

6.551J/HST714J ACOUSTICS OF SPEECH AND HEARING

Lecture ~~25~~ 24

12/07/04

Features for consonants in English

	p	t	k	b	d	g	f	v	s	z	š	ž	č	ǰ	θ	ð	m	n	ŋ	w	y
cont	-	-	-	-	-	-	+	+	+	+	+	+	-/+	-/+	+	+	-	-	-		
strid.	-	-	-	-	-	-	+	+	+	+	+	+	+	+	-	-	-	-	-		
lips	+			+			+	+									+				
blade		+			+				+	+	+	+	+	+	+	+		+			
body			+			+													+		
ant		+			+				+	+	-	-	-	-	+	+		+			
nas	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	+	+		
voic	-	-	-	+	+	+	-	+	-	+	-	+	-	+	-	+					

Table 1. Articulator-free features for some consonants in English

	t,d	s,z	θ,ð	n
continuant	-	+	+	-
sonorant	-			+
strident		+	-	

Table 2. Seven articulators and the features that specify phonetic distinctions that can be made with each articulator. The articulators are divided into three groups: those that can form a constriction in the oral cavity, those that control the shape of the airway in the laryngeal and pharyngeal regions, and the aspect of vocal-fold adjustment relating to stiffness.

lips	[round]
tongue blade	[anterior] [distributed] [lateral] [rhotic]
tongue body	[high] [low] [back]
soft palate	[nasal]
pharynx	[advanced tongue root]
glottis	[spread glottis] [constricted glottis]
vocal folds	[stiff vocal folds]

voicing:

+: b d g v ð z ʒ ĵ
-: p t k f θ s š č

coronal (blade)

+: t d θ ð s z š ʒ č ĵ n
-: p k b g f v m ŋ

strident

+: s z š ʒ č ĵ f v
-: t d g ŋ θ ð m n p k b

English plural

<u>əz</u>	<u>s</u>	<u>z</u>
kiss	hat	cab
bush	cake	cad
fez	cap	cog
rouge	cuff	cave
church		quay
judge		eye
		ham
		hen

