

Demestre, J., Llisterri, J., Riera, M., & Soler, O. (2006).
La percepció del llenguatge. In O. Soler (Ed.), *Psicologia
del llenguatge*. (pp. 35-114). Barcelona: Editorial UOC.
http:
//liceu.uab.cat/~joaquim/publicacions/Demestre_
Llisterri_Riera_Soler_06_Percepcio_Parla.pdf

Capítol II

La percepció del llenguatge

Josep Demestre Viladevall, Joaquim Llisterra Boix, Montserrat Riera Masjoan i Olga Soler Vilageliu

El segon capítol se centra en el processament del llenguatge que es duu a terme des de la captació del senyal sonor fins a l'anàlisi sintàctica de l'oració, passant pel reconeixement de les paraules. L'hem anomenat "Percepció del llenguatge" perquè es considera que en aquests tres estadis les representacions mentals que hi intervenen són essencialment de tipus lingüístic.

En el primer apartat, "Percepció de la parla", s'estudia aquest procés des d'una perspectiva fonamentalment fonètica. Veurem com s'analitzen els indicis acústics de l'ona sonora i se'n fan diferents abstraccions, per acabar amb la presentació de diferents models que descriuen com el continu sonor se segmenta en unitats lingüístiques. Aquest procés s'il·lustra al llarg de l'apartat fent servir les representacions del senyal sonor habituals en els estudis fonètics.

L'apartat següent, "Reconeixement de paraules", presenta com duem a terme el reconeixement de les paraules. En primer lloc, es descriuen les unitats de representació lèxiques i la informació que contenen, per passar després a la presentació de diferents models teòrics que descriuen el procés de reconeixement de paraules auditiu i visual i la interacció d'aquestes dues modalitats sensorials durant la lectura. Diferents gràfics contribueixen a aclarir les explicacions.

El tercer apartat, "Sintaxi" presenta el processament de la sintaxi. Els primers apartats presenten el nivell sintàctic de processament i les característiques més importants dels enfocaments teòrics que han dirigit la recerca empírica en aquest àmbit i els models proposats. De seguida s'exposen dos models cabdals de marcs teòrics oposats. Finalment es descriuen detalladament alguns dels efectes detectats mitjançant investigacions empíriques. Nombrosos exemples ajuden a la comprensió dels continguts.

1. Percepció de la parla

1.1. La parla com a senyal transmissor d'informació lingüística

1.1.1. La parla com a senyal sonor

Emetre un enunciat requereix, en la darrera etapa del procés de producció de la parla, el moviment coordinat de tota una sèrie d'estructures anatòmiques com ara els pulmons i els músculs implicats en la respiració, la laringe i les cordes vocals, la llengua, els llavis, el vel del paladar i la mandíbula inferior. El resultat és una ona sonora, producte del corrent d'aire originat als pulmons i modulada pel tracte vocal. Des d'aquesta perspectiva, la parla es pot considerar com un senyal sonor que emetem per tal de transmetre informació a un receptor.¹

D'una banda, aquesta ona sonora conté tota la *informació lingüística* necessària perquè el nostre interlocutor compregui el missatge que li volem fer arribar. De l'altra, també aporta dades sobre l'estat d'ànim de la persona, el grup social al qual pertany, la seva procedència geogràfica, les seves intencions comunicatives, etc. En aquest sentit, diem que l'ona sonora vehicula *informació extralingüística*.

No tota la informació lingüística que conté l'ona sonora té la mateixa importància. Hi ha diferències entre els sons que fan canviar el significat de les paraules i que s'anomenen *fonològiques*. Quan en una llengua un determinat so fa variar el significat d'una paraula en substituir-lo per un altre, diem que aquest so és un *fonema* i el representem entre barres inclinades (/ /).

Considerem les paraules *bata, bada, bava, bassa, barra* i *vaca*. Fixeu-vos que canviant únicament un so, en podem variar el significat. Les diferències entre *t /t/, d /d/, v /b/, s /s/, rr /r/* i *c /k/* són fonològiques, ja que substituint un so per un altre s'altera el significat de la paraula. Per tant, /t, d, b, s, r, k/ són fonemes en català.

En canvi, hi ha casos en què un so es pot pronunciar de maneres lleugerament diferents sense que per això variï el significat de la paraula que el conté. En aquest cas, parlem de diferències fonètiques, ens referim als sons com a *al·lòfons* i els representem entre claudàtors ([]).

1. En l'apartat 1.4 del capítol V trobareu una explicació dels mecanismes articularis en el marc dels processos implicats en producció de la parla.

Pronuncieu en veu alta la paraula *dida* i observeu les diferències entre la primera *d* i la segona. Notareu que en la primera premeu la punta de la llengua contra la zona que queda al darrere de les dents incisives, mentre que en la segona no la premeu del tot, tan sols l'aproveu a aquesta zona, de manera que queda una mica d'espai per on passa l'aire. Però si substituïu una *d* per l'altra, el significat de la paraula no varia. Per això, les diferències entre aquests dos sons són fonètiques i diem que el fonema /d/ té dos al·lòfons: un d'oclusiu [d] (en què es tanca el pas de l'aire) i un d'aproximant [ð] (en què deixem passar una mica d'aire pel canal que queda mig obert entre els articuladors que s'apropen).

En resum, la parla és una ona sonora, resultat dels moviments articuladoris, que vehicula informació lingüística i extralingüística. Entre els sons de la parla hi ha diferències significatives –fonològiques– i diferències que no fan canviar el significat –fonètiques.²

1.1.2. Problemes en la percepció de la parla

La funció de l'ona sonora que produïm en parlar és transmetre el contingut que ha d'arribar al receptor. Com a emissors, codifiquem en un senyal sonor la informació lingüística que volem donar; en canvi, quan actuem com a receptors, hem de dur a terme el procés invers i descodificar la informació que ens transmet l'ona sonora que ha emès el nostre interlocutor.

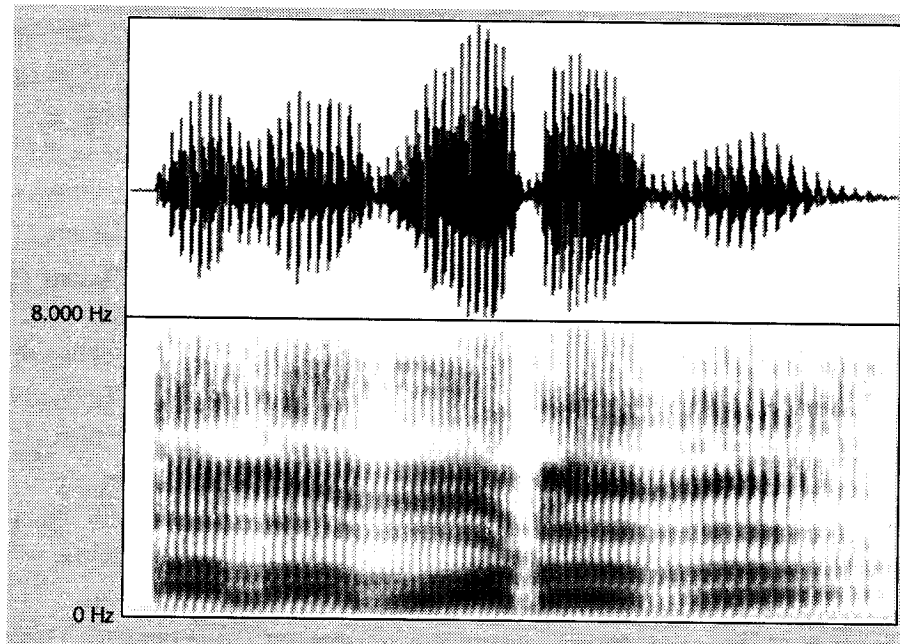
Aquest procés de *descodificació* o, altrament dit, d'extracció d'informació lingüísticament significativa, presenta dos problemes que exposarem a continuació: el de la segmentació i la linealització, d'una banda, i el de la invariància i la variació, de l'altra.

En primer lloc, cal ser conscients que contràriament al que ocorre en la llengua escrita, en la parla no hi ha separacions entre els sons ni entre les paraules. És cert que en ocasions fem pauses per a respirar, per a organitzar el contingut del que diem o per a tenir temps de planificar el discurs, però el material fònic que queda entre dues pauses sol consistir en més d'una paraula.

2. En les pàgines del Grup de Fonètica del Departament de Filologia Espanyola de la Universitat Autònoma de Barcelona, podreu trobar els inventaris de fonemes i al·lòfons tant del català (vocals: http://liceu.uab.es/~joaquim/publicacions/IPA_Gifs/IPA_vow_cat.html i consonants: http://liceu.uab.es/~joaquim/publicacions/IPA_Gifs/IPA_cons_cat.html) com del castellà (vocals: http://liceu.uab.es/~joaquim/publicacions/IPA_Gifs/IPA_vow_sp.html i consonants: http://liceu.uab.es/~joaquim/publicacions/IPA_Gifs/IPA_cons_sp.html), la seva transcripció mitjançant l'alfabet fonètic internacional: <http://www.arts.gla.ac.uk/IPA/ipachart.html>, que és el que fem servir en aquest apartat, i una bibliografia que us orientarà pel que fa a la descripció fonètica i fonològica de totes dues llengües.

La figura 2.1. mostra la forma de l'ona sonora i l'espectrograma corresponents a l'enunciat *a la vora de*. En aquesta il·lustració i en les següents, la forma d'ona representa com canvia l'energia sonora o la intensitat (eix vertical) d'un enunciat al llarg del temps (eix horitzontal), mentre que l'espectrograma permet de visualitzar en quines bandes freqüencials (eix vertical) hi ha energia sonora (representada mitjançant una escala de grisos, de manera que les zones amb més energia es veuen en un to més fosc), i combina aquesta informació amb el temps (eix horitzontal). Observeu que ni entre paraules ni entre sons no es donen interrupcions.

