

Acoustic Phonetic Representation of Speech Sounds



Speech Analysis Tools

- Wavesurfer
- Audacity
- Praat
- Sonic Visualiser
- Speech Signal Processing Toolkit (SPTK)
- Waveform Wizard



Silence, Unvoiced and Voiced Portions of Speech

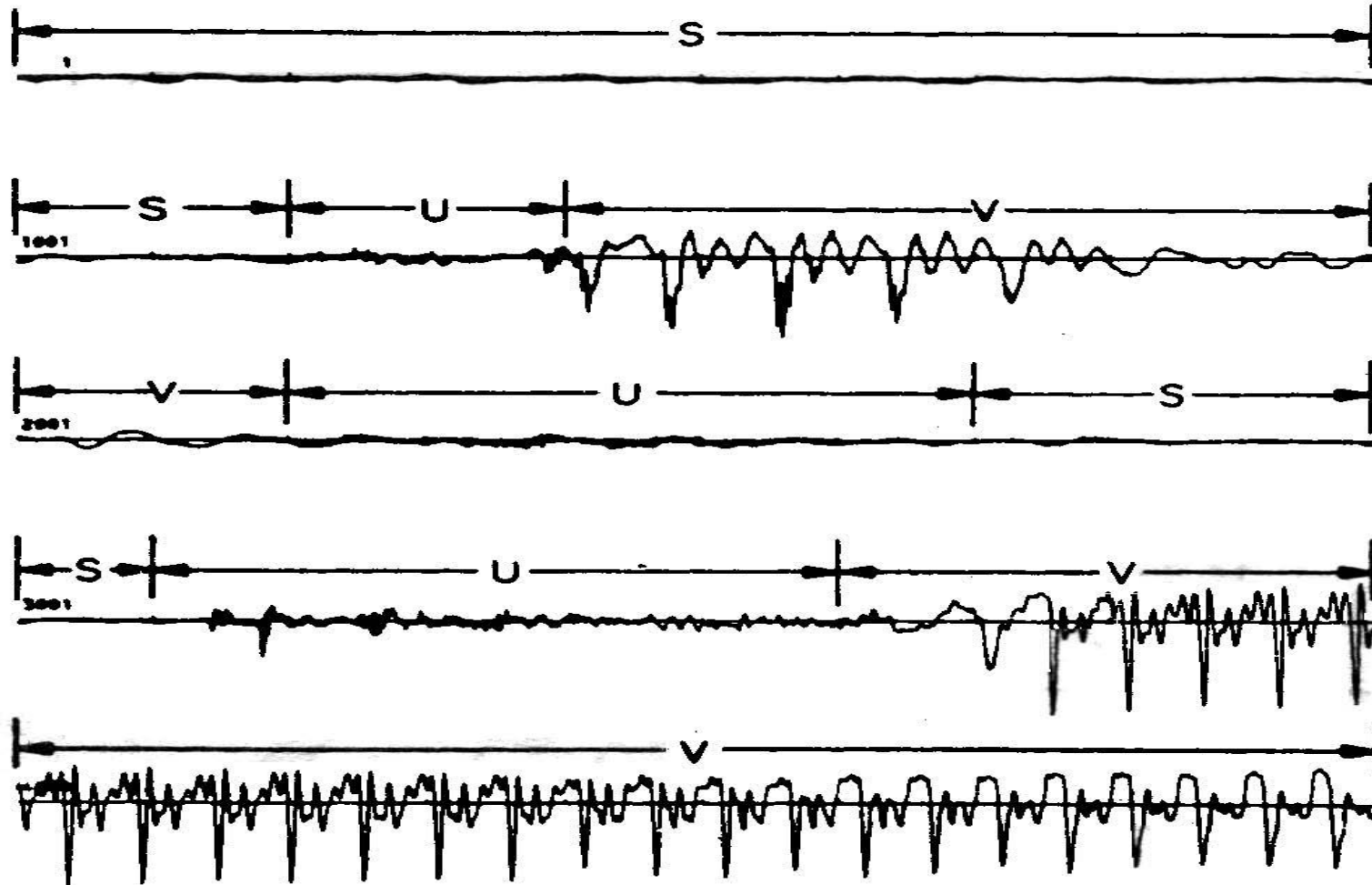


Figure 2.7 Waveform plot of the beginning of the utterance "It's time."

Narrowband & Wideband Spectrograms

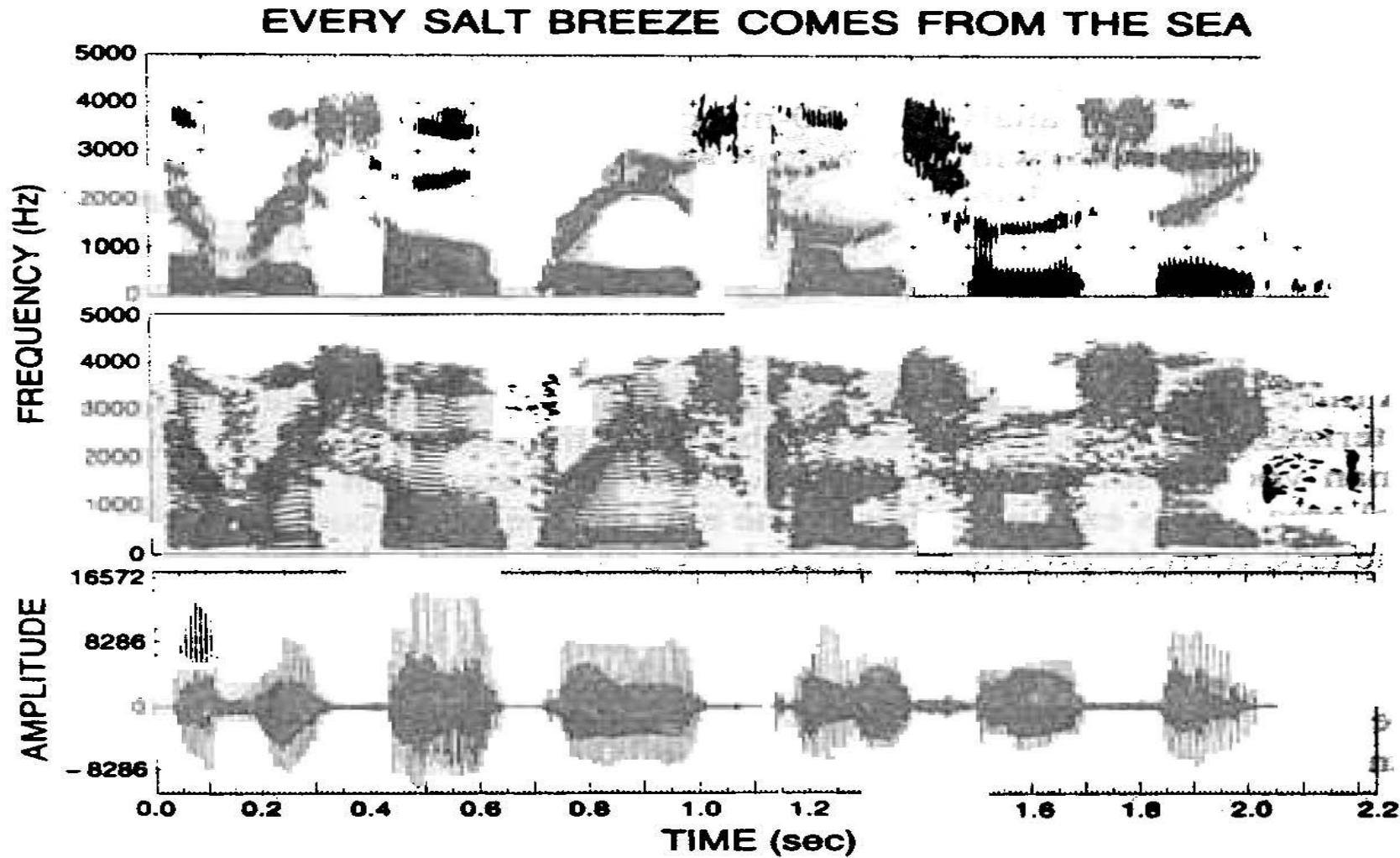
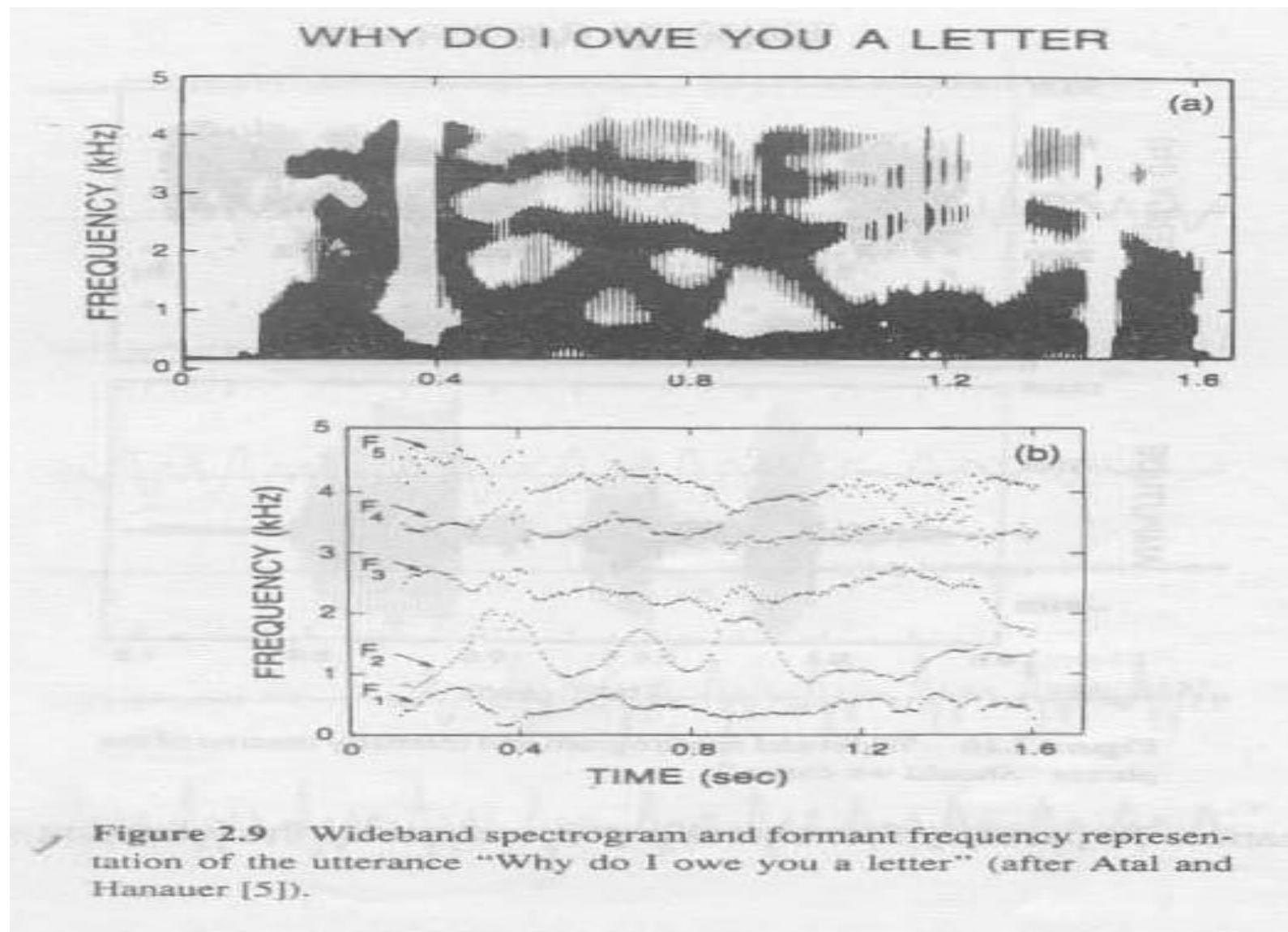


Figure 2.8 Wideband and narrowband spectrograms and speech amplitude for the utterance "Every salt breeze comes from the sea."