Nasal Assimilation

land lamp laŋk

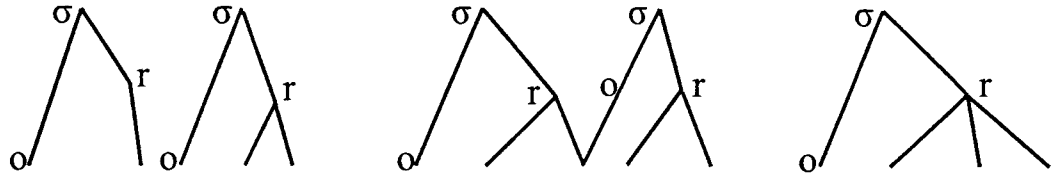
Suffix /jən/

commune	communion
hell	hellion
	million

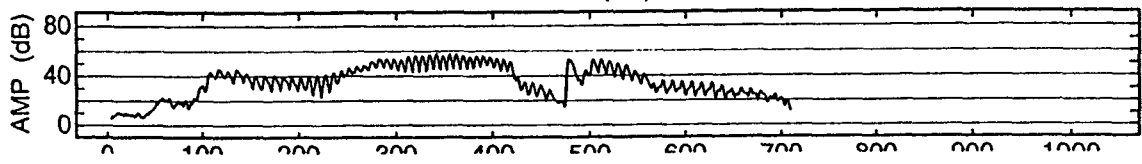
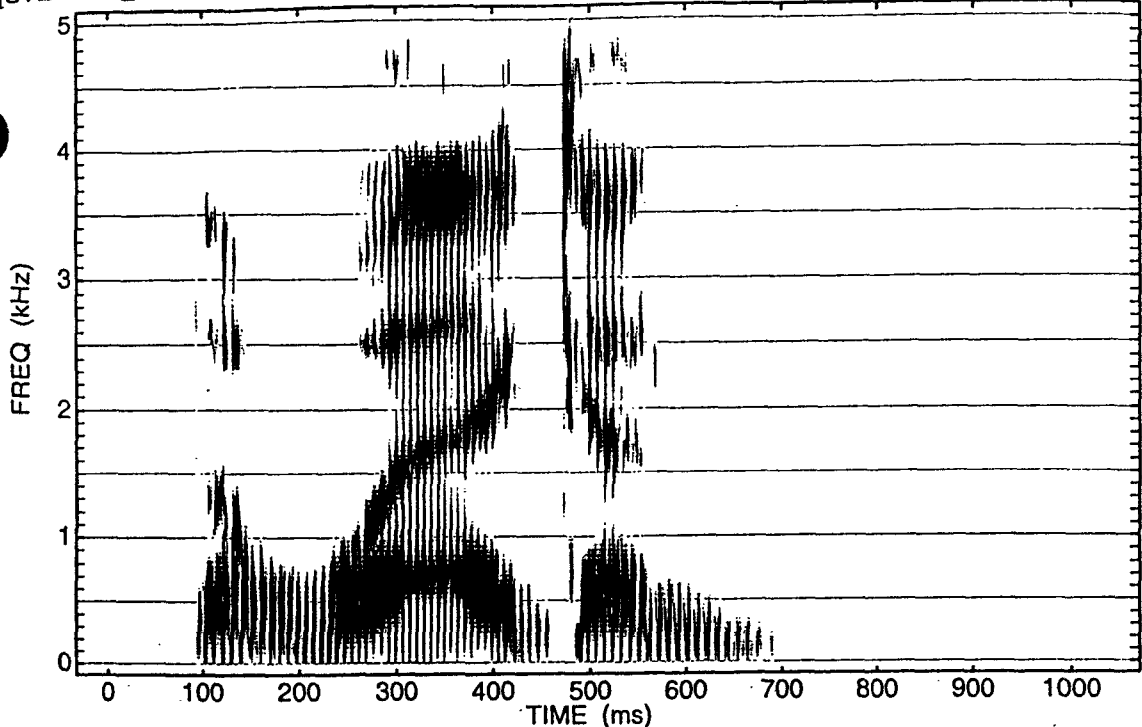
confess + jən	confession
deride	derision
contrite	contrition
Egypt	Egyptian
divide	division

4

Table 5. Lexical representations for the words help, debate and wagon. The syllable structure of each word is schematized at the top (σ = syllable, o = onset, r = rime).

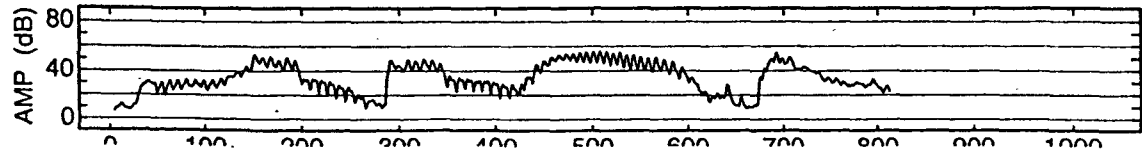
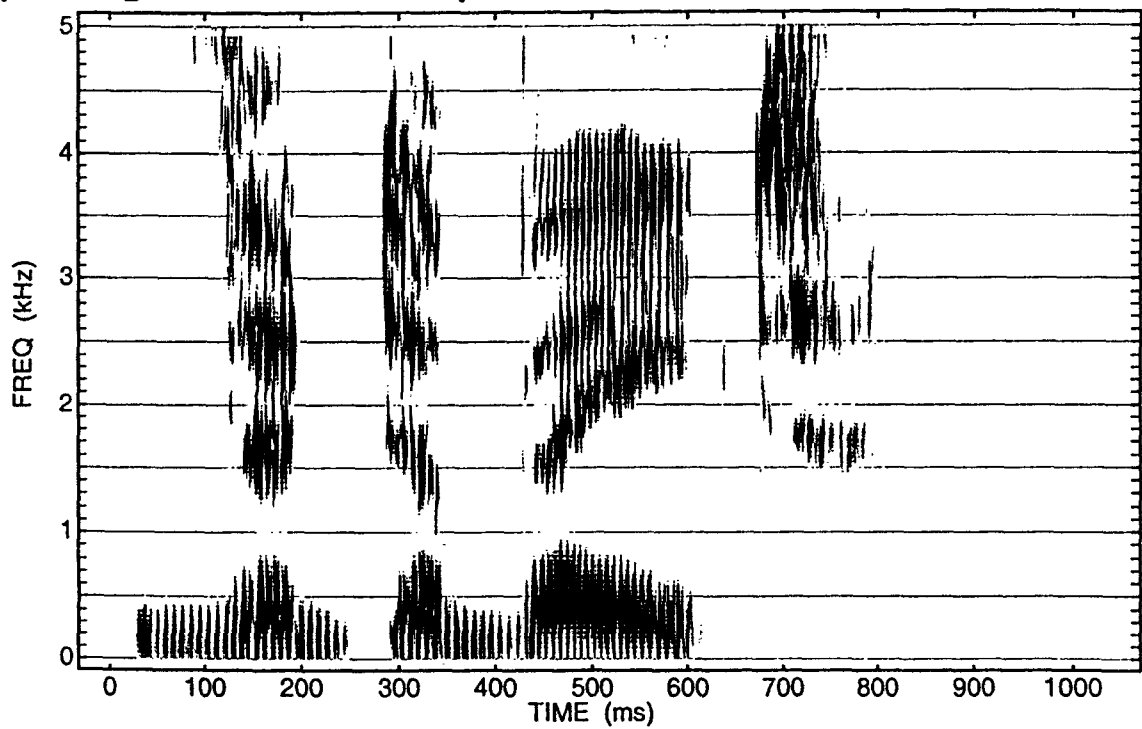


