitat a què poden vibrar les cordes vocals és la pròpia fisiologia d'aquests músculs: la velocitat que puguin atènyer serà inversament proporcional a la seva longitud i al seu volum o gruix. Com més curtes i estretes, més ràpida podrà ser la seva vibració i més agut el sostre superior de l'interval tonal del parlant. Com sabem, la veu de cada individu presenta una **tessitura** determinada, és a dir, un espai o **camp tonal** que sol cobrir en situacions normals:<sup>2</sup> típicament, les dones presenten una tessitura que abraça des dels 180 fins als 400 Hz, considerablement més aguda que la dels homes (dels 80 fins als 200 Hz). Els nens poden fins i tot superar els nivells tonals dels parlants femenins. Els dos gràfics següents mostren els traçats de freqüència fonamental de la frase declarativa *La Maria viu a Vilamallà* emesos per dos locutors de català central, un de femení (esquerra) i un de masculí (dreta). La forma de les dues corbes és essencialment la mateixa: les síl·labes pretòniques es realitzen un to mitjà de la tessitura de cada parlant, que s'eleva considerablement durant la tònica (cf. *Maria*);

<sup>2</sup> "Terme musical derivat de l'italià *tessitura* (teixit, trama) que designa el fragment de l'escala sonora que millor convé a una veu determinada. (...) Cal no confondre-la amb l'extensió total o interval de les dues notes extremes." (de Candé 1967:182). Així, la classificació usual de les veus per al cant (soprano, contralt, tenor,...) es basa en els valors que presenta la tessitura mitjana del cantant en condicions normals. De forma general, el terme *camp tonal* designa l'espai de tonalitat que ocupa generalment un parlant (en anglès s'utilitza el terme "key").

després es produeix una davallada progressiva del to durant el cos central i terminal de la frase, cadència descendent que representa un dels trets més característics de l'entonació declarativa. La diferència principal entre els dos contorns, doncs, rau en l'espai tonal que ocupen: mentre que el del locutor femení ocupa un interval aproximat que va dels 135 als 200 Hz, el del locutor masculí ocupa un espai tonal més greu (dels 100 als 160 Hz) —el valor màxim de l'espai tonal de cada parlant es correspon amb l'altura del primer cim tonal i el valor mínim amb el darrer valor d' $F_0$  del contorn:



El fet que cada parlant ocupi un espai tonal propi evidencia que, a dreta llei, no podem comparar els valors absoluts d' $F_0$  de diferents parlants. En aquest sentit és alligador el consell que dóna Peter Ladefoged (1975:227) als dialectòlegs interessats en la descripció de les llengües tonals: "Quan treballem amb parlants o amb enregistraments de llengües tonals tingueu cura de no reproduir els tons exactes del parlant llevat que aquest presenti la vostra mateixa tessitura. Un dels problemes que tenia si imitava exactament els tons del meu informador era que sovint eren malinterpretats pels parlants d'aquesta llengua".<sup>3</sup> Per tal d'evitar el problema apuntat per Ladefoged, una pràctica força habitual en els estudis d'entonació ha estat la de comparar les corbes melòdiques d'un sol informador, atès que d'aquesta manera es garanteix una tessitura tonal estable que serveix com a punt de referència i facilita la interpretació de les diferències entre contorns.

