

University of California
Berkeley

College of Engineering
Department of Electrical Engineering
and Computer Sciences

Professors : N.Morgan / B.Gold
EE225D

Spring, 1999

Speech Perception

Lecture 18

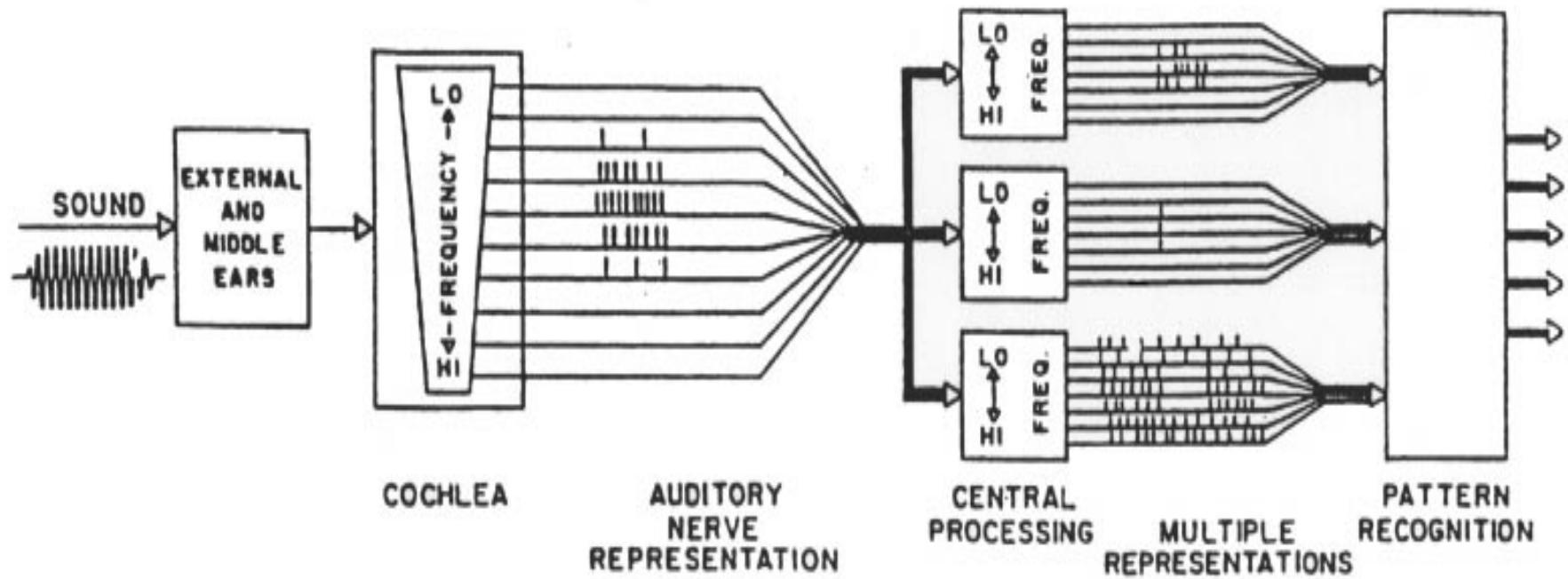


Figure 17.1 : Block diagram of sound representation in the auditory system.

Speech Perception

Production ————— Perception
 Physiology Psychophysics

Consonant Perception

Consonants are more important than vowels.

Cnsnnts r mr mprtnt thn vwls
 ooa ae oe ioa a oe

	p	t	k	f	θ	s	ʃ	b	d	g	v	ð	z	ʒ	m	n
p	240		41	2	1											
t	1	252	1	1						i						
k	18	3	219													
f				225	24			5			2					
θ	9		1	69	185			3				1				
s						232										
ʃ							236									
b					1			242			24	12	1			
d									213	22			1			
g					1				33	203		3				
v								6			171	30				1
ð					1			1		3	22	208	4			!
z									2	4	1	7	238			
ʒ														244		
m															274	1
n																252

Figure 17.5 : Confusion matrix for S/N=+12dB and frequency response of 200-6500Hz.