	d	ɛ	b	e	t		w	æ	g	ə	n		h	ɛ	l	p
vowel		+		+				+		+				+		
glide							+						+			
consonant	+		+		+				+		+				+	+
stressed		-		+				+		-				+		
reducible		+		-				-		+				-		
continuant	-		-		-				-		-				-	-
sonorant	-		-		-				-		+				+	-
strident																
lips			+													+
tongue blade	+				+						+				+	
tongue body									+							
round			-				+									-
anterior	+				+						+				+	
lateral															+	
high		+		-			+	-	+	-				-		
low		-		-			-	+	-	-				-		
back		-		-			+	-	-	+				-		
adv. tongue root				+			+	-						-		
spread glottis													+			
nasal											+					
stiff vocal folds	-		-		+				-							+

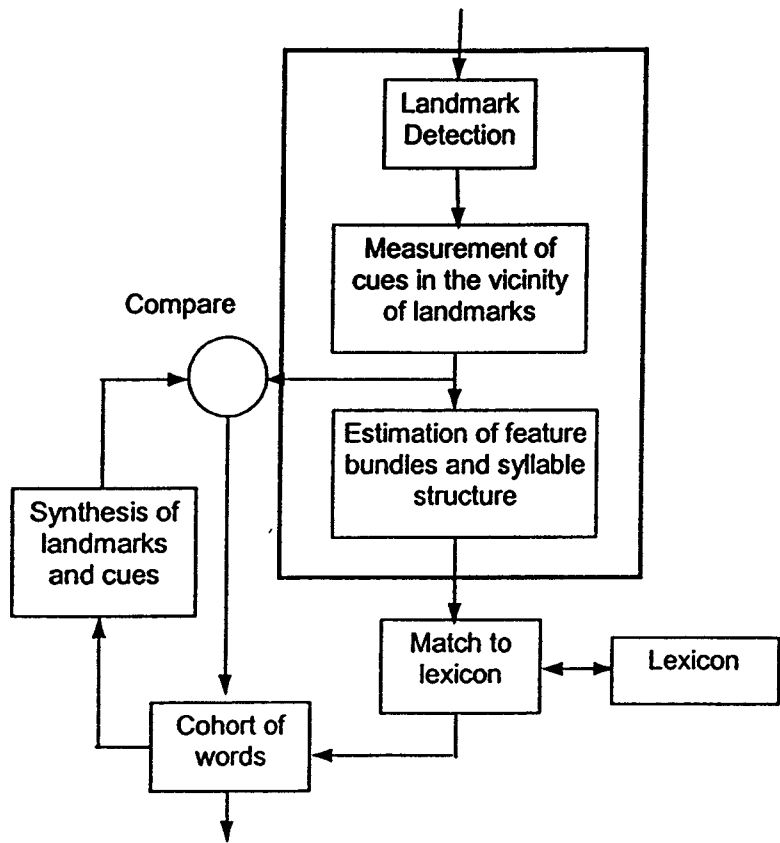


LSPECTO: 256-pt DFT, smart AGC 6.4-ms Hamming window every 1 ms

36



Possible model for lexical access



Speech intelligibility for various speech units

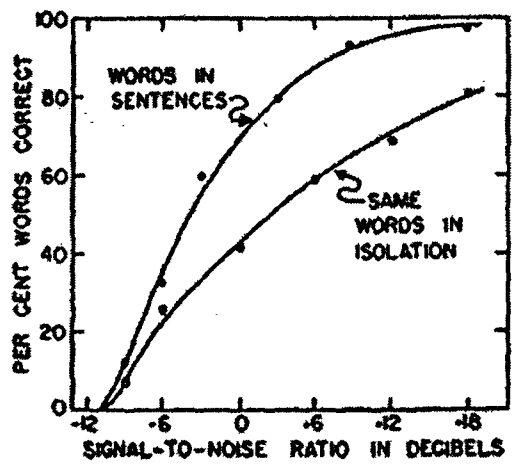


Figure 2: Dependence of word identification score on speech to noise ratio for keywords spoken in sentences (five per sentence) and for the same words spoken in isolation (from Miller, Heise, and Lichten, 1951).

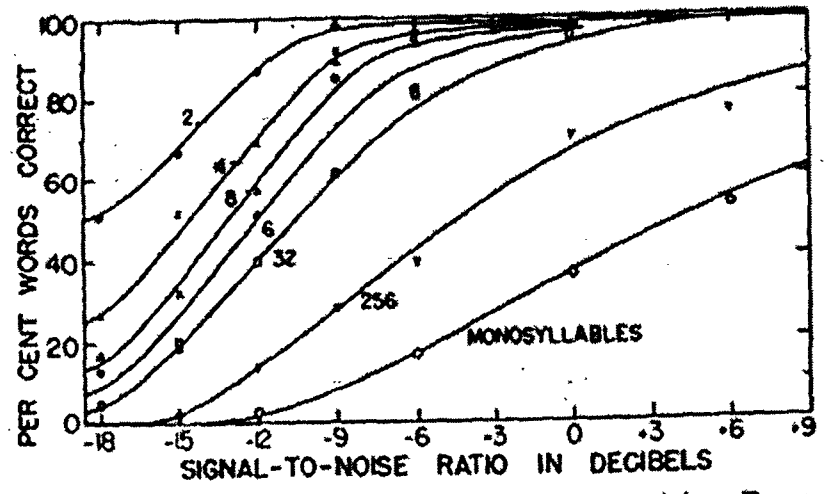