Figura 2.1. Forma d'ona (part superior) i espectrograma (part inferior) de l'enunciat "a la vora de"



Per aquest motiu, se sol caracteritzar la parla com un *continu sonor*. Una de les operacions que ha de fer el receptor per a descodificar el missatge és la *segmentació*; és a dir, ha de descompondre aquest continu en unitats lingüístiques més petites, tant si són paraules, com síl·labes o fonemes, per tal d'extreure'n la informació rellevant que permeti de comprendre el missatge.³

En segon lloc, també cal adonar-se que els òrgans de la producció de la parla es mouen de manera contínua, passant gradualment de la posició necessària per

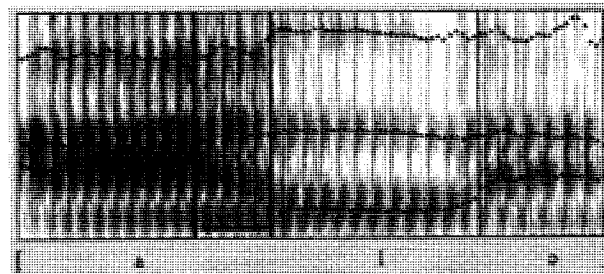
3. En l'apartat 1.4.1. d'aquest capítol s'estudia quines són aquestes unitats.

a produir un so a la pròpia del que ve a continuació. De fet, hi ha una influència mútua entre els sons, que es coneix com a *coarticulació*.⁴

La coarticulació es manifesta també en el senyal sonor, ja que, com hem explicat, aquest és conseqüència dels moviments del tracte vocal.⁵ La continuïtat d'aquests moviments fa que en un mateix instant del senyal hi hagi simultàniament informació corresponent a més d'una unitat lingüística. Aquest fet planteja una nova dificultat a l'hora de descodificar el missatge, ja que la informació que necessitem no apareix de manera seqüencial o lineal, sinó que podríem dir que es troba superposada. Aquest fenomen es coneix com el problema de la *linealització*.

La figura 2.2 mostra l'espectrograma de la síl·laba *ala*. Els punts assenyalen la trajectòria dels formants, que són la manifestació acústica dels moviments graduals del tracte vocal en passar de [a] a [l] i de [l] a [ə]. La zona entre les línies negres es coneix com a *transició*, i conté informació conjunta de la [l] i de la [a], ja que reflecteix el moviment dels articuladors en desplaçar-se des de la configuració pròpia de [a] a la configuració característica de [l].

Figura 2.2. Espectrograma de la síl·laba *ala*



Les línies gruixudes assenyalen la transició entre [a] i [l], una zona de l'ona sonora que conté informació simultània sobre els dos sons.

Tant el fet que no hi hagi separacions entre les unitats lingüístiques que configuren el senyal sonor portador d'un missatge –el problema de la *seg-*

4. Trobareu més explicacions sobre la coarticulació en l'apartat 1.4. del capítol V.

5. Podeu observar la filmació amb raigs X d'un parlant suec pronunciant el mot *pion* ('peònia', en català) que trobareu en les pàgines del Departament de Lingüística de la Universitat de Lund (http://www.ling.su.se/staff/ericdotter/projects/pion_fp11.htm). Fixeu-vos en la continuïtat del moviment dels articuladors. També podeu veure les imatges una per una (http://www.ling.su.se/staff/ericdotter/projects/pion_all.htm). Trobareu altres filmacions similars en les pàgines de la Queen's University a Kingston (http://psyc.queensu.ca/~munhallk/05_database.htm).

mentació– com la superposició d'informació sobre més d'una unitat en un determinat punt del senyal –el problema de la *linealització*–, constitueixen dos fenòmens que s'han d'explicar en el marc de les teories de la percepció de la parla.⁶

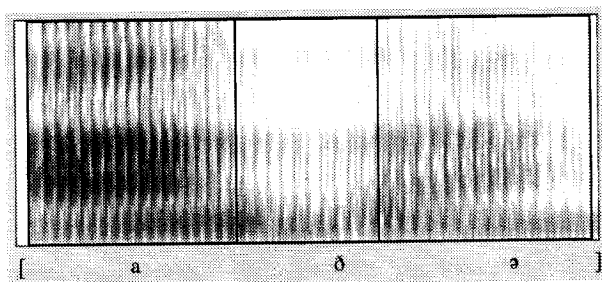
Per tot el que acabem d'explicar, la parla és un continu sonor. La coarticulació entre sons fa que en un mateix segment del senyal sonor es trobi informació simultània sobre més d'una unitat lingüística.

Com assenyalàvem en l'apartat anterior, hi ha diferències fonètiques entre els sons que no afecten el significat. A l'hora de comprendre un enunciat, aquestes són, precisament, les que d'alguna manera hem d'abstreure. La informació important no és pas la diferència entre una [b] oclusiva –com a *vas*– o una [β] aproximant –com a *cava*–, sinó la que hi ha entre una /b/ i una /k/, que ens pot fer entendre *vas* o *cas*.

Des d'aquest punt de vista, doncs, saber una llengua vol dir dues coses: en primer lloc, ser capaç de processar les diferències entre sons que són significatives; en segon lloc, agrupar en una única categoria –els fonemes– sons que s'articulen de manera diferent –els al·lòfons. Aquesta darrera capacitat està estretament relacionada amb el fenomen de la percepció categòrica.

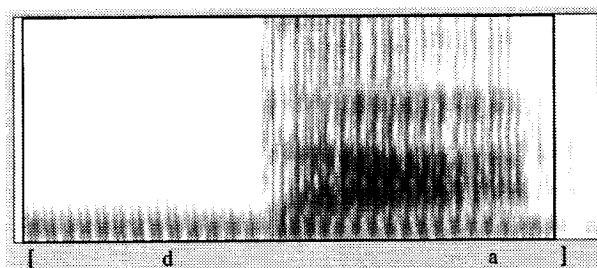
Les diferències en l'articulació també es manifesten en l'ona sonora i, per tant, incideixen en la percepció. En les figures 2.3 i 2.4 podeu observar la representació acústica en forma d'espectrograma de [d] oclusiva –com la primera de la paraula *dau*– i de [ð] aproximant –com la que pronunciaríeu a *cada*.

Figura 2.3. Espectrograma de la síl·laba *ada*, que mostra una [ð] aproximant



6. Podeu consultar l'apartat 1.5.1 d'aquest capítol.

Figura 2.4. Espectrograma de la síl·laba *da*, que mostra una [d] oclusiva



Hi ha diferències entre els sons que es deuen al context fonètic en què apareixen. Una consonant, per exemple, s'articula cap a la part més anterior del tracte vocal quan va seguida de [i] i cap a una part més posterior quan va seguida de [u], per influència del lloc d'articulació anterior o posterior de la vocal, o dit d'una altra manera, per efecte de la coarticulació amb el so que la segueix.

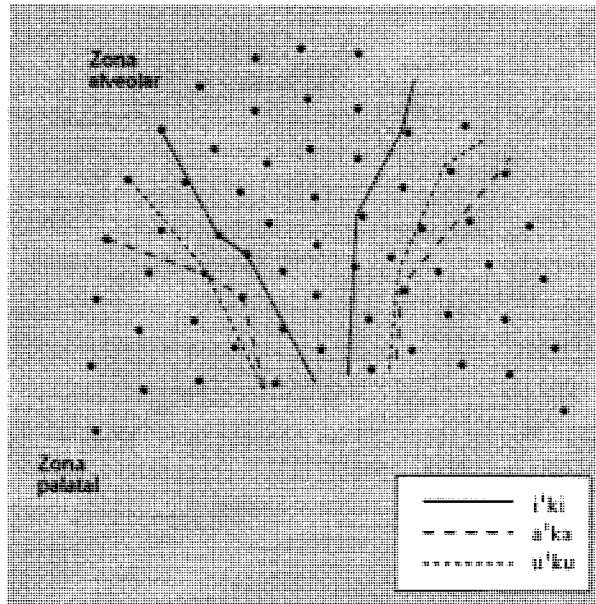
En la figura 2.5 podeu veure un electropalatograma, que és un document d'anàlisi articulatòria que mostra el contacte entre la llengua i el paladar en el moment de pronunciar un so. La zona que queda per darrere de cadascuna de les línies correspon al contacte entre la llengua i el paladar en pronunciar [k] amb les vocals [a, i, u]. Observeu que en [iki], [aka] i [uku] la zona de contacte és diferent, més anterior (és a dir, cap als alvèols) en [iki] i més posterior (és a dir, cap al vel del paladar) en [uku]. Les diferències en les zones de contacte es deuen a la coarticulació amb les vocals que envolten la consonant. Proveu de pronunciar vosaltres mateixos aquestes combinacions de sons i noteu com s'avança o s'endarrereix la llengua segons la vocal que acompanya la [k].⁷

Un mateix so també es pot pronunciar de manera diferent –i, per tant, tindrà propietats acústiques diferents– si es troba al principi o al final d'un enunciat, si està accentuat o no, o si es parla més ràpid o més lentament, per esmentar alguns factors que incideixen en la *variació fonètica*. També són diferents des del punt de vista acústic les realitzacions fonètiques d'un mateix so que es troben en parlants de les diferents variants geogràfiques d'una mateixa llengua, i ho són igualment les veus d'homes, de dones i de nens. En canvi, malgrat aquesta

7. Trobareu més informació sobre l'electropalatografia en les pàgines del Laboratori de Fonètica de la Universitat de Califòrnia a Los Angeles: <http://www.linguistics.ucla.edu/faciliti/facilities/physiology/epg.html>.

aparent variació, no tenim dificultats per entendre el contingut del missatge que rebem.

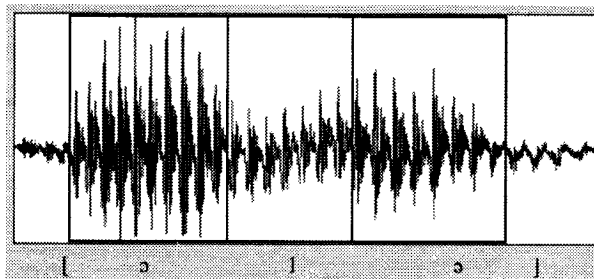
Figura 2.5. Electropalatograma de les seqüències [iki] [aka] i [uku]



Font: Recasens, D. (1986). *Estudis de fonètica experimental del català oriental central* (p. 45). Barcelona: Publicacions de l' Abadia de Montserrat (Biblioteca Milà i Fontanals, 6).