	p	t	k	f	θ	s	ʃ	b	d	g	r	θ	z	ʒ	m	n
p	80	43	64	17	14	6	2	1	1			1			2	
t	71	84	55	5	9	3	8	1				1	2		2	3
k	76	76	107	12	8	9	4					1			1	
f	18	12	9	175	48	11	1	7	2	1	2	2				
θ	19	17	16	104	64	32	7	5	4	5	6	4	5			
s	8	5	4	23	39	107	45	4	2	3	1	1	3	2		1
ʃ	1	6	3	4	6	29	195		3							1
b	1			5	4	4		136	10	9	47	16	6	1	5	4
d							8	5	80	45	11	20	20	26	1	
g					2			3	63	66	3	19	37	56		3
r				2		2		48	5	5	145	45	12		4	
θ					6	1	1	31	6	17	86	58	21	5	6	4
z					1	1	1	7	20	27	16	28	94	44		1
ʒ								1	26	18	3	8	45	129		2
m	1							4			4	1	3		177	46
n					4			1	5	2		7	1	6	47	163

Figure 17.5 : Confusion matrix for S/N=-6dB and frequency response of 200-6500Hz.

Condition	S/N	Band
1	-18	200 - 6500
2	-12	200 - 6500
3	- 6	200 - 6500
4	0	200 - 6500
5	6	200 - 6500
6	12	200 - 6500
7	12	200 - 300
8	12	200 - 400
9	12	200 - 600
10	12	200 - 1200
11	12	200 - 2500
12	12	200 - 5000
13	12	1000 - 5000
14	12	2000 - 5000
15	12	2500 - 5000
16	12	3000 - 5000
17	12	4000 - 5000

Figure 17.6 : Seventeen conditions of S/N and filtering.

Miller Nicely System

Voicing

s - z

Nasality

m, n, ng, ——— other Sounds

Affrication

fricatives ——— m,n

Duration

p, b

Place of Articulation

[p, k, t
b, g, d

PHONEMES

Features	[m	n	v	ð	z	ʒ	ʒ̂	b	d	g	w	r	l	j	f	θ	s	ʃ	ʃ̂	p	t	k	h	
Voicing	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	-	-	-	-	-	-	-	-	-
Nasality	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Sustention	-	-	+	+	+	+	-	-	-	-	+	+	+	+	+	+	+	+	-	-	-	-	-	+
Sibilation	-	-	-	-	+	+	+	-	-	-	-	-	-	-	-	-	+	+	+	-	-	-	-	-
Graveness	+	-	+	-	-	○	○	+	-	○	+	-	○	○	+	-	-	○	○	+	-	○	○	○
Compactness	-	-	-	-	-	+	+	-	-	+	-	-	○	+	-	-	-	+	+	-	-	+	+	+

Stimulus Words used in the DRT

VOICING

Voiced---Unvoiced

veal---feel
 bean---peen
 gin---chin
 dint---tint
 zoo---Sue
 dune---tune
 voal---foal
 goat---coat
 zed---said
 dense---tense
 vast---fast
 gaff---calf
 vault---fault
 daunt---taunt
 jock---chock
 bond---pond

NASALITY

Nasal---Oral

meat---beat
 need---deed
 mitt---bit
 nip---dip
 moot---boot
 news---dues
 moan---bone
 note---dote
 mend---bend
 neck---deck
 mad---bad
 nab---dab
 moss---boss
 gnaw---daw
 mom---bomb
 knock---dock

SUSTENTION

Sustained---Interrupted

vee---bee
 sheet---cheat
 vill---bill
 thick---tick
 foo---pooh
 shoes---choose
 those---doze
 though---dough
 then---den
 fence---pence
 than---Dan
 shad---chad
 thong---tong
 shaw---chaw
 von---bon
 vox---box

Stimulus Words used in the DRT

SIBILATION

Sibilated---Unsibilated

zee---thee
 cheep---keep
 jilt---gilt
 sing---thing
 juice---goose
 chew---coo
 Joe---go
 sole---thole
 jest---guest
 chair---care
 jab---dab
 sank---thank
 jaws---gauze
 saw---thaw
 jot---got
 chop---cop

