

Francesc Alamon i Queralt

**APROXIMACIÓ FONÈTICA ALS PROCESSOS
D'INTEGRACIÓ PERCEPTIVA DE SENYALS DE PARLA
DE BAIXA INTENSITAT**

Tesi doctoral

Director: Dr. Joaquim Llisterrí i Boix

*Departament de Filologia Espanyola
Facultat de Lletres
Universitat Autònoma de Barcelona
Setembre de 1998*

Universitat Autònoma de Barcelona
Servei de Biblioteques



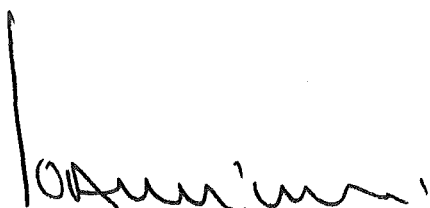
1500459809

Francesc Alamon i Queralt

APROXIMACIÓ FONÈTICA ALS PROCESSOS
D'INTEGRACIÓ PERCEPTIVA DE SENYALS DE PARLA
DE BAIXA INTENSITAT

Tesi doctoral

Director: Dr. Joaquim Llisterra i Boix



A handwritten signature in black ink, appearing to read 'Alamon i Queralt', is written above a horizontal line.

*Departament de Filologia Espanyola
Facultat de Lletres
Universitat Autònoma de Barcelona
Setembre de 1998*

APROXIMACIÓ FONÈTICA ALS PROCESSOS
D'INTEGRACIÓ PERCEPTIVA DE SENYALS
DE PARLA DE BAIXA INTENSITAT

Francesc Alamon i Queralt

Director: Dr. Joaquim Llisterra i Boix

Departament de Filologia Espanyola

Facultat de Lletres

Universitat Autònoma de Barcelona

Setembre de 1998

*"Nature has designed the speech communication process so that it
can operate under a great variety of adverse conditions."*

Moore (1982) p. 236

Agraïments:

Al doctor Joaquim Llisterri, per haver volgut ser el director d'aquesta tesi, per haver confiat en mi i per haver-me esperonat i il·lusionat en els moments més difícils. Perquè ha estat una font de recursos quasi inexhaurible i perquè els seus suggeriments i la seva ajuda en l'elaboració i la correcció d'aquesta tesi han estat valuosíssims.

A David Puigvi, per haver prestat la seva veu en l'enregistrament del material fonètic que vam utilitzar en la preparació de la prova.

A Josep M. Vigatà i els altres informants que es van prestar a dur a terme la prova pilot i que ens van indicar les dificultats amb què s'anaven trobant durant la seva realització.

Als informants de la Facultat de Lletres de Bellaterra i als altres informants que van participar en la prova experimental, per la seva col·laboració pacient i desinteressada durant les llargues estones d'escolta en la cambra insonoritzada.

A Josep M. Fernández, per ajudar-me en la selecció dels informants.

A Francesc Mayós i a Jaume March, pel seu suport i ajuda en el plantejament de l'anàlisi estadística de les dades i en les múltiples consultes per a la seva interpretació.

A Rosa Sol i a Jordi Sorribes, pel seu ajut amb els problemes d'interpretació idiomàtica.

A Miquel Mateus, per la seva col·laboració amb les eines informàtiques i d'edició.

A la doctora Elena Franco, que em va introduir, molt al començament d'aquesta tesi, en els camins de la Neurologia i de la Fisiologia de les vies auditives en el Sistema Nerviós Central

A l'Escola d'Enginyeria La Salle, de la Universitat Ramon Llull de Barcelona, per haver-me deixat efectuar diverses proves amb l'instrumental del seu Laboratori d'Acústica.

A l'Escola de Ciències de l'Educació de Lleida, per haver pogut disposar de la seva sala d'Informàtica per al processament del senyal digitalitzat.

A les persones que treballen en el Laboratori de Fonètica Experimental de la Universitat Autònoma de Barcelona i, en especial, a Lourdes Aguilar, Marcel·la Andreu, Beatriz Bleuca, Juan María Garrido, Maria Jesús Machuca, Carme de la Motta, Dolors Poch. A tots ells el meu agraïment per les facilitats que m'han donat durant les meves estades en el Laboratori, per l'atenció prestada davant dels meus dubtes o desconeixements i, sobretot, per la seva amistat

A totes les persones que, d'alguna o altra manera, han mostrat la seva confiança en mi i m'han animat a continuar endavant. Als bons amics i a tota la família que s'han interessat incansablement pel meu treball durant tot aquest temps i m'han mostrat la seva solidaritat i afecte.

A la Leo i al Francesc, a qui dedico aquestes pàgines, perquè sempre he rebut d'ells el màxim suport i perquè han acceptat amb enorme paciència les incomoditats que ha comportat l'elaboració de la meva (nostra) tesi.

ÍNDEX

	pàg.
0.-Introducció.....	1
1.- La perspectiva de l'oient i el procés auditiu.....	5
1.1.-Introducció.....	5
1.2.-Els elements "passius" del procés comunicatiu verbal.....	6
1.3.-Les característiques de la locució.....	6
1.4.-Les característiques del senyal.....	9
1.4.1.-Els trets de la parla que es manifesten en el senyal.....	10
1.4.1.1.-Els trets intrínsecs del senyal.....	11
1.4.1.2.-Els trets prosòdics del senyal.....	13
1.4.2.-Interferències en el senyal: distorsions i emmascaraments.....	15
1.4.2.1.-El mitjà transmissor del senyal.....	17
1.4.2.2.-L'ambient acústic que envolta l'emissor i el receptor.....	18
1.5.-Les característiques de l'audició.....	18
1.5.1.-L'audició.....	20
1.5.2.-La mesura del llindar d'audibilitat.....	23
1.6.- Conclusions.....	24
2.- La percepció i la comprensió de la parla.....	28
2.1.-Introducció.....	28
2.2.-La percepció com a procés psicològic.....	28
2.3.-La percepció de la parla.....	31
2.4.- La invariança en el senyal de la parla.....	32
2.5.-Característiques generals de la percepció.....	33
2.6.-Selectivitat perceptiva i atenció selectiva.....	34
2.7.-Processos en la percepció de la parla.....	35
2.7.1.- Nivell auditiu.....	36
2.7.1.1.-Anàlisi auditiva perifèrica.....	36
2.7.1.2.-Anàlisi auditiva central.....	36
2.7.2.-Nivell fonètic.....	38
2.7.3.-Nivell fonològic.....	41

2.8.-La memòria en la percepció de la parla.....	43
2.8.1.-Memòria auditiva.....	43
2.8.2.-Memòria fonètica.....	44
2.9.-Models de percepció de la parla.....	44
2.9.1.-Els models actius.....	45
2.9.1.1.-L'anàlisi per síntesi.....	45
2.9.1.2.-La Teoria Motora de la percepció de la parla.....	46
2.9.1.3.-L' Aproximació Directa i Realista a la Percepció de la Parla.....	47
2.9.2.-Els models passius.....	48
2.9.2.1.-El model de Klatt.....	49
2.10.- De la percepció a la comprensió.....	50
2.11.-La percepció de la parla contínua.....	51
2.12.-La comprensió de la parla.....	53
2.13.-Models de reconeixement de paraules i accés al lèxic.....	55
2.13.1.-Models interactius del processament lèxic.....	58
2.13.2.-Model autònom de comprensió de paraules.....	60
2.13.3.-Models mixtos de comprensió de paraules.....	61
2.14.-Conclusions.....	62
3.- La teoria de la Hiper- i la Hipo-Articulació (<i>HYPER- & HYPO- Speech</i>).....	65
3.1.-La teoria de la H & H i l'estratègia de la variabilitat adaptativa.....	66
3.2.-Evolució i aplicacions de la teoria de H & H	71
3.3.-Discussió del model i hipòtesis.....	73
3.3.1.-Consideracions i hipòtesi 1: la relació entre la intel.ligibilitat del senyal i la predicció de la informació.....	74
3.3.2.-Consideracions i hipòtesi 2: la relació entre la informació present en el senyal i la informació independent del senyal.....	80
3.3.3.-Consideracions i hipòtesi 3: el paper que juguen els diferents tipus de context.....	83
3.4.- Conclusions.....	85
4.- Disseny Experimental.....	89
4.1.-Aspectes generals del disseny experimental.....	89
4.1.1.-Objectius i metodologia experimental.....	89

4.1.2.-Esquema experimental.....	92
4.2.-Configuració general de l'experiment.....	95
4.2.1.-Confecció del corpus de la prova.....	96
4.2.1.1.-El corpus de paraules amb sentit.....	96
4.2.1.2.-El corpus de paraules sense sentit.....	109
4.2.1.3.-Les frases portadores.....	112
4.2.2.-L'enregistrament del corpus.....	120
4.2.2.1.-Locutor i locució.....	120
4.2.3.-Tractament de la parla enregistrada.....	120
4.2.3.1.-Digitalització.....	121
4.2.3.2.-L'aplicació informàtica de tractament acústic.....	123
4.2.3.3.-Anàlisi acústica del corpus de l'experiment.....	126
4.2.3.4.-Atenuació.....	128
4.2.3.5.-El repertori d'estímuls i el mètode Ramp invers.....	130
4.2.4.-Realització de l'experiment.....	132
4.2.4.1.-Els subjectes: selecció i característiques.....	135
4.2.4.2.-La prova audiomètrica.....	136
4.2.5.-Factors i variables que intervenen en la prova.....	139
4.3.-Procediments d'anàlisi dels resultats de l'experiment.....	143
4.3.1.-Procediment d'extracció i tabulació dels resultats.....	144
4.3.1.1.-La classificació dels nivells de variació.....	145
4.3.1.2.-La tabulació dels resultats.....	150
4.3.2.-Els resultats i el procediment estadístic d'anàlisi.....	154
4.3.2.1.-El procediment d'anàlisi estadística.....	155
5.- Presentació, anàlisi i discussió dels resultats.....	159
5.1.-Presentació dels resultats referits a identificació fonètica de la vocal tònica en paraules amb sentit i sense sentit.....	160
5.1.1.-Influència del sentit dels mots de l'estímul en la identificació correcta de les vocals tòniques.....	171
5.1.2.-Influència del sentit i el nivell d'intensitat de l'estímul en la identificació correcta de les vocals tòniques.....	172

5.1.3.-Influència del sentit i el nombre de síl.labes de l'estímul en la identificació correcta de les vocals tòniques.....	174
5.1.4.-Influència del sentit i l'estructura de la síl.laba tònica de l'estímul en la identificació correcta de les vocals tòniques.....	180
5.1.5.-Influència del nombre de síl.labes, el sentit dels mots i el nivell d'intensitat de l'estímul en la identificació correcta de les vocals tòniques.....	183
5.1.6.-Influència de l'estructura síl.làbica, el sentit dels mots i el nivell d'intensitat de l'estímul en la identificació correcta de les vocals tòniques.....	187
5.1.7.-Influència de l'estructura síl.làbica, el sentit dels mots i el nombre de síl.labes de l'estímul en la identificació correcta de les vocals tòniques.....	192
5.1.8.-Les vocals.....	197
5.1.9.-Incidència del sentit de les paraules en la identificació vocàlica.....	198
5.1.10.-Relació entre la identificació de les vocals i el nivell d'intensitat.....	200
5.1.11.-Incidència en la identificació de les vocals de la interacció entre el nivell d'intensitat i el sentit del mot.....	203
5.1.12.-Relació entre la identificació de les vocals i el nombre de síl.labes.....	205
5.1.13.-Incidència en la identificació de les vocals de la interacció entre el nombre de síl.labes i el sentit del mot.....	206
5.1.14.-Relació entre la identificació de la vocal tònica del mot clau i l'estructura síl.làbica en la qual apareix.....	208
5.1.15.-Incidència en la identificació de les vocals de la interacció entre l'estructura síl.làbica i el sentit del mot.....	209
5.1.16.-Les confusions fonètiques.....	212
5.1.17.-Les confusions vocàliques.....	215
5.1.18.-Les confusions consonàntiques.....	221
5.1.18.1.-La classificació de les confusions consonàntiques segons les característiques articulatòries.....	229
5.1.18.2.-La classificació de les confusions consonàntiques segons els trets distintius dels sons que hi intervenen.....	233
5.1.18.3.-La classificació de les confusions consonàntiques per segments fonològics.....	244
5.2.-Anàlisi de la identificació fonètica i del reconeixement lèxic en paraules amb sentit i sense sentit.....	250
5.2.1.-Introducció.....	250
5.2.2.-Resultats relacionats amb la identificació fonètica correcta dels mots presentats.....	252
5.2.3.-Resultats relacionats amb el reconeixement lèxic.....	258
5.2.4.-Anàlisi dels resultats obtinguts.....	267
5.2.5.-Influència del nivell d'intensitat en la identificació i en el reconeixement..	268
5.2.6.-Influència del sentit de les paraules en la identificació i en el reconeixement.....	273
5.2.7.-Relació entre el nivell d'intensitat i el sentit dels mots.....	276
5.2.8.-Incidència del factor nombre de síl.labes en el procés d'integració lèxica amb paraules amb sentit i sense sentit.....	278
5.2.9.-Les confusions de base fonètica durant el procés d'integració lèxica.....	285

5.3.-Presentació, anàlisi i discussió dels resultats referits a la incidència del context en la identificació fonètica i en el reconeixement lèxic.....	292
5.3.1.-Les respostes en blanc.....	296
5.3.2.-Les respostes correctes.....	297
5.3.3.-Les respostes errònies o les confusions.....	301
5.3.3.1.-Confusions que porten com a resultat mots sense sentit (CMSS)	302
5.3.3.2.-Confusions produïdes per la presentació de mots sense sentit que porten com a resultat mots amb sentit (CMAS).....	303
5.3.3.3.-Confusions lèxiques per la presentació de paraules amb sentit que porten com a resultat altres mots amb sentit (CMAS)	305
5.4.-Conclusions relacionades amb els objectius del procediment experimental.....	307
5.4.1.-Referència a l'objectiu 1.....	307
5.4.1.1.-Conclusions generals obtingudes a partir de l'objectiu 1.....	307
5.4.1.2.-Conclusions sobre la percepció de les vocals.....	309
5.4.1.3.-Conclusions sobre l'estudi de les confusions vocàliques i consonàntiques.....	310
5.4.2.-Referència a l'objectiu 2.....	311
5.4.2.1.-Conclusions generals obtingudes a partir de l'objectiu 2.....	311
5.4.3.-Referència a l'objectiu 3.....	314
5.4.3.1.-Conclusions generals obtingudes a partir de l'objectiu 3.....	314
6.- Conclusions.....	317
6.1.-Resultats relatius a la hipòtesi 1a: La predicció que fa l'oient sobre el contingut semàntic del senyal guia els processos d'integració de la informació.....	318
6.2.-Resultats relatius a la hipòtesi 2a: Els elements característics dels dos tipus d'informació.....	328
6.2.1.-Nivell auditiu	329
6.2.2.-Nivell fonètic	330
6.2.3.-Nivell lèxic	335
6.3.-Resultats relatius a la hipòtesi 3a: Mecanismes diferents d'anàlisi segons el sentit de les paraules.....	339
7.- Bibliografia.....	347

ANNEXOS:

ANNEX 1: Corpus de les oracions utilitzades en la prova.....	3
ANNEX 2: Taula de càlcul dels factors d'anivellament.....	8
ANNEX 3: Instruccions per a la correcta aplicació del test.....	14
ANNEX 4: Full de respostes utilitzat durant l'experiment.....	16
ANNEX 5.1: Anàlisi dels resultats basades en les taules de l'apartat 5.1.....	23
ANNEX 5.2: Anàlisi dels resultats basades en les taules de l'apartat 5.2.....	99

INTRODUCCIÓ

Mitjançant els processos de percepció de la parla, el receptor és capaç d'extreure la informació que es troba codificada en el senyal que li arriba de l'emissor i que és portador d'un missatge. La naturalesa d'aquesta informació, en el sentit més estricte, és de caire acústic i fonètic, i la tasca inicial del receptor s'encamina cap a la identificació fonètica de les unitats més senzilles en què es pot segmentar la parla. Partint d'aquests supòsits, en aquests darrers anys s'han desenvolupat nombrosos estudis (vegeu, més endavant, l'apartat 2.4) per tal de destriar els elements essencials i invariables que es troben, pressumptament, en l'estímul acústic, els quals són susceptibles de ser reconeguts i categoritzats per tots els usuaris d'una determinada llengua. Els resultats d'aquests treballs han demostrat que el senyal de la parla és altament variable entre els diferents parlants i, fins i tot, en un mateix parlant.

Partint d'un enfocament més ampli d'aquesta qüestió, es pot considerar que la tasca del receptor consisteix a elaborar una informació de tipus semàntic a partir del senyal, donat que la finalitat última de tot el procés comunicatiu s'adreça a la comprensió del missatge. En aquest sentit, la identificació de les unitats acústiques i fonètiques esdevé menys rellevant, ja que, per mitjà d'un seguit d'operacions mentals el receptor procedeix a elaborar una representació de caire cognitiu. Probablement, aquest procediment no consisteix en l'aplicació d'un patró perceptiu constant, idèntic en totes les situacions comunicatives, sinó en una adaptació continuada al senyal, atès el seu caràcter variable.

Per tal de poder elaborar models que expliquin com es produeixen els processos per a la comprensió del missatge parlat, ens cal contraposar el paper que juguen cadascuna de les dues fonts d'informació: la que procedeix del senyal de parla (font acústica, fonètica i processament prelèxic) i la que procedeix del coneixement que té, de la llengua i del món, el propi receptor. Gràcies a aquestes dues fonts informatives, es pot adaptar a la situació comunicativa concreta del missatge (accés al lèxic i reconeixement del context semàntic i comunicatiu).

El present treball s'emmarca en l'exploració dels processos que activa el receptor en l'intent de comprendre el missatge parlat i té per objectiu l'estudi de la interrelació de dues informacions: la dependent del senyal i la independent del senyal. El paper que suposadament juguen aquestes dues fonts informatives ha estat descrit per Björn Lindblom en diversos treballs (1987b, 1990, 1991, 1996), en els quals es desenvolupa el que es coneix com la "Teoria de la parla hiper i hipoarticulada" (*Hyper- & Hypo- Speech Theory*).

En aquesta teoria, Lindblom es centra en l'adaptació contant que exerceix l'emissor en la producció de la parla per tal que, amb el mínim esforç articuladori s'aconsegueixi la màxima discriminabilitat i, en conseqüència, el receptor pugui integrar el missatge. Aquesta adaptació, que varia constantment segons les necessitats comunicatives, es desenvolupa al llarg del contínuum entre dos extrems: la "parla clara" i la "parla no-clara". Així, doncs, el referent del model teòric de Lindblom es centra bàsicament en la producció de la parla, a la qual atorga màxima importància per tal que el missatge sigui suficientment discriminable. No obstant, en el present treball hem volgut comprovar l'adequació d'aquest model des del punt de vista de la percepció, en el benentès que la producció de la parla està teleològicament destinada a la seva

descodificació i comprensió.

Així, doncs, suposem que l'adaptació que s'exerceix en la producció del missatge ha de ser corresposta, en la seva descodificació, per part del receptor. Altrament atribuiríem un paper excessivament passiu a la tasca perceptiva. L'aproximació a aquests processos d'adaptació constant del receptor centra la recerca que desenvolupem en el present estudi.

La metodologia de recerca utilitzada en l'experiment es basa en l'estudi de la percepció d'elocucions atenuades fins a intensitats pròximes al llindar de l'audició de la parla. A més, en aquestes condicions perceptives tan difícils per als receptors, hem impedit la possible interrelació parlant-oient, és a dir, la qualitat del senyal que produeix l'emissor no pot ser condicionada per la percepció, bona o dolenta, del receptor. Així, doncs, aïllant la possible interrelació del receptor hem pogut analitzar algunes de les estratègies perceptives que aquest posa a disposició per a fer front a un senyal de parla tan degradat.

Hem basat l'estudi d'aquestes estratègies tot classificant-les en tres grans plans perceptius: 1) la identificació fonètica (establint quins són els sons més resistents a la disminució de la intensitat); 2) el reconeixement lèxic (establint quines propietats fonològiques i semàntiques afavoreixen l'avaluació del lèxic en paraules amb sentit i sense sentit); 3) el context semàntic de la frase (establint el grau de facilitació dels diferents contextos en l'accés al lèxic i en la reducció dels errors de comprensió).

Les condicions perceptives que hem manipulat durant l'experiment poden semblar excepcionals, no obstant són ben presents en alguns actes comunicatius reals. De fet, es poden trobar exemples de situacions en les quals l'emissor no pot avaluar (i, per tant, no pot satisfer) les necessitats perceptives del receptor. Ens referim a casos com el de les locucions radiofòniques deficientment percebudes, el dels parlaments públics escoltats a una certa distància, el de les converses telefòniques amb diversos receptors (amb l'opció de "mans lliures"), el de les ordres donades a diverses persones alhora (com ara, el cas del realitzador de televisió als operadors de càmera). Així mateix, es poden donar situacions en les quals la parla ha estat enregistrada amb una qualitat deficient o bé es produeix mitjançant un sintetitzador de veu, el qual evidentment no pot contemplar la situació comunicativa més enllà d'uns supòsits limitats.

Els resultats d'aquesta tesi poden utilitzar-se en la recerca d'aquests i d'altres camps de la comunicació humana. Així mateix, alguns àmbits de la tecnologia de la parla necessiten eines que permetin valorar els avenços que constantment es van assolint. En aquest sentit, les conclusions del present treball poden proporcionar informació útil per a l'elaboració de proves d'avaluació de conversors de text a veu amb elocucions semànticament impredecibles (vegeu Benoît, 1990), o bé, per al reconeixement de la parla en ambients adversos, amb una relació senyal/soroll molt baixa, o bé en l'avaluació dels sistemes de transmissió de veu en les telecomunicacions. En relació amb aquest àmbit de les comunicacions, també pot ser d'utilitat en els sistemes de codificació i descodificació de la veu per a ser transmesa a través de la línia telefònica. En aquest cas es plantegen problemes de reconstrucció del senyal amb una qualitat suficient perquè pugui ser percebut i comprès (vegeu Juola, 1996 a,b). Finalment, també pot ser útil per a confeccionar eines de diagnòstic clínic per tal d'atendre persones amb deficiències auditives o amb deficiències de sordesa verbal.

Pel que fa als continguts que presentem en aquest treball, els hem organitzats en quatre grans apartats:

1.-En la primera part (capítols 1 i 2), fem un repàs dels elements que intervenen en l'acte comunicatiu i dels factors que poden influir en el procés que s'inicia en la producció de la parla i que recorre els diferents estadis de l'audició, la percepció i la comprensió. Atès que el nostre interès es centra en la perspectiva de l'oient, hem prestat més atenció en aquells aspectes que incideixen d'una manera o altra en el procés d'integració de la informació continguda en el missatge. Així mateix, hem fet una revisió dels treballs més rellevants que han estat elaborats al voltant dels diferents elements comunicatius, i el d'aquells que tenen una relació directa amb l'estudi que hem realitzat. També destaquem el repàs dels principals models perceptius, d'accés al lèxic i al reconeixement de paraules que s'han formulat en els darrers anys i que actualment són motiu de múltiples treballs de recerca per a la seva confrontació.

2.- En la segona part (capítol 3), presentem el model de Lindblom que inspira la present recerca (la teoria de la parla hiper i hipoarticulada), així com el desenvolupament i les aplicacions que n'han fet diversos autors. Posteriorment es desenvolupa la discussió al voltant de la teoria i es formulen les consideracions i les hipòtesis de treball que es deriven de l'enfocament perceptiu i que posen a prova la validesa d'aquest model.

3.-En la tercera part (capítol 4), s'estructura una proposta metodològica per tal de poder afrontar la comprovació de les hipòtesis plantejades en el capítol 3. A fi d'establir una guia de treball que articuli l'obtenció de resultats, es formulen tres grans objectius i es valoren els possibles factors que incideixen de forma més o menys directa en l'execució de l'experiment i en el posterior tractament de les dades.

4.-Finalment, en la quarta part (capítols 5 i 6), es presenten els resultats de l'experiment, d'acord amb els objectius marcats en el capítol 4, tot valorant-los un a un respecte de la bibliografia presentada en la primera part i respecte de les hipòtesis formulades. El comentari i la discussió dels resultats és constant al llarg de tota la presentació de resultats, la qual cosa forneix la base per a la validació de les hipòtesis que s'exposen en el capítol 6, de la qual es desprenen les conclusions del treball.

En un annex, a part, es recullen els materials utilitzats durant l'experiment i el tractament estadístic de les dades que es va realitzar per tal d'interpretar els resultats.

Finalment, ens cal assenyalar que l'estudi que plantegem en aquest treball intenta afrontar els processos de percepció i de comprensió de la parla des d'un punt de vista global, és a dir, apropant-nos a cadascun dels diferents nivells d'anàlisi partint d'un únic experiment. Estem convençuts que qualsevol intent de tractar els diferents processos de forma individual, sense tenir present la interacció amb els altres nivells superiors i inferiors, suposa un enfocament incomplet. Som conscients que el risc que assumim en tractar el procés d'integració en el seu conjunt és que la complexitat de la recerca no permet aprofundir alguns dels aspectes del procés. L'avantatge rau en la possibilitat de fer una observació globalitzadora.

1.- LA PERSPECTIVA DE L'OIENT I EL PROCÉS AUDITIU

1.1.-Introducció

En aquest primer capítol farem un recorregut pels diferents elements que componen tot acte comunicatiu parlat entre l'emissor i el receptor. L'enfocament d'aquesta revisió tindrà la perspectiva de l'oient, és a dir, la d'aquell que no pot intervenir en el senyal que li arriba i per tant es veu abocat a integrar-ne la informació des d'una posició estàtica. Entre els elements que analitzem remarcarem els que venen condicionats pel locutor, els de la pròpia naturalesa del senyal i els corresponents al sistema auditiu del receptor, presentant així les tres etapes clàssiques de l'acte comunicatiu.

És evident que abans de començar a definir qualsevol treball empíric cal conèixer el comportament de les variables que poden influir, a priori, en els plantejaments experimentals. Tractant-se de la parla, seria molt difícil d'enumerar tots els factors que, d'una manera o altra, tenen a veure amb la variació que pot produir-se en els diferents moments de l'acte comunicatiu que estudiem en aquest capítol.

La producció de la parla està sotmesa a una gran plasticitat. Aquesta propietat es fa palesa en la producció de sons diferents que corresponen a manifestacions fonètiques diverses de categories conceptuals com els fonemes i en la possibilitat d'emetre enunciats adaptats a cada situació ¹. Si considerem que el resultat d'aquest procés (*output*) és susceptible de ser percebut i comprès per oients diferents que pertanyen a la mateixa comunitat lingüística, per tal de poder plantejar un experiment en aquest camp, es fa imprescindible ajustar i concretar els factors de variació que afecten de manera més directa la tasca de percebre i de comprendre.

Molts estudis sobre la percepció de la parla s'han centrat en el control d'aquelles característiques considerades essencials en la configuració dels diferents sons, renunciant a unes altres que aporten una informació secundària.² Per aquest motiu s'han portat a terme experiments utilitzant sistemes de generació de senyals sintetitzats, suficientment semblants a la parla natural, els quals permeten un control acurat d'algunes variables que resultaria impossible de fer amb parlants humans. Des del nostre punt de vista, però, hem volgut fer un plantejament de recerca a partir de la parla natural perquè les variables que cal controlar en aquest estudi afecten globalment cadascun dels al·lòfons que integren el corpus de l'experiment, en comptes

1.-Lindblom (1987b)

2.-Els trets que més s'han tingut en compte fins ara són: En el domini freqüencial, la distribució formàntica (sobretot els formants que es troben entre les freqüències de 200 a 5600 Hz., que és on es troba la màxima informació en la percepció de la parla (Pavlovic i Studebaker, 1984)), la freqüència fonamental, els harmònics més intensos, etc., i s'ha renunciat a les freqüències més altes, que aporten la naturalitat i ajuden més a la intel·ligibilitat dels sons (Pastore, 1981). En el domini del temps, els efectes de la coarticulació i els seus correlats acústics: Black, 1949; House, 1961; Kimbrough, 1973; Klatt, 1973; Stevens, 1980; Carlson i Grantrom, 1986; Beardsley i Cullinam, 1987; Sachs *et alt*, 1988; etc. Finalment, el domini de la intensitat ha estat tractat per autors com Lehiste i Peterson, 1959.

de centrar-se en l'estudi dels diversos indicis acústics característics de cada classe de sons.

Per una necessitat pràctica hem reduït els condicionants que afecten la parla com a acte comunicatiu a un nombre limitat de variables, conscients que la simplificació pot comportar errors i conscients també que aquestes variables són sovint difícils de controlar. Així, doncs, hem distingit entre alguns dels trets característics de la parla que arriben al receptor sense que aquest hagi de intervenir de manera voluntària i alguns dels trets que requereixen la seva col·laboració intencionada, conscient i atenta. Dels primers n'anomenarem els elements "passius" del procés comunicatiu verbal (apartat 1.2) i, dels segons, els elements "actius" d'aquest procés (capítol 2).

1.2.-Els elements "passius" del procés comunicatiu verbal

Des de la perspectiva del receptor que participa en el procés de la comunicació, existeixen unes característiques implícites en el senyal o en els estadis inferiors de la percepció que tenen origen o bé en l'emissor, o bé en el context extralingüístic³, o bé en el propi receptor i que exerceixen, tots ells, una influència sobre les condicions amb què es percep el missatge. No obstant, el receptor no pot intervenir sobre aquestes característiques per tal de modificar-les o d'adaptar-les a les seves pròpies necessitats com a destinatari, per la qual cosa se solen definir com a elements "passius".

Per tal de presentar-les breument, les hem dividit entre aquelles característiques de la locució (apartat 1.3), del senyal (apartat 1.4) i de l'audició (apartat 1.5) que, d'una manera o altra, tenen una certa incidència sobre el procés de la comprensió.

1.3.-Les característiques de la locució

En la producció de la parla cal distingir dos punts de vista clarament diferenciats. Per una banda, la realització aïllada de sons de la parla més o menys estables i, per altra, la construcció de seqüències dinàmiques en les quals els sons es solapen i s'influencien els uns amb els altres, degut al fenomen que anomenem coarticulació, responsable de bona part de la variabilitat fonètica de la parla.

El parlant regula aquesta coarticulació entre els sons i controla la manera de presentar-los a l'oient en funció de la situació que capta amb els sentits i del tipus d'intervenció que pretén aconseguir. Per a porta-ho a terme, de vegades presenta les seqüències sobrearticulantes, a fi que el missatge tingui les màximes garanties acústiques i articulatòries (*hyperspeech*). D'altres, en canvi, considera que les circumstàncies ambientals són prou bones i no s'esforça especialment per a produir la locució (*hypospeech*)⁴.

3 .-D'acord amb Jakobson (1960), també considerarem com a sinònims d'aquest terme la "situació contextual" o, simplement, el "context". Vegeu-ho més endavant en els apartats 1.4.1.1 i 1.4.2. Cal distingir aquest context de l'anomenat "context fònic" o "context estructural" (*environment*), exposat en l'apartat 1.4.1.2.

4 .-La teoria de la *hyper-hypo speech*, desenvolupada per Lindblom (1987b, 1990), està fonamentada en els

En tot cas, el parlant no utilitza més esforços que els necessaris per a la situació concreta en què es produeix l'acte comunicatiu, segons sosté la hipòtesi de l'economia fònica. El criteri que regula la necessitat d'articular amb més o menys precisió en cada moment el determina la finalitat més important de la parla, és a dir, l'elaboració d'un estímul amb significat, posant-hi les condicions més favorables per tal que l'oient pugui desxifrar i comprendre el missatge. Tot plegat fa suposar que, a l'igual com s'esdevé en tots els moviments biològics, el control motor de la parla està teleològicament organitzat de manera que s'adapti a les circumstàncies i a la finalitat que pretén el parlant, que és comunicar-se (Lindblom, 1983b). Això comporta que la parla es manifesti amb una important variabilitat, a fi que es pugui adaptar a les condicions pròpies de cada situació comunicativa. (d'aquesta darrera consideració en parlarem més endavant, en el capítol 3.)

Per tot el que s'acaba de dir, no hi ha dubte de la importància que tenen les característiques de l'acte de producció de la parla en l'èxit de la percepció del senyal. Els estudis per sistematitzar aquestes característiques sovint es basen en les anomenades proves d'intel·ligibilitat⁵, mitjançant les quals s'intenta d'establir una sèrie de categories discretes relacionades amb la facilitat de percebre i entendre el missatge, així com les condicions acústiques òptimes de la seva producció i transmissió⁶.

Així, es pot determinar la influència que exerceixen sobre la percepció, tant les característiques que imprimeix el locutor en el moment de la producció de la parla com les característiques de la codificació de la parla en relació amb la intenció comunicativa. De fet, qualsevol missatge parlat aporta una informació àmplia i que s'extén més enllà de l'estricta contingut que es transmet. La forma de la pròpia locució ja aporta molta informació d'aquest tipus, atès que revela trets com per exemple els següents:

-Trets derivats de les característiques fisiològiques particulars del parlant, com la presència o absència de defectes de la parla (com per exemple, la hiper-nasalització), sovint produïts per defectes físics dels òrgans articuladoris o per afeccions transitòries en aquests òrgans (inflamacions, refredats, afonies, etc.). (Stevens, 1972a)

-La hiper/hipo-articulació, de la qual ja hem parlat anteriorment en aquest mateix apartat, i que s'adapta a les condicions ambientals en què es produeix l'acte comunicatiu.

-L'esforç articuladori en la realització dels sons, conegut com "tensió", el qual tindrà un correlat físic que es podrà mesurar i que s'anomena intensitat acústica i un correlat psicoacústic que produirà una sensació subjectiva en l'oient anomenada sonoritat (Fairbanks *et al*, 1950; Lehiste i Peterson, 1959)

estudis de recerca sobre la invariança fonètica realitzats per MacNeilage (1970, 1980). Posteriorment, s'han realitzat experiments sobre la compensació articulatòria que han reforçat notablement aquesta teoria. (vegeu-ho en el cap.3).

5.- Miller i Nicely (1955); American Standard Association (1960); Gerber (1974b); Voiers (1977, 1981); Pastore (1981); Schmidt-Nielsen (1983); Walver i Byrne (1985).

6.- Inicialment les proves d'audició de la parla foren utilitzats per Flechter i Steinber (1929) en els laboratoris de Bell Telephone com una eina d'avaluació de l'efectivitat de diversos sistemes de comunicació electrònics, inclòs el telèfon. Posteriorment, al 1942, Hughson i Thompson van ser els primers que enfocaren aquestes proves com una eina de diagnosi en aplicacions clíniques.

-L'altura tonal habitual o excepcional que decideix el locutor segons el missatge que es vol transmetre

-Els trets de domini i adequació lingüístics en relació amb factors prosòdics com l'accent, l'entonació, la durada, el ritme i les pauses.

-La velocitat de locució. Segons Levelt (1989), la velocitat de la parla normal s'aproxima a les 150 paraules per minut. Això representa una mitjana d'una paraula cada 400 milisegons. No obstant, sota causes estressants o en situacions de limitació del temps es pot doblar fàcilment i aconseguir una mitjana d'una paraula cada 200 milisegons⁷.

Des del punt de vista exclusiu de la producció de la parla, existeixen descripcions molt acurades sobre aspectes com la qualitat de la veu (Laver, 1980), el mode de fonació (Ladefoged, 1971), els hàbits articulatoris (Nolan, 1983) o el funcionament no patològic de cadascun dels components de l'aparell articulatori (Stevens, 1972b).