Formant Frequencies



Time and Time-Frequency Representation of Speech

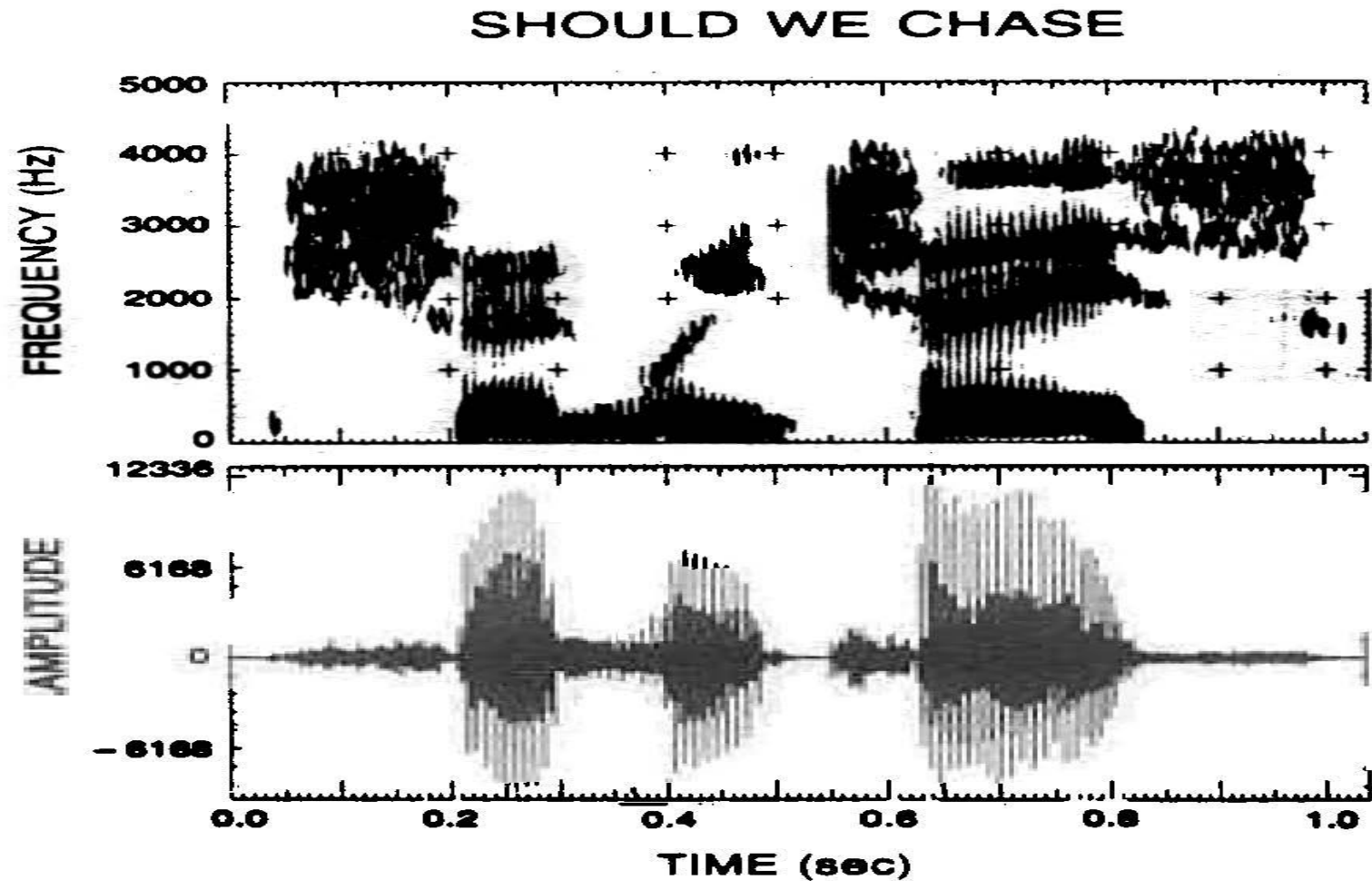
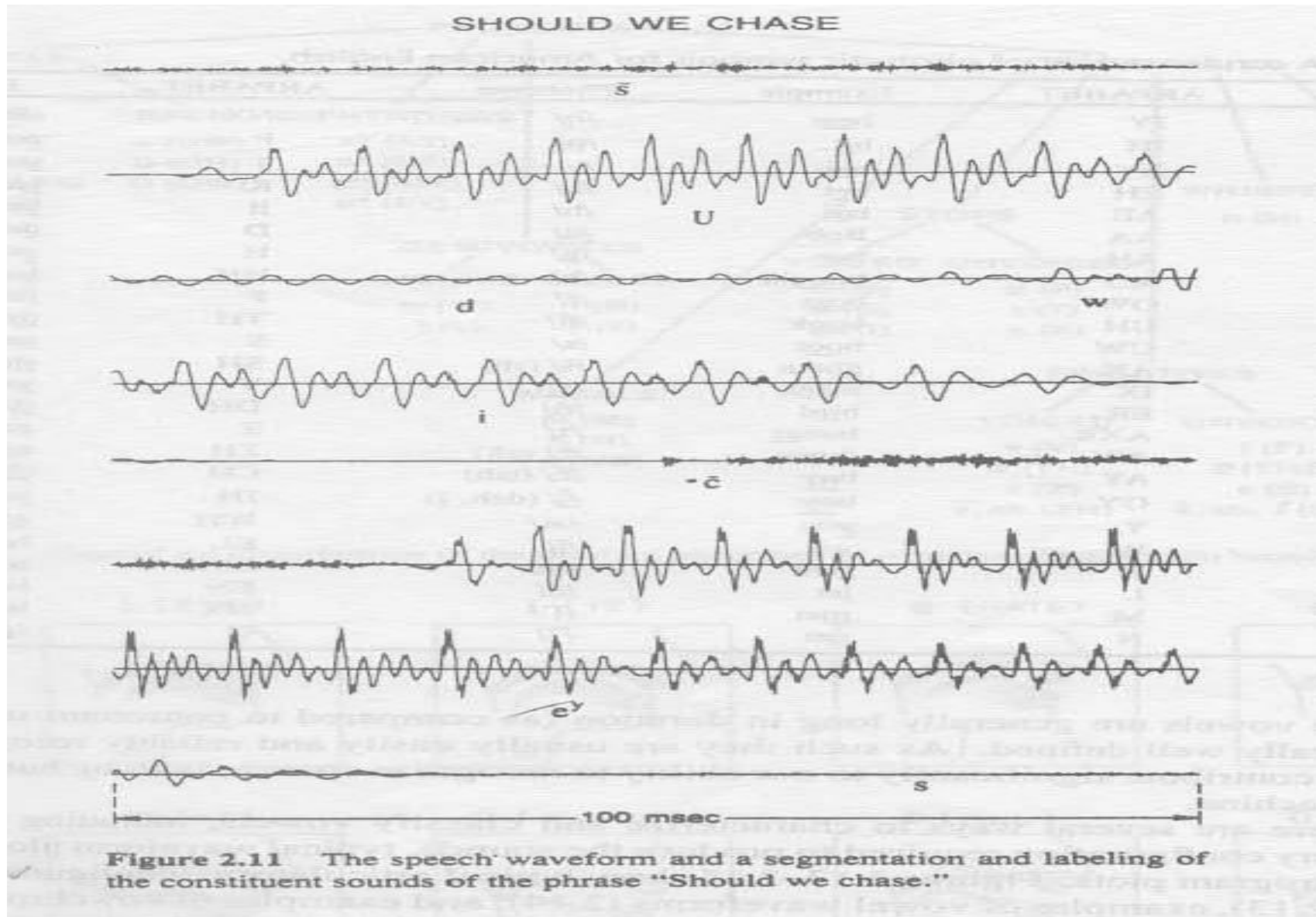


Figure 2.10 Wideband spectrogram and intensity contour of the phrase "Should we chase."

Voiced and Unvoiced Portions of Speech



Significance of Vowels vs. Consonants

Text : **Speech signal is produced** by exciting the time varying vocal tract system with time varying excitation

Vowels: - - ee - - - i - - a - i - - - o - u - e -

Consonants: Sp - - ch s - gn - l - s pr - d - c - d

Classification of English Phonemes

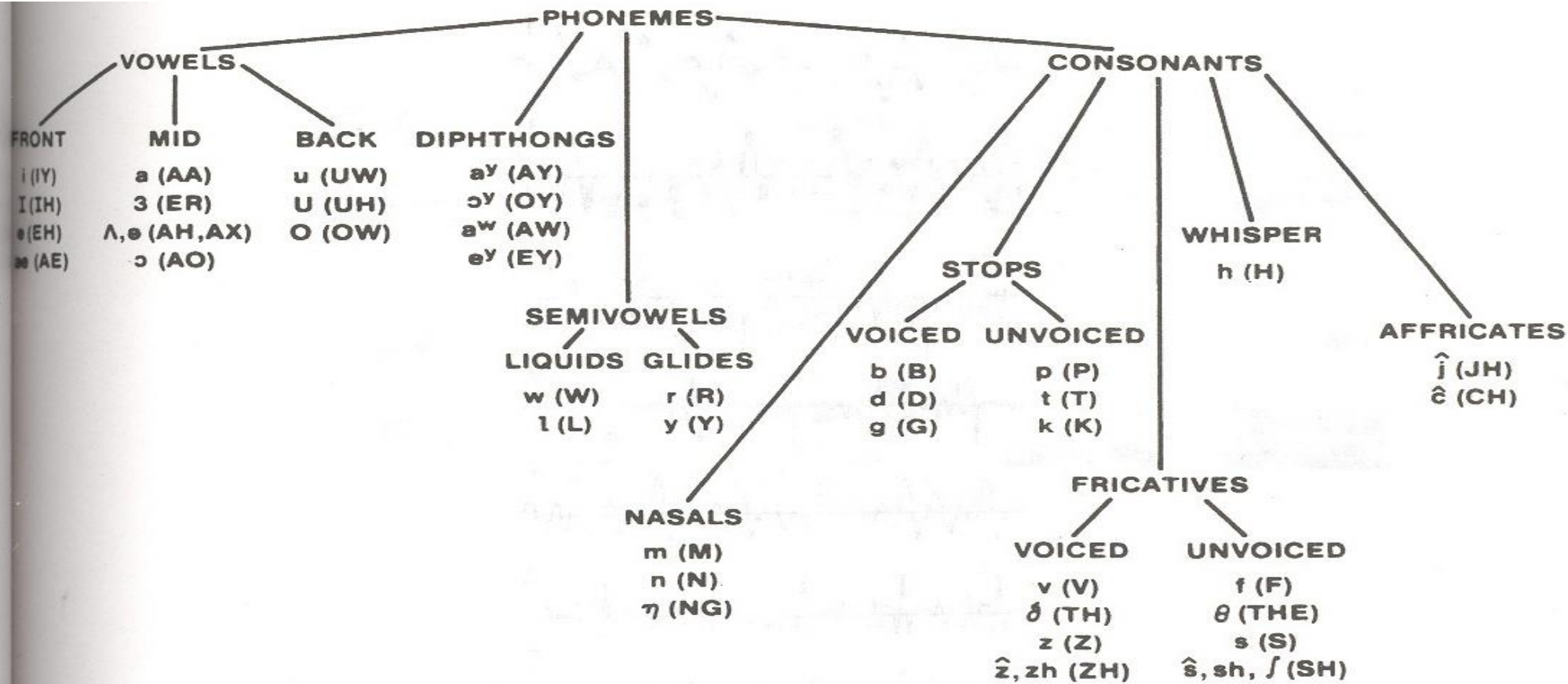


Figure 2.12 Chart of the classification of the standard phonemes of American English into broad sound classes.

Hindi Sound Units

Set of Characters representing Vowel speech sounds

Vowel	Short	/a/(अ)	/i/(इ)	/u/(उ)	/e/(ए)	/o/(ओ)
	Long	/a:/(आ)	/i:/(ई)	/u:/(ऊ)	/e:/(ऎ)	/o:/(औ.)
	Diphthongs	/ai/(ऐ)		/au/(औ)	/ /	/ /

Set of Characters representing Consonant-Vowel Combinations, where Consonant is any plosive and Vowel is /a/(अ)

Place of articulation	Manner of articulation				Nasals	Semivowels	Fricatives
	Unvoiced		Voiced				
	Unaspirated	Aspirated	Unaspirated	Aspirated			
Velar	/ka/(क)	/kha/(ख)	/ga/(ग)	/gha/(घ)	/kna/(ङ)	--	/ha/(ह)
Palatal	/cha/(च)	/chha/(छ)	/ja/(ज)	/jha/(झ)	/cna/(ञ)	/ya/(य)	/sha/(श)
Retroflex	/ṭa/(ट)	/ṭha/(ठ)	/Ḍa/(ड)	/Ḍha/(ढ)	/ṭna/(ण)	/ra/(र)	/shha/(ष)
Denti-Alveolar	/ta/(त)	/tha/(थ)	/da/(द)	/dha/(ध)	/na/(न)	/la/(ल)	/sa/(स)
Bilabial	/pa/(प)	/pha/(फ)	/ba/(ब)	/bha/(भ)	/ma/(म)	/va/(व)	--