GRAVENESS

Grave---Acute

weed---reed
 peak---teak
 bid---did
 fin---thin
 moon---noon
 pool---tool
 bowl---dole
 fore---thor
 met---net
 pent---tent
 bank---dank
 fad---thad
 fought---thought
 bond---dong
 wad---rod
 pot---tot

COMPACTNESS

Compact---Diffuse

yield---wield
 key---tea
 hit---fit
 gill---dill
 coop---poop
 you---rue
 ghost---boast
 show---so
 keg---peg
 yen---wren
 gat---bat
 shag---sag
 yawl---wall
 caught---taught
 hop---fop
 got---dot

Place	p	k	t	b	d	g	f	thin	s	sh	v	the	z	azure	m	n	ng	l	r	w	h
bilabial	+	-	-	+	-	-	-	-	-	-	-	-	-	-	+	-	-	-	-	+	-
labiodental	-	-	-	-	-	-	+	-	-	-	+	-	-	-	-	-	-	-	-	-	-
dental	-	-	-	-	-	-	-	+	-	-	-	+	-	-	-	-	-	-	-	-	-
alveolar	-	-	+	-	+	-	-	-	+	-	-	-	+	-	-	+	-	+	-	-	-
palatal	-	-	-	-	-	-	-	-	-	+	-	-	-	+	-	-	-	-	+	-	-
velar	-	+	-	-	-	+	-	-	-	-	-	-	-	-	-	-	+	-	-	-	-
pharyngeal	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+
Manner																					
glide	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	+	+	-
nasal	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	+	+	-	-	-	-
stop	+	+	+	+	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
fricative	-	-	-	-	-	-	+	+	+	+	+	+	+	+	-	-	-	-	-	-	-
voicing	-	-	-	+	+	+	-	-	-	-	+	+	+	+	+	+	+	+	+	+	+

Figure 17.7 : Binary distinctive feature set of Jakobson et al.

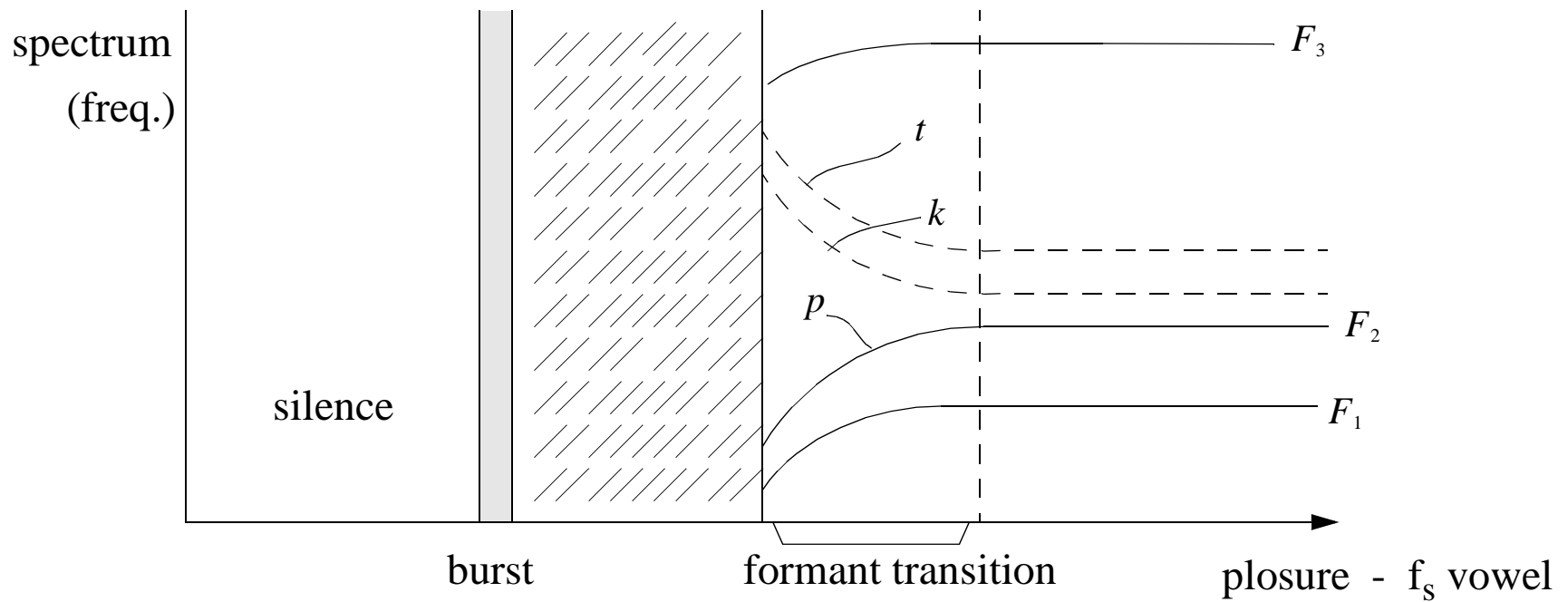
Jacobson, Fant, Halle.