Entre els elements que configuren l'acte de la producció de la parla, també hi apareixen condicionants més generals, propis del context ambiental, social i cultural en el marc del qual el locutor desenvolupa la seva acció. N'assenyalem només uns quants:

-El coneixement i domini que té el locutor sobre el tema o aspecte que tracta en el seu missatge. Aquest punt està molt estretament relacionat amb el nivell socio-cultural dels interlocutors i amb l'àmbit d'ús lingüístic on es troben immersos.

- El nivell d'educació i de cultura del parlant també influeix en l'ús d'un corpus lèxic més o menys ampli. Segons Levelt (1989), un parlant adult normal i educat escull, en la seva parla habitual, d'entre 30.000 mots que conformen el seu vocabulari actiu. Per altra banda, Levelt cita l'estudi d'Oldfield (1963) en el qual es va poder comprovar que els estudiants de grau mig d'Oxford disposaven d'un conjunt de 75.000 mots en el seu vocabulari passiu, que és aquell en què un individu comprèn les paraules encara que no les utilitzi en el moment de produir missatges.

-També depèn del parlant l'elecció dels constituents estructurals més amplis del discurs. L'organització de la parla segons unes regles de la sintaxi més o menys normatives pot adoptar diverses formes, combinant-les adequadament amb la morfologia de les paraules. Per aquest motiu, i a criteri del locutor, es poden produir, per exemple, formes més o menys emfatitzades capgirant l'ordre lògic dels elements del discurs o es pot allargar o escurçar el missatge segons la voluntat de precisió del parlant (vegeu Clark & Clark, 1977; Levelt, 1989). La decisió del parlant és força més imprevisible com menys definit sigui el context, de manera que el receptor ha d'estar disposat a rebre i interpretar un nombre molt elevat de possibles missatges parlats i de possibles estructures diferents.

7.-Sobre l'estructura fonètica de la parla ràpida vegeu, per exemple, Dalby (1986).

- La plasticitat motora que, com ja hem dit abans, es relaciona estretament amb la situació contextual, amb la intenció i amb la finalitat expressiva del locutor. (Lindblom, 1983a)

- Els estils de parla⁸ que qualsevol parlant és capaç de variar segons la situació comunicativa en què es troba, tot adoptant formes ben diferenciades: parla ràpida, lenta, fluixa, forta, més entenedora, menys entenedora, adaptada a un auditori ampli, a una situació d'intimitat, a estrangers, a nens, a persones sordes, a un ordinador, etc. Tot parlant canvia aquestes maneres de parlar en funció de les regles socials que governen les interaccions entre el parlant i l'oient

Notem altra vegada la dificultat d'establir un barem calibrador capaç de determinar com s'adapten tots aquests aspectes a les condicions que requereix una situació comunicativa concreta. Es la convenció social i l'èxit comunicatiu (que és finalment el que es pretén) els que determinen el seu bon ús o mal ús.

En el nostre cas entendrem que es donen bones condicions de locució quan la major part d'aquests defectes físics de l'aparell articuladori hi són absents i si els subjectes que escolten el discurs són capaços de repetir-lo correctament, tot i que algunes parts d'aquest discurs no tinguin significat conegut (ex: síl.labes aïllades i sense sentit, logotomes, etc.).

1.4.-Les característiques del senyal

Per poder encetar la discussió sobre els efectes del senyal sobre l'oient o receptor i les seves interrelacions, convé deixar clars alguns termes que s'utilitzaran al llarg d'aquest treball. Per raons òbvies caldrà fer-ho breument.

Un senyal acústic és una funció temporal en la qual es concreta un valor d'amplitud per a cada instant del temps. Segons les seves característiques acústiques, els senyals poden classificar-se en sons (si tenen trets o elements periòdics en la seva composició) i sorolls (si són totalment aperiòdics). Per l'estructura de la seva forma d'ona també es poden classificar en simples (si estan constituïts per una sola freqüència, i aleshores s'anomenen tons) o compostos (si és el resultat de la suma de diverses ones simples).

Els components físics d'aquests senyals són la freqüència, mesurada en cicles per segon (Hz.), la intensitat, mesurada en decibels (dB.) i el temps, mesurat en segons, o fraccions de segon (s.).

Finalment, també podem classificar els senyals segons si són parla o no-parla (també

8 .- Sobre aquesta qüestió es poden consultar, entre altres, els següents treballs: Aguilar, Machuca i Martínez, 1981; Aguilar i Machuca, 1984; Engstrand i Krull, 1988a, 1988b; Eskénazi, 1993; Gregory i Carroll, 1978; Labov, 1972a; Llisterri, 1992; Schmidt-Nielsen, 1983.

anomenats fonètics i no-fonètics, respectivament).

El terme parla s'usa per caracteritzar l'estímul produït per l'aparell vocal humà i que té un significat abstracte. Aquest estímul està construït mitjançant l'encadenament de diversos sons discrets, ordenadament estructurats, susceptibles de ser segmentats en ells mateixos o en unitats lingüístiques més grans, conformant un continuïum sonor. A aquests sons de la parla s'aplica el terme fonètic i al seguit de sons, cadena fònica. (Gerber, 1974b)

La tos, el plor, el xiulet, el xiscle, i tots els altres sons que no puguin ser components d'aquesta cadena fònica per la impossibilitat de combinar-los de manera sistemàtica amb els altres, encara que siguin produïts per l'aparell vocal humà, s'han de considerar com no-parla.

Però la parla és molt més complexa que un seguit de sons que es corresponen amb un missatge neutre. En tot cas, els missatges no són mai neutres i comporten un coneixement ampli del codi per part de qui els usa. Se suposa, doncs, que el senyal ha estat construït amb coneixements sobre la llengua, sobre els sons que s'hi empren, sobre les regles que regeixen la seva possible combinació, sobre els matisos que li aporten els trets suprasegmentals com l'entonació, els canvis d'intensitat, el tempo, etc. (Sobre l'especificitat de la parla vegeu l'apartat 2.7.2 del capítol següent)

Tots aquests elements es poden identificar en un senyal de parla. Són tangibles i mesurables des de punts de vista i disciplines tant diferents com l'acústica, la lingüística, la tecnologia electrònica, etc. Són els elements que depenen únicament del senyal i estan estretament relacionats amb altres característiques que són presents en l'acte de parla (acte comunicatiu). Aquestes altres característiques extralingüístiques representen les condicions independents del senyal, ja que envolten la situació on es produeix aquest acte de parla. Els interlocutors que de manera activa (emissor) o passiva (receptor) operen amb el senyal tenen una certa informació extralingüística sobre el que s'està esdevenint al seu voltant. Són els coneixements sobre la temàtica, la situació, l'interlocutor, el moment comunicatiu, l'adequació del missatge, les relacions personals, els components no verbals, la intenció última del missatge, etc. La disciplina que es fa càrrec de la seva descripció (la semiòtica) abraça un camp molt difús i més difícil de descriure (un recull força complet sobre aquest tipus de components que es troben en la parla es pot consultar a Scherer - Giles (eds.), 1979)

Tot plegat, tant la informació que arriba al receptor a través del senyal (l'estímul) com la informació independent del senyal (el context) són les dues cares del mateix esdeveniment, que en definitiva és la comunicació. La combinació de totes dues proporciona al receptor els elements necessaris per a la comprensió del sentit complet del missatge (Lindblom, 1987b)

1.4.1.-Els trets de la parla que es manifesten en el senyal

El punt de partida en l'estudi acústic es dona en el fet que el senyal de la parla està format per un seguit de segments acústics (anomenats al·lòfons), encadenats els uns amb els altres i que varien segons les característiques generals de l'estil de parla adoptat i del context comunicatiu en què es produeix.

Les representacions abstractes i ideals d'aquests segments acústics s'estableixen en categories fonològiques. Cadascuna d'elles conté uns trets essencials i definidors de la categoria, en oposició a la resta. El conjunt d'aquests trets per a cada categoria configura el que coneixem per **fonema** (Jakobson i Halle, 1956; Anderson, 1985). La successió de sons emesos per un parlant per confeccionar un enunciat s'anomena **cadena fònica o cadena de la parla**.

Jakobson & Halle (1956) estableixen una diferenciació entre els trets intrínsecs, que apareixen en qualsevol fonema independentment de la seva funció dins de la cadena de la parla, i els trets prosòdics, que estan basats en la funció del temps i els trets definidors dels quals es fonamenten en els contrastos que hi ha dins de la cadena fònica entre un segment i els pròxims.

El fonema representa l'agrupació d'aquests trets intrínsecs i distintius, en tant que oposats, que tenen tots els sons de la parla. La definició que tradicionalment es dona d'aquests trets no està basada en la percepció; no obstant això, les confusions que es donen en la percepció s'hi correlacionen força, tal com intentarem fer veure en l'apartat 5.1.16 i ss..

Tant els trets intrínsecs com els prosòdics estan correlacionats amb unes característiques acústiques que poden ser analitzades de manera tangible, ateses les seves condicions físiques⁹.

1.4.1.1.- Els trets intrínsecs del senyal

Les qualitats físiques de la parla es troben, òbviament, emmarcades dins dels límits de la capacitat auditiva del receptor. Tant pel que fa a la freqüència com a la intensitat o a la durada, els òrgans de producció i de recepció de la parla condicionen les característiques dels sons.

La distribució freqüencial d'aquests sons abraça un marge que s'extén des dels 30-40 Hz. fins als 10.000 Hz. No obstant, no tots els sons distribueixen la seva energia en tot l'espectre. Així, per exemple, les vocals no tenen freqüències significativament importants per a la seva identificació més amunt dels 4000 Hz., mentre que els sons fricativs mantenen l'energia fins als 10.000 Hz.

Segons les característiques dels sons es pot apreciar que certes bandes de freqüència contenen més energia que d'altres. Aquesta distribució no es troba sempre exactament en el patró de cada so (*sound pattern*), sinó que varia depenent del parlant (el seu tracte vocal, els seus trets particulars, etc.), de les condicions acústiques en què s'ha produït el so (el medi), del context fonètic i de la situació contextual on es produeix i com es produeix.

En els mesuraments amb filtres de banda ampla (de més de 300 Hz.), aquestes concentracions d'energia apareixen clarament distribuïdes i identificables, sobretot, en els sons vocàlics i en les consonants líquides i nasals. S'anomenen formants i tenen una amplada freqüencial d'entre 40 i 250 Hz. Cadascuna de les vocals té una amplada específica en els seus

9.-En relació amb els correlats acústics dels trets distintius són importants els treballs de Blumstein i Stevens (1981); Chomsky i Halle (1982); Delattre (1970), Delgutte (1982); Jakobson, Fant i Halle (1952); Jakobson i Halle (1956); Stevens (1983); Stevens i Blumstein (1981).

formants. (Dunn, 1961)

Els mesuraments amb filtres de banda estreta (de 45 Hz. o menys), aporten un registre de la freqüència fonamental sobre la que es basen tots els altres harmònics de l'espectre i que determina el to intrínsec de cada segment sonor del senyal i l'altura tonal. Segons Gerber (1974b), la freqüència fonamental de la veu és rarament inferior a 85 Hz. La mitjana de la veu masculina adulta es troba al voltant dels 125 Hz i la de la veu femenina adulta al voltant dels 212 Hz.

Una altre paràmetre del senyal és la intensitat. L'abordem des de dos punts de vista diferents:

1.- Segons la distribució de l'energia al llarg de l'espectre freqüencial, sobretot dels seus formants

En estudis comparatius de vocals amb un nivell d'intensitat total equiparable, on l'energia global de so és la mateixa, la distribució freqüencial d'aquesta energia permet distingir els diferents sons (Fant *et alt.*, 1963; Delattre, 1969; Carlson *et alt.*, 1970; Rossi, 1971). L'energia pot concentrar-se a la banda més baixa i llavors es diu que és un so greu en oposició amb un d'agut que la tindria a la part alta de l'espectre. O bé, es pot parlar de so vocàlic si disposa d'una estructura formàntica a diferència dels sons consonàntica que tenen l'energia dispersa. (Jakobson - Halle, 1956).

2.- Segons la intensitat intrínseca de cada so

Si s'observa la representació de la corba d'intensitat (envolupant mitjana d'intensitat sobre l'eix temps) d'un segment de parla s'aprecia força canvis de nivell entre els sons que formen part de la cadena de sons. Hom podria aduir que aquestes fluctuacions d'intensitat s'ajusten exclusivament a raons d'accentuació o a motius més generals de tipus prosòdic. No obstant, la mesura del nivell d'intensitat dels sons, sobretot dels vocàlics pronunciats amb el mateix esforç, ha estat motiu d'estudi des dels primers treballs de Black (1949) i Fairbanks *et alt.* (1950). Aquests autors han demostrat que, independentment de l'accent i dels trets prosòdics, es pot establir una intensitat específica per a cadascun dels sons de la parla (de l'ordre de 4'5 dB de diferència entre el so vocàlic més intens i el menys intens).

Lehiste & Peterson (1959), partint dels resultats de diversos experiments, van suggerir que:

a.- Les vocals produïdes amb el mateix esforç són isòfones per als oients, malgrat llur diferència d'intensitat.

b.- Les mateixes vocals produïdes amb un esforç articulatori diferent ja no són isòfones, ni fins quan tenen una intensitat igual.

Rossi (1971) arriba a la conclusió que la intensitat és un fenomen lligat a la fonia o

nivell d'intensitat sonora de les vocals i que depèn de la distribució de l'energia dins de l'espectre. Llavors estableix dos tipus de vocals: les de banda estreta (amb un nombre de bandes crítiques inferiors a 10): /a, ə/ i les de banda ampla (amb més de 10 bandes crítiques) /i, e, ε, o, u/. Si l'energia espectral d'aquestes vocals es troba en la zona afavorida per l'oïda (entre 250 i 800 Hz d'una banda i 2000 i 4000 Hz d'una altra) la vocal percebuda augmenta la fonia: /i, u/. Mentre que les que tenen l'energia concentrada en la zona de 800 a 1600 Hz. no augmenten la fonia: /a/.

En relació amb la intensitat intrínseca de les vocals en llengua catalana i en llengua castellana, Blecua - Acín (1995) conclouen en el seu treball que les variables que més influeixen en la intensitat vocàlica són l'accent, la posició dins de l'enunciat (inicial o final) i si es troben en posició prepausal o no. Aquests resultats poden atribuir-se a les característiques pròpies de cada llengua, atès que en el cas del català i el castellà i degut a les diferències entre els seus inventaris vocàlics, es comprova que les propietats acústiques de les vocals no es comporten de la mateixa manera.

1.4.1.2.- Els trets prosòdics del senyal

La parla, ja s'ha dit abans, és un senyal serial distribuït al llarg del temps. Per aquest motiu, als principals atributs acústics del senyal (la intensitat i la freqüència), cal afegir-hi la pròpia evolució i interrelació d'aquests atributs en segments més o menys llargs. Això implica la presència de transicions, solapaments, fluctuacions i adaptacions d'un so a un altre al llarg d'un procés de coarticulació.

Jakobson & Halle (1956) prenen el concepte "prosòdic" com un conjunt d'atributs de la sensació (força, quantitat i to) que es dona només en aquells fonemes que formen la cresta de la síl.laba. No obstant això, la parla té estructures més llargues que la síl.laba. Atenent als tres atributs anteriors, hom troba en la parla qualitats com l'entonació, la velocitat de la parla, els canvis de velocitat en funció del missatge, o la relació entre durada i intensitat. La situació o context en què es troben els interlocutors determina el nivell de la intensitat general del missatge. Situacions com l'exaltació o el dolor (el crit), també l'enuig, la discreció, la llunyania o proximitat dels interlocutors, etc. configuren un context íntimament lligat amb el nivell d'intensitat. Tots aquests elements suprasegmentals difícilment es poden descriure des de la síl.laba

Tot plegat revela que el senyal no pot mirar-se, només, des del punt de vista de qualitats discretes de cada un dels seus components, sinó que dins de l'estructura acústica de la cadena fònica, cada so i cada segment depenen molt estretament del context fonètic que els envolta (Miller, 1981).

Donat, doncs, que la parla és un fenomen complex, es fa força difícil de descriure acústicament aquestes relacions entre els sons. Per exemple, és complicat d'establir, en termes de freqüència-temps, el que s'esdevé en una transició entre dues vocals o els canvis sobtats d'entonació en un discurs. Malgrat aquesta dificultat, hi ha l'evidència que la parla està organitzada al llarg de l'eix temporal de manera que pugui ser percebuda per l'oient, el qual no

té massa problemes per detectar-ne els matisos i extreure'n la informació. D'entre les característiques de la naturalesa de la parla continuada, se'n destaquen les següents:

-Els components de la parla són altament redundants. Segons Casacuberta i Vidal (1987a), en un sistema digital només caldria una informació de 50 bits/segon per transmetre el missatge bàsic que conté el senyal mentre que la parla normal és de 100.000 bits/segon. Malgrat la variabilitat i els ambients de soroll, el suplement redundat d'informació conté dades per fer comprendre el missatge, per donar trets d'identitat del parlant o del tipus i qualitat de la parla, etc.

Aquestes trets redundants estan associats, també, amb la sobrearticulació o subarticulació per part dels parlants (hipo/hiper articulació) (vegeu-ho en el cap.3).

-Els sons rarament es produeixen aïllats, sinó que els uns influeixen sobre els altres produint solapaments i encadenaments. Tot plegat configura un context fònic (*environment*), en el qual el comportament de cada so pot ser explicat per la conjunció dels sons adjacents i dels elements prosòdics que evolucionen conjuntament¹⁰. Per aquest motiu un so es percep, sovint, gràcies a les evolucions dels so veïns.

-La constància dels fonemes es desprèn de la generalització dels atributs no constants del senyal (al·lòfons) i permet d'establir unes categories¹¹ que aglutinen les múltiples realitzacions de cada patró articulatori.

-La informació que es vol transmetre sempre és més gran que l'estrictament continguda en el senyal de la parla, ja que va acompanyada d'elements que no són pròpiament de parla (gestualització, èmfasi, context, etc.).

-Una síl·laba no es pronuncia dues vegades exactament igual. Això es deu principalment a variacions circumstancials com ara la intenció lingüística (l'entonació, la sonoritat o el sentit del missatge), variacions temporals (la fluïdesa, la coarticulació, el tempo, la sobre o subarticulació), entorn acústic (medi), condicions físiques i psicològiques del locutor (el cansament, la malaltia, la fisiologia o els aspectes culturals) (Caelen, 1986; Stevens & House, 1963). Aquest tema el tractarem més endavant, a l'apartat sobre variabilitat adaptativa (vegeu el capítol 3).

-Molt sovint el senyal conté imprecisions o elements diferents als estrictament fonètics; per falta de part de la informació, per distorsions o per paràsits superposats. Es el que anomenem "interferències en el senyal".

10 .-Sobre el context consonàntic vegeu Delattre (1969); sobre la coarticulació vegeu Daniloff i Hammarberg (1978), Magno-Caldognetto (1980), Recasens (1984); sobre la dependència del context estructural vegeu Miller (1981)

11 .-Vegeu Lisker i Abramson (1970); Eimas i Corbit (1973)

1.4.2.-Interferències en el senyal: distorsions i emmascaraments

Ultra les característiques i peculiaritats que li imprimeix el locutor, el senyal pot deformar-se degut al mitjà que el transporta o a les condicions ambientals amb què es troba al llarg del seu recorregut fins al receptor. Pastore (1981) explica aquestes circumstàncies com un fenomen d'emascarament, en el qual es troba un estímul amb contingut informatiu i la interferència produïda per la presència d'un segon estímul anomenat emmascarador¹². Aquesta definició implica un gran nombre de tipus diferents d'interaccions entre els dos sons. En cada situació comunicativa es donen unes condicions ambientals específiques que fan que l'estímul comunicador prevalgui sobre l'emascarament, o viceversa.

Considerem que les condicions del senyal són bones si arriba al receptor amb una configuració similar a la que tenia en el seu origen, menyspreant la variació produïda per una lleu pèrdua d'intensitat¹³. En canvi, si la intensitat del senyal ha quedat reduïda de manera considerable, l'estímul resultant arriba a l'interlocutor en condicions de baixa qualitat.

L'avaluació de la qualitat de la parla es realitza mitjançant una mesura que relaciona la mitjana de l'energia del senyal i la mitjana de l'energia del soroll de fons del medi de transmissió. S'anomena "Relació senyal-soroll" (SNR= Signal-Noise Ratio) i s'expressa en dB segons la fórmula

$$SNR = 10 \log \frac{\bar{x} \text{ senyal}}{\bar{x} \text{ soroll}}$$

Evidentment, com més elevada és la relació resultant més alta serà la qualitat de l'estímul¹⁴. Vegem-ho esquematitzat a continuació, en la figura 1.1

12.-Vegeu Steeneken (1992), Studebaker (1977)

13.- Per a la consideració de "parla clara" des de diversos enfocaments, vegeu Bond i Moore (1994), Cutler i Butterfield (1990), Picheny, Durlach i Braidà (1986)

14.-Existeixen altres sistemes d'avaluació de la qualitat de la parla i que han estat desenvolupats des de l'enginyeria industrial per a aplicacions de sonorització d'ambients. S'anomenen STI (*Speech Transmission Index*) o RASTI (*Rapid STI*). Vegeu Houtgast i Steeneken (1985a, 1985b), Steeneken i Houtgast (1980, 1985); Steeneken (1992).

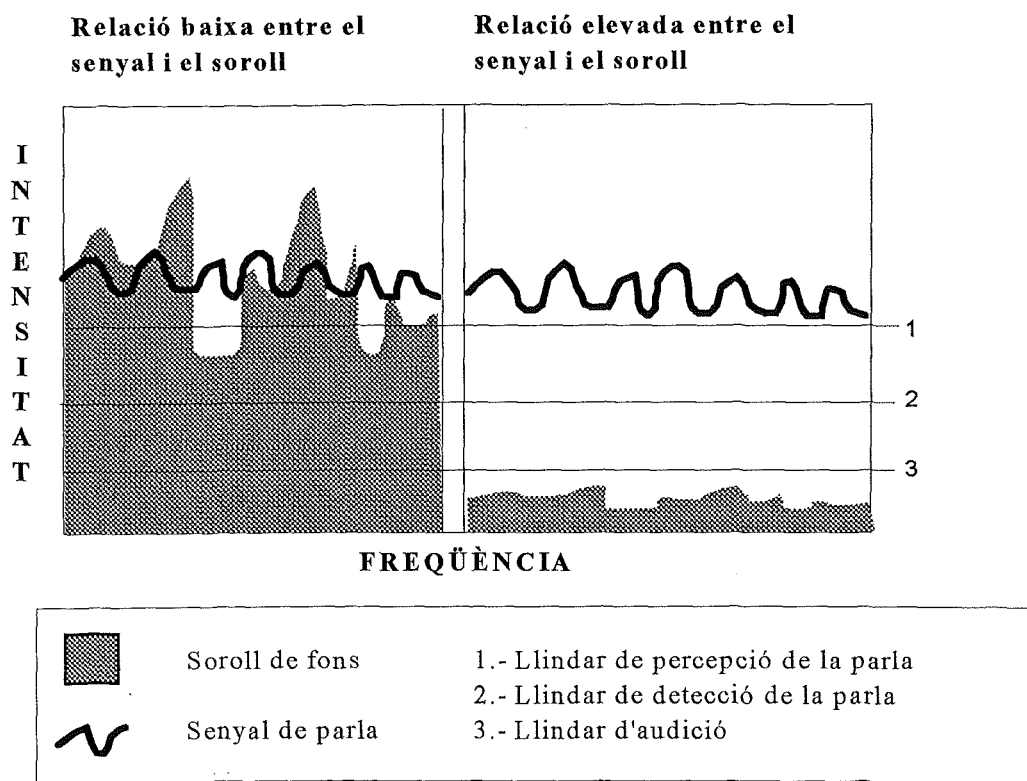


Figura 1.1: Percepció del senyal en condicions naturals, amb dificultats de comprensió si el soroll de fons representa un emmascarament excessiu (esquerra) i amb possibilitat de comprensió donat que el soroll de fons no representa cap emmascarament (dreta)

Arran d'aquesta relació, alguns autors com ara Flanagan *et al.* (1979) o Laver (1980) han classificat les diferents qualitats de la parla. Una d'aquestes classificacions, basada en quatre graus, es descriu a continuació:

1.-Qualitat de comentari o de retransmissió (*commentary or broadcast quality*): Es refereix a la parla d'alta qualitat amb una banda ampla de freqüències (0-7000 Hz) i sense soroll perceptible.

2.-Qualitat telefònica (*toll quality*): Descriu la qualitat de la parla a través de la xarxa telefònica (marge de 200-3200Hz., amb una SNR de més de 30 dB i menys de 2-3% de distorsió harmònica). En aquest cas, la major part de les vegades es perd la freqüència fonamental (F_0) i la percepció d'aquest tipus de parla es realitza per les diferències entre els harmònics.

3.-Qualitat de comunicacions (*communications-quality*): Es una parla prou intel.ligible però amb notables distorsions comparada amb la qualitat telefònica.

4.-Qualitat de parla sintetitzada (*synthetic-quality*): Té una intel.ligibilitat superior

al 80-90% però presenta una degradació considerable respecte de la parla natural (s'anomena, sovint, so de màquina o so metàl·lic), amb soroll de fons i sense trets identificadors de parlant.

En la parla natural, el senyal de la veu difícilment es transmet en condicions tan òptimes com per no atenuar-se o interferir-se amb altres senyals acústics ambientals. Les interferències i distorsions poden classificar-se per dos aspectes:

- a.-Pel mitjà transmissor del senyal (canal)¹⁵.
- b.-Per l'ambient que envolta l'emissor i el receptor (context)¹⁵.

Atès que la influència d'aquests dos factors sobre les condicions de percepció esdevenen realment importants, hem dedicat els dos següents apartats a presentar-los i a descriure'ls amb detall.

1.4.2.1.- El mitjà transmissor del senyal

Les característiques del mitjà transmissor del missatge determinen, en gran manera, que el senyal arribi en bones condicions al receptor. Així, sovint es pot detectar la presència d'emascaraments en el senyal produïts per les condicions del medi. En els mitjans habituals en què es produeix la parla (l'aire) s'hi donen dos tipus d'interferències que condicionen la qualitat del senyal. Per una banda el soroll ambiental de fons (relació senyal/soroll) que pot tenir procedències molt diverses. Per una altra la reverberació produïda pels propis elements que configuren l'entorn on es produeix l'acte de parla.

En locals tancats s'ha establert un temps mitjà òptim de reverberació (temps que triga en reduir-se el nivell de pressió sonora en 60 dB) segons el qual es considera que existeix una bona sonoritat. Per calcular-ho s'aplica la fórmula de Sabine:

$$\text{Temps de reverberació (en segons)} = k \frac{V}{S}$$

on k és un coeficient amb valor 0.165, V és el volum de la sala en m³ i S són les unitats d'absorció del so segons el coeficient de cada material que configura la superfície de les parets. (Setó, 1971)

A tall d'exemple es pot dir que una sonoritat acceptable per a un recinte de menys de 200 m² tindria un temps de reverberació de 0.8 segons. Evidentment els materials amb que està construït el local tenen molt a veure amb el que acabem de dir i permeten que sigui possible el

15.-Aquestes termes estan extrets de la clàssica descripció de l'esquema de la comunicació de Jakobson (1960)

control de la reverberació¹⁶.

Pel que fa als mitjans artificials per a la transmissió del senyal, per exemple en els electrònics, el canal delimita els marges freqüencials del senyal i la seva amplitud, i desestima els valors que es troben per sobre d'aquestes bandes. La informació transmesa també esdevé reduïda. En aquest cas la qualitat també es mesura segons la relació senyal-soroll.

1.4.2.2.- L'ambient acústic que envolta l'emissor i el receptor

El soroll ambiental neutralitza parcialment o totalment el senyal portador del missatge. Sovint els sorolls de presència continuada (*continuous simultaneous masking*) són els que necessiten menys intensitat per tal d'emascarar el senyal acústic de la parla, mentre que els sorolls puntuals (*gated simultaneous masking*) i de curta durada interfereixen breument i, per tant, no tenen tanta incidència sobre el senyal de la parla (Pastore, 1981).

Un altre soroll emmascarador habitual és el produït per la parla d'altres persones en el mateix recinte o canal per on es transmet el senyal. Aquest soroll de fons, anomenat '*Babbling noise*'¹⁷, és especialment emmascarador donat que abasta totes les freqüències de la parla i, a més, és difícil d'aïllar per part de l'oient quan l'ha de distingir dels altres sons.

1.5.-Les característiques de l'audició

En els estudis articulatoris, l'experimentador pot realitzar les tasques experimentals d'una manera força directa sobre els òrgans de la parla, moltes vegades, fins i tot, mitjançant l'observació visual. En canvi, els estudis sobre el funcionament de l'audició es basen, inevitablement, en procediments de recerca indirectes, ja que, en el sistema auditiu i el sistema

16.-Coeficients d'absorció d'alguns materials per a 512 Hz segons Allue (1968)

<u>Material</u>	<u>Coefficient d'absorció per m2</u>
Paret de ciment	0.015
Paret de totxo	0.034
Tela de cotó sobre la paret	0.150
Cortinatges gruixuts	0.300
Revestiment suro comprimit	0.300
Una persona	0.470
Fibra de vidre	0.650

També es pot trobar a Diamant (1967). Aquestes mesures han estat estandarditzades mitjançant les normes ISO 140/7 1978, UNE 74-040-84/7 i NBE-CA-81 que tracten de les condicions acústiques d'edificis.

17.-També conegut com '*cocktail party noise*'. Intenta recrear el soroll de molta gent parlant alhora. Per exemple, es pot aconseguir artificialment enregistrant un discurs, de manera desordenada, 24 vegades sobre una cinta magnetofònica, la qual no és esborrada en cada enregistrament. Vegeu McAllister i Dufberg, (1989), McBurney i Collings (1977)

central d'interpretació i comprensió, tots els òrgans són interns i és força complex aconseguir una observació significativa del seu funcionament, per la qual cosa ha calgut desenvolupar mètodes força sofisticats per a accedir-hi¹⁸.

Donades aquestes característiques, Clark & Clark (1977) apunten que les possibilitats d'estudi de l'audició i el sistema de percepció només són possibles, actualment, mitjançant:

- 1.- L'estudi de les característiques acústiques dels sons identificats.
- 2.- L'estudi de la identificació de sons produïts per sistemes de parla sintetitzada
- 3.- Estudi del comportament de la identificació en condicions d'emascarament i modificació de les propietats dels sons tant en el domini freqüencial com en el temporal.

Aquest tipus d'estudis proporcionen unes dades que estan, poc o molt, influïdes per la participació del subjecte i pels seus condicionants psicològics i, molt sovint, subjectius. Per això ha calgut relacionar dues disciplines, la psicologia i l'acústica, per aconseguir un treball conjunt i per assegurar l'obtenció de bons dissenys de recerca i bones possibilitats d'interpretació de les dades obtingudes. La **psicoacústica** recull i estudia la resposta subjectiva als estímuls acústics, particularment, en relació a la tasca de detecció i discriminació. (Vegeu Beagley i Knight, 1979; Hyde & Stephens, 1979; Knight, 1979; Moore, 1982; Pastore, 1981; Pickles, 1982; Tyler, 1988).

Els estímuls acústics que arriben cada dia i a totes hores a les oïdes de les persones contenen un tretsonors que són separats en funció de la informació que aporten a cada individu en aquell moment concret. D'aquesta manera, s'estableix una selecció entre allò que resulta d'interès i el que no. Atès que no podem evitar la constant arribada a l'oïda de les ones acústiques, els sistemes auditiu i perceptiu humans han desenvolupat una estratègia de selecció d'aquests sons. I si, a més, centrem l'atenció en la recepció dels sons de la parla, encara hi ha un mecanisme més especialitzat en aquesta tasca de selecció, per tal com la parla és estrictament un sistema complex de sons amb informació. (Moore, 1982)

A diferència d'altres anàlisis auditives, l'anàlisi de la parla, doncs, no és un procés purament auditiu. Hi ha una feina d'extracció d'indícis fonèmics del flux continu que arriba a l'oient. Tot això implica uns mecanismes molt més complexos, que no els merament auditius. Cal que es classifiquin aquests indicis i s'elevin a àrees del cervell per ser interpretats i integrats adequadament en el coneixement (Seneff, 1988).

A la vista de tot el que s'acaba de dir, diversos estudis sobre la comprensió de la parla (vegeu-los en el capítol 2) han establert uns estadis molt amplis on es conjuguen aquests processos que comencen amb l'arribada del senyal acústic a l'oïda i acaba en el moment de ser comprès. Els podem anomenar de manera genèrica com l'audició, la percepció i la comprensió.

En aquest primer capítol ens centrarem, sobretot, en el primer estadi i deixarem per més endavant el segon i el tercer. Aquesta divisió dins del procés es deu al fet que l'audició implica una activació automàtica i totalment involuntària (Studdert-Kennedy, 1974) dels centres

18.- Des de fa uns quants anys alguns autors han realitzat estudis sobre el procés de codificació de l'estímul acústic en el nervi auditiu. Concretament cal destacar, entre d'altres, els treballs de Delgutte, 1982; Delgutte i Kiang, 1984a,b,c,d,e; Geisler, 1988; Ghitza, 1988; Greenberg, 1988a; Greenberg *et al*, 1986; Kiang, 1975; Sachs *et al*, 1975

perifèrics de l'oïda (processos sensorials), mentre que la percepció i la comprensió s'estableixen mitjançant una participació activa per part de l'oient per integrar la nova informació al seu coneixement, tot fent intervenir, en aquesta integració, estadis més centrals dins del procés cognitiu.

Si bé és cert que els processos cognitius i els sensorials són contigus i interdependents (Schwartz, 1986), per una qüestió metodològica òbvia cal tractar-los establint aquesta divisió.

1.5.1.- L'audició

La primera part del procés auditiu, que anomenem audició, no implica que l'oient hagi d'efectuar cap mena d'interpretació del senyal que li arriba. L'aparell auditiu funciona com un transductor elèctric que transforma, de manera automàtica, les vibracions de l'ona acústica en energia elèctrica. L'energia aportada per l'ona acústica dins de l'oïda es converteix, mitjançant una excitació electro-química de les cèl·lules de la membrana basilar (Moore, 1982; Pickles, 1982), en un seguit d'estímuls nerviosos que poden ser tractats pels centres d'interpretació del cervell. En cap cas podem pensar que aquí, a l'oïda interna, es produeixi l'extracció del sentit de l'enunciat.

La descripció anatòmica i fisiològica de l'aparell de l'oïda és àmpliament i detalladament coneguda (Darwin, 1983; Diehl, Kluender i Walsh, 1990; Gerber, 1974b; Greenberg, 1988a, 1988b, 1996; Moore, 1982; Netter, 1987; Shamma, 1985; entre d'altres) i aquí no hi aprofundirem.

Sobre la manera com es produeix aquesta transformació del so en un codi neurofuncional, s'han proposat dues teories: la teoria de la localització i la teoria de les freqüències. La primera suposa que la membrana basilar està formada per un gran nombre d'elements cel·lulars fixos que responen específicament a les freqüències seleccionades per a cadascun. La segona proposa l'existència d'impulsos nerviosos que són susceptibles de respondre a la mateixa freqüència que els estímuls aferents i, per tant capaços de respondre a freqüències diferents. Totes dues teories han estat comprovades i desenvolupades per Georg von Békésy, el qual va arribar a la conclusió que el funcionament de l'aparell de l'audició havia de combinar aspectes de totes dues (von Békésy, 1960)¹⁹.