Tongue Hump in Vowel Sounds

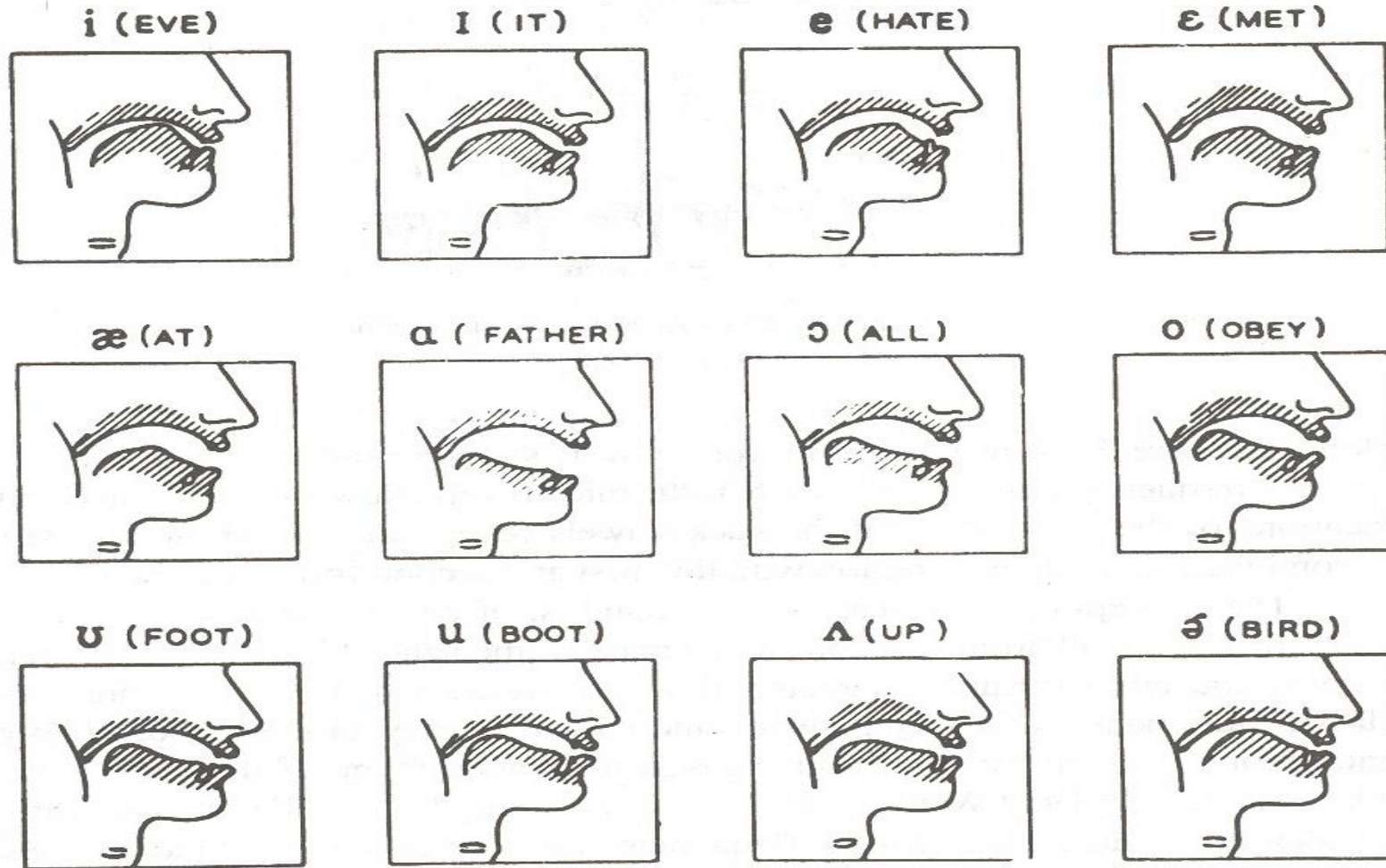


Figure 2.13 Articulatory configurations for typical vowel sounds (after Flanagan [3]).

Waveform Plots of Vowel Sounds

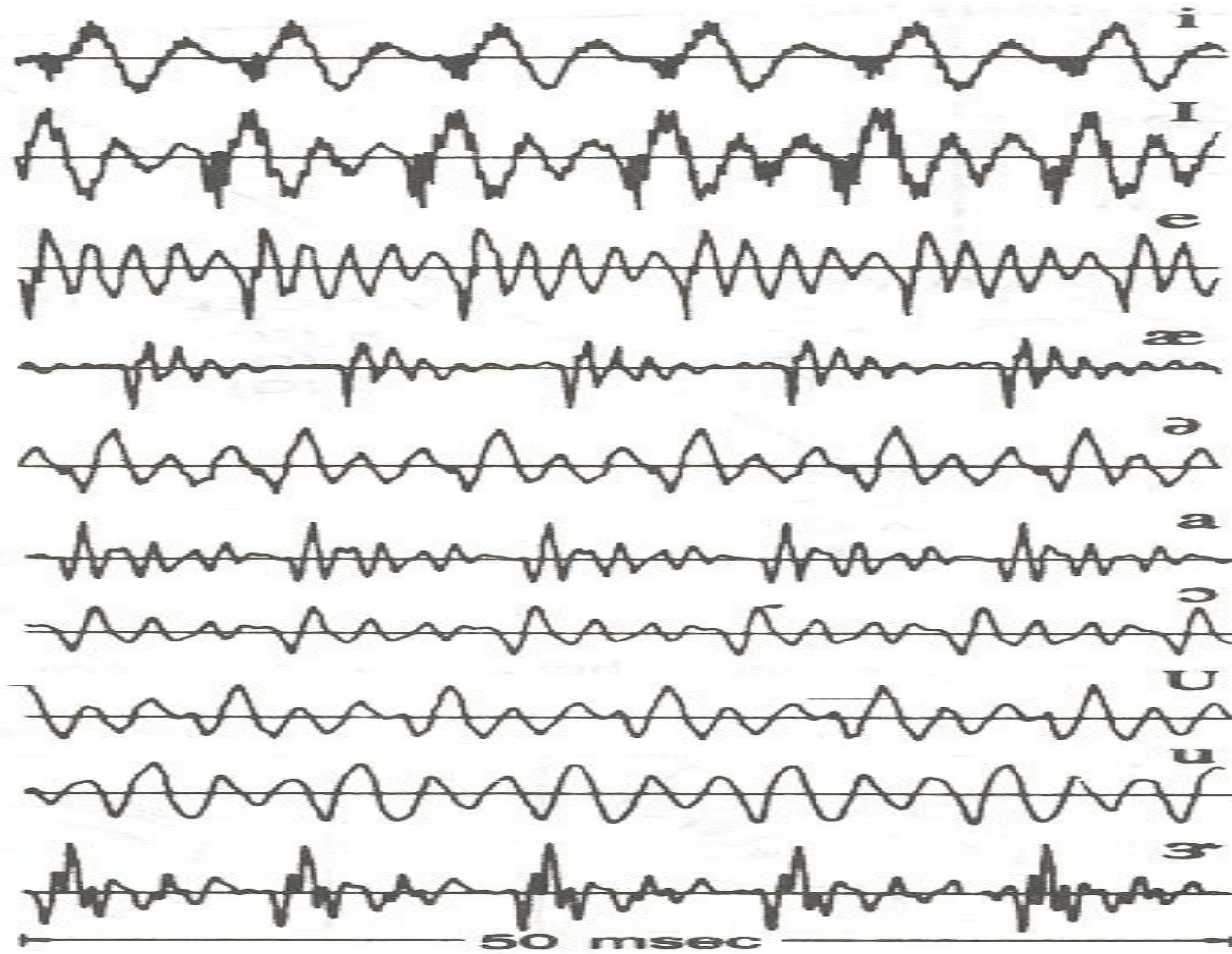


Figure 2.14 Acoustic waveform plots of typical vowel sounds.

Spectrograms of the Vowel Sounds

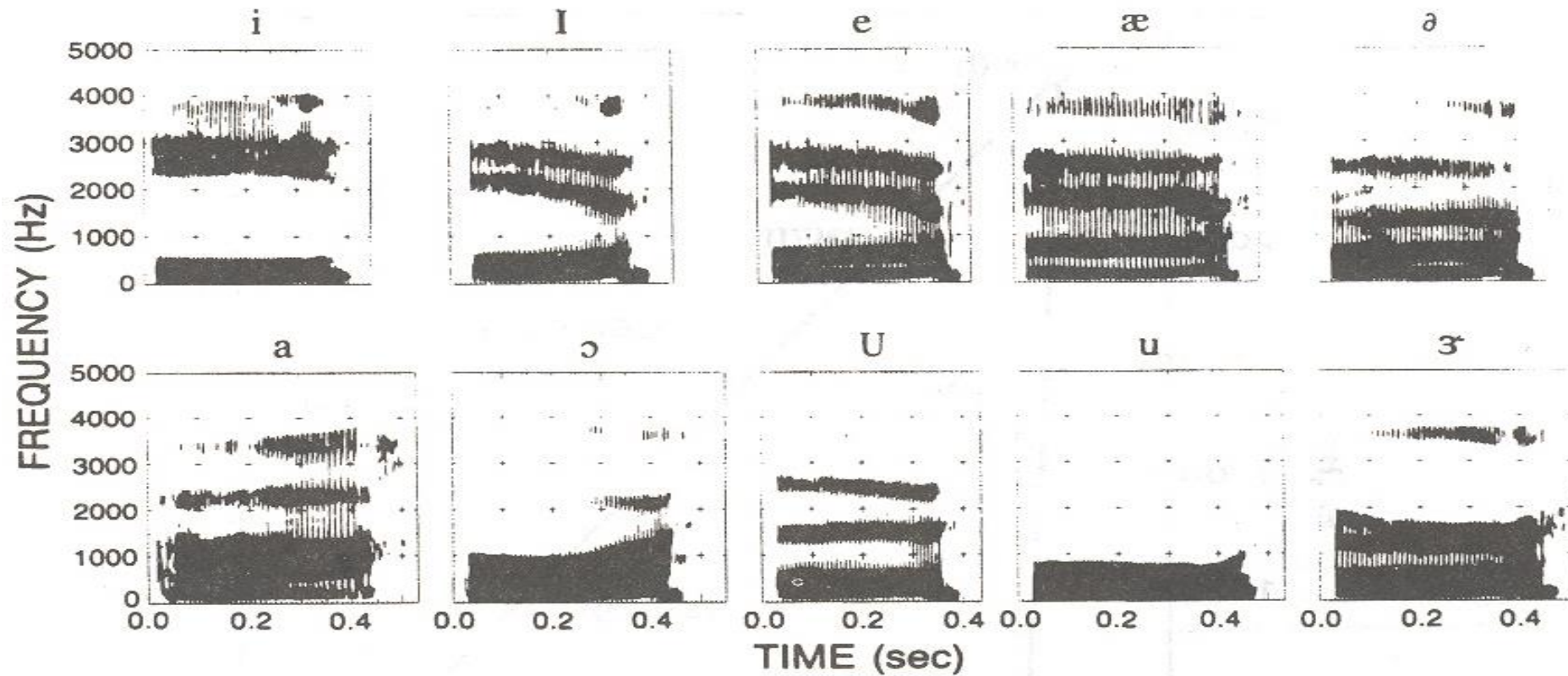


Figure 2.15 Spectrograms of the vowel sounds.

F1-F2 Clusters of the Vowel Sounds

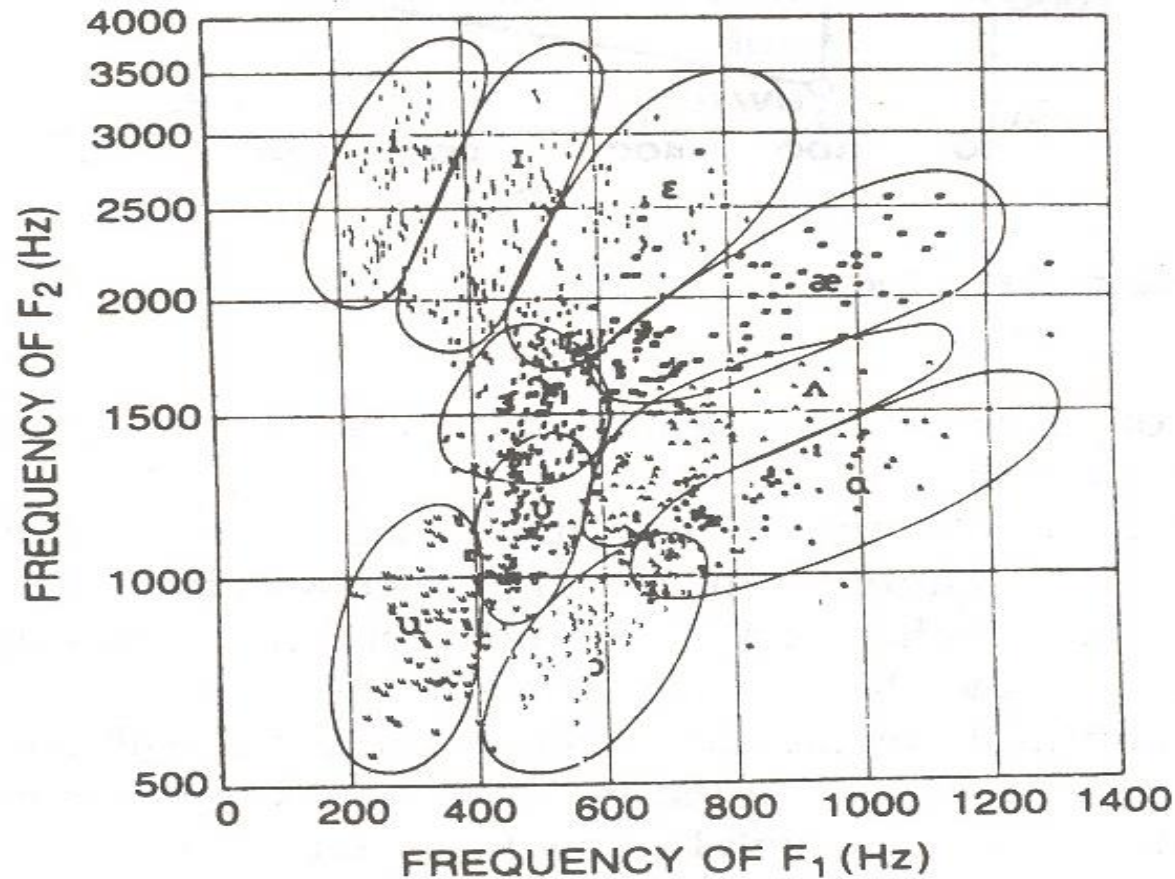


Figure 2.16 Measured frequencies of first and second formants for a wide range of talkers for several vowels (after Peterson & Barney [7]).