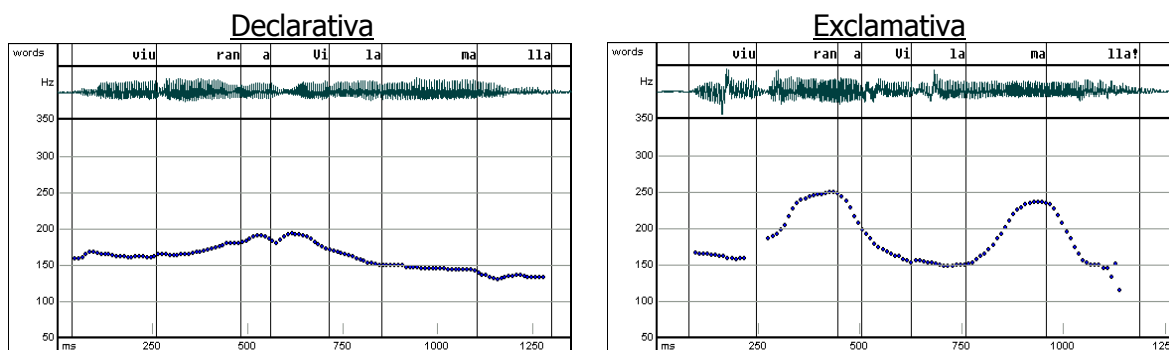
Cada individu sol usar només una part del seu espai tonal total quan parla normalment —típicament, el terç inferior—, encara que també el pot controlar voluntàriament i modificar-ne l'amplada amb motius lingüístics. Per exemple, en una pronúncia emfàtica o imperativa, el sostre superior de la tessitura se sol elevar mentre que el límit inferior (també anomenat línia de suspensió)<sup>4</sup> es manté gairebé invariable. Habitualment, una pronúncia eloqüent es caracteritza per l'ús d'excursions tonals àmplies, i a l'inrevés, una parla monòtona i uniforme sol presentar intervals reduïts i cims d'entonació baixos. Tècnicament, el **camp tonal** d'una inflexió (en anglès, "accent range" o "pitch range") es defineix com l'interval que hi ha entre la vall i el cim d'una inflexió ascendent o entre el cim i la vall d'una inflexió descendent.<sup>5</sup> Una de les diferències més òbvies entre l'entonació declarativa i l'exclamativa rau en el camp tonal que ocupen les inflexions. Fixeu-vos que contorn exclamatiu següent conté dues

<sup>3</sup> "If you are working with a friend or with recordings of a speaker of a tone language, be careful not to imitate his exact pitches, unless he has just the same pitch range as you normally do. (...) One of my problems in doing fieldwork with speakers of tone languages is that they often say I am mispronouncing a word when I imitate them fairly exactly." (Ladefoged 1975:227).

<sup>4</sup> Mascaró i Pons (1986:20) emprà el terme *línia de suspensió* per referir-se al límit inferior de la tessitura de cada parlant. En anglès s'empenen els termes "speaker baseline" o " $F_0$  minimum".

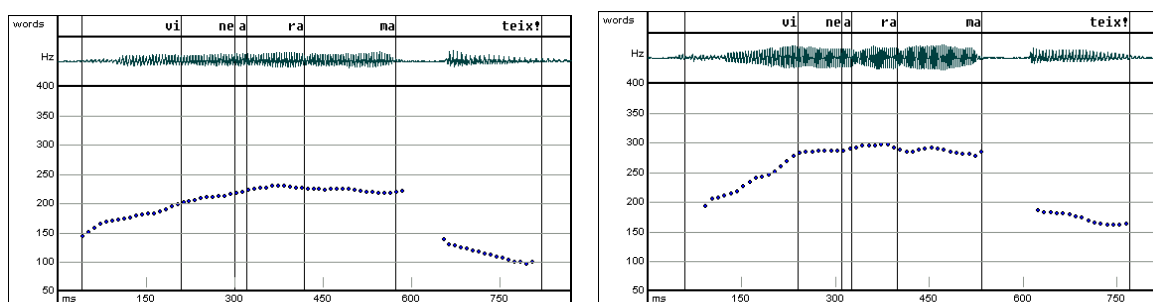
<sup>5</sup> Recordeu que el terme *camp tonal* també pot designar l'espai de tonalitat que sol ocupar un parlant en la seva parla habitual.

prominències melòdiques d'un camp tonal força ampli. A més, l'últim accent té un to baix (descendent) en la declarativa i alt (ascendent) en la declarativa emfàtica.



Convé remarcar que una pronúncia emfàtica no sempre es tradueix en un increment del sostre superior del camp tonal, sinó que també es pot manifestar amb un abaixament del nivell inferior. Un bon exemple d'aquest tipus podria ser el de l'entonació exhortativa transcrita en l'apartat 1.1: tant les síl·labes pretòniques com la tònica es pronuncien en un to greu, que esdevé més greu com més emfàtic és l'enunciat.

