

Speech Sounds of American English

- There are over 40 speech sounds in American English which can be organized by their basic manner of production

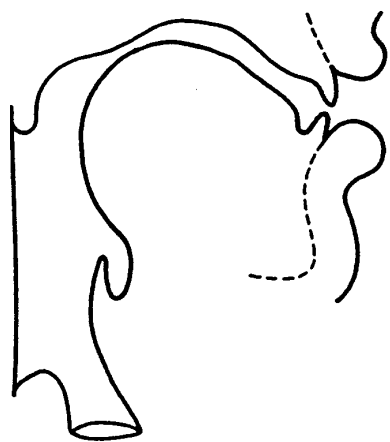
Manner Class	Number
Vowels	18
Fricatives	8
Stops	6
Nasals	3
Semivowels	4
Affricates	2
Aspirant	1

- Vowels, glides, and consonants differ in degree of constriction
- **Sonorant** consonants have no pressure build up at constriction
- **Nasal** consonants lower the velum allowing airflow in nasal cavity
- **Continuant** consonants do not block airflow in oral cavity

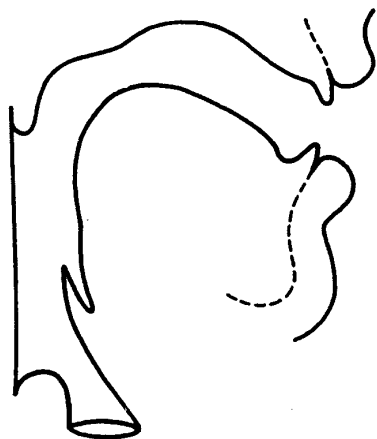
Vowel Production

- No significant constriction in the vocal tract
- Usually produced with periodic excitation
- Acoustic characteristics depend on the position of the jaw, tongue, and lips

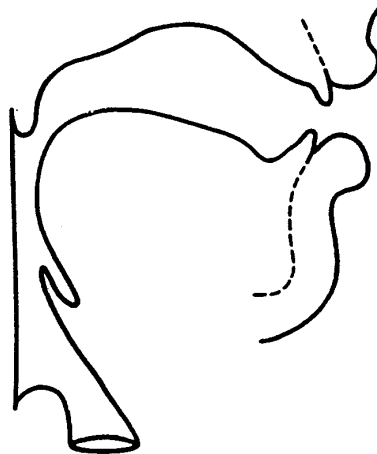
[i]



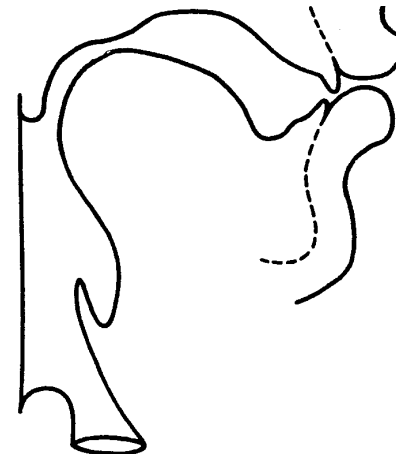
[æ]



[ɑ]



[u]



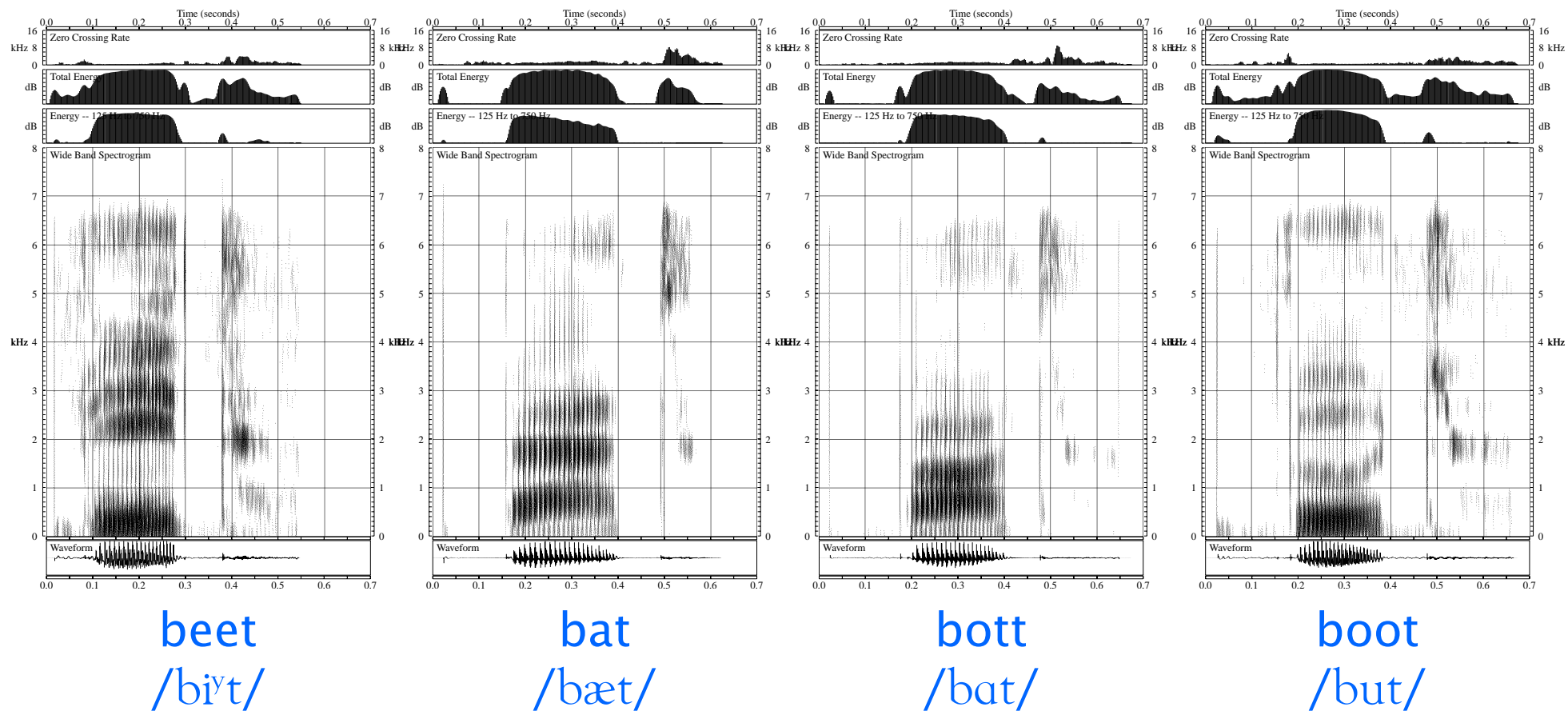
Vowels of American English

- There are approximately 18 vowels in American English made up of monothongs, diphthongs, and reduced vowels (schwa's)

/i ^y /	iy	beat	/ɔ/	ao	bought	/ɑ ^y /	ay	bite
/ɪ/	ih	bit	/ʌ/	ah	but	/ɔ ^y /	oy	Boyd
/e ^y /	ey	bait	/o ^w /	ow	boat	/ɑ ^w /	aw	bout
/ɛ/	eh	bet	/ʊ/	uh	book	[ə]	ax	about
/æ/	ae	bat	/u/	uw	boot	[ɪ]	ix	roses
/ɑ/	aa	Bob	/ɜ̃/	er	Bert	[ə̃]	axr	butter

- They are often described by the articulatory features: **High/Low**, **Front/Back**, **Retroflexed**, **Rounded**, and **Tense/Lax**

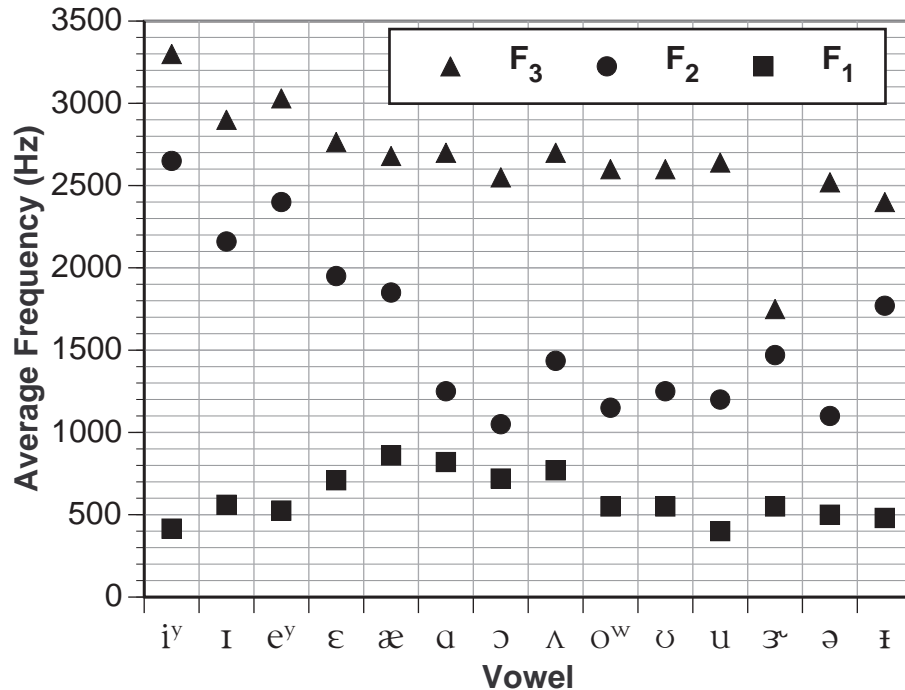
Spectrograms of the Cardinal Vowels



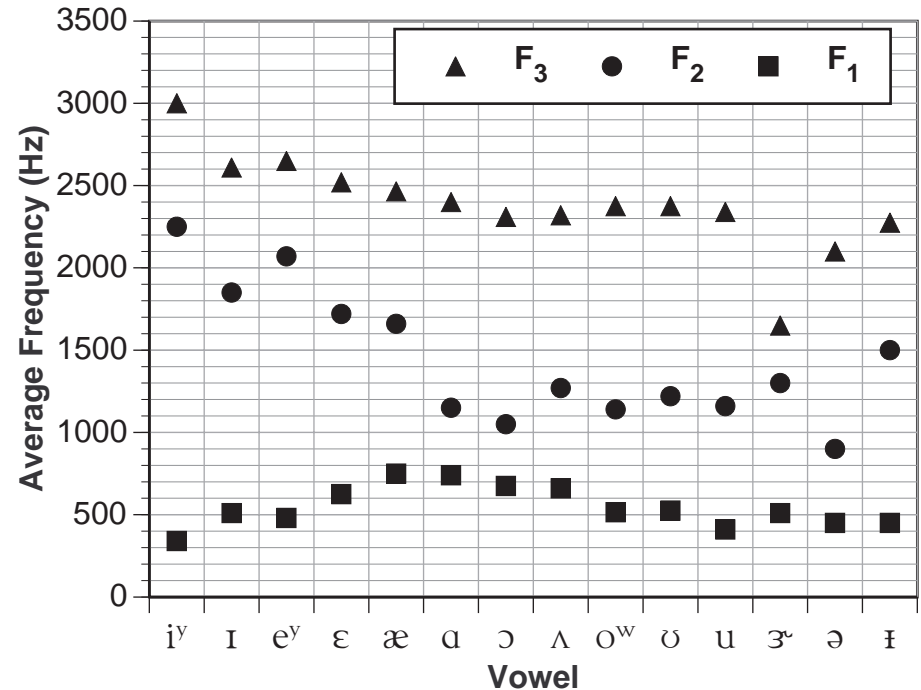
Vowel Formant Averages

- Vowels are often characterized by the lower three formants
- **High/Low** is correlated with the first formant, F1
- **Front/Back** is correlated with the second formant, F2
- **Retroflexion** is marked by a low third formant, F3

Female Speakers



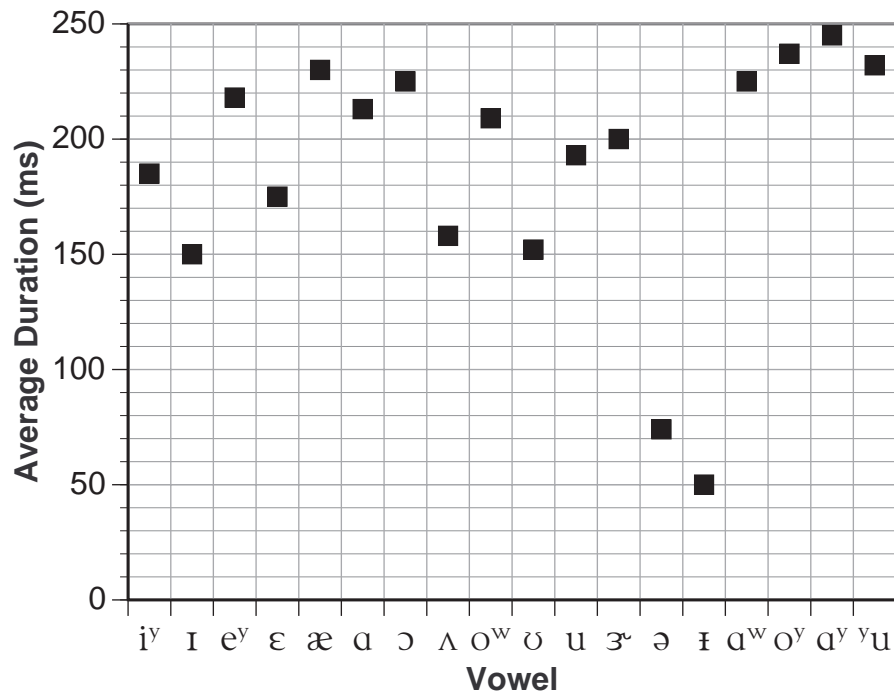
Male Speakers



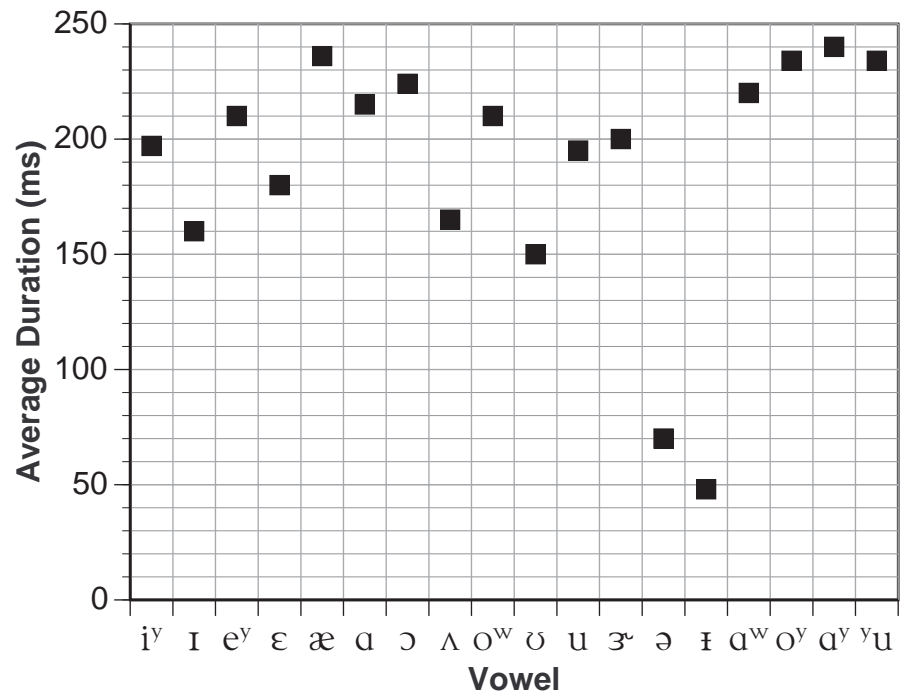
Vowel Durations

- Each vowel has a different intrinsic duration
- Schwa's have distinctly shorter durations (50ms)
- /ɪ, ɛ, ʌ, ʊ/ are the shortest monothongs
- Context can greatly influence vowel duration

