



Lecture 6: Basics of Phonetics

Zhizheng Wu

Outline

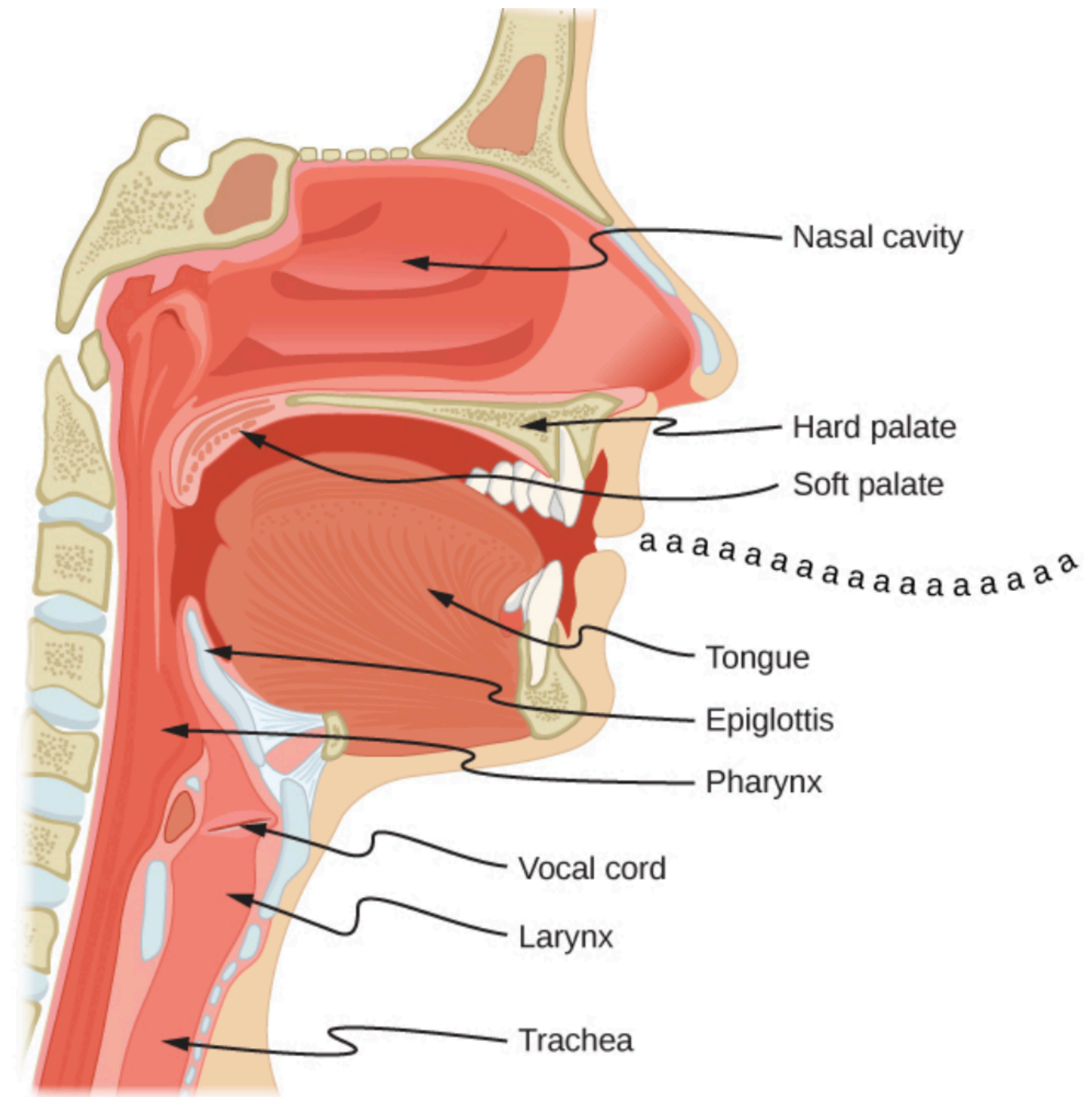
- ▶ Recap of speech production
- ▶ Phone, International Phonetic Alphabet, and Grapheme-to-phoneme conversion
- ▶ Articulatory phonetics