Production

Perception

Haskin's Lab.

Pattern Playback



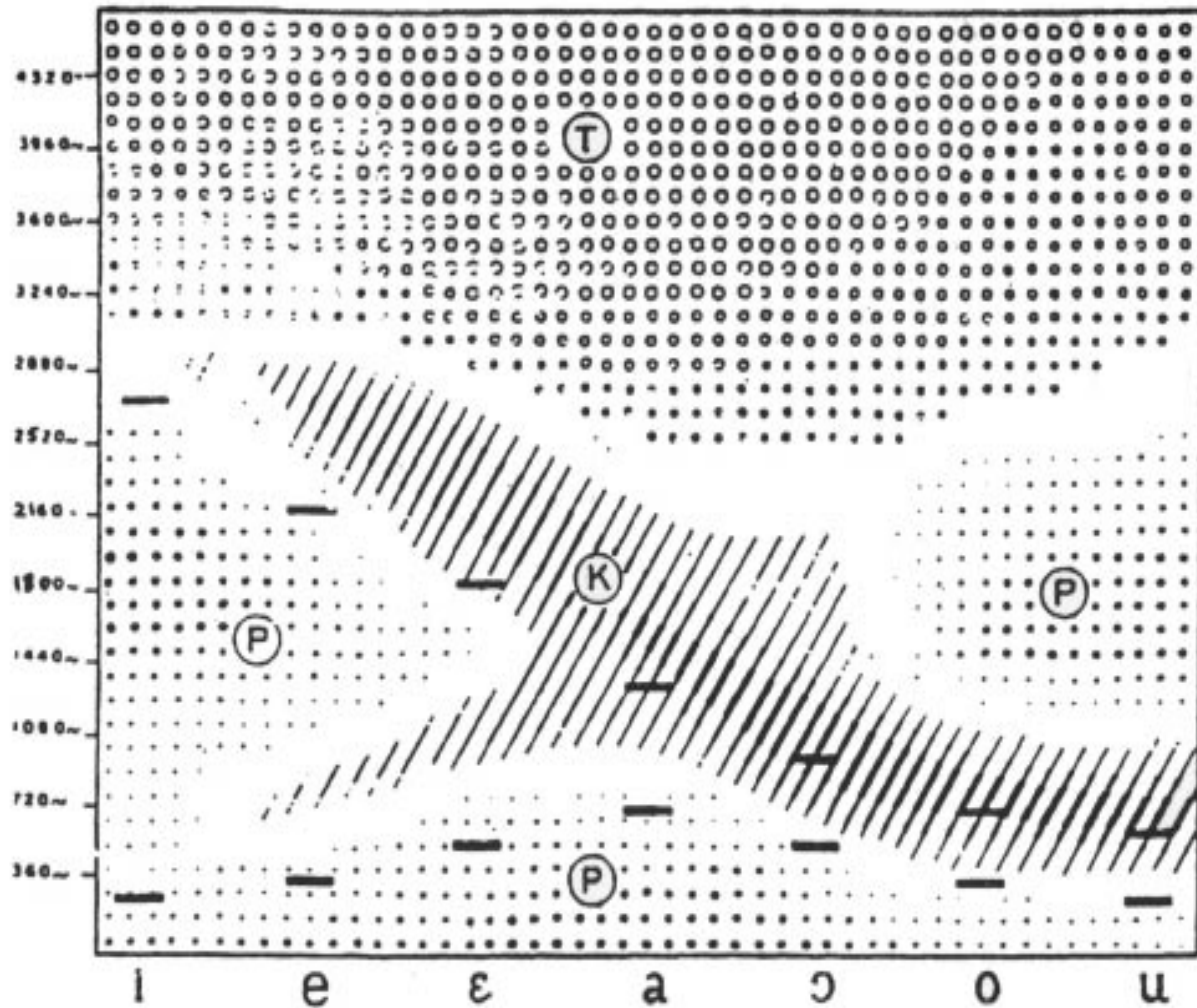


Figure 17.9 : Perceptual responses to different burst frequencies and vowels for the voiceless plosives.

