

English Phonetics

English Obstruents

- we can divide English consonants into two subgroups, *obstruents* and *sonorants*
- obstruents* are consonants that are formed by obstructing the airflow, causing increased air pressure in the vocal tract

Plosives	vl.	p	t	k	(ʔ)
	vd.	b	d	g	

- sonorants* involve no turbulent airflow in the vocal tract
- the English obstruent consonants include the plosives, fricatives and affricates

Fricatives	vl.	f	θ	s	ʃ	h
	vd.	v	ð	z	ʒ	

- obstruents involve either a total closure of the vocal tract or a partial closure, i.e., a stricture causing friction

Affricates	vl.	tʃ
	vd.	dʒ

- obstruents may be subdivided into:
 - plosives*, which involve a complete closure followed by an "explosive" release of air
 - fricatives*, with only limited closure, i.e., a sufficient degree of closure to cause friction
 - affricates*, with a total closure followed by a fricative release

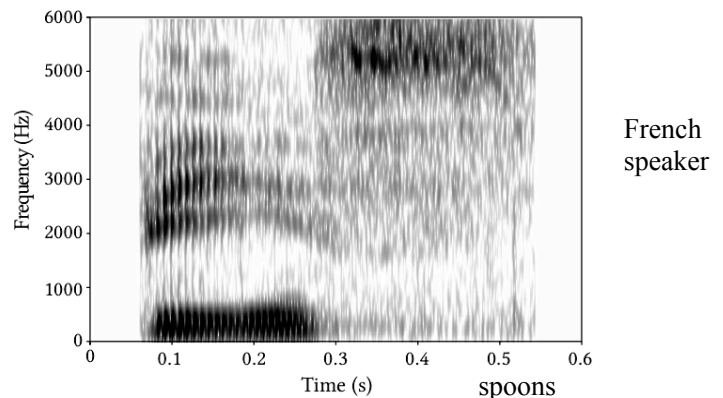
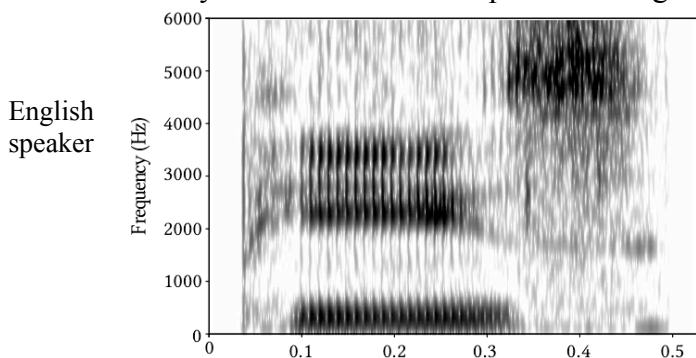
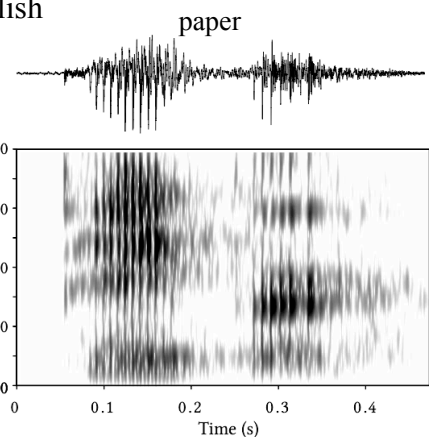
- obstruents are typically voiceless, though voiced obstruents are common

English Plosives

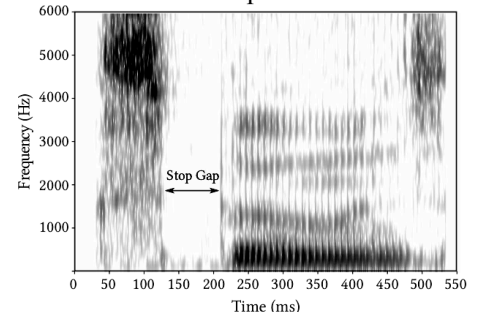
- English plosives consist of voiced and voiceless pairs of consonants
- they occur at the labial, alveolar, and velar points of articulation
- there is also a glottal stop, that occurs in various positions in different dialects