Alguns estudis actuals enfoquen la seva recerca cap a la comprovació del principi de la teoria de la localització, és a dir, de quina manera la freqüència, que és una entitat física, és codificada en la perifèria auditiva segons una representació tonotòpica i com es produeix la descodificació en parts més centrals del sistema auditiu per formar una entitat subjectiva que és el to. (Pantev *et alt.*, 1989)

Altres enfocaments en la mateixa direcció estudien els petits sistemes neuronals independents necessaris per a l'extracció de l'estructura espectral, la intensitat, la durada i la seva relació amb les dimensions subjectives de qualitat, timbre, to, estridència, etc. (Studdert-Kennedy, 1974; Seneff, 1988)

19.-Altres referències sobre el tema són: Geisler, 1988; Pickles, 1982; Schwartz, 1986; Zemlin, 1968.

Des d'un punt de vista psicoacústic, l'etapa auditiva inclou l'audició però integra també la memòria auditiva (el seu funcionament serà tractat més endavant, en el cap. 2) que no es dóna, evidentment, en l'aparell auditiu. En aquesta etapa s'identifiquen els sorolls que no són de la parla, donat que la parla implica estadis superiors d'interpretació.

L'audició, però, no és només una capacitat de recepció del senyal acústic sinó que també suposa d'altres conceptes que cal tenir en compte, perquè hi tenen una relació molt estreta: la detecció i la sensació.

Tant l'una com l'altra impliquen una base d'intensitat per poder-les definir i s'estableixen al voltant del concepte de llindar d'audició (vegeu-ho més endavant, a l'apartat 1.5.2). La **detecció o la detectabilitat** es vincula al fet que el subjecte s'adona, mitjançant el seu aparell auditiu, de la presència o absència de l'estímul acústic, però degut a la seva baixa intensitat, no té prou elements per reconèixer-lo (Pastore, 1981; Perelló, 1980).

En la recerca sobre l'audició i la percepció sovint s'ha treballat amb estímuls que s'han dirigit als oients-subjectes de manera que fos possible anotar-ne i descriure'n les reaccions consegüents²⁰. Els experimentadors han basat el disseny dels seus estudis en la resposta que es produïa en l'aparell de l'audició a partir d'un cert nivell d'intensitat de l'estímul. En els termes més elementals, la detecció es produeix en el moment que la intensitat de l'estímul supera un llindar d'excitació de l'òrgan sensorial.

Evidentment, la intensitat necessària per a la detecció augmenta a mesura que la durada de l'estímul decreix (Moore, 1982) (recordem que la intensitat és una mesura d'energia per unitat de temps). Això fa pensar que el llindar, com diu Bilger (1977), és una característica d'alguns tipus d'estímuls, i no una característica de l'òrgan sensorial.

Realment, l'òrgan de l'audició no respon de la mateixa manera a tots els estímuls. D'aquí que ja, des del començament de la psicoacústica, s'hagués treballat amb els llindars d'audibilitat. Concretament, els primers d'establir les corbes d'audició mínima van ser Sivian & White (1933) i les van expressar en termes de "mínima pressió audible" (MAP) i "mínim camp audible" (MAF) (Citat a Gerber, 1974b).

Més endavant, cap a finals de la dècada dels 50 i començament de la del 60, fou desenvolupada la "Teoria de la Detecció del Senyal" (SDT), presentada per Birdsall (1955); Tanner (1955) i Swets (1964). Aquesta teoria fou desenvolupada inicialment en el camp de l'enginyeria elèctrica i expressa l'ideal matemàtic o el procés de detecció òptim. D'aquesta manera va resultar ser un recurs molt valuós per descriure la detecció sensitiva (Clarke i Bilger, 1977). No intenta descriure el subjecte pròpiament dit, sinó la millor actuació possible, teòricament possible, per a la detecció de senyals concrets. Amb aquesta teoria es va proporcionar una eina molt potent per a l'estudi de l'actuació humana en la detecció i el reconeixement de senyals acústics i per desenvolupar models descriptius del sistema auditiu.

Per tot el que s'ha dit, es podria pensar que els experiments sobre els quals es basa la detectabilitat plantegen, només, la capacitat decisòria del subjecte sobre la recepció o no d'estímuls aïllats. Però la detecció del senyal comporta altres tipus de tasques, com ara

20 Sobre les tecnologies emprades en aquesta recerca vegeu Beagley (ed.) (1979); Hyde i Stephens (1979)

determinar els canvis d'intensitat d'un estímul acústic.

Portman i Portman (1967) planteja una gradació d'intensitat de menys a més per tal d'establir el darrer moment infraliminar de la detecció del so fins al moment en què la distinció del so és clara. L'experiment es féu amb un to pur, que és un so que no existeix lliurement a la natura però que estableix un nivell estable i homogeni. Partint de nivells totalment inaudibles, s'augmentà la intensitat del to fins que es féu palesa una lleu sensació, interrompuda i irregular, en la qual no era possible de reconèixer la freqüència. En aquest punt es defineix la detecció del so i la informació que rep el subjecte és força imprecisa. A mesura que s'anà augmentant la intensitat, la sensació del so anà adquirint més trets identificadors, fins que s'arribà a un nivell de so continu. En aquest punt comença el que anomenem audibilitat o sensació. La diferència de nivells d'intensitat entre els dos llindars fou de 5 dBs. en aquest punt²¹.

L'oïda humana presenta un llindar absolut de sensació diferent per a cada freqüència. La representació gràfica sobre un eix de coordenades Freqüència-Intensitat dibuixa una corba amb un màxim d'audibilitat a baixa intensitat entre els 200 Hz. i els 5600 Hz., on arriben a ser audibles sons de -5 dB SPL. El llindar de sensació també es troba al voltant de 20 Hz i dels 16.000 Hz en Freqüència i dels 140 dB SPL en Intensitat. Aquests límits conformen el que s'anomena l'àrea d'audició o camp auditiu humà i la representació en l'eix de coordenades es coneix com la gràfica de Wegel. (Gerber, 1974b; Moore, 1982; O'Shaughnessy, 1983; Pavlovic i Studebaker, 1984, citat a O'Shaughnessy, 1987; entre d'altres.)

Els atributs físics de l'ona acústica, com ara la intensitat, són objectius i es poden mesurar mitjançant unitats d'energia o pressió. En canvi, la sensació auditiva es defineix com una apreciació subjectiva d'aquesta energia o intensitat, amb les magnituds de la qual no sembla que existeixi una relació de proporcionalitat.

En els treballs de Weber (1846) es va establir la llei que porta el seu nom i que bàsicament adverteix que notem els canvis relatius de la intensitat de l'estímul i no els absoluts, i, a més, qualsevol estímul cal que sigui augmentat en una proporció constant de la seva magnitud per tal que es noti un canvi de sensació. També estableix que el creixement de l'estímul en proporció geomètrica determina un augment de la sensació en proporció aritmètica (la intensitat de la sensació és una funció logarítmica de la intensitat de l'estímul)²².

A partir d'aquest principi, i d'altres estudis posteriors (Fechner, 1860; Sabine, 1910; Fletcher & Munson, 1933; Stevens, 1951, 1961; etc.) es va ajustar i generalitzar la llei de Weber. Com a conseqüència del coneixement de les lleis de la sensació, es va poder descriure i predir la relació entre la intensitat de l'estímul acústic i la seva sonoritat o sensació en tota l'àrea d'audició. (Vegeu els treballs de Robinson i Dadson, 1956).

21 .-Des d'un punt de vista neurosensorial, els receptors del sistema auditiu s'activen quan l'estímul arriba a un cert nivell, anomenat potencial receptor. Aquest potencial és gradual, és a dir, el valor del potencial és proporcional a la intensitat de l'estimulació sensorial. Si aquest potencial gradual arriba a un valor determinat, anomenat el valor del llindar, es propagarà un potencial d'acció a través de les fibres nervioses fins al sistema nerviós central. Aquests potencials d'acció estan regits per una llei de "tot o res", és a dir, s'activen en superar-se el llindar o no s'activen gens. (Schwartz, 1986). Malgrat aquest comportament del sistema sensorial, el subjecte mai no té aquesta sensació de trencament entre un nivell d'estimulació i un altre. Aquests potencials són totalment involuntaris i apareixen, normalment, en estats d'inconsciència com ara la son.

22 .-D'aquí que s'utilitzin unitats logarítmiques com el decibel (per a la intensitat) o el mel (per a l'altura tonal).

1.5.2.- La mesura del llindar d'audibilitat

Les mesures fonamentals i més importants que s'han realitzat al llarg de la història de la psicoacústica han estat la determinació dels llindars auditius. Les dades que han proporcionat aquests estudis són numèriques i revelen el nombre de respostes encertades o equivocades que un subjecte ha donat davant d'un estímul i d'una tasca enfocada cap a la possible distinció entre dos nivells de sensació diferents. Recordem que les dades s'expressen sobre les propietats de l'estímul, més que no pas del subjecte.

La unitat subjectiva que van utilitzar Weber i Fechner i els seus seguidors per expressar aquesta distinció entre nivells o magnituds diferents fou el que s'anomena "diferència justament perceptible" (*just-noticeable difference* [JND] que és sinònim del terme *difference limen* [DL]) (Flanagan, 1957; Gerber, 1974b).

Per tal de determinar aquestes petites diferències d'apreciació entre dos sons que difereixen en una sola característica, es dissenyaren uns experiments en què els subjectes havien de decidir si els dos sons successius eren iguals o diferents. És el que s'anomena procediment AX, on X és el so que va canviant. Altres variants d'aquesta tècnica són les conegudes com ABX o AXB, amb les quals es pregunta al subjecte si X és semblant (o igual) a A o a B²³.

Es considera que AX és la tècnica més comuna i, davant de dos senyals diferents, produeix un percentatge més elevat de respostes identificant-los com a "senyals diferents" que com a "senyals idèntics". A més, a mesura que es milloren les condicions acústiques del senyal-estímul, també va augmentant el nombre de respostes que el consideren com "senyals diferents". En el moment que aquestes respostes arriben al 75% es considera que s'ha assolit la diferència justament perceptible, és a dir el llindar. El mateix plantejament es reproduïx en les altres tècniques.

El llindar d'audibilitat, i per raons d'aplicació en medicina, s'ha establert en uns nivells concrets, a partir de l'experimentació amb molts subjectes d'audició normal. És el que s'anomena la normo-acúsia²⁴. Evidentment aquest llindar s'estableix en termes absoluts, com un model de comparació per poder jutjar la normalitat dels subjectes. Aquest llindar rep el nom de llindar absolut (*threshold*) (Gerber, 1974b; Moore, 1982). Això pressuposa que hi haurà un llindar relatiu (anomenat *limen*). Efectivament, segons Gerber (1974b) són més interessants els estudis comparatius entre sons perquè permeten veure comportaments diferents entre estímuls diferents. És el que s'anomena discriminació auditiva. Segons aquest autor, quan es treballa sobre el llindar absolut, sempre s'estudia la comparació del so amb el silenci, mentre que amb el llindar relatiu s'estableix una relació més àmplia: amb la sonoritat, amb la variable temps, amb el to, etc.

A partir dels llindars relatius s'han establert, per exemple, les relacions d'igual nivell de

23.- Sobre els tipus de prova emprats en detecció de la parla es pot consultar la revisió que fa Pastore (1981) on presenta un bon nombre de mètodes de treball (S/N, AX, ABX, AXB, 2AFC,...). Per a la intel·ligibilitat de la parla vegeu les revisions sobre els tests disponibles que en fan Gerber (1974), Tillman (1977) i Mackie *et al* (1987). En comprensió, McAllister i Dufberg (1989) presenten els tests anomenats Threshold, Ramp i Helen.

24.- Per a l'aplicació en audiometria clínica, s'usa un nivell d'audició establert en les normes ISO 389-1975

sonoritat entre sons de diferents freqüències i diferents intensitats (Robinson & Dadson, 1956)²⁵. La seva unitat de mesura és el fon (Gerber & Bauer, 1974) i les corbes que assenyalen el mateix grau de sonoritat al llarg del marge freqüencial s'anomenen "isofòniques".

Un altre exemple fruit dels treballs sobre aquests líndars relatius és la relació entre freqüència i altura tonal amb què es percep un so o una sèrie de sons. En aquest cas la unitat de mesura és el mel. (Stevens & Davis, 1966)

1.6.- Conclusió

Els processos d'integració de la parla que duu a terme l'oient passen, en primer lloc, pel control dels nivells d'audició, tant si els límits venen determinats per causes externes al subjecte (com ara la naturalesa i les característiques del senyal) com si provenen dels condicionants del propi oient (capacitat auditiva, coneixement de la llengua, líndars auditius per a la percepció de la parla,...). D'acord amb aquestes circumstàncies, tot disseny experimental sobre percepció i comprensió de la parla haurà de tenir-les present per tal que no resultin uns factors distorsionadors dels resultats de la prova.

25.- Normalitzada segons la norma ISO/R 226-1961, ISO/R 131, ISO/R 357 i UNE 74-003-78 i 74-010-76. En aquestes normes es recullen les diferents corbes d'isofonia i el procediment per al seu càlcul.

2.- LA PERCEPCIÓ I LA COMPRESIÓ DE LA PARLA

2.1.-Introducció

En els apartats anteriors hem parlat pròpiament de "la sensació auditiva", i l'hem tractada com la resposta dels receptors auditius als estímuls acústics. Hem tingut present, també, les característiques i les circumstàncies en les quals es produeixen els estímuls i en les quals actuen els dits receptors. En conjunt, hem tractat del caràcter físic i fisiològic del sistema d'audició humana. A partir d'ara, i seguint el camí de la integració de la parla, caldrà focalitzar l'atenció en diversos estadis superiors que fan que allò que arriba al subjecte com una ona acústica, pugui acabar esdevenint una imatge amb significat dins de la seva ment. Es el que anomenarem, genèricament, percepció.

Després de l'audició, els òrgans responsables de la percepció i de la integració de la parla es concentren en el Sistema Nerviós Central (SNC). A partir d'aquest nivell es decideix si la informació que arriba de l'activitat auditiva és important o si es pot rebutjar. En aquest sentit, Forgas (1972) propugna que el cervell intervé en la percepció dels estímuls amb dues funcions, que són la recepció i la selecció, i planteja l'existència de dos sistemes anatòmico-fisiològics correlatius:

a.- L'activació, que es concentra en el tronc cerebral i el cervell mig del subcòrtex, s'estableix a partir d'un nivell d'excitació nerviosa superior a un nivell mínim, anomenat llindar perceptiu, i permet que es comenci a tractar la informació que arriba des de les etapes externes.

b.- La selecció, exercida pel sistema d'activació reticular, és una funció encarregada de distribuir cadascun dels impulsos sensitius que arriben de l'exterior cap als "centres de relleu", els quals reparteixen la informació a les àrees específiques per tal que sigui tractada segons l'especialització de cadascuna d'elles i després sigui enviada als centres de resposta.

Aquestes dues funcions estan molt més disseminades dins del Sistema Nerviós Central (SNC) i, fins a l'actualitat, no se'n coneix una localització tan definida com s'esdevé amb els òrgans de l'audició. Tampoc no en podem predir un funcionament tan exacte, ja que l'*input* auditiu assolit en el primer estadi és utilitzat en uns altres nivells d'anàlisi, cadascun dels quals es presenta amb una major complexitat.

De fet, fins fa poc, el funcionament del cervell s'ha vist com un tot conjuntat que s'identifica amb la tasca general de l'individu, com si fos una caixa tancada. L'aproximació al seu estudi funcional, des del punt de vista de la fisiologia de la recepció-interpretació dels estímuls acústics, s'havia hagut de realitzar mitjançant mètodes indirectes. L'aplicació d'aquests mètodes comportava, per exemple, que les àrees del cervell que participen en l'acte comunicatiu s'estudiessin a partir de l'observació de les reaccions i respostes a estímuls

presentats a subjectes sans o a partir de l'observació d'algunes reaccions produïdes per malalts afectats en els centres cerebrals de la parla, a causa de lesions d'origen traumàtic o patològic ¹.

Actualment, l'estudi de la percepció de la parla s'aborda, si més no, des de dos punts de vista: l'un des de la neuro-anatomia i la neuro-fisiologia, i l'altre des de la investigació psicològica del llenguatge i la psicolingüística. Es probable que, a la llarga, els dos hagin d'anar units per poder arribar a extreure conclusions ben fonamentades (Crick, 1980), però des d'ara ja podem puntualitzar que el present treball es troba alineat amb el segon enfocament, en una vessant més pròxima a la lingüística que a la psicologia, i només ens referirem al funcionament de les àrees cerebrals de llenguatge quan sigui necessari ².

Malauradament, per poder analitzar l'activitat perceptiva que possibilita que l'oient interpreti el missatge, ens haurem de moure sobre un àmbit força especulatiu, atès que qualsevol reacció seva a un estímul de parla pot estar influïda per múltiples condicionants comportamentals no controlables.

2.2.-La percepció com a procés psicològic

Per poder continuar parlant dels diferents passos que realitza el subjecte per a la integració de la parla, se'ns fa indispensable l'ús del concepte de "procés psicològic". Per definir-lo ens basem en la noció de procés que expliquen García-Albea (coord.) (1986) i Belinchón *et alt.* (1992). Segons aquests autors cal partir d'un enfocament mentalista o cognitiu pel qual l'activitat lingüística és:

...un sistema de procesos abstractos que operan sobre la información representada en la competencia y en otros sistemas de conocimiento.
(Belinchón *et alt.*, 1992, p.292)

Aquest enfocament pressuposa dues característiques distintives: en primer lloc, l'individu humà és un subjecte actiu que elabora significativament els estímuls del medi, organitzant la seva activitat segons plans i estratègies que controlen i guien el seu comportament. En segon lloc, aquest comportament està organitzat de forma jeràrquica i recursiva.

1.- Lúria (1947) va estudiar uns 400 casos de lesions a l'hemisferi esquerre del cervell, produïdes per ferida de guerra en soldats que havien combatut a la Segona Guerra Mundial, i les conseqüències que comportaven en la capacitat comunicativa de l'individu. La major part d'aquestes lesions produïren afàsies en un grau relatiu a l'alteració produïda en àrees corticals determinades.

2.-Per conèixer aspectes concrets sobre el primer punt de vista es pot consultar una àmplia bibliografia d'entre la qual destaquem Blumstein (1988); Caplan (1987, 1988); Damasio i Damasio (1992); Duhamel i Poncet (1986); Gazzaniga *et alt.* (1979, 1982, 1987); Geschwind (1970, 1972, 1980); Hillyard i Kutas (1983); Kutas i Hillyard (1980); Lenneberg (1967); Lúria (1973, 1976); Manning (1988); Novak *et alt.* (1989); Ojemann, Fried i Lettich (1989); Ostrosky-Solís i Ardila (1994); Pantev *et alt.* (1989); Rugg i Barrett (1987); Schwartz (1976, 1978).

Vegeu, també, l'àmplia exposició sobre la tecnologia i els mètodes utilitzats en aquests treballs de recerca que es recullen a Beagley (ed.) (1979).

A partir d'aquesta perspectiva es pot entendre el procés mental com

...cualquier operación que, aplicada sobre un estado mental que le sirve de entrada (input), produce como resultado (output) un estado mental diferente al anterior. Así pues, según esta definición, cualquier actividad que requiera la mediación de representaciones mentales puede definirse como proceso. (Belinchón et al, 1992, p.292)

En el procés de la percepció de la parla hem d'entendre que, a partir d'un senyal acústic de parla es desencadenen diverses operacions (instruccions) que transformen els estats mentals d'entrada en noves representacions més elaborades i, alhora, més abstractes. Ara bé, aquesta transformació no es produeix de forma instantània, sinó que es requereix un cert temps perquè s'activin les diverses actuacions del procés. (Cutting i Pisoni, 1978)

De forma general, els components d'aquest procés poden ser considerats des de diversos nivells de processament. En cadascun d'aquests nivells, els diferents processos que hi intervenen parteixen de les representacions de sortida dels processos pròxims i elaboren noves representacions que, alhora, serveixen de punt de partida en nous processos del mateix nivell o d'un altre.

Aquesta perspectiva del processament de la parla s'ha desenvolupat en Psicologia des de dues tendències dins de l'enfocament cognitiu: el paradigma computacional o de processament de la informació (Cutting i Pisoni, 1978; García-Albea, 1983, 1986, entre d'altres autors) i el paradigma connexionista (McClelland i Rumelhart, 1986). Aquests dos paradigmes estan abastament resumits i explicats a Belinchón *et al.* (1992).

El present treball segueix el primer paradigma i es desenvolupa tot adaptant el patró del processament de la informació de la parla. A continuació reproduïm un esquema que sintetitza el processament de la informació que es dona durant la integració i la comprensió del senyal de parla.

NIVELLS

CONCEPTUAL

Representació
conceptualProcessos
semàntics

SINTÀCTIC

Representació
sintàcticaProcessos
sintàctics

FONO LòGIC

Representació
fonològicaProcessos
fonològicsFONÈTIC
AUDITIU
ACÚSTICRepresentació
acústico-fonètica

Figura 2.1.- Diagrama de flux que mostra els principals components del procés de comprensió de la parla organitzats en nivells de processament. Els processos apareixen tancats en caixes, mentre que les representacions d'entrada i sortida de cada procés es mostren obertes (Belinchón *et al.*, 1992)

Segons aquest esquema, la integració del senyal de parla es comporta com un flux, dirigit de baix a dalt, a partir del qual es van produint diverses representacions al llarg dels diferents nivells d'integració. En aquest aspecte, cal que considerem alguns postulats dels models interactius (vegeu el capítol següent) que completen eficaçment el model clàssic.

En efecte, no podem pensar que la direcció del procés d'integració de la parla sigui unidireccional de baix cap a dalt, com es mostra a la figura (vegeu Norman, 1969), ni que calgui acabar el procés en un nivell determinat perquè se n'enceti un altre en el nivell següent superior, sinó que el processament de la recepció de la parla integra múltiples processos que es troben interconnectats en diverses direccions, sovint d'anada i tornada, que interactuen de forma simultània (Rumelhart, 1989).

Per altra banda, entre alguns d'aquests nivells es produeix la intervenció de diversos tipus de memòria que permeten accedir de nou a la informació processada anteriorment, a fi de reestructurar-la, reprendre-la i reincorporar-la a nivells superiors. (Cutting i Pisoni, 1978). Tot això, però, serà desenvolupat en pròxims apartats.

A continuació ens centrarem en un seguit d'operacions que configuren una part important del procés d'integració de la parla. És el que s'anomena genèricament, **percepció**.

2.3.-La percepció de la parla

Abans de continuar presentant els processos de percepció caldria fer una separació entre el procés perceptiu i el procés auditiu. El concepte d'audició està relacionat, com ja s'ha vist, amb qualsevol tipus d'estímul acústic que provoca un acte de resposta en les terminacions nervioses dels òrgans sensorials externs. La percepció que aquí ens interessa, per contra, implica una resposta a un tipus determinat d'aquests estímuls acústics: els sons de la parla. La resposta a aquest tipus d'estímul és d'ordre intern i requereix de la consciència de l'individu, la qual no era necessària en l'audició.

A partir d'aquest punt és quan es poden començar a aplicar els criteris lingüístics en el processament de la parla i quan hom pot començar a parlar de processos perceptius del llenguatge. De fet l'audició informa mitjançant sensacions auditives, però és l'individu qui percep i interpreta. Per tant, la percepció no depèn tant del tipus d'estímul, com del processament d'aquest estímul, començant en l'oïda i acabant en el còrtex cerebral.

Per poder acotar el concepte de percepció s'ha de tenir en compte tot el que s'ha dit abans sobre els processos mentals i llurs elements interconnectats. Aquests elements actuen conjuntament amb una finalitat única: servir a la comprensió del missatge que duu codificat el senyal acústic. En aquest sentit, més que de conjunt d'elements perceptius cal parlar de sistema perceptiu.

Durant aquest procés d'integració, tot el sistema perceptiu es disposa en funció de la intencionalitat del subjecte que percep; tant si la tasca rau en la comprensió del significat d'una elocució, com si el que es pretén és simplement identificar els seus sons. Per aquest motiu s'ha

dit que els processos de la parla (tant en la seva producció com en la seva percepció) presenten unes característiques teleològicament organitzades (Lindblom, 1983a). Cal afegir que algunes d'aquestes característiques actuen per una predisposició del subjecte a distingir entre els sons de la parla i tots els altres sons (no-parla). (Studdert-Kennedy, 1982).

De fet, alguns estudis sobre la resposta coclear davant d'estímuls acústics demostren que algunes terminacions nervioses de l'oïda interna depassen abans el seu llindar d'activació davant de clics (sons simples i curts) que davant de sons complexos com ara la parla. (Delgutte & Kiang, 1984)³

Aquesta selecció dels sons no es dona solament durant el procés físic de la integració de la parla en els òrgans de l'oïda, sinó que també es produeix en els processos interns que s'originen a partir d'aquesta preselecció d'estímuls. En efecte, segons diversos autors, existeix una predisposició biològica a percebre els sons del llenguatge com a diferents dels altres sons (Lenneberg, 1967; Mattingly *et al.*, 1971; Studdert-Kennedy, 1974 i 1982). Fins i tot, diversos estudis neuro-fisiològics han fet veure que la pròpia ment s'organitza des d'abans del naixement segons estructures modulars i funcionals que predisposen a certes adquisicions fonamentals, entre elles el llenguatge (Galaburda *et al.*, 1978⁴; Melher i Dupoux, 1990).

Segons s'ha demostrat reiteradament (Lieberman *et al.* 1967; Studdert-Kennedy, 1974; Berlin i McNeil, 1976; Barbanoj, 1982; Gazzaniga *et al.* 1982; entre altres autors), el cervell humà està organitzat de forma asimètrica per respondre a les funcions del llenguatge. En realitat existeix un predomini de l'oïda dreta sobre l'esquerra, que no depèn del significat d'allò que sent el subjecte, sinó que respon a una especialització de l'hemisferi cerebral esquerre per processar el llenguatge i l'altre hemisferi (el dret) per processar sons que no són pròpiament de parla. No obstant, tal com Darwin (1971)⁵ va demostrar, l'especialització de l'hemisferi esquerre no respon només davant d'un estímul de parla, sinó que és capaç d'analitzar certs trets acústics del senyal de la parla, tals com els moviments ràpids dels formants o l'atac d'intensitat en les consonants sordes.

2.4.- La invariança en el senyal de la parla

S'ha plantejat, també, que han d'existir certs trets invariants en el senyal de manera que puguin ser comparats i identificats, en la ment del receptor, com a elements propis del codi lingüístic utilitzat⁶. No obstant, tant en els treballs de Jakobson i Halle (1956) com, una mica més tard, en els de Liberman i col.laboradors (1967) es comença a insinuar que difícilment hi poden haver trets invariants en el senyal de la parla, ja que les característiques d'aquest senyal són molt variables en funció dels trets individuals del parlant (diferent freqüències fonamentals en funció del tipus de veu, de si és masculina, femenina o d'infants, etc.), de la velocitat de la

3.-Citat a O'Shaughnessy (1987)

4.-Citat a Geschwind (1980)

5.-Citat a Studdert-Kennedy (1974)

6.-Blumstein i Stevens han publicat diversos treballs en els quals intenten demostrar la possibilitat d'aïllar certes característiques invariants del senyal de parla. (Blumstein i Stevens, 1979, 1981; Stevens i Blumstein, 1981). Per a un debat sobre el tema de la invariança i la variabilitat en els processos de la parla, vegeu Perkell i Klatt (eds.) (1986).

parla (amb tots els efectes de coarticulació i, per tant, també de superposició acústica), del *tempo*, de la intensitat, etc.

La parla espontània també presenta una gran quantitat de variacions degudes a cada situació de parla, a la relació amb l'àmbit d'ús o a l'actitud general del parlant (parla acurada o relaxada, forta o dèbil, etc.) (Engstrand i Krull, 1988a i 1988b).⁷

Ara bé, si això és així, ¿com s'identifiquen els trets dins del senyal?, ¿La identificació es realitza mitjançant la comparació dels diferents components del senyal i uns patrons fonètics que es troben en la ment del subjecte?.⁸ La resposta a aquestes preguntes continua sent la clau del problema, sobretot perquè s'han elaborat diversos models perceptius que intenten donar-hi una solució des de punts de vista prou allunyats entre ells (Klatt, 1988). Aquests models es revisen en l'apartat 2.9 d'aquest mateix capítol.

2.5.-Característiques generals de la percepció

En el que sembla que hi ha un acord és sobre algunes de les característiques generals de la percepció. Per començar, cal dir que la percepció no és un procés automàtic, sinó una activitat cognoscitiva, basada en les impressions del sentit auditiu però lligada a un sistema de decisió que relaciona aquestes impressions amb d'altres ja experimentades i estableix si són o no significatives. Així, doncs, es requereix que el subjecte que percep manifesti una predisposició general per a dur a terme el procés. Aquesta implicació més o menys activa del subjecte ha de comportar, com a mínim, una postura d'atenció cap al senyal i a les circumstàncies contextuals que l'envolten.

En efecte, el subjecte s'implica en el procés tot generant expectatives sobre el contingut del senyal que li està arribant. Diversos estudis han demostrat que la naturalesa d'aquestes expectatives pot determinar les característiques de la resposta dins del procés de percepció i de comprensió. Així, un estímul que és esperat amb unes determinades característiques i que es presenta amb unes altres, provoca respostes que estan d'acord amb les expectatives creades inicialment (Schwartz, 1986). Per un altre costat, Keller (1985) il·lustra aquest fenomen amb l'exemple dels subjectes bilingües, que adapten i ajusten les expectatives perceptives i la distinció de les categories lingüístiques segons quina de les dues sigui la llengua que es parla al seu voltant.

Aquesta intervenció activa del subjecte davant dels estímuls perceptius depèn de dos factors que incideixen clarament durant el procés perceptiu. Per un costat, els indicis informatius que es troben en el propi senyal i en el context comunicatiu, i per un altre, la tasca que desenvoluparà el subjecte davant d'aquest senyal.

Si la tasca d'integració del senyal comportés la detecció, classificació i ordenació d'absolutament totes les característiques del senyal, el volum d'informació que s'hauria de processar constantment en la parla contínua seria enorme i la velocitat de processament s'alentiria (Lieberman, 1967). Per tant, cal suposar que existeixen altres procediments paral·lels

7.-Per a una revisió dels factors que condicionen la variació de la parla vegeu Murillo (1981) i Llisterra (1987). Aquest darrer autor presenta una classificació dividida en quatre tipus de factors: individuals, de producció, contextuals i sociolingüístics.

8.-Per a una revisió del tema de la invariança, vegeu Llisterra (1987)

que estalvien de processar tota aquesta informació i imprimeixen una alta velocitat de processament.

Per una banda, el senyal de la parla conté molta informació que és redundant, ja que, per exemple, la percepció d'un determinat fonema està condicionada pels fonemes adjacents, la manifestació fonètica dels quals comporta la presència d'unes determinades transicions en el senyal com a conseqüència de la coarticulació. Aquesta informació dependent del context fonètic és una part fonamental de la percepció (Miller, 1981)

Per una altra banda, el receptor sovint s'anticipa, si més no mentalment, al contingut del missatge que s'està integrant i que encara no ha estat emès en la seva totalitat. Aquesta situació és deguda al coneixement de l'emissor, del context i del coneixement general de la situació.

Finalment, l'aplicació de regles d'eliminació de la redundància i de restauració fonològica del senyal (vegeu, més endavant, el nivell d'anàlisi fonològica), i el coneixement conceptual del referent del missatge també contribueixen a la interpretació del senyal (Lindblom, 1983a).

Tot plegat fa suposar que existeixen dues direccions en els processos d'integració de la parla que funcionen de forma harmònica. Es el que Norman (1969) va anomenar "processos de dalt-a-baix (*top-down*) i de baix-a-dalt (*bottom-up*)" ⁹. Segons aquest autor, el subjecte descodifica els elements del senyal de la parla només fins al punt que no pot extreure'ls del context. En aquest darrer sentit, hi ha proves que els *output* del nivell lèxic, semàntic i sintàctic (nivells superiors) poden afectar la percepció fonètica i fonològica (nivells inferiors). (Studdert-Kennedy, 1976)

2.6.-Selectivitat perceptiva i atenció selectiva

Uns altres processos propis de la percepció són els que tenen a veure amb la selectivitat perceptiva. Evidentment, els estímuls acústics que arriben a l'oïda són molt nombrosos i d'índole molt diversa. Qualsevol acte de parla també es troba immers en aquest maremàgnum acústic. Això es pot constatar en els enregistraments de parla realitzats fora de les condicions de laboratori. Si el procés de percepció de la parla es realitzés de forma indiscriminada, semblantment a com s'esdevé amb un micròfon omnidireccional, la varietat tan gran d'estímuls faria necessaris grans esforços de processament per poder extreure una quantitat petita d'informació. Ultra això, els estímuls diferents de la parla representarien un emmascarament constant.

L'estratègia perceptiva que estalvia esforços té a veure amb l'atenció anomenada selectiva (Berlin - McNeil, 1976). Per mitjà de l'atenció el subjecte és capaç de focalitzar les seves estratègies perceptives cap a uns determinats estímuls de parla, tot rebutjant els altres sons contextuals. De fet, la pròpia natura del senyal de la parla facilita aquesta focalització, en tant que ja existeix una predisposició del subjecte a respondre a aquest tipus d'estímul per

9.-Sobre l'existència dels processos *top-down* i *bottom-up*, vegeu Ohala (1986)

damunt de qualsevol altre estímul acústic d'intensitat similar. Segons alguns autors¹⁰, aquesta predisposició es basa en una determinada manera de processar el senyal de la parla (*speech mode*) deguda, sobretot, a la seva estructura fonètica que es relaciona amb les propietats articulatòries amb què ha estat produït el senyal (vegeu l'apartat 2.9.1.2, sobre la teoria motora). De fet, això fa pressuposar l'existència d'un processador per a la percepció de la parla diferent del que s'utilitza en altres sons (Mattingly *et alt.*, 1971)

Aquesta focalització perceptiva també s'esdevé en el cas que la majoria dels senyals acústics contextuals siguin de parla. En aquest cas, les estratègies d'atenció selectiva encara s'exerceixen de forma més intensa, ja que tots els sons tenen característiques similars i només un dels senyals que arriben al subjecte és del seu interès. Així, la selecció que es produeix en el procés perceptiu permet d'acollir un senyal tot menystenint i ignorant els altres. La possibilitat de percebre la parla contínua és més efectiva que la identificació de sons de forma isolada. El fenomen del '*cocktail party*' (vegeu l'apartat 1.4.2.2) posa de manifest una vegada més la importància que tenen, per a la percepció de la parla, les característiques estructurals en què es troba organitzada.

Existeixen diversos estudis sobre l'atenció selectiva que utilitzen una tècnica experimental anomenada 'escolta dicòtica'¹¹. Mitjançant aquesta tècnica s'ha pogut fonamentar l'existència i el funcionament d'aquest tipus de selectivitat perceptiva.

2.7.-Processos en la percepció de la parla

No es coneix encara el nombre exacte d'estadis per al processament de la parla. No obstant, i malgrat que hi hagi diversos punts de vista per explicar el funcionament dels processos de percepció de la parla, sembla que tots els autors accepten el fet que existeixen, almenys, tres nivells d'anàlisi diferents. Seguint Studdert-Kennedy (1974), aquests nivells són els següents:

- | | |
|---------------------|-----------------------------------|
| 1.-Nivell auditiu | (<i>auditory analysis</i>). |
| 2.-Nivell fonètic | (<i>phonetic analysis</i>). |
| 3.-Nivell fonològic | (<i>phonological analysis</i>). |

Aquest autor presenta un quart estadi (que per a nosaltres ja no és estrictament perceptiu, sinó propi del procés de la comprensió) que és l'estadi lèxic, sintàctic i semàntic. Totes aquestes fases formen una jerarquia conceptual tot i que poden ser successives i simultànies. Per altra banda, aquests estadis no tenen una importància fixa, sinó variable en funció de les condicions de l'oient.