Els parlants també poden decidir de traslladar el **registre tonal**<sup>6</sup> i emprar una tonalitat més greu o més aguda. Si es canvia el registre, tant el límit inferior com superior de la tessitura habitual del parlant es traslladen de forma significativa cap amunt o cap avall —és per aquesta raó que Mascaró i Pons (1986:27) anomena aquest fenomen **transport de tessitura**. Els enunciats pronunciats en un registre agut s'han relacionat en diverses llengües amb la manifestació de tensió i d'agitació, i, per contra, els enunciats articulats en un registre greu solen tenir efectes més aviat tranquil·litzadors, i en alguns casos reprovatoris. Per exemple, podem emetre una oració imperativa com *iVine ara mateix!* utilitzant el nostre camp tonal habitual o bé desplaçant-lo cap amunt: fixeu-vos que el contorn emfàtic es realitza uns 100Hz més amunt, la qual cosa intensifica el to imperatiu del missatge.



Una altra de les funcions del transport de tessitura és el marcatge de canvis estructurals en el discurs: en català, els elements externs a la frase com els vocatius o els incisos no interrogatius (com *La Maria —la nena dels Aymeric— fa temps que no*

<sup>6</sup> Cal no confondre el terme *registre* amb el terme (potser més conegut) de l'àmbit de la pragmàtica que fa referència a la varietat lingüística usada en diferents situacions comunicatives (registre formal, registre informal,...).

ve) solen pronunciar-se amb un registre significativament més greu que la resta de l'enunciat; i un canvi de tema (com *Fa temps que no ve: ¿que li passa res?*) sol comportar una elevació local del registre tonal. De tota manera, cada llengua pot assignar significats específics als trasllats de registre tonal: en japonès, per exemple, l'ús d'una tonalitat aguda constitueix una marca d'educació.

Tot i que la configuració articulatòria supraglotal no modifica substancialment el to fonamental del so, s'ha observat que alguns moviments articuladoris sí que influeixen de forma lleu sobre la velocitat de vibració de les cordes vocals i produeixen el que es coneix amb el nom d'**efectes segmentals** o **efectes microprosòdics** (t Hart *et al.* 1990, de Pijper 1987). Tal com indica el seu nom, els efectes microprosòdics solen ser pertorbacions de poca magnitud —pràcticament negligibles— que se sobreposen als moviments més evidents i rellevants del contorn (també anomenats "macroprosòdics" o "macromelòdics"). Un dels efectes segmentals més ben estudiats és el de l'altura vocàlica: en igualtat de condicions, les vocals altes ([i], [u]) presenten valors superiors d' $F_0$  als de les vocals baixes ([a], [ə]), diferència que pot arribar a ser d'uns 25 Hz (Lehiste 1970:68-83, Ladd i Silverman 1984, Silverman 1987, Steele 1986). És per aquesta raó que alguns estudis han defensat que cada vocal es caracteritza per un **to fonamental** o **to intrínsec** ( $IF_0$ ). L'increment tonal durant l'articulació de les vocals altes s'explica per una raó fisiològica: els efectes que produeix l'elevació de la llengua sobre la glotis. Com que els músculs que formen la llengua estan units a la part superior de l'os hioide i, al seu torn, alguns músculs de la laringe estan units a la part inferior d'aquest os, aleshores una elevació de la llengua comporta l'elevació de la laringe i l'augment de la tensió dels seus músculs, fet que produeix una petita acceleració de la velocitat de vibració. Tot i això, una investigació recent de Ladd i Silverman (1984:34) constata que la influència de l'altura vocàlica sobre els valors d' $F_0$  disminueix a mesura que l'enunciat s'allarga: aquest efecte és més pronunciat en mots emesos aïlladament i es va atenuant progressivament en mots inserits en frases marc i en un discurs fins que és pràcticament nul en síl·labes àtones.