Labial Plosives

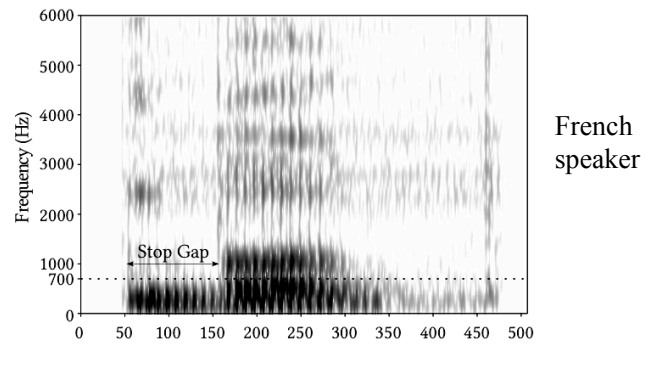
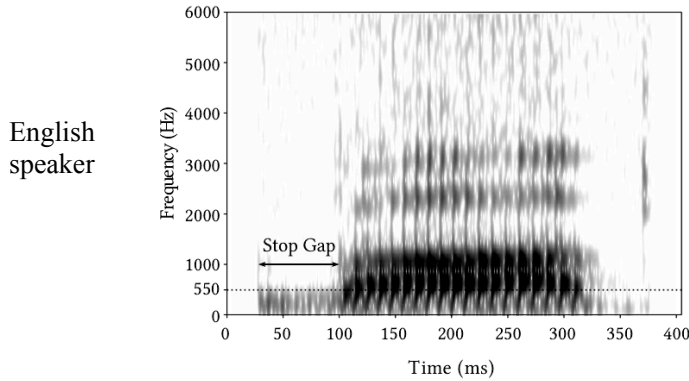
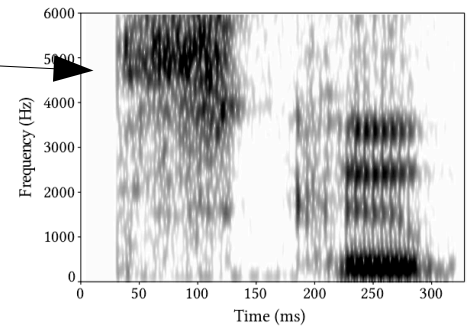
- the labial plosives, /p/ and /b/, may have several different realizations in English
- this depends on the position where they occur
- when voiceless /p/ is initial in a stressed syllable, as in *paper* or *popular*, it is fairly strongly aspirated, symbolized [p^h]
- the aspiration is clear in the following spectrogram at 0.5-0.6s
- it appears as high-intensity noise in the range of 3,000-5,000Hz
- the presence of aspiration in such forms is a crucial aspect of English
- the word 'peas', as pronounced by a native speaker of English, clearly shows the aspiration from 0.05-0.1s
- it is absent in the pronunciation of the [p] (~0.06s) in the word 'peas' spoken by the French speaker
- this clearly marks a non-native speaker of English



- the English voiceless bilabial plosive is not always aspirated: after [s], it shows no aspiration
- the stop gap from 140-210ms, indicates the [p]
- there is almost no turbulence following this
- this is typical of an English voiceless plosive after [s]
- there is a third variety that appears at the end of the syllable, referred to as unreleased [p̚]



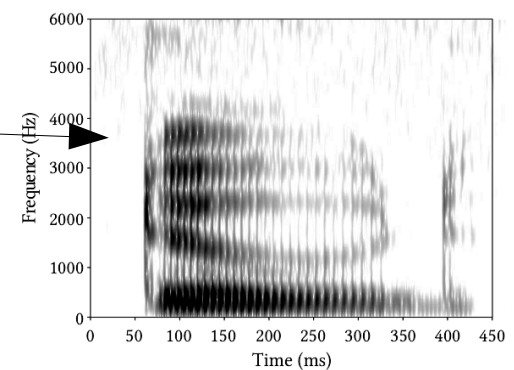
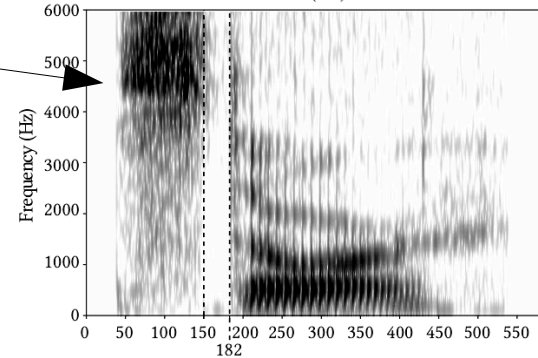
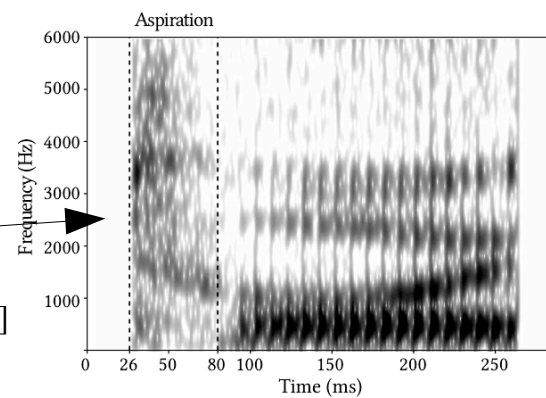
- examine the word *scoop*: from 300ms, there is nothing; the plosive is unreleased; all we can see is the labial effect on the vowel
- English also has a voiced bilabial plosive, [b]
- voicing is not strong in English, as shown in the spectrogram of *Bob*
- compare the voicing with that of a French speaker
- the stop gap is much larger in the French pronunciation (English 75ms, French 120ms)
- also the voicing bar is higher (English 550Hz, French 700Hz)