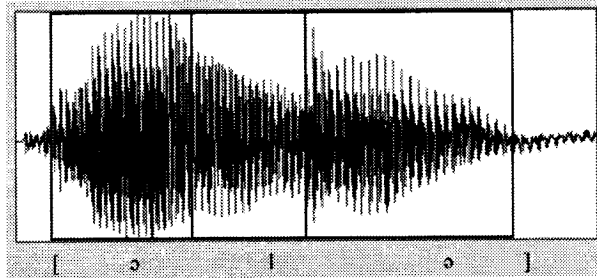
Les figures 2.6 i 2.7 mostren la forma d'ona de la paraula *hola* pronunciada per dos parlants diferents, un de masculí (figura 2.6) i un de femení (figura 2.7). Podeu observar que hi ha diferències notables en la forma d'ona, tot i que se segueix un patró general.

Figura 2.6. Forma d'ona de la paraula *hola* pronunciada per un parlant masculí



Les marques verticals indiquen la separació entre els tres sons que la formen.

Figura 2.7. Forma d'ona de la paraula *hola* pronunciada per un parlant femení



Les marques verticals indiquen la separació entre els tres sons que la formen.

Com assenyàvem al principi, la variació fonètica en els sons i la manera com som capaços de fer-ne abstracció és un fet que han d'explicar els diferents models de percepció de la parla. Aquest fenomen, que es coneix com el problema de la invariància, també s'ha estudiat des de la perspectiva de la fonètica, disciplina que s'ocupa de descriure les propietats generals dels sons de la parla i la manera com canvien segons el context, el parlant i altres factors que hem esmentat. Se sol distingir entre la *variació interlocutor*, que es dona entre diferents parlants, i la *variació intralocutor*, que es troba en el mateix so emès pel mateix parlant en contextos o en moments diferents.

En resum, en el procés de percepció de la parla hem d'identificar com a elements invariants fragments del senyal que articulatòriament i acústicament són diferents.

1.1.3. Audició i percepció de la parla

El senyal sonor que resulta del moviment dels articuladors en el procés de producció de la parla arriba a l'oïda del receptor i perquè es pugui dur a terme la percepció, ha de passar primer per una sèrie de transformacions. El sistema auditiu és l'encarregat de convertir l'estímul sonor en informació assimilable per les àrees del cervell especialitzades en el processament de la parla, de manera que sense audició, no hi pot haver percepció.

En primer lloc, l'oïda interna recull el so, l'oïda mitjana l'amplifica o el redueix, segons calgui, i d'allà passa a l'oïda interna, on el moviment vibratori de les molècules d'aire que confegeixen l'ona sonora es transforma en moviments dels

líquids que es troben a l'interior de la còclea i que fan oscil·lar la membrana basilar, una estructura anatòmica que conté l'òrgan de Corti, essencial en l'audició. Allà és on es produeix l'excitació de les cèl·lules nervioses connectades al nervi auditiu, encarregat de portar l'estímul fins al cervell. En conjunt, l'audició es pot entendre com el resultat de la transformació del senyal sonor –molècules d'aire en vibració– en estímuls que processa el còrtex auditiu –senyals elèctrics–, i en aquest sentit podem dir que és un procés fisiològic.⁸

Escoltar, però, no vol dir necessàriament entendre. Quan sentim una llengua que no coneixem, l'audició es duu a terme normalment, però en canvi no podem parlar de percepció, ja que no som capaços d'identificar i discriminar els sons per a extreure el significat que porta l'ona sonora. L'audició és responsable, per exemple, que ens adonem que les ones sonores dels segments [a] i [u] són físicament diferents; en canvi, quan donem valor lingüístic a aquests dos sons i els identifiquem com a vocals d'una determinada llengua que poden diferenciar paraules –*sac* i *suc* en català–, estem activant les nostres capacitats perceptives. La percepció de la parla és, per tant, un procés cognitiu que implica coneixement lingüístic i que es fa en àrees especialitzades del cervell.

Caldria encara distingir la percepció, entesa en termes purament fonètics i fonològics, de la comprensió, que requereix, per part del receptor del missatge, coneixement lèxic, sintàctic, semàntic i pragmàtic, i també un cert coneixement del context extralingüístic en què s'emet un enunciat. En la resta d'apartats d'aquest capítol s'explica de quina manera es fa la comprensió.

Cal recordar, doncs, que l'audició és un procés fisiològic, mentre que la percepció és un procés cognitiu.

1.2. La percepció de la parla

1.2.1. Els indicis acústics en la percepció de la parla

Com hem explicat, la tasca que l'oïdor ha de fer quan li arriba una ona sonora portadora d'un missatge és descodificar-la amb la finalitat d'identificar les unitats

8. Si voleu conèixer més detalls sobre el procés de l'audició, consulteu l'apartat 5.3 de Hayward (2000) o l'apartat Hearing del capítol 6 de Borden et al. (1994).

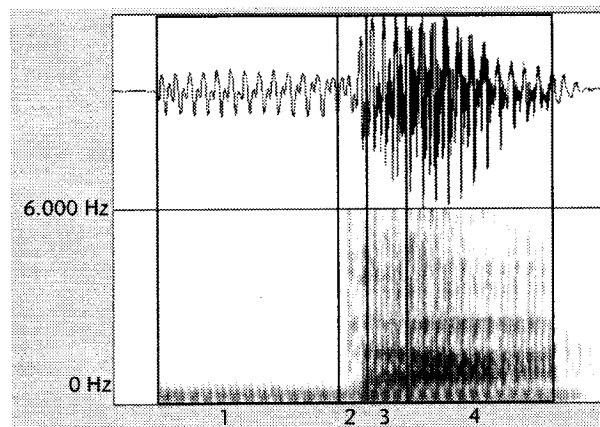
lingüístiques que conté. Per a dur a terme aquesta operació, a més de tenir en compte tota la informació que proporciona el context en què s'emeta un enunciat, el receptor s'ha de basar en la informació acústica que és present en el senyal.

Les recerques dels Laboratoris Haskins a New Haven, Connecticut,⁹ van posar de manifest, ja al final dels anys cinquanta, que cada classe de sons de la parla es caracteritza per un conjunt d'*indicis acústics* que permeten d'identificar-lo en el senyal sonor i distingir-lo d'altres sons.

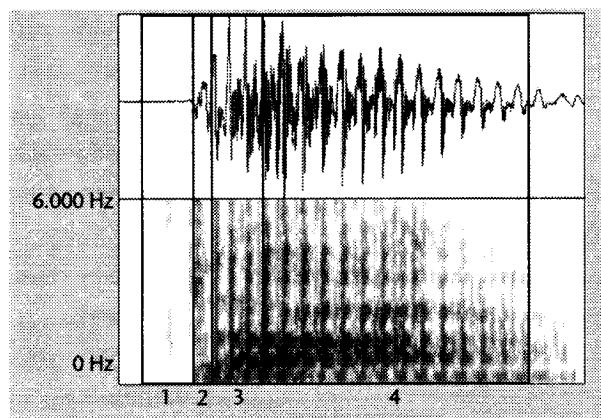
Ens intentarem apropar a la definició d'índex acústic exemplificant-la amb dos indicis clàssics: la sonoritat i les transicions dels formants.

Començarem per la *sonoritat*, observant les consonants [t] i [d], l'espectrograma de les quals es representa en les figures 2.8 i 2.9. En la figura 2.8 es pot veure que en [da] apareix primer una zona amb energia sonora (1), que correspon a la vibració de les cordes vocals mentre s'articula la [d], amb la llengua tocant la part posterior de les dents incisives. Després apareix el que s'anomena *l'explosió de la consonant* (2), que és el moment en què l'aire surt dels llavis en obrir-se el canal bucal. A (3) s'aprecien els moviments dels formants, que corresponen als del tracte vocal per passar de la configuració articulatòria de [d] a la de [a]. I finalment, a (4) es troben els formants corresponents a la vocal.

Figura 2.8. Forma d'ona i espectrograma de la síl·laba [da]



9. Podeu consultar les pàgines dels Laboratoris Haskins en què es descriuen els experiments fets amb el Pattern Playback, treballs que constitueixen l'inici de les recerques sobre els indicis acústics emprats en la percepció de les diferents classes de sons: <http://www.haskins.yale.edu/featured/pat-play.html>.

Figura 2.9. Forma d'ona i espectrograma de la síl·laba [ta]

Si compareu la representació acústica de [da] (figura 2.8) amb la de [ta] (figura 2.9), us adonareu que a l'inici de [ta] (1) no hi ha energia acústica, ja que les cordes vocals no vibren mentre tanquem el pas de l'aire per la cavitat bucal amb la llengua en contacte amb la part posterior de les incisives; (2) correspon al moment de l'explosió de [t], quan la llengua deixa d'interrompre el pas de l'aire i tot el que s'havia acumulat en la fase de tancament surt d'una manera relativament sobtada, i (3) correspon als moviments dels articuladors en passar de [t] a [a].

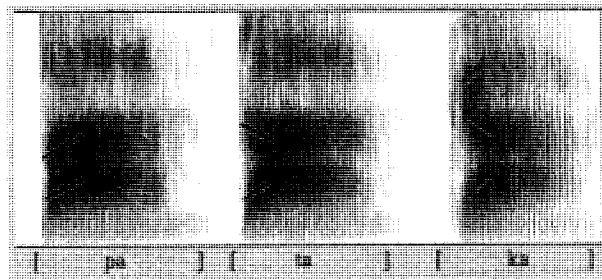
Aquesta energia sonora –anomenada *barra de sonoritat*– que s'observa a [d] i que, en canvi, no apareix a [t], constitueix un índex acústic que, juntament amb d'altres, fem servir en la percepció per a distingir les consonants sonores –les que es produeixen amb vibració de les cordes vocals– de les sordes –aquelles en què no vibren les cordes vocals.