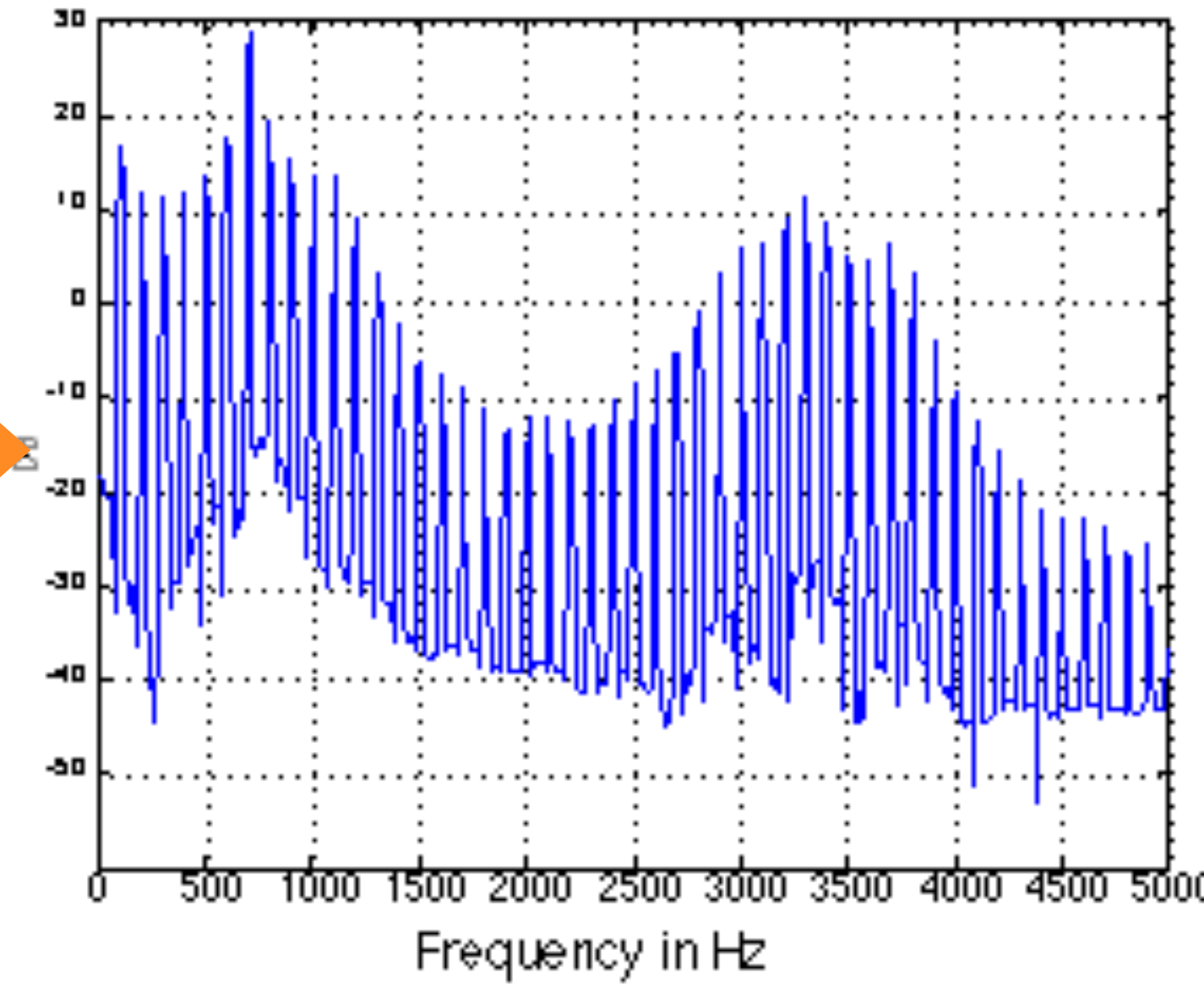
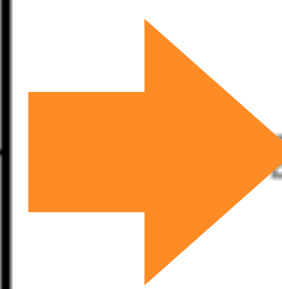