- for English, aspiration is a more significant marker of the distinction between voiced and voiceless obstruents

Alveolar Plosives

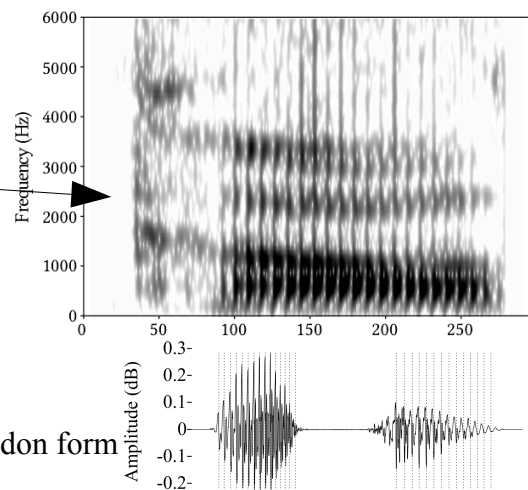
- like the labial plosives, English has a contrastive voicing distinction among the alveolar plosives, /t/ vs /d/
- voiceless /t/ in initial position is aspirated, [t^h]
- the same comments regarding aspiration of the [p^h] hold for [t^h]
- similarly, /t/ following [s] is unaspirated, thus [t], as shown in the word *store*
- notice the absence of aspiration on the [t] in *store*, just as with the [p] in *spoons*
- the stop gap is from 150ms-182ms, followed by a very brief burst
- if you compare this with the aspiration in *toy*, it is easy to see the difference
- as with the voiceless bilabial, there is also an unreleased version of /t/
- found at the end of the syllable
- it is represented as [t̚]
- there are also other variants that appear in certain dialects
- these include the glottal stop, [ʔ], commonly found in British English varieties, and the alveolar flap, [ɾ], found in North American dialects
- before leaving the alveolar plosives, we should examine the voiced alveolar plosive, /d/
- /d/ is not strongly voiced, much less voiced than French /d/
- an example of English [d] is the word *dude*
- the voicing, and the stop overall, is very brief
- compare the initial [d] in *dude* with the final [d]
- the final [d] is about 75ms long and its voicing bar reaches approximately 500Hz, much like [b]
- the voiced alveolar plosive may also appear as a flap [ɾ] in intervocalic position in NA English



Velar Plosives

- the velar plosives, voiceless /k/ and voiced /g/, appear in the same varieties as the bilabial plosives
- in word- or syllable-initial stressed position, the voiceless velar plosive is aspirated, [k^h]

- when preceded by [s], it has the same properties as the other plosives discussed so far, as shown by the spectrogram of *scoop*
- note the brief aspiration from 190-215ms
- this much less than that for *call*, where aspiration ranges from approximately 35-90ms



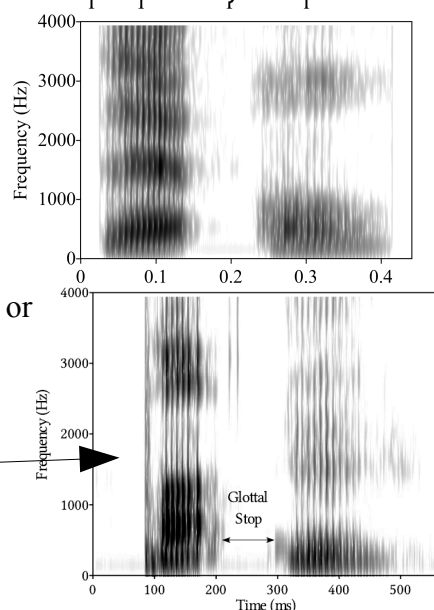
Glottal Stop

- one final stop acts as an allophonic variant of /t/ in some contexts
- the glottal stop is invariably voiceless
- when producing [ʔ], the vocal cords are held tightly together, preventing vibration
- the glottal stop is realized as a gap in the flow of sound, as in the London form for *little*