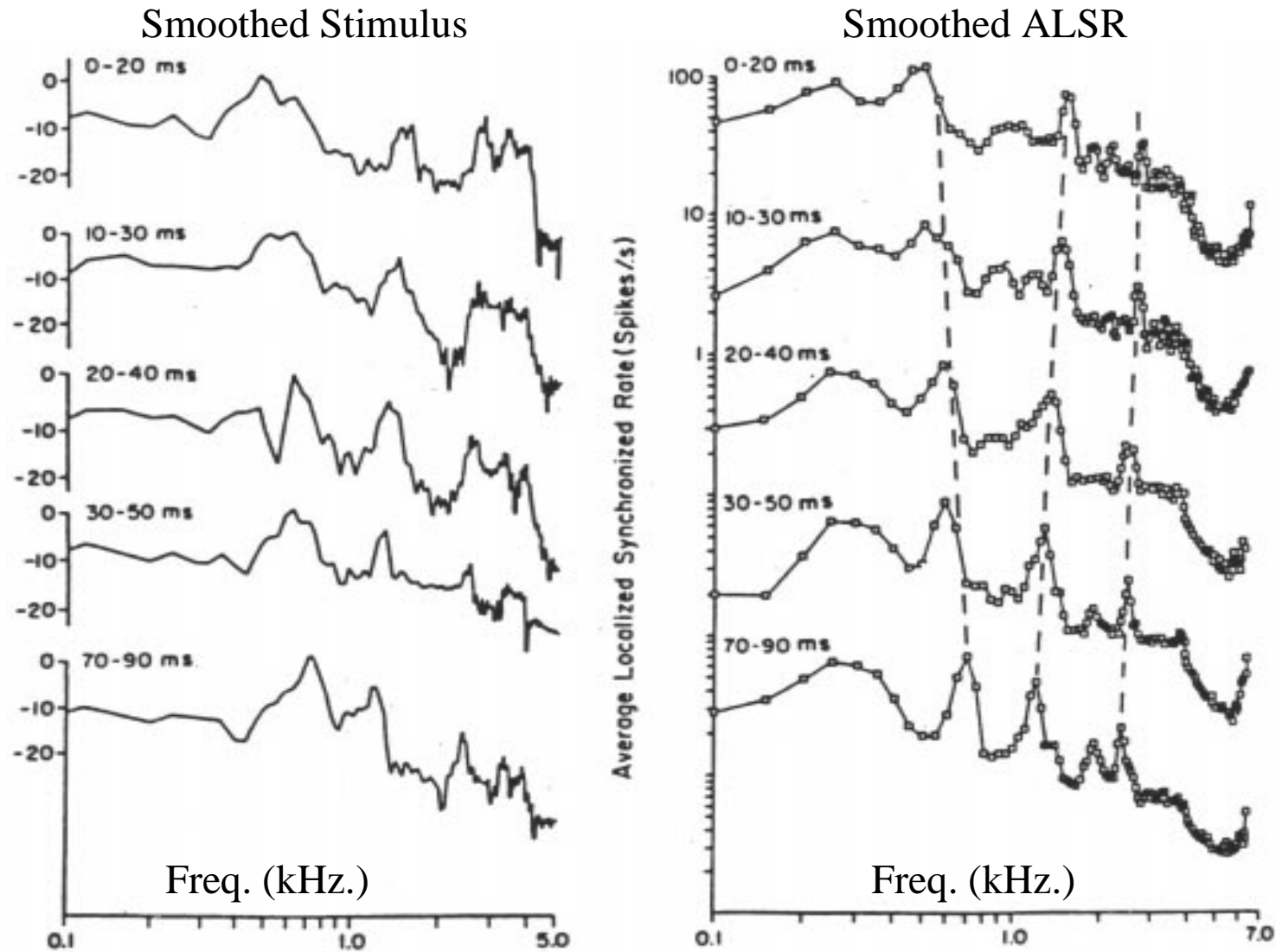


Figure 17.11 : Smoothed spectra and corresponding ALSR's for the fine 20ms. Intervals at the beginning of the - da - stimulus.