A continuació, veurem un segon índex acústic, relacionat en aquest cas amb el lloc d'articulació de les consonants: les *transicions* dels formants.

En la figura 2.10 es mostra l'espectrograma de les síl·labes [pa, ta, ka], formades per una consonant oclusiva sorda i una vocal, en què s'han marcat les transicions –moviments dels formants– de la vocal a la consonant. Adoneu-vos que els moviments dels formants són diferents en les tres consonants, ja que reflecteixen els canvis en la configuració del tracte vocal en moure's des de la posició pròpia de cada consonant a la de la vocal: en [p] el tancament del pas de l'aire propi dels oclusius es produeix en els llavis; en [t], amb el

contacte de la part anterior de la llengua darrere les dents incisives, i en [k], amb la part posterior de la llengua tocant el vel del paladar. Per tant, els moviments que cal fer per a arribar fins a la posició necessària per a l'articulació de [a] són diferents per a cadascuna de les tres consonants i també ho són les transicions que observem en l'espectrograma.

Figura 2.10. Espectrograms de les síl·labes [pa, ta, ka]



Les transicions dels formants s'han assenyalat amb una línia.

Les transicions són, doncs, un indicatiu acústic que, en combinació amb altres indicis, ens permeten de recuperar en el senyal sonor informació sobre el lloc d'articulació de la consonant.

En conclusió, els indicis acústics que es troben en el senyal sonor són elements que ens permeten d'identificar i discriminar els diferents sons de la parla.¹⁰

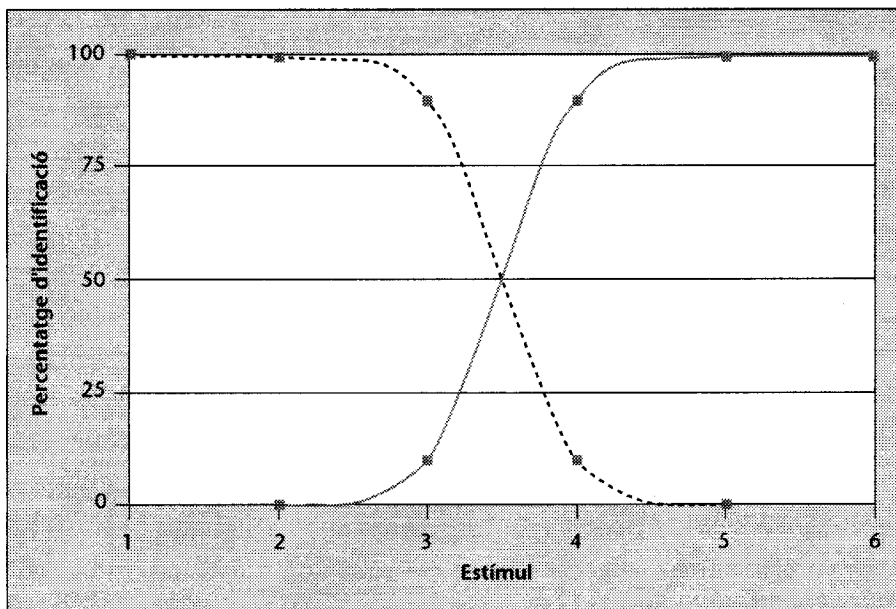
1.2.2. La percepció categòrica de la parla

Un dels aspectes més estudiats en el camp que ens ocupa és la percepció categòrica. Com ja us hem fet notar, conèixer una llengua implica saber agrupar en una única categoria –fonemes– sons que acústicament poden ser diferents. En els experiments sobre la percepció dels sons de la parla, això es manifesta en una discontinuïtat en les respostes a estímuls que corresponen a fonemes diferents.

10. Si voleu saber més sobre els indicis acústics de les vocals i les consonants, consulteu l'annex que apareix al final del llibre i els capítols 6 i 7 de Hayward (2000). Pel que fa als elements suprasegmentals, podeu consultar, a més de l'annex esmentat, el capítol 5 de Pickett (1999).

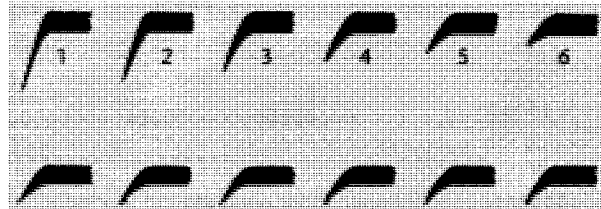
Els primers estudis que van mostrar l'existència de la percepció categòrica en la parla es van fer als Laboratoris Haskins. Un dels experiments més coneguts consistia a demanar a un grup de persones que identifiquessin /b/, /d/ o /g/ en una sèrie de síl·labes produïdes artificialment, en què es canviava gradualment la direcció de les transicions del segon formant (vegeu la figura 2.12), un dels indicis acústics que, com hem vist, aporta informació sobre el lloc d'articulació de les consonants. Mentre que els estímuls variaven gradualment, les respostes van mostrar, en canvi, que hi ha un punt en què en la percepció es produeix un salt d'una categoria –/b/, per exemple– a una altra –/d/. Això es manifesta en respostes com les que s'esquematitzen en la figura 2.11:

Figura 2.11. Resultats (hipotètics) d'una prova d'identificació que mostren la classificació dels estímuls en dues categories



El que cal notar aquí és que mentre que els tres primers estímuls s'identifiquen com a, posem per cas, /b/, els tres darrers s'identifiquen com a pertanyents a una categoria diferent, /d/, per exemple. En canvi, els estímuls canvien de manera gradual, com podeu observar en la figura 2.12.

Figura 2.12. Sèrie d'estímuls emprats en una prova de percepció categòrica, de manera que formen un continu acústic de la síl·laba [ba] a la síl·laba [da]



Si el nostre sistema perceptiu no agrupés les consonants en categories, en les respostes no es trobaria el canvi que en la figura 2.11 es veu entre l'estímul 3 i el 4.¹¹

El fenomen de la percepció categòrica implica també que en la nostra llengua distingim millor les diferències entre els sons que pertanyen a categories diferents –per exemple, /b/ i /d/– que no pas les que es donen entre els sons d'una mateixa categoria –per exemple, [b] i [β], que són al·lòfons del fonema /b/. Expressat en altres termes, percebem més fàcilment les diferències fonològiques que les fonètiques.¹²

Per tot això, la percepció categòrica posa de manifest que en la percepció de la parla agrupem dins d'una mateixa categoria sons que poden ser acústicament diferents.

1.3. L'estudi de la percepció categòrica de la parla

En aquest apartat ens centrarem en els treballs que han tractat la percepció categòrica de la parla, i l'estudiarem des de diferents perspectives. En primer lloc, veurem els estudis que s'han fet sobre persones adultes; en segon, els que

11. Consulteu les pàgines dels Laboratoris Haskins (<http://www.haskins.yale.edu/featured/bdg.html>) i escolteu els estímuls que es van utilitzar per a estudiar els indicis acústics que permeten de diferenciar el punt d'articulació de les consonants [b, d, g] en anglès, similars als que es presenten en la figura 2.12. En les pàgines del Departament de Lingüística de la Universitat d'Umeå (<http://www.ling.gu.se/~anders/KatPer/Applet/index.eng.html>) trobareu una prova de percepció categòrica amb les síl·labes [ba, da, ga].

12. En l'apartat *Categorical perception* del capítol 6 de Borden et al. (1994) i en l'apartat 5.2 de Hayward (2000), podreu llegir una explicació més detallada dels experiments sobre percepció categòrica de la parla.

s'han portat a terme amb infants, i en tercer, els que han tingut com a subjectes altres espècies animals. Tot això ens portarà a reflexionar sobre el paper de l'audició i el paper del coneixement fonètic i fonològic en la percepció de la parla.

1.3.1. La percepció categòrica de la parla en adults

Diferents experiments han posat de manifest que els adults percebem categòricament els contrastos en el lloc d'articulació –per exemple, entre /b/ i /d/–, en el mode d'articulació –entre /p/ i /m/– i en la sonoritat –entre /p/ i /b/. Aquest fenomen també es dona entre vocals, especialment si es presenten en síl·labes senceres formades per una consonant, una vocal i una consonant. Els estudis que es fan per a observar la manera de percebre els sons d'una mateixa llengua s'anomenen *intralingüístics*.

Una pregunta que han provat de respondre diferents investigadors és si la divisió en categories es manifesta de la mateixa manera en totes les llengües. S'ha vist clarament que això no és així i que la manera com categoritzem els sons depèn molt de la nostra primera llengua. Un mateix continu acústic pot ser segmentat en categories diferents per parlants de llengües diferents, segons l'inventari de fonemes de cada llengua. Quan s'investiga aquest fenomen, se sol parlar d'estudis *interlingüístics*.

Penseu, per exemple, en el vocalisme del català i del castellà. Mentre que el català en posició accentuada té set fonemes vocàlics –/i/ sí, /e/ cec, /ɛ/ sec, /a/ sac, /ɔ/ soc, /o/ sóc (del verb ser), /u/ suc–, el castellà en presenta cinc –/i/ piso, /e/ peso, /a/ paso, /o/ poso, /u/ puso. Això implica que l'espai vocàlic està dividit en set categories en català i en cinc en castellà, de manera que un castellanoparlant tendeix a confondre les vocals catalanes /e/-/ɛ/ i /o/-/ɔ/, ja que en el seu sistema hi ha únicament una categoria /e/ i una categoria /o/.

És un fet ben conegut que quan aprenem una llengua diferent de la nostra, especialment si ho fem quan ja som adults, gairebé sempre la parlem amb l'"accent" propi de la primera llengua que hem adquirit. Una de les hipòtesis que s'han proposat per a explicar aquesta experiència tan comuna és que interpretem els sons de les altres llengües segons el conjunt de categories fonològiques de la nostra primera llengua. En aquest sentit, s'ha dit que hi ha una mena de

“filtre” perceptiu –el sistema fonològic– que ens fa interpretar els sons d’una llengua estrangera emprant les categories de la primera llengua. La producció també es veu afectada, ja que, pel que sembla, és relativament difícil produir diferències que no percebem.¹³

En català, /s/ (sorda) i /z/ (sonora) són fonemes perquè distingeixen paraules com ara *caça* /kasə/ i *casa* /kazə/. En castellà, la *s* de la paraula *desde* [dezðe] és, des del punt de vista fonètic, una consonant sonora –ho podeu comprovar si la pronuncieu i us poseu les puntes dels dits a la zona de la laringe com se suggereix en un exercici de l’apartat “Teoria de la interacció social (sociointeraccionisme)”–, però en canvi no hi ha mots que es distingeixin pel contrast entre /s/ i /z/ com en català.