F1-F2 Centroids of the Vowel Sounds

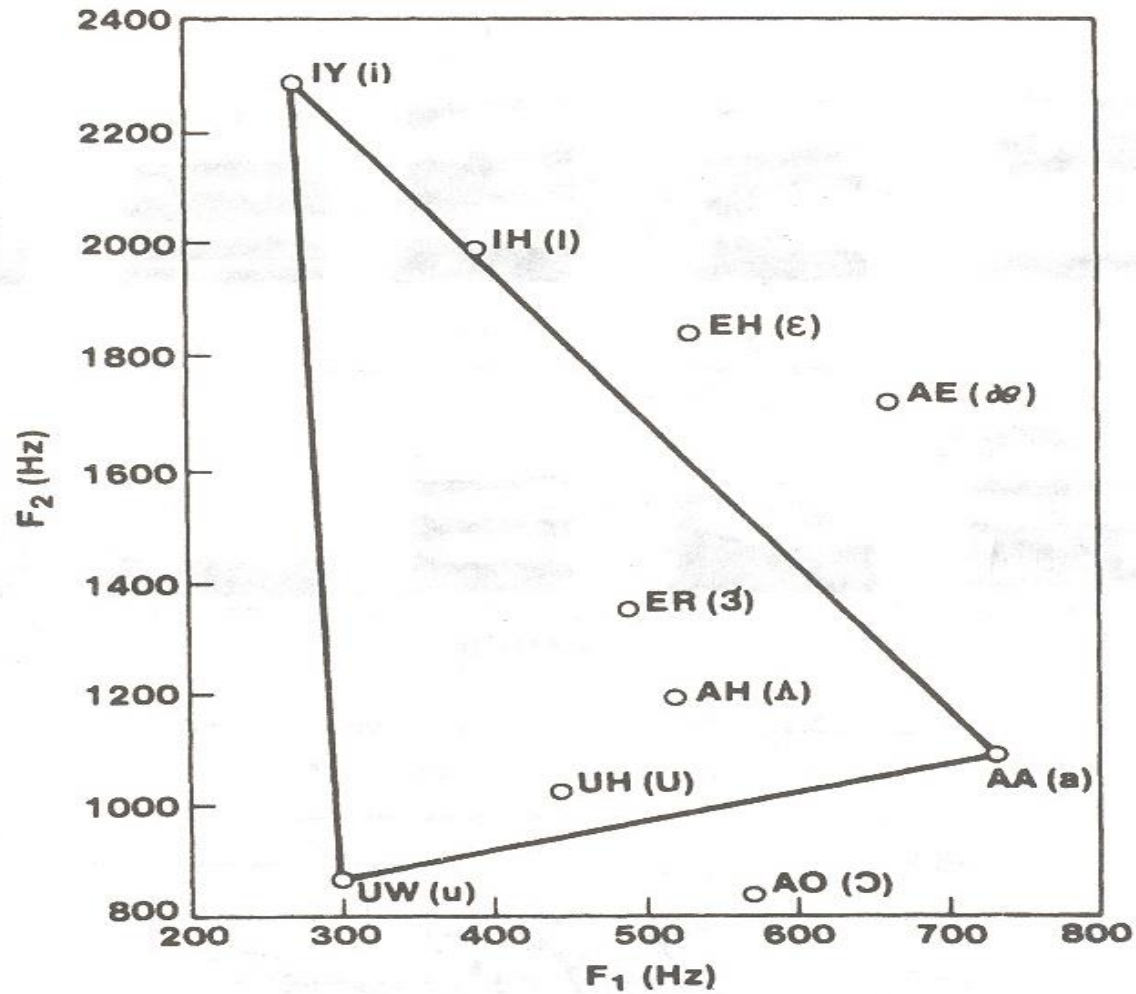


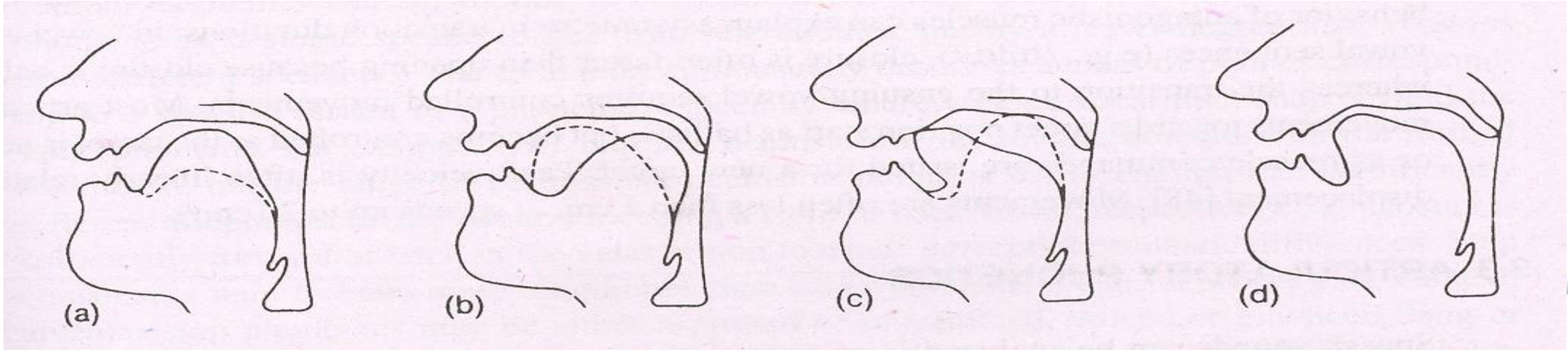
Figure 2.17 The vowel triangle with centroid positions of the common vowels.

Formant Frequencies for Typical Vowels

TABLE 2.2. Formant frequencies for typical vowels.

ARPABET Symbol for Vowel	IPA Symbol	Typical Word	F ₁	F ₂	F ₃
IY	/i/	beet	270	2290	3010
IH	/ɪ/	bit	390	1990	2550
EH	/ɛ/	bet	530	1840	2480
AE	/æ/	bat	660	1720	2410
AH	/ʌ/	but	520	1190	2390
AA	/ɑ/	hot	730	1090	2440
AO	/ɔ/	bought	570	840	2410
UH	/ʊ/	foot	440	1020	2240
UW	/u/	boot	300	870	2240
ER	/ɜ/	bird	490	1350	1690

Articulatory Positions for Production of Various Sound Units



- Articulatory positions:
- (a) vowel with two tongue heights
 - (b) high vowel for front and back positions
 - (c) stop for alveolar and velar places
 - (d) an alveolar fricative

Vowel Diagram

(position of the tongue hump in oral cavity,
F-front, C-central and B-back positions)

F

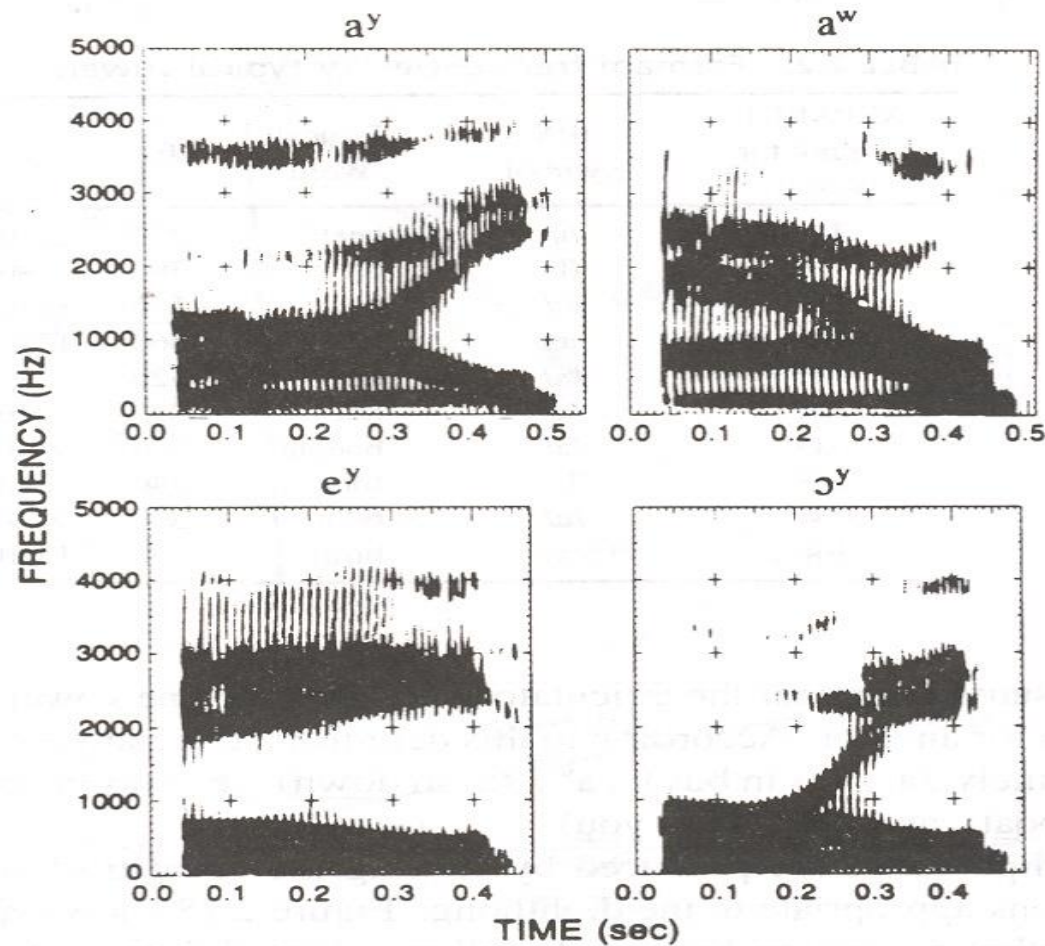
C

B

H	<i>/i/</i>		<i>/u/</i>
M	<i>/e/</i>		<i>/o/</i>
L		<i>/a/</i>	

(height of the tongue hump,
H-high, M-medium and L-low)

Spectrograms of Diphthongs

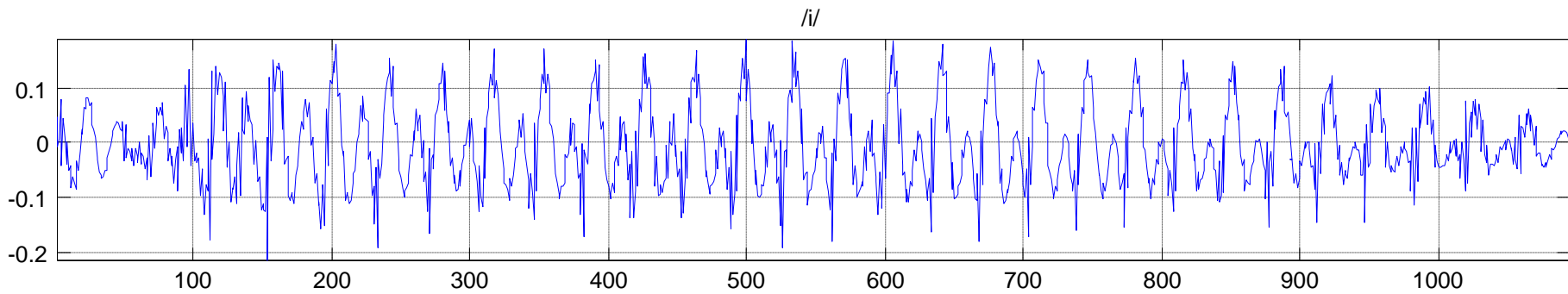
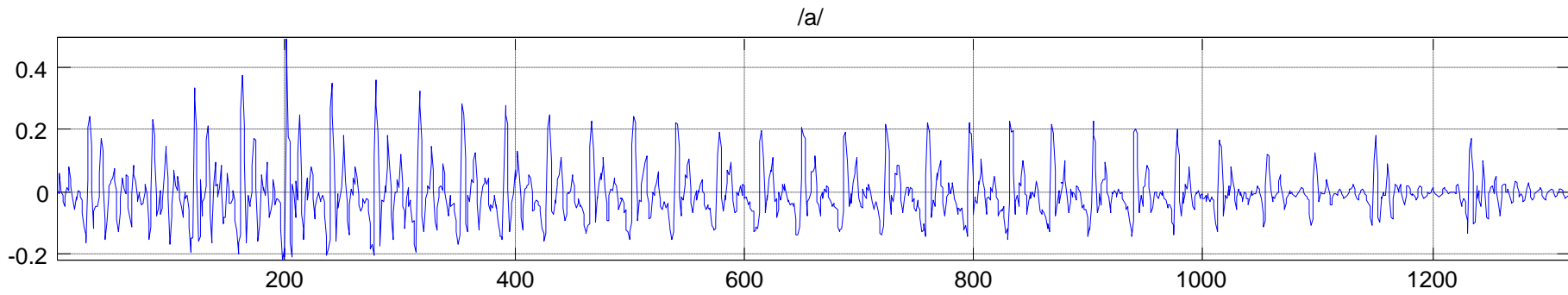
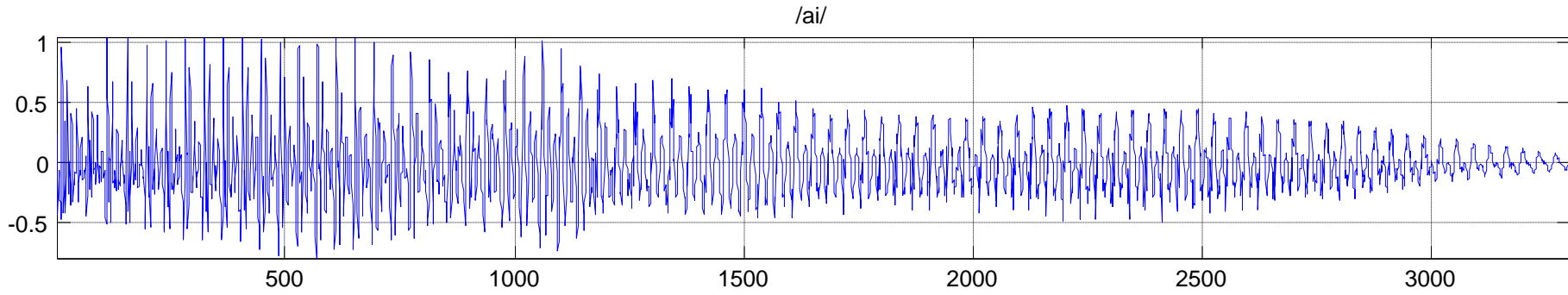