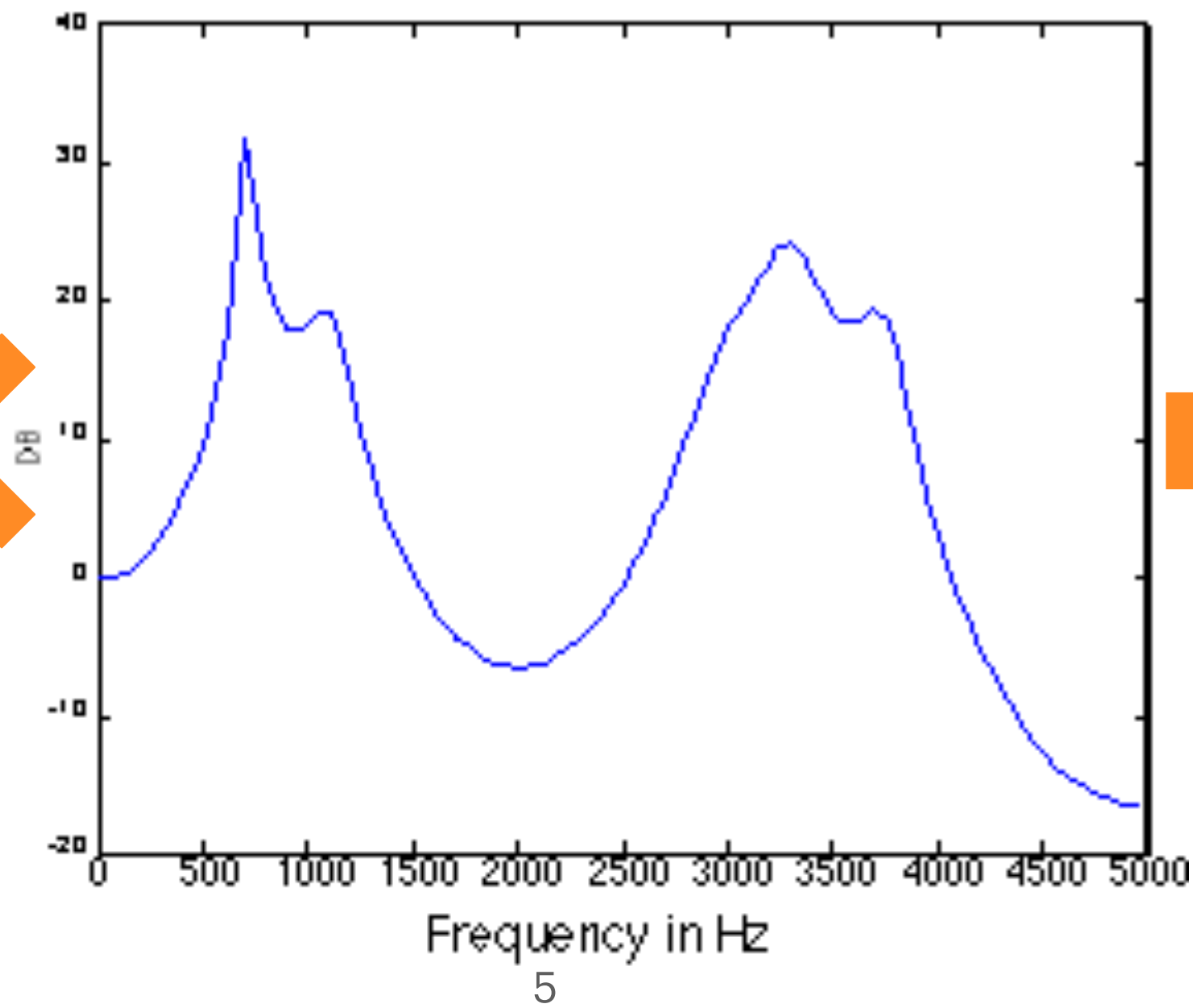
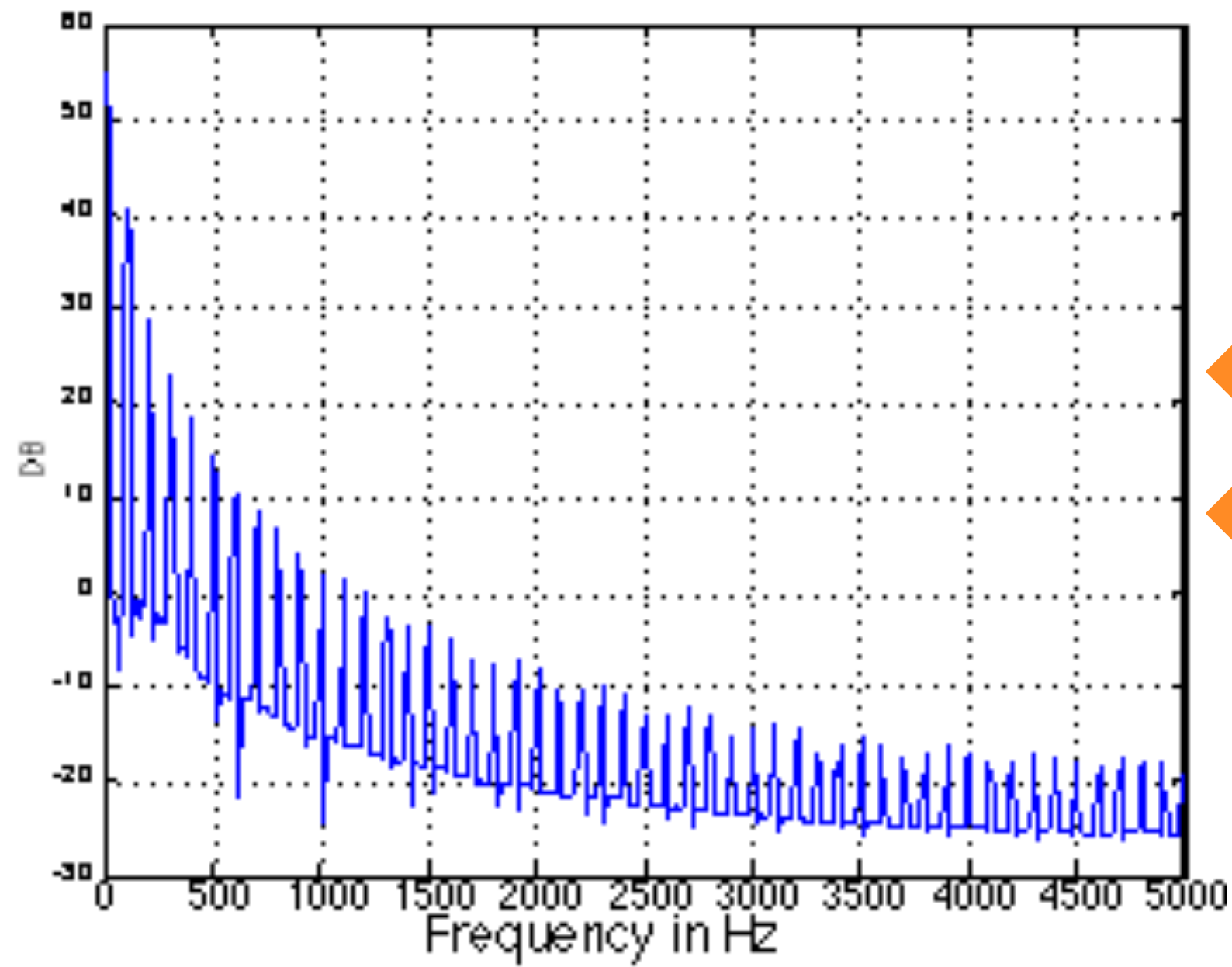