Female Speakers

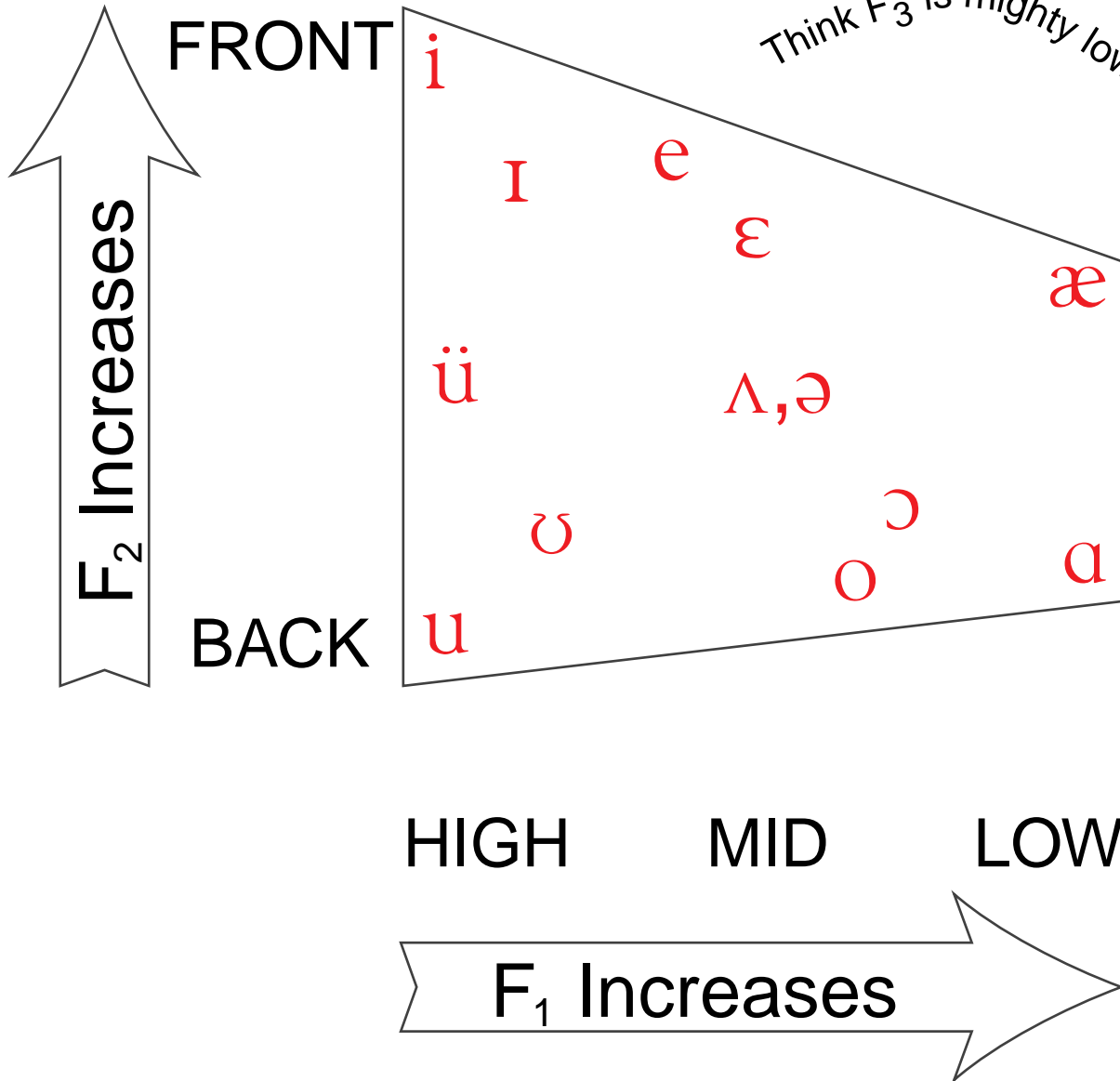


Male Speakers



Happy Little Vowel Chart

"So inaccurate, yet so useful."



Think F_3 is mighty low? Your pal æ is the way to go!

TENSE = Towards Edges
tends to be longer

LAX = Towards Center
tends to be shorter

SCHWAS:

Plain [ə] About [əbɑʊt]

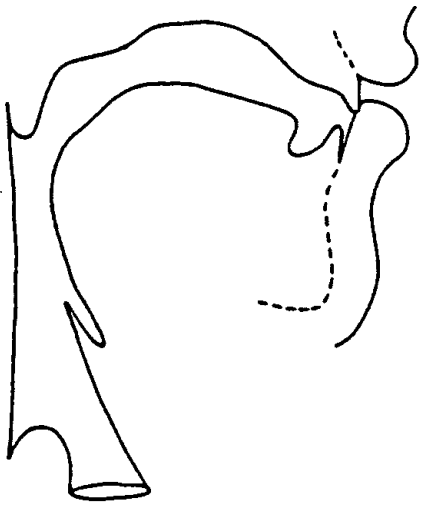
Front [ɪ] Roses [rɔʊzɪz]

Retroflex [ɻ] Forever [fæʀevə]

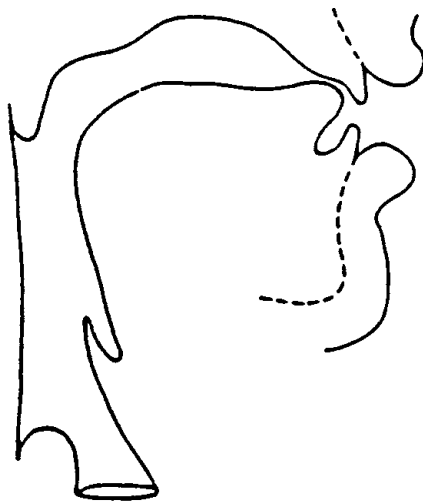
Fricative Production

- Turbulence produced at narrow constriction
- Constriction position determines acoustic characteristics
- Can be produced with periodic excitation

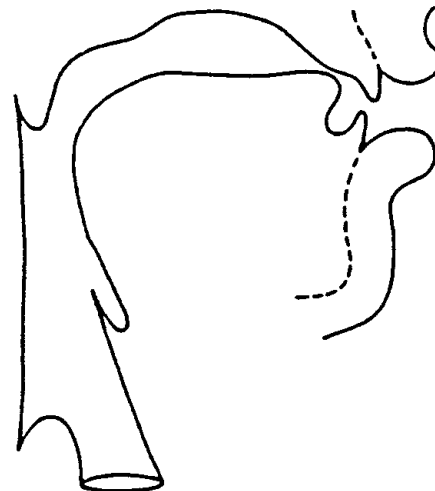
[f]



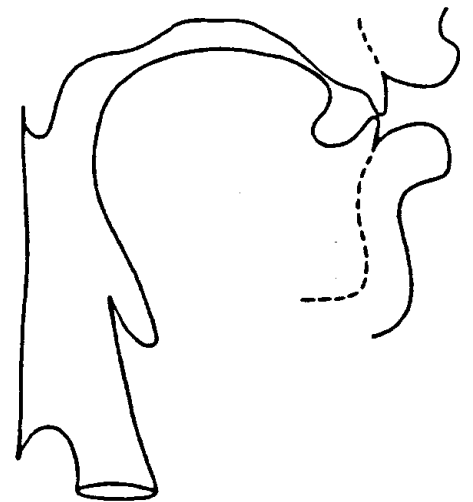
[θ]



[s]



[š]

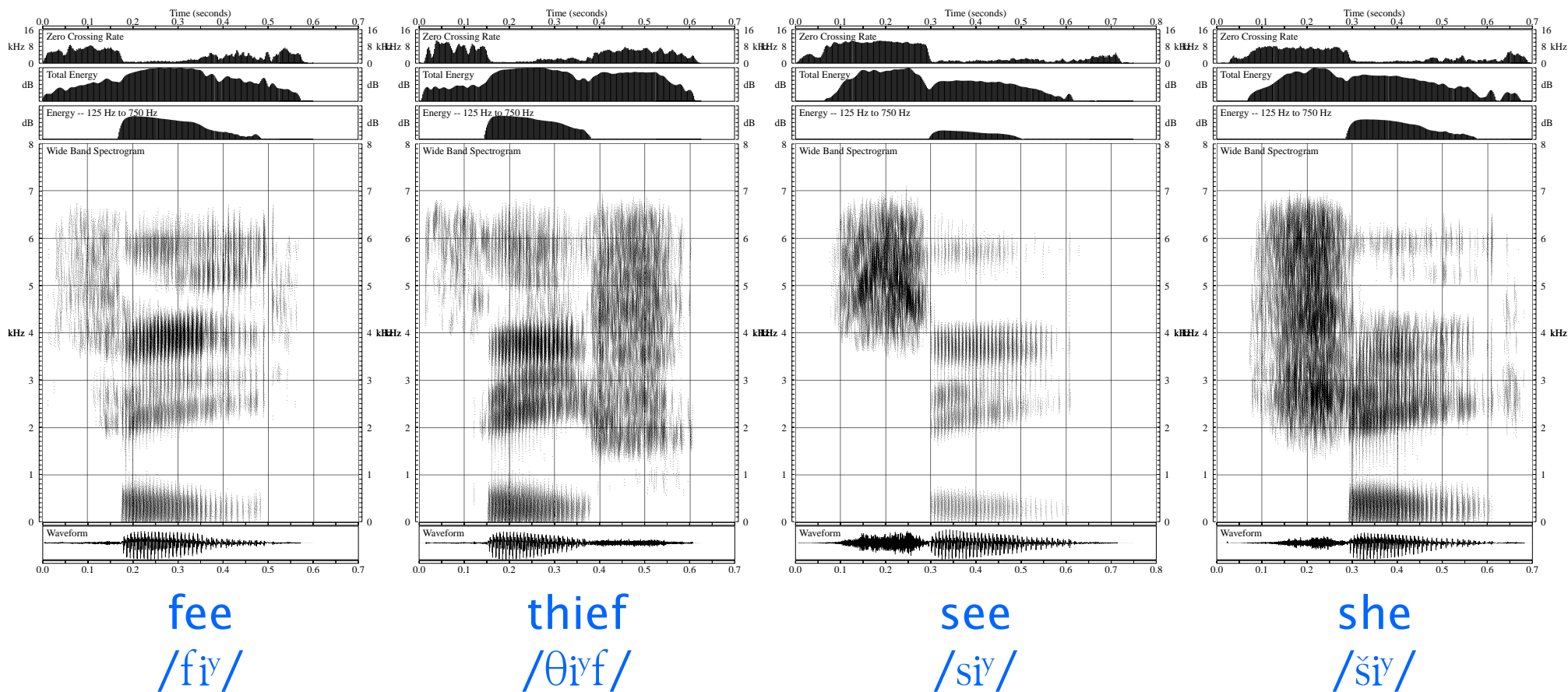


Fricatives of American English

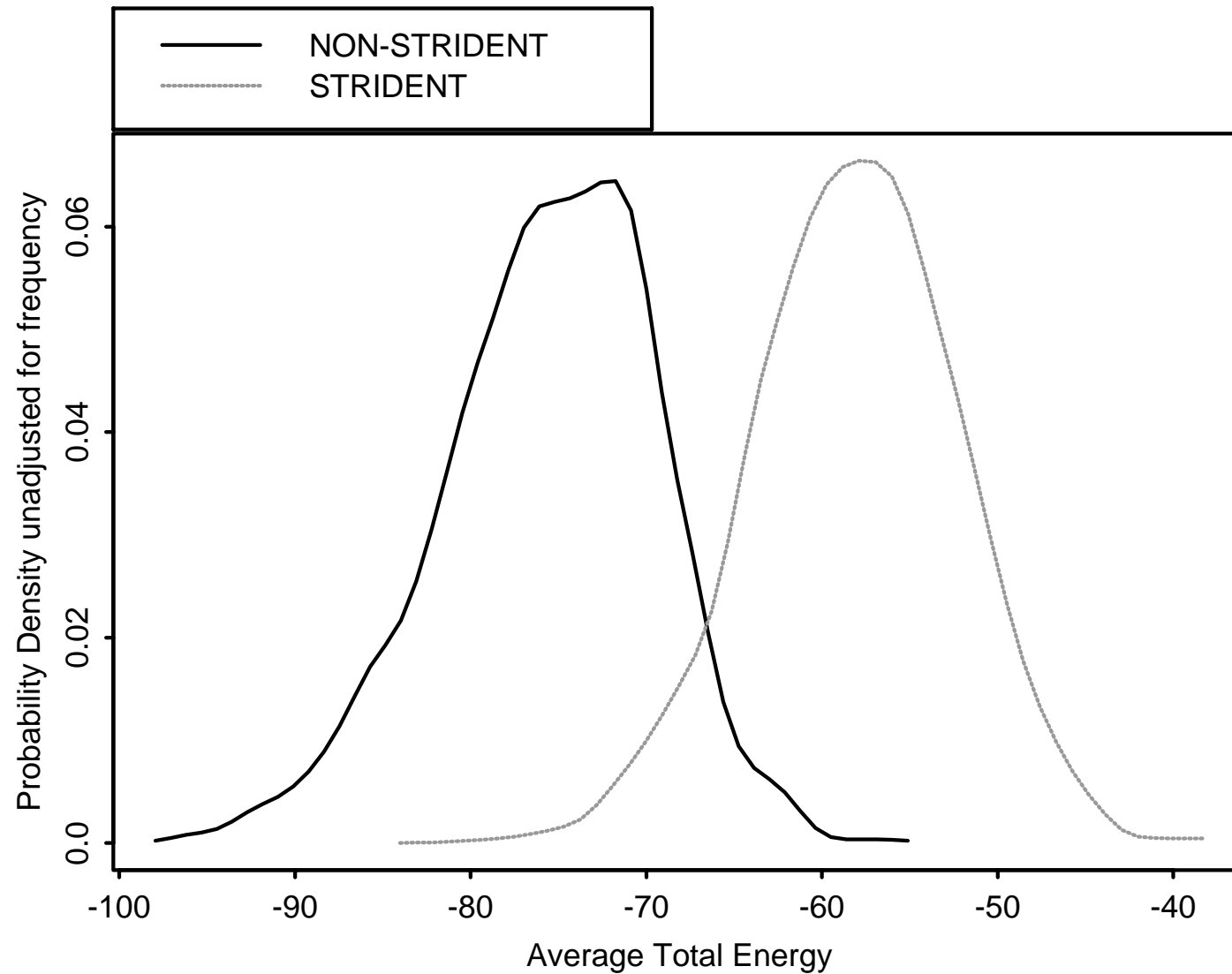
- There are 8 fricatives in American English
- Four places of articulation: **Labio-Dental** (Labial), **Interdental** (Dental), **Alveolar**, and **Palato-Alveolar** (Palatal)
- They are often described by the features **Voiced/Unvoiced**, or **Strident/Non-Strident** (constriction behind alveolar ridge)

Type	Unvoiced			Voiced		
Labial	/f/	f	fee	/v/	v	v
Dental	/θ/	th	thief	/ð/	dh	thee
Alveolar	/s/	s	see	/z/	z	z
Palatal	/ʃ/	sh	she	/ʒ/	zh	Gigi