Diphthongs in English

1. buy
2. down
3. bait
4. boy
5. boat
6. you

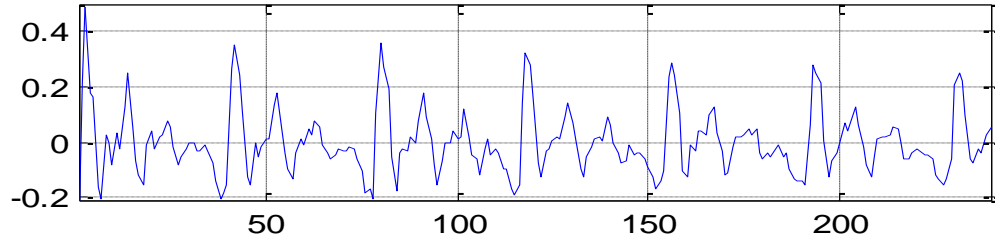
Figure 2.18 Spectrogram plots of four diphthongs.

Diphthongs : Time Domain Waveforms

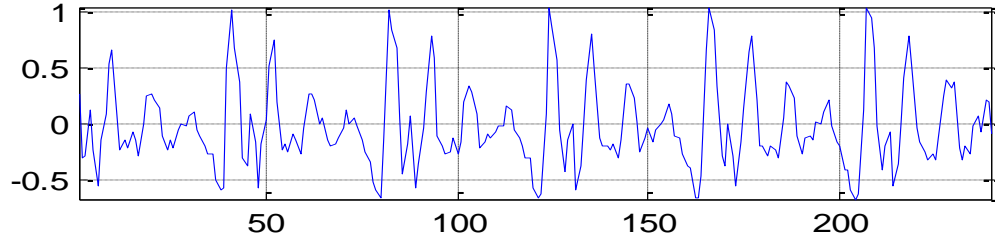


Diphthongs

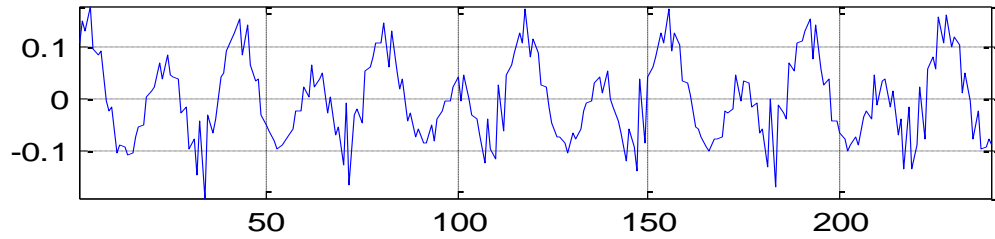
/a/



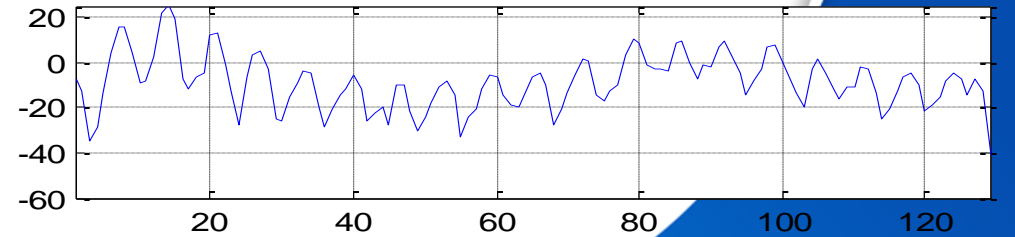
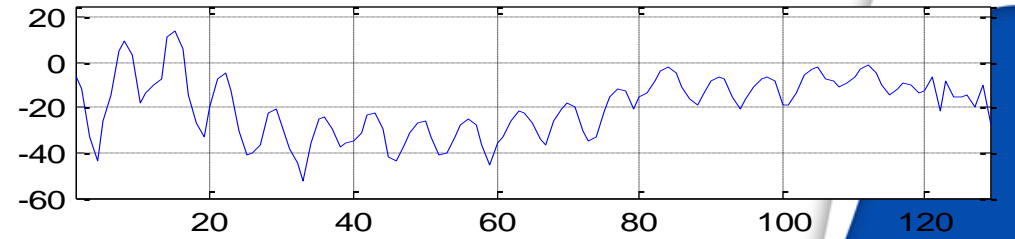
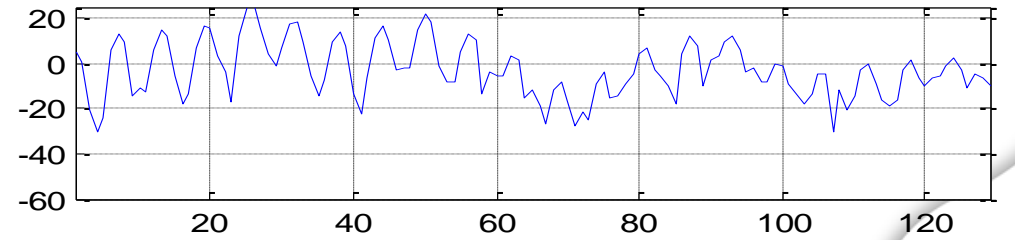
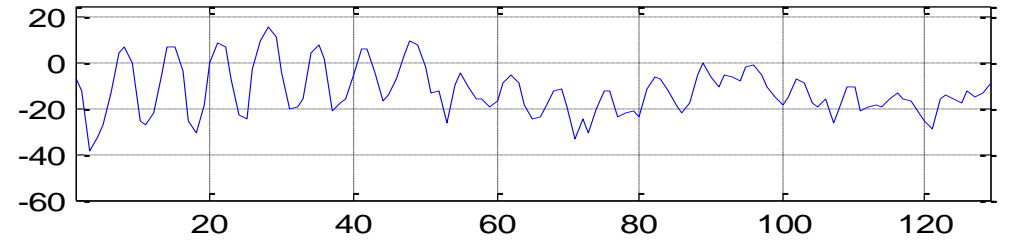
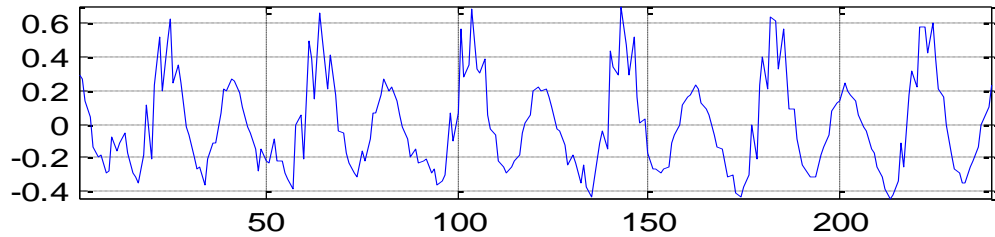
/a/ part of /ai/



/i/



/i/ part of /ai/



Time-Domain Waveforms of Nasal Sounds

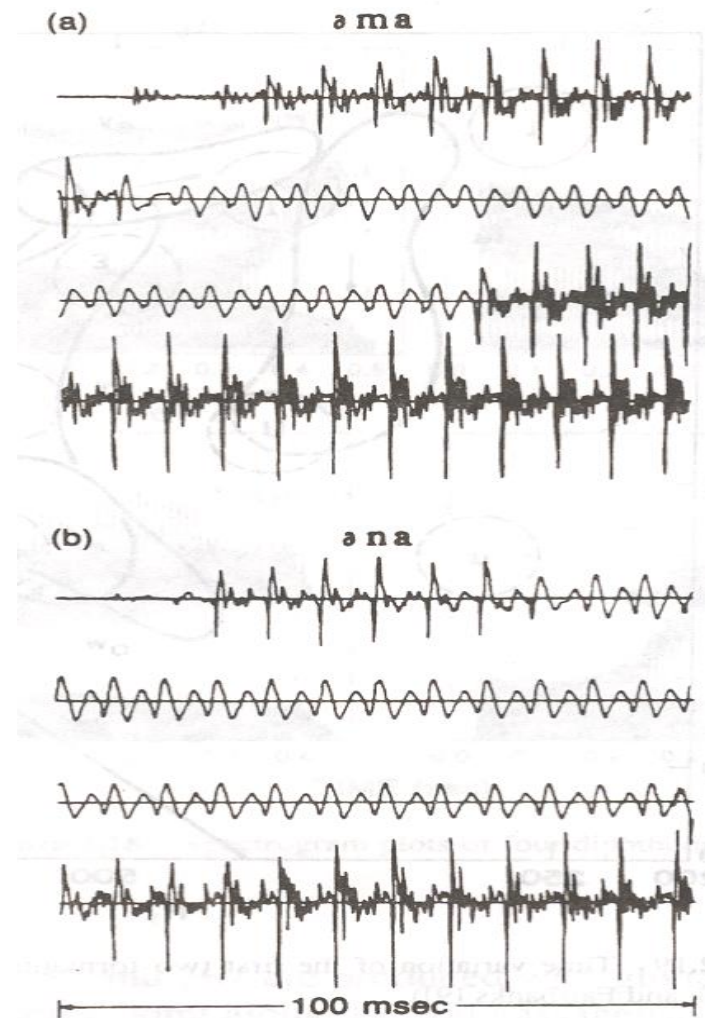


Figure 2.20 Waveforms for the sequences /ə-m-a/ and /ə-n-a/.