Molts castellanoparlants tenen problemes per a distingir, tant en la percepció com en la producció, entre mots com *caça* i *casa* quan aprenen català, tot i que en castellà articulen el so [z] cada vegada que la consonant /s/ va seguida d’una altra consonant sonora (*mismo*, *rasgo*, *esbelto*, etc.). Això mostra que, ben probablement, els problemes no vénen de la capacitat d’articular un determinat so, sinó del conjunt de categories fonològiques de la llengua materna: mentre que per a un castellanoparlant la diferència entre [s] i [z] és purament fonètica i condicionada pel context, per a un catalanoparlant és una diferència fonològica, ja que serveix per a distingir paraules amb significats diferents.

Podem concloure, per tant, que la percepció de les categories fonològiques en l’edat adulta depèn de l’inventari de fonemes de la llengua i està condicionada per l’estructura fonològica de la primera llengua que hem adquirit.

1.3.2. La percepció categòrica de la parla en infants

Ja que en l’edat adulta la percepció de les categories fonètiques està clarament influenciada per la llengua que primer hem après, a partir dels anys setanta es va començar a estudiar com perceben els infants. Fent servir tècniques basades en els canvis en la freqüència amb què els infants xuclaven un xumet en escoltar estímuls que percebien com a nous i altres procediments similars, Peter Eimas a la Universitat de Brown, Peter Jusczyk a la Universitat John Hopkins

13. En l’apartat *The Effect of Experience on Speech Sound Discrimination* del capítol 13 de Pickett (1999) trobareu més informació sobre la percepció de contrastos fonètics i fonològics segons l’experiència lingüística.

i els seus col·laboradors van demostrar que infants d'un mes eren capaços de discriminar diferències fonètiques d'una manera semblant a la dels adults, de manera que diferenciaven estímuls anàlegs a /pa/ i /ba/ i mostraven una millor discriminació entre aquestes dues síl·labes que entre diferents realitzacions fonètiques de la mateixa síl·laba.¹⁴

Aquesta capacitat de discriminació en els infants s'ha confirmat en molts experiments posteriors i ha plantejat una qüestió important: es tracta simplement d'una distinció auditiva, o d'una capacitat innata de distingir contrastos que són significatius des del punt de vista lingüístic?

Pel que sembla, els infants neixen amb una gran capacitat de discriminació, que progressivament es va limitant segons les categories de la llengua que escolten en el seu entorn. D'aquesta manera, el sistema perceptiu es configura al llarg dels primers vuit o deu mesos de vida, de manera que es perd la capacitat de distinció entre categories fonològiques que no es donen en la primera llengua que s'adquireix i s'augmenta la sensibilitat envers aquestes. Això explicaria la dificultat que esmentàvem en el cas dels adults per a distingir entre contrastos – com ara /s/ i /z/– que no són fonològics en la seva llengua.

Per tal d'investigar com s'estableixen les categories fonològiques segons la primera llengua que s'aprèn, Patricia Kuhl, de la Universitat de Washington, i el seu equip van estudiar la discriminació de les vocals /i/ de l'anglès i /y/ del suec (una vocal com la de *sud* en francès) en infants anglesos i suecs de sis mesos. Mitjançant un sintetitzador, van crear dues versions de cadascuna: una que un grup de subjectes adults va considerar que era el millor exemple de la categoria /i/ o /y/ –és a dir, una vocal que podríem anomenar *prototípica*, en el sentit que es dona habitualment a aquest terme en els estudis sobre cognició– i una altra que el mateix grup va considerar com un exemple no massa bo de cada categoria. Tant els infants anglesos com els suecs van presentar dificultats per a discriminar la vocal prototípica d'altres similars de la mateixa categoria en la seva llengua. En canvi, quan als infants anglesos se'ls va presentar el prototip suec i als infants suecs el prototip anglès, la discriminació va millorar.

14. Els primers experiments clàssics de percepció de la parla en infants es descriuen a Eimas (1985). Podeu ampliar la informació amb el capítol 13 de Pickett (1999). En les pàgines de l'Infant Lab del Laboratoire de Sciences Cognitives et Psycholinguistique de l'École des Hautes Études en Sciences Sociales a París, trobareu una descripció dels mètodes emprats actualment per a l'estudi de la percepció de la parla en infants: <http://www.ehess.fr/centres/lscp/babylab/indexuk.html>.

Aquests resultats van dur a postular l'existència d'un *efecte imant* en la percepció: la vocal més prototípica atreu perceptivament altres vocals similars, que es confonen amb la prototípica, mentre que les realitzacions més allunyades del prototip es confonen menys entre si. En altres termes, l'experiment de Kuhl mostra que als sis mesos s'ha perdut en certa manera la capacitat de discriminar entre sons fonèticament diferents que s'apropen molt al representant prototípic d'una categoria fonològica.¹⁵

Dels estudis experimentals es desprèn, per tant, que els infants menors de deu o dotze mesos poden percebre categòricament contrastos fonològics que no són propis de la llengua del seu entorn. Més endavant, aquesta capacitat es restringeix segons la llengua que s'adquireix.

1.3.3. Processament auditiu i processament fonètic

Una altra manera de contestar a la pregunta de si la capacitat de discriminar millor entre categories fonològiques que entre sons fonèticament semblants és innata o respon a les característiques del sistema auditiu, és mitjançant els estudis sobre percepció d'estímuls lingüístics en animals.

S'ha treballat, per exemple, amb xinxilles, ja que tenen un sistema auditiu molt similar a l'humà, i s'ha mostrat que aquests rosegadors discriminen de manera categòrica entre sons que són fonològicament diferents en les llengües, segmentant d'un mode equivalent al dels humans continus sons com ara /pa/-/ta/-/ka/. És evident que en aquest cas, i en molts altres que s'han observat en primats o en aus, el coneixement lingüístic no té cap mena de paper i, per tant, la discriminació es fa en un nivell de processament purament auditiu.¹⁶

En resum, els estudis amb infants i amb animals han portat a pensar que les fronteres perceptives entre categories fonològiques poden estar, en part, relaci-

15. En l'apartat Speech Sound Classification by Babies: Prototypes and the 'Perceptual Magnet Effect' del capítol 13 i l'apartat The Development of Speech Sound Categories: The Native Language Magnet Model del capítol 15 de Pickett (1999), trobareu una descripció més completa de l'efecte de l'imant perceptiu en el desenvolupament de les categories fonològiques i de les seves implicacions per a un model de percepció de la parla.

16. Trobareu més informació sobre la percepció de la parla en altres espècies en l'apartat Speech Sound Discrimination by Animals del capítol 13 de Pickett (1999).

onades amb la capacitat de discriminació del sistema auditiu, de manera que aquelles diferències acústiques que la nostra oïda pot discriminar millor serien precisament les que es farien servir en les llengües per a diferenciar entre categories fonològiques. Alguns contrastos entre fonemes que es troben a les llengües serien, doncs, els naturals segons les capacitats del sistema auditiu humà.

S'ha proposat, per tant, que hi ha una manera auditiva i una manera fonètica de percebre, i que aquesta darrera s'aplicaria quan ens trobem en presència del que interpretem com un senyal sonor del qual hem d'extreure informació lingüística.

Un problema que s'ha investigat també en el camp de la percepció de la parla és el de la relació entre els contrastos fonètics i fonològics que som capaços de produir i els que tenim la capacitat de percebre. Els estudis s'han dut a terme comparant la percepció de contrastos en llengües diferents en adults monolingües, en parlants bilingües, en adults i nens que aprenen una segona llengua i també amb nens que presenten dificultats d'articulació.

Els resultats obtinguts estan lluny encara de ser concloents. Se sap que la influència de la primera llengua és molt rellevant i que la intensitat del contacte amb la segona llengua exerceix igualment un paper, però també s'ha vist que des del punt de vista del desenvolupament tant de la primera llengua com de les que s'aprenen com a segones o terceres, és difícil establir la precedència de la producció o de la percepció.

La percepció intermodal: l'efecte McGurk

Tot i que fins ara ens hem centrat en la informació acústica, cal no oblidar que en moltes ocasions la comunicació comporta també un contacte visual entre l'emissor i el receptor. Un experiment dissenyat originalment per McGurk i MacDonald el 1976 va mostrar que si, per exemple, un subjecte escolta la síl·laba /ba/ i simultàniament veu l'enregistrament en vídeo de la cara d'una persona pronunciant /ga/, és molt possible que respongui que ha sentit /da/. Així es demostra que en la percepció de la parla fem també indicis visuals que complementen els acústics, integrant informació procedent de diferents canals o modes, i d'aquí ve el nom de *percepció intermodal*.¹⁷

17. En les pàgines del Perceptual Science Laboratory de la Universitat de Califòrnia a Santa Cruz, trobareu una breu explicació i un vídeo de demostració de l'efecte McGurk: <http://mambo.ucsc.edu/psl/dwmdir/da.html>. Podeu llegir també les seccions 1 i 2 de Massaro i Stork (1998).

1.4. La segmentació i les unitats de la percepció de la parla

1.4.1. La segmentació de la parla

Tot el que hem estudiat fins aquest punt porta a pensar que la primera tasca que cal fer en el procés de la percepció és la *segmentació* del senyal sonor que, en l'apartat 1.1.2, caracteritzàvem com un continu. Dèiem que a diferència de la llengua escrita, si observem l'ona sonora d'un enunciat, no hi trobem, d'entrada, l'equivalent als espais en blanc que separen les paraules.