Fig. 2. Intelligibility of monosyllables as a function of the size of the test vocabulary. (Data are not corrected for effects of chance.)

Figure 1: Dependence of intelligibility on the number of items (monosyllabic words) in the test vocabulary (from Miller, Heise, and Lichten, 1951).

Spin Test: influence of context

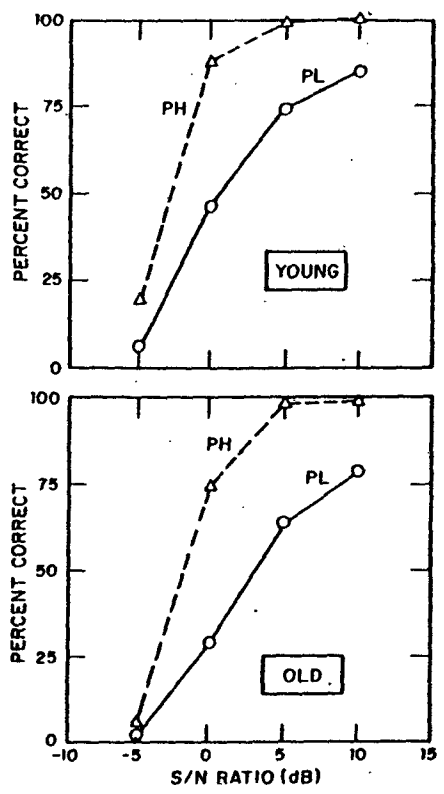


FIG. 4. PH and PL scores versus S/N ratio for the original 10 test forms (250 items of each type), for normally hearing young and old listeners.

examples: PH: The boat sailed across the bay

PL: We were talking about the bay

Confusion matrices for consonants in CV syllables

TABLE V. Confusion matrix for $S/N = +6$ db and frequency response of 200-6500 cps.

	p	t	k	f	θ	s	ʃ	b	d	g	v	ʋ	z	ʒ	m	n
p	162	10	55	5	3											
t	8	270	14									1				
k	38	6	171	1												
f	5	1	2	207	57			3			1					
θ	5	1	2	71	142	3					2	2				
s				1	7	232	2			1						
ʃ						1	239									
b				1	2			214			31	12				
d								11	206	14	9	1	2			
g									64	194	4	2	1			
v				1	1			14	2	2	205	39	5			1
ʋ								2		4	55	179	22	2		
z									3	10	2	20	198	3		
ʒ									3	4			2	215		
m															217	3
n									1						2	285