Spectrograms of Nasal Sounds

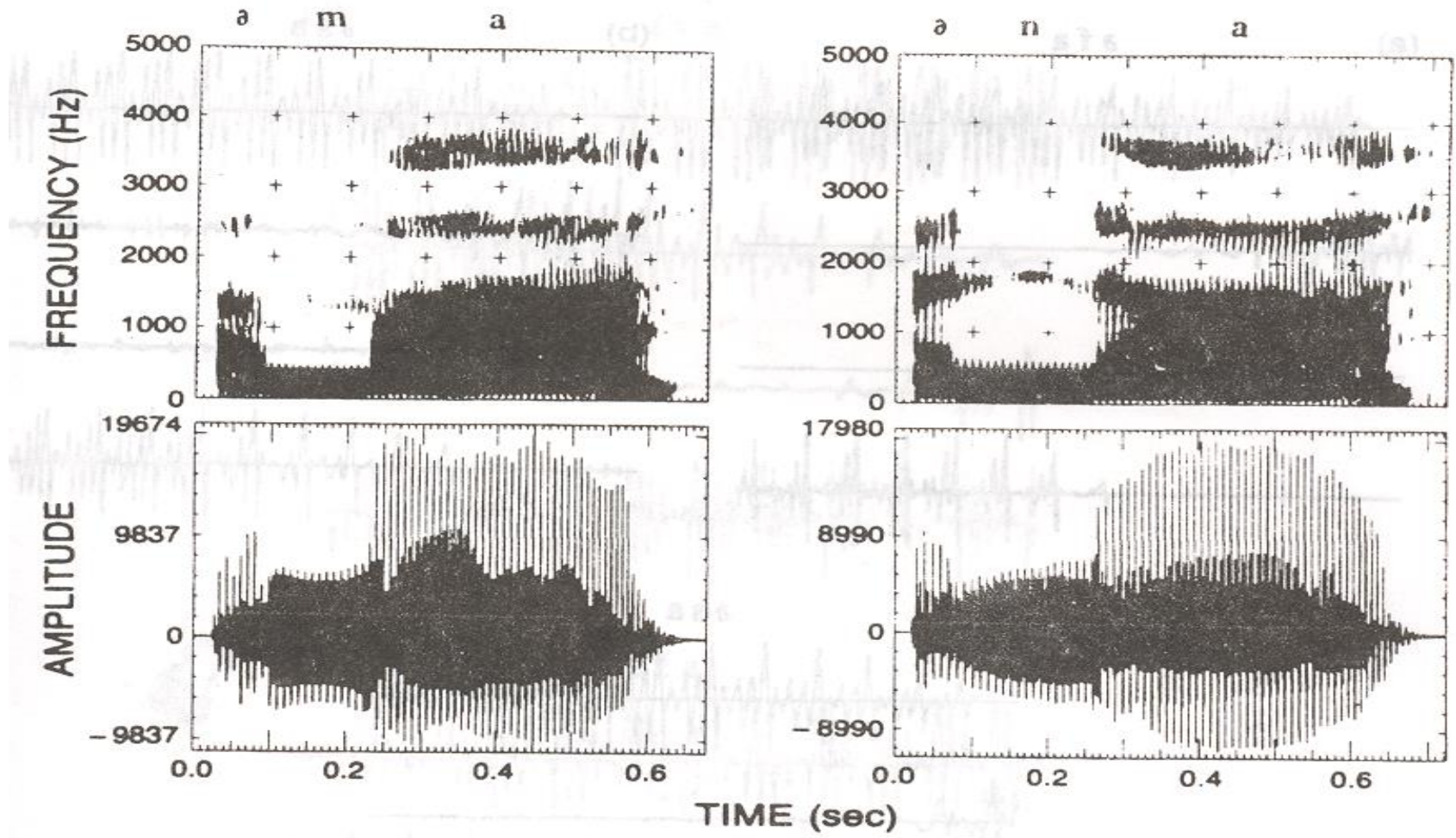


Figure 2.21 Spectrograms of the sequences /ə-m-a/ and ə-n-a/.

Waveforms of Unvoiced Fricatives (/f/, /s/, /sh/)

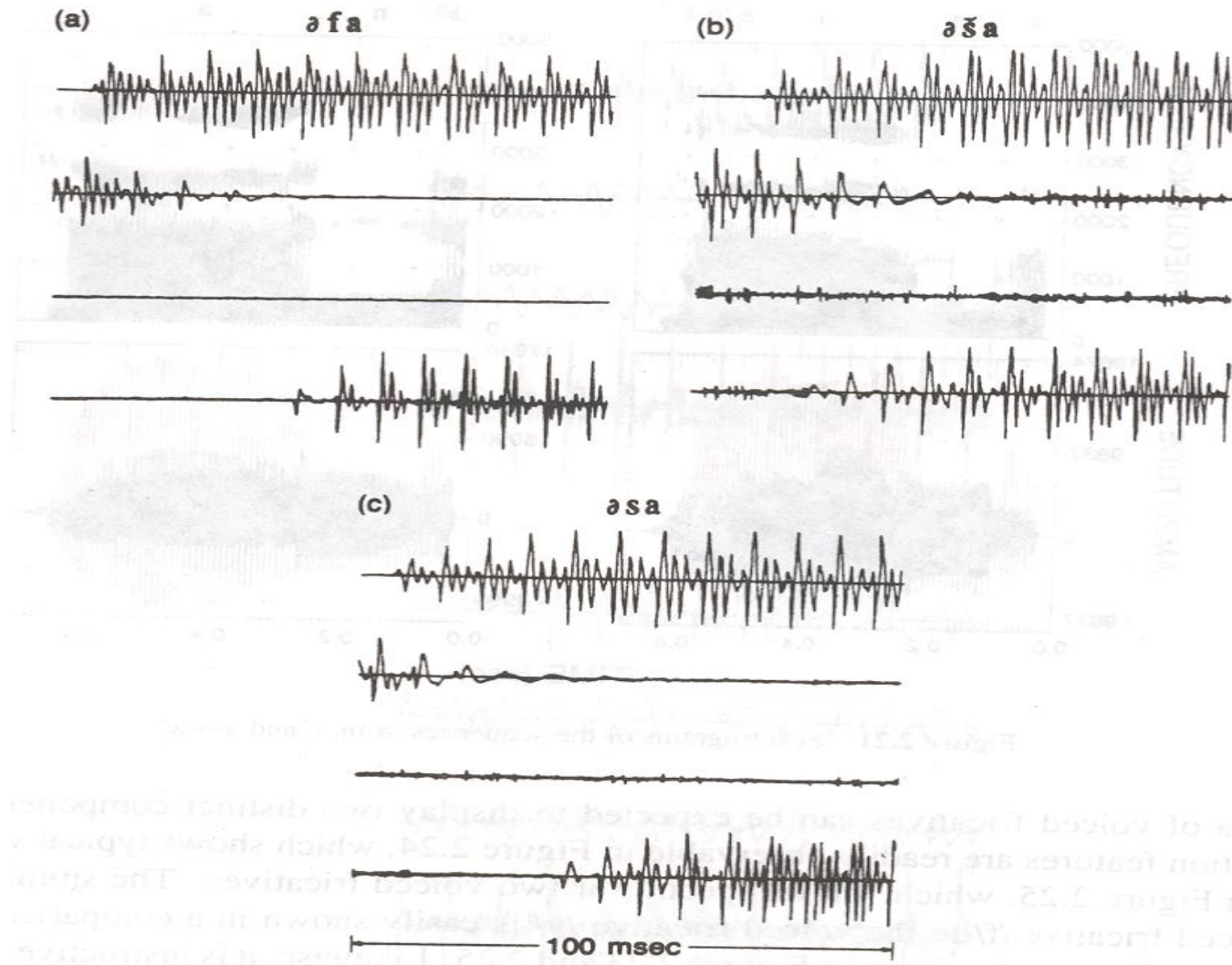


Figure 2.22 Waveforms for the sounds /f/, /s/ and /sh/ in the context /ə-x-a/ where /x/ is the unvoiced fricative.

Spectrograms of Unvoiced Fricatives (/f/, /s/, /sh/)

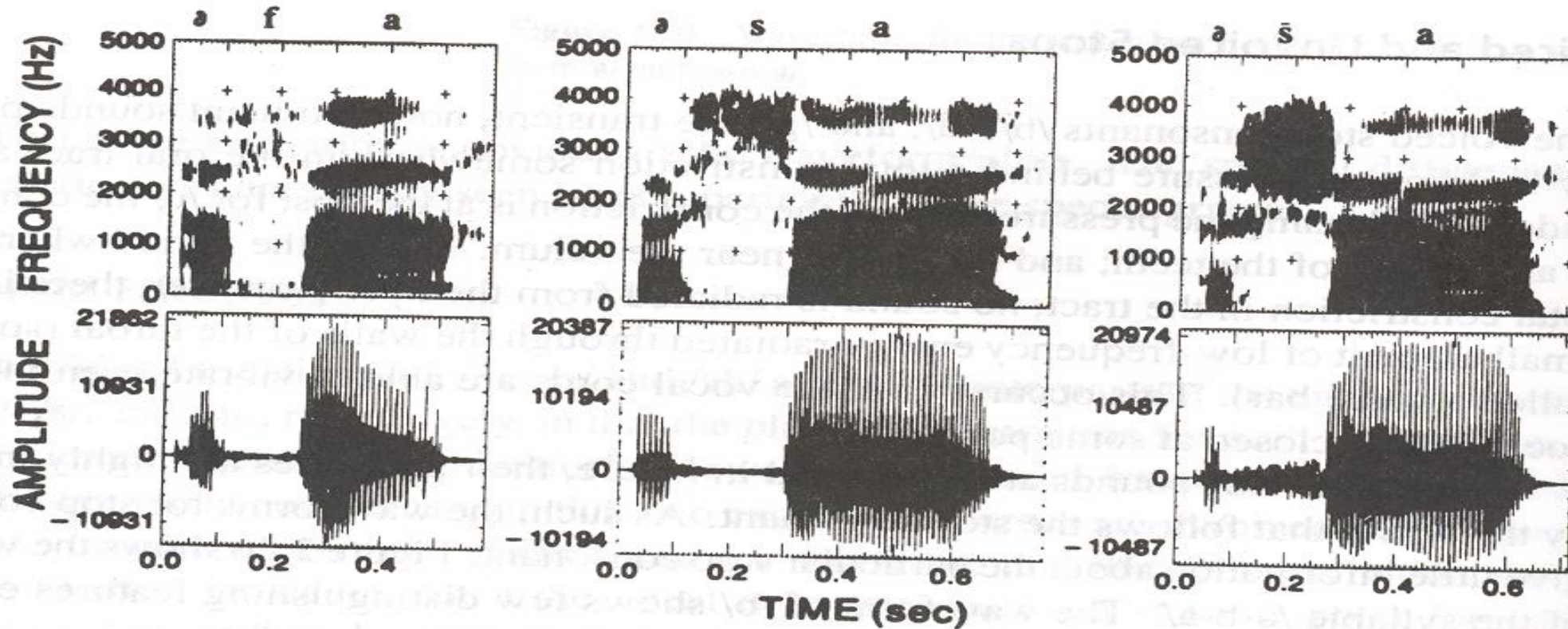


Figure 2.23 Spectrogram comparisons of the sounds /ə-f-a/, /ə-s-a/ and /ə-sh-a/.

Waveforms of Voiced Fricatives (/v/, /z/)

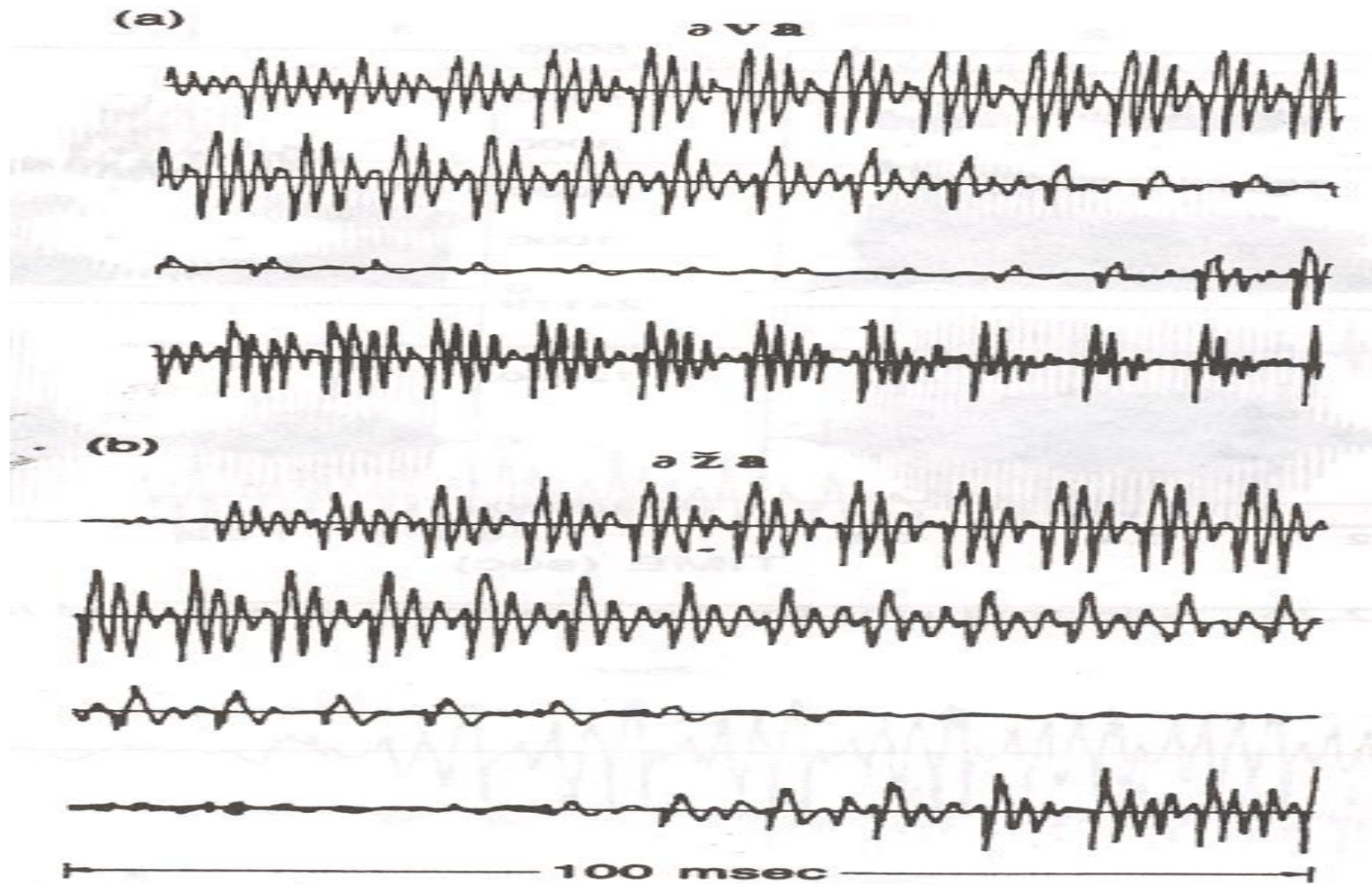


Figure 2.24 Waveforms for the sequences /ə-v-a/ and /ə-zh-a/.