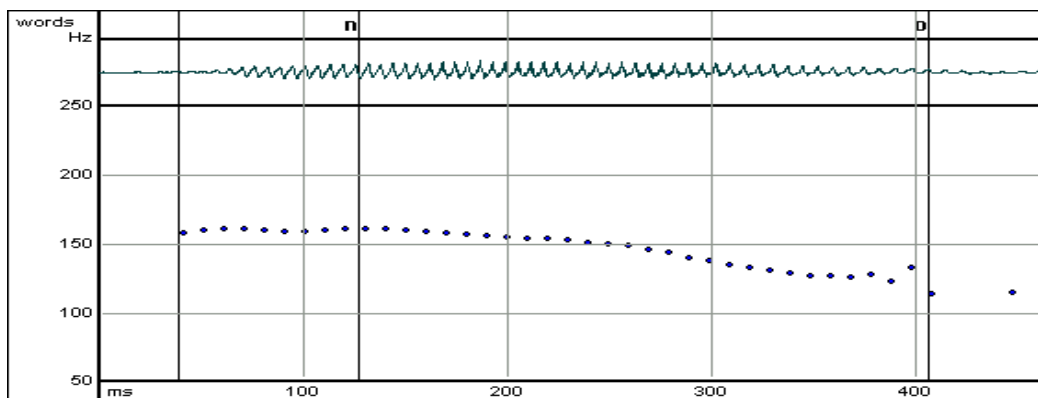
Un altre efecte segmental conegut és el que produeixen les consonants oclusives sonores ([b, d, g]): consisteix en una reducció momentània de la velocitat de vibració de les cordes en el moment posterior a l'oclusió. La formació d'una oclusió en la cavitat bucal durant el període vibratori fa que el nivell de pressió supraglotal augmenti considerablement, la qual cosa produeix un alentiment de la velocitat de vibració de les cordes (Quilis 1981:357, Silverman 1987). Normalment la dificultat de mantenir la sonoritat durant l'articulació dels segments oclusius es pot visualitzar en els traçats de to fonamental, que mostren una davallada local dels valors d' $F_0$  en algun punt després de l'oclusió.

## 2.2. La manifestació acústica de l'entonació

El resultat físic de l'obertura i el tancament de les cordes vocals és la vibració de les molècules d'aire que passen a través de la glotis, la qual produeix un senyal acústic complex i quasiperiòdic. Com és sabut, l'ona sonora complexa que genera la vibració de les cordes arriba a les cavitats supraglotal, on els articuladors adopten diverses configuracions articulatòries. Cadascuna d'aquestes cavitats actua com a ressonador i com a filtre, és a dir, amplifica la intensitat d'algunes freqüències de l'ona glotal i n'atenua d'altres, de manera que, segons la posició dels articuladors, s'obtenen senyals de qualitats vocàliques i consonàntiques diverses.

A finals del segle passat, la invenció del quimògraf va permetre la inscripció de les primeres ones sonores en un cilindre giratori (encara que sense poder sentir-les posteriorment). El parlant emetia els enunciats a través d'un tub connectat a un tambor (el qual començava a vibrar) mentre una agulla de gravació anava inscrivint les vibracions resultants en un full col·locat sobre un cilindre giratori. Com que aquest tambor actua com un filtre de freqüències altes, el traçat quimogràfic només enregistra les freqüències més baixes del so (el primer i el segon harmònics). Això fa que aquesta tècnica no sigui prou adequada per estudiar els timbres dels sons i en canvi permeti explorar els trets prosòdics: a principis de segle, Navarro Tomás i Pere Barnils la van utilitzar àmpliament en els seus estudis sobre duració i entonació del castellà i del català, respectivament.

L'oscil·loscopi o l'oscil·lògraf, d'aparició posterior, va permetre de captar les freqüències més altes de les vibracions sonores i dur a terme anàlisis més acurades del senyal.<sup>7</sup> L'**oscil·lograma** (també anomenat **forma d'ona**) representa l'amplitud del senyal acústic en l'eix de les ordenades i el temps en l'eix de les abscisses (vg. l'oscil·lograma de la seqüència *no*). En els segments sonors, cada obertura o tancament de les cordes vocals origina un pic de pressió d'aire que es tradueix en un cim d'amplitud. Fixeu-vos que en l'oscil·lograma següent s'aprecien els pics d'amplitud a distàncies més o menys regulars. Si volem calcular la freqüència vibratòria aproximada de les cordes comptarem el nombre de pics que ocorren en una determinada fracció de temps: 15 pics d'amplitud en un interval de 100 ms. —vg. l'espai delimitat per dues línies verticals grises entre 200-300 ms.—, la qual cosa ens indica que la vocal s'ha pronunciat amb una freqüència fonamental aproximada de 150 Hz (és a dir, 150 pics d'amplitud cada segon).