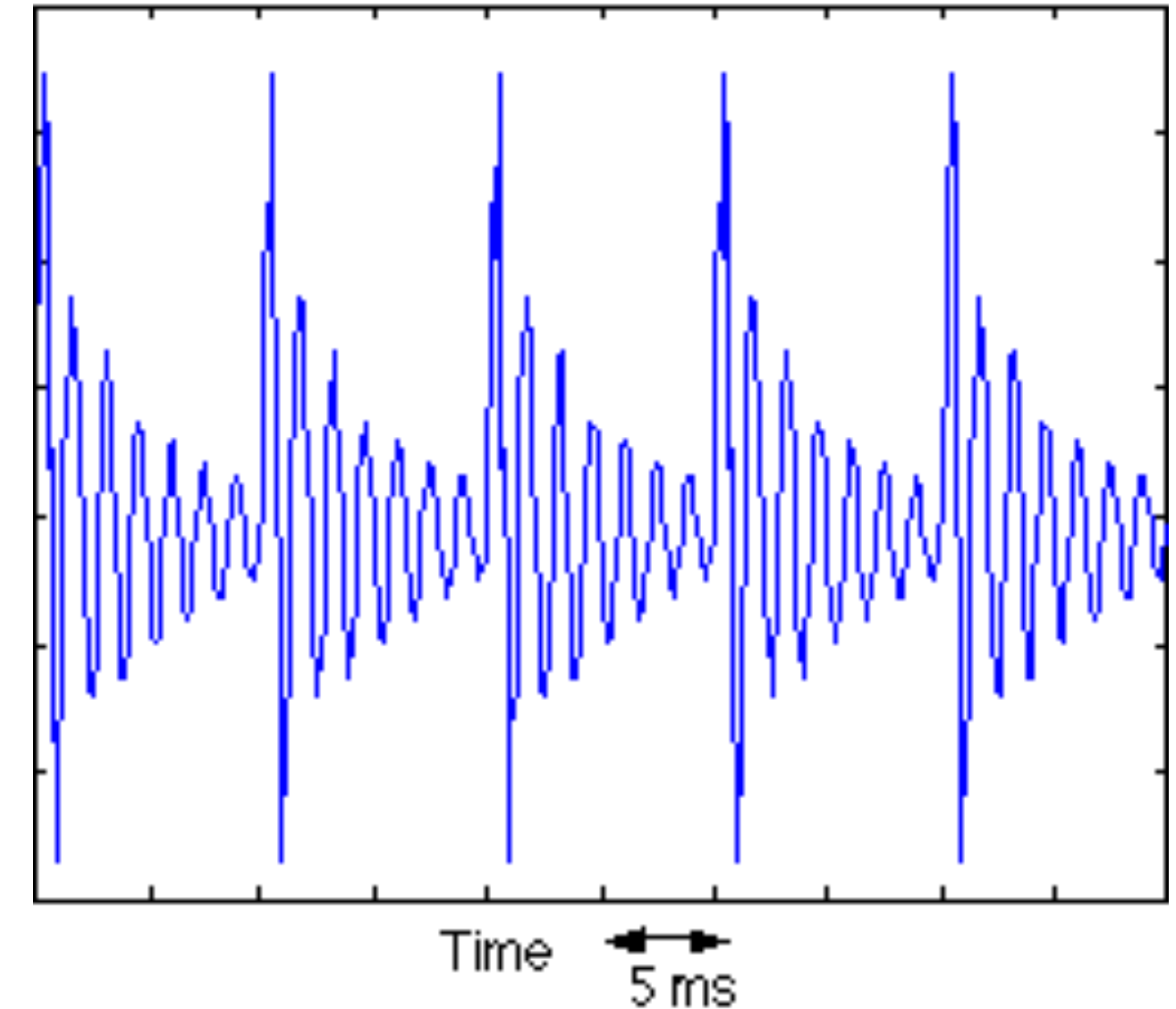
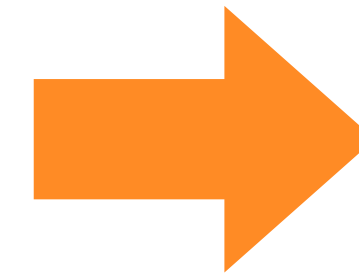