Spectrograms of Voiced Fricatives (/v/, /z/)

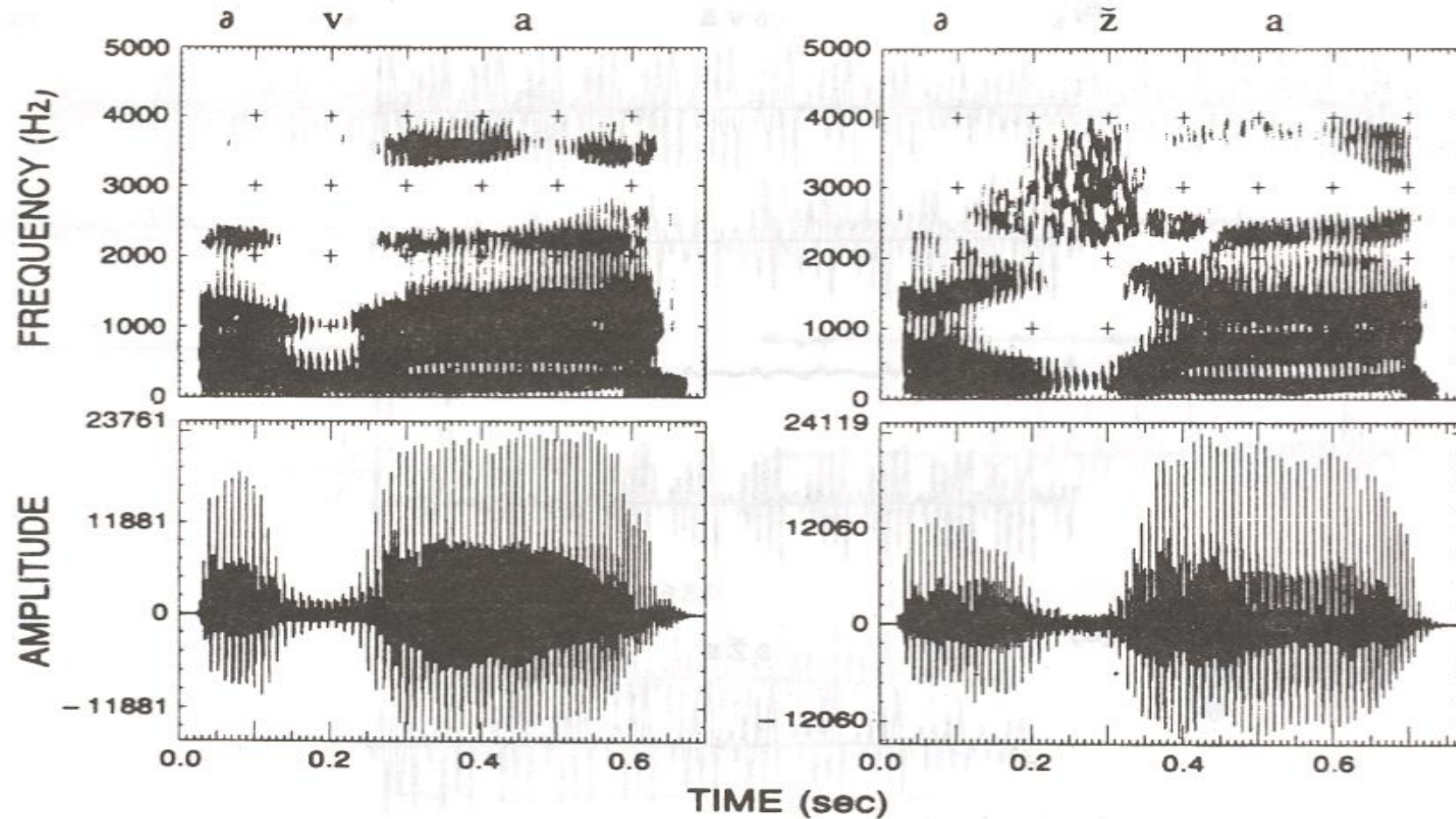


Figure 2.25 Spectrograms for the sequences /ə-v-a/ and /ə-zh-a/.

Waveform of a Voiced Stop (/b/)

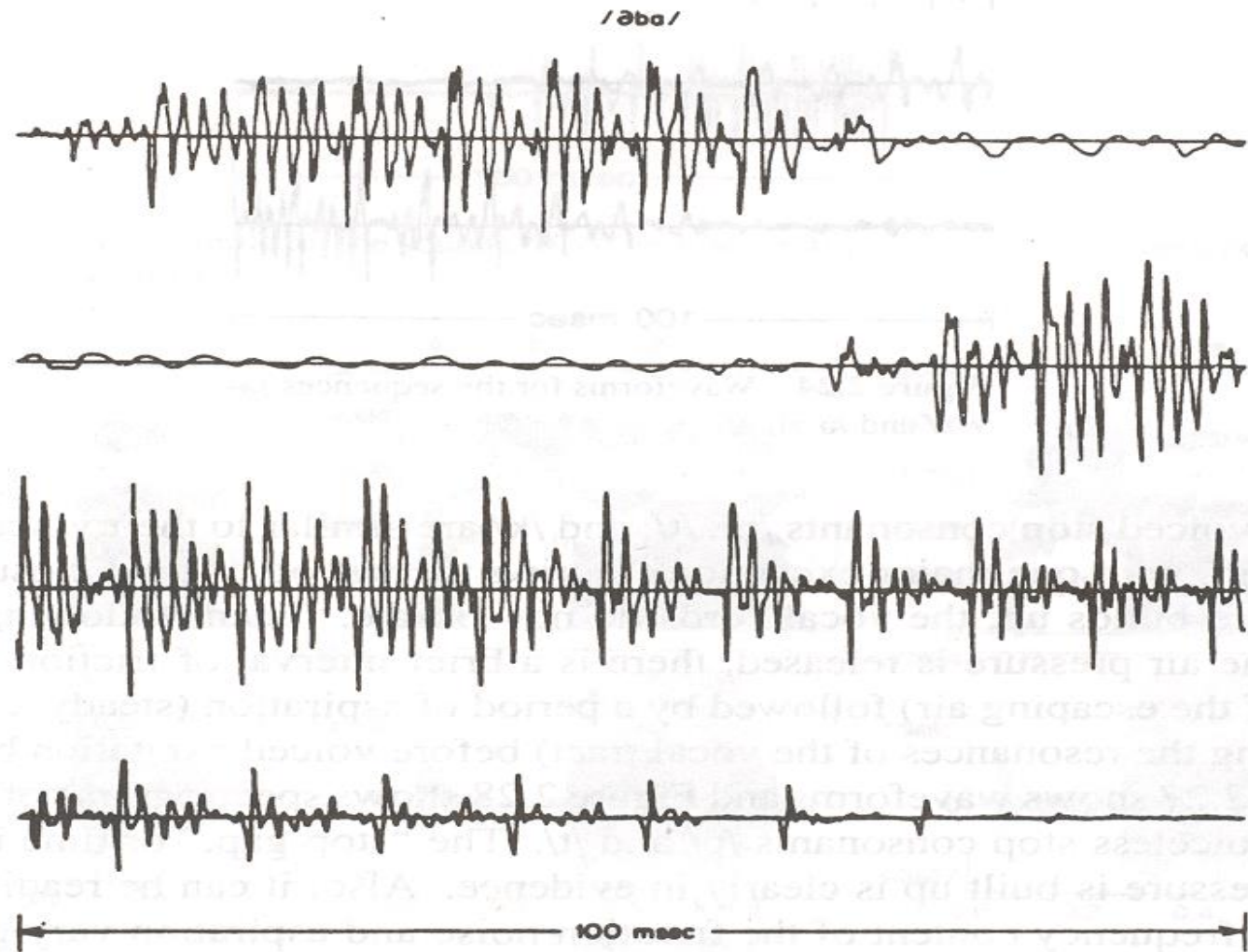


Figure 2.26 Waveform for the sequence /ə-b-a/.

Waveforms of Unvoiced Constants (/p/, /t/)

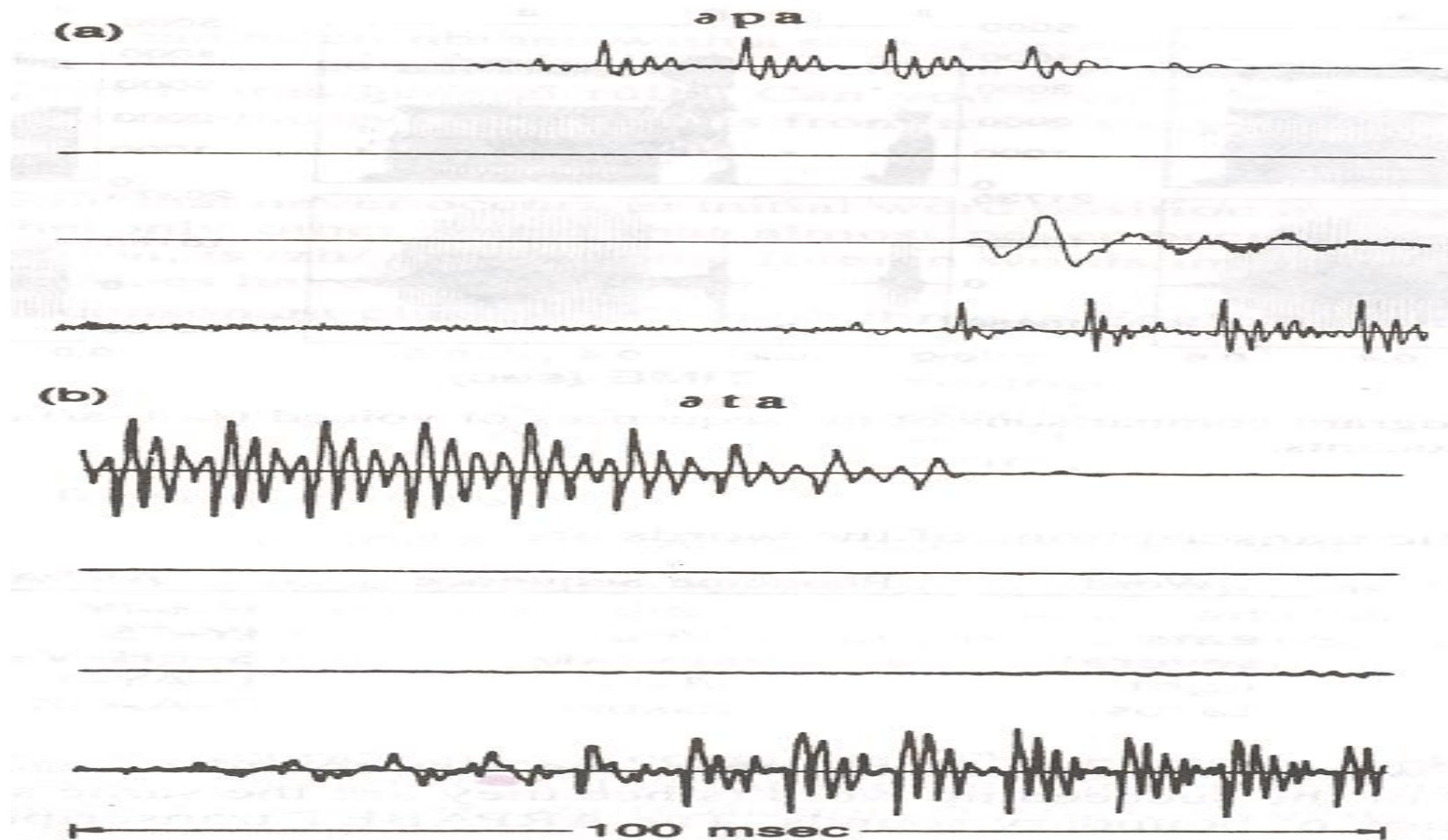


Figure 2.27 Waveforms for the sequences /ə-p-a/ and /ə-t-a/.