10 .-Lieberman *et alt.*(1967), Liberman i Mattingly (1985). En l'actualitat és viva una forta polèmica en la qual es discuteix aquesta predisposició de l'oient a percebre la parla (fonamentada en la imatge de les articulacions que han produït els sons -*gestures*- i en la imatge de la seva pròpia producció-*distals events*-), davant dels que argumenten que el senyal de la parla és diferent dels altres senyals acústics i que, en tot cas existeixen uns mecanismes perceptius específics per a processar aquesta informació acústica. Per a resseguir aquesta polèmica vegeu Diehl *et alt.*(1996), Fowler (1991, 1996), Lindblom (1996), Ohala (1996), O'Shaughnessy (1996), Remez (1996) i Stevens (1996). Sobre el model del *Direct-Realism* de Fowler, vegeu l'apartat 2.9.1.3.

11 .-Berlin i McNeil (1976) presenten aquest paradigma experimental i fan una llarga llista de treballs en els quals se'n fa ús.

2.7.1.- Nivell auditiu

Del primer nivell d'anàlisi ja n'hem parlat en l'apartat 1.5 del capítol 1. No obstant hi ha autors que el consideren com un estadi propi del procés perceptual (Pisoni i Luce, 1987). Concretament aquest dos autors hi suggereixen l'existència d'una subdivisió, de manera que les funcions que desenvolupen aquests subestadis són:

- 1.-Anàlisi auditiva perifèrica
- 2.-Anàlisi auditiva central

Tot i que tots dos són nivells prefonètics, i per tant no lingüístics, Carlson i Granström (eds.) (1982) han exposat la necessitat de considerar-los connectats al procés de percepció, per tal com aquest sistema auditiu perifèric utilitza certes estratègies per codificar el senyal de parla.

2.7.1.1.-Anàlisi auditiva perifèrica

Searle *et alt* (1979)¹² van trobar certes proves que el sistema auditiu perifèric humà constantment analitza els estímuls acústics amb una resolució aproximada d'un terç d'octava. Aquesta capacitat de filtrar els senyals acústics no serveix per si sola per identificar sons de la parla, però permet aïllar certs components bàsics que possibiliten la tasca dels analitzadors dels estadis superiors. Amb els seus experiments, Searle *et alt.* van demostrar el grau en què el reconeixement artificial de la parla es pot millorar una vegada s'apliquen certes transformacions psico-físiques en les representacions perifèriques de la parla.

Pastore (1981), per altra banda, planteja que l'anàlisi acústica dels estímuls de la parla probablement pot explicar molt millor la percepció de la parla que el que suposen els investigadors de la parla i molt menys del que s'assumeix en Psicoacústica. En general està acceptat que el senyal de la parla està subjecte als mateixos principis generals de l'anàlisi d'altres senyals acústics i només després d'una transformació, la informació obtinguda en aquest nivell d'anàlisi pot ser integrada en els nivells superiors (Eimas *et alt.*, 1985).

2.7.1.2.-Anàlisi auditiva central

A partir de la informació processada en l'anàlisi perifèrica, el sistema auditiu és l'encarregat d'analitzar-la i d'emmagatzemar els resultats en una memòria auditiva. Aquesta memòria s'anomena ecoica (també *memòria acústica precategòria*) i té una durada molt curta, de pocs segons. Els experiments suggereixen que l'etapa auditiva és selectiva, és a dir, fixa l'anàlisi de la parla i no la dels altres sons. Evidentment, aquest model comporta necessàriament l'escolta selectiva (Mattingly *et alt.*, 1971).

12 .-Citat a Pisoni i Luce (1987)

La funció més important d'aquest estadi consisteix en analitzar la parla per tal d'extreure'n les seves característiques acústiques. S'identifiquen sota aquesta designació les estructures espectrals, la freqüència fonamental, la intensitat mitjana, la durada del senyal, l'amplitud dels atacs i decaïments, etc. Normalment, diverses característiques acústiques poden configurar un tret fonètic simple. (Eimas i Corbit, 1973)

Liberman *et alt* (1967) van posar al descobert l'existència de les característiques acústiques dels sons de la parla en front dels patrons ¹³ acústics dels fonemes. Així, no tots els fonemes /d/ són iguals, ni totes les /a/ ho són. Segons els sons adjacents, cada fonema tindrà una realització diferent, que serà identificada segons les seves característiques acústiques. En el cas de la /a/, depenent del context, adoptarà una 'forma' espectrogràfica concreta en les seves transicions formàntiques, en relació amb el lloc d'articulació en què s'han produït els segments fonètics adjacents. D'aquestes característiques es diu que són dependents del context. (Per a una revisió d'aquest punt, vegeu Cole i Scott, 1974b)

Existeixen, també, segons Cole i Scott, (1974 a, b), certes consonants les característiques acústiques de les quals es mostren inalterables siguin quins siguin els sons que les precedeixen o segueixen. Les característiques d'aquestes consonants es manifesten inalterables al context acústic on es troben. Concretament són les consonants estridents i es diu que les seves característiques són independents del context.

Altres treballs sobre la dependència i independència del context fonètic (Miller, 1981) han arribat a la conclusió que la percepció del sistema de característiques acústiques no opera sempre depenent del context, sinó que utilitza procediments diversos. Així, durant els primers estadis del processament del senyal de parla, es produeix una anàlisi de les característiques dels sons des d'una vessant acústica, la qual proporciona informació sobre les distincions de tipus fonètic. No obstant, en estadis superiors es concreta una interrelació entre les característiques pròpies dels diferents sons (per exemple, l'anàlisi de la durada de les transicions determina la manera d'articulació mentre que la forma com es desenvolupen les transicions determina el lloc d'articulació).

Alguns autors com Eimas *et alt*. (1985) han aconseguit aïllar els efectes de dependència d'algunes d'aquestes característiques, tot establint que certes interaccions són de tipus asimètric, d'altres interaccions són mútues i d'altres, mútues i asimètriques. També han demostrat que la relació de dependència que existeix durant el processament de dues característiques depèn directament de la combinació concreta d'aquelles característiques.

En resum, l'anàlisi auditiva que es produeix en aquest primer estadi del processament de la parla proporciona les característiques acústiques del senyal. A partir d'elles s'esdevindran les representacions auditives que, un cop emmagatzemades en la memòria, seran utilitzades en la posterior classificació fonètica.

13 .-Liberman *et alt*. (1967) van indicar que en un sistema de percepció de la parla que estigués basat en la suma de sons (un darrera un altre), en algunes situacions es requeriria la percepció de 30 fonemes per segon, la qual cosa no podria ser discriminada per l'oïda. Per altra banda, la percepció per patrons acústics a la manera dels caràcters tipogràfics comportaria la disponibilitat constant de 40 patrons (en una llengua com l'anglès), els quals no podrien abastar més de 50 ms. Aquesta condició no permetria identificar alguns dels sons de la parla. Això suggereix l'existència d'un nombre més reduït de patrons i d'unes característiques més discretes.

2.7.2.-Nivell fonètic

Aquest és el primer nivell del procés d'integració de la parla en què s'acompleixen tasques de tipus lingüístic. L'anàlisi que es realitza en aquest nivell consisteix, per una banda, en identificar els trets fonètics distintius de la parla a partir dels indicis acústics (característiques) que s'han obtingut en les anàlisis de l'estadi anterior i, per una altra, en establir un ordre correcte en l'aparició d'aquests trets. (Miller *et alt.* 1983)

Els trets distintius poden ser descrits en termes d'un conjunt relativament petit de codis lingüístics que serveixen per diferenciar uns segments fonètics d'uns altres. Aquests codis es sistematitzen en una classificació binària de les propietats dels sons de la parla (Jakobson, Fant i Halle, 1952; Jakobson i Halle, 1956; Chomsky i Halle, 1968). Aquests trets tenen dues particularitats: per un costat, es correlacionen amb uns atributs de tipus acústic i per un altre costat, els seus antecedents són de tipus articulatori (Stevens, 1980; Miller *et alt.*, 1983) juntament amb els correlats perceptius corresponents (Jakobson i Halle, 1956)¹⁴

Després de diversos treballs sobre la confusió perceptual (Miller i Nicely, 1955), la percepció d'estímuls amb característiques acústiques conflictives (Massaro i Cohen, 1976, citat a O'Shaughnessy, 1987) i la percepció consonàntica categorial (Lisker i Abramson, 1970), es començà a suggerir la possibilitat de l'existència d'uns detectors de trets distintius en aquest nivell fonètic d'anàlisi, ja que les situacions de confusió perceptiva es correlacionen força amb aquests trets. Aquests detectors funcionarien com estructures especialitzades i serien les responsables d'extreure les característiques fonètiques del senyal de la parla.

Evidentment, l'existència d'aquests detectors de trets es fonamenta sobre el supòsit que el senyal de la parla no es presenta al subjecte com qualsevol altre senyal acústic, sinó que es caracteritza com un senyal que conté una informació codificada de manera diferent als altres sons, i per aquest motiu es fa necessària la presència d'analitzadors específics per poder descodificar aquesta informació (Eimas *et alt.*, 1973, 1985)¹⁵

A partir d'experiments que havien realitzat alguns investigadors sobre els detectors de trets en la visió, Eimas i Corbit (1973) van indicar que aquests detectors també existien en la percepció auditiva i van basar la seva existència en dos tipus de resultats experimentals: a) en primer lloc, en els estudis de la discriminació entre parla i no-parla efectuada per nadons d'entre un i quatre mesos; la troballa fou que distingeixen els dos tipus de sons de manera similar a com ho fem els adults i que distingeixen els sons entre diferents categories molt millor que dins de la pròpia categoria. (Colombo & Bundy, 1983; citats a Mehler & Dupoux, 1990); b) en segon lloc, en els estudis basats en "l'adaptació selectiva a una determinada característica fonètica"; aquesta adaptació consisteix a "fatigar" el suposat detector de trets (o unitat neural), a fi que no sigui tan eficaç davant d'un estímul que contingui aquella característica (Eimas i Corbit, 1973; Eimas, Cooper i Corbit, 1973). Els resultats de l'experiment van demostrar la bona sensibilitat d'aquests detectors a una determinada estimulació selectiva i per una altra banda van reforçar el model de percepció categorial (Miller, 1981; Pastore, 1981). Altres experiments sobre diverses característiques fonètiques també van arribar a resultats similars (Cooper 1974, 1975)

14 .-En l'apartat 1.4.1 del capítol anterior ja s'ha introduït aquest aspecte.

15 .-Sobre les característiques del senyal vegeu l'apartat 1.4 del capítol anterior i el 2.6 d'aquest..

La tècnica de l'adaptació selectiva, però, va obrir un controvèrsia, encara no resolta, entre els que pensaven que l'efecte d'adaptació provenia del cansament dels detectors en un nivell fonètic (Eimas i Corbit, 1973; Sawusch i Jusczyk¹⁶, 1981) i els que pensaven que podria produir-se com una forma de contrast o de resposta auditiva similar a la que es produeix davant de qualsevol altre senyal acústic en un nivell auditiu (Diehl *et alt.*, 1978, 1980).

Malgrat que els resultats obtinguts per Eimas i Corbit apuntaven inicialment en la direcció que els analitzadors de trets havien de ser de tipus lingüístic, en tant que només responien a estímuls de parla similars als fonemes i, per tant, se situaven en un estadi fonètic del procés perceptiu, alguns autors van aportar dades per demostrar el contrari.

Per una banda, Sawusch i Jusczyk (1981) van demostrar que l'adaptació seguia les característiques espectrals del so que es repetia (so adaptador) i no s'efectuava, en cap cas, una identificació fonètica. Aquests resultats situaven els efectes obtinguts de l'adaptació selectiva en el nivell de l'anàlisi auditiu.

Per una altra, McGurk i MacDonald (1976) i Roberts i Summerfield (1981)¹⁷ van realitzar uns experiments basats en l'adaptació selectiva, però hi van afegir un estímulo visual en el qual un parlant movia els llavis articulant una síl·laba diferent de les dues que s'estaven examinant. Els resultats, coneguts com l'*efecte McGurk*, van ser contraris als que es podien esperar en l'escolta selectiva. La conclusió a què s'arribà va refermar amb contundència l'argument que els detectors de trets no són específics de la parla sinó que responen a qualsevol estímulo complex i, per tant, es troben en un estadi auditiu i no fonètic.

En una altra vessant de la mateixa anàlisi fonètica, i com una extensió del paper que juguen els trets distintius en la percepció de la parla, s'han efectuat múltiples experiments que suggereixen que la percepció dels estímuls es realitza d'una forma categòrica, és a dir, que l'habilitat de discriminar dos estímuls diferents de parla depèn només de la capacitat de classificar-los com lingüísticament diferents.

Malgrat que aquesta capacitat de situar els sons presentats en categories diferents no és innata a l'individu, sinó que aquest l'incorpora a mesura que va aprenent una llengua, sembla que es pot demostrar una certa predisposició a desenvolupar habilitats de percepció lingüística des d'una etapa molt precoç de la vida. En aquest sentit Kulh (1979, 1986) ha demostrat en els seus estudis que els nens menors de 6 mesos són capaços de distingir entre sons de la mateixa categoria fonètica i que tenen l'habilitat de distingir-los malgrat els canvis en les dimensions del tracte vocal dels parlants i malgrat els canvis d'entonació que produeix cada parlant. Fins i tot són capaços de reconèixer diferències entre les diverses realitzacions d'una categoria fonètica determinada.

Per una altra banda, sembla clar que hom no pot usar coneixements lingüístics per categoritzar els sons que no tenen correspondència fonètica (en el cas dels sons de no-parla). Tampoc no ho pot fer si els sons es presenten en una llengua no coneguda ¹⁸, ja que no pot

16.-Citats a Pisoni i Luce (1987)

17.-Citats a Miller *et alt.* (1983)

18.-Per exemple, si el senyal de parla que arriba no pot proporcionar un contingut semàntic i ha estat generat en un idioma del qual el subjecte en desconeix els components fonètics, la seva atenció es centrarà a esbrinar les característiques fonètiques del senyal (nivell fonètic) a fi de representar, si més no, una imatge fonètica el

trobar bases auditives per a la discriminació entre els sons i hauria de poder comptar, doncs, amb assignacions categorials. (Studdert-Kennedy, 1974; 1976; Liberman i Mattingly, 1985)

S'han donat diverses proves sobre l'existència d'aquest procés perceptiu de tipus categorial. Una d'elles es constata a partir del fet que, des d'un punt de vista acústic, els sons de la parla poden variar considerablement en les seves propietats, no solament quan són produïts per parlants diferents, sinó fins i tot quan un mateix parlant els produeix en moments diferents de l'elocució. Tals sons, doncs, podrien ser percebuts com diferents en una percepció amb bases estrictament auditives. No obstant, aquests sons són classificats com idèntics pels individus perceptors. Es coneguda l'experiència realitzada en aquest sentit per part de Lisker i Abramson (1970), en la qual van demostrar la categorització dels sons [p, b] al llarg d'un continuïum en el qual es variava la durada del *Voice Onset Time* (VOT) de les consonants oclusives. En VOT curts o llargs, els subjectes determinaven la sonoritat dels oclusius fins a una precisió de 10 ms. prop de la frontera del canvi de la categoria sord/sonor, mentre canvis en la durada del VOT molt superiors no provocaven un canvi en la categoria si es trobaven allunyats d'aquesta frontera.

Pel que fa al comportament d'aquestes categories, i aplicant la tècnica experimental de l'adaptació selectiva, Miller *et al* (1983) es van adonar que, a mesura que l'adaptació es movia cap al límit de la categoria, hi havia un increment inicial en l'adaptació, seguida per una subsegüent reducció. Aquesta constatació els féu arribar a la conclusió que, per a un contrast concret entre dues categories, existeix un límit en l'interval d'un estímul que es considera com a rellevant. Per altra banda, un estímul, considerat com a pertanyent a una categoria concreta, varia la seva efectivitat com a membre d'aquella categoria al llarg del continuïum, cosa que fa suposar l'existència d'estructures internes dins de les categories. Finalment, aquestes estructures internes proporcionen les bases per establir correctament les diferències entre categories en el nivell fonètic.

Un altre procediment experimental que va donar suport a l'existència de la percepció categorial es coneix com paradigma AX i ABX o AXB. Amb aquestes proves es presentava als subjectes un so (en el cas de AX) o dos (en els casos de ABX i AXB) i se'ls preguntava si el so X era igual a A o no i si era igual a B o no. També es podia plantejar si el so X era igual a A o a B. No obstant, Pastore (1981) ha posat en dubte l'efectivitat d'aquest mètode pel fet que la tasca de categoritzar el so X es recolza en l'activació de la memòria de curt termini (degut a l'interval de temps entre els estímuls), la qual podria ser la responsable de la classificació del so més que no el propi procés d'anàlisi fonètica. Per altra banda, en el curs de l'experiència es constata que el subjecte desestima la seva participació quan la tasca de discriminació és difícil i la sobreestima quan la tasca és fàcil.

Malgrat tot el que s'ha exposat fins aquí, la percepció categorial com una capacitat lingüística exclusiva dels humans no està, ni de bon tros totalment acceptada¹⁹. Tampoc no està resolta la qüestió de si la percepció categorial és una característica general o no de la percepció auditiva. Aquestes qüestions sobre la percepció categorial i la seva discussió resten exposades a Belinchón *et al* (1992). Per aprofundir en d'altres aspectes sobre la percepció de contrastos fonètics, vegeu Pastore (1981).

més pròxima possible a les imatges mentals dels sons de l'idioma propi (fonemes). De fet, no pot anar més enllà en el procés d'integració d'aquest senyal.

19.-Sobre la controvèrsia al voltant del problema de l'inatisme en la detecció dels trets, vegeu la revisió que fa Keating (1988)

2.7.3.-Nivell fonològic

En el tercer nivell d'anàlisi es produeix el pas dels trets distintius emmagatzemats en la memòria categorial (de curt termini) a la identificació dels fonemes corresponents (que és la unitat mínima lingüística). Aquest pas suposa la eliminació de la variabilitat fonètica en favor dels components fonològics, que són invariants. (Studdert-Kennedy, 1976; Pisoni i Luce, 1987).

En aquest nivell s'apliquen les regles fonològiques que determinen si una distribució fonètica concreta es pot integrar com una combinació fonològica de la pròpia llengua o no. El resultat s'emmagatzema en una memòria de curt termini perquè estigui disponible per a processos posteriors. Existeixen dos aspectes que expliquen aquest procediment.

El primer considera que el contingut de la memòria fonètica accedeix al següent estadi superior (nivell fonològic) i allí és avaluat mitjançant les regles fonològiques pròpies de la llengua. A partir del resultat d'aquesta avaluació es produeix l'acceptació o el rebuig d'una determinada cadena fonològica. Si la combinació de fonemes és acceptada, aquesta informació s'emmagatzema en una memòria de curt termini i segueix el camí en el procés d'integració per posar-se al servei de nivells superiors que n'extreuen el contingut semàntic.

Si, per contra, no és acceptada com una combinació pertinent, es posen en funcionament diverses estratègies que permeten reprendre la informació de la memòria de curt termini i, si cal, és restaurada fonològicament (tot reordenant els seus components o restablint i substituint aquells trets distintius que no responen a la combinació fonològica esperada) a fi d'adequar-la a les característiques necessàries del segment de parla perquè es pugui aconseguir, més endavant, la comprensió del missatge. Aquestes estratègies de restauració es produeixen a partir d'una actuació conjunta i simultània dels nivells fonètic i fonològic, entre els quals es porten a terme operacions de retroalimentació i d'ajustament. El nivell fonètic interpreta les característiques acústiques mentre que el fonològic, l'ordre d'aparició dels fonemes. (Studdert-Kennedy, 1974, 1976)

Per exemplificar aquest segon cas, considerem la situació en la qual hom se situa davant d'un senyal de parla sense contingut semàntic, com és el cas dels logotomes. El subjecte té la capacitat de repetir els sons de l'enunciat que li ha arribat (nivell fonètic), ja que ha identificat una cadena fònica que segueix les regles fonotàctiques de la seva pròpia llengua (nivell fonològic). Aquesta interconnexió permet al subjecte de configurar-se una imatge mental d'aquella cadena parlada. Malgrat tot, el procés d'integració no pot sobrepassar aquest nivell fonològic, donat que no existeix cap imatge lèxica en nivells superiors.

Pisoni i Luce (1987), apunten altres aspectes que cal considerar en aquest nivell fonològic, com ara les estructures més grans que el fonema. Segons aquests autors, històricament molts lingüistes van considerar l'*output* del nivell fonològic com una cadena de fonemes ordenats linealment i les síl·labes no jugaven cap funció en aquesta organització fonològica ²⁰. No obstant, a partir d'estudis fets recentment en el camp de la fonologia s'ha

20.-Chomsky i Halle, 1968

indicat l'existència d'una representació jeràrquica de l'estructura interna de la síl.laba. Els seus dos components principals són l'inici de la síl.laba o *onset* (que està constituït opcionalment per una consonant o grup consonàntic) i per una rima o *rhyme* (constituïda per la vocal o diftong -nucli- i les consonants -coda- que també són opcionals). Algunes experiències (Treiman, 1983; Halle, 1985) apunten que aquests constituents de la síl.laba poden ser psicològicament reals i que tant la percepció de la síl.laba com del fonema poden ser processos del nivell d'anàlisi fonològica.²¹

La importància de la síl.laba en aquest estadi rau en la seva condició d'organitzadora de les unitats segmentals. Per tal d'assenyalar l'ordre dels diferents fonemes, estableix una estructura basada en l'oposició "consonant - vocal". Aquesta circumstància permet que el subjecte predigui la situació del nucli sil.làbic durant el procés de segmentació fonològica.

En diversos treballs experimentals (Barry, 1984; Cutler *et al.*, 1986; Cutler i Norris, 1988; Mehler *et al.*, 1981; Pallier *et al.*, 1993; Sebastián-Gallés *et al.*, 1992; Sendlmeier, 1995) s'han buscat proves sobre la segmentació de la parla i els processos de percepció de la parla basats en la unitat sil.làbica. En aquests experiments s'aplicà el paradigma experimental conegut com "detecció de sons", en el qual es demana al subjecte que reconegui un determinat segment en un estímul presentat (paraules aïllades).

Davant d'estímuls que començaven amb la mateixa seqüència de sons, es va observar que el subjecte reconeixia el segment que se li havia assenyalat, significativament més aviat en paraules amb les quals coincidia amb la síl.laba inicial que amb les que no hi coincidia. Així la paraula francesa PALMIER es reconeixia abans que PALAIS quan el segment que calia reconèixer era PAL i a l'inrevés quan era PA. Això mateix, però, no va succeir en experiments similars amb estímuls en llengua anglesa. En aquesta llengua l'estructura CV era sempre la que es reconeixia abans. Les conclusions a què s'arribà foren de dos tipus. En primer lloc, els subjectes només utilitzen estratègies de segmentació sil.lábiques quan la seva llengua és fàcilment segmentable sil.làbicament. En segon lloc, les estratègies de segmentació de la llengua pròpia dels subjectes són ràpidament utilitzables davant d'estímuls que no es corresponen amb la llengua pròpia. Per tant, els patrons de segmentació no responen a la llengua que s'escolta sinó a les estratègies apreses per a segmentar la llengua pròpia.

Per altra banda, Cutler i els seus col.laboradors²² han demostrat el paper que juguen els components prosòdics de la parla. En els seus treballs han arribat a la conclusió que l'oient és capaç de predir el punt exacte on es trobarà una determinada paraula tònica dins de l'elocució. La síl.laba accentuada, doncs, esdevé un indicador important per determinar el punt de màxima informació. És evident que aquest enfocament, anomenat *metrical segmentation strategy*, contravé els models que es basen en el desenvolupament temporal de la parla i que troben el punt d'activació del lèxic un cop s'han comparat les paraules candidates des de l'esquerra cap a la dreta, sense comptar amb les característiques accentuals (vegeu la *Cohort Theory*, en l'apartat 2.13.3).

21 .-Clark i Clark (1977) suggereixen que la identificació d'un segment mínim de parla pressuposa l'anàlisi de tota la síl.laba i no només del segment aïllat, ja que tot segment de parla depèn dels sons que l'envolten en el context. Els sons de la parla no s'identificarien de forma seqüencial, sinó conjunta.

22 .-Cutler i Foss (1977), Cutler i Darwin (1981), Cutler i Norris (1988), Cutler i Butterfield (1990), Conine *et al.* (1987)

2.8.-La memòria en la percepció de la parla

No es podria descriure cap sistema de percepció de la parla si només es tingués en compte un procés de transformació de la informació entre els diferents nivells d'anàlisi i no es considerés el paper fonamental que juguen diversos tipus de memòria intermitja entre aquests estadis. Segons Cutting i Pisoni (1978), hi ha molts motius que avalen aquesta afirmació, tots ells basats en la naturalesa mateixa del procés perceptiu. Les característiques principals que requereixen la intervenció d'operacions de la memòria es podrien resumir com segueix:

-La percepció és un procés que es desenvolupa molt ràpidament, tot i que s'activen diversos estadis d'anàlisi.

-La parla és un senyal dinàmic en el qual les interrelacions entre les parts de l'espectre acústic són constantment variants i en varia també constantment la seva velocitat.

-Alguns segments del senyal poden ser integrats més fàcilment que d'altres

Malgrat tot, aquestes variacions del senyal de la parla poden ser controlades i processades. L'única manera d'aconseguir descodificar el senyal en temps real és mitjançant l'ús continuat d'una sèrie de memòries que constantment s'actualitzen i es reutilitzen per a cada nova informació, i tot plegat a una velocitat força elevada. Les memòries que afecten directament els processos perceptius -auditiva i fonètica- s'exposen a continuació.

2.8.1.-Memòria auditiva²³

La durada de la memòria auditiva és de pocs segons, just el temps suficient perquè la informació sobre les característiques acústiques que s'han obtingut en el nivell auditiu puguin ser traspassades al nivell d'anàlisi següent. Si no s'extreu ràpidament aquesta informació el contingut de la memòria desapareix definitivament.

La finalitat d'aquesta memòria no està enfocada només a les característiques acústiques de la parla, sinó que els *ítem* que pot emmagatzemar poden ser, també, de tipus sil·làbic, lèxic o, fins i tot petit grups de paraules. La seva capacitat està limitada, segons diferents autors, a 7 unitats (Miller, 1956) o, encara menys, a 5 unitats (Broadbent, 1975)²⁴.

Concretant en el procés de la percepció de la parla, aquesta memòria funciona de forma selectiva, tot distingint la naturalesa dels diferents sons que arriben al nivell auditiu. Així, doncs, afina molt en l'anàlisi de la parla, i no en l'anàlisi d'altres sons de no-parla, provinents d'una font acústica que no és de l'interès del subjecte (Crowder, 1981).

23.- La memòria auditiva també té d'altres denominacions: memòria immediata, de curt termini (*short-term memory* [STM]), ecoica, precategorial i prefonètica.

24.- Ambdós autors estan citats a Cutting i Pisoni (1978)

Per altra banda, la informació que emmagatzema, que és de tipus precategòria, té a veure amb les característiques acústiques del senyal; és especialment sensible a la informació sobre la manera d'articulació, el punt d'articulació i la sonoritat, i no ho és tant a d'altres aspectes del senyal de la parla.

2.8.2.-Memòria fonètica

Així com la memòria auditiva té una durada molt curta i ràpidament desapareix, la memòria fonètica pot ser consultada durant algun temps. Segons O'Shaughnessy (1987), els resultats d'algunes experiències sobre intel·ligibilitat suggereixen que aquesta memòria té una durada màxima de set síl·labes en el decurs de la parla contínua. Més enllà d'aquesta llargada es fan necessaris els processos de tipus cognitiu i, per tant, una memòria de tipus lèxic o semàntic.

La memòria fonètica, també anomenada memòria tardana o de llarg termini, es fa càrrec de la informació que li proporciona l'estadi fonètic. Aquesta informació no es basa en les característiques acústiques com en el cas de l'auditiva, sinó que a través de l'anàlisi fonètica, es basa en els sons continguts en el senyal de parla que s'han pogut categoritzar. La memòria fonètica emmagatzema aquestes identificacions, tot assignant un nom a cadascun dels segments que conformen el senyal. Ara ja no es distingeixen les diferents realitzacions d'un mateix so, sinó que totes elles tenen assignada una única denominació en la memòria fonètica. Aquesta capacitat de discriminar els sons de la parla i d'assignar-los diferents denominacions en aquesta memòria fonètica depèn, segons Pisoni (1973), de l'aprenentatge i de l'experiència en una llengua.

Segons Crowder (1981), les denominacions dels sons s'emmagatzemen en aquesta memòria com "paquets" de trets distintius, cadascun dels quals és preservat independentment dels altres. D'aquesta manera el subjecte pot tornar enrera i valorar fonèticament un senyal que ha comportat un missatge defectuós o mal interpretat. Així mateix, també li serveix per poder resseguir l'ordre dels diferents sons i les característiques estructurals de l'enunciat (disposició accentual, entonació, etc.).

Es important remarcar que ambdues memòries, l'auditiva i la fonètica, estan totalment integrades entre els diferents estadis de processament del senyal de la parla i totes dues operen seqüencialment

2.9.-Models de percepció de la parla ²⁵

A mesura que s'han anat dissenyant nous models experimentals, paral·lelament s'ha fet necessari d'anar incorporant els resultats que s'anaven obtenint en models generals, a fi d'explicar els processos de la percepció de forma global. Així mateix, calia que aquests models

25 .-Per a una revisió general dels principals models de percepció es pot consultar, entre d'altres, Altmann (1990); Borden i Harris (1980); Clark i Clark (1977); Cutting i Pisoni (1978); Darwin (1987); Goldinger, Pisoni i Luce (1996); Klatt (1986, 1988); Lass (Ed.) (1976, 1979, 1984); Pisoni i Luce (1987, 1988); Stevens (1986)

mantinguessin la coherència a la llum dels resultats experimentals. Malgrat que s'han exposat múltiples possibilitats per explicar el conjunt de processos de la percepció, totes elles es poden encabir en dos grans grups de models perceptius: els models actius i els models passius, seguint la divisió proposada per Borden i Harris (1980).

2.9.1.-Els models actius

Els models actius sorgeixen com a conseqüència de la dificultat que suposa explicar la percepció de la parla a partir de l'anàlisi del senyal. Ja hem explicat, a l'apartat 2.4 d'aquest mateix capítol, el problema que representa el fet que no es trobin elements invariants en el senyal que permetin aplicar el reconeixement de patrons per a identificar els diferents elements de la parla. És difícil, doncs, presentar la percepció de la parla com un mecanisme que, d'una forma automàtica, compara el senyal amb els patrons fonològics emmagatzemats en la memòria.

Per altra banda, com ja hem explicat a l'apartat 2.7.1.2, alguns autors suggereixen que la parla s'atén d'una manera diferent als altres senyals acústics que ens envolten. Quan un subjecte identifica (o espera identificar) un senyal de parla, invoca certs mecanismes específics per a la percepció de la parla i es centra en aquells paràmetres del senyal que aporten una informació determinada. Ara bé, com es produeix aquesta anàlisi de les característiques del senyal de la parla?

Ja hem explicat els diferents estadis o nivells en què es produeix l'anàlisi del senyal de la parla (apartat 2.2). El traspàs d'informació entre cada nivell i el següent o l'anterior es produeix constantment, atès que la naturalesa d'aquests estadis no és tancada, sinó que cadascun es recolza en els altres per a poder avaluar les tasques que es donen en cada estadi. No obstant cada estadi té la seva pròpia tasca per a reconèixer els elements propis que l'afecten en aquell nivell. Per fer-ho, no es pot parlar d'una aplicació de plantilles fixes, ja que se'n necessitaria una quantitat molt gran, atesa la variabilitat del senyal de la parla i, tot i així, sempre restarien situacions comunicatives que no estarien previstes en aquest conjunt de patrons. (Clark i Clark, 1977)

2.9.1.1.-L'anàlisi per síntesi

Per resoldre el problema de la variabilitat de la parla i de com en resten afectats els processos de percepció, Halle i Stevens (1962) van proposar una teoria activa que van anomenar "Anàlisi per Síntesi" (*Analysis-by-Synthesis*).

Segons aquesta teoria, el subjecte, de forma inconscient, produeix mentalment una versió sintètica de l'*input* de parla. Aquesta nova versió representa una hipòtesi construïda a partir dels indicis obtinguts en les diferents anàlisis que es produeixen en els estadis auditiu, fonètic i fonològic. Si les dues versions, la provinent de l'anàlisi i l'elaborada a tall d'hipòtesi concorden, el subjecte considera que la integració del senyal és correcta. Si no concorden, cal

donar més importància als processos d'anàlisi del senyal. Amb aquest funcionament es potencia la doble direcció en el flux de la informació: *top-down* i *bottom-up*.

Segons Clark i Clark (1977), el model de l'"anàlisi per síntesi" té un avantatge important, ja que permet adaptar-se tant a la variabilitat inter-locutor com intra-locutor. El sintetitzador intern pot ser ajustat per adaptar-se a diverses velocitats, tons, qualitats de la veu, dialectes i impediments de la parla. D'aquesta manera el sintetitzador reté totes les característiques fonètiques i fonològiques més importants que es troben inherents en el llenguatge, però fa concessions a certes qualitats de la parla dels locutors en una situació i en un moment determinats.

2.9.1.2.-La Teoria Motora de la percepció de la parla

A partir d'un plantejament general similar al que s'havia mantingut en el model d'anàlisi-per-síntesi, s'han formulat diverses explicacions per concretar el funcionament del procés intern pel qual es genera una representació sintetitzada. Un equip d'investigadors, format per Liberman i els seus col.laboradors (1967), van postular que tot subjecte disposa d'uns patrons articulatoris comuns per a la producció i per a la percepció de la parla. D'aquesta manera, la integració perceptiva del senyal provoca una activació dels mateixos mecanismes que s'utilitzen en la producció de la parla. Es a dir, pressuposa l'existència d'una única correlació (*mapping*) entre els fonemes i els seus corresponents comandaments articulatoris.

La característica més important d'aquest "mapa" és que els seus components es configuren a partir dels elements invariants de la parla, que no es troben en el senyal acústic pels problemes de la coarticulació i el solapament articulatori, però si que es troben en un nivell cognitiu més elevat. Aquest plantejament pren el nom de Teoria Motora (*Motor Theory*) de la parla, en referència als comandaments motors responsables de l'articulació dels sons.

Aquesta teoria va ser reelaborada per Liberman i Mattingly (1985). En la nova formulació els autors varen proposar nous postulats generals, coneguts amb la denominació de Teoria Motora Revisada (*Revised Motor Theory*). De fet representen diversos reajustaments respecte de la Teoria Motora inicial. Entre aquests postulats destaquen els quatre següents:

- 1.-La producció i la percepció de la parla es troben lligades de tal manera que comparteixen una representació comuna i una estratègia comuna de processament
- 2.-La unitat bàsica de processament és el gest articulatori subjacent (*underlying articulatory gestures*), que té una naturalesa abstracta i es troba associat amb un segment fonètic que es correspon amb la producció de la parla.
- 3.-La percepció es produeix de forma directa, per a la qual cosa existeix un mòdul innat especialment destinat a processar les ones acústiques i percebre directament les articulacions subjacents candidates. Posteriorment es realitza la selecció del patró articulatori correcte mitjançant un procés de decisió automàtica ràpida, al qual no tenim accés conscient

4.-No hi ha cap altre model que sigui tan complet com per explicar globalment el conjunt de circumstàncies que caracteritzen la parla, és a dir, el fenomen de la coarticulació, l'intercanvi de trets, etc.