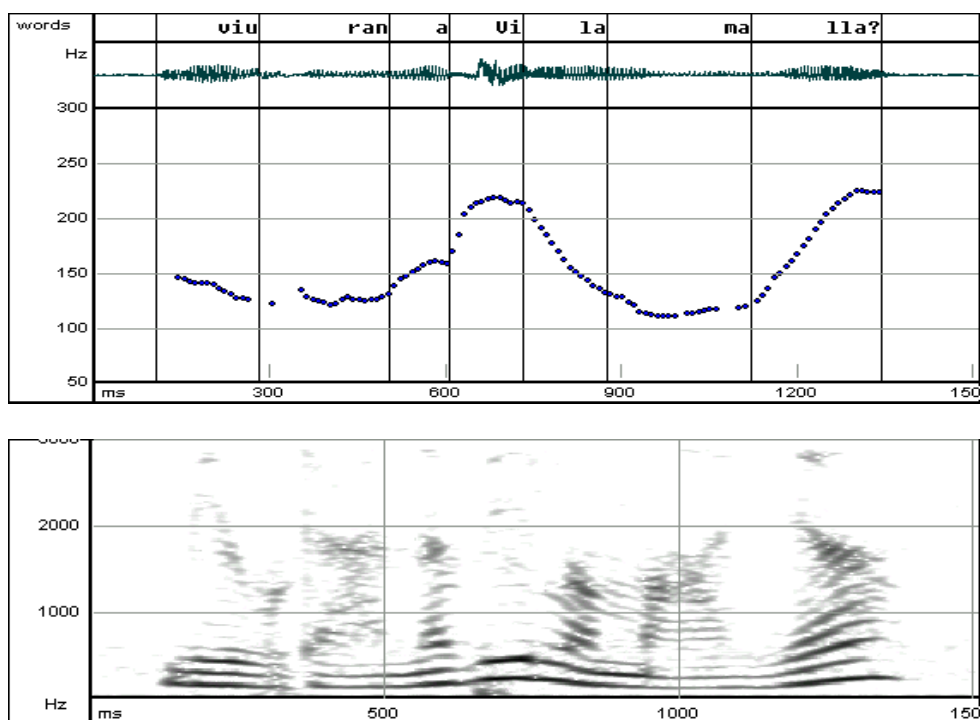


L'ona sonora és una ona complexa i periòdica que es pot descompondre matemàticament en una sèrie d'ones simples d'una determinada freqüència i amplitud mitjançant l'algorisme anomenat transformada de Fourier. Un dels trets distintius d'aquestes ones simples és el fet que presenten un valor de freqüència múltiple de l'ona de freqüència més baixa, també anomenada **fonamental** o **F<sub>0</sub>**. Les ones múltiples del fonamental s'anomenen **harmònics**. Així, si un so té una freqüència fonamental de 120 Hz, aleshores el primer harmònic tindrà un valor de 240 Hz, el segon de 360 Hz, etc., configurant així una seqüència harmònica que es va transformant constantment al llarg de l'emissió de l'enunciat. En la parla, tots els sons

<sup>7</sup> Els lectors poden trobar més referències sobre l'evolució d'aquests mètodes a León i Martín (1972), Llisterra (1991, 1996) i Quilis (1981).

sonors estan formats per ones periòdiques i presenten les propietats que hem esmentat. En canvi, els sons sords (com ara les fricatives sordes) són sons aperiòdics.<sup>8</sup>