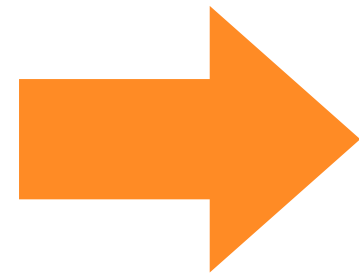
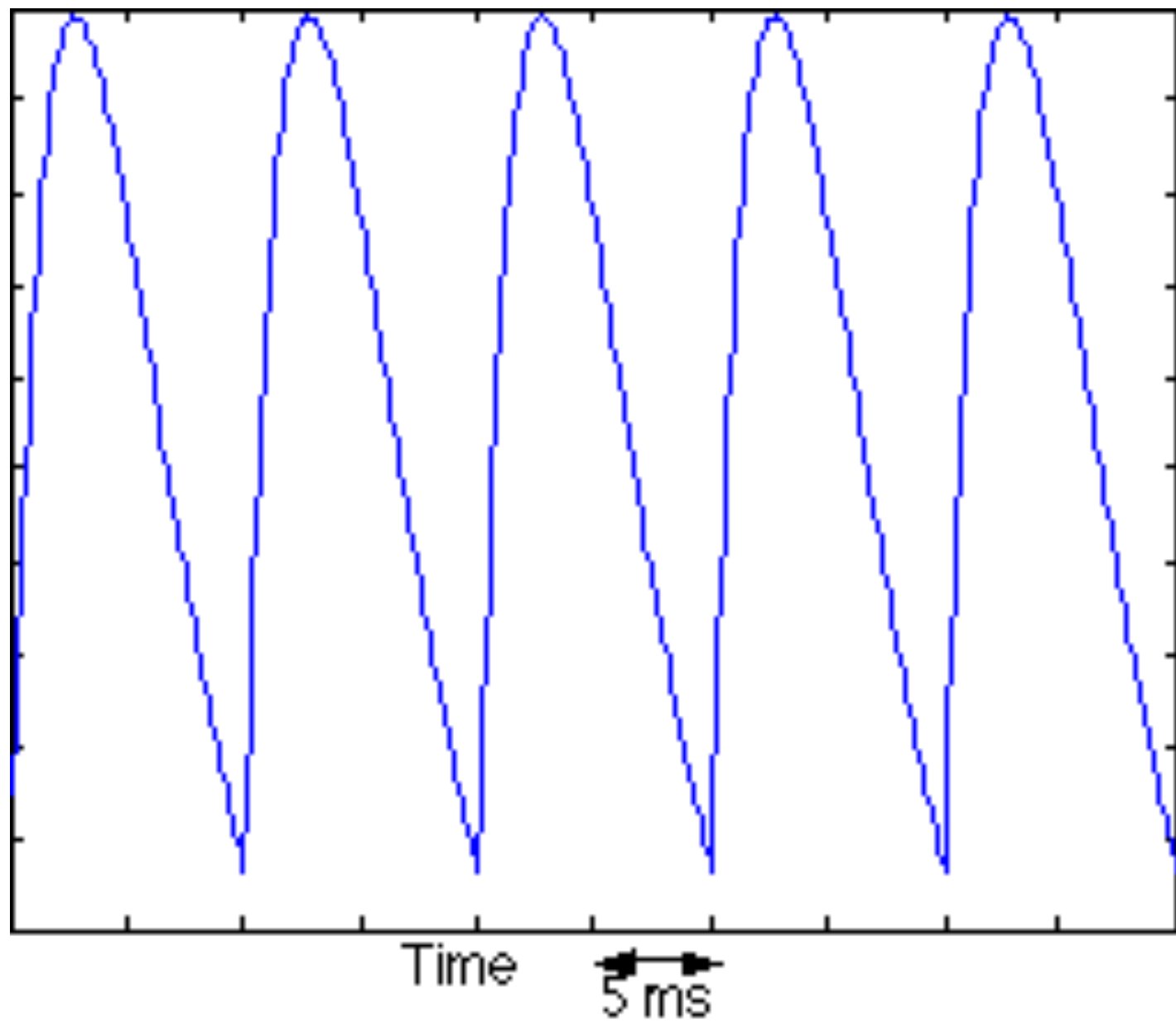
Recap