Spectrograms of Unvoiced Fricatives

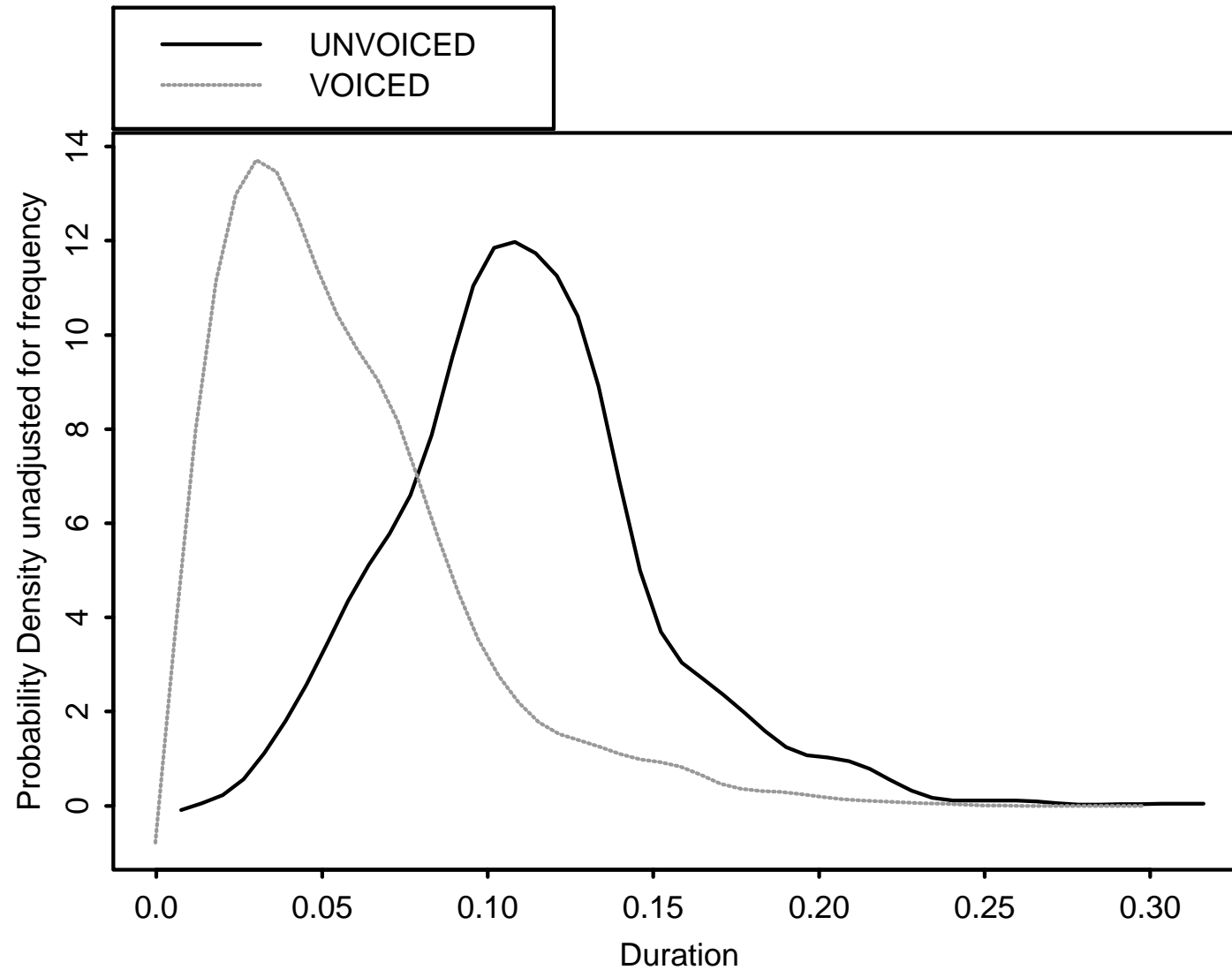


Fricative Energy



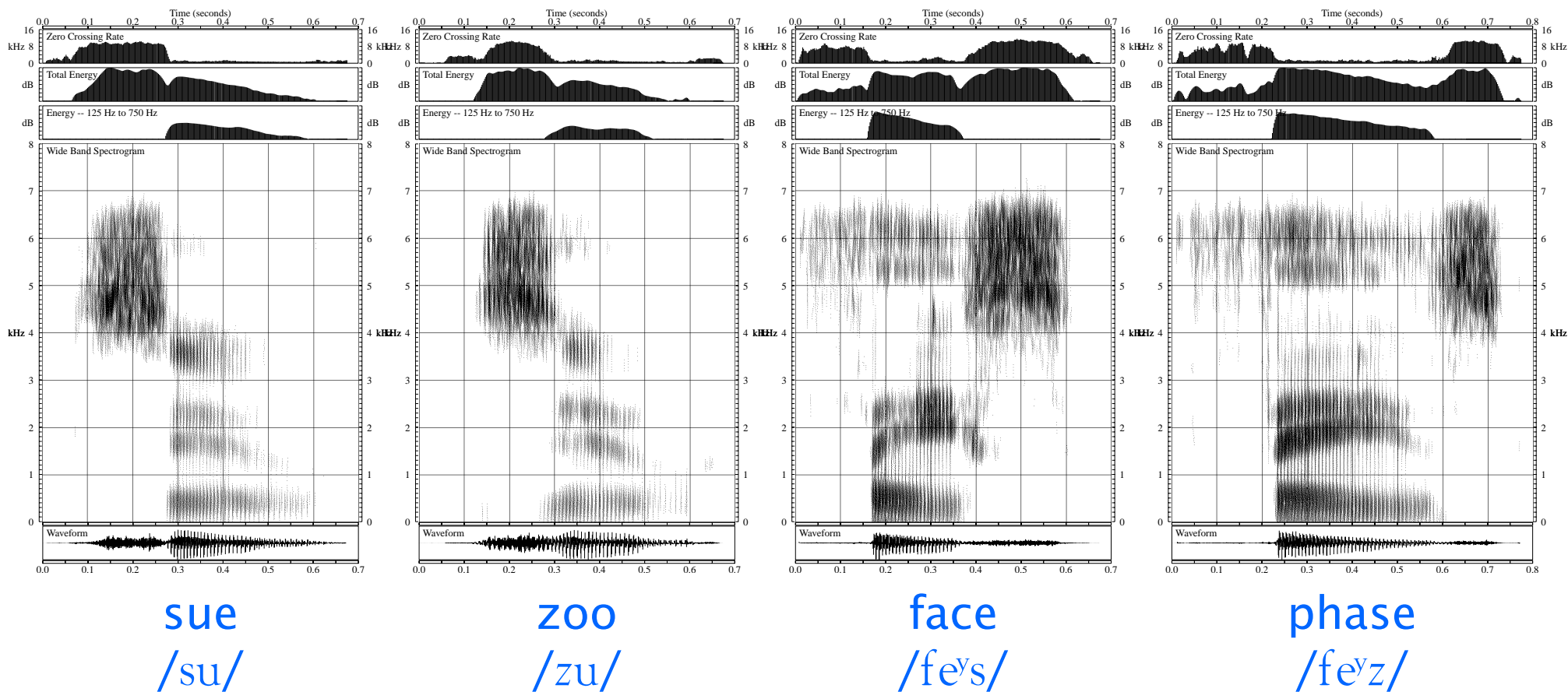
Strident fricatives tend to be stronger than non-strident fricatives.

Fricative Durations



Voiced fricatives tend to be shorter than unvoiced fricatives.

Examples of Fricative Voicing Contrast



Friendly Little Consonant Chart

"Somewhat more accurate, yet somewhat less useful."

		Place of Articulation				
		Labial	Dental	Alveolar	Palatal	Velar
Manner of Articulation	Stop	p b		t d		k g
	Fricative	f v Weak (Non-strident)	θ ð Non-strident	s z Strong (Strident)	š ž Strident	
	Nasal	m		n		ŋ

Voicing: Unvoiced Voiced

The Semi-vowels:

- y is like an extreme i
- w is like an extreme u
- l is like an extreme o
- r is like an extreme ɜ

The Odds and Ends:

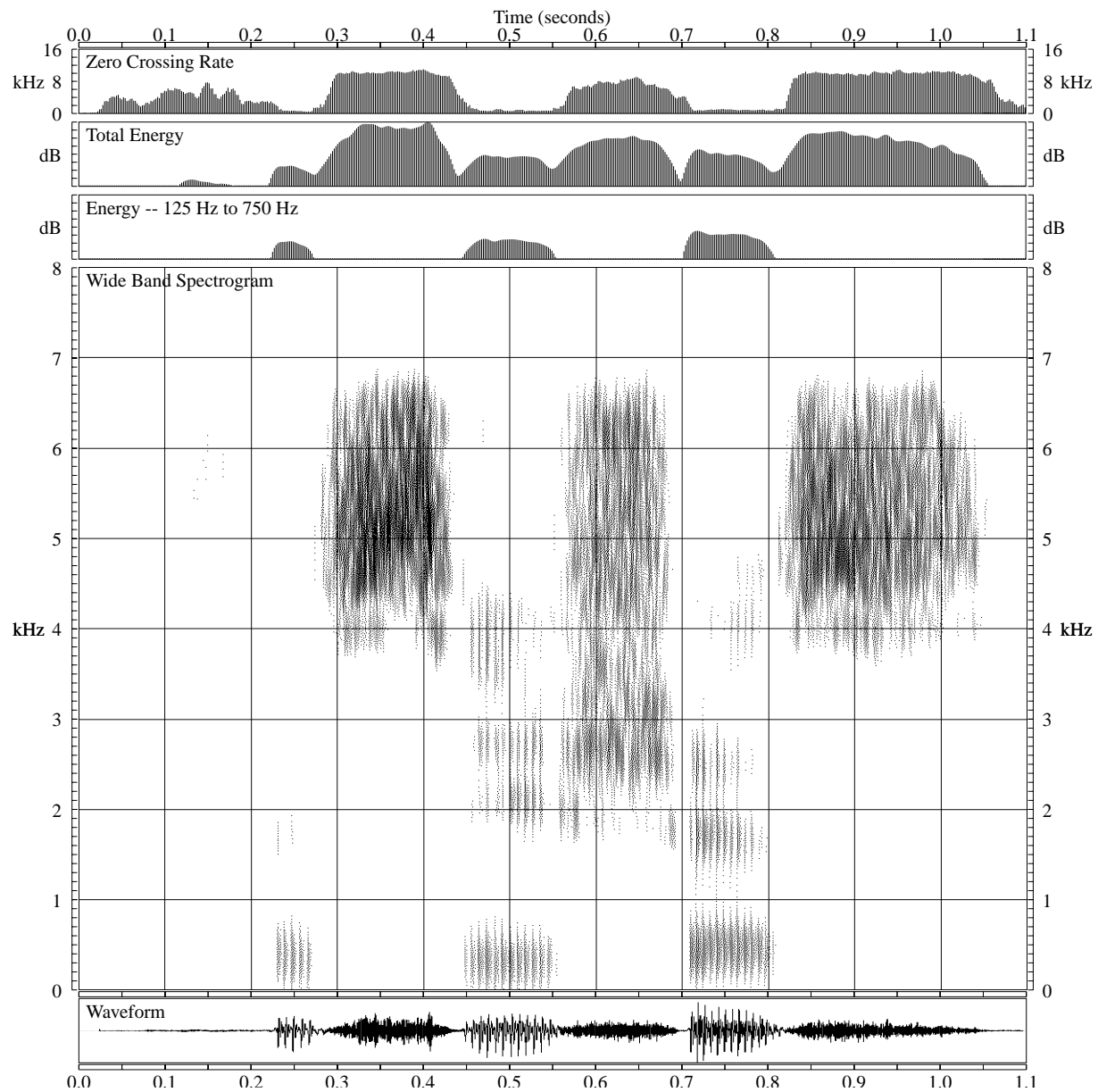
- h (unvoiced h)
- ɦ (voiced h)
- ɾ (flap) ɽ (nasalized flap)
- ʔ (glottal stop)

The Affricates:

- č is like t+š
- ǰ is like d+ž

MIT

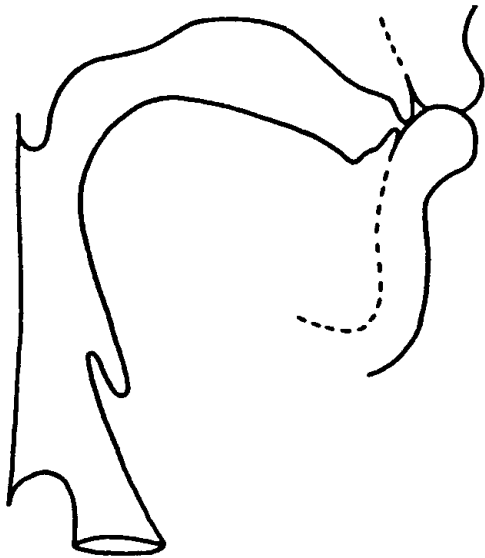
What is this word?



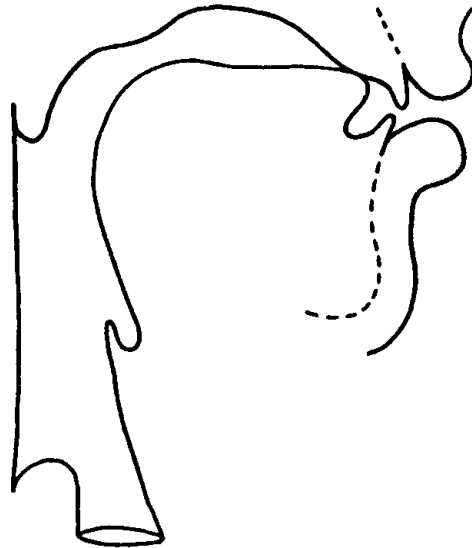
Stop Production

- Complete closure in the vocal tract, pressure build up
- Sudden release of the constriction, turbulence noise
- Can have periodic excitation during closure

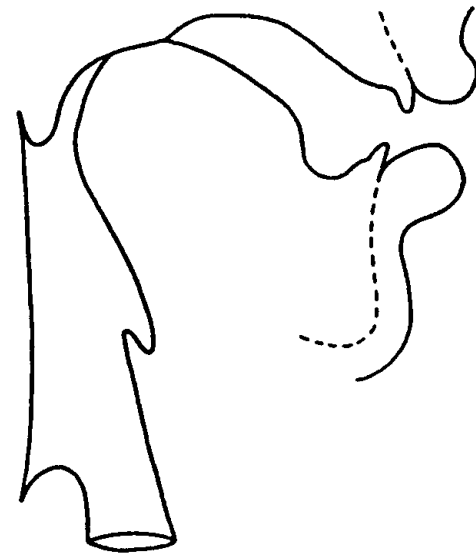
[b]



[d]



[g]



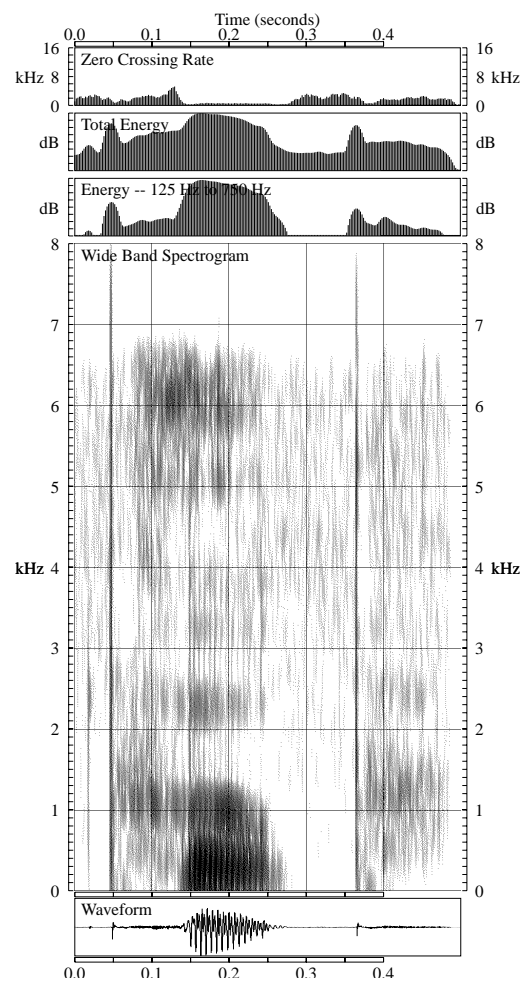
MIT Stops of American English

- There are 6 stop consonants in American English
- Three places of articulation: **Labial**, **Alveolar**, and **Velar**
- Each place of articulation has a voiced and unvoiced stop

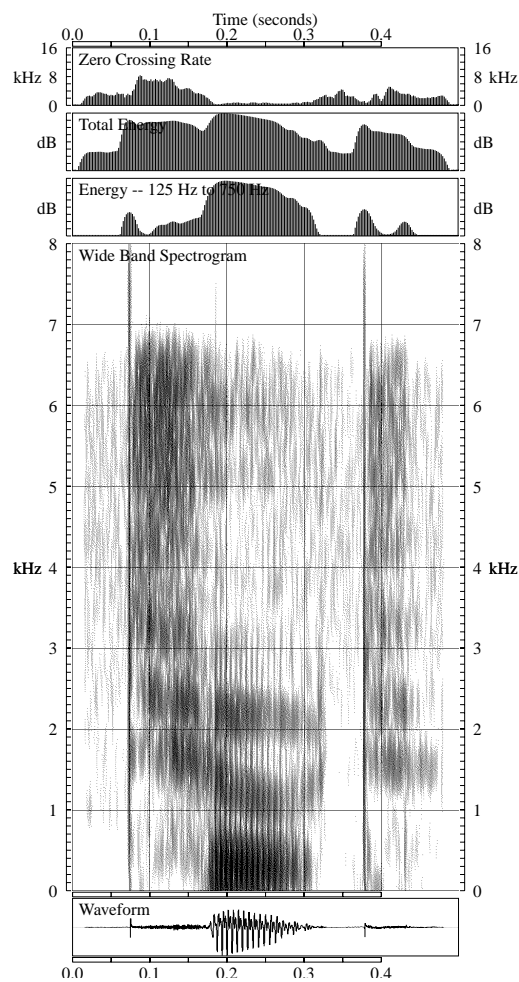
Type	Voiced			Unvoiced		
Labial	/b/	b	bought	/p/	p	pot
Alveolar	/d/	d	dot	/t/	t	tot
Velar	/g/	g	got	/k/	k	cot

- Unvoiced stops are typically aspirated
- Voiced stops usually exhibit a “voice-bar” during closure
- Information about formant transitions and release useful for classification