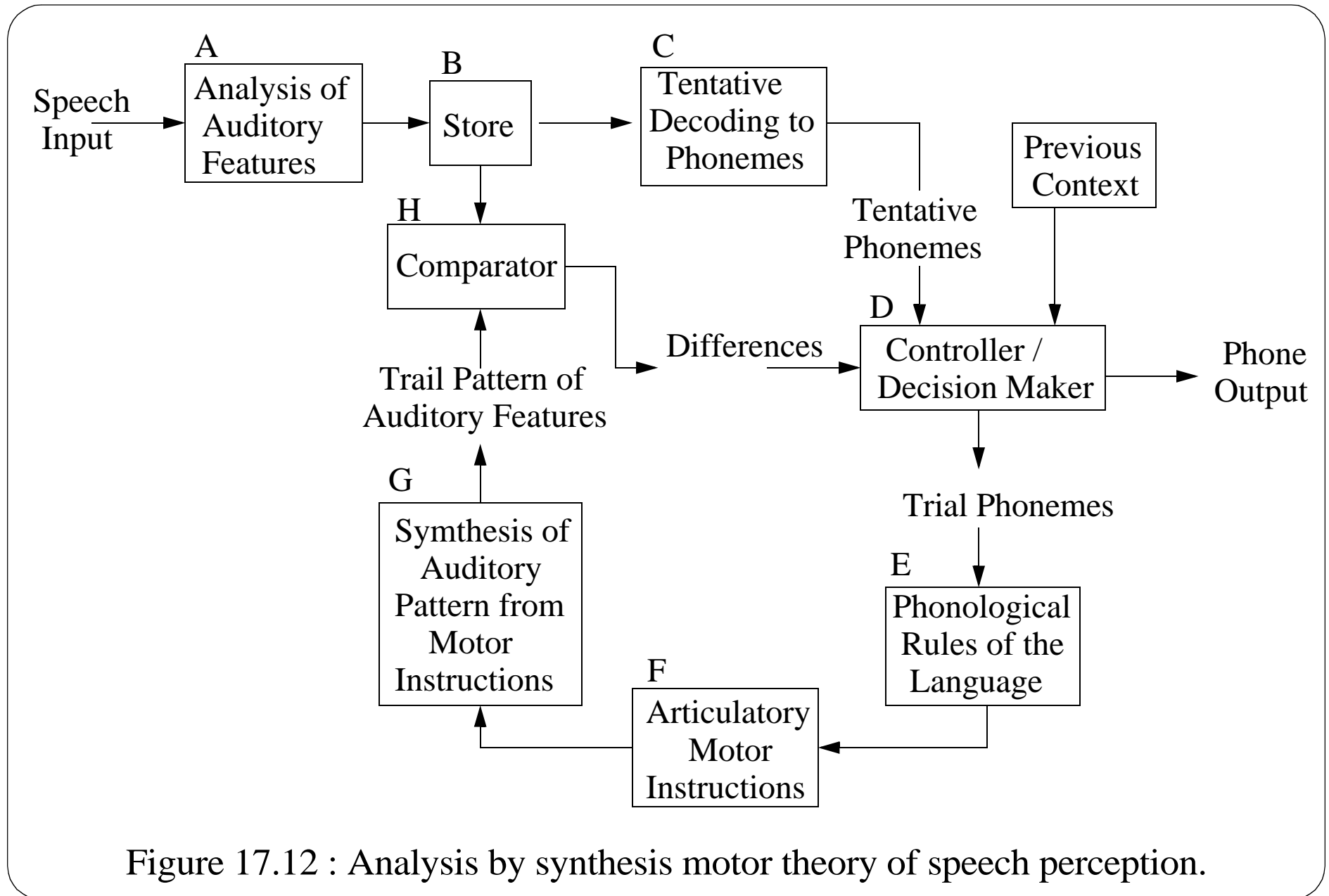


Figure 17.12 : Analysis by synthesis motor theory of speech perception.