Preglottalization

- [ʔ] may occur as preglottalization on consonants at the end of the syllable
- the final use of glottal stop appears in North American English, and is similar to the case of preglottalization
- it applies to forms having a syllabic alveolar nasal preceded by a voiceless alveolar plosive
- words like *button*, *cotton*, and *kitten*, but not in *sudden*, *happen*, *bottom*, *little*, or *butter*
- it only happens with syllabic alveolar nasals preceded by voiceless alveolar plosives
- a spectrogram of the word *button*, [bʌʔn] shows this:

- a. quite good [kwaiʔt gʊ:d]
- b. look down [lʊʔk daʊn]
- c. happen [hæʔpɪn]

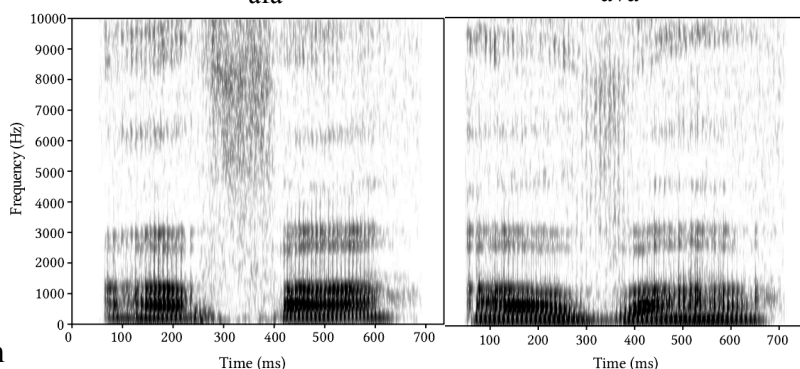


English Fricatives

- English fricatives include [f, θ, s, ʃ, h, v, ð, z, ʒ]
- this is a large set of fricatives typologically: few languages have so many contrastive fricatives
- Korean has only two, or three, depending on the dialect: [s, s', h]

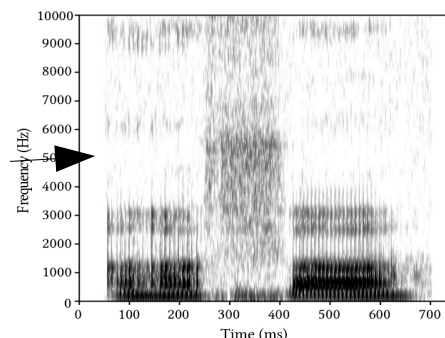
Labiodental Fricatives

- there are two contrastive labiodental fricatives in English, the voiceless [f] and the voiced [v]
- their typical properties include high frequency turbulence concentrated above 4,000 Hz
- words such as *fox*, *file*, and *frame* begin with [f], while *tough*, *half*, and *stuff* end with [f]
- there is no voicing bar with [f]
- the voiced labiodental fricative also shows high frequency turbulence focused above 4,000 Hz
- there is a substantial voicing bar occupying approximately the lower 400 Hz
- words beginning with [v] include *vie*, *valve*, *view*, while words ending with [v] include *halve*, *live*, *cove*



Interdental Fricatives

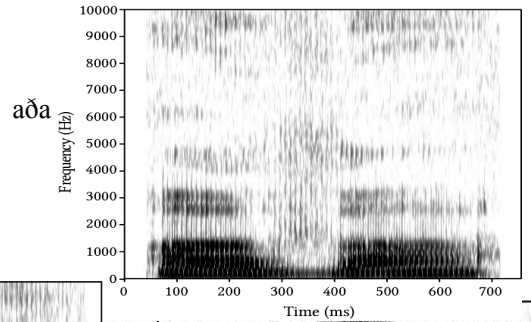
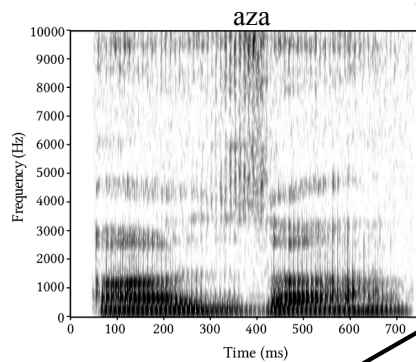
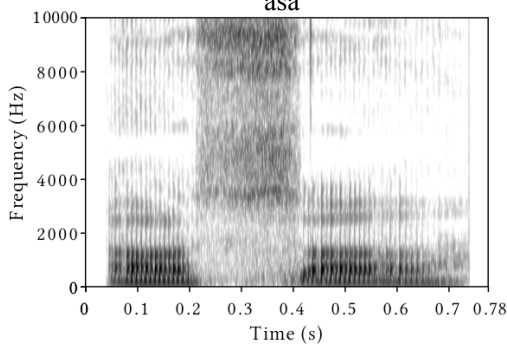
- Interdental fricatives are not common in the languages of the world
- English has both voiceless and voiced variants
- [θ] is found as the first sound in words such as *think*, *thigh* and *thought*
- as the last sound of words such as *both*, *path* and *with*
- energy begins low (2500hz)
- the voiced counterpart, [ð], occurs as the first sound in words such as *though*, *that* and *they*