Des dels anys quaranta als setanta, l'espectrògraf ha estat un dels instruments més emprats en la disciplina de la fonètica acústica. Aquest aparell és essencialment un analitzador espectral que obté un tipus de representació del senyal acústic conegut amb el nom d'**espectrograma** o **sonograma**, el qual permet visualitzar l'ona sonora en els dominis de temps, freqüència i amplitud. Un espectrograma enregistra l'evolució espectral del senyal en el temps: l'eix de les ordenades representa els valors de freqüència del so —que sol incloure l'interval de freqüència que ocupa la veu humana, de 0 fins a 8000 Hz— i l'eix de les abscisses l'evolució d'aquesta informació en el temps. Finalment, el grau de negror indica el nivell relatiu (i no mesurable) d'intensitat del senyal. Segons el filtre que s'utilitzi, hom obté espectrograms de banda ampla (amb filtres de 300 Hz) i de banda estreta (amb filtres de 45 Hz). La figura següent mostra el traçat d' $F_0$  i l'espectrograma de banda estreta de l'enunciat interrogatiu *¿Viuran a Vilamallà?* Com veiem, els espectrograms de banda estreta permeten d'apreciar l'evolució del fonamental i de cada harmònic per separat, i per això s'ha emprat de forma habitual per a l'exploració de l'entonació. Ara bé, com que l'ona fonamental és difícil d'inspeccionar visualment a causa dels seus valors baixos, les anàlisis espectrogràfiques de l'entonació solen prendre com a punt de referència harmònics més alts. En aquest cas, els valors absoluts del fonamental es calculen de forma indirecta dividint pel nombre múltiple corresponent, segons l'harmònic que es triï.



<sup>8</sup> Per a una explicació sobre la representació acústica de les ones sonores, vg. Ladefoged (1975) i Quilis (1981). També es pot trobar un resum en català a Recasens (1991b) i Pradilla (1998).

El càlcul manual dels períodes d' $F_0$  en un oscil·lograma o en un espectrograma és una tasca feixuga i no massa acurada. Actualment es disposa d'un bon nombre de programes informàtics d'extracció automàtica de corbes d' $F_0$  que representen un gran ajut per al transcriptor. Una de les tècniques més emprades, la tècnica d'autocorrelació, consisteix a analitzar un fragment curt del senyal (d'uns 40 ms.) i correlacionar-lo amb el resultat de l'anàlisi d'un fragment posterior, que sol situar-se a una distància de 2-20 ms. de l'inici de l'interval inicial.<sup>9</sup> L'anàlisi del senyal acústic extret del micròfon és el mètode més emprat actualment per a la medició del to fonamental, sobretot per la relativa comoditat que suposa el simple enregistrament de la parla i l'anàlisi acústica posterior en un ordinador personal. No es pot pas afirmar, però, que aquesta tècnica estigui exempta de problemes o que els algorismes de medició automàtica del fonamental siguin totalment satisfactoris. De fet, els traçats d' $F_0$  que s'obtenen amb aquests mètodes solen contenir errors. Una de les principals dificultats que tenen aquests programes és la delimitació dels períodes sords i els sonors —com sabem, la vibració de les cordes només és possible durant l'articulació de segments sonors. Un fenomen que també complica l'extracció dels valors del fonamental són els anomenats efectes segmentals, que, com ja hem advertit, introdueixen "distorsions" en el traçat d' $F_0$  (vg. §2.1). És per aquest motiu que hom aconsella que l'anàlisi acústica de l'entonació es basi en seqüències formades únicament per vocals i consonants sonants (semivocals i consonants líquides i nasals). Finalment, el nivell de soroll ambiental, la qualitat de l'enregistrament i el procés de digitalització posterior també influeixen en el grau de fiabilitat del càlcul de la corba fonamental.

Actualment una de les tècniques que proporciona resultats més acurats i precisos en la medició de la velocitat de vibració de les cordes vocals és l'electroglotografia (EGG) o laringografia. Aquesta tècnica es basa en l'anàlisi d'un senyal acústic que s'enregistra directament sobre la part externa de la laringe. El laringògraf (o glotògraf) és un aparell format per un parell d'elèctrodes que es col·loquen sobre el coll del parlant (per sobre del cartílag tiroide, el principal cartílag de la laringe). Quan s'envia un corrent elèctric molt feble d'un elèctrode a l'altre, l'aparell detecta les modificacions d'impedància que causen les cordes vocals sobre aquest corrent i n'enregistra les variacions d'amplitud en una representació anomenada glotograma. Aquest procediment d'extracció del fonamental es considera un dels més fiables actualment, atès que les dades que s'obtenen no presenten interferències o distorsions causades pel soroll ambiental o pels efectes supraglotsals.