Spectrograms of Stop Consonants

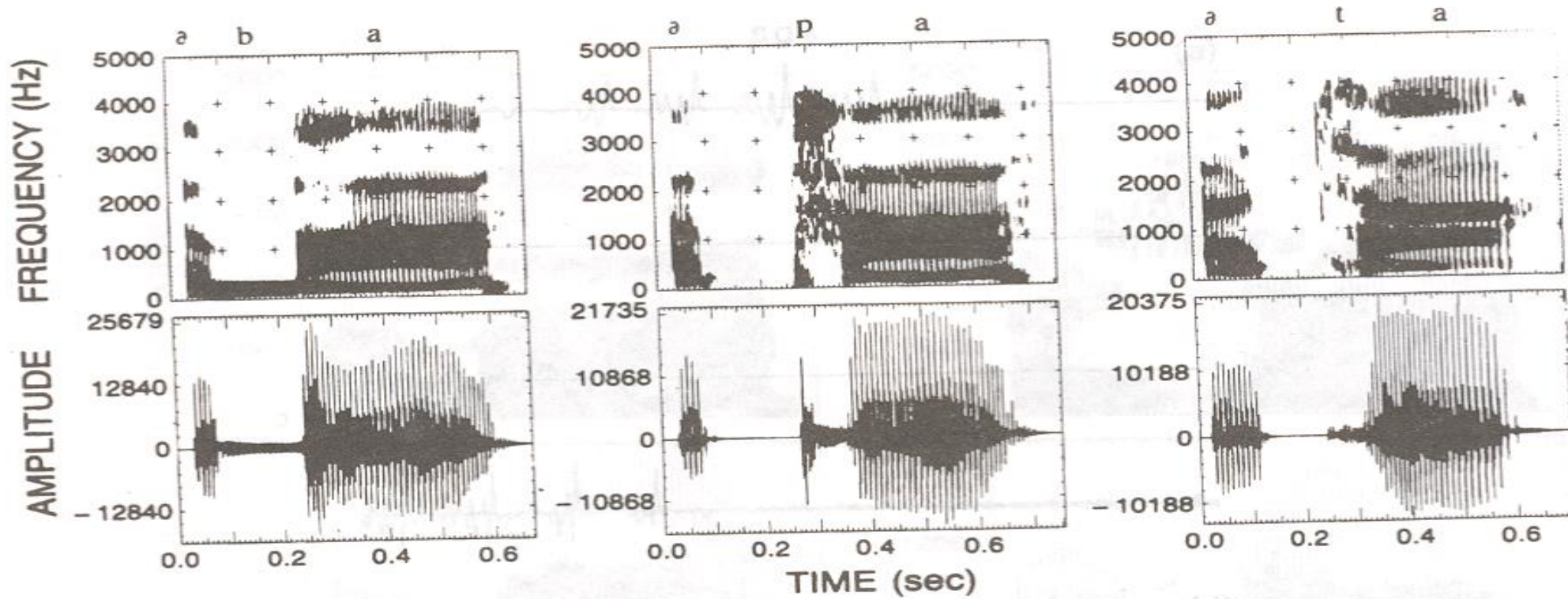
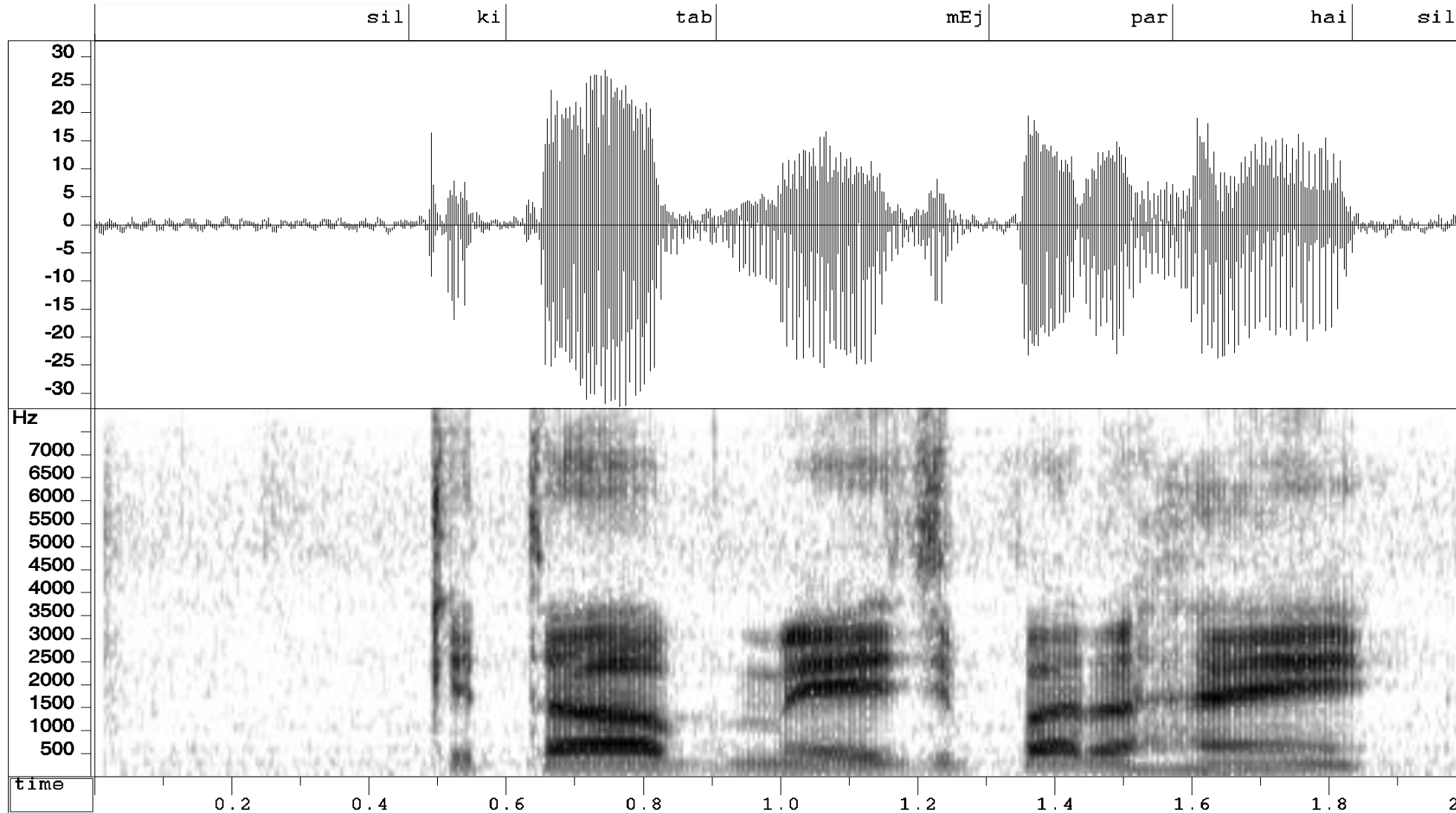


Figure 2.28 Spectrogram comparisons of the sequences of voiced (/ə-b-a/) and voiceless (/ə-p-a/ and /ə-t-a/) stop consonants.

Time & Freq Domain Rep. of Speech of a Sentence



Spectrograms of Isolated Digits

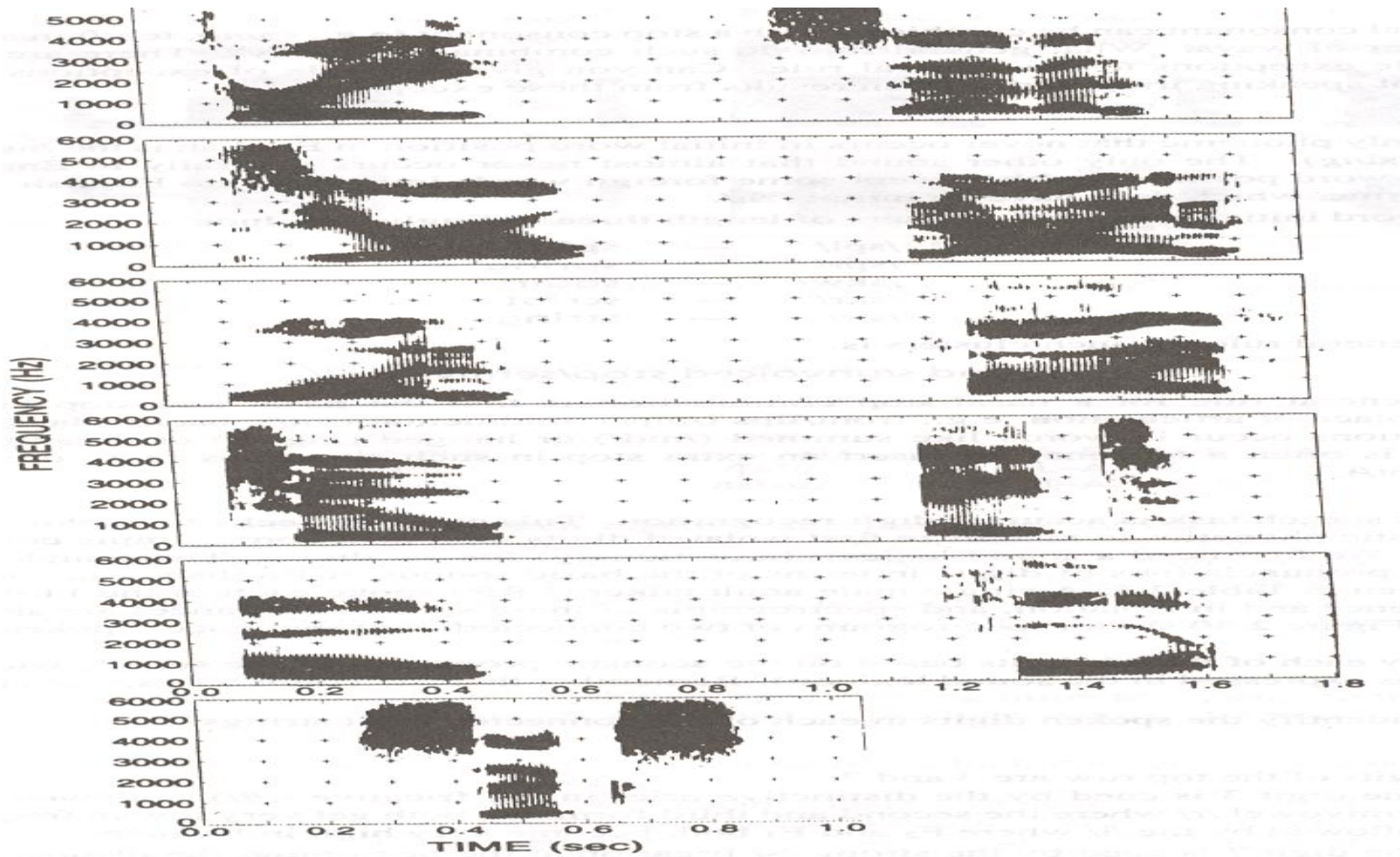


Figure 2.29 Spectrograms of the 11 isolated digits, 0 through 9 plus oh, in random sequence.

Spectrograms of Two Connected Digit Sequences

TABLE 2.3. Sound Lexicon of Digits

Word	Sounds	ARPABET
Zero	/z I r o/	Z-IH-R-OW
One	/w ʌ n/	W-AH-N
Two	/t u/	T-UW
Three	/θ r i/	TH-R-IY
Four	/f o r/	F-OW-R
Five	/f aʲ v/	F-AY-V
Six	/s I k s/	S-IH-K-S
Seven	/s ɛ v ə n/	S-EH-V-AX-N
Eight	/eʲ t/	EY-T
Nine	/n aʲ n/	N-AY-N
Oh	/o/	OW

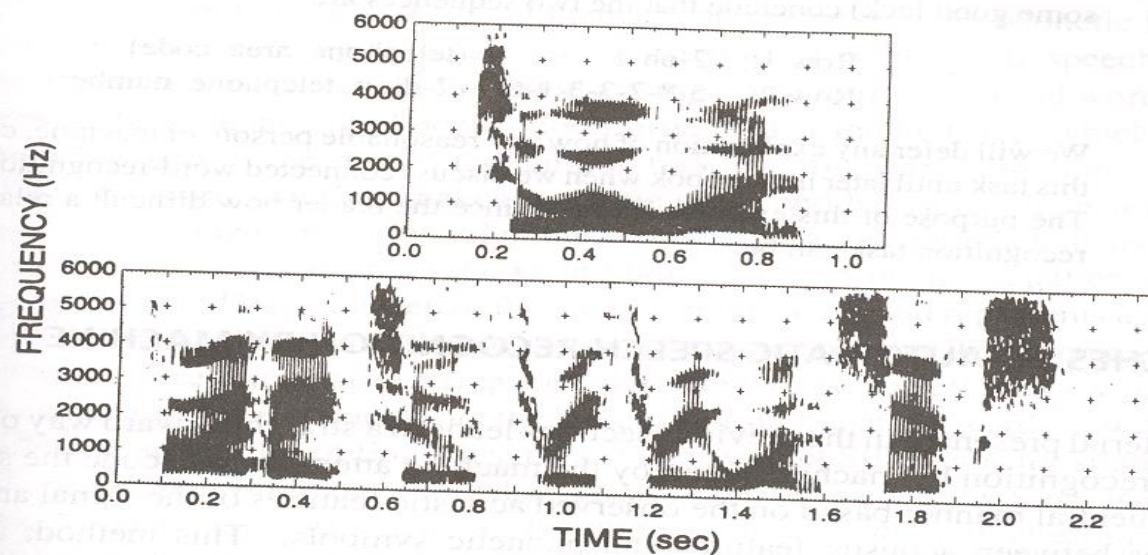


Figure 2.30 Spectrograms of two connected digit sequences.