TABLE III. Confusion matrix for $S/N = -6$ db and frequency response of 200-6500 cps.

	p	t	k	f	θ	s	ʃ	b	d	g	v	ʋ	z	ʒ	m	n
p	80	43	64	17	14	6	2	1	1		1	1			2	
t	71	84	55	5	9	3	8	1					2		2	3
k	66	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
v				2		2		48	5	5	145	45	12		4	
ʋ					6			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

TABLE II. Confusion matrix for $S/N = -12$ db and frequency response 200-6500 cps.

	p	t	k	f	θ	s	ʃ	b	d	g	v	ʋ	z	ʒ	m	n
51	53	65	22	19	6	11		2		2	3	3	1	5	8	5
64	57	74	20	24	22	14		2	3	1	1	2	1	1	5	1
50	42	62	22	18	16	11		4	1	1	1	2			4	2
31	22	28	85	34	15	11		3	5		8	8	3		3	
26	22	25	63	45	27	12		6	9	3	11	9	3	2	7	2
16	15	16	33	24	53	48		3	5	6	3	1	6	2	1	
23	32	20	14	27	25	115		1	4	5	3		6	3	4	2
4	2	2	18	7	7	1		60	18	18	44	25	14	6	20	10
3	1	1	4	7	4	11		18	48	35	16	24	26	14	9	12
3	1	1	1	4	5	7		20	38	29	16	29	29	38	10	9
	1	1	12	5	4	5		37	20	23	71	16	14	4	14	9
		1	4	17	2	3		53	31	25	50	33	23	5	13	6
6	1	2	2	2	6	14	8	23	29	27	24	19	40	26	3	6
3	2	2	1		6	7		7	30	23	9	7	39	77	5	14
1	1		1	1	1	1		11	3	6	8	11	1	1	109	60
								2	2	6	7	1	1	9	84	145

Simplified model for speech production

10

Figure 2

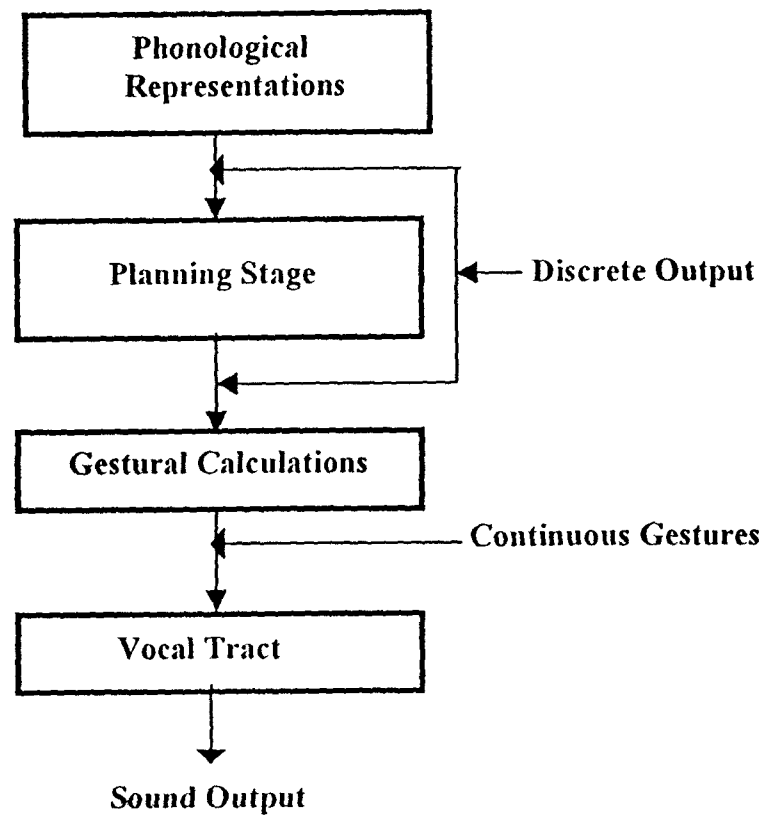


Figure 4

Model for speech production with "enhancement" and "overlap"