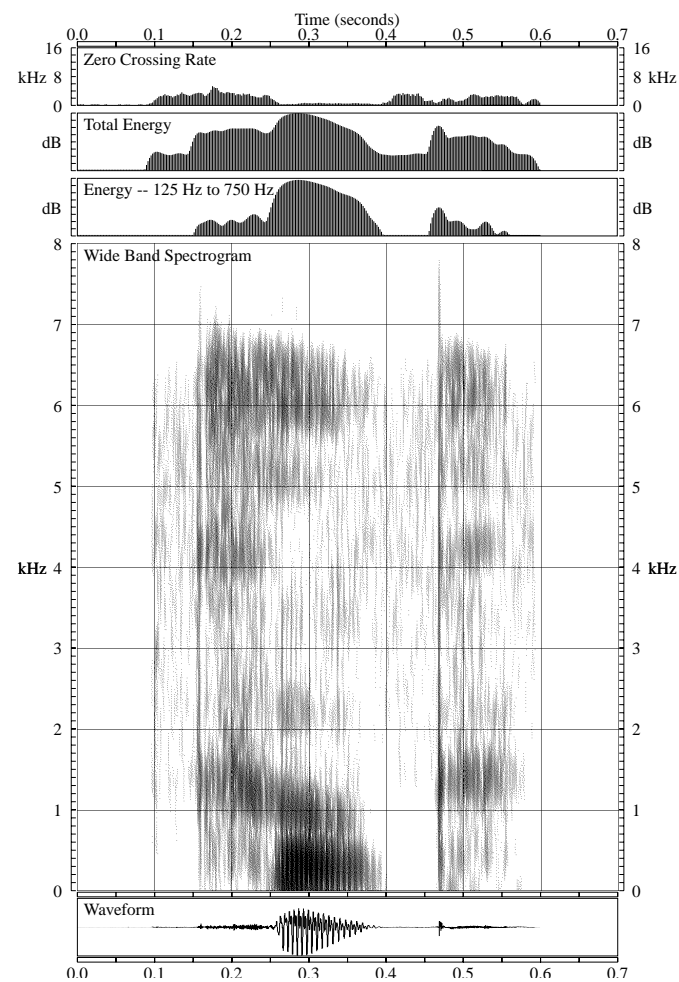
Spectrograms of Unvoiced Stops



poop
/pup/

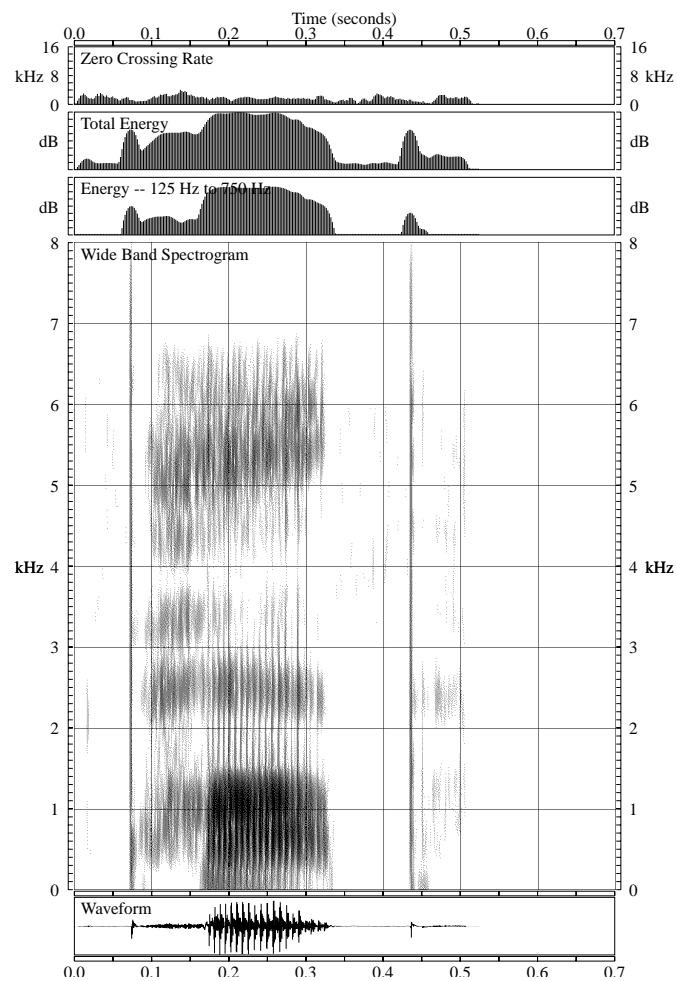


toot
/tut/

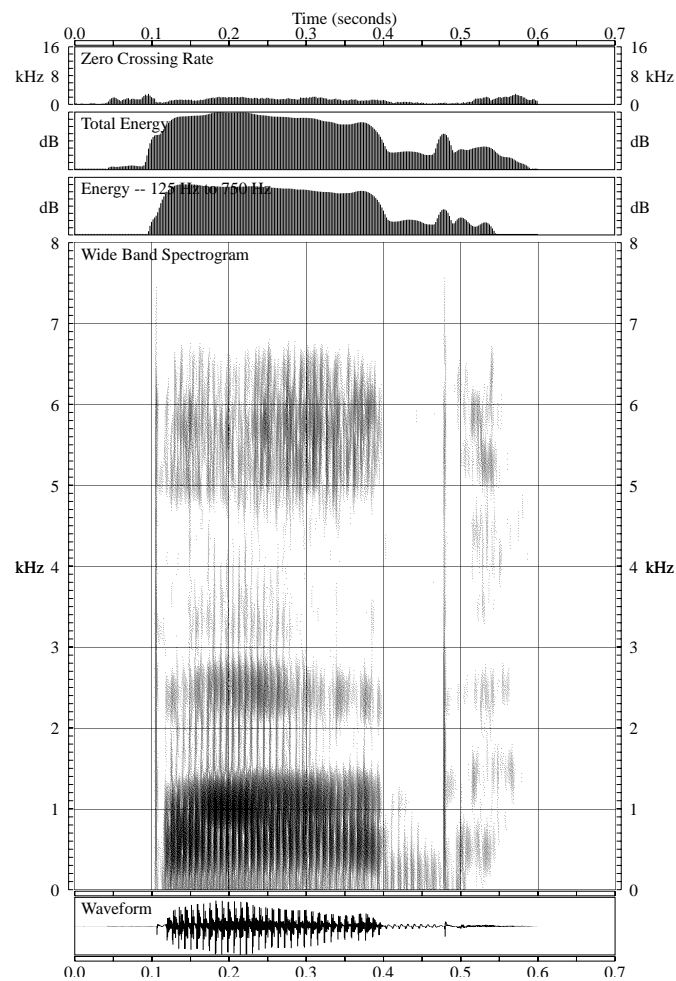


kook
/kuk/

Examples of Stop Voicing Contrast

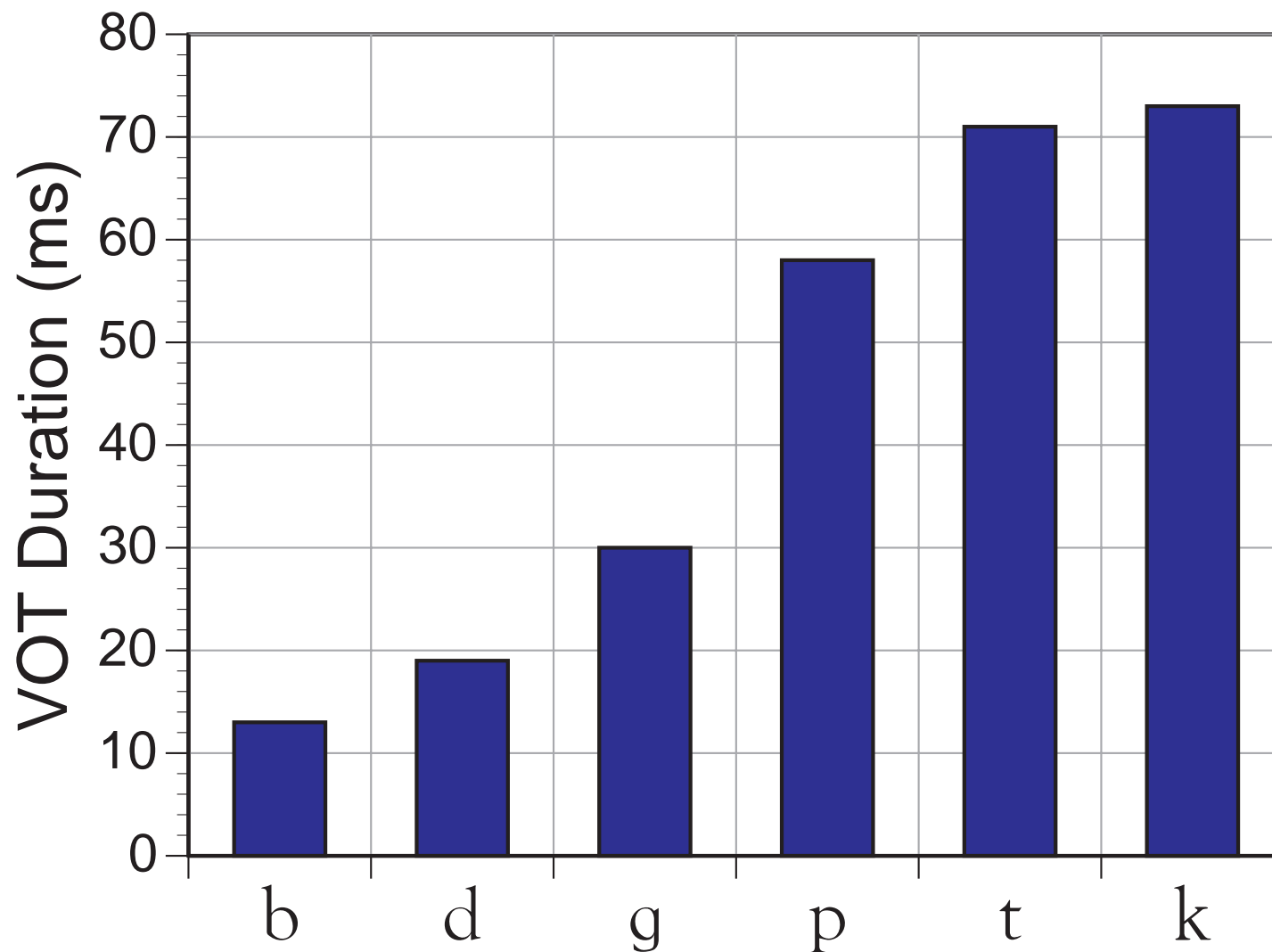


pop
/pap/



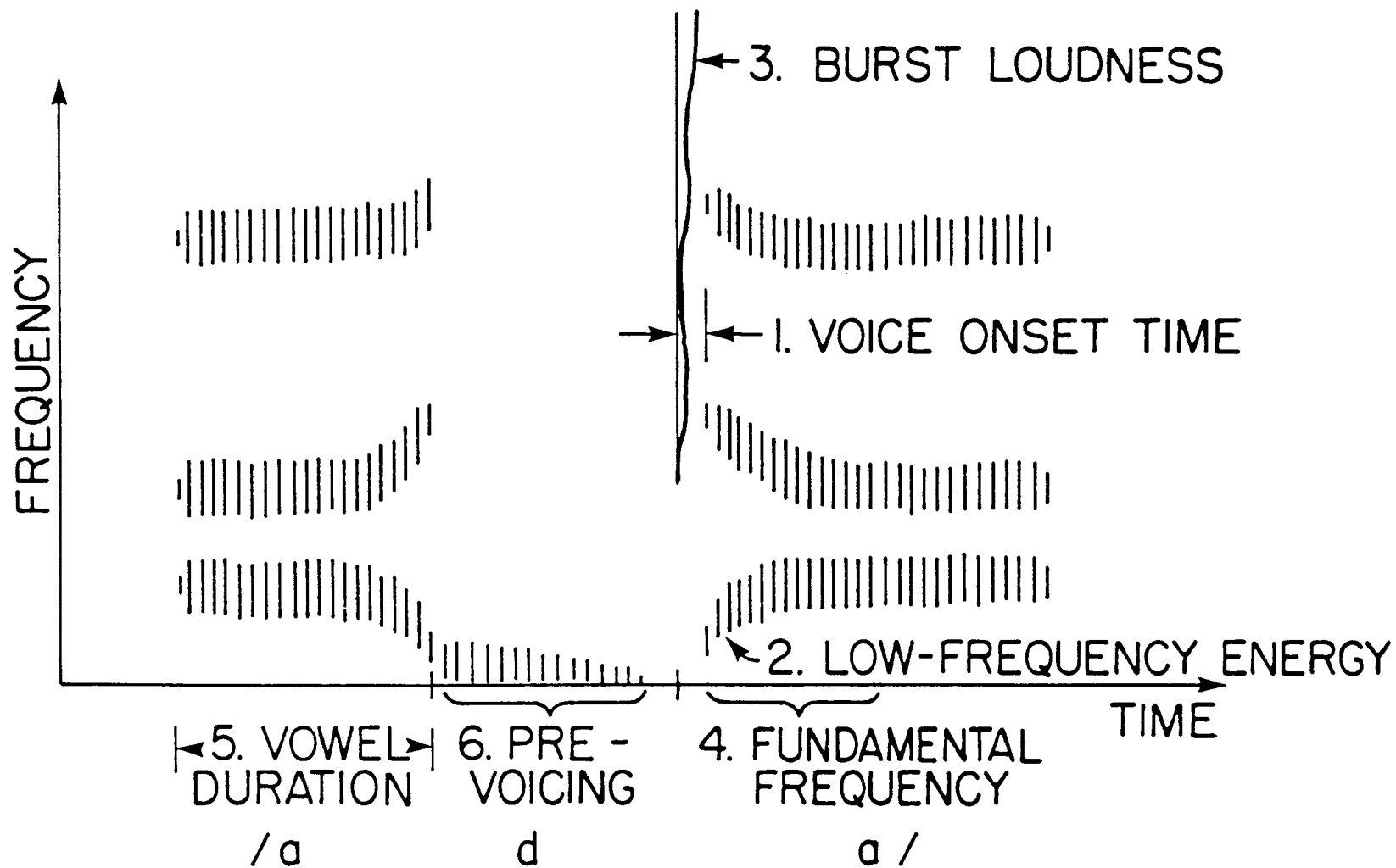
bob
/bab/

Singleton Stop Durations



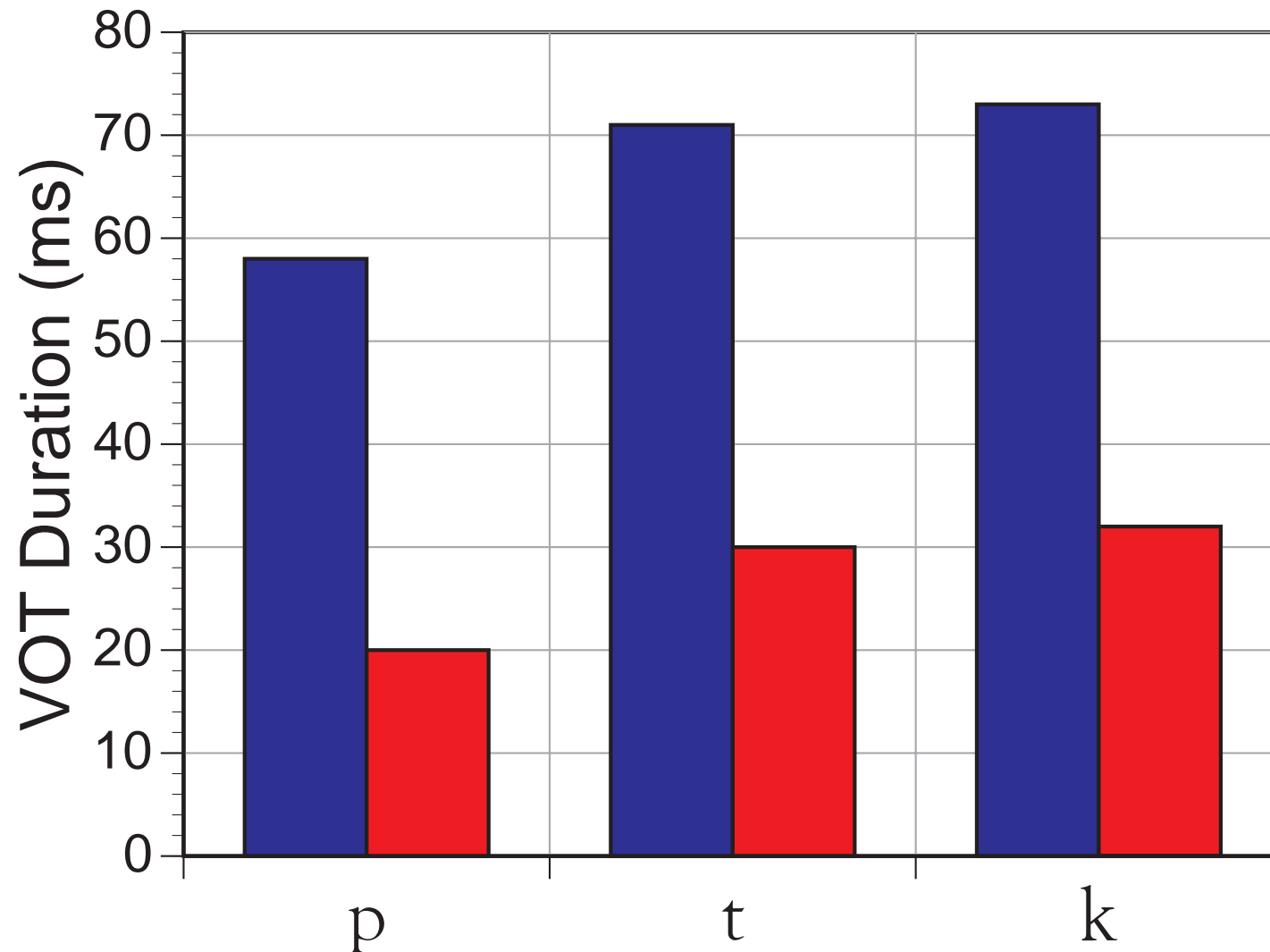
Voice onset times (VOTs) are longer for unvoiced stops.