El mateix Klatt (1988) comenta aquests quatre aspectes de la Teoria Motora Revisada i diu que no formen part necessàriament d'una teoria especial de la percepció de la parla, ja que no es presenta particularment vinculada a l'articulació, a no ser que es vulgui postular l'existència d'un estadi intermig de processament de la configuració del tracte vocal (*vocal tract shape*) com un component essencial per a la interpretació de les dades acústiques en termes lingüístics.²⁶ Per altra banda, suggereix que aquesta nova versió té alguns punts de contacte amb un altre model sobre la percepció de la parla (anomenat LAFS), desenvolupat per ell mateix i que exposem a continuació (en l'apartat 2.9.2.1).

2.9.1.3.-L' Aproximació Directa i Realista a la Percepció de la Parla

En una línia molt semblant a la teoria Motora de la percepció de la parla, Fowler (1986, 1991, 1994, 1995 i 1996) ha proposat una perspectiva diferent. Aquest nou model, que és conegut amb el nom de *Direct-Realist Approach to Speech Perception*, proposa que els elements que s'integren en la percepció no són altra cosa que les imatges mentals de les articulacions realitzades pel parlant que emet el senyal. La raó d'aquesta afirmació rau en el fet que la parla està estructurada segons la configuració articulatòria de la producció de la parla. En contrast amb altres models, el senyal acústic produït pel parlant no esdevé la finalitat de la informació per ell mateix sinó que simplement és el mitjà que serveix com a referent per a evocar aquests gestos articuladoris (*articulatory gesture*), els quals es perceben de forma directa a partir del senyal acústic.

En aquesta manera de percebre la parla, basada en la percepció general dels elements físics que ens envolten en el nostre món, es deixen de banda els diferents processos de tipus cognitiu atès que la percepció es produeix de forma directa, és a dir, únicament evocada, però no elaborada mitjançant diversos estadis de processament. Aquesta és la diferència més important entre aquest model i la Teoria Motora Revisada, amb la qual manté uns plantejaments força coincidents.

Darrerament, aquest model ha estat fortament rebutjat per alguns autors que han presentat proves sobre la presència de diversos processos de tipus fonètic i fonològic en la percepció del senyal acústic (sobretot quan es detecta com un senyal de parla *-speech mode-*), argumentant que no hi ha proves suficients que els oients es serveixin dels patrons articuladoris per percebre la parla. (Ohala, 1996; O'Shaughnessy, 1996). Per a aquests autors, el Model de Fowler està poc contrastat experimentalment i en alguns casos es basa en resultats experimentals poc rigorosos (Diehl *et al.*, 1996)

26 .Per a una discussió àmplia sobre la teoria Motora revisada i les característiques modulares de la parla, vegeu Mattingly i Studert-Kennedy (Eds.) (1991)

2.9.2.-Els models passius

Per poder donar una explicació coherent als processos de percepció, els models actius recalquen la importància que té el fet que el subjecte conegui els mecanismes de producció de la parla. Contràriament, els models passius parteixen del plantejament oposat; en aquests models adquireix importància l'estadi sensorial, és a dir, aquells mecanismes de filtrat de què disposa el subjecte per abordar el senyal de parla.

Diversos investigadors (Klatt, 1979; Massaro i Oden 1980, entre d'altres) han deixat de banda la línia de recerca que preveu l'existència d'un nivell fonològic distintiu, sobretot, a causa de la dificultat amb què s'han trobat a l'hora d'identificar els segments invariants en el senyal acústic corresponents als fonemes. Aquests autors parteixen de la base que, si no es pot fer evident una representació a nivell fonètic, podria ser gratuït el voler postular un nivell fonològic d'anàlisi en un model conceptual de la percepció de la parla. Aquesta concepció del problema planteja dos focus d'atenció.

D'una banda, alguns d'aquests estudis se centren més en els procés auditiu per tal d'extreure els elements acústics invariants que permetin una correspondència entre les propietats acústiques i les representacions fonètiques (Blumstein i Stevens, 1980, 1981).

D'una altra, autors com Klatt (1979) consideren que els elements invariants no afloren en un pla auditiu, sinó en un estadi molt més central, en el qual el reconeixement de patrons invariables es fa a través del reconeixement de "blocs" lèxics, evidentment molt més grans que el fonema.

Stevens (1986,1995) presenta un model que difereix de Klatt perquè suposa un lexicó²⁷ representat per una matriu de trets fonètics (precompilats) que facilitaria la comparació mitjançant seqüències d'atributs extrets del senyal en cada síl.laba. En canvi, Lahiri i Jongman (1990) propugnen l'absència d'un nivell d'anàlisi fonològica intermig entre la representació dels trets fonètics i el lèxic, i apunten que aquestes representacions fonològiques afloren directament en un "mapa" situat en la "ruta" del lexicó.

Tal com suggereixen aquestes discussions, la tendència d'estudi dels models passius no es recull en un model més o menys aglutinador (tal com succeeix en els models actius), sinó que es presenten com resultats experimentals que responen a diversos aspectes del processament de la parla.

27 .-El lexicó es defineix com un diccionari intern, emmagatzemat en la memòria de llarg termini, que està compost per entrades lèxiques. Aquestes entrades estan a disposició tant dels processos de producció com dels de comprensió de la parla. Existeixen dos models d'emmagatzematge del lèxic en el lexicó: el magatzem únic i els magatzems múltiples (Sebastián-Gallés, 1986). En paraules de Emmorey i Fromkin (1988), *"the mental lexicon is that component of the grammar that contain all the information -phonological, morphological, semantic, and syntactic- that speakers know about individual words and/or morphemes"* (p. 124).

Altres referències sobre l'estructura i el funcionament del lexicó es poden consultar a Aitchinson (1987), Cutler *et alt.* (1987), Forster (1976), Lahiri i Marslen-Wilson (1991), Luce (1986), McQueen (1991), Morton (1982).

2.9.2.1.-El model de Klatt

Entre els treballs que apunten en aquesta direcció, potser el més significatiu és el que es coneix com el model LAFS (*Lexical Acces from Spectra*) desenvolupat per Klatt (1979, 1986).²⁸

Aquest model està basat en l'esquema de funcionament dels sistemes artificials de reconeixement de parla i proposa que el subjecte, durant el procés d'integració del senyal de parla, genera hipòtesis lèxiques per tal de predir les paraules que componen el senyal. La unitat mínima de percepció en aquest sistema de percepció és el difonema (segment compost d'una vocal i una consonant o grup consonàntic), del qual la informació rellevant és la transició entre el segment consonàntic i el vocàlic. Cada transició possible es troba a disposició del subjecte com un patró diferent de reconeixement, i cadascun d'ells possibilita que el sistema pugui caracteritzar les variacions acústiques esperades dins dels segments fonètics que formen part del senyal. La naturalesa d'aquestes variacions ve determinada pel context fonètic que s'estableix entre els components acústics dels difonemes.

La percepció de la parla, en aquest model, consisteix a trobar el millor patró de comparació entre la representació espectral de l'input i les possibles aproximacions que es donen a través d'una xarxa de combinacions difonèmiques.

Les dimensions d'aquesta xarxa de difonemes poden contenir 4000 estats que es troben interconnectats a través de 6000 camins possibles. El marcador per a cada camí a través de la xarxa es basaria en certs aspectes del senyal, als quals es donaria un pes específic superior a d'altres. Aquests aspectes podrien ser l'expectació generada sobre la durada i el to de cada segment de parla.

Finalment, la xarxa no és de naturalesa fixa, sinó que s'adapta, en cada circumstància, a les condicions de la parla o a les característiques del parlant.

Tot i que en aquesta xarxa estan representades les transicions fonèmiques, no està configurada per a explicitar els trets fonètics o les decisions segmentals, sinó que trasllada directament les característiques acústiques del senyal a les hipòtesis lèxiques. Aquest procés es desenvolupa sense identificar els segments fonètics, sinó que aquests són precompilats mitjançant una xarxa de plantilles que conté les seqüències d'espectres acústics de la llengua en ús i que és capaç de controlar el lèxic sencer.

D'aquesta manera, no és necessari d'analitzar en cada moment el senyal de la parla, sinó que l'anàlisi es produeix durant l'aprenentatge de la llengua i, posteriorment resta emmagatzemat com un patró de reconeixement.

Aquest plantejament pressuposa que no existeixen processos *top-down* per a l'anàlisi dels estadis inferiors del procés, sinó que la base de la percepció es fonamenta en un procés de baix-a-dalt (*bottom-up*)

28 .-D'altres models passius d'accés directe, com el Logogén de Morton (1969, 1979) o el TRACE de Elman i McClelland (1984, 1986) s'exposen en l'apartat 2.13.1.

Segons O'Shaughnessy (1987), el model de Klatt proporciona material per una especulació interessant sobre la percepció de la parla, però, com en el cas d'altres models de percepció, és difícil de dissenyar experiments que puguin avaluar críticament el model. És improbable que un model amb tants estats i camins correspongui als processos mentals reals durant la percepció de la parla natural. El seu valor rau en el fet que ens proporciona un model comprensible per al disseny de sistemes de reconeixement automàtic de la parla.

2.10.- De la percepció a la comprensió

En els apartats anteriors hem exposat certes experiències que intentaven explicar el funcionament d'alguns aspectes molt concrets de la percepció. Els estímuls utilitzats en aquests experiments consistien en conjunts de sons que no estaven contextualitzats dins d'un discurs, sinó que es presentaven als subjectes en forma de logotomes o paraules pronunciats aïlladament (parla aïllada).

Aquest tipus d'estímul es pot considerar com força diferent de la parla que es troba en situacions naturals (en converses, en elocucions, etc.). Per aquest motiu, hom va suggerir que els mecanismes de percepció podrien estar desenvolupant tasques diferents davant d'estímuls corresponents a diferents modalitat de parla. D'aquesta manera, les teories construïdes sobre la parla aïllada podrien no ser extrapolables a qualsevol tipus de parla, entre els quals hi hauria la parla contínua natural (Clark i Clark, 1977). Segons aquests autors, la parla natural de qualsevol conversa es presenta als oients en forma de "substància" i "significat", i aquests oients es fixen en el significat i no en els sons de la parla, mentre que en parla aïllada l'atenció es fixa més en el so que en el significat.

Miller, Heise i Lichten (1951)²⁹ van desenvolupar una experiència en la qual avaluaven la resposta dels oients als dos tipus d'estímuls, parla contextualitzada i parla aïllada, i van arribar a la conclusió que els subjectes responien de forma diferent, tot atribuint aquests efectes a la major capacitat de predicció que té el subjecte davant d'estímuls contextualitzats en oracions.

A partir d'aquesta mateixa pressuposició, Pollack i Pickett (1964), van demostrar experimentalment que les estratègies en la percepció de parla contextualitzada i en parla aïllada eren, per força, diferents. L'experiment va consistir a enregistrar diverses elocucions produïdes per uns quants parlants. A aquests mateixos subjectes se'ls varen presentar seqüències (paraules o síl·labes) retallades dels seus propis enregistraments. Els percentatges de reconeixement dels segments descontextualitzats van ser molt més baixos que quan es presentaven dins de l'elocució.³⁰

En experiments sobre intel·ligibilitat, Schmitdt-Nielsen (1983) va constatar la resistència de la parla contextualitzada a ser perceptualment desfigurada en front de la parla

29.-Citada a Belinchon *et al* (1992). Sobre una experiència similar vegeu també Lieberman (1963)

30.-Segons Sotillo *et al* (1995), existeix una relació entre la intel·ligibilitat i l'assimilació pel fet que un mot extret del context perd certes referències necessàries per a la intel·ligibilitat. Sense el suport del context costa més de reconèixer una assimilació.

aïllada ³¹, la qual cosa fa suposar que els oients activen mecanismes diferents en funció de la tasca perceptiva que han de desenvolupar. La predisposició habitual de l'oient es concentra a percebre parla contínua, ja que és el tipus de parla que utilitza normalment per satisfer les seves necessitats comunicatives.

Per altra banda, en un estudi sobre restauració fonèmica, Warren (1970, 1976) va demostrar que, tot i canviar certs sons dels estímuls de parla contínua que es presentaven als subjectes per sons de no-parla de la mateixa durada, aquests "sentien" clarament el fonemes i els consideraven "físicament presents". Qualsevol intent d'identificar el so "retallat" a través de la localització del so estrany, fracassava. Els subjectes eren incapaços de localitzar el so intrús.³²

Des d'un altre punt de vista, les observacions fetes per Luria (1949) en pacients que havien perdut traumàticament les capacitats perceptives de la parla van demostrar que alguns d'ells, que no eren capaços de distingir fonològicament dos sons concrets, podien distingir perfectament dues paraules, entre les quals només hi havia la diferència, en la mateixa posició del mot, dels dos sons anteriorment presentats.

La conclusió que s'extreu d'aquestes constatacions apunta al fet que la capacitat de distinció fonològica entre sons aïllats podria ser un procés específic de la percepció de la parla i no necessàriament imprescindible en el procés d'integració de l'estímul de parla contínua (Clark i Clark, 1977).

2.11. -La percepció de la parla contínua³³

Entre les característiques de la parla contínua destaca la seva variabilitat. A diferència de la parla aïllada-acurada, les constants alteracions del ritme, de la velocitat, de la intensitat, etc., fan que la distribució dels diferents sons en l'eix temporal sigui molt poc homogènia ³⁴. A

31.-Sobre aquest tema vegeu també Moore (1982). Segons aquest autor, el senyal de la parla no té un únic aspecte sobre el qual es basa la percepció sinó que existeix un conjunt d'elements que aporten la mateixa informació. Es a dir, la informació que transporta el senyal de la parla és suficientment redundant perquè pugui ser distorsionada per un ampli ventall de condicions i, tot i així, romangui intel·ligible.

Sobre aquesta qüestió es basen alguns models de percepció de la parla com la teoria Motora que hem vist més amunt. Per a autors com Jenkins (1991) el senyal de la parla té unes característiques acústiques especials sobre les quals es fixa l'oient per a poder interpretar-lo. En canvi, autors com Miller (1982) han proposat models de percepció basats en un concepte totalment auditiu, tot negant que el senyal de la parla sigui "especial" i donant una explicació psico-acústica a la suposada codificació diferenciada entre senyals de parla i de no-parla.

32.-Sobre les il·lusions auditives i les restauracions fonèmiques es poden consultar els treballs de Eggen (1989), Eimas *et alt.* (1996), Repp (1992), Warren (1970, 1976). Una bona revisió del tema es troba a Warren (1996).

33.-Alguns autors fan referència al caràcter continuat del senyal de la parla en contrast amb l'anàlisi segmental que s'atribueix als processos perceptius, i que es realitza mitjançant unitats discretes. Sota aquest enfocament s'emmarquen diversos treballs, entre els que destaquem els següents: Aguilar *et alt.* (1991, 1993), Cole (Ed.) (1980), Crystal i House (1990), Fowler (1988), Klatt (1975), Marslen-Wilson i Welsh (1978), Pollack i Pickett (1964), Schmidt-Nielsen (1983).

34.-En aquest sentit, estudis com els realitzats per Aguilar i Machuca (1995) han demostrat que les variacions que produeixen els parlants en un acte de parla estan més relacionades amb la tasca que realitzen els interlocutors que amb altres factors com la familiaritat entre ells o la seva intencionalitat.

més, els elements ambientals emmascaren i deterioren el senyal i el context fonètic aboca a la difuminació dels límits entre els sons, tot encavalcant els uns amb els altres. Els límits de paraules no estan marcats en el senyal, i quan ho estan, sol coincidir amb el límit sintàctic. Com a conseqüència de tot això, és molt més difícil de segmentar i identificar els seus components, exclusivament, a través de patrons fonètics estables.

Per altra banda, en l'anàlisi del senyal que es realitza durant el procés perceptiu no és necessària la identificació exhaustiva, un-a-un, de tots els segments fonètics i llur correspondència fonològica. La participació del subjecte-receptor (*top-down*) proporciona informació immediata per suplir o complementar la que prové de l'anàlisi del senyal acústic.

En aquest sentit hi ha proves experimentals que el reconeixement de paraules pot no requerir la identificació de fonemes individuals com un graó intermedi³⁵. En algunes experiències basades en el temps de reacció (TR) a estímuls de parla, s'ha posat de manifest que les paraules compostes de diversos fonemes són reconegudes abans que els fonemes aïllats (Marslen-Wilson i Welsh, 1978).

Una altra característica de la percepció de la parla contínua és la presència de processos que ajuden a la segmentació del senyal. Un d'aquests processos es basa en la segmentació per síl.labes, que es manifesten com unitats de processament intermedi. D'aquesta manera els subjectes reconeixen els indicis sil.làbics en el senyal i poden recolzar-s'hi per reconèixer els límits de les paraules. En un estudi sobre la unitat de segmentació de la parla, efectuat per Savin i Bever (1970), es varen obtenir resultats que demostraven que el TR per a detectar la síl.laba era inferior al que calia per detectar un fonema. La conclusió d'aquest estudi proposava la síl.laba com la unitat de segmentació. En un altre estudi, realitzat per Cutler i Norris (1988), s'ha arribat a conclusions que advoquen la síl.laba com la unitat fonamental per a la segmentació de la parla, sobre la qual s'inicien les bases d'accés al lèxic (sobre aquesta qüestió vegeu, també, l'apartat 2.7.3).

En un estudi realitzat per Sendlmeier (1995), es conclou que els oients tenim a disposició diferents representacions mentals de diverses dimensions. El tipus d'unitat de representació utilitzat pels oients en qualsevol experiment depèn de les demandes de la tasca. No obstant, la síl.laba es manifesta com la unitat perceptiva que actua de forma predominant quan no es requereixen d'altres representacions específiques per a desenvolupar la tasca encomanada.

Un altre element, que actua associat amb l'anterior, és el domini de les regles fonotàctiques de la llengua. Aquest mecanisme permet de reconèixer una seqüència de sons com a pertanyents a una síl.laba d'una paraula o com a pertanyents a una seqüència de sons que no es correspon a cap possible síl.laba de la llengua i, per tant, esdevé una no-paraula (logotoma). En experiments amb TR, s'ha demostrat que les seqüències que violen les restriccions sil.lábiques de la llengua són considerades com paraules sense sentit molt més ràpidament que les pròpies paraules de la llengua (Garnham, 1985).

35 .-A l'apartat 2.10, ja s'han esmentat els experiments sobre restauració fonèmica desenvolupats per Warren (1970, 1976)

Tot plegat condueix a un seguit de reflexions sobre les característiques de la percepció de la parla contínua:

- 1.-La variabilitat és molt gran en la parla espontània natural i no sembla que distorsions tan importants com hi ha en alguns senyals sonors molt defectuosos impedeixin la comprensió. Per tant, cal buscar la invariança en l'activitat cognitiva de reconeixement més que en l'ona sonora. (Lindblom 1987b)
- 2.-La percepció de la parla contínua no es correspon al fet de posar junts els sons de la parla identificats aïlladament, sinó que existeix, també, la intervenció de nivells superiors que s'anticipen a l'anàlisi dels sons individualment. Tots dos processos interactuen, de manera que l'un es basa en la informació que aporta l'altre i a l'inversa.
- 3.-Tot el procés de percepció de la parla es troba subordinat a les restriccions de tipus semàntic, sintàctic, i lèxic que s'estableixen en estadis superiors de processament. Aquestes restriccions actuen normalment de forma inconscient i depenen en gran mesura del context, tant estructural com semàntic. No obstant, el procés d'integració i interpretació de la informació en els estadis superiors requereix paral·lelament un procés atencional.

2.12.-La comprensió de la parla

Definir la comprensió de la parla representa una tasca difícil pel fet que implica suposar l'existència de múltiples factors que operen conjuntament en un nivell d'anàlisi encara molt més central i, per tant, més inaccessible. Així, doncs, es molt més complicat de discernir la naturalesa d'aquestes operacions i com interactuen entre elles.

Com s'ha explicat en el capítol anterior, la percepció té la finalitat de segmentar, classificar i categoritzar els elements de la parla, tot recorrent uns estadis i desenvolupant en cadascun d'ells un procés mental. En canvi, en la comprensió, intervenen un conjunt d'operacions que tenen per objectiu la interpretació d'un missatge a partir de les informacions dels estadis perceptius. Hom acumula aquestes dades sensorials que, combinades amb les dades de tipus verbal, permeten d'excloure el sentit estricte dels enunciats i les intencions improbables del locutor. Es a partir d'una anàlisi de la motivació profunda de l'emissor que es poden descartar les diverses hipòtesis errònies per fer-se amb el sentit final del missatge (Keller, 1985).

Aquestes operacions estan organitzades en un tipus de processament similar al perceptiu, amb diversos estadis distribuïts de manera seqüencial, amb una base temporal constatable per a cadascun dels subprocessos (Belinchón *et alt.*, 1992). A més, les operacions actuen mitjançant símbols lingüístics i els resultats que en deriven es concreten en nous estats mentals.

Existeix, encara, un altre element intrínsec del procés: es tracta del coneixement que té el subjecte sobre el significat de les unitats lingüístiques (paraules o frases) que es van generant al llarg de la descodificació del senyal. És una competència atemporal i es manifesta sempre present per atendre i ajustar els diferents nivells d'interpretació pels quals es propaga el senyal

de parla. Mitjançant aquest coneixement, que es troba emmagatzemat en la memòria de llarg termini, existeix la possibilitat de fer progressar aquesta informació dins de cada nou estadi i entre els estadis que operin conjuntament.

A pesar del que s'ha dit fins ara, la distinció entre el que anomenem percepció i la comprensió pròpiament dita no té una delimitació clara, ja que depèn de la tasca que es proposi a l'oient i de si aquesta tasca activa totes les operacions superiors d'interpretació del missatge (Keller, 1985). Per exemple, si la tasca consisteix a distingir les paraules amb contingut semàntic de les que no en tenen (logotomes) en una llista on es troben barrejades, la frontera entre la percepció i la comprensió estarà força definida, ja que els logotomes no poden superar un cert nivell d'ordre semàntic i, tot i que poden ser repetits per un possible oient, no seran interpretats perquè estan buits de sentit. Si la tasca consisteix simplement a reconèixer una paraula dins d'un enunciat sintàcticament defectuós, el límit entre la percepció i la comprensió dependrà de la motivació del subjecte per fer una avaluació més o menys aprofundida del sentit de l'enunciat, al qual podria arribar per mitjà dels processos de restauració de la frase.

Segons Keller (1985) la comprensió del sentit final del missatge comporta la interacció de tres grans fonts d'informació provinents de diversos sub processos. Totes tres operen en paral·lel i de forma simultània i no és fins a la fase final, en la qual s'ensamblen tots tres, que es produeix pròpiament la comprensió. Aquestes fonts serien:

- 1.-Comprensió lèxica (que proporciona informació sobre la forma dels lexemes i de les expressions fetes i aporta el sentit estricte de cadascun d'ells)
- 2.-Comprensió gramatical o sintàctica (que proporciona informació sobre les relacions entre els lexemes)
- 3.-Comprensió prosòdica (que proporciona informació sobre la entonació, ritme, accentuació, l'èmfasi, la funció comunicativa de la locució, el coneixement del context comunicatiu, etc.).

En aquests tres processos es desenvolupen estratègies de descodificació diferents i independents durant la fase inicial. No obstant tots tres estableixen una estratègia activa en el procés de descodificació per la qual, d'acord amb la informació que prové dels estadis perceptius, es formulen hipòtesis sobre els elements lingüístics que aporta el senyal de parla i, posteriorment, es verifiquen aquestes hipòtesis tot comparant-les amb la nova informació que s'integra contínuament.

En el present treball no es pretén d'esbrinar els elements que configuren aquestes fonts d'informació ni com es produeix la interacció de totes tres. No obstant, hem cregut que era convenient de tractar, ni que només fos com un esment, el primer dels tres temes: l'accés i el reconeixement al lèxic. Actualment, aquest punt és motiu de recerca des de la psicologia del llenguatge. Els mètodes que s'hi estan usant estableixen el temps de reacció d'un individu a un cert estímulo lingüístic. Potser els objectius d'aquests treball difereixen poc o molt dels nostres. En tot cas, ens ha semblat que no es podien obviar per l'aportació tan important que representen per a entendre alguns aspectes del funcionament d'aquest estadi, al qual nosaltres també ens referim en el nostre experiment. Per aquest motiu, a continuació donem una visió dels models que han estat i són, en l'actualitat, els més actius en aquesta recerca.

Respecte al reconeixement de frases i als aspectes de comprensió semàntica i prosòdica, només citarem alguns treballs que ajuden a situar el tema, donat que aquesta part no té una repercussió immediata en l'experiment que pretén validar aspectes de la teoria del *hyper- & hypo- speech* (vegeu-ho més endavant, a l'apartat 3.1).

2.13.-Models de reconeixement de paraules i accés al lèxic

Segons Belinchón *et al* (1992), el reconeixement lèxic se situa en un dels primers estadis del processament de la parla. Malgrat això, però, no es tracta, ni de bon tros, d'un procés senzill, ja que hi convergeixen, per una part, les representacions dels estadis inferiors de la percepció, que contenen les informacions referides a l'estadi fonològic i a la segmentació del senyal de parla i, per una altra, les representacions dels estadis cognitius, com ara, la informació disponible en el lèxicó de l'oient i el coneixement del context, tant l'estructural (trets morfològics, sintàctics i semàntics) com el no estructural (és a dir, el coneixement del context comunicatiu i el coneixement del món que es troben a disposició del subjecte).

Sembla que l'anàlisi lèxica és un factor important d'entre els que incideixen en la intel·ligibilitat de la parla i que fins i tot resulta més important que els factors de tipus fonètic o fonològic. Segons Sotillo *et al* (1995), l'estructura del lèxicó i la freqüència d'aparició de certes estructures morfològiques hi juguen un paper cabdal.

Abans de continuar ens cal fer una distinció entre dos conceptes referits al processament lèxic, que podrien comportar algunes confusions: el reconeixement de paraules i l'accés al lèxic.³⁶

El reconeixement de paraules (o funció d'accés) fa referència al procés pel qual es compara la informació provinent dels nivells acústics i fonològics amb els patrons acústics i fonològics que sobre un determinat mot es troben emmagatzemats en la memòria. La comparació només s'efectua segons la forma de la paraula i està guiada pel senyal.

L'accés al lèxic (o funció d'integració) alludeix a un procés cognitiu que té la funció d'evocar una paraula concreta del lèxicó (diccionari mental compost de totes les paraules que es troben guardades en la memòria de l'oient). Aquest accés es produeix quan s'ha acomplert el reconeixement de la paraula en un estadi subordinat. Simultàniament es produeix un ajust entre el contingut semàntic de les paraules i el context del discurs..

Vegem-ho representat a la figura 2.2.

36 .-Alguns autors com Klatt (1979, 1986); Morton (1969, 1979, 1982); Elman i McClelland (1984, 1986), que postulen models d'accés directe, no accepten aquesta diferenciació, ja que està basada en aspectes formals que no són tinguts en compte en els processos lèxics (vegeu els models interactius, a l'apartat 2.13.1)

Esquema del processament lèxic

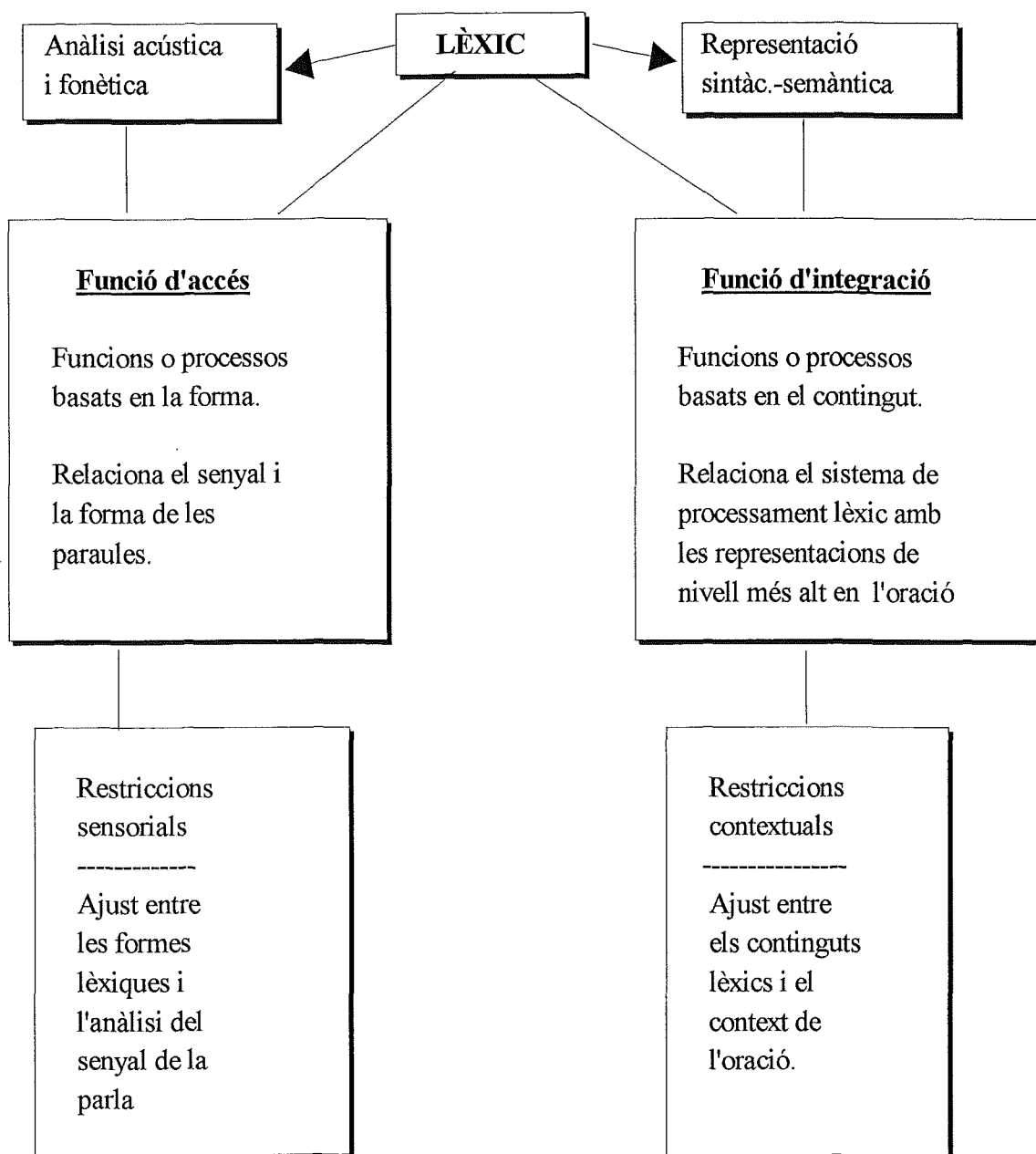


Figura 2.2.- Representació de l'esquema de processament lèxic en el qual es tenen presents tant els aspectes basats en la forma de les paraules com en el seu contingut semàntic (segons Marslen-Wilson, 1989)

En els darrers anys hi ha hagut diversos estudis sobre aquest tema que han mirat d'explicar, amb resultats experimentals, de quina manera l'entrada acústica entra en contacte amb la informació enregistrada en el lèxic i de quina manera es produeixen certes representacions internes que funcionen mediatitzant el procés. D'aquesta manera s'ha investigat si es produeix competència entre dos mots candidats davant d'un ítem acústic ³⁷; o si la freqüència d'ús dels mots afecta al seu reconeixement ³⁸ o si el context semàntic té alguna mena d'influència en aquest procés ³⁹.

Ara bé, el processament del lèxic no pot ser estudiat des d'una òptica aïllada del procés general de comprensió. Qualsevol intent per tractar-lo com un procés autònom, sense tenir present la interacció amb els altres nivells superiors i inferiors, suposaria un enfocament incomplet. (Pisoni i Luce, 1987).

Així, doncs, existeixen pocs models que intentin donar una explicació de l'accés al lèxic parlat des d'una visió conjunta de la comprensió, entre altres raons per la dificultat d'integrar, globalment, resultats experimentals que aporten proves a cadascun dels models, tot i que defensen postulats oposats.

Simplificant, la classificació que es pot fer d'aquests models està fonamentada, una vegada més, en la doble direcció que pren la informació (*top-down* i *bottom-up*) i en l'arquitectura en què estan disposades les diferents funcions de processament a través de les quals es dona més pes específic a algunes d'aquestes funcions i se'n resta a les altres. Així mateix, es distingeixen diversos graus d'interacció entre les funcions responsables del reconeixement lèxic i les funcions d'estadis superiors i inferiors.

Segons aquesta classificació⁴⁰ distingim entre:

- 1.-Models interactius del processament lèxic
- 2.-Models autònoms de comprensió de paraules
- 3.-Models mixtos de comprensió de paraules

Com és lògic, aquests models pretenen donar explicació al reconeixement i a la comprensió del lèxic a partir de la informació que arriba al subjecte a través dels seus sentits, tant si aquesta es dona des d'un nivell auditiu com des d'un nivell visual o tàctil (lectura). A més, alguns d'ells, connecten els processos de percepció i de producció de la parla, i fan referència a un estadi superior del coneixement des d'on opera, de forma general, el llenguatge. No obstant això, en aquest treball només ens referirem a la part d'aquests models que expliquin els processos de percepció i comprensió del lèxic des de la perspectiva del receptor. En aquest sentit exposem, a continuació, els postulats dels models que es corresponen amb la classificació anterior.

37.-Vegeu els estudis fets per Marslen-Wilson i col.laboradors (Marslen-Wilson i Tyler, 1980; Marslen-Wilson, 1984; Marslen-Wilson (Ed.), 1989; Marslen-Wilson, 1990; Lahiri i Marslen-Wilson, 1991; etc.)

38.-Sobre l'efecte facilitador que exerceix la freqüència d'ús de les paraules en el reconeixement lèxic, vegeu Forster (1976, 1990), Luce (1986), Marslen-Wilson (1990), Morton (1979, 1982), Onifer i Swinney (1981), Van Bergem (1996).

39.-Vegeu Connine (1990), Garrett (1978), Swinney (1979), Tyler (1990)

40.-Seguim la diferenciació que efectuen Belinchón *et al.* (1992)

2.13.1.-Models interactius del processament lèxic

La característica més important d'aquests models interactius és l'intercanvi constant d'informació entre els diferents nivells d'anàlisi (lèxic, sintàctic i semàntic) i la seva distribució en paral·lel i en múltiples direccions. Aquesta premissa comporta la presència simultània de tots els processos superiors de descodificació del senyal i, per tant, la disponibilitat d'informació semàntica en els nivells més baixos del reconeixement de paraules i l'accés al lèxic.

Dues teories importants s'emmarquen en aquest plantejament: El model Logogén de Morton i el model TRACE d'Elman i McClelland.

-El model Logogén (Morton 1969, 1979, 1982; Patterson i Shewell, 1987) postula l'existència, en el processament lèxic, d'elements passius que tenen la capacitat de representar cada paraula en el lèxic mental. Cada un d'aquests elements, que s'anomenen "logogens", conté tota la informació disponible en la memòria del subjecte sobre una paraula (el seu significat, la possible relació amb altres significats, les possibles relacions sintàctiques i morfològiques, la seva estructura formal, etc.).

En el cas de paraules polisèmiques (mateixa forma i diferent significat), existiria un logogén per a cada significat diferent. Així, la comprensió de qualsevol matís sintàctic, semàntic, etc. en el discurs suposaria l'existència d'un logogén específic.

En el procés d'accés al lèxic, tots els logogens atenen les característiques fonològiques i contextuals dels *input* lèxics del senyal. D'acord amb aquestes característiques, cada logogén estableix un nivell de resposta a la informació provinent de les altres anàlisis del sistema. Si el nivell de resposta del logogén supera un llindar determinat, aleshores s'activa i es troba disponible per als altres processos del sistema en forma de patró fonològic. No tots els logogens tenen el mateix llindar d'activació, sinó que els diversos llindars venen determinats per factors com ara la freqüència d'ús de les paraules.