### 2.3. La percepció de l'entonació

Com la intensitat o la durada, el **to** és una propietat perceptiva. Encara que el correlat acústic per excel·lència de la impressió auditiva del to és la freqüència fonamental (o la seva freqüència, si es tracta d'un to pur), convé tenir present que la sensació perceptiva de canvi melòdic pot ser originada no només per variacions en el contorn melòdic, sinó també per modificacions de la intensitat o d'altres paràmetres acústics (vg. Bolinger 1955, Crystal 1969:108). Conèixer els mecanismes subjacents de la percepció del to és una part important de l'estudi de l'entonació: pot ajudar, per exemple, a delimitar quins són els moviments tonals realment significatius de les corbes melòdiques, és a dir, aquells que l'oïent processa per extraure la informació que transmet l'entonació. Sabem que no totes les variacions de freqüència fonamental dels patrons entonatius són percebudes per l'oïda humana, que actua com a filtre

---

<sup>9</sup> Per a una revisió d'altres tècniques d'extracció automàtica del fonamental, vg. 't Hart *et al.* (1990:22-23).



transformador de l'ona i desestima bona part de les variacions que presenta el senyal acústic: presumiblement, doncs, les variacions fonètiques que resultin "filtrades" no seran rellevants per a la comprensió ni es relacionaran amb la transmissió d'informació.

Sabem que la percepció del nivell tonal per part dels oients actua de forma relativa i que només es pot dur a terme satisfactòriament en el cas que es puguin posar en relació tons contigus: així, els oients són pràcticament incapaçs de reconèixer en quin to s'emet una seqüència si aquesta conté un únic to estable (els errors poden arribar a ser d'entre 5 i 9 semitons). Però, quin és el canvi mínim d' $F_0$  que pot percebre l'oïda humana, és a dir, el **llindar diferencial**? Els resultats de diversos experiments demostren que els oients són capaços de discriminar diferències tonals tan petites com d'1 a 5 Hz (i els oients més sensibles, de menys d'1 Hz). Tot i això, els valors de llindar que s'obtenen varien en funció de les propietats del senyal acústic. La precisió en la sensació de modificació del to és directament proporcional a la durada de l'estímul: a mesura que l'estímul s'allarga la precisió augmenta. La complexitat del senyal també és important en la percepció tonal: d'una banda, si el senyal s'emet en un to constant el llindar obtingut és inferior que si s'emet en un to ascendent o descendent; igualment, mentre en un diftong el llindar sol ser d'uns 2 Hz, en una vocal de timbre invariable aquest pot arribar a ser de 0,3 Hz (Moore 1997:557). Finalment, el grau de sensibilitat a la variació tonal també depèn de la regió freqüencial en què se situa el senyal: en particular, és menor en senyals de freqüències inferiors a 1 kHz i superiors a 10 kHz. Convé assenyalar que les variacions melòdiques lingüísticament pertinents de la parla presenten una magnitud molt superior als nivells mínims de discriminació tonal enregistrats en els experiments psicoacústics: per exemple, l'excursió tonal típica d'un to ascendent en una oració declarativa sol ser d'uns 25 o 30 Hz, i, si és emfàtica, de 50 o 60 Hz (en funció també de factors com la tessitura del parlant).<sup>10</sup>

Tot i que cada parlant pronuncia els contorns melòdics amb la seva tessitura particular, aquests són percebuts de forma totalment equivalent pels oients de la llengua en qüestió. Lingüísticament, doncs, el que realment caracteritza un determinat contorn és la distància relativa que hi ha entre nivells tonals adjacents. Com que l'oïda humana és més sensible a les variacions de to que es produeixen en freqüències baixes, percebem com a equivalents els intervals tonals que hi ha entre 100 i 200 Hz, 200 i 400 Hz i 400 i 800 Hz, etc. És per aquesta raó que els contorns pronunciats per locutors amb tessitures greus es caracteritzen per excursions i intervals més reduïts que els pronunciats per locutors amb tessitures agudes. Encara que la unitat de freqüència més emprada és l'hertz (Hz), els investigadors interessats en la percepció prefereixen convertir-la a una base logarítmica: una de les més usades és el **semito** ('t Hart *et al.* 1990:24). La conversió a una escala logarítmica permet de comparar directament dues corbes tonals pronunciades per dos parlants amb tessitures de veu distintes: així, els tres intervals anteriors són equivalents a tretze semitons (o una octava).