Voicing Cues for Stops



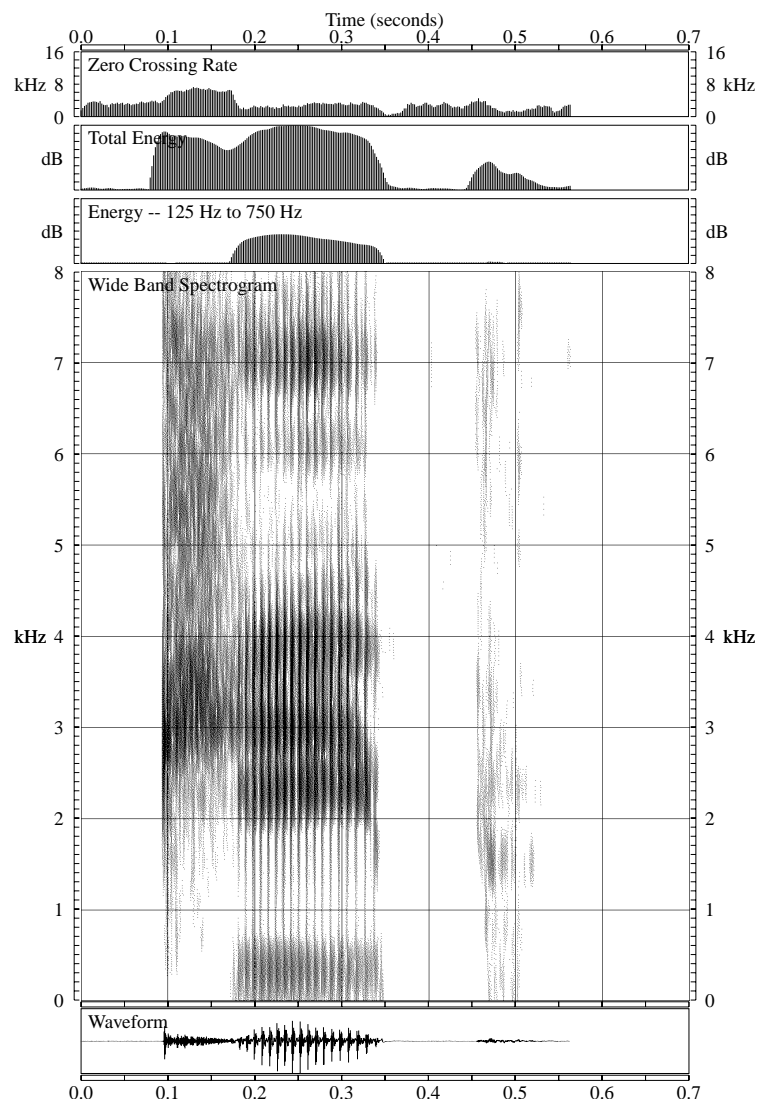
There are many voicing cues for a stop.

MIT /s/-Stop Durations

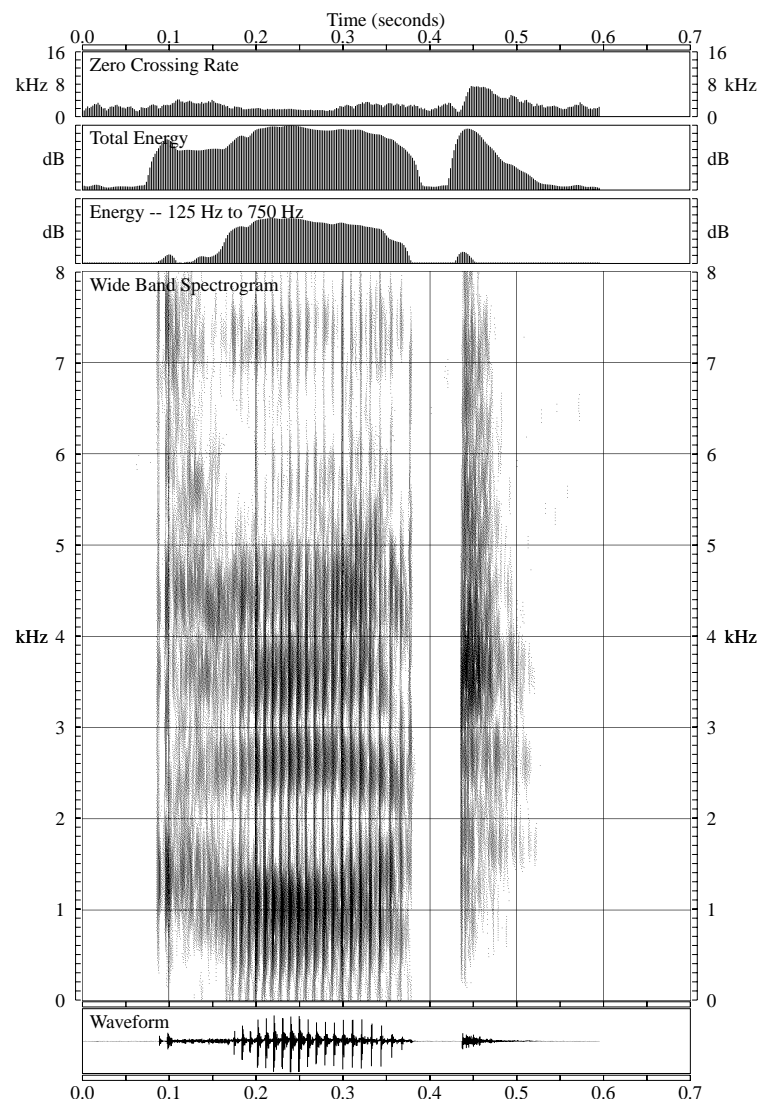


Unvoiced stops are unaspirated in /s/ stop sequences.

Examples of Front and Back Velars



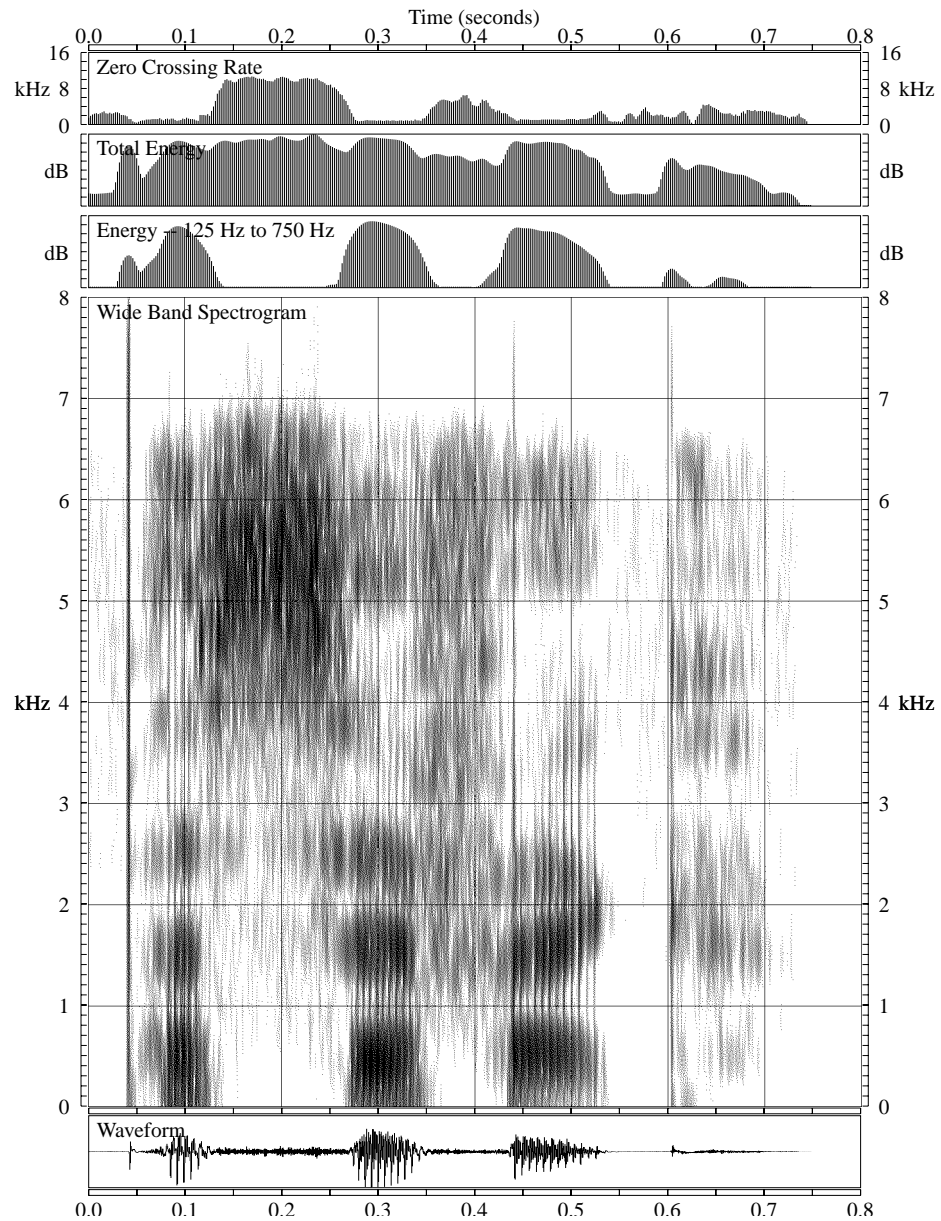
keep
/kiːp/



cot
/kɒt/

MIT

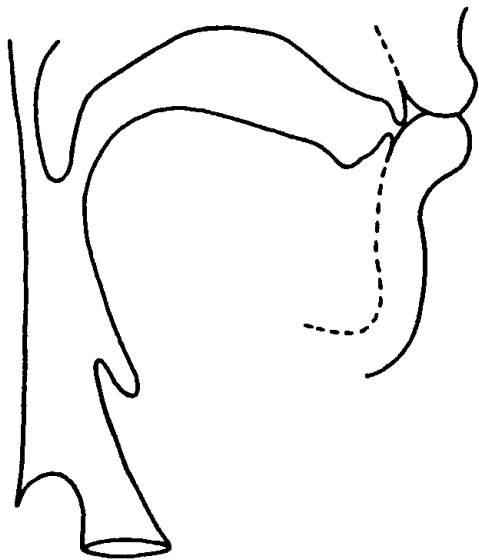
What is this word?



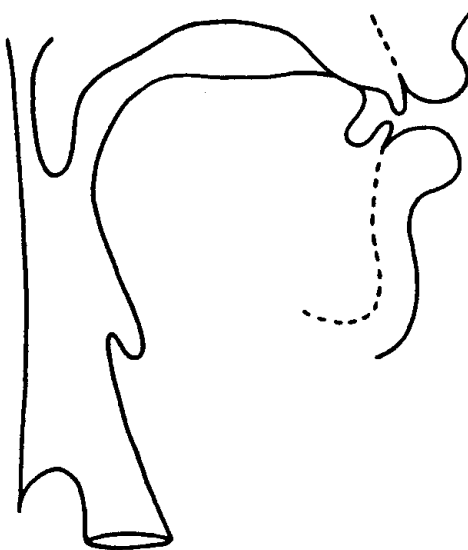
Nasal Production

- Velum lowering results in airflow through nasal cavity
- Consonants produced with closure in oral cavity
- Nasal murmurs have similar spectral characteristics

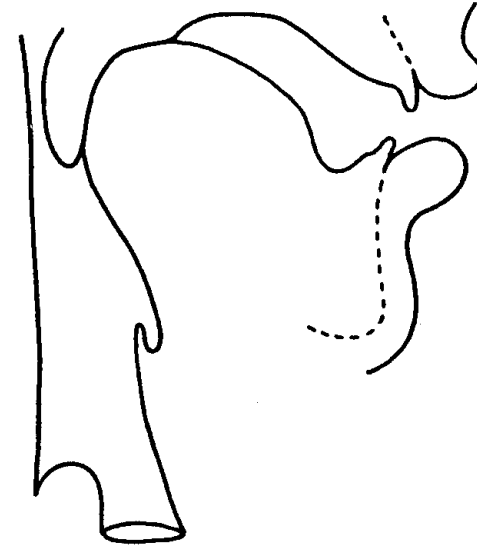
[m]



[n]



[ŋ]



MIT

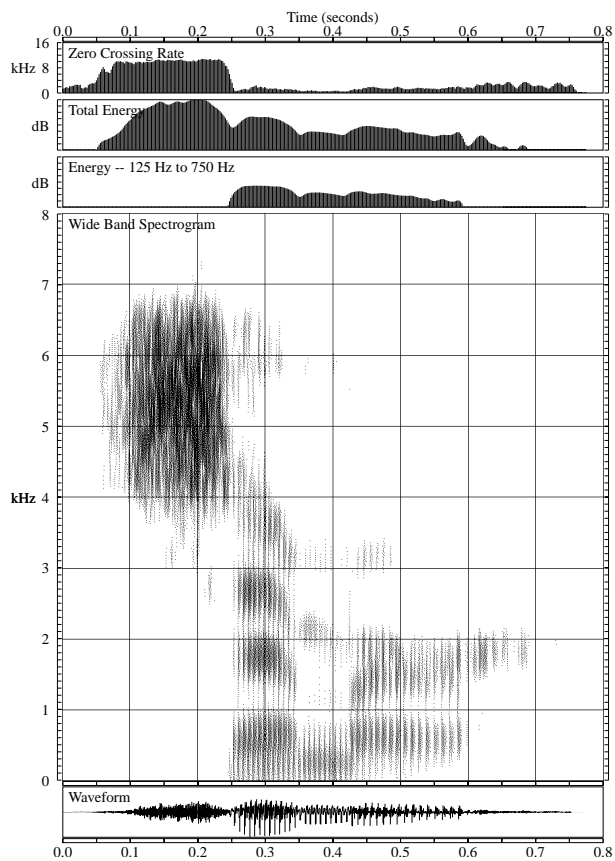
Nasal of American English

- Three places of articulation: **Labial**, **Alveolar**, and **Velar**

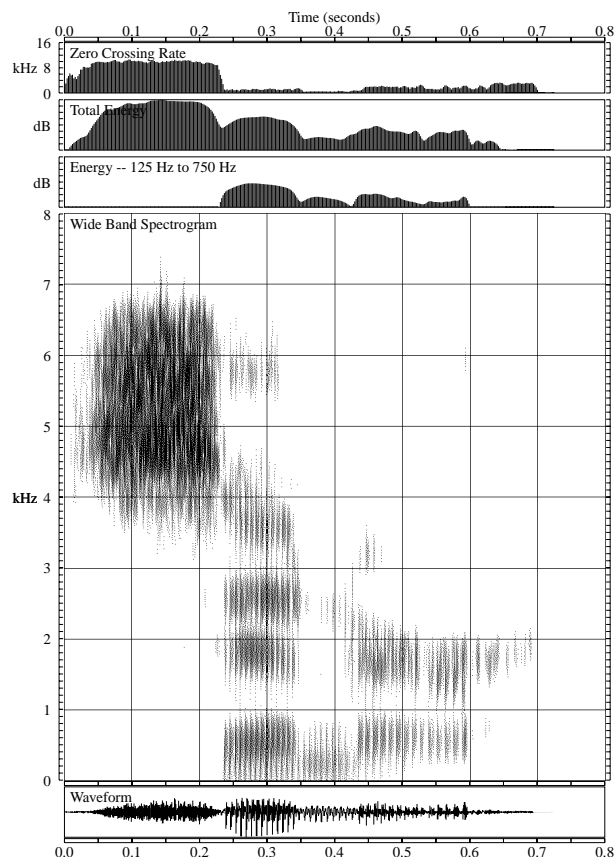
Type	Nasal		
Labial	/m/	m	me
Alveolar	/n/	n	knee
Velar	/ŋ/	ng	sing

- Nasal consonants are always attached to a vowel, though can form an entire syllable in unstressed environments ([n̩], [m̩], [ŋ̩])
- /ŋ/ is always post-vocalic in English
- Place identified by neighboring formant transitions