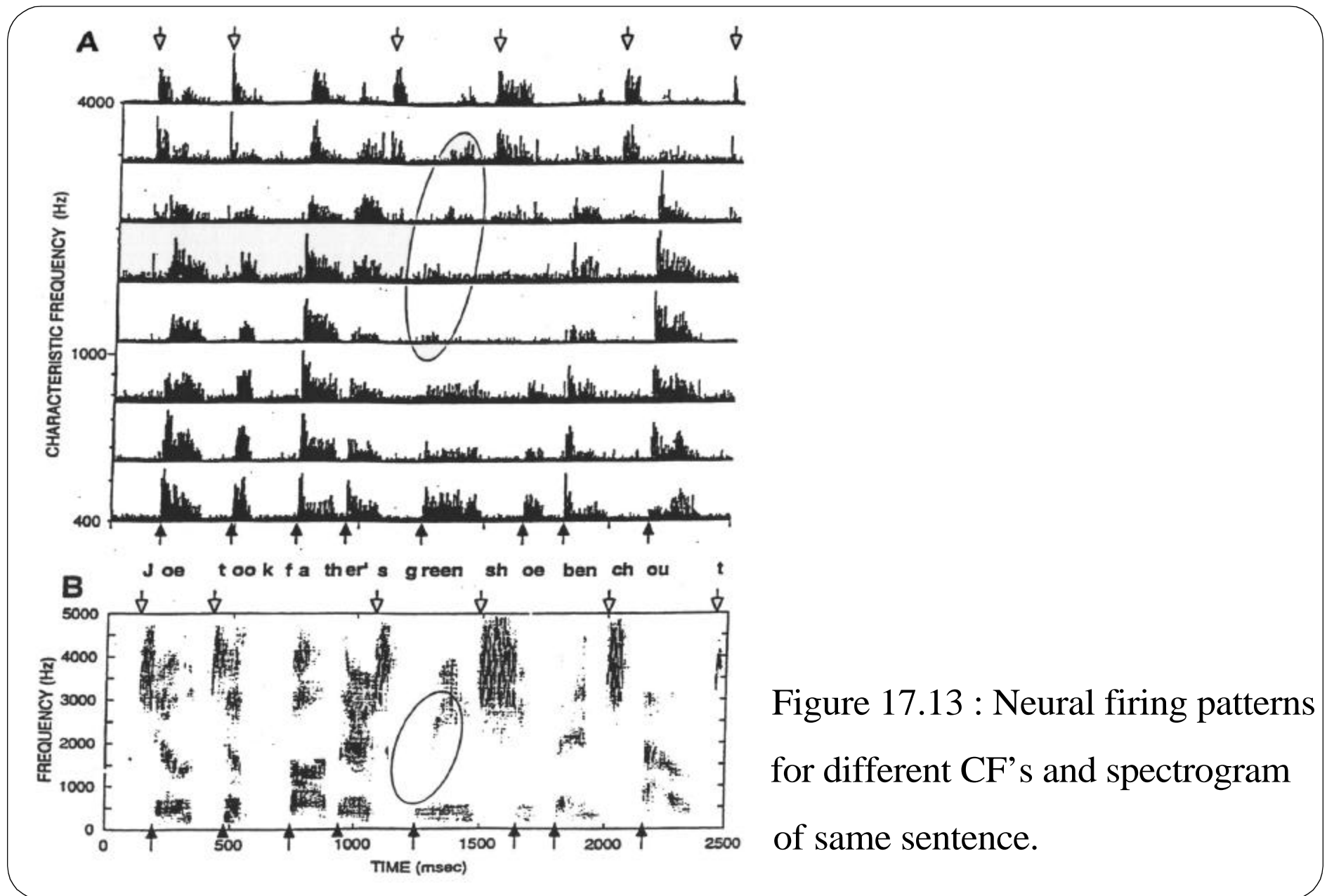


Figure 17.13 : Neural firing patterns for different CF's and spectrogram of same sentence.