- as the last sound in words such as *bathe*, *betroth* and *soothe*

Alveolar Fricatives

- English also has alveolar fricatives, [s] and [z]
- the most common crosslinguistically
- the bulk of the turbulence occurs above 3500Hz
- With [z] there is a voicing bar

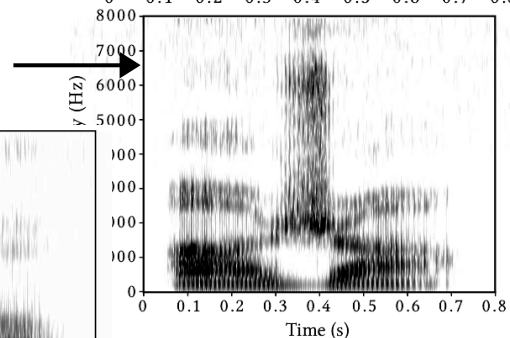
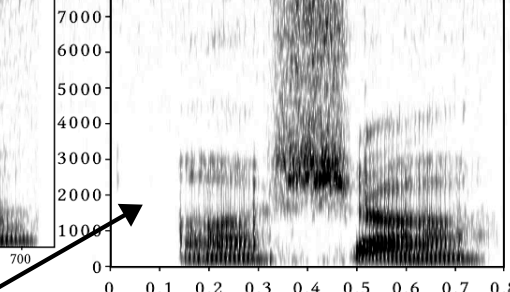
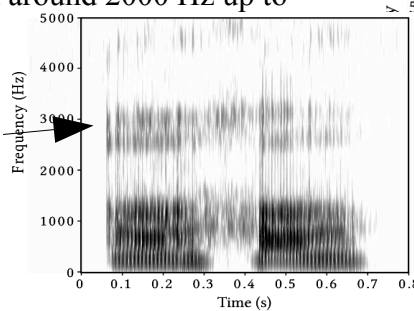


Alveo-Palatal Fricatives

- English has both voiceless and voiced alveopalatal fricatives, [ʃ] and [ʒ]
- the range of turbulence for both of these is from around 2000 Hz up to 10,000 Hz

Glottal Fricative

- the final English fricative is voiceless glottal [h]
- there is no voicing bar for [h]
- its turbulence appears to be strongest around 1,000 Hz



Sibilants

- fricatives can be divided into sibilants versus non-sibilants
- this distinction appears in the rules for forming the plural and other rules involving a suffix with /-s/ or /-z/
- sibilants involve a turbulent airstream that strikes an obstacle, such as the teeth
- non-sibilants involve turbulence at the site of constriction
- sibilants tend to be louder than non-sibilants
- most of their acoustic energy occurs at higher frequencies
- for instance, [s] has acoustic energy starting at around 3,500 Hz, and reaching as high as 10,000 Hz
- [ʃ] has most of its acoustic energy at around 4,000 Hz, extending up to around 8,000 Hz
- the English sibilants include [s,ʃ, z, ʒ]

Fricative Frequency Range Intensity Sibilant Voicing Bar

Fricative	Frequency Range	Intensity	Sibilant	Voicing Bar
f	5000hz and above	low	no	none
v	5000hz and above	low	no	yes
θ	2,500hz and above	low	no	none
ð	2,500hz and above	low	no	yes
s	3,500 and above	high	yes	none
z	3,500 and above	high	yes	yes
ʃ	2,000 and above	high	yes	none
ʒ	2,000 and above	high	yes	yes
h	750 hz-3000	low	no	none

English Affricates

- the English affricates include both voiceless, [tʃ], and voiced, [dʒ], alveopalatal affricates
- *lecher* illustrates the voiceless alveopalatal affricate, [tʃ] and *ledger* shows the voiced variant, [dʒ]