La segmentació implica, com veurem en l'apartat següent, detectar unitats lingüístiques en l'ona sonora. Cal considerar, però, que a més de la informació de tipus segmental –fonemes i al·lòfons– que hem presentat fins ara, el senyal sonor és també portador d'informació que en fonètica s'anomena *suprasegmental* o *prosòdica*: l'accent, l'entonació i el ritme són elements que contribueixen, en el nivell fonètic, a la segmentació dels enunciats.

L'entonació o melodia és el resultat dels canvis de vibració de les cordes vocals durant la producció d'un enunciat, i es manifesta acústicament en variacions de la freqüència fonamental –la més baixa d'una ona sonora– al llarg del temps, que es representen en l'anomenada *corba melòdica*. Observeu en les figures 2.13, 2.14 i 2.15 la corba melòdica de l'enunciat "La Maria menja melindros" pronunciat com una afirmació, com una pregunta i com una exclamació.

Figura 2.13. Corba melòdica de l'enunciat "La Maria menja melindros" pronunciat com una afirmació

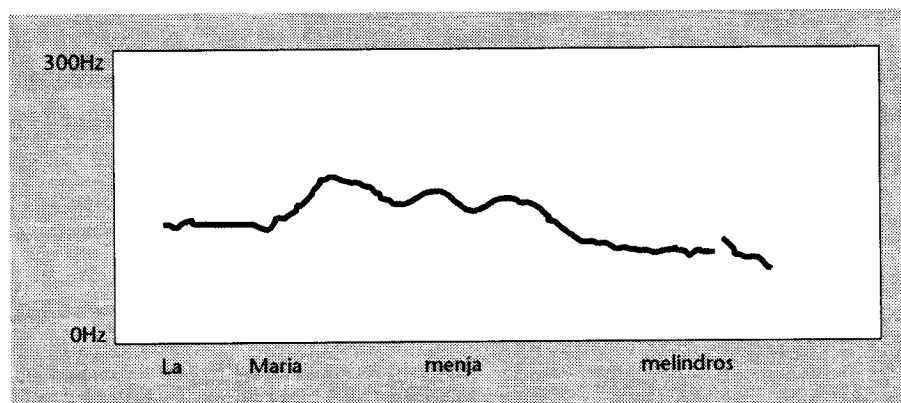


Figura 2.14. Corba melòdica de l'enunciat "La Maria menja melindros" pronunciat com una interrogació

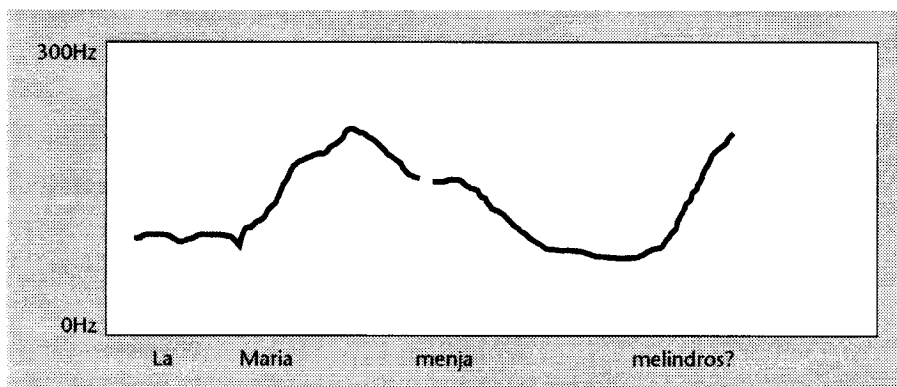
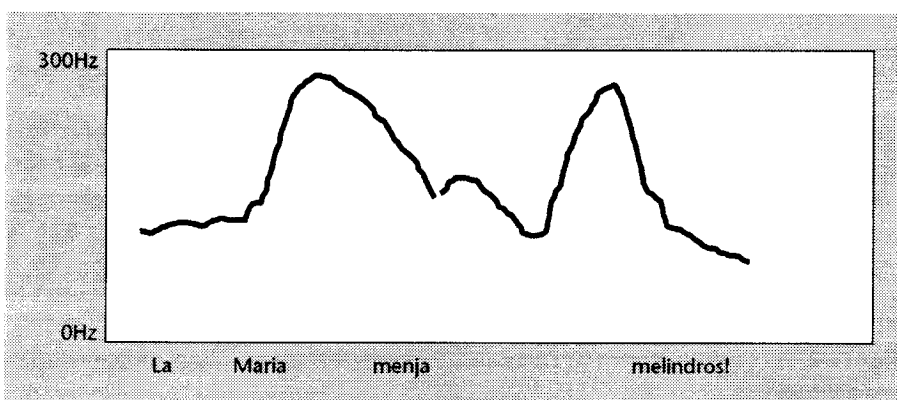


Figura 2.15. Corba melòdica de l'enunciat "La Maria menja melindros" pronunciat com una exclamació



En el moment de la segmentació és útil també la informació sobre les possibles combinacions entre sons, tant a l'interior d'una paraula com entre paraules; és el que s'anomena *informació fonotàctica*. Aquestes combinacions són específiques de cada llengua.

Recordem, doncs, que la segmentació del senyal sonor es fonamenta en indicis acústics segmentals i suprasegmentals (o prosòdics) i també en informació fonotàctica.

1.4.2. Les unitats de percepció de la parla

Els diferents autors que han considerat el problema de la segmentació han proposat una sèrie d'unitats que els receptors identificaríem en el senyal sonor com a part del procés de la percepció.

En primer lloc, s'han postulat els *trets acusticofonètics*, que serien uns elements presents directament en el senyal sonor i que el nostre sistema perceptiu aprendria a extreure i a processar a partir dels indicis acústics que caracteritzen les diferents classes de sons.¹⁸ Tot i així, cal recordar que la detecció d'aquests indicis acústics està estretament lligada a les capacitats del sistema auditiu.¹⁹

En segon lloc, un altre candidat seria el *fonema*, entès com un so capaç de contrastar significats entre mots, com hem explicat anteriorment. S'ha proposat també que els *al·lòfons* que apareixen condicionats pel context podrien constituir les unitats bàsiques de la percepció de la parla, i així es té en compte la influència del context fonètic i de la coarticulació que són inherents, com hem vist, a la producció dels sons.

En tercer lloc, s'ha postulat una unitat més gran, la *síl·laba*, un conjunt de fonemes organitzats entorn d'un nucli –en general, una vocal– que els parlants reconeixen sovint de manera intuïtiva, fins i tot encara que no estiguin alfabetitzats. Aquesta hipòtesi es basa, en part, en el resultat d'experiments que van mostrar que el temps de reacció es més breu quan es demana que un subjecte identifiqui, per exemple, *sis* en un mot com *sistema* (la divisió sil·làbica del qual és *sis-te-ma*), que en una paraula com *sisè* (*si-sè*) en què *sis* no constitueix una síl·laba. Malgrat tot, el paper de la síl·laba en la percepció podria variar segons la llengua i les seves característiques rítmiques.

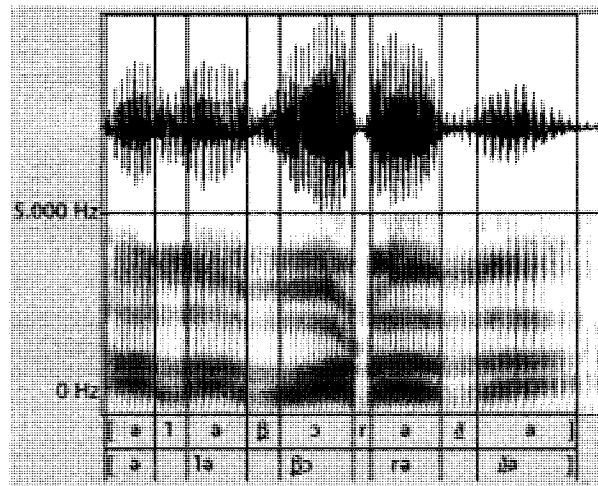
Sembla, però, que no tenim una resposta definitiva a la pregunta de quina és la unitat bàsica en la percepció de la parla. Hi ha, això sí, un consens en què abans d'arribar a la fase d'accés al lèxic cal analitzar i descodificar d'alguna manera la informació acústica i fonètica present en un enunciat.

Observeu la segmentació en al·lòfons i en síl·labes de l'enunciat “a la vora de” que es representa en la figura 2.16.

18. Podeu consultar l'apartat 1.2.1 d'aquest capítol.

19. Podeu consultar l'apartat 1.3.3 d'aquest capítol.

Figura 2.16. Forma d'ona (part superior) i espectrograma (part inferior) de l'enunciat "a la vora de"



En resum, les principals unitats de percepció de la parla que han estat proposades són els trets acusticofonètics, els fonemes, els al·lòfons i les síl·labes.²⁰

1.5. Models de percepció de la parla: les teories motrius

1.5.1. Els models de percepció de la parla

Per tal de donar una solució als problemes que han sorgit en els estudis sobre la percepció de la parla, s'han proposat diferents models que intenten explicar aquest procés. Un model de percepció de la parla ha d'explicar, almenys, dos fenòmens que ja hem estudiat.

- La variació entre parlants (interlocutor) i la variació en un mateix parlant (intra locutor).

20. Trobareu un resum sobre les unitats de percepció de la parla i les referències bibliogràfiques rellevants en l'apartat 2.1 de McQueen i Cutler (1997).

- La capacitat de segmentar el continu sonor en unitats lingüístiques, tenint en compte el següent:
 - La coarticulació fa que en un mateix segment del senyal es trobi informació sobre més d'una unitat lingüística.
 - Senyals acústicament diferents poden portar informació sobre la mateixa unitat lingüística.