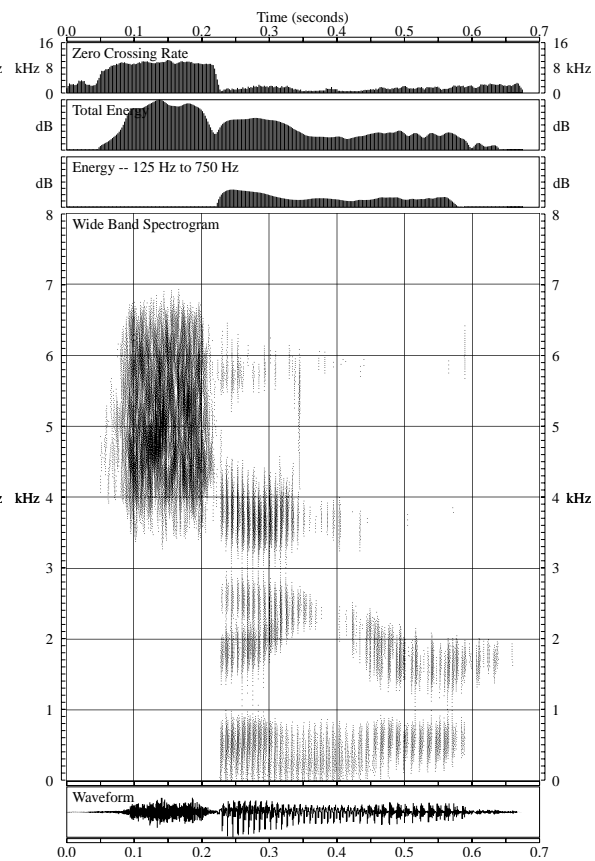
MIT Spectrograms of Nasals



simmer
/sɪmɜː/



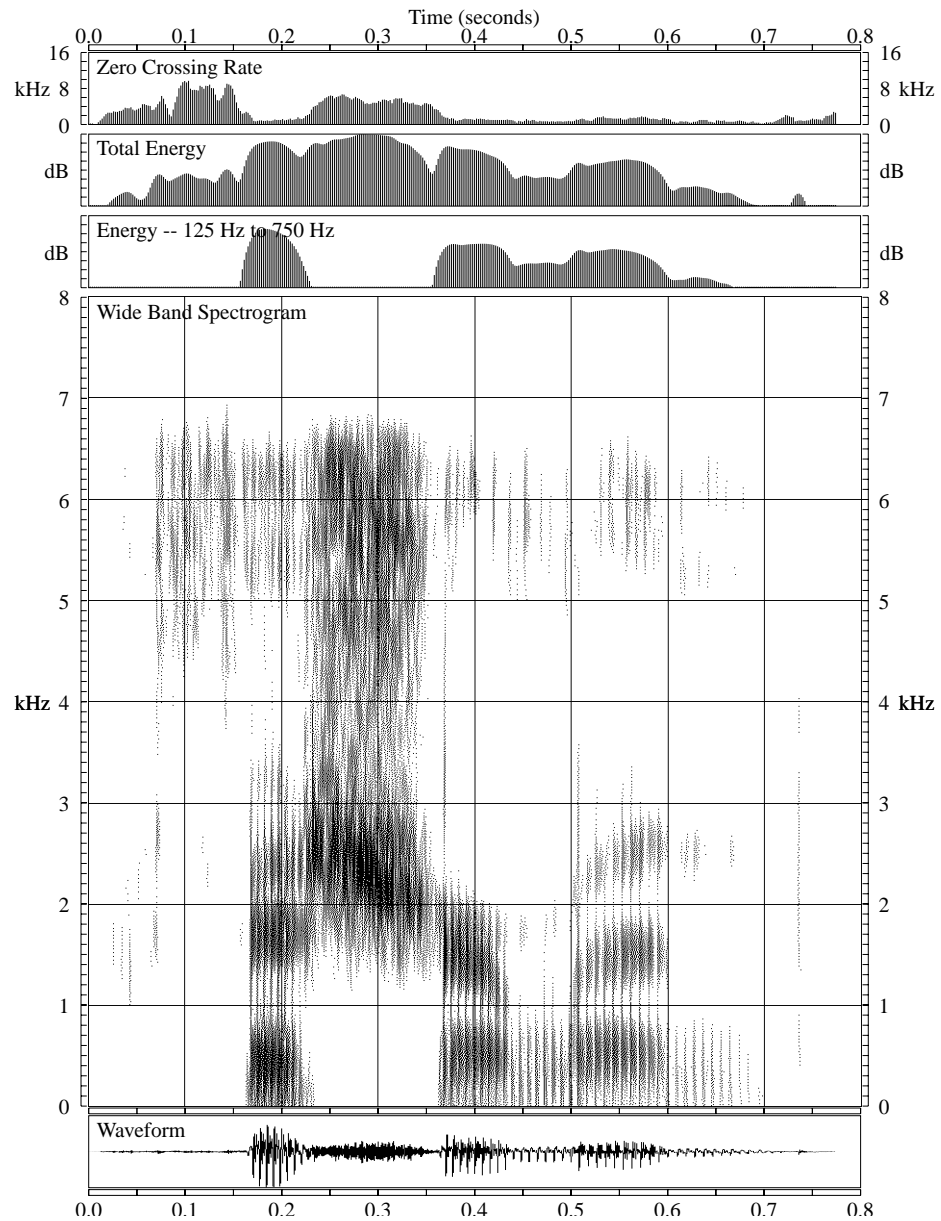
sinner
/sɪnɜː/



singer
/sɪŋɜː/

MIT

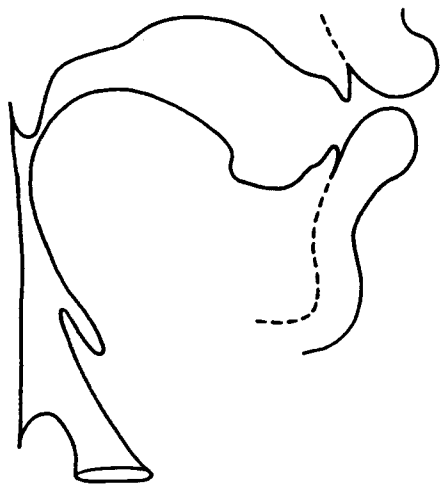
What is this word?



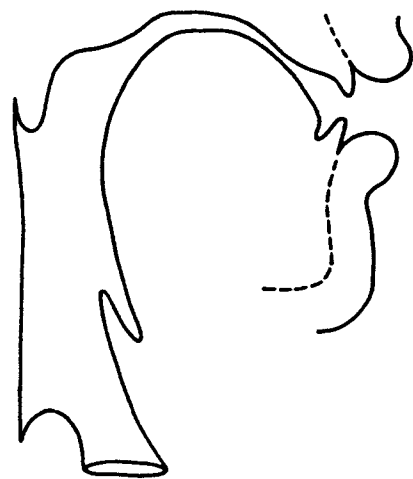
Semivowel Production

- Constriction in vocal tract, no turbulence
- Slower articulatory motion than other consonants
- **Laterals** form complete closure with tongue tip, airflow via sides of constriction

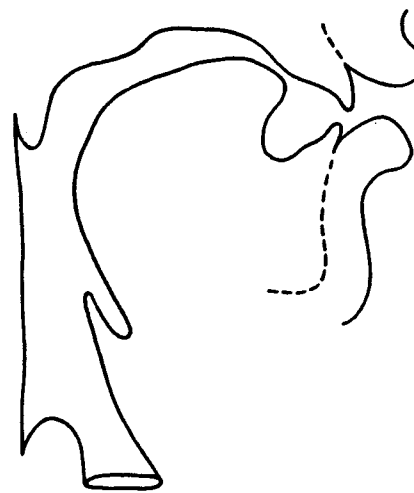
[w]



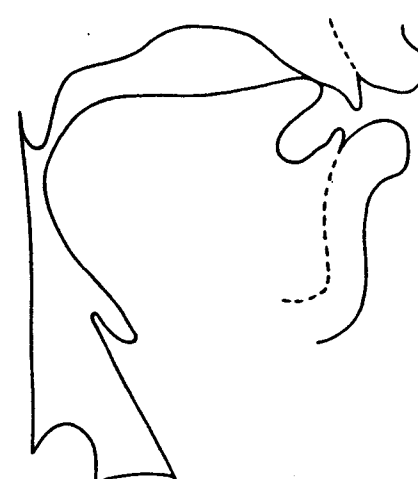
[y]



[r]



[l]



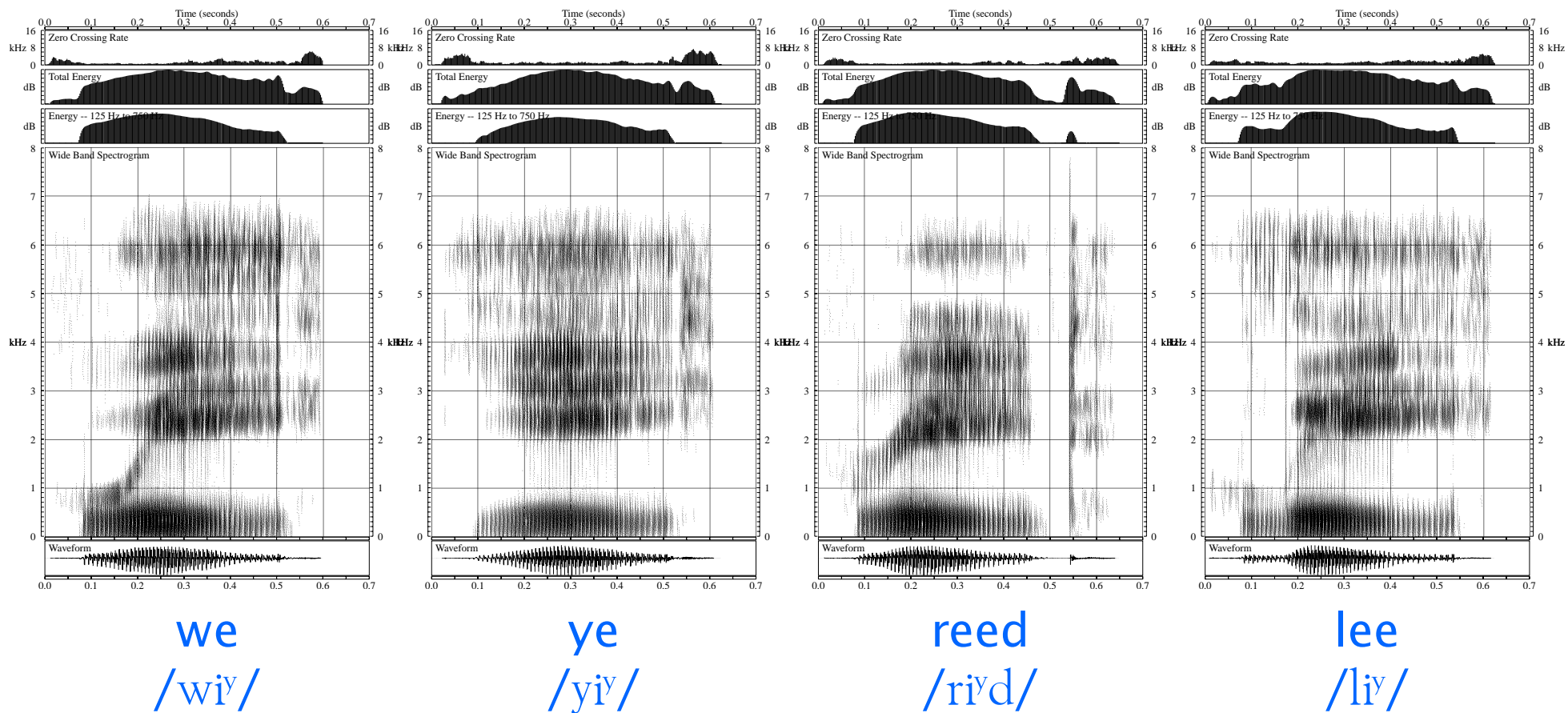
Semivowels of American English

- There are 4 semivowels in American English
- Sometimes referred to as **Liquids** or **Glides**

Type	Semivowel			Nearest Vowel
Glides	/w/	w	wet	/u/
	/y/	y	yet	/i/
Liquids	/r/	r	red	/ɜ̃/
	/l/	l	let	/o/

- Glides are a more extreme articulation of a corresponding vowel
 - Similar, though more extreme, formant positions
 - Generally weaker due to narrower constriction
- Semivowels are always attached to a vowel, though /l/ can form an entire syllable in unstressed environments ([l̩])

Spectrograms of Semivowels

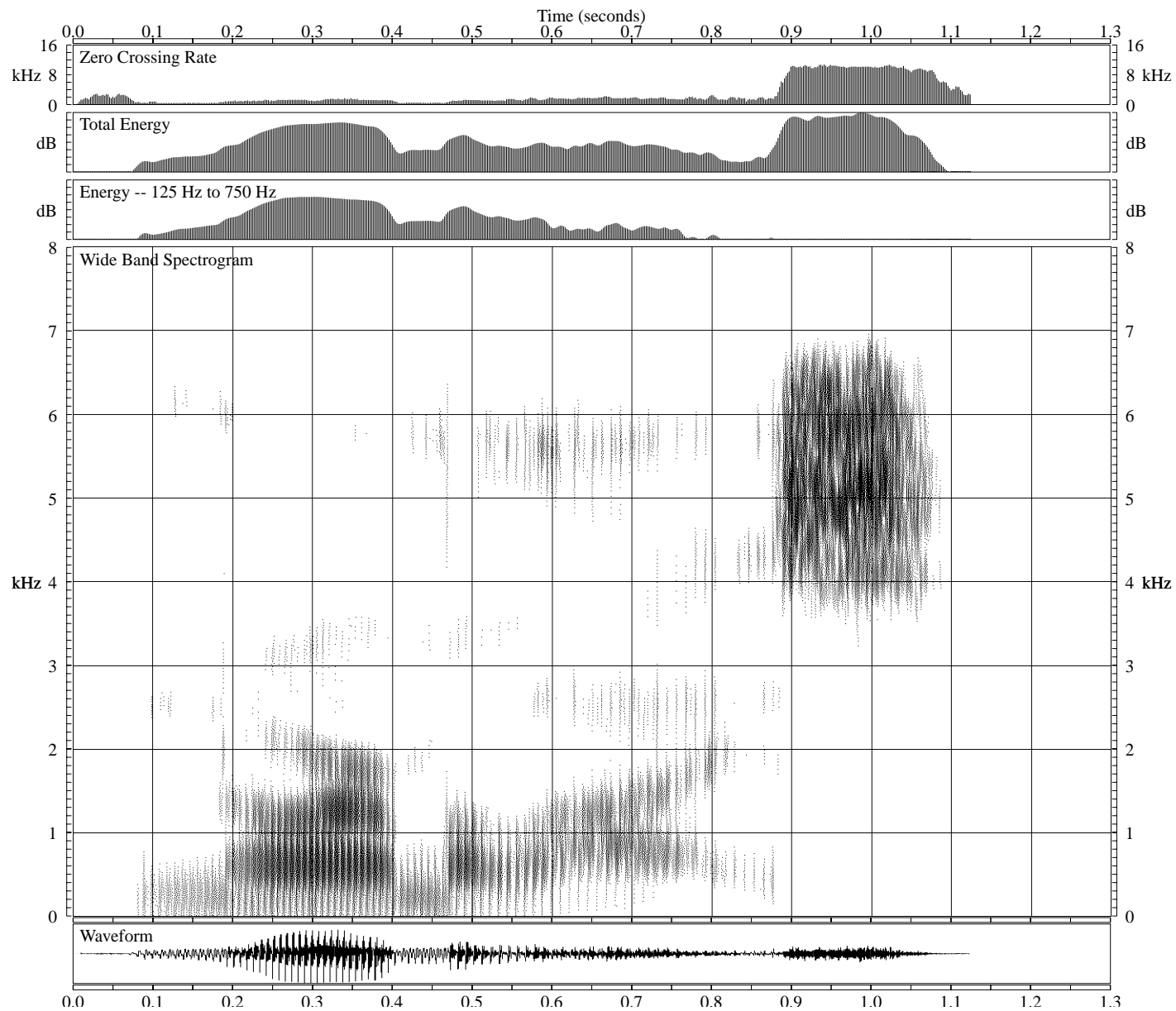


Acoustic Properties of Semivowels

- /w/ and /l/ are the most confusable semivowels
- /w/ is characterized by a very low F1, F2
 - Typically a rapid spectral falloff above F2
- /l/ is characterized by a low F1 and F2
 - Often presence of high frequency energy
 - Postvocalic /l/ characterized by minimal spectral discontinuity, gradual motion of formants
- /y/ is characterized by very low F1, very high F2
 - /y/ only occurs in a syllable onset position (i.e., pre-vocalic)
- /r/ is characterized by a very low F3
 - Prevocalic F3 < medial F3 < postvocalic F3

MIT

What is this word?