Respecte al dos processos lèxics, reconeixement de paraules i accés al lèxic, el primer es produeix amb l'assoliment del llindar d'activació del logogén, mentre que el segon es comporta d'una manera automàtica tot posant a disposició del sistema tot el contingut informatiu que es troba dipositat en el logogén un cop reconegut el mot en qüestió.

Aquest és un model altament interactiu, en el qual tots els nivells d'anàlisi es combinen i interactuen a partir de la informació que tenen a disposició. Així, si una determinada informació, provinent d'un nivell d'anàlisi, resulta insuficient per activar un logogén, un altre nivell pot completar el llindar d'activació tot aportant informació d'una altra índole.

D'aquesta manera, el model explica els efectes del context i de la freqüència d'ús de les paraules, però escassament explica com es produeix la integració entre la informació fonètica i la informació dels nivells més elevats, quin és el curs temporal del reconeixement de paraules, la natura de les unitats perceptuals o el rol del lèxic en el reconeixement de paraules (Pisoni i Luce, 1987)

-El model TRACE de reconeixement de paraules (Elman i McClelland 1984, 1986) es basa en els principis d'activació interactiva, segons els quals, el processament de la informació es produeix en paral·lel a través d'interaccions excitadores i inhibidores d'un gran nombre de petites unitats de processament, el comportament de les quals es troba determinat per l'excitació o inhibició de les altres que té al voltant i amb les quals es troben connectades. Cadascuna d'aquestes unitats s'anomenen "nodes".

Totes aquestes interaccions es troben organitzades en tres nivells d'anàlisi: el nivell de trets, el fonològic i el lèxic. En el primer nivell existeix un gran conjunt de detectors de trets (diferents dels trets distintius de Jakobson), un per a cada una de les característiques fonètiques de la parla i s'activen constantment al llarg de l'eix temporal (vegeu l'apartat 2.7.2). En el nivell fonològic s'hi concentren detectors per a cada un dels possibles fonemes de la llengua. En el nivell lèxic hi ha detectors per a cadascun dels mots. Cadascun d'aquests detectors, sigui quin sigui el nivell d'anàlisi, es troba representat per un node.

El funcionament del model es basa en tres característiques fonamentals:

-En primer lloc, el model presenta un estructura que funciona jeràrquicament, de manera que els seus nivells produeixen representacions progressivament més abstractes. Cada possible representació és potenciada o inhibida en funció de la informació disponible. Cada representació potenciada serà susceptible de ser-ho més, mentre que la inhibida no podrà ser activada. Això representa que el sistema funciona de manera progressiva i incremental, per tal que el patró d'activació esdevingui cada cop més consolidat.

-En segon lloc, es produeix un processament de la informació en paral·lel. La informació s'emmagatzema de forma distribuïda en xarxes (semblantment a l'estructura de les xarxes neuronals) compostes per unitats subsimbòliques que s'estableixen i s'activen a partir de senyals d'excitació o inhibició. Aquestes xarxes actuen tot atenent, simultàniament, múltiples restriccions de tipus formal i de significat en cadascun dels nivells d'anàlisi.

-En tercer lloc, el processament es produeix de forma interactiva, és a dir, el flux de la informació no es processa en una única direcció sinó en múltiples direccions, de manera que, constantment es produeixen realimentacions en cada nivell provinents dels nivells superiors i inferiors. Així, les unitats de representació de la xarxa modifiquen el seu grau d'activació en funció de la informació que es transmet pels nodes amb els quals estan connectats. Aquesta informació tant pot generar una excitació com una inhibició de l'activació de cada node.

El model TRACE presenta diversos avantatges, entre els quals podem destacar el fet que la xarxa de detectors pot ajustar-se per reconèixer fonemes en la parla contínua. A més, el model considera la coarticulació entre consonants i vocals com una font important d'informació. Per altra banda, el fet que existeixin connexions entre el nivell lèxic i els altres nivells menys elaborats permet de donar explicació dels dos tipus de tasca: del reconeixement

lèxic i del reconeixement dels fonemes que componen la paraula. Finalment, el model interactiu reconeix la influència del context en els processos de reconeixement de paraules.

Però també planteja diversos problemes importants. Per una banda, els autors han adoptat detectors de fonemes com a unitats lingüístiques per formar paraules però no han explicat els fenòmens de la variabilitat de la parla (aspectes fonètics). Per exemple, no han precisat si aquests detectors són capaços d'actuar amb tots els possibles parlants o si cal un detector de fonemes per a cada parlant, cosa que limitaria enormement l'eficàcia del sistema (Sebastián-Gallés, 1986).

2.13.2.-Model autònom de comprensió de paraules

-El principal model de processament autònom del lèxic és el que es coneix com Model de recerca autònoma de Forster (1976, 1979, 1990). Aquest model planteja un funcionament serial del procés de reconeixement de la parla i d'accés al lèxic. Proposa que els diferents processos que es desenvolupen en aquest model operen de forma autònoma i basen la recerca lèxica en la freqüència d'ús de les paraules, organitzades dins del lexicó segons aquest criteri. D'aquesta manera, les que són més freqüents afloren més de pressa que les que no ho són.

Forster proposa l'existència de tres processadors que funcionen en paral·lel i de forma separada: un processador lèxic (que recull els indicis fonètics i fonològics per determinar l'entrada lèxica i assenyalar-la en el lexicó), un processador sintàctic (que recull les informacions lèxiques fins que en pot determinar l'estructura sintàctica) i un processador semàntic (que construeix l'estructura conceptual del missatge). A més, el model consta d'un processador no lingüístic anomenat *General Processing System* (GPS) que recull la informació elaborada en qualsevol dels tres processadors i incorpora el coneixement general necessari per prendre una determinada decisió o resposta. De totes maneres, el GPS no pot influir sobre els estadis inferiors, de manera que la informació només es transmet en una sola direcció. Així es suggereix que el procés està dirigit per les dades que provenen de l'exterior (*data-driven*) i impedeix que exerceixin la seva influència els nivells superiors del coneixement (procés *bottom-up*). L'única concessió, en aquest sentit, es concentra en el fet que la freqüència d'aparició de les paraules, d'alguna manera, configura un determinat coneixement del món, recollit en el lexicó de l'oient.

En els experiments de TR en els quals apareix un accés més ràpid a paraules d'un mateix camp semàntic en contraposició a paraules de camps semàntics diferents, el model apunta que aquestes relacions lèxiques es produeixen en nivells superiors del coneixement i, per tant, són posteriors als processos d'anàlisi lingüística. D'aquesta manera es manté la suposada autonomia dels tres processadors esmentats.

Aquest model ha explicat força eficaçment l'accés de paraules en posició aïllada, però la natura autònoma dels seus processos no permet explicar, amb prou convicció, l'accés del lèxic en frases o discursos en els quals les paraules tenen una relació estreta amb el context. D'aquesta manera, l'accés al lèxic se situa en una tasca concreta de recerca de paraules (designació de paraules) i divergeix d'una comprensió de la parla en la qual el context facilita, de forma important, la predicció de les unitats lingüístiques, entre les quals es troba el lèxic.

2.13.3.-Models mixtos de comprensió de paraules

Existeixen uns models que consideren certes propietats dels models autònoms i certes altres dels models interactius. Per aquest motiu s'anomenen "mixtos". El més representatiu d'aquests models és el proposat per Marslen-Wilson (1984, 1987; Marslen-Wilson i Tyler, 1980), anomenat Model de "cohort".

A diferència dels models anteriors, que estan plantejats per donar resposta a l'accés lèxic tant per via auditiva com visual, el model de "cohort" està pensat exclusivament per explicar com s'activa i es processa el lèxic procedent d'un senyal parlat. En el seu funcionament existeixen dues etapes, ordenades una darrera de l'altra.

-La primera etapa té una base fonològica autònoma ⁴¹ i la seva finalitat és la d'activar totes aquelles paraules del lèxic que coincideixen amb les característiques fonològiques del senyal en una comparació que es produeix, *on line* ⁴², des de l'esquerra cap a la dreta. Aquesta activació de candidats s'anomena "cohort inicial" i va seguida per un procés de selecció de la paraula candidata, gràcies a l'eliminació progressiva de les altres paraules de la cohort, que van esdevenint fonològicament diferents de la que sorgeix de l'anàlisi de les característiques del senyal. Finalment, només una paraula de totes les activades inicialment coincideix completament, en les seves característiques fonològiques, amb la que ha estat escoltada. ⁴³

-La segona etapa parteix del reconeixement lèxic de l'etapa anterior però no s'estructura de forma seqüencial sinó interactiva. Així, el model preveu l'existència d'un punt òptim de reconeixement que es produeix quan la paraula del senyal només té un candidat en la cohort de paraules activades. Normalment, aquest reconeixement es produeix molt abans que no s'hagi acabat la comparació de tots els sons del senyal. Segons les característiques de la paraula (el nombre de síl·labes, la distribució accentual, etc.), l'oient pot fer una predicció de la paraula candidata amb més o menys temps, però, realment, el factor que determina aquesta anticipació és el context. S'han trobat proves experimentals que demostren que les paraules contextualitzades sintàcticament o semànticament presenten un TR inferior en el reconeixement de la paraula de la cohort que les paraules presentades aïlladament (Marslen-Wilson i Tyler, 1980). Així, doncs, cal pressuposar que tant el procés que atén la forma del senyal com el que aporta la informació del context es produeixen de forma paral·lela, ja que no cal esperar la informació procedent de la via sensorial perquè s'activi la que es troba disponible pel context. Segons s'ha pogut demostrar experimentalment (Marslen-Wilson, 1987, 1989), els mecanismes de la comprensió de la parla de tipus sintàctic, semàntic i pragmàtic no interactuen directament amb els mecanismes

41 .-El procés es desencadena a partir d'un flux d'informació que prové del senyal i va de baix cap a dalt (*bottom-up*).

42 .-Aquest processament es produeix simultàniament a l'entrada auditiva del senyal. Per tant implica la impossibilitat, a priori, de reconèixer els components formals del senyal de la parla de forma conscient.

43 .-Sobre els efectes de competició entre diverses paraules aspirants dins de la cohort, vegeu Altmann (1990) i Marslen-Wilson (1990)

d'anàlisi acústica en el mòdul d'accés lèxic, sinó que més aviat competeixen en paral·lel per tal de localitzar la paraula dins del lèxic amb més eficàcia i velocitat ⁴⁴.

Aquest model presenta, com tots els altres, alguns inconvenients. Un d'ells és que no té present l'efecte que produeix la freqüència d'ús de les paraules i que ha estat contemplat com un factor important en els altres models vistos anteriorment. Un altre és que no explica com es produeix el contacte entre l'entrada lèxica i el seu contingut semàntic. En aquest sentit, Marslen-Wilson (1990) ha explicat, experimentalment, que els efectes de la freqüència d'ús no incideixen en l'activació lèxica, sinó que són efectes postperceptuals, ja que es donen en la visió, però no en l'audició.

Un altre problema que no pot explicar el model, tal com exposen Goldinger *et al.* (1996), és la correcció *on line*, dels errors que es produeixen en l'articulació de la parla i que, sovint, es concentren en la part inicial de la paraula. Si l'activació de la cohort es produeix d'esquerra a dreta, un error a començament de la paraula implica una activació errònia d'una cohort i, en aquesta cohort, no s'hi activarà mai un possible candidat per a ser comparat amb la paraula entrada en el senyal.

Finalment, com assenyala Sebastian-Gallés (1986), és curiós que un model que atén exclusivament la informació auditiva basi el funcionament de la cohort en models desenvolupats per al llenguatge escrit.

2.14.-Conclusions

En aquest capítol hem exposat que el concepte "percepció" designa un procés mental organitzat en un conjunt d'operacions que s'activen en el moment en què el senyal de parla arriba a l'oïda del receptor i que tenen la funció de processar la informació des d'un estadi purament auditiu fins a la detecció, classificació i organització dels trets de parla que són significatius per al desenvolupament de la comprensió del missatge.

Hem distingit tres nivells d'anàlisi del senyal: l'auditiu, el fonètic i el fonològic; i hem assenyalat el paper important que juguen els diferents tipus de memòria al llarg del procés de percepció. Així mateix, hem presentat els diferents models teòrics que han formulat diversos autors per tal de donar explicació a múltiples estratègies que utilitza el receptor per poder integrar el "senyal acústic" com un "senyal lingüístic".

Finalment, també hem volgut fer una aproximació al que diuen alguns autors sobre el pas de la informació des dels processos de percepció del senyal als de comprensió del missatge. En aquest punt hem vist diversos models que donen explicació sobre l'accés del lèxic i la integració de les paraules, i el paper que juga el context, tan el de tipus estructural (propri del missatge), com el corresponent al contingut semàntic del lèxic en què es configura el senyal.

44.-Per a una discussió sobre aquest punt, vegeu els articles de Samuel (1990), Connine (1990) i Tyler (1990).

Tot plegat és un valuós marc de referència a partir del qual s'organitza l'experiment (vegeu-ho més endavant, en el capítol 4) i es desenvolupa la discussió dels resultats (capítol 5). Abans, però, ens cal explicar quins són els processos que estudiem en aquest treball, qui els formula i com, i quines hipòtesis se'n poden derivar. Tot això és el que tractem a continuació, en el capítol 3.

3.- LA TEORIA DE LA HIPER- I LA HIPO-ARTICULACIÓ (HYPER- & HYPO- SPEECH)

El nostre treball té com a objectiu considerar la interrelació entre la informació que es troba en el senyal i la que es troba en el context en el procés de percepció de la parla. Per aquest motiu hem presentat els elements acústics del senyal (cap.1) i els elements perceptius i d'accés al lèxic que incideixen en la interpretació del missatge (cap. 2). No obstant, el nostre objectiu es centra principalment en establir el paper que juguen els factors relacionats amb la informació present en el senyal i en estudiar com interactua aquesta informació amb aquella que s'evoca en la ment del subjecte. Per altra banda, ens interessa saber si els elements que hem vist fins aquí poden condicionar, i en quin grau ho fan, el procés de comprensió del senyal comunicatiu. Tot plegat ens ajudarà a definir dos tipus de parla (seguint Lindblom, 1987b):

- 1.- La que presenta un desplegament d'informació excessiu en relació a les necessitats que exigeix la situació comunicativa per tal que un determinat missatge sigui comprès (*Hyper-speech*)
- 2.- La que presenta un desplegament d'informació insuficient en relació al que demana la situació comunicativa i, per tant, dificulta la comprensió final del missatge (*Hypo-speech*)

Per aquest motiu passem a presentar un model formulat per Lindblom que intenta donar explicació als processos coordinats de la producció i la comprensió del senyal de la parla. De fet, el model de Lindblom parteix d'un enfocament lingüístic, basat en la relació comunicativa parlant-oient i té en consideració, tant els elements que configuren el senyal, com la seva interacció amb d'altres elements de tipus contextual.

Aquest model es denomina així perquè formula, tal com veurem a continuació, l'existència d'un equilibri entre els factors de producció de la parla, més o menys articulada, i les expectatives que té el parlant perquè aquesta articulació sigui suficientment discriminable per part de l'oient i no produeixi una parla amb informació redundant i excessiva (parla hiperarticulada) o es quedi per sota de les expectatives de l'oient i la informació sigui deficient (parla hipoarticulada). La línia que separa aquests dos tipus de parla és constantment variable i inestable. Per aquest motiu aquest model es coneix com a teoria de la *Hyper- & Hypo- Speech* (d'ara en endavant l'anomenarem H&H), que nosaltres hem denominat, en català, com la "Teoria de la parla hiper i hipoarticulada".

3.1.-La teoria de H & H i l'estratègia de la variabilitat adaptativa

En el capítol 2, ja hem exposat l'interès dels investigadors per trobar trets invariants en el senyal de la parla. Durant anys, s'han desenvolupat treballs encaminats a aïllar aquests elements i a conèixer i comprendre el comportament dels participants en un intercanvi comunicatiu parlat. Per altra banda, en el camp de la enginyeria, el desenvolupament de sistemes de reconeixement automàtic de parla també demanava la presència d'aquests elements invariants en el senyal per aplicar-los en el disseny de patrons de reconeixement artificial. Els resultats obtinguts a partir d'aquells experiments no han proporcionat unes conclusions tant sistemàtiques com s'esperava i la suposada invariança ha estat minimitzada al llarg dels diversos estadis del procés de producció i percepció de la parla.

Un primer camp en què ha estat important la recerca de trets invariants és el procés d'articulació de la parla. La constància de situacions de solapament articulatori (per exemple, en parla ràpida) o de situacions de compensació articulatòria (en articulació forçada), demostraren que si la invariança es trobava en el nivell articulatori, havia de ser, forçosament, en un estadi abstracte, en un nivell des del qual es troben disponibles els patrons que controlen els moviments articulatoris. Aquests patrons s'anomenen "objectius articulatoris espacials" (*spatial articulatory targets*) (MacNeilage, 1970, 1980). Aquesta idea entronca directament amb un model de producció de la parla relacionat estretament amb el procés de percepció, com en el cas de la Teoria Motora (vegeu l'apartat 2.9.1.2 del capítol anterior)

La invariança també s'ha buscat en el senyal de la parla. Concretament Blumstein i Stevens (1979, 1981) i Stevens i Blumstein (1978, 1981) han suggerit que el senyal de la parla ha de contenir, de forma més o menys codificada, una certa informació invariant constant. El fet que la recerca actual encara no hagi aconseguit de desxifrar-la no hauria de suposar la negació de la seva existència. Els treballs realitzats per aquests autors han evidenciat alguns indicis acústics que revelen certs trets invariants, com ara la configuració de l'espectre del VOT en les consonants oclusives o certs trets de les consonants nasals. Tot plegat, però, no representa, en opinió de Lindblom (1983a), una gran contribució en aquesta línia de treball per resoldre la variació de la parla a nivell fonètic.

Més encara, el senyal acústic de la parla es presenta com una funció dinàmica de base temporal. En aquest dinamisme, l'element que aporta més informació és aquell que és capaç de contrastar entre les realitzacions fonètiques successives. En aquest sentit, la coarticulació entre consonants i vocals esdevé molt important per la informació que comporta (Lieberman i Mattingly, 1985). De tota manera, aquesta informació es presenta gràcies al contrast entre les propietats espectrals i no per les propietats fixes d'aquests segments. Lindblom l'anomena "contrast suficient" (*sufficient contrast*) o "discriminabilitat suficient" (*sufficient discriminability*), i sosté que està orientat a aconseguir el nivell d'intel·ligibilitat suficient perquè es donin les condicions comunicatives adequades. La variabilitat, doncs, està subjecta a aconseguir aquests contrastos fonètics. La realització més representativa, en aquest sentit, es dona en la coarticulació CV o VC, cosa que contribueix fortament a remarcar la variabilitat fonètica en contra de l'existència de trets purament invariants (Lindblom, 1987b; 1990).

Pel que sembla, doncs, la invariança en la parla només pot trobar-se com una

representació abstracta en els processos de la producció i de la percepció (vegeu el capítol 2). És a dir, a un nivell fonològic del processament. Llavors, la variabilitat fonètica podria ser la conseqüència d'una disposició en el senyal de la parla d'elements acústics, més o menys referits a llur representació mental, però sotmesos de forma incontrolable a múltiples factors de tipus estructural i ambiental.

Tant la insistència de Stevens i Blumstein en trobar elements invariants codificats en el senyal (*speech code*), com una possible configuració circumstancial del senyal acústic són dos aspectes a què ha intentat donar resposta la teoria anomenada *Hypo- & Hyper- Speech* (Teoria H & H) de Lindblom (1987b, 1990, 1991)

Aquesta teoria rebutja la idea que la invariança tingui una base fonètica i que, com a tal, pugui trobar-se més o menys codificada en el senyal de la parla. A partir d'aquest principi, pressuposa l'existència d'altres fonts d'informació que intervenen i col·laboren en el moment en què el senyal entra en el sistema de percepció de la parla i faciliten la seva interpretació i comprensió final.

Es considera que la naturalesa d'aquestes fonts és d'ordre intern i es fa referència als processos *top-down* (vegeu capítol 2), a partir dels quals es produeix un efecte facilitador en la interpretació del senyal d'entrada (*bottom-up*). Així, es combinen la informació que aporta el senyal de la parla, la informació de què disposa el subjecte sobre la situació comunicativa, el tema, l'interlocutor, la circumstància i la informació de tipus gramatical i lèxica. Fruit d'aquesta interrelació sorgeix un estat dinàmic, que fluctua constantment en funció dels canvis esdevinguts en el procés comunicatiu i que està guiat, en bona part, per les necessitats d'interpretació del senyal. Mitjançant aquests mecanismes facilitadors es possible d'enfrontar-se a una possible degradació del senyal o a una informació que es presenta com incompleta o reduïda (Lindblom, 1991).

Per altra banda, el parlant juga un paper important, ja que pot dirigir els seus esforços cap a aconseguir un senyal que sigui adequat a les circumstàncies comunicatives i, sobretot, que estigui orientat a l'oient i a les seves possibilitats per rebre i interpretar el senyal amb garanties (*listener-oriented behavior*)¹. Així, doncs, el parlant crea una interrelació amb l'oient tot adaptant l'articulació (estructurada pels patrons de producció) i reorganitzant-la constantment en funció de les variacions successives d'aquesta interrelació. D'aquesta manera, encara contribueix més a la variabilitat del senyal (Lindblom, 1990).

Aquesta variabilitat del senyal de la parla i, per tant, la falta d'invariança, és fruit de l'organització adaptativa de la producció de la parla envers un objectiu comunicatiu: la comprensió del missatge (Lindblom, 1983a). En aquest sentit MacNeilage (1970)² afirmava:

"... the essence of the speech production process is not an inefficient response to invariant central signals, but an elegantly controlled variability of response to the demand for a relatively constant end."

La presència o absència d'elements invariants en el senyal de la parla no representa el

1 -Per exemple, Bard *et al* (1995) han demostrat que els parlants pronuncien les paraules menys clarament quan els referents comunicatius són presents que quan no hi són.

2 .-Citat a Lindblom (1987b)

problema principal de la invariança fonètica, a la qual se subordinaria el fenomen de la variabilitat. Al contrari, la variabilitat fonètica és el quid de la qüestió en tant que reflecteix els diferents objectius i propòsits que condicionen el senyal de la parla davant de les circumstàncies reals de la comunicació parlant-oient.

L'adaptació constant del parlant a cada situació depèn de circumstàncies molt diverses, com acabem d'explicar. En funció d'aquestes circumstàncies, el senyal de la parla haurà estat construït pel parlant de forma més o menys relaxada. En part, la variabilitat adaptativa es pot explicar i predir des de les interaccions entre els diferents nivells de la producció del missatge. Així, una parla molt acurada, segurament tindrà molts elements invariants i esdevindrà una parla força intel·ligible que anomenem parla clara o hiperarticulada (*hyper-speech*), mentre que una parla relaxada presentarà molta més variabilitat, amb probables reduccions articulatòries i contrastos deficients entre els seus components fonètics. Llavors, la seva intel·ligibilitat serà baixa i l'anomenem parla no-clara o hipoarticulada (*hypo-speech*).

La teoria *H & H speech* fa referència a les condicions del senyal respecte a la informació fonètica que transmet. Les característiques d'aquesta informació se situen entre la parla clara, en un extrem, i la parla no-clara, en l'altre extrem. Les constants variacions entre aquests dos extrems s'organitzen en funció de l'adaptació parlant-oient. Aquesta adaptació, teòricament, pot ser graduada i distribuïda al llarg d'un contínuum, de més a menys, (tal com veurem a continuació) i el seu correlat perceptiu és la intel·ligibilitat.

La qüestió essencial d'aquest plantejament sobre el dinamisme del senyal és la següent: ¿de què depèn el grau d'adaptació entre el parlant i l'oient i quins són els factors que intervenen en aquesta interacció?

La resposta a aquesta qüestió està lligada als processos interns que desenvolupen tant el parlant com l'oient. En el procés de producció, el parlant estructura i organitza la seva parla de manera que no representi un esforç articulador superior al necessari per a generar un senyal de parla concret. És el que s'anomena "economia fònica" (*talker-oriented simplification*)³ i està regulada pel parlant en cada moment de l'elocució, d'acord amb les seves necessitats comunicatives. El senyal que se'n genera és ric o pobre en informació segons aquest ajust.

Per altra banda, el parlant té també un coneixement del seu interlocutor i de les seves característiques; en funció d'aquesta informació i de la pròpia intenció comunicativa, elabora un tipus senyal que considera idoni per a la situació comunicativa i per a l'interlocutor amb qui es comunica⁴. Aquesta possibilitat d'adaptació a l'oient s'anomena "plasticitat"⁵ (*listener-oriented reorganization*) i determina les propietats articulatòries de segments d'hiperarticulació o d'hipoarticulació.

Tant l'economia fònica com la plasticitat esdevenen primordials en la interacció parlant-oient; el parlant imprimeix unes característiques en la producció del senyal que es troben pròximes a un límit teòric, per sobre del qual l'oient pot desenvolupar els processos perceptius. Si el senyal es troba per sota d'aquestes característiques mínimes, l'oient no rep la suficient informació per poder començar el processament del senyal de la parla. Si, per contra, el senyal

3.-Lindblom (1983a, 1983b, 1987b, 1996). Sobre l'economia fònica, vegeu també l'apartat 1.3 del capítol 1

4.-Vegeu els estudis citats a la nota 8 del capítol 1 sobre els estils de parla.

5.-Gazzaniga et al. (1979), Lindblom (1983a, 1983b, 1987b, 1996)

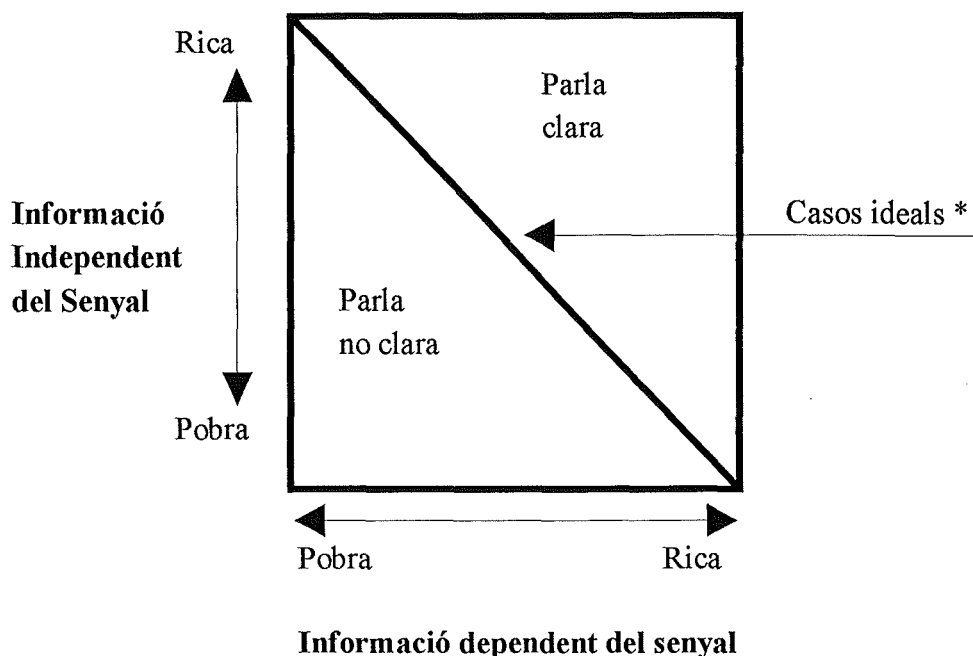
sobrepassa aquest límit, s'haurà gastat més energia de la necessària, malgrat que l'oient pugui processar perfectament la informació. Lindblom (1990) parla, fins i tot, d'una lluita (*struggle*) entre aquests dos factors que ell anomena "orientats a la producció" (*production-oriented*) i "orientats a l'oient" (*output-oriented*) i que apunten cap a la dimensió *hypospeech-hyperspeech*.

En aquest sentit, Nootboom (1991) suggereix que el parlant ideal ha de mostrar un comportament orientat cap a l'oient a tots els nivells de la producció de la parla i ha de ser capaç de fer una predicció estimativa de la informació implícita que comporta el missatge i que haurà de tenir en compte l'oient en el procés d'interpretació i comprensió (*listener-generated*). Si els parlants generen la parla exclusivament basats en la producció (*production-oriented*) i no tenen en la seva ment les circumstàncies perceptives de l'oient (*output-oriented*) per a poder-la-hi adaptar, possiblement es produiran múltiples dificultats en la percepció i interpretació del senyal. No obstant, Nootboom insinua que, malgrat la llibertat que té el parlant per a generar el seu missatge, les convencions lingüístiques exerceixen una inhibició important o una limitació dels efectes no lingüístics en la producció de les paraules deguda als patrons prosòdics de la llengua⁶.

Tots aquests factors depenen fonamentalment del parlant i es troben reflectits, d'una manera o altra, en el senyal. És el que Lindblom ha anomenat "informació dependent del senyal" (*signal-dependent information*) i és, sobre aquesta informació, que s'inicien i es desenvolupen els processos perceptius (*bottom-up*). Ara bé, si el parlant ha portat tota la iniciativa en la codificació d'aquest senyal, existeix un altre tipus d'informació que depèn de les capacitats comprensives de l'oient, anomenada "informació independent del senyal" (*signal-independent information*), la qual sorgeix directament dels processos cognitius de dalt a baix (*top-down*).

La relació entre aquests dos tipus d'informació sol representar-se mitjançant una línia inclinada que separa la "parla clara" de la "parla no-clara" i que assenyala els casos ideals en els quals es produeix la interrelació parlant-oient. Aquesta línia també representa el límit per sobre del qual existeix una parla suficientment intel·ligible per permetre la comunicació. Evidentment, les interrelacions que es produeixen per sota d'aquesta línia de casos ideals no tindran la intel·ligibilitat necessària perquè l'acte comunicatiu pugui arribar a bon fi (vegeu la figura 3.1).

6 .-Sobre aquesta qüestió vegeu també Nootboom *et al* (1978)



* Línia límit per a la probabilitat de comprensió

Figura 3.1: Interrelació entre la informació dependent i independent del senyal per a la comprensió de la parla, segons Lindblom (1987b)

Els dos eixos de l'esquema mantenen una relació directa en funció de la capacitat predictiva que pugui tenir l'oient i que se centra en la informació continguda en el context comunicatiu i no en el senyal.⁷ Així, davant d'un grau elevat de coneixement contextual (informació independent del senyal), l'oient pot desplegar unes expectatives (sobre la temàtica, sobre els possibles camps lèxics, etc.) que li permetin anticipar-se a la informació que conté el senyal. En aquest cas, l'oient no requereix de tots els atributs informatius del senyal per interpretar el missatge, sinó que pot assolir-los per mitjà d'una relació suficient entre un context ric i un senyal pobre. Aquesta mateixa situació és igualment efectiva en sentit invers, és a dir, amb una estreta relació entre un senyal ric en informació i un context poc explícit. En aquest cas, també existiria la possibilitat d'interpretar correctament el missatge.

La teoria de H & H està desenvolupada per donar resposta al fenomen de la variabilitat del senyal. El seu plantejament assumeix la interrelació de diversos dels components que intervenen en un acte comunicatiu. Així, la comprensió es produeix per la combinació fluctuant de la informació que es troba en el senyal, i que contínuament és adaptada pel parlant, i la informació independent del senyal, que contínuament és adaptada per l'oient.

7.-Segons Pinillos (1975) la percepció va precedida per les expectatives que ha generat el receptor. Així mateix els processos anticipatoris adquirits són intrínsecs de la activitat perceptiva.

Segons Lindblom (1990), aquesta adaptació que desenvolupa l'oient inclou la discriminació entre tots els components que integren el lèxic per tal d'obtenir l'*ítem* correcte. Això converteix l'accés al lèxic en una funció de distinció i caracterització de l'estímul acústic més que no de comparació de trets invariants. Aquest procés de discriminació està regit pels processos complementaris al senyal, que són bàsicament cognoscitius.

Per bé que el procés que descriu la teoria H & H vincula tant el parlant com l'oient, el present treball se centra en el procés d'aquest darrer sota la perspectiva que acabem de descriure: la discriminació regulada pel coneixement. Els processos que resseguim contemplen tant les característiques del senyal com els diferents estadis de la percepció, l'accés al lèxic i la comprensió final del missatge.

3.2.- Evolució i aplicacions de la teoria de H & H

En els darrers anys s'ha intentat un aplicació explícita d'aquesta teoria. Alguns investigadors, com McAllister i Dufberg (1989), han intentat encarar aquesta perspectiva en la mesura de la comprensió de la parla. Concretament van basar el seu experiment en parlants nadius normals (com a grup de control), en estudiants de L2 i en parlants amb dèficits auditius. Els resultats dels dos darrers grups mostren que la informació independent del senyal es manifesta força rendible per als subjectes que estudien la segona llengua, que es mostren poc sensibles a l'emascarament. En canvi, el grup de persones amb discapacitats auditives obtenen uns resultats significativament inferiors quan se'ls presenta parla emmascarada. En el cas dels primers, la informació que es troba disponible en el senyal és suficient per activar la informació independent del senyal i, per tant, s'assoleix suficientment la línia dels casos ideals. Per altra banda, els subjectes del segon grup, que disposen de molt poca informació de parla en el senyal a causa de la seva incapacitat auditiva, obtenen resultats inferiors perquè els és més difícil d'assolir la informació continguda en el senyal i, per tant, sovint es troben per sota de la línia dels casos ideals.

Altres autors, com Benoît i Abry (1995), plantegen una revisió de l'esquema de Lindblom per tal de bastir una escala de mesura de la complexitat lingüística de la informació independent del senyal i de la qualitat acústica de la informació dependent del senyal. Segons aquests autors, l'eix corresponent a la informació dependent del senyal podria ser quantificat en termes de paràmetres (com ara la relació senyal-soroll) i, per tant, podria establir-se una escala per determinar, en cada moment, el nivell que es correspon amb aquesta informació. En l'eix de la informació independent del senyal, en canvi, és molt més difícil de definir una unitat de mesura, ja que els factors que hi prenen part són de tipus cognoscitiu i psicològic. Com a exemple es pot citar, el grau de decisió de l'oient davant d'un estímul de parla o la previsió del parlant sobre la capacitat perceptiva del seu interlocutor.

Ja hem explicat més amunt que el parlant adapta la parla en funció de les possibilitats que creu que té l'oient d'extreure la informació del senyal. D'aquí que el parlant avalui constantment la importància del context sintàctic i semàntic del seu discurs. Tant és així que el parlant sovint realitza una hiperarticulació (*hyperspeech*) quan ha de pronunciar mots que, dins d'un context, són impredecibles i contenen una càrrega semàntica superior als altres que

l'envolten⁸. No obstant, aquest esforç articulatori està estrictament limitat a aquelles situacions en les quals la complexitat lingüística o la intel·ligibilitat acústica ho aconsellin. En tots els altres casos, i en funció del principi del mínim esforç, l'articulació de la parla esdevé hipoarticulada (*hypospeech*).