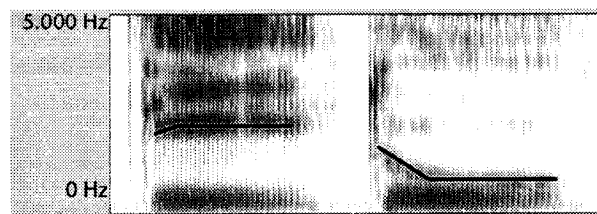
Les teories sobre la percepció de la parla se solen agrupar en dues categories: les que fan referència al procés de producció –conegudes com a teories actives o *teories motrius*, ja que el receptor ha de recórrer a informació sobre la manera com produeix els sons, és a dir, sobre els moviments dels articuladors– i les que es basen en l'audició i en l'extracció de trets fonètics –conegudes com a teories passives o *teories auditives* pel pes que es dona al procés d'audició i de percepció fonètica.

1.5.2. Les teories motrius de la percepció de la parla

La teoria motriu de la percepció de la parla, com la coneixen actualment després de successives revisions, arrenca dels treballs d'Alvin Liberman i els seus col·laboradors als Laboratoris Haskins. Un dels principals problemes que es pretén explicar és la manca d'invariància en el senyal acústic que ja hem assenyalat en diferents ocasions.

Observeu l'espectrograma de la figura 2.17, corresponent a les síl·labes [ti] (a l'esquerra) i [tu] (a la dreta), en què s'ha marcat la trajectòria de les transicions del segon formant, que mostren el moviment dels articuladors en passar de [t] a [i] i de [t] a [u].

Figura 2.17. Trajectòries del segon formant en els espectrograms de [ti] (esquerra) i [tu] (dreta)



Com veieu, en el cas de [tu] la transició del segon formant és descendent (de les freqüències altes a les baixes), mentre que en [ti] és ascendent (de les freqüències baixes a les altes). Tot i tenir direccions diferents, totes dues constitueixen un índex acústic per a la percepció de [t]. Això mostra que parts del senyal acústicament diferents –les transicions– poden portar informació per a la identificació d'una mateixa unitat lingüística –el fonema /t/.

El fet que els experiments sobre percepció categòrica fessin veure que la resposta a un continu d'estímuls com [ba-da-ga] presentava salts abruptes entre una categoria i una altra, com ja heu tingut oportunitat d'estudiar, va fer pensar als proponents de la teoria motriu que una possible explicació es trobava en l'articulació: de la mateixa manera que no percebem sons intermedis entre [b] i [d], tampoc no som capaços de produir-los.²¹

La recerca dels invariants es va desplaçar, doncs, del senyal acústic als moviments articuladoris. Tot i així, tampoc no és possible trobar una única configuració articulatòria o un únic conjunt de moviments dels articuladors associats en qualsevol circumstància a la producció d'un so. Recordeu, per exemple, els efectes de la coarticulació que mostràvem en la figura 2.5.

La solució a aquest problema va ser postular l'existència de representacions abstractes dels moviments articuladoris necessaris per a produir un so, i pensar que el que fem realment en la percepció és intentar, d'alguna manera, "recuperar" la intenció articulatòria de l'emissor. Cal assenyalar que aquestes representacions corresponen a un conjunt de moviments que es fan simultàniament per a produir un so –*gest articuladori*–, però que en la producció d'un enunciat es poden ordenar de maneres diferents en el temps a causa de la coarticulació. En aquest sentit, doncs, caldria afegir el gest articuladori abstracte a les unitats que s'han proposat per a explicar la percepció de la parla exposades anteriorment.

Com podeu deduir, un dels avantatges d'aquest model és que estableix un lligam directe entre la producció i a la percepció, ja que la percepció consistiria en el reconeixement del gest articuladori que dona lloc a una unitat lingüística codificada en el senyal sonor, emprant el coneixement interioritzat que el receptor té de la manera com ell mateix produeix els sons. Aquest coneixement i la capacitat d'activar-lo en la percepció formarien part d'un mòdul especialitzat en el processament del llenguatge, de manera que la teoria

21. Per a una revisió d'aquest enfocament, podeu repassar l'apartat 1.2 d'aquest capítol.

motriu²² encaixaria amb altres plantejament sobre la modularitat del llenguatge.²³

Encara hi ha models que podríem qualificar de més radicals, com ara la *teoria del realisme directe* de Carol Fowler i els seus col·laboradors als Laboratoris Haskins. Segons aquests investigadors, el fet que, per exemple, siguem capaços de repetir exactament un enunciat del qual no comprenem el significat fa pensar que no fem pas una detecció de categories fonètiques o fonològiques ni de representacions articulatòries abstractes, sinó que directament recuperem informació sobre els gestos articuladoris necessaris per a la producció dels sons.²⁴

Podem dir, en resum, que les teories motrius de la percepció de la parla postulen una relació directa entre la percepció i la producció, i l'existència d'un mòdul específic de processament fonètic de la parla.

1.6. Models de percepció de la parla: les teories auditives

Contràriament al que hem vist en el cas de les teories motrius, les teories que s'agrupen sota la denominació d'*auditives* donen un paper preponderant als mecanismes perceptius i atorguen molta menys importància al coneixement dels mecanismes articuladoris.

La majoria dels models auditius es fonamenten en l'existència d'un mecanisme de detecció d'indicis acústics o de trets fonètics basat en la comparació entre el senyal sonor que arriba al receptor i un conjunt de plantilles o patrons acústics prèviament emmagatzemats.

Un model que intenta combinar la percepció fonètica amb l'accés al lèxic i que es basa en la idea d'unes plantilles emmagatzemades en el nostre sistema perceptiu és el LAFS (*Lexical Access From Spectra*), desenvolupat per Dennis Klatt a l'Institut de Tecnologia de Massachusetts el 1979. La idea d'aquest model és

22. En l'apartat Theories of Speech Perception del capítol 6 de Borden et al. (1994), podreu llegir un resum de diferents experiments que donen suport a la teoria motriu de la percepció de la parla. Hi ha una presentació més detallada en l'apartat The Motor Theory of Speech Perception del capítol 14 de Pickett (1999).

23. En l'apartat 3.4 d'aquest capítol trobareu una explicació més extensa sobre la modularitat.

24. En l'apartat Direct-Realist Approach to Speech Perception de Goldinger et al. (1996) i en l'apartat Invariance in the percept but not the object: The theory of Direct Realism del capítol 15 de Pickett (1999), trobareu una descripció d'aquest model.

que en la nostra memòria disposem d'un diccionari d'espectres prototípics corresponents als sons de la nostra llengua –més concretament a les combinacions de dos sons, com ara les síl·labes formades per una consonant i una vocal, ja que així es té en compte la coarticulació– i que en el moment de la percepció els comparem amb l'anàlisi que la nostra oïda fa de l'estímul lingüístic que li arriba en forma de senyal sonor. En aquest model, el lèxic també estaria representat en una xarxa d'espectres de combinacions de sons.²⁵

Un espectre és una representació d'un instant del senyal sonor en què s'observa la distribució d'energia acústica (eix vertical) en diferents zones freqüencials (eix horitzontal). Com podeu observar en la figura 2.18, els formants que en l'espectrograma apareixen en forma de barres (mostrant el temps en l'eix horitzontal, la freqüència en el vertical i l'energia o amplitud en forma d'una escala de grisos), en l'espectre es manifesten com a pics, cadascun dels quals es caracteritza per la seva freqüència i la seva amplitud.

En el cas de les vocals, la distribució dels formants en l'espectre és responsable de la percepció de les diferències de timbre que es donen, per exemple, entre /i/, /a/ i /u/.

Mentre que el LAFS se centra en representacions espectrals de combinacions de sons, altres models consideren la percepció com un procés d'extracció d'indicis acústics. Aquest és el cas del FLMP (*Fuzzy Logical Model of Speech Perception*) proposat per Dominic Massaro i els seus col·laboradors a la Universitat de Califòrnia a Santa Cruz des del final dels anys setanta.

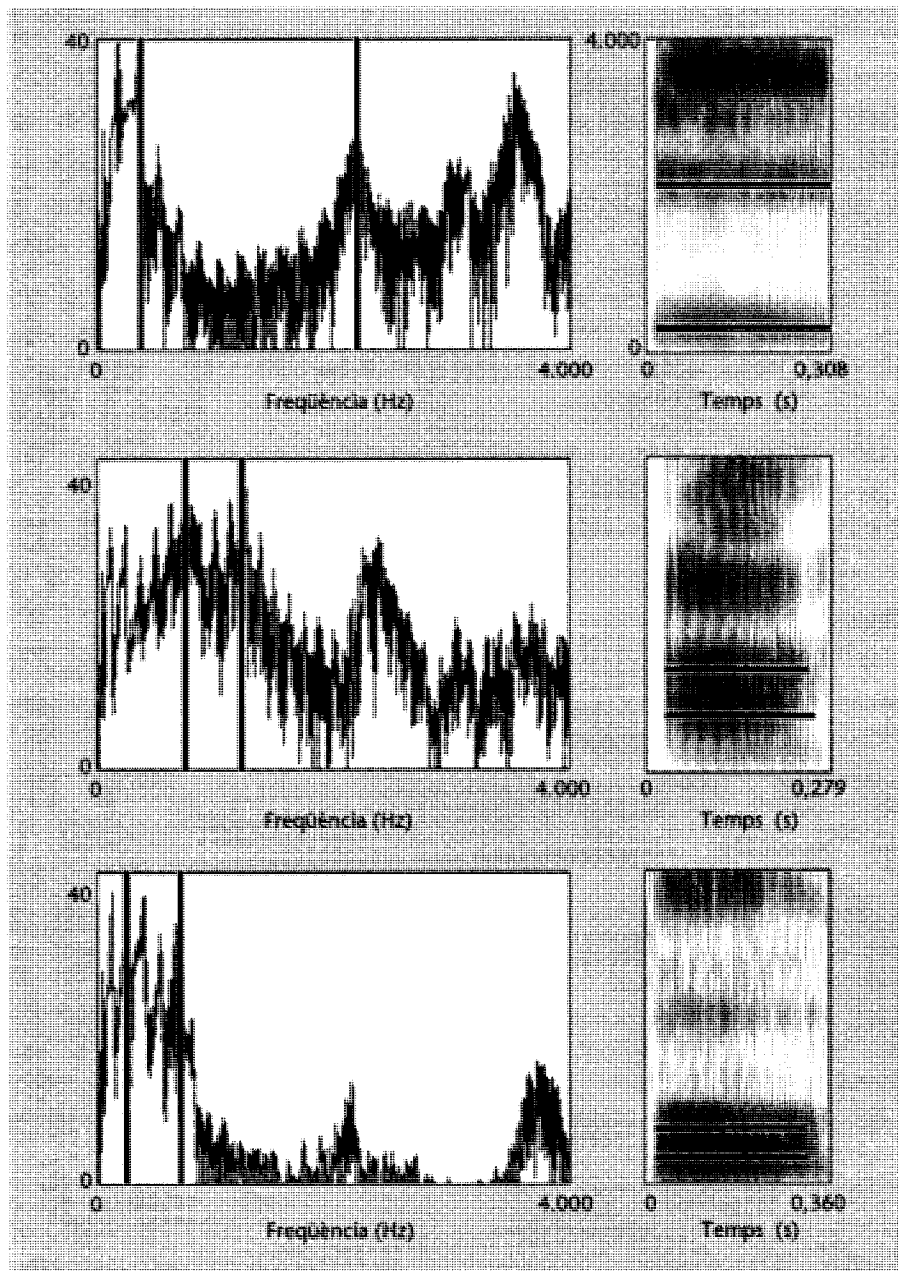
Segons l'FLMP, la percepció comença per determinar de quins trets acusticofonètics es troben presents en el senyal sonor. La denominació de *difús* (*fuzzy*) ve donada pel fet que l'oïdor no busca si un tret és present o no ho és, sinó que avalua la seva probabilitat d'aparició, fent servir totes les fonts d'informació possibles, incloent-hi altres canals com ara el visual (com es posa de manifest en l'efecte McGurk).²⁶