- ▶ Speech representation in time and frequency domain
- ▶ Speech production and the source-filter model

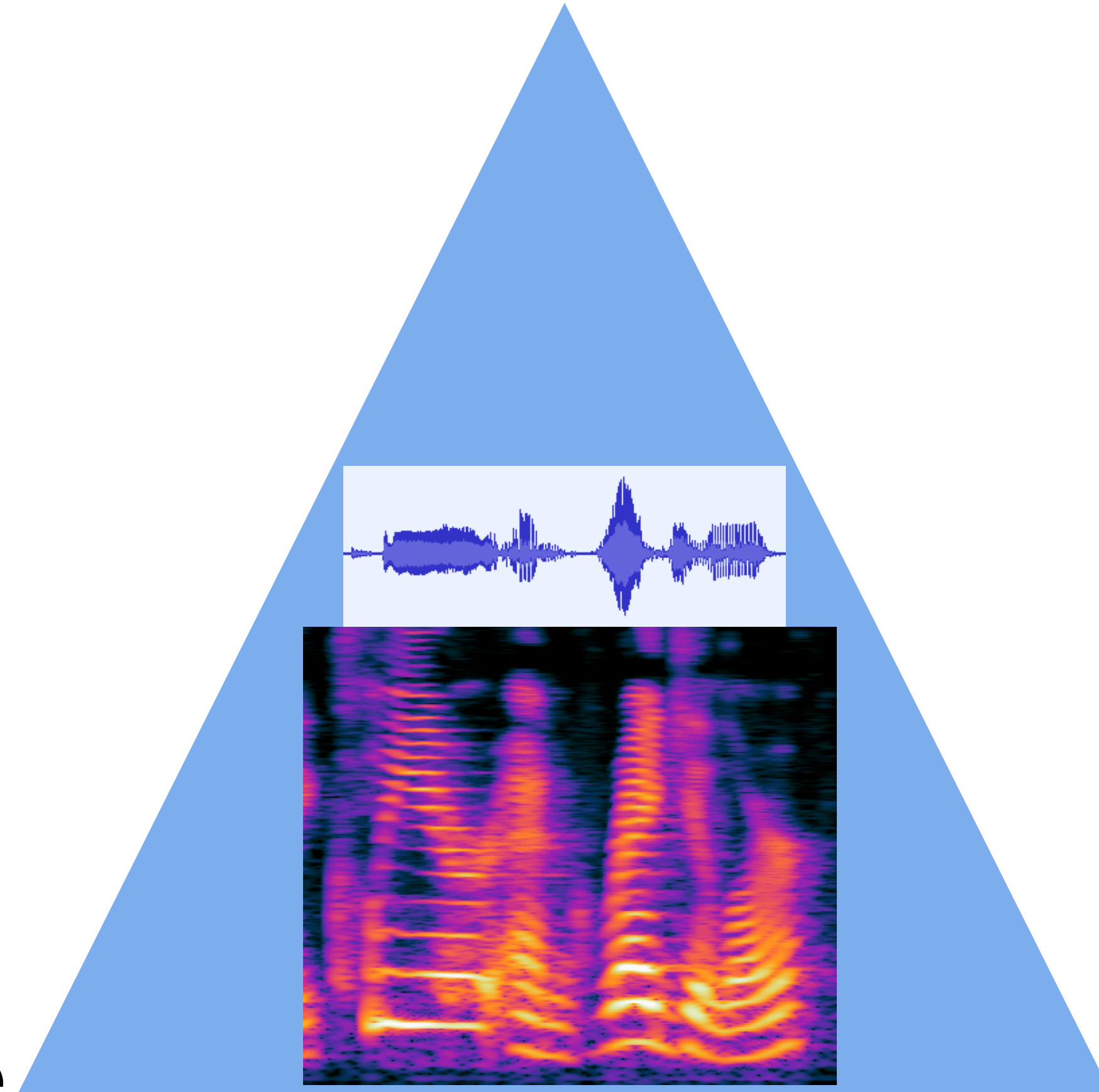
Speech production

- ▶ Source-filter model
 - Source produces an initial sound
 - Vocal tract filter modifies it
- ▶ Source
 - An input of acoustic energy into the speech production system
- ▶ Vocal tract filter
 - Articulators: tongue, teeth, lips, velum etc