Benoît i Pols (1992)⁹ revisen aquesta interrelació entre els dos eixos de la informació i suggereixen que el diagrama de Lindblom és deliberadament esquemàtic i, per aquest motiu, en proposen una nova versió que es pot representar com segueix:

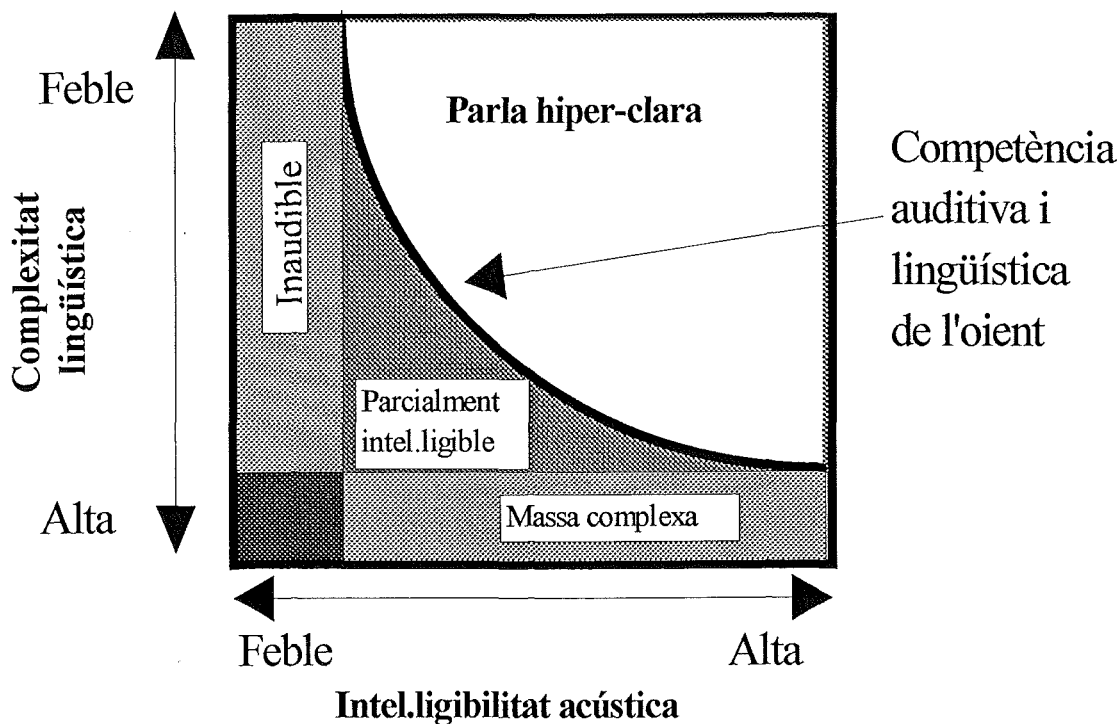


Figura 3.2: Esquema de Benoît i Pols (1992), que modifica el de Lindblom (1987b). En aquesta versió es distingeixen dues zones extremes en les quals la parla no pot ser compresa. Així mateix s'assenyalen diversos graus d'intel·ligibilitat en funció de les característiques acústiques del senyal i de la complexitat de la situació comunicativa.

En aquest nou esquema se sintetitzen tres pressupòsits que no es prenen en

8 .-En aquest sentit, Cutler i els seus col·laboradors (Cutler - Foss, 1977; Cutler - Darwin, 1981) van demostrar que els trets prosòdics guien l'atenció de l'oient, fins al punt que els accents oracionals li permeten de predir els mots que contenen la informació més important dins de l'elocució (paraules-clau). A partir d'aquesta troballa, Cutler considera que els elements prosòdics són la primera font d'informació, fins i tot per davant dels elements que presenten una variabilitat molt petita.

9 .-Citat a Benoît i Abry (1995)

consideració en l'esquema de Lindblom. En primer lloc es té present l'existència d'un llinar acústic per sota del qual l'oient no pot percebre res (la parla esdevé inaudible). En aquest punt la comprensió és impossible. Per altra banda es pot donar el cas que el missatge, tot i ser construït en una llengua coneguda per l'oient, contingui una informació excessivament complexa, de manera que sature la capacitat memorística o comprensiva de l'oient. En aquest segon cas tampoc no hi pot haver comprensió de la parla. Finalment, la darrera variació que introdueixen aquests autors respecte de l'esquema inicial de Lindblom consisteix a situar els casos ideals per a la comprensió del missatge en una línia corba que es trobaria per sota dels casos ideals definits per Lindblom. Aquesta novetat ve donada per la suposició que un missatge mitjanament complex tolera millor la degradació en combinació amb unes condicions acústiques mitjanament bones. El resultat d'aquesta combinació pot conduir a una comprensió del missatge per sota de les condicions que establí Lindblom en el seu diagrama. De la demostració que en fan els autors s'arriba a la conclusió que la interrelació entre els dos factors de la negociació parlant-oient no és fruit d'una addició entre el component dependent del senyal i el component independent del senyal, com proposa Lindblom amb la línia recta del seu diagrama, sinó fruit d'una multiplicació en la qual qualsevol índex d'un factor és suficientment important per operar amb els factors més o menys definits de l'altre eix informatiu.

Així, doncs, l'esquema revisat de Lindblom està construït sobre la base que l'eix de la informació independent del senyal es pot relacionar amb la complexitat lingüística del missatge i la informació dependent del senyal amb la intel·ligibilitat acústica. Aquesta correspondència, però, es manifesta en tots els casos en què existeix intenció comunicativa, tant pel que fa al parlant com pel que fa a l'oient.

No obstant, en un plantejament experimental en el qual s'intenten aïllar els factors que intervenen en la comprensió dins de l'acte comunicatiu, difícilment es podria pensar en l'existència d'un ajustament de la producció del senyal a la complexitat lingüística del missatge i a les condicions acústiques ambientals. Prenent com a base d'estudi un dels interlocutors de l'acte comunicatiu (l'oient), cal que centrem l'atenció en els indicis informatius de què pot disposar en escoltar una cadena parlada, encara que no s'hagi produït com a resultat d'un acte comunicatiu d'interrelació parlant-oient.

3.3.-Discussió del model i hipòtesis

En els capítols anteriors hem descrit diversos aspectes que han estat, i encara són, el motiu de la recerca dels mecanismes que intervenen en la percepció i la comprensió de la parla. Tots ells configuren un conjunt complex d'accions i d'interaccions que difícilment poden trobar explicació de forma aïllada. Al contrari, la finalitat última de cadascun d'aquests elements puntuals es troba en la combinació amb els altres per a assolir la comprensió del missatge comunicatiu. Així mateix, el seu funcionament esdevé organitzat amb aquest objectiu (Lindblom, 1983a).

No resulta senzill de confeccionar un model que doni explicació de totes les tasques que es produeixen en el procés d'integració del senyal parlat. Ja hem vist més amunt que existeixen diversos models que intenten explicar els processos perceptius de manera global. No obstant, fins al moment actual no en coneixem cap que doni una resposta totalment

satisfactòria davant dels resultats experimentals que es generen en intentar validar cadascun d'aquests models. De manera similar s'esdevé amb els models que intenten explicar l'accés i integració del lèxic i amb els models sobre els processos comprensius.

La complexitat del procés augmenta a mesura que hi intervenen nivells més centrals i, per tant, encara es fa més difícil de definir un patró que abracci la totalitat del procés de comprensió de la parla. De fet, hauria de ser tan ampli i preveure tantes situacions comunicatives que deixaria de ser un model teòric per passar a ser una descripció de la realitat.

Per aquest motiu la teoria de Lindblom sobre la interrelació dels elements en el fet comunicatiu es presenta com un marc teòric desproveït de les explicacions puntuals de cada situació comunicativa. És, però, en aquesta manca de precisió on es troba el seu potencial de predicció. Si és capaç de suportar i de donar recer a les diverses interrelacions que esmentàvem suara, voldrà dir que ens trobem, si més no inicialment, davant d'un possible model que aglutina els processos que intervenen en la producció i els que intervenen en la integració del senyal de la parla.

És en aquest sentit que el present treball pretén de desenvolupar i comprovar experimentalment diversos aspectes de l'esquema de *hyper- & hypo- speech*. En concret, l'experiment es centra en dos aspectes: 1) l'estudi de la percepció de senyals de parla que es presenten en condicions de baixa intensitat (IDS pobra); i 2) el comportament que desplega el receptor per tal d'adaptar-se a aquestes circumstàncies tan desfavorables. Per a dur-ho a terme intentarem abordar l'acte comunicatiu des d'un nombre ampli d'elements i funcions que intervenen en la comprensió de la parla. Mitjançant el control sobre el seu funcionament podrem aproximar-nos a la verificació i/o reestructuració del model. La metodologia que utilitzarem està basada en la verificació d'hipòtesis de treball, a partir de les quals s'ha dissenyat l'experiment (vegeu el capítol 4). En els propers apartats (3.3.1, 3.3.2 i 3.3.3) presentem cadascuna de les hipòtesis que pretenem validar, les quals se situen en la línia de la teoria que acabem de descriure. Per tal de contextualitzar-les, abans de cadascuna d'aquestes hipòtesis, desenvoluparem unes consideracions prèvies que pretenen ser breus discussions d'alguns aspectes del model de Lindblom sobre les quals es bastirà el disseny de l'experiment.

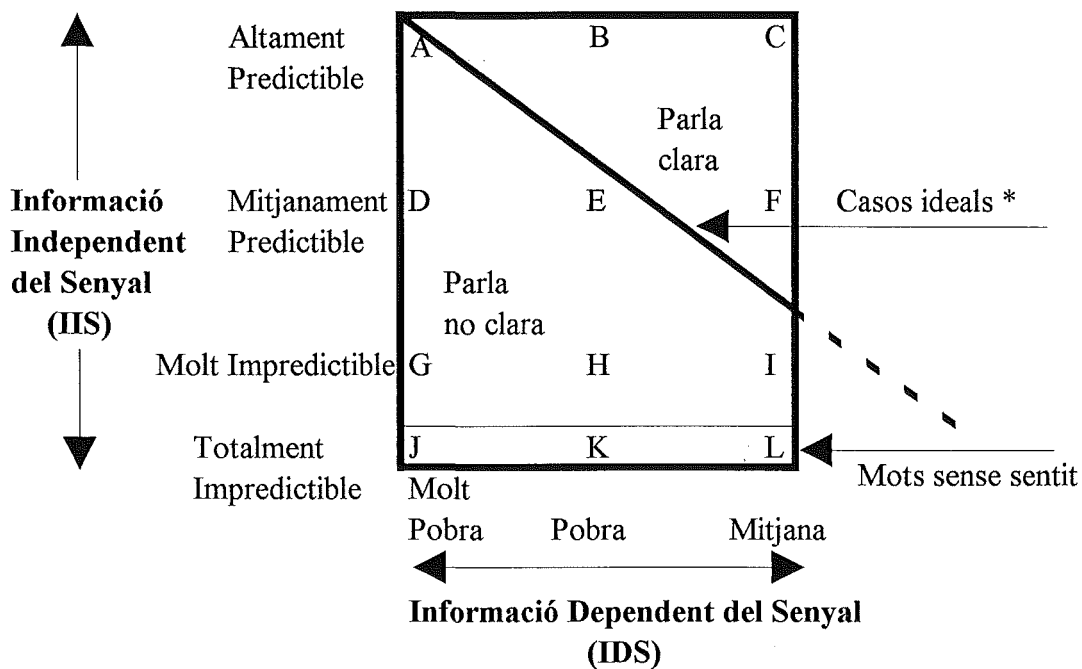
3.3.1.-Consideracions i hipòtesi 1: la relació entre la intel·ligibilitat del senyal i la predicció de la informació

L'esquema de la interrelació entre la informació dependent del senyal i la independent del senyal (figura 3.1) assenjala una línia que defineix els casos ideals en els quals es realitza la comprensió sense que hi hagi una despesa innecessària d'energia per part de qui produeix el senyal. Pel que es desprèn dels límits d'aquesta línia, existeix una diversitat de situacions comunicatives que tenen com a resultat una parla clara (si es donen per sobre de la línia) o una parla no-clara (si es donen per sota). Segons les circumstàncies en què es produeixi la interrelació dels múltiples factors que intervenen en el fet comunicatiu, el resultat es trobarà per sobre o per sota de la línia de casos ideals.

En una primera aproximació a aquests factors presentem diversos casos teòrics, de tal manera que la seva distribució dins de l'esquema sigui equitativa. Per poder formular aquesta primera hipòtesi fem dues consideracions prèvies:

- 1.-La gradació dels dos eixos és plantejada per Lindblom de forma qualitativa, ja que no defineix, a priori, els factors determinants de cadascun d'ells.
- 2.-Les situacions comunicatives que es trobin per sota de la línia de casos ideals són forçosament inapropiades per a aconseguir de comprendre el missatge.

A partir d'aquestes dues premisses hem establert 12 possibles situacions de comprensió d'un missatge i les hem situat en la gràfica de Lindblom segons la influència de la informació que s'aporta en cadascun dels eixos. La distribució resultant es mostra en la figura següent:



* Línia límit per a la probabilitat de comprensió

Figura 3.3.- Diagrama de les expectatives d'interrelació i probabilitat de comprensió en funció de les condicions genèriques del senyal i del context.

Cada cas representat en la figura 3.3 prové del producte de múltiples factors. Alguns d'ells determinen l'escala de l'eix de la IDS (informació dependent del senyal) i d'altres de la IIS (informació independent del senyal).

Hem concretat els factors que intervenen en la IDS amb la denominació d'intel·ligibilitat acústica del senyal. Sota aquest concepte tenim presents les condicions en les quals s'ha produït

el senyal de parla i en les quals s'ha tramès fins al receptor.

- Per una banda considerem que el locutor ha produït el senyal de parla amb una articulació suficientment clara perquè sigui avaluada com a hiperarticulació.
- Per altra banda considerem, com un factor determinant, la intensitat general que té el senyal acústic quan arriba al subjecte. Així, malgrat que la possible variació d'aquest darrer factor és molt gran, en la nostra aproximació prenem tres punts de l'eix per poder configurar una hipòtesi de treball: 1) Informació molt pobra, 2) Informació pobra i, 3) Informació mitjana. Hem adoptat aquesta gradació en el benentès que ens situem en l'extrem baix de la intensitat acústica a partir del qual tractarem de determinar el límit dels casos ideals. Si establim la condició d'un senyal acústicament molt pròxim al llindar d'audició, tindrem un factor dependent del senyal força controlat i podrem examinar el paper que juga la informació independent davant d'un senyal deficient.

Donat que en el present experiment ens centrem en les intensitats baixes, en aquesta figura 3.3 tan sols representem la informació dependent del senyal fins al grau "mitjana", en el benentès que només reproduïm una part de l'esquema de Lindblom representat en la figura 3.1. La línia de punts discontinus assenyala la tendència cap al grau d'IDS "alta", en la part dreta de la figura.

Tanmateix, entenem que la intensitat no és l'únic factor que determina la informació dependent del senyal, sinó que moltes altres variables, exposades ja en l'apartat 1.4, estableixen relacions entre elles i doten al senyal d'una informació extralingüística. No obstant, aquesta informació no serà considerada en aquest estudi i, en tot cas, es mantindrà estable en totes les presentacions del senyal de parla. Per altra banda, pressuposem que no hi ha impediments fisiològics en l'oient, del qual assegurem que disposa d'una bona capacitat auditiva, la qual cosa facilita la integració del senyal acústic.

En relació amb els factors que intervenen en la IIS, ens cal assenyalar que la llista és forçosament molt més llarga, per tal com hi intervenen processos cognoscitius. No obstant hem destriat alguns aspectes que, d'una forma genèrica, incideixen en la comprensió del discurs (vegeu l'apartat 2.12):

- En primer lloc hem distingit entre el discurs construït amb paraules amb significat, conegut per l'oient, i sense significat conegut. En el primer cas hi ha la possibilitat d'accedir a la comprensió del discurs, donat que els diversos components lingüístics (sobretot els mots) són susceptibles de ser predits en la ment de l'oient. En el segon cas, malgrat que la informació provinent del senyal sigui molt elevada, no hi ha possibilitat de desxifrar el missatge, atès que aquest no existeix dins del marc del coneixement que té l'oient.
- En segon lloc, considerem que la informació sobre el context i el referent comunicatius que posseeix el subjecte, i que va completant a mesura que integra el senyal, permet predir els mots que tenen més contingut semàntic dins del discurs.

Si prenem la frase com un marc que es capaç d'establir un context semàntic, hem de suposar que la seva estructura i els components que la integren deuen determinar el grau de predictibilitat dels mots semànticament claus¹⁰. El factor "predicció" és fruit, doncs, d'un conjunt de processos mentals a partir dels quals s'avalua l'encaix d'un determinat nombre de mots "candidats" dins d'un discurs o frase. Per aquest motiu hem fet la gradació de la IIS en tres grans blocs: Les frases que permeten que el mot semànticament clau sigui altament predictable, mitjanament predictable i molt impredecible.

Es important de notar que una frase que estableix un context pel qual un mot resulta molt impredecible es troba classificada dins dels missatges susceptibles de ser compresos, malgrat que la dificultat sigui elevada. Per altra banda, existeixen discursos que no permeten la predicció del mot perquè aquest no té cap contingut semàntic conegut. En el primer cas, la frase estableix un context estructural i semàntic que facilita la predicció del mot clau i del seu significat. En el segon cas és el mot el que s'erigeix com l'element actiu que impedeix tota predicció per falta de significat.

Aquesta gradació comporta que un discurs que facilita que un mot sigui altament predictable estigui catalogat com "lingüísticament poc complex"¹¹. En canvi, un discurs que fa molt impredecibles els mots clau s'ha de considerar com d'una complexitat lingüística molt elevada, encara que la seva comprensió sigui possible.

La taula que resulta de la combinació de tots aquests factors està ressenyada a continuació:

10 .-Vegeu la nota 8 d'aquest mateix capítol. Entenem per *mot-clau* aquell que presenta una informació sintàctica i semàntica rellevant dins d'un enunciat, de tal manera que hi determina la coherència semàntica dels altres mots.

11 .-En la nomenclatura de Benoît i Pols (1992)

SITUACIONS D'INTERRELACIÓ IDS-IIS	INTELLIGIBILITAT ACÚSTICA (Condicions del senyal) (1)	MOTS (2)	LA FRASE DETERMINA QUE EL MOT SIGUI ... (3)	COMPLEXITAT LINGÜÍSTICA PEL QUE FA A LA INFORMACIÓ DEL CONTEXT SEMÀNTIC (4)	RELACIÓ AMB ELS CASOS IDEALS (6)
A	MOLT POBRA	AMB SENTIT	ALTAMENT PREDICTIBLE	BAIXA	PARLA HIPER-ARTICULADA
B	POBRA	AMB SENTIT	ALTAMENT PREDICTIBLE	BAIXA	PARLA HIPER-ARTICULADA
C	MITJANA	AMB SENTIT	ALTAMENT PREDICTIBLE	BAIXA	PARLA HIPER-ARTICULADA
D	MOLT POBRA	AMB SENTIT	MITJANAMENT PREDICTIBLE	MITJANA	PARLA HIPO-ARTICULADA
E	POBRA	AMB SENTIT	MITJANAMENT PREDICTIBLE	MITJANA	PARLA HIPO-ARTICULADA
F	MITJANA	AMB SENTIT	MITJANAMENT PREDICTIBLE	MITJANA	PARLA HIPER-ARTICULADA
G	MOLT POBRA	AMB SENTIT	MOLT IMPREDICTIBLE	ALTA	PARLA HIPO-ARTICULADA
H	POBRA	AMB SENTIT	MOLT IMPREDICTIBLE	ALTA	PARLA HIPO-ARTICULADA
I	MITJANA	AMB SENTIT	MOLT IMPREDICTIBLE	ALTA	PARLA HIPO-ARTICULADA
J	MOLT POBRA	SENSE SENTIT	TOTALMENT IMPREDICTIBLE	NO (5) HI HA MISSATGE	PARLA SENSE SENTIT
K	POBRA	SENSE SENTIT	TOTALMENT IMPREDICTIBLE		PARLA SENSE SENTIT
L	MITJANA	SENSE SENTIT	TOTALMENT IMPREDICTIBLE		PARLA SENSE SENTIT

Taula 3.1.- Relació de la suposada influència que, de forma genèrica, exerceixen els diferents factors de la IDS i de la IIS. Les diverses situacions expressades en aquesta taula estan representades en la figura 3.3

Notes per a la interpretació de la taula:

(1) Considerem una possible variació en el senyal (si més no, pel que fa a la seva intensitat global) i pressuposem unes condicions favorables pel que fa a la locució (suficientment clara) i a l'audició (oients normals en un ambient sense emmascaraments notables). La gradació que hem establert (molt pobra, pobra i suficient) es correlaciona amb el nivell d'intensitat de 33 dB, 36 dB i 39 dB amb què es presentarà el senyal en l'experiment. Això, però, serà explicat extensament en el capítol 4.

(2) Paraules incloses o no en el lèxic de l'oient, el qual es caracteritza com un usuari normal de la llengua en què està construït el discurs. El concepte "usuari normal" es desenvolupa en l'apartat 4.2.4.1 del pròxim capítol.

(3) Pressuposem que les estructures sintàctica i semàntica indiquen la situació i el

significat de la informació rellevant dins de la frase

(4) El context discursiu facilita el desplegament d'un determinat camp semàntic

(5) Impossibilitat d'accedir a un context semàntic pertinent. El subjecte se situa davant d'una seqüència parlada i sense sentit que, a més, es troba en els límits de la qualitat auditiva i amb un context pràcticament nul que no permet de predir el possible missatge. Si es donés un context molt més ric, l'oient podria començar a elaborar hipòtesis preveient possibles mots amb sentit per tal de completar el context (són els casos K i L). Malgrat tot, les expectatives d'aquests casos tampoc no es veuen acomplides perquè falla el sentit del lèxic.

(6) En aquesta columna s'indica la situació de cada una de les interrelacions IDS-IIS respecte de la línia de casos ideals (vegeu la figura 3.3), tot assenyalant, de forma aproximada, la consideració de parla hiperarticulada (*hyper-speech*), hipoarticulada (*hypo-speech*) i senyal de parla sense sentit (*no speech*)

Els casos A, B, C requereixen poca informació dependent del senyal per tal que sigui possible la comprensió (el context facilita la predicció), mentre que en els casos G, H, I requereixen molta més informació dependent del senyal, ja que els elements contextuais són molt pobres. En d'altres paraules, les situacions A, B i C es basen en els continguts dels mots i la relació amb la frase i les G, H i I es basen més en les característiques fonètiques del senyal.

Vistes totes aquestes consideracions, definim la hipòtesi 1 amb dues formulacions consecutives:

a) Els processos d'integració de la informació procedent del senyal estan més condicionats per la possibilitat que té l'oient de predir el contingut semàntic d'aquesta informació que no pel nivell informatiu que aporta directament el senyal

b) En cas de confirmar-se la hipòtesi anterior, la interrelació dels factors que intervenen en els processos de la comprensió de la parla permetria definir com a "parla hiperarticulada" determinades situacions d'interrelació de la informació que, en l'esquema de Lindblom, es consideren "parla hipoarticulada", tot acostant-nos, així, a les consideracions de Benoît i Pols (1992). La verificació d'aquesta interrelació hipotètica dependrà de si aquests factors assoleixen els nivells de funcionament que hem previst en la taula 3.1.

3.3.2.-Consideracions i hipòtesi 2: la relació entre la informació present en el senyal i la informació independent del senyal.

L'esquema de Lindblom (1987b) no precisa quins són els elements que, des d'un punt de vista lingüístic, intervenen en cadascun dels eixos. En conseqüència, tampoc no explica quins tenen un pes més específic per a la finalitat comunicativa i com es produeix la interacció entre tots ells.

Seguint els diversos estadis del procés d'integració de la parla que hem exposat en els capítols anteriors (cap. 1 i 2), hem estructurat una segona hipòtesi a partir dels factors suposadament transcendents de cadascun dels dos eixos. La formulació de la hipòtesi 2 diu així:

Dins de cada estadi de processament del senyal de la parla existeixen interrelacions, més o menys intenses, entre a) els elements informatius continguts en el senyal (IDS) i b) els corresponents analitzadors específics de què disposa l'oient (IIS). En conseqüència, suposem que totes dues vessants estan constituïdes per un nombre determinat de components a través dels quals es concreta la informació rellevant que es troba en el missatge comunicatiu.

En la taula 3.2, presentem el llistat hipotètic d'aquests components, distribuïts en diversos nivells d'accés a la informació.

Nivell d'accés al senyal de la parla	Informació Dependent del Senyal	Informació Independent del Senyal
<p><u>ACCÉS AUDITIU:</u></p> <p>Aspectes físics de caràcter general</p>	<p>-Intensitat global del senyal</p> <p>-Qualitat del senyal: (Relació senyal/soroll)</p>	<p>-Avaluació auditiva del senyal i adaptació posterior</p>
<p><u>ACCÉS FONÈTIC:</u></p> <p>Classes de sons</p>	<p>-Elements de l'inventari fonètic i la seva posició dins de la síl.laba:</p> <p>Vocals Con. Prevocàliques Con. Postvocàliques</p>	<p>-Coneixement de l'inventari fonètic</p> <p>-Coneixement de les restriccions fonotàctiques</p>
<p><u>ACCÉS LÈXIC:</u></p> <p>Tipologia lèxica</p>	<p>-Estructura lèxica del mot-clau dins de la frase (estructura sil.làbica i morfològica.)</p> <p>Longitud del mot-clau:</p> <p>Monosíl.labs Bisíl.labs</p>	<p>-Coneixement, per part de l'oient, del contingut semàntic:</p> <p>Mot amb significat Mot sense significat</p> <p>-Freqüència d'aparició en la llengua</p>
<p><u>ACCÉS CONTEXTUAL:</u></p> <p>Tipologia oracional</p>	<p>Segons l'estructura oracional:</p> <p>Corba d'entonació Pics d'intensitat Límits sintàctics marcats fonèticament Estructura sintàctica¹²</p>	<p>-Segons la predictibilitat del mot clau en funció de la semàntica de la frase, el context pot ser:</p> <p>Neutre Obert Restringit</p>

Taula 3.2.-Seqüenciació discreta d'alguns factors que intervenen en la IDS i en la IIS i que constitueixen variables de l'experiment descrit en el capítol 4.

12 .-En tant que assenyalava la posició del mot-clau o del punt de màxima informació dins del discurs.

La taula 3.2 està organitzada sota el pressupòsit que existeixen uns factors externs que incideixen en el procés d'integració del senyal i que tenen un origen extern, tant respecte del parlant com de l'oient. En tot cas, influeixen de manera global sobre el senyal, sigui quina sigui la seva estructura interna. Aquest factors tenen un caràcter quantitatiu (el nivell d'intensitat global del senyal de parla) i qualitatiu (la relació senyal-soroll, que té una relació directa amb la intel·ligibilitat). Tots dos tipus de factors han estat classificats dins de l'accés auditiu en relació amb la informació aportada pel senyal. No obstant, aquesta classificació cal entendre-la en termes de condicions acústiques més que com una informació pròpiament dita. Per altra banda, tenim present una altra pressuposició sobre els aspectes de tipus ambiental que, en aquest cas, es troben implícits en el propi oient. En primer lloc, l'oient té un coneixement del món i dels referents comunicatius en què es dona el senyal i, en segon lloc, existeix una complicitat entre els interlocutors en tant que el missatge forçosament conté un significat coherent i lògic dins del context en què ha estat emès i no és només un seguit de paraules. En cas contrari la situació comunicativa sempre es troba per sota de la línia dels casos ideals de la figura 3.3.

Tornant a l'estructura de la taula, l'hem organitzada entenent que cadascun dels nivells d'accés comporta una doble vessant en relació amb els processos *top-down* i *bottom-up*. Per una banda existeix una informació tangible en el senyal que ascendeix de baix cap a dalt. És la informació que es pot descriure en termes lingüístics i que configura l'estructura interna del senyal. Per altra banda és necessària la intervenció de l'oient en cadascun d'aquests nivells, tant pel que fa al reconeixement de cadascun dels aspectes de la IDS, com pel que fa a la capacitat de desxifrar i d'interpretar el significat global que es pot extreure del senyal (processament de dalt cap a baix).

Per altra banda, els factors que configuren tant la IDS com la IIS no interactuen en una relació "d'un-a-un", sinó que dins de cada eix, els factors es troben relacionats com si formessin una xarxa, de manera que un element d'un dels eixos no pot considerar-se aïlladament, sinó integrat en una estructura més àmplia, amb múltiples interrelacions entre ells. Per exemple, la variable 'vocal', cal considerar-la com contextualitzada en una estructura sil·làbica, la qual, a la vegada, es troba integrada en un mot d'un nombre determinat de síl·labes.

Pel que fa als elements que presentem dins de l'eix de la IDS, cal aclarir que són el resultat d'una selecció determinada. Som conscients que no són tots els factors possibles. No obstant, i davant de la dificultat de definir-los tots, hem cregut convenient de fer una anàlisi d'aquests que hem escollit amb el benentès que calia tenir-ne presents uns quants davant de la dificultat enorme de considerar-los tots. En aquest sentit, en la taula esmentem, també, alguns aspectes que clarament completen el plantejament de la hipòtesi, però que no poden ser contemplats, posteriorment, ni en l'experiment ni en l'anàlisi de les dades.

En la segona vessant (IIS), els factors que la integren es pressuposen com formant part de la competència lingüística dels parlants coneixedors de la llengua i usuaris habituals. Així, tant les restriccions fonotàctiques, com el coneixement de l'inventari fonètic, com la capacitat d'avaluar fefaentment el significat dels mots i el seu contingut semàntic es consideren implícits en el propi oient.

Pel que fa a la classificació i avaluació de la pertinència del context semàntic per a

cadascuna de les frases i del lèxic que les integren, l'hem simplificat en tres tipus de situacions: la que es concreta a partir d'un context neutre, d'un context obert i d'un context restringit. Totes tres situacions són descrites a continuació.

3.3.3.-Consideracions i hipòtesi 3: el paper que juguen els diferents tipus de context.

A partir del que acabem de definir en la segona hipòtesi, podem suposar l'existència de diversos contextos que incideixen en el procés d'integració del senyal i de comprensió del missatge. Per una banda, els anomenats contextos estructurals i, per una altra, els contextos semàntics.

Els primers es concreten en la major part dels aspectes exposats en la columna de la Informació dependent del senyal de la taula 3.2. Segurament es pot establir, en cadascun d'aquests aspectes pròpiament estructurals (fonològics, sintàctics, etc.), els casos que possibiliten una millor integració del senyal de la parla. Per altra banda, els contextos semàntics determinen el marc en el qual els diferents components lèxics estableixen una relació de significat que dona coherència a tota l'elocució i permet que el seu contingut sigui comprès. Els elements responsables de generar aquest marc de referència són els diferents mots que integren la frase. Sanford i Garrod (1981), entre d'altres, van desenvolupar una explicació sobre el procés de comprensió del discurs que es coneix amb el nom de 'Models mentals o escenaris'. Segons aquests autors, a partir de la informació que es fa present per mitjà d'una paraula, s'activen diversos camps semàntics (tòpics) amb els seus corresponents camps lèxics. En el moment que el mot es relaciona amb les paraules del context i desactiva tots els camps semàntics menys un, es verifica un sol camp lèxic i es concreta una única possibilitat de comprensió.

Tot aquest processament del discurs es produeix de forma continuada i l'avaluació del contingut del senyal proporciona informació 'sobre la marxa' (*on line*), és a dir, a mesura que es van integrant el significat de les paraules en el suposat context que s'ha desplegat en la ment de l'oient (Van Dijk & Kintsch, 1983).

En el funcionament natural d'aquests processos es poden produir alteracions importants en la relació continuada entre el contingut del lèxic i el context semàntic. En aquests casos és molt probable que s'accedeixi a avaluar i a considerar la informació provinent del senyal i, per tant, a tenir en compte uns nivells més baixos d'integració del senyal.

Moltes poden ser les circumstàncies per les quals la comprensió no es pot dur a terme amb totes les garanties d'èxit. De totes maneres, les raons principals es poden emmarcar, en el context del nostre experiment, en tres grans causes:

- a.-perquè el senyal no pot proporcionar tota la informació fonètica,
- b.-perquè la paraula clau no té significat conegut per a l'oient
- c.-perquè el context no estableix suficientment el marc semàntic del qual parlàvem en els paràgrafs anteriors.

Tots aquests casos es troben sistematitzats en la taula 3.3 que presentem a continuació.

	FRASES AMB MOTS AMB SENTIT			FRASES AMB MOTS SENSE SENTIT		
	CONTEXT NEUTRE	CONTEXT OBERT	CONTEXT RESTRINGIT	CONTEXT NEUTRE	CONTEXT OBERT	CONTEXT RESTRINGIT
ACCÉS ACTIU	Fonètic Lèxic *	Fonètic Lèxic *	Fonètic Lèxic Context oracional	Fonètic * *	Fonètic * *	Fonètic * Context oracional
ACCÉS SEMIACTIU	* * *	* * Context oracional	* * *	* * *	* * Context oracional	* * *
ACCÉS INACTIU	* * Context oracional	* * *	* * *	* Lèxic Context oracional	* Lèxic *	* Lèxic *

Taula 3.3.-Relació entre els nivells d'accés per a cada tipus de context establert en l'experiment durant el processament del senyal

Els diferents contextos que es presenten en la taula 3.3 fan referència a les possibles restriccions que efectua una frase o un discurs sobre un o diversos camps semàntics. Aquestes restriccions, com ja s'ha dit més amunt, són de tipus semàntic, ja que relacionen el significat global del discurs amb el dels mots-clau. D'acord, doncs, amb la funció que exerceix el context el podem definir com neutre, obert i restringit.

Entenem per context neutre aquell que admet qualsevol mot, sigui quin sigui el seu significat. El resultat d'un discurs amb un context neutre és considerat per l'oient com un discurs vàlid, malgrat que hi hagi múltiples possibilitats de combinació i, per tant, múltiples candidats lèxics provinents, a la vegada, de múltiples camps lèxics. El nivell de restricció lèxica

és zero, tant pel que fa a la seva forma com al seu contingut. En aquest context la integració de les paraules-clau estarà eminentment basada en la forma.

El context obert delimita el possible accés de candidats lèxics a uns quants camps, de manera que els elements que es troben en una frase exerceixen uns condicionants semàntics sobre les paraules-clau que apareixen al llarg del discurs.

Finalment, el context restringit estableix un únic camp semàntic actiu del qual procediran els candidats lèxics per a ocupar la posició de les paraules-clau. Els elements que apareixen al llarg del discurs exerceixen una reducció molt important sobre les possibilitats lèxiques per a travar i completar el discurs.

A més sabem que el processament del senyal s'estableix com una tasca destinada, en darrer terme, a l'extracció de la informació que conté (Lindblom, 1983a). Aquest tipus d'informació es basa en un reconeixement dels mots-clau a partir del context semàntic de l'enunciat, la qual cosa permet de generar prediccions sobre el significat final del missatge. No obstant, degut que es poden presentar missatges amb paraules desconegudes per l'oient, aquest ha d'afrontar-les a partir de la informació de tipus fonètic, que és l'única informació de què disposa. Aleshores, suposem que les estratègies d'integració de la parla s'organitzen segons si els enunciats contenen paraules amb sentit (integrables per comparació amb les paraules que es troben en el lèxic de l'oient) o paraules sense sentit (integrables només per processos fonètics). D'acord amb tot aquest plantejament pressuposem, en aquesta hipòtesi 3, que en els enunciats compostos de mots amb sentit s'activen procediments d'anàlisi del senyal de la parla diferents als que s'utilitzen quan els enunciats estan formats per mots sense sentit.

3.4.- Conclusions

En aquest capítol hem presentat el marc conceptual sobre el qual hem desenvolupat l'experiment d'aquest treball: La teoria de la parla hiper i hipoarticulada, que fou desenvolupada per B. Lindblom durant els anys 1987 i 1996.