En una segona etapa, la hipòtesi que s'ha fet sobre els trets es compara amb un conjunt de prototips de fonemes, consistents en una descripció dels trets acústics i fonètics que idealment caracteritzarien cada fonema. Finalment, es determina el grau d'adequació entre la representació emmagatzemada dels prototips de fonemes i els trets detectats en el senyal, i s'arriba així a una decisió sobre les unitats lingüístiques presents en l'ona sonora.

25. Podeu consultar l'apartat 2.2 d'aquest capítol.

26. Recordeu l'efecte McGurk que podeu consultar en l'apartat 1.3.3 d'aquest capítol.

Figura 2.18. Espectres (a l'esquerra) i espectrogrames (a la dreta) de les vocals [i] (fila superior), [a] (fila central) i [u] (fila inferior)



Les línies horitzontals marquen la posició del primer (F1) i del segon formant (F2) en l'espectrograma i les verticals, en l'espectre.

Noteu que en aquest model no es fa referència a un mòdul especialitzat en la percepció de la parla, com ocorria amb la teoria motriu,²⁷ sinó que es recorre a representacions acústiques i fonètiques emmagatzemades en la memòria.

Els dos models que hem descrit fins ara intenten resoldre el problema de la invariància, o bé postulant que tenim emmagatzemades combinacions de sons en comptes de segments aïllats (LAFS) o bé proposant prototipus que podrien ser diferents segons el context en què apareix cada fonema (FLMP).²⁸

En canvi, altres autors com Kenneth Stevens, de l'Institut de Tecnologia de Massachusetts, defensen més aviat que hi ha indicis acústics invariants tant en el senyal sonor com en la seva representació auditiva. A partir d'aquests indicis, es portaria a terme l'extracció de trets fonològics abstractes que permetrien d'identificar les unitats lingüístiques codificades en el senyal sonor. Per aquest motiu, el model de Stevens i els seus col·laboradors, desenvolupat al final dels anys vuitanta, es coneix com a LAFF (*Lexical Access From Features*).

Segons el LAFF, els mots estarien representats en forma de patrons de trets distintius organitzats en categories equivalents al mode i el lloc d'articulació. En la percepció, es postula que l'oïdor busca els correlats auditius o acústics dels trets distintius i els compara amb la representació en trets de les paraules.²⁹

Els fonemes –sons que poden diferenciar significats de paraules– es poden descompondre en trets distintius que serveixen per a diferenciar-los (d'aquí ve el nom de *distintius*) i classificar-los. Així, /b/ tindria els trets sonor (ja que vibren les cordes vocals), labial (en produir-lo es dona un tancament dels llavis) i oral (pel fet que l'aire surt per la cavitat bucal). En canvi, /p/ tindria els trets sord (per la manca de vibració de les cordes vocals), labial i oral. Al seu torn, /m/ es caracteritzaria pels trets sonor, labial i nasal (perquè en la seva producció l'aire surt per la cavitat nasal). D'aquesta manera,

27. Sobre la teoria motriu, podeu consultar el que dèiem en l'apartat 1.5.2 d'aquest capítol.

28. Els models LAFS de Klatt i FLMP de Massaro s'expliquen més detalladament a Goldinger et al. (1996) i en els apartats Invariance, and Lexical Access from Features (LAFF) del capítol 14 i FLMP: The Fuzzy-Logic Model of Perception del capítol 15 de Pickett (1999).

29. Trobareu una descripció més detallada dels plantejaments de Stevens i del LAFF en l'apartat Acoustic invariants: the Quantal Theory of Speech, Acoustic Invariance and Lexical Access from Features (LAFF) del capítol 14 de Pickett (1999).

podem classificar els fonemes en matrius de trets distintius emprant els signes + o - segons la presència o absència de cada tret:

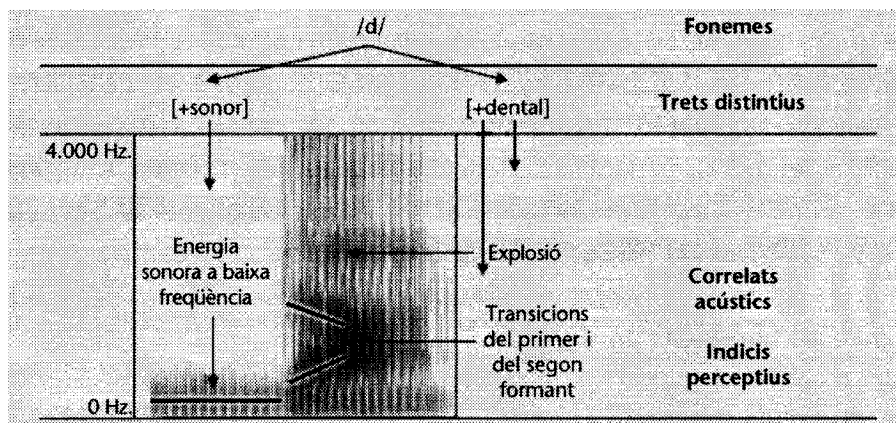
Taula 2.1.

	Labial	Sonor	Oral
/p/	+	-	+
/b/	+	+	+
/m/	+	+	-

En el camp de la fonètica i la fonologia, s'han definit diferents conjunts de trets distintius i s'han dedicat molts esforços a buscar els correlats de cada tret. Per correlats entenem aquelles característiques físiques, tant en el senyal sonor com en l'articulació, que ens permeten de postular la presència d'un tret. Per exemple, el correlat articulatori del tret sonor seria la vibració de les cordes vocals, mentre que el seu correlat acústic seria l'existència d'una zona de freqüències baixes en el senyal sonor -visible en forma de barra de sonoritat a l'espectrograma, com podeu comprovar en la figura 2.19.

La figura 2.19 mostra de manera esquemàtica i simplificada les relacions entre els correlats acústics que actuen com a indicis perceptius, els trets distintius als quals s'associen i els fonemes, entesos com a conjunts o feixos de trets.

Figura 2.19. Relació entre els correlats acústics que actuen com a indicis perceptius, els trets distintius i els fonemes en la síl·laba [da]



Resumint el que hem vist fins ara, els models auditius de percepció de la parla postulen l'existència de mecanismes de detecció d'indícis o de trets acusticofonètics i de procediments de comparació amb representacions emmagatzemades en la memòria.

1.7. Models de percepció de la parla: el paper del senyal i el paper del context

Acabarem aquesta revisió de les principals propostes teòriques per a explicar la percepció de la parla amb el plantejament de Björn Lindblom, conegut com a *teoria de la hiperarticulació i la hipoarticulació* o H&H, desenvolupada des del començament de la dècada dels noranta a les universitats d'Estocolm i d'Austin.

Lindblom parteix del fet essencial en què insistíem al principi: l'objectiu de la comunicació és transmetre un significat de l'emissor al receptor i això es fa, en el cas de la parla, mitjançant una ona sonora. L'intercanvi comunicatiu, però, té lloc en un context que condiciona el mode de producció de la parla. Així, en un entorn molt sorollós que dificulti la intel·ligibilitat, en adreçar-se a una persona amb dificultats d'audició o, en general, quan es vol estar segur que la informació es transmet adequadament, l'emissor tendirà a parlar d'una manera molt clara o "hiperarticulada". En canvi, quan el context permet de predir parcialment el missatge o en situacions en què és més important el fet d'establir i mantenir el contacte que no pas la informació que es transmet, l'emissor optarà per un estil "hipoarticulat" propi de la parla familiar o col·loquial.³⁰

El que és important d'aquest model és que la percepció de la parla no es concep com únicament lligada al senyal sonor, sinó com el producte de la interacció entre el senyal i la situació comunicativa en què es produeix. També té en compte el fet que els parlants adaptem el nostre estil de parla a la situació per tal d'assegurar-nos que la informació es transmeti correctament i amb el mínim esforç possible tant per part de l'emissor com del receptor, de manera que la invariància es trobaria en l'objectiu final del procés de la comunicació, que és justament el de transmetre un significat.³¹

Des d'aquesta perspectiva, la percepció de la parla es pot considerar com el resultat de la interacció entre el senyal sonor i la situació comunicativa, que es complementen per tal que la transmissió d'informació sigui com més eficient millor.

30. Podeu consultar l'apartat 1.1.1 d'aquest capítol.

31. En l'apartat *Toward a more comprehensive theory of speech perception* del capítol 15 de Pickett (1999), trobareu una explicació més detallada del model H&H de Lindblom.