Nombrosos estudis han constatat que, encara que els sons de la parla presentin una gran variació acústica, la percepció actua de forma discreta, "simplificant" l'input sonor. Per exemple, una vocal pot tenir una gran varietat de realitzacions fonètiques en funció del context lingüístic en què apareix (nivell accentual, entorn segmental, posició sil·làbica en l'enunciat, entre d'altres), i, no obstant això, la nostra oïda identifica tot aquest conjunt de realitzacions fonètiques amb una mateixa vocal. Aquesta

---

<sup>10</sup> Per a una explicació més detallada d'alguns aspectes perceptius de l'entonació, vg. Ladefoged (1975), Quilis (1981), Lieberman i Blumstein (1990) i Moore (1997), entre d'altres.

característica de la percepció humana es coneix amb el nom de **percepció categorial**. Un dels casos més ben estudiats de percepció categorial és el de la distinció entre oclusives sordes i sonores en funció del paràmetre de VOT (Voice Onset Time). Aquest paràmetre es defineix com la diferència temporal entre l'explosió d'una consonant oclusiva i l'inici de la vibració de les cordes vocals. Les consonants oclusives sordes i sonores varien al llarg d'un *continuum* de valors de VOT: per regla general, les oclusives sordes tenen un VOT positiu i les sonores un VOT negatiu (és a dir, les cordes comencen a vibrar abans de l'explosió). Doncs bé, quan els oients senten estímuls resintetitzats amb increments progressius dels valors de VOT, els classifiquen en dos grups: tots els senyals que presenten un determinat grup de valors de VOT els identifiquen amb una consonant sonora i, passat un determinat xamfrà, amb una consonant sorda. Així doncs, es posa de manifest que, malgrat la variació gradual que trobem en els valors de VOT, la percepció opera de forma binària i només és capaç de separar el contrast sord/sonor.

Pel que fa a l'entonació, mentre els valors d' $F_0$  van variant de forma gradual al llarg de l'emissió de l'enunciat, l'oient en canvi percep nivells tonals estables sobre cadascuna de les síl·labes que el componen. Fins fa poc manipular artificialment el contorn d' $F_0$  no era una tècnica a l'abast de tothom, i això explica en part la manca d'experiments sobre la percepció entonativa —per contra, la percepció accentual ha estat estudiada més a fons; vg. §4.2. No obstant això, podem destacar la tasca duta a terme per l'escola holandesa de l'entonació, un dels models que més s'ha interessat pels mecanismes subjacents de la percepció del to (vg. §7.3.3). Un dels projectes més reeixits d'aquest corrent ha consistit en la recerca de les unitats fonològiques del danès per mitjà d'una tècnica d'estilització de traçats de to fonamental: aquest mètode consisteix a anar suprimint aquelles fonts de variació tonal que són irrellevants (és a dir, aquelles que no produeixen cap efecte perceptiu evident), fins arribar als moviments tonals més elementals de la corba. Per exemple, les variacions micromelòdiques, modificacions pràcticament imperceptibles que sorgeixen com a conseqüència de la interacció entre el component segmental i la melodia, són les primeres que s'eliminen en el procés d'estilització. Per altra banda, estudis de percepció com els de Ward i Hirschberg (1985), Hirschberg i Ward (1992) i Grice (1995) han ajudat a avançar en la identificació dels trets tonals perceptivament significatius de la corba melòdica. Remetem el lector al capítol 7, en què es descriuen alguns d'aquests experiments i la seva importància en la demostració de la rellevància fonològica de determinats trets melòdics.