Aquesta teoria cal emmarcar-la en la discussió sobre l'existència d'elements invariants en el senyal de parla. Lindblom, davant de fenòmens com la coarticulació o els estils de parla, remarca que la variabilitat és intrínseca de la parla. Així, els trets invariants es remeten a un pla fonològic i, per tant, teòric. Per contra, suggereix l'existència d'altres fonts d'informació que es troben a disposició del receptor per tal de facilitar la percepció i comprensió del missatge comunicatiu. Segons diu aquest autor, la informació que aporta el senyal de parla es combina amb la informació de què disposa el subjecte sobre la situació comunicativa o el coneixement del tema i la capacitat de relacionar-ho amb altres coneixements.

Fruit d'aquesta interrelació és possible adaptar-se a un senyal per extreure'n la informació, encara que les condicions del senyal estiguin deteriorades. Això planteja l'existència d'una adaptació constant entre parlant i oient de manera que es pugui transmetre una informació amb la màxima economia fònica, sense que se'n perdi la possibilitat de comprensió per part del receptor. La relació entre la informació que aporta el senyal i la capacitat de l'oient per a integrar-lo (fruit del seu coneixement del món) estableix dos extrems

anomenats "parla hiperarticulada" i "parla hipoarticulada". La línia divisòria entre totes dues circumstàncies s'anomena "línia de casos ideals", on la possibilitat d'integració informativa és òptima amb la màxima economia fònica.

Posteriorment a la formulació d'aquesta teoria s'han desenvolupat algunes modificacions per part d'altres autors, com ara Benoît i Pols (1992), en el sentit que la línia de casos ideals podria variar en funció de tres factors: el llindar acústic, la complexitat del missatge i la comprensió del missatge per sota de la línia de casos ideals de l'esquema de Lindblom.

Vist tot aquest marc teòric, en el present treball volem accedir a comprovar alguns aspectes d'aquesta teoria i les seves modificacions, centrant-nos especialment en els casos en els quals el senyal és presentat al receptor en condicions de baixa intensitat, propera al llindar de l'audició, per tal d'estudiar els mecanismes que aquest utilitza per tal d'adaptar-s'hi. Per tal d'abordar aquest enfocament, hem dissenyat tres hipòtesis que poden ajudar a validar alguns dels plantejaments que se'n desprenen. Aquestes hipòtesis es defineixen, de forma resumida, com segueix:

Hipòtesi 1 a) Els processos d'integració de la informació procedent del senyal estan més condicionats per la possibilitat que té l'oient de predir el contingut semàntic d'aquesta informació que no pel nivell informatiu que aporta directament el senyal.

b) En cas de confirmar-se la hipòtesi anterior, la interrelació dels factors que intervenen en els processos de la comprensió de la parla permetria definir com a "parla hiperarticulada" determinades situacions d'interrelació de la informació que, en l'esquema de Lindblom, es consideren "parla hipoarticulada", tot acostant-nos, així, a les consideracions de Benoît i Pols (1992).

Hipòtesi 2: Les dues vessants que existeixen dins de cada estadi de processament del senyal de la parla estan constituïdes per un nombre determinat de components a través dels quals es concreta la informació rellevant que es troba en el missatge comunicatiu.

Hipòtesi 3: En els enunciats compostos de mots amb sentit s'activen procediments d'anàlisi del senyal de la parla diferents als que s'utilitzen quan els enunciats estan formats per mots sense sentit.

A partir d'aquestes hipòtesis s'ha dissenyat l'experiment que es presenta, a continuació, en el capítol 4.

4.-DISSENY EXPERIMENTAL

4.1.-Aspectes generals del disseny experimental

A l'igual que en altres sistemes complexos, l'estudi de la parla com un sistema comunicatiu planteja el dilema de si cal abordar-lo des d'un punt de vista global, cosa que resulta o molt difícil o necessàriament superficial, o si és millor aïllar aspectes parcials que permeten una aproximació més detallada i profunda, però, també, més deslligada de la totalitat del procés comunicatiu.

Dit en d'altres paraules, el fet de separar variables interrelacionades en les situacions de percepció natural de la parla podria arribar a proporcionar resultats comportamentals poc fiables o parcialment distorsionats de la realitat.

L'opció metodològica que hem adoptat en aquest estudi intenta que el tractament de les variables, fora del seu estat natural sigui interpretat, en la discussió dels resultats, dins d'una visió global del sistema. De fet, si tenim en compte el postulat de Lindblom pel qual el senyal és constantment adaptat a fi que hi hagi comunicació entre l'emissor i el receptor, hem de suposar, d'entrada, que en l'acte comunicatiu es produeixen un gran nombre d'interrelacions entre tots els elements que hi intervenen. En el nostre cas, només fixarem l'atenció en el paper que juguen en el receptor, i analitzarem quins són els mecanismes que posa a disposició en cadascun dels estadis perceptius per tal que s'assoleixi la comprensió del missatge.

Donat que en el model de la H & H no es plantegen problemes de comprensió del missatge davant de parla hiperarticulada, nosaltres enfocarem l'experiment des de l'altre costat, és a dir, des de la reacció que provoca un senyal deficient que s'aproxima, amb més o menys mesura, a la línia d'equilibri dels casos ideals, representada en la figura 3.1.

4.1.1.- Objectius i metodologia experimental

Tal com hem exposat en el capítol 2, en els diversos nivells en què s'organitza el procés d'integració de la parla per a la seva percepció i comprensió es realitzen múltiples operacions. Habitualment no es tracta d'operacions simples amb una única tasca sinó de complexes interrelacions entre els diferents analitzadors que avaluen la cadena fònica que ha arribat al subjecte perceptor. Totes aquestes anàlisis són efectuades diverses vegades, de forma reiterada, fins que se n'extreu la informació dins de cada estadi. I aquesta informació encara és comparada amb la dels altres estadis i amb la que es trobava inicialment a disposició del subjecte. Aquesta xarxa tan complexa obliga a determinar una estratègia d'anàlisi que permeti de descriure i de preveure el funcionament de cadascun dels suposats nivells de processament.

De fet, la determinació del nombre de nivells i la categoria de cadascun d'ells no és, ni de bon tros, ni evident ni delimitada. Per tant, el punt de partida del disseny experimental del present treball s'estableix a partir dels dos grans eixos de l'esquema de Lindblom presentat en el capítol 3, subdividits encara en diversos aspectes per a cadascun d'ells.

Així, doncs, l'anàlisi dels resultats obtinguts haurà de versar al voltant dels nivells relacionats amb:

1.-La informació dependent del senyal

2.-La informació procedent de la interrelació entre la informació dependent i la independent del senyal.

Donat que el procés d'integració de la parla s'inicia de baix cap a dalt (*bottom-up*)¹, cal considerar la informació dependent del senyal com una informació delimitada i autònoma, almenys a l'inici del procés. Això, però, no pot ser aplicat de la mateixa manera a la informació independent del senyal, ja que aquesta, en el subjecte receptor, sempre està condicionada per la primera. Per aquest motiu els dos nivells sobre els quals es dissenya l'experiment no són equiparables en l'anàlisi. Per tant, en el segon cas s'estableix com a objectiu l'anàlisi de la interrelació entre els dos tipus d'informació (dependent i independent), ja que la informació independent del senyal no pot ser tractada de forma autònoma, tal com s'estableix per a la dependent.

Per altra banda, a fi de distingir si la percepció del senyal s'ha realitzat mitjançant l'ús de la informació continguda estrictament en el senyal o si s'ha utilitzat informació de tipus semàntic, hem organitzat l'anàlisi diferenciant dos tipus d'integració del senyal: la identificació i el reconeixement².

La identificació implica la percepció dels trets fonètics d'un so o d'una cadena fònica i comporta que el subjecte de la prova repeteixi correctament tots els sons que formaven part de l'estímul presentat.

En canvi, el reconeixement es pot definir com la percepció dels trets fonètics d'una cadena fònica tot assignant-li el contingut semàntic que li correspon en l'ús de la llengua. Evidentment, implica la valoració correcta del sentit de la paraula per part del subjecte de la prova.

La identificació es concreta quan la resposta verbal que emet el subjecte davant de l'estímul coincideix completament amb els sons que se li han presentat. En el recompte d'aquestes ocurrències no es té en compte la completa comprensió del missatge per part del subjecte i, per tant, s'ignoren les possibles confusions produïdes en altres nivells del procés d'integració del senyal de la parla. Per altra banda, el reconeixement comporta que la identificació fonètica dels sons que hagi realitzat el subjecte estigui corroborada amb la correcta assignació del sentit del mot clau, cosa que s'obté a partir de la resposta que emet el subjecte per escrit.

Concretant el que acabem de dir, i d'acord amb aquesta estratègia d'estudi, hem

1 .- Tal com hem explicat en el segon capítol, aquest procés d'integració és progressiu de baix cap a dalt i, evidentment, l'origen del procés és físic i s'inicia a partir de la informació continguda en el senyal (vegeu apartat 2.1 i 2.2).

2 .- Cal no confondre aquesta nomenclatura amb els termes "reconeixement de paraules" i "accés al lèxic" que hem presentat en el capítol 2 i que són d'aplicació exclusiva al procés d'integració del lèxic. En el nostre cas, es prenen en un sentit més genèric, tal com s'explica en la seva definició.

determinat que el procés experimental s'organitzi en tres grans objectius, que es resumeixen en les taules presentades a continuació:

OBJECTIU 1: IDENTIFICACIÓ FONÈTICA

TIPUS D'INFORMACIÓ	Informació dependent del senyal
NIVELLS EN EL PROCÉS	Auditiu i fonètic
FINALITAT	Establir el grau d'identificació fonètica de les vocals en relació amb diferents factors que formen part de l'estructura del senyal i determinar la possible interacció entre ells, tant amb paraules amb sentit com sense sentit
SUBOBJECTIU 1	Resistència de la identificació fonètica de la Vocal Tònica (V_T) a la gradació d'intensitat (rampa d'intensitat), en mots amb sentit i sense sentit.
SUBOBJECTIU 2	Resistència de la identificació fonètica de la Vocal Tònica (V_T) a la gradació d'intensitat en paraules monosil·làbiques i paraules bisil·làbiques amb sentit i sense sentit.
SUBOBJECTIU 3	Resistència de la identificació fonètica de la Vocal Tònica (V_T) a la gradació d'intensitat amb estructures sil·làbiques CV - CVC - CCV - CCVC en mots amb sentit i sense sentit.
SUBOBJECTIU 4	Resistència a la confusió dels trets fonètics i les característiques articulatòries dels sons vocàlics i consonàntics.

OBJECTIU 2: IDENTIFICACIÓ FONÈTICA I RECONeixEMENT LÈXIC

TIPUS D'INFORMACIÓ	Informació dependent i independent del senyal
NIVELLS EN EL PROCÉS	Fonètic i lèxic
FINALITAT	Establir la influència del sentit del mot en la identificació fonètica del mot sencer, en funció dels diferents components de la informació dependent del senyal, tot determinant la possible interacció entre ells.
SUBOBJECTIU 1	Resistència de la identificació fonètica de la paraula sencera a la gradació d'intensitat, en mots amb sentit i sense sentit.
SUBOBJECTIU 2	Resistència del reconeixement lèxic a la gradació d'intensitat, en mots amb sentit i sense sentit.
SUBOBJECTIU 3	Resistència a la confusió fonètica de les paraules durant la integració lèxica

OBJECTIU 3: IDENTIFICACIÓ FONÈTICA I RECONeixEMENT LÈXIC

EN FUNCió DEL CONTEXT

TIPUS D'INFORMACIó	Informació dependent i independent del senyal
NIVELLS EN EL PROCÉS	Lèxic i contextual
FINALITAT	Establir una jerarquia dels tipus de context que afavoreixen identificacions correctes en paraules amb sentit i en paraules sense sentit i determinar la possible interrelació entre ells.
SUBOBJECTIU 1	Avaluació de la influència dels diferents tipus de context en el reconeixement de paraules amb sentit, corroborada semànticament (resposta del subjecte sobre el sentit del mot), en funció de la gradació d'intensitat .
SUBOBJECTIU 2	Avaluació de la influència dels diferents contextos en la identificació de paraules sense sentit, corroborada semànticament (resposta del subjecte sobre el sentit del mot), en funció de la gradació d'intensitat.

4.1.2.-Esquema experimental

Tal com hem dit més amunt, tenim interès en tractar les variables des d'un punt de vista global. No obstant, i per una raó pragmàtica, hem organitzat l'experiment en tres fases, a fi que puguem controlar amb més efectivitat els factors que incideixen en els processos de percepció i de comprensió de la parla. La separació del model experimental en fases, no necessàriament ha de proporcionar resultats aïllats per a cadascuna d'elles, sinó que uns esdevindran la base per a la interpretació dels altres.

La fase prèvia representa un primer contacte amb el comportament de les diverses variables que intervenen en la percepció de la parla. A partir dels resultats s'ha de corregir el disseny de l'experiment per tal d'assolir l'objectiu final. Els condicionants sobre els quals es presta més atenció són per una banda la intensitat i la qualitat del senyal, i per una altra les característiques auditives del subjecte i la seva capacitat de reacció a l'estímul. La preparació de l'experiment i l'estudi previ de les seves condicions en determinaran la correcta aplicació.

FASE PRÈVIA

- Coneixement de les característiques auditives dels subjectes mitjançant proves audiomètriques.
- Calibratge dels aparells d'anàlisi.
- Prova pilot d'enregistrament i de confecció i aplicació de la prova.

La prova pilot d'aquesta fase prèvia pressuposà un mínim disseny experimental que fou avaluat després d'aplicar-lo a diversos subjectes. A partir dels resultats d'aquesta avaluació s'elaborà el disseny experimental definitiu. La informació que s'extragué de la prova pilot fou essencial per a la construcció de tot el disseny, perquè permeté evidenciar el comportament particular d'algunes variables que, en un principi, es menystenien i que resultaren ser significatives. Per altra banda, es pogué corregir la dinàmica de la sessió experimental, tot calculant els temps per a cada part de la prova.

FASE PRIMERA (enunciats sense sentit)

Determinació dels factors que intervenen en la identificació fonètica:

- La intensitat intrínseca mitjana de les vocals tòniques en les paraules del corpus.
- El llindar d'identificació fonètica (llindar d'intensitat necessari per a la discriminació dels sons consonàntics i vocàlics en enunciats sense sentit).
- La influència de l'estructura sil·làbica en el procés d'identificació (llindar d'intensitat necessari per a la discriminació lèxica en cada estructura).

Confecció i aplicació de la prova d'identificació fonètica

Anàlisi dels resultats

En la primera fase es pretén esbrinar quina és la intensitat mínima necessària perquè un so vocàlic o una síl·laba siguin percebuts i reconeguts. Donat que tot oient percep en funció

d'allò que coneix i pot identificar, el material parlat utilitzat aquí no té significat en la llengua catalana. D'aquesta manera s'elimina la variable de l'expectativa per part del subjecte receptor.

La identificació fonètica dels sons que componen les síl·labes ens acosta molt més al procés d'identificació fonètica que es produeix en la parla natural que si s'haguessin utilitzat al·lòfons aïllats.

La segona fase representa un intent d'aproximació a la comprensió tenint en compte la intervenció del receptor en tant que cal la seva participació (els seus coneixements, la seva capacitat de relació dels mots i els seus significats, etc.) per a la interpretació del missatge. Des d'un altre punt de vista, també s'intenta descobrir la intensitat necessària per a la comprensió de mots amb sentit conegut pels subjectes. En aquesta fase els subjectes hauran de confirmar l'existència del mot en el seu vocabulari personal i si això els permet d'establir un context que doni significació al mot en qüestió.

FASE SEGONA
(enunciats amb sentit)

Determinació dels factors que intervenen en la comprensió de la parla:

-Aproximació al llindar de comprensió en la qual hi intervé la selecció lèxica de paraules conegudes i la seva integració dins de frases marc coherents.

Confecció i aplicació de les proves de reconeixement lèxic i contextual.

Anàlisi dels resultats.

En aquesta fase cal tenir present la notable importància d'alguns aspectes de caire psicolingüístic que analitzarem més endavant.

Arran d'aquestes 2 fases experimentals es poden estudiar, d'una banda, els nivells d'intensitat auditiva necessaris per a la percepció de la "informació dependent del senyal" i, per una altra, es pot realitzar una aproximació als aspectes contextuais i de participació del receptor que configuren la "informació independent del senyal", seguint el model de Lindblom exposat en l'apartat 3.1.

4.2.-Configuració general de l'experiment

En el disseny i realització de l'experiment s'ha tingut molt en compte de tractar separatament cadascun dels estadis consecutius que hi intervenen, de manera que, per una banda, es pugui contrastar l'esquema de la percepció i la comprensió de la parla exposats més amunt (en el capítol 2) i, per una altra, sigui possible de tenir el màxim control sobre les variables experimentals.

Els diversos passos del procés es podrien esquematitzar tal com es presenta a la figura 4.1, on s'assenyalen els diversos estadis que s'ha seguit en l'elaboració i realització de l'experiment:

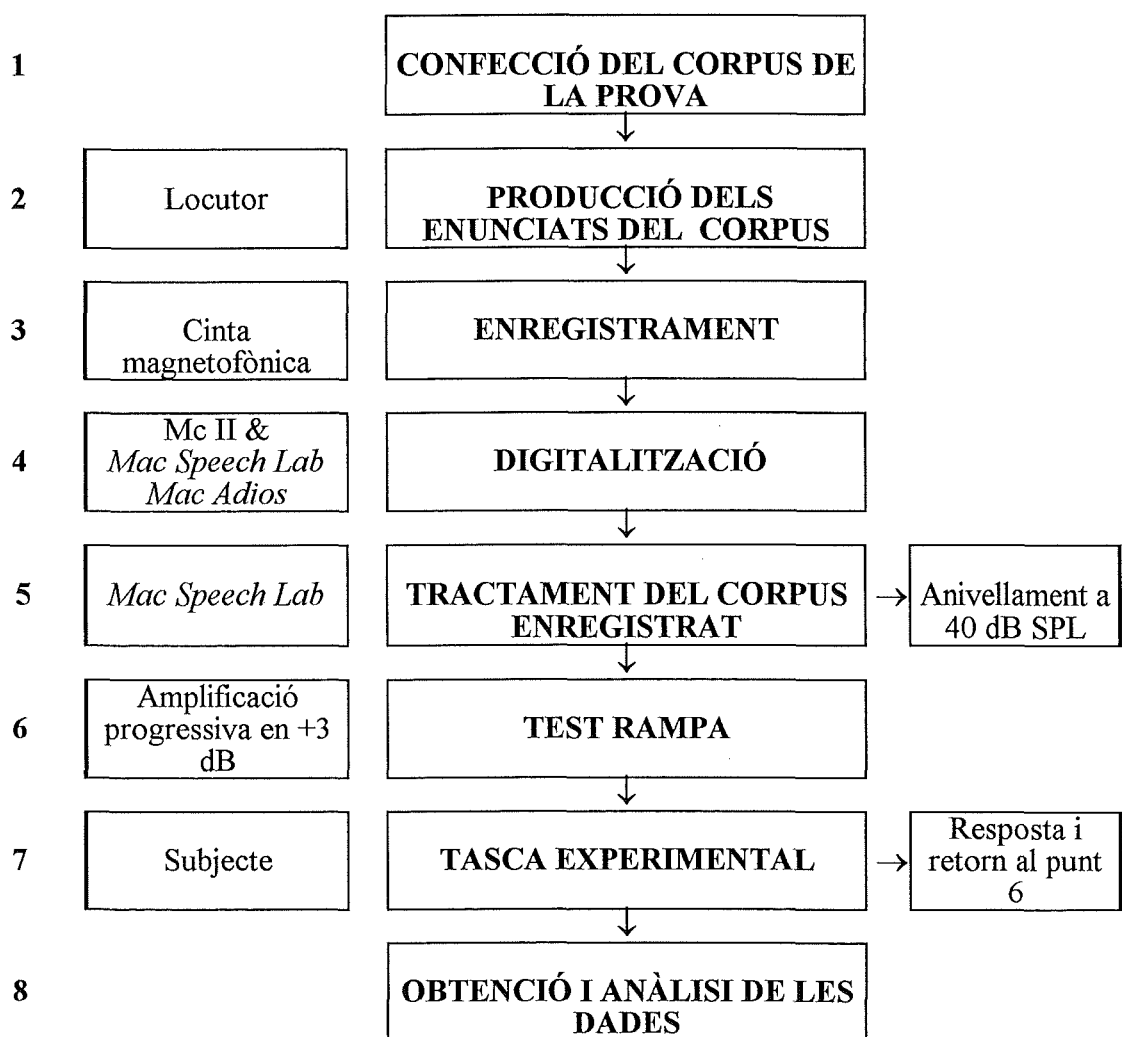


Figura 4.1: Representació de les diverses etapes del disseny experimental

A continuació es presenten i s'expliquen detalladament cadascuna d'aquestes etapes.

4.2.1.-Confecció del corpus de la prova

El material verbal utilitzat com a estímul en la prova consta d'un conjunt de frases pronunciades per un parlant en un entorn acústic òptim, amb les característiques del que s'anomena habitualment parla de laboratori³. En un primer moment, es va pensar en confeccionar la prova a partir de les produccions espontànies d'un parlant, però hauria resultat molt difícil d'obtenir tot el ventall de característiques estructurals requerides per l'experiment i, segurament, s'hi haurien afegit altres condicionants externs (sobretot els elements acústics que s'afegeixen al senyal i l'emascaren) que haurien fet molt difícil de tractar-la de manera controlada. Per aquest motiu, i sabent que el tipus de material verbal que s'utilitzaria tindria un component d'artificialitat, es va optar per confeccionar els estímuls de parla de manera que aquests fossin estructurats segons uns determinats criteris, que més endavant descriurem.

Per una altra banda, era important de tenir un corpus de paraules sense sentit que permetés l'estudi del procés pel qual l'oient accedeix a un estadi estrictament fonètic. També calia aconseguir un altre corpus de paraules, fàcilment identificables, que s'utilitzarien en l'estudi del procés d'accés al lèxic.

Tots dos corpus, un cop elaborats, han format part d'una sola prova i s'han presentat al subjecte de manera conjunta, però en la seva confecció hi han intervingut factors diferents. Per aquest motiu es fa necessari de tractar-los per separat. Els anomenem corpus de paraules amb sentit i sense sentit, respectivament.

4.2.1.1.-El corpus de paraules amb sentit

El corpus de paraules per a l'estudi del procés d'accés al lèxic està confeccionat a partir de mots d'ús habitual que poden trobar-se en el vocabulari bàsic d'una persona de cultura mitjana. Per a la seva selecció s'ha fet ús dels vocabularis bàsics de la Llengua Catalana⁴ i s'ha tingut cura de no incloure en el corpus, paraules que no poguessin formar part d'aquests repertoris. La composició d'aquest corpus consta de dues llistes de mots amb sentit. L'una, amb paraules monosil·làbiques i, l'altra, amb paraules bisil·làbiques planes.

En aquestes darreres, i per tal que no hi hagués un nombre excessiu de variables incontrolades, es va creure necessari de centrar l'atenció en la primera síl·laba, la tònica, i no tenir en compte la vocal amb què acaba cada mot. Així, doncs, a fi d'homogeneïtzar el llistat d'aquestes paraules, es va creure oportú adoptar com a criteri que totes elles acabessin en vocal neutra [ə]. D'aquesta manera va quedar controlada una possible variable d'influència de la segona vocal en el procés d'identificació del mot.

Pel que fa a l'estructura de la síl·laba tònica, en tots els casos comença en consonant, i aproximadament un 50% de monosíl·labs i un 75% de bisíl·labs estan formats per síl·labes obertes. Les estructures sil·làbiques utilitzades en el corpus són:

CV, CCV, CVC, CCVC(C)

3.-Sobre les característiques de la "parla de laboratori" i les diferències amb la "parla espontània", es pot consultar Lindblom (1987b)

4.-Vegeu Llobera (1968), CELC (1978), Mas *et al.* (1992)

Totes les síl.labes àtones també tenen estructura CV.

En un estudi sobre l'estructura sil.làbica del català central, Yzaguirre (1991) aporta la freqüència absoluta d'aparició de les diferents estructures a partir d'una mostra de vora 2 milions i mig de síl.labes. Els percentatges resultants són els següents:

Segment inicial de síl.laba

	V	CV	CCV	TOTAL
V	6,53	48,00	4,83	59,36
VC	9,23	27,08	1,45	37,77
VCC	0,44	1,08	0,19	2,63
VCCC	0,01	0,20	0,00	0,22
TOTAL	16,23	77,27	6,48	100,00

Segment final de la síl.laba

Taula 4.1.- Percentatge absolut d'aparició de les estructures sil.làbiques del català central (Extret d'Yzaguirre 1991).

Tal com apareix en aquesta taula, les síl.labes seleccionades per a formar part de les paraules del corpus, enquadrades amb doble línia, representen el 81,36 % del total d'estructures sil.làbiques. De fet, el conjunt d'aquestes quatre estructures sil.làbiques és molt significatiu dins del conjunt de les estructures que apareixen en el començament de paraula, ja que les començades per vocal són un 16,23% del total.

En conseqüència, doncs, hem confeccionat el corpus de la prova a partir d'un doble criteri: a) seleccionant les paraules que comencen amb consonant, ja que són més freqüents que les que comencen amb vocal i, b) entre les síl.labes que comencen amb consonant, triant per al corpus aquelles estructures que apareixen en la taula 4.1 com a més freqüents. Així, doncs, la distribució de les estructures sil.làbiques a l'hora de confeccionar les llistes s'ha fet de manera que resultin el màxim d'equilibrades, sobretot pel que fa a les consonants o grups consonàntics inicials i a la vocal tònica. Per aconseguir aquest equilibri hem elaborat la llista de paraules mantenint les estructures sil.làbiques que apareixen en els registres lingüístics més habituals, tot

respectant les proporcions expressades en la taula 4.1.

El corpus de paraules amb sentit, doncs, ha estat confeccionat a partir de la freqüència d'aparició de les consonants de la varietat estàndard del català oriental en la posició inicial de paraula, tant si van seguits de la vocal sil·làbica ("posició inicial aïllada") com si entre totes dues hi apareix una altra consonant ("posició inicial agrupat primer element")⁵

La mostra lèxica que utilitzada per obtenir la distribució d'aquests sons consonàntics fou el Diccionari manual castellà-català català-castellà (1974) atès que, pel fet de ser un diccionari manual, aconseguí reunir el lèxic més habitual de l'usuari de la llengua (vocabulari mitjà) i no inclou les paraules de registres lingüístics diferents al de l'estàndard ⁶. Per altra banda, tots els mots inventariats en aquest diccionari s'acompanyen de la seva transcripció fonètica, la qual cosa proporciona un criteri en la tasca de considerar l'estàndard fonètic dels mots del català oriental.

A partir d'aquest recull, s'ha fet la revisió de 18712 mots, que són tots aquells que comencen per consonant. A continuació s'han seleccionat tots els que presentaven la vocal tònica a la primera síl·laba de paraula, que resultaren ser 2.727 mots. A partir d'aquesta darrera reducció s'ha fet l'índex de freqüència d'aparició de consonants en les combinacions CV i CCV.

D'entrada no s'han inclòs en el càlcul d'aquesta distribució els sons [tʃ] i [dʒ], atès que la realització d'aquests sons es deriven a principi de mot per fonètica sintàctica i no es troben en el sistema fonemàtic estàndard del català ⁷. També s'han rebutjat els diftongs creixents tòncics (del tipus #[j]+V o del tipus #[kw]+V) per tal de simplificar l'estructura consonant + vocal desestimant, així, una variable afegida que és la presència de les semiconsonants.

Pel que fa als grups consonàntics a inici de mot, s'han tingut en compte totes les combinacions C+[r]+V i C+[l]+V. Concretament s'han revisat les paraules amb síl·laba inicial tònica que començaven per [pr, tr, kr, br, dr, gr, fr, pl, kl, bl, gl, fl]. El resultat del nombre de paraules que reuneixen les condicions anteriors es mostra en la taula 4.2. També s'hi han calculat el total de paraules que concorren per a cadascuna de les diferents consonants i grups consonàntics (columna 1), el nombre global de paraules que comencen per cada consonant (columna 2), el nombre total de mots seleccionats de cada consonant respecte el total de mots amb síl·laba inicial tònica i estructura CV o CCV (columna 3) i, finalment, el nombre total de mots de cada consonant respecte el total de mots del diccionari amb consonant inicial (columna 4). Al davall de cada columna es totalitzen les aparicions de cada vocal com a nucli de les síl·labes inicials tòniques precedides de consonant.

5.- Aquesta nomenclatura és la utilitzada a Rafel (1979)

6.- Badia i Margarit, autor del pròleg d'aquest diccionari, parla de la reducció de lèxic que s'hi ha operat i reconeix el següent criteri: "... hom ha procurat de retenir els mots més corrents i els indispensables per als lectors corrents d'avui" (pàg. XVI del pròleg.). Sobre la repercussió de la freqüència d'ús de les paraules en la predisposició dels subjectes a reconèixer-les vegeu l'estudi que han realitzat García Albea *et alt.* (1982).

7.-López del Castillo (1976), p. 249. Encara que ell mateix reconeix que, en la norma de força parlants del dialecte oriental, la realització dels fonemes fricatius a principi de mot sovint s'esdevé en al·lòfons africats.

								1	2	3	4
	[a]	[e]	[ɛ]	[i]	[o]	[ɔ]	[u]	Suma total	Mots presents en el diccionari	% de la columna 1 sobre 2727 mots	% de la columna 2 sobre 18712 mots
[p]	49	18	36	26	19	47	23	218	2278	8	12
[t]	42	16	37	29	30	25	11	190	1332	7	7
[k]	133	2	2	18	24	101	19	299	3144	11	17
[b]	64	42	44	33	30	44	18	275	1918	10	10
[d]	18	8	16	13	11	16	10	92	2118	3	11
[g]	40	0	4	7	15	5	1	72	824	3	4
[f]	40	16	20	16	22	17	18	149	1118	5	6
[s]	44	13	37	29	66	33	10	232	1314	8	7
[ʃ]	16	2	3	13	1	2	5	42	188	2	1
[z]	1	1	5	2	1	0	0	10	30	0	0
[ʒ]	7	10	14	4	6	7	11	59	317	2	2
[r]	47	35	21	14	14	22	18	171	1546	6	8
[l]	17	3	2	17	1	6	6	52	251	2	1
[λ]	22	18	6	15	6	12	12	91	384	3	2
[m]	119	17	27	22	23	29	23	260	1554	9	8
[n]	19	9	11	12	5	20	11	87	385	3	2
[ɲ]	5	0	0	1	1	0	0	7	11	0	0
[pr]	3	11	9	4	6	15	1	49		2	
[tr]	21	5	16	16	5	12	7	82		3	
[kr]	5	5	12	11	3	5	1	42		2	
[br]	12	4	5	6	7	9	5	48		2	
[dr]	5	0	2	2	1	1	1	12		0	
[gr]	13	7	5	6	2	4	6	43		2	
[fr]	9	4	6	3	2	0	2	26		1	
[pl]	20	1	7	0	2	3	2	35		1	
[kl]	20	1	4	4	0	6	1	36		1	
[bl]	6	1	4	0	1	1	1	14		1	
[gl]	4	0	1	0	1	7	1	14		1	
[fl]	7	2	3	1	1	5	1	20		1	
	808	251	359	324	306	454	225	2727	18712	100	100

Taula 4.2.- Nombre absolut de síl.labes inicials tòniques i distribució de les consonants inicials i de les vocals nucli de síl.laba (extret del Diccionari manual castellà-català català-castellà, 1974)

Per tal de validar aquestes dades i poder avaluar el grau de representativitat de cadascun dels casos respecte al que Badia i Margarit denomina ús quotidià de la llengua, es va creure oportú de fer una comparació amb altres estudis que s'han realitzat sobre la freqüència

d'aparició de les síl·labes en el conjunt dels mots de la llengua catalana. Concretament farem referència a dos treballs: Cerdà (1970) i Rafel (1979). Donat que tots dos presenten les dades en percentatges, hem fet els càlculs adients sobre les dades de la taula anterior i en presentem els resultats a continuació, a la taula 4.3:

	VOCALS							(1)	(2)	(3)
	[a]	[e]	[ɛ]	[i]	[o]	[ɔ]	[u]	% TOTAL PRESENT ESTUDI	% TOTAL RAFEL (1979)	Diferència entre els dos treballs
[p]	0.26	0.10	0.19	0.14	0.10	0.25	0.12	1.17	1.38	-0.21
[t]	0.22	0.09	0.20	0.15	0.16	0.13	0.06	1.02	0.86	0.16
[k]	0.71	0.01	0.01	0.10	0.13	0.54	0.10	1.60	2.15	-0.55
[b]	0.34	0.22	0.24	0.18	0.16	0.24	0.10	1.47	0.95	0.52
[d]	0.10	0.04	0.09	0.07	0.06	0.09	0.05	0.49	3.04	-2.55
[g]	0.21	0.00	0.02	0.04	0.08	0.03	0.01	0.38	0.19	0.19
[f]	0.21	0.09	0.11	0.09	0.12	0.09	0.10	0.80	0.63	0.17
[s]	0.24	0.07	0.20	0.15	0.35	0.18	0.05	1.24	1.39	-0.15
[ʃ]	0.09	0.01	0.02	0.07	0.01	0.01	0.03	0.22	0.08	0.14
[z]	0.01	0.01	0.03	0.01	0.01	0.00	0.00	0.05	0.00	0.05
[ʒ]	0.04	0.05	0.07	0.02	0.03	0.04	0.06	0.32	0.15	0.17
[r]	0.25	0.19	0.11	0.07	0.07	0.12	0.10	0.91	0.40	0.51
[l]	0.09	0.02	0.01	0.09	0.01	0.03	0.03	0.28	1.93	-1.65
[λ]	0.12	0.10	0.03	0.08	0.03	0.06	0.06	0.49	0.23	0.26
[m]	0.64	0.09	0.14	0.12	0.12	0.15	0.12	1.39	1.30	0.09
[n]	0.10	0.05	0.06	0.06	0.03	0.11	0.06	0.46	0.59	-0.13
[ɲ]	0.03	0.00	0.00	0.01	0.01	0.00	0.00	0.04	0.00	0.04
[pr]	0.02	0.06	0.05	0.02	0.03	0.08	0.01	0.26	----	-----
[tr]	0.11	0.03	0.09	0.09	0.03	0.06	0.04	0.44	----	-----
[kr]	0.03	0.03	0.06	0.06	0.02	0.03	0.01	0.22	----	-----
[br]	0.06	0.02	0.03	0.03	0.04	0.05	0.03	0.26	----	-----
[dr]	0.03	0.00	0.01	0.01	0.01	0.01	0.01	0.06	----	-----
[gr]	0.07	0.04	0.03	0.03	0.01	0.02	0.03	0.23	----	-----
[fr]	0.05	0.02	0.03	0.02	0.01	0.00	0.01	0.14	----	-----
[pl]	0.11	0.01	0.04	0.00	0.01	0.02	0.01	0.19	----	-----
[kl]	0.11	0.01	0.02	0.02	0.00	0.03	0.01	0.19	----	-----
[bl]	0.03	0.01	0.02	0.00	0.01	0.01	0.01	0.07	----	-----
[gl]	0.02	0.00	0.01	0.00	0.01	0.04	0.01	0.07	----	-----
[fl]	0.04	0.01	0.02	0.01	0.01	0.03	0.01	0.11	----	-----
TOTAL	4.34	1.34	1.92	1.73	1.65	2.43	1.21	14.57	16.07	-1.5

Taula 4.3.- Percentatge d'aparició de les consonants sobre els 18712 mots inventariats i comparació amb les dades de Rafel (1979) ⁸

8 .-Dels dos estudis esmentats anteriorment, només el de Rafel presenta la distribució de freqüències de les unitats consonàntiques a principi de mot. Per aquest motiu, només apareix la columna de dades referents a aquest estudi.