Affricate Production

- There are two affricates in American English:

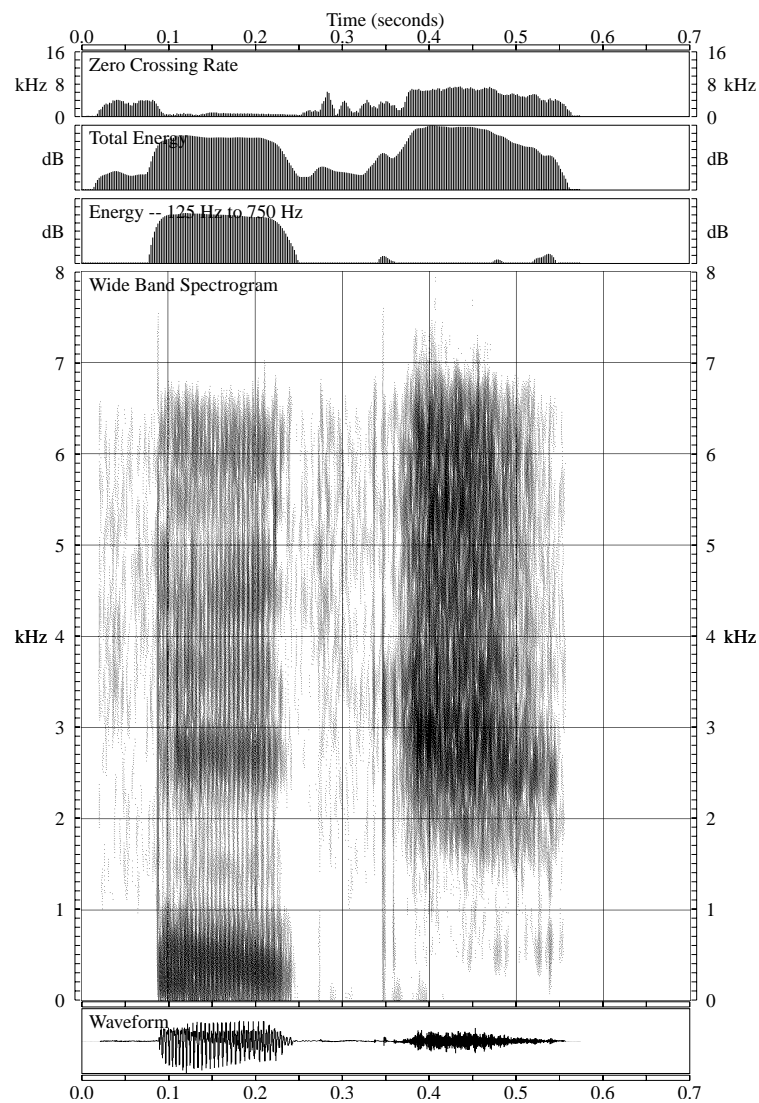
Voiced	Unvoiced
/j/ jh judge	/tʃ/ ch church

- Alveolar-stop palatal-fricative pairs
- Sudden release of the constriction, turbulence noise
- Can have periodic excitation during closure

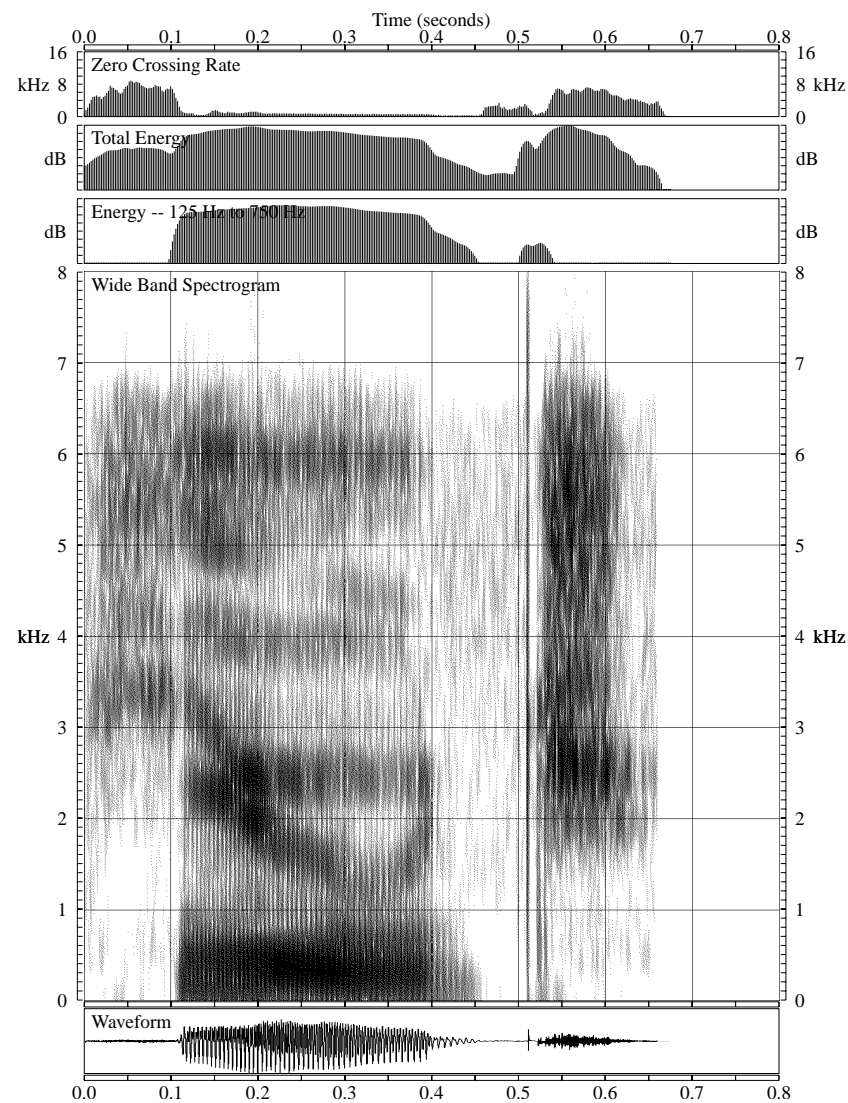
Aspirant Production

- There is one aspirant in American English: /h/ (e.g., “hat”)
- Produced by generating turbulence excitation at glottis
- No constriction in the vocal tract, normal formant excitation
- Sub-glottal coupling results in little energy in F1 region
- Periodic excitation can be present in medial position

Spectrograms of Affricates and Aspirant



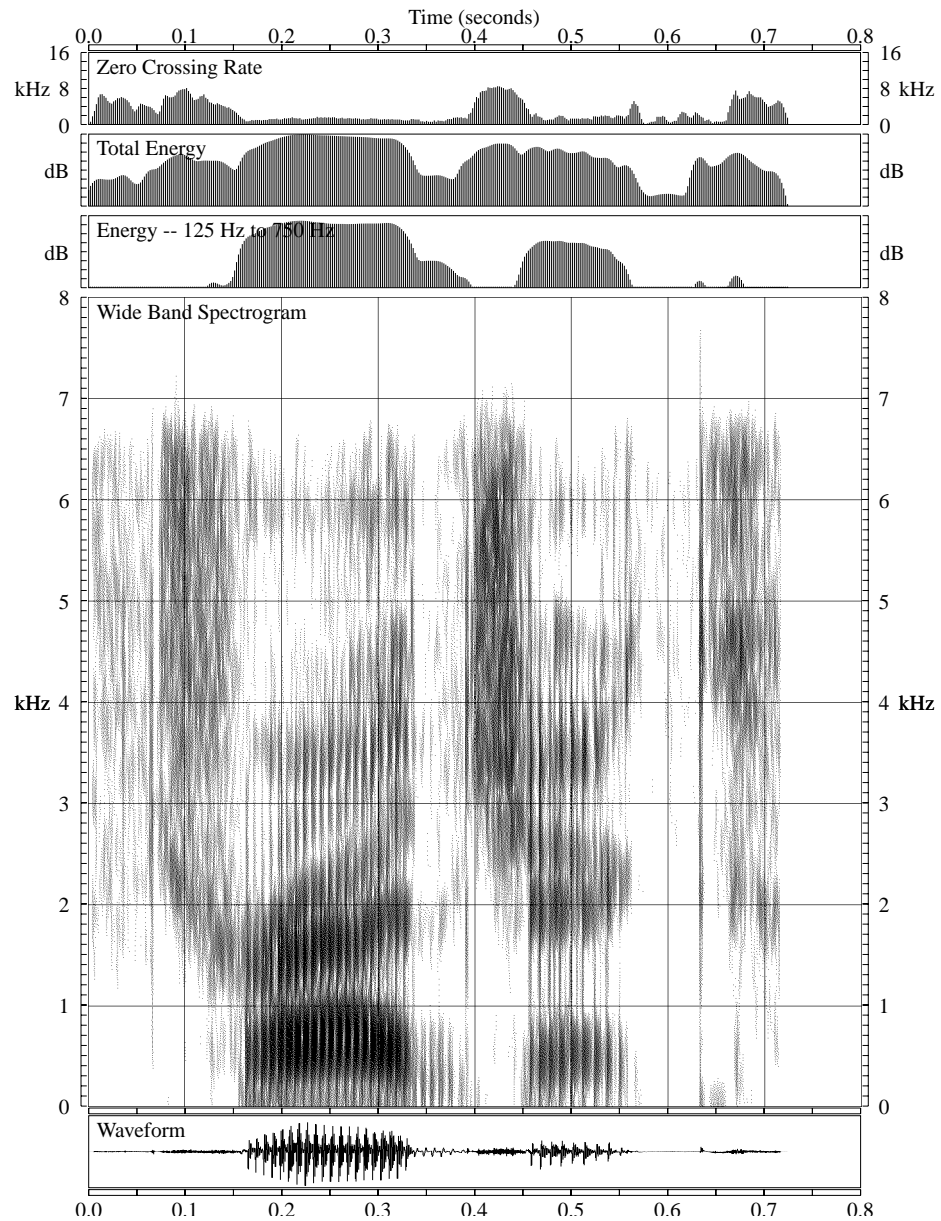
each
/iʃ/



huge
/hyʃ/

MIT

What is this word?



Phonotactic Constraints

- Phonotactics is the study of allowable sound sequences
- Analyses of word-initial and -final clusters reveal:
 - 73 distinct initial clusters (about 10 “foreign” clusters)
 - 208 distinct final clusters
- Can be used to eliminate impossible phoneme sequences:
 - /tk/ can't end a word, and
 - /kt/ can't begin a word,
 - Therefore, */... t k t .../ is an **impossible** sequence

Word-Initial Consonants from MWP Dictionary

-	of	hy	human	sf	sphere	tr	true
b	be	ĵ	just	sk	school	ts	tsunami
bl	black	k	can	skl	sclerosis	tw	twenty
br	bring	kl	class	skr	screen	ty	tuesday
by	beauty	kr	cross	skw	square	θ	thief
č	child	kw	quite	sky	skewer	θr	through
d	do	ky	curious	sl	slow	θw	thwart
dr	drive	l	like	sm	small	ð	the
dw	dwel	m	more	sn	snake	v	very
f	for	mw	moire	sp	special	vw	voyager
fl	floor	my	music	spl	split	vy	view
fr	from	n	not	spr	spring	w	was
fy	few	p	people	spy	spurious	y	you
g	good	pl	place	st	state	z	zero
gl	glass	pr	price	str	street	zl	zloty
gr	great	pw	pueblo	sw	sweet	zw	zweiback
gw	guava	py	pure	š	she	ž	genre
h	he	r	right	šr	shrewd		
hw	which	s	so	t	to		

MIT

The Syllable

- Syllable structure captures many useful generalizations
 - Phoneme realization often depends on syllabification
 - Many phonological rules depend on syllable structure
- Syllable structure is predicated on the notion of ranking the speech sounds in terms of their **sonority** values

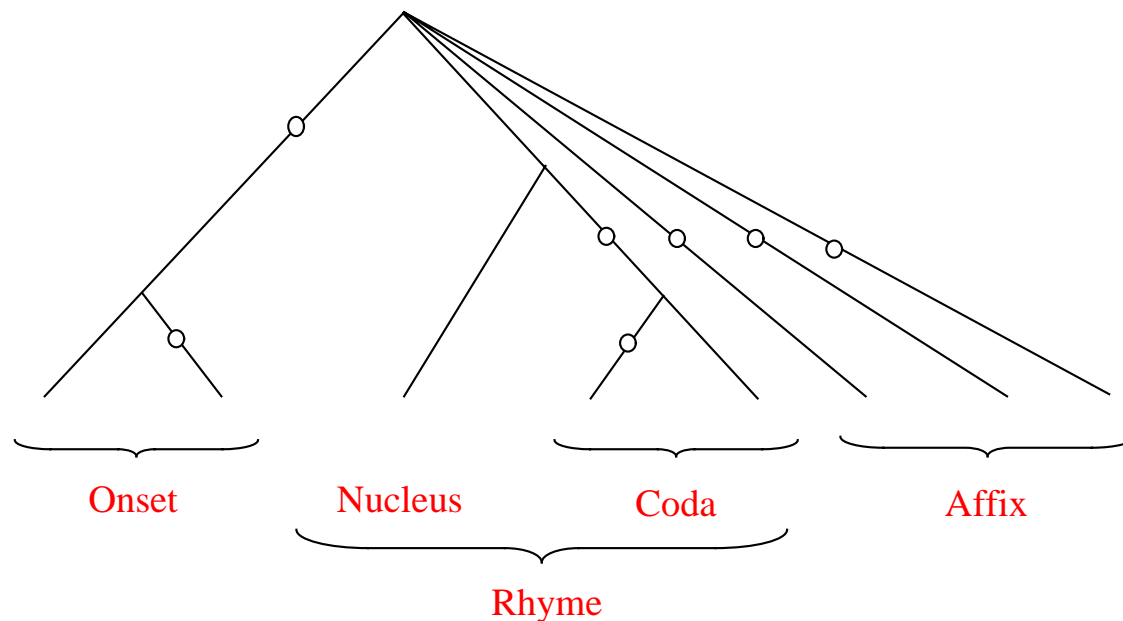
Sounds	Sonority Values	Examples
Low Vowels	10	/a, ɔ/
Mid Vowels	9	/e, o/
High Vowels	8	/i, u/
Flaps	7	/r/
Laterals	6	/l/
Nasals	5	/m, n, ŋ/
Voiced Fricatives	4	/v, ð, z/
Unvoiced Fricatives	3	/f, θ, s/
Voiced Stops	2	/b, d, g/
Unvoiced Stops	1	/p, t, k/

Syllables and Sonority

- Utterances can be divided into syllables
- The number of syllables equals the number of sonority peaks
- Within any syllable, there is a segment constituting a sonority peak that is preceded and/or followed by a sequence of segments with progressively decreasing sonority values

suprasegmental													
s	u	p	r	Λ	s	ε	g	m	ε	n	t	ə	l
3	8	1	7	9	3	9	2	5	9	5	1	9	6
minimization													
	m	ɪ	n	ɪ	m	ɑ ^y	z	e	ʃ	ə	n		
	5	8	5	8	5	10	4	9	3	9	5		
fire													
					f	ɑ ^y			ə				
					3	10	(8)		9				

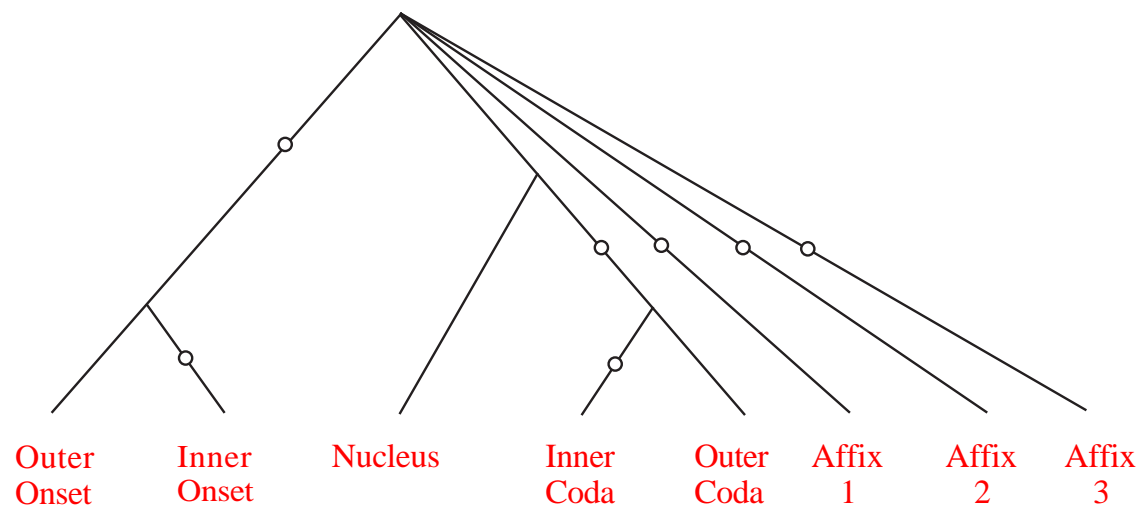
MIT The Syllable Template



- Branches marked by ◦ are optional
- Nucleus must contain a non-obstruent
- Sonority decreases away from nucleus
- Affix contains only coronals: /s, z, t, d, θ, ð, č, ĵ/
- Only the last syllable in a word can have an affix
- /sp/, /st/, and /sk/ are treated as single obstruents

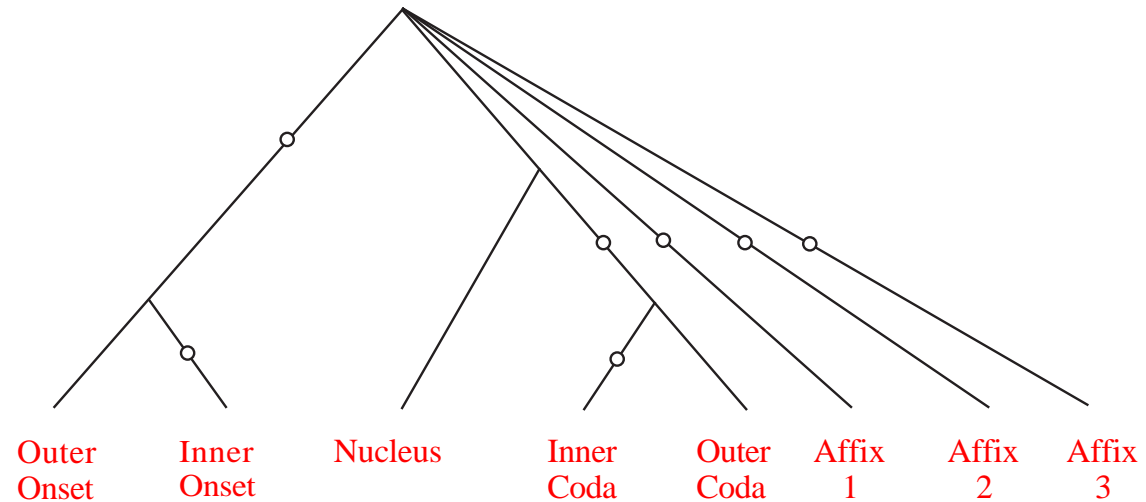
MIT

Some Examples



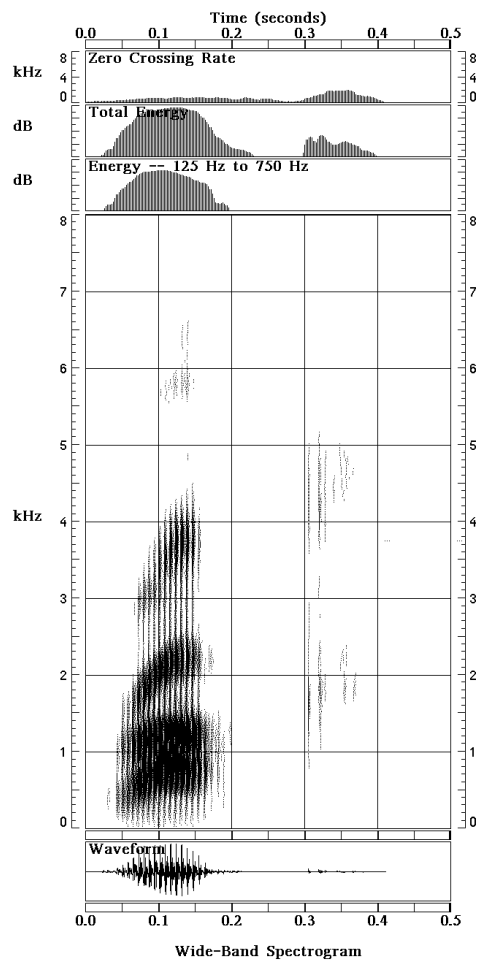
	Outer Onset	Inner Onset	Nucleus	Inner Coda	Outer Coda	Affix 1	Affix 2	Affix 3
crown	k	r	a	w	n			
fledged	f	l	ε		ʃ	d		
links	l		ɪ	ŋ	k	s		
dwarves	d	w	a	r	v	z		
stick	st		ɪ		k			
sixths	s		ɪ		k	s	θ	s

Words Containing /r/ and /l/

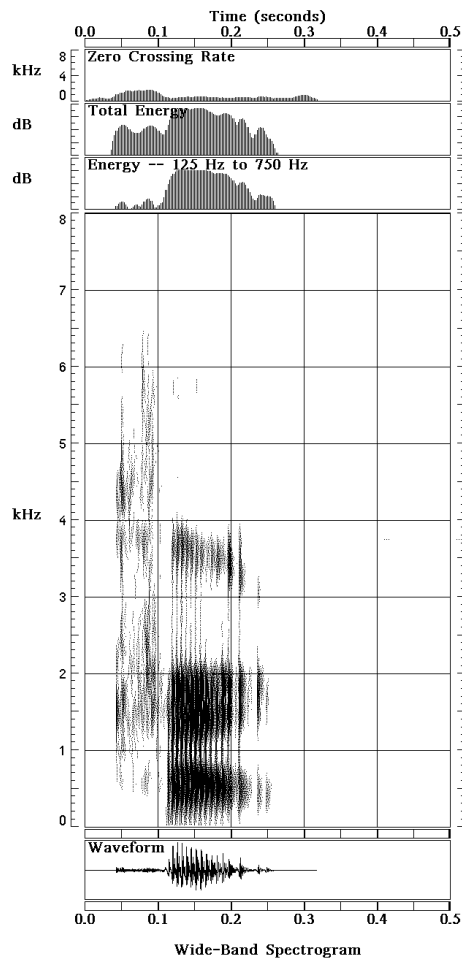


	Outer Onset	Inner Onset	Nucleus	Inner Coda	Outer Coda	Affix 1	Affix 2	Affix 3
rock	r		a		k			
crock	k	r	a		k			
curt	k		ɜ̃		t			
cart	k		a	r	t			
car	k		a		r			
lick	l		ɪ		k			
bottle	b		a, l		t			
kill	k		ɪ		l			

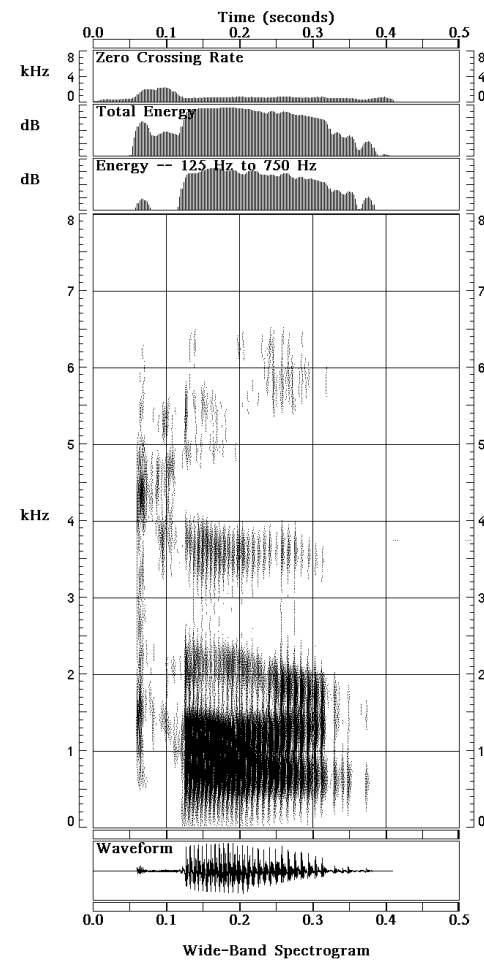
Acoustic Realizations of /r/



rock
/rak/

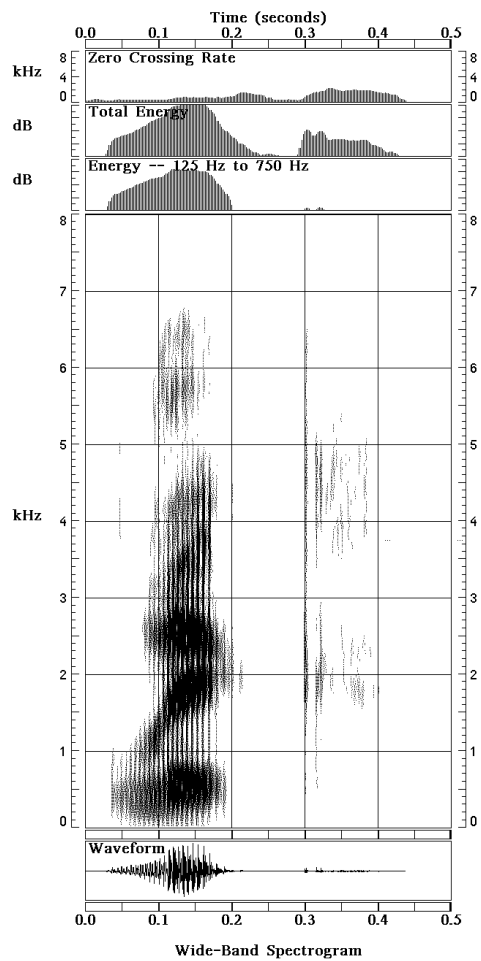


curt
/kɜrt/

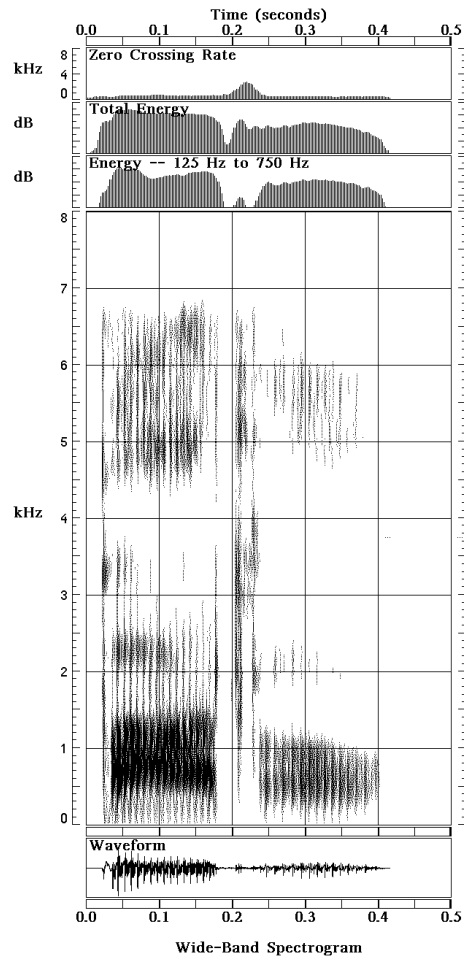


car
/kar/

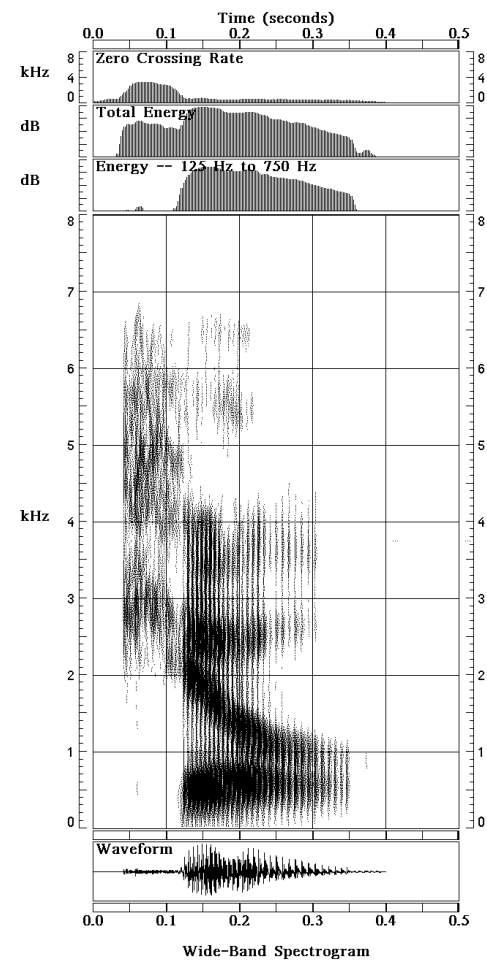
MIT Acoustic Realizations of /l/



lick
/lɪk/

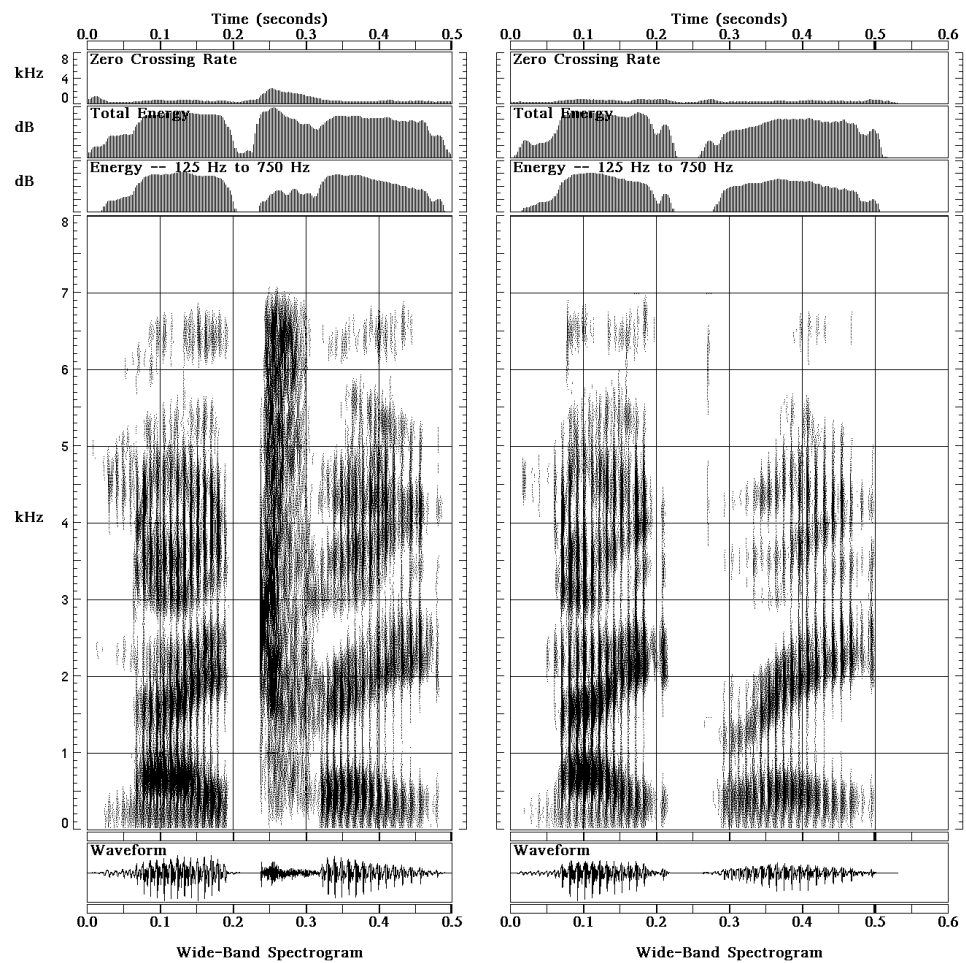


bottle
/bɑtəl/



kill
/kɪl/

Allophonic Variations at Syllable Boundaries



nitrate
/nɑ^y tre^yt/

night rate
/nɑ^yt re^yt/

MIT Assignment 2

Crystal Query

Search String: Search Type: phonemic Ignore syllable boundaries

Lexicon History

Lexicon 1 : 19837 entries

Search Results

Spelling	Pronunciation	Stress Pattern	BCF	Rank	<input type="checkbox"/> Sort by BCF
AWOL	'e wɒl	SU	0	0	0.0%
Advent	'æd vɛnt	SU	5	5891	0.0%
African	'æ frɪ kɪn	SUR	28	2257	0.0%
Afrikaans	æ frɪ k'ɑns	UUS	0	0	0.0%
Albanian	æɫ b'e ni ən	USUR	2	8801	0.0%
Alexandrian	æ liɔ z'æn dri ən	UUSUR	0	0	0.0%
Algerian	æɫ j'ɪ ri ən	USUR	4	6583	0.0%
Allah	'æ lə	SR	0	0	0.0%
Allhallows	ɔɫ h'æ lɔz	USU	0	0	0.0%
Amazonian	æ mə z'o ni ən	URSUR	0	0	0.0%
American	ə m'e rɪ kɪn	RSRR	596	151	0.1%
Americanization	ə m'ɛ rə kɪ nɪ z'e Ńɪn	RURRRSR	0	0	0.0%

Statistics

Weight by BCF Phoneme Occurrences Phonemes per Word Syllables per Word Stress Pattern Occurrences