Content

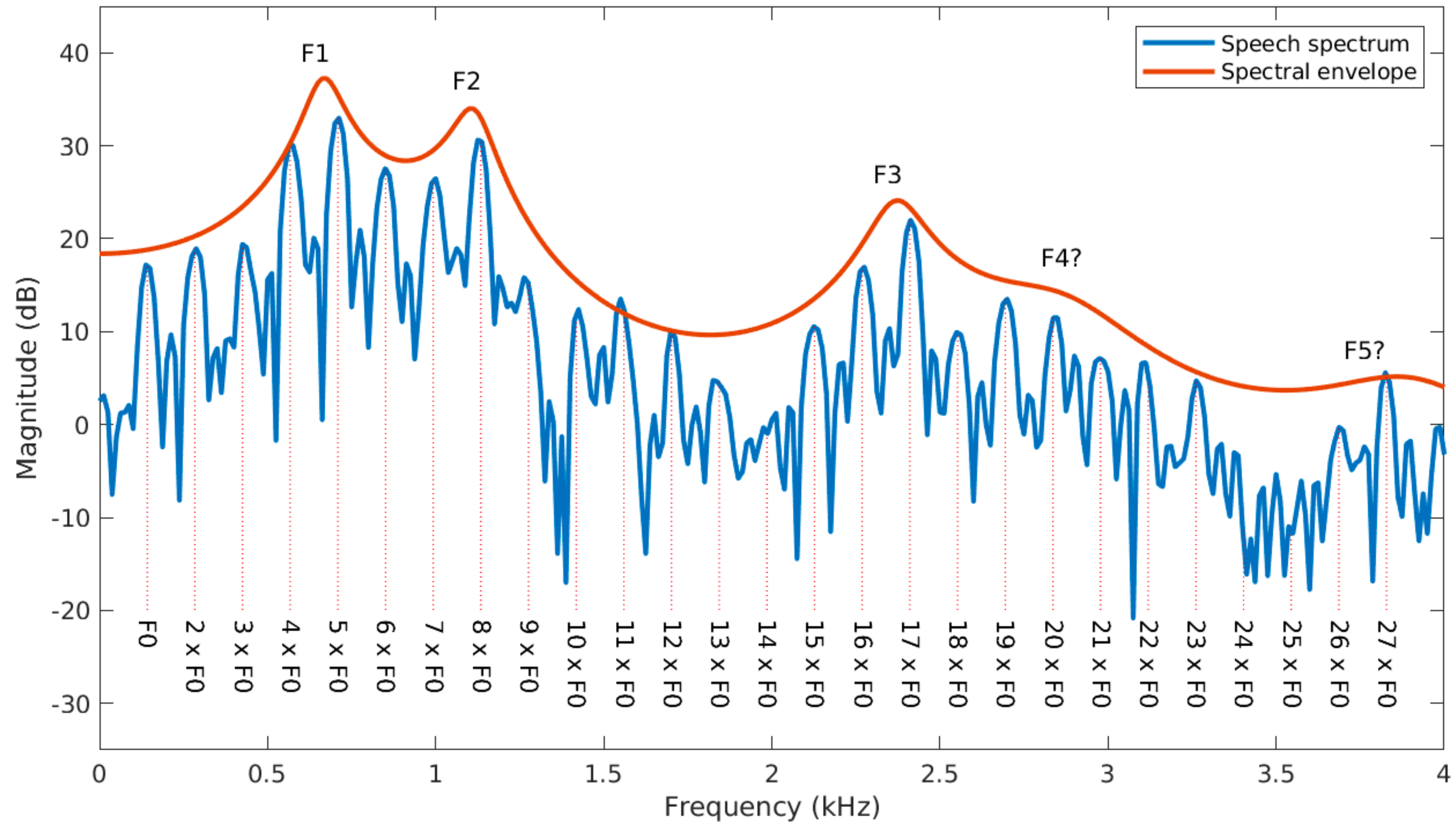


Timbre

Prosody

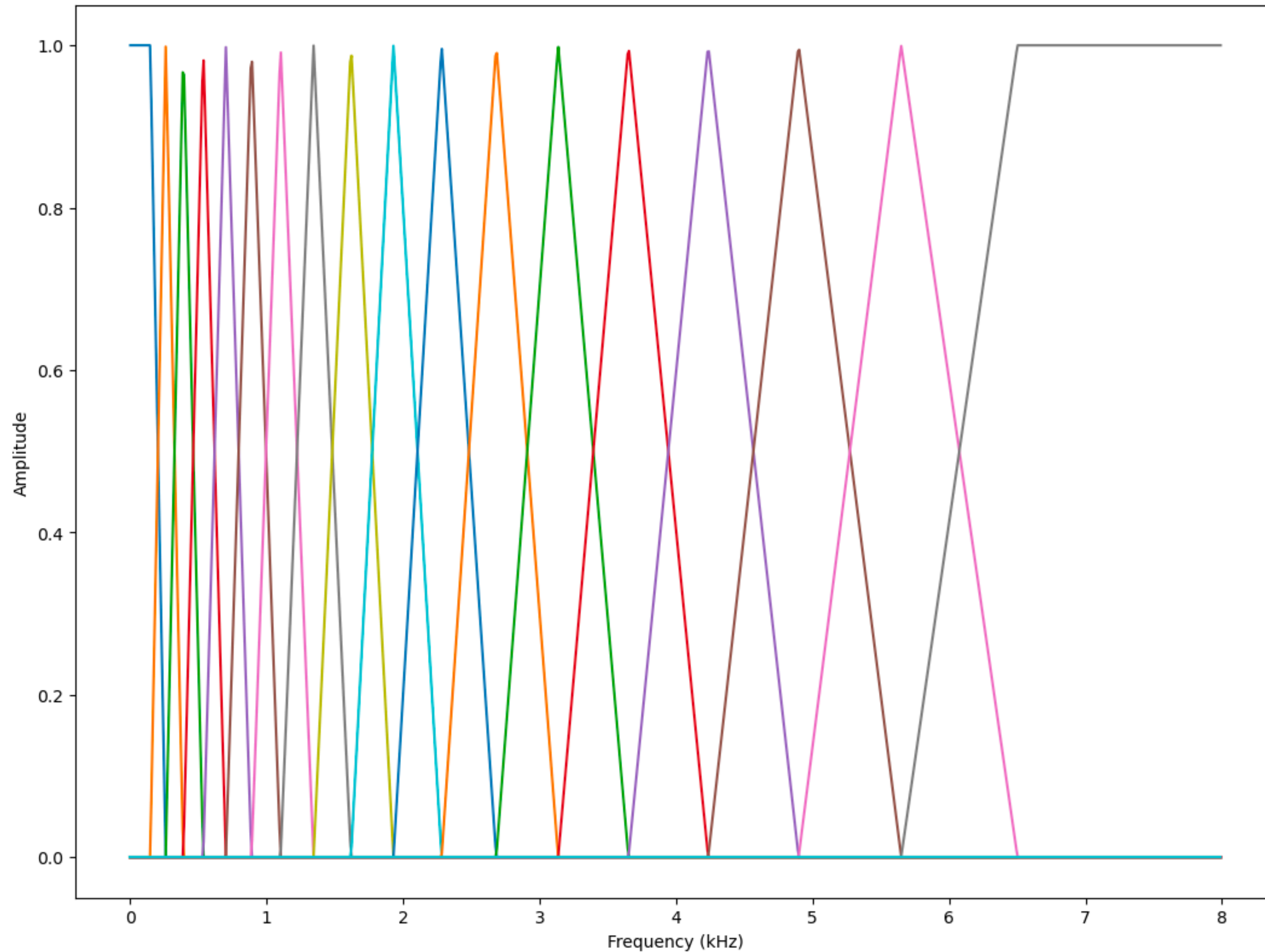
Fundamental frequency

- ▶ F_0 and harmonics kF_0



Mel filterbank

- ▶ Filterbank
 - triangle-centres are at the frequencies corresponding to equal distance steps on the mel scale
- ▶ Higher frequencies, above 6.5 kHz in particular, are poorly modelled



Here are the words for "mom" in several different languages:

- English: Mom
- Spanish: Mamá
- French: Maman
- German: Mama
- Italian: Mamma
- Portuguese: Mãe
- Dutch: Moeder
- Russian: Mama (Mama)
- Chinese: 妈妈 (Māma)
- Japanese: 母 (Haha)
- Korean: 엄마 (Eomma)
- Arabic: أم (Umm)
- Hindi: माँ (Maan)
- Bengali: মা (Ma)

Phone

- ▶ The pronunciation of a word can be represented as a sequence of phones
- ▶ The standard phonetic representation for transcribing the world's languages is the International Phonetic Alphabet (IPA)



tomato

[t^hə'meɪrəʊ]

西红柿

xī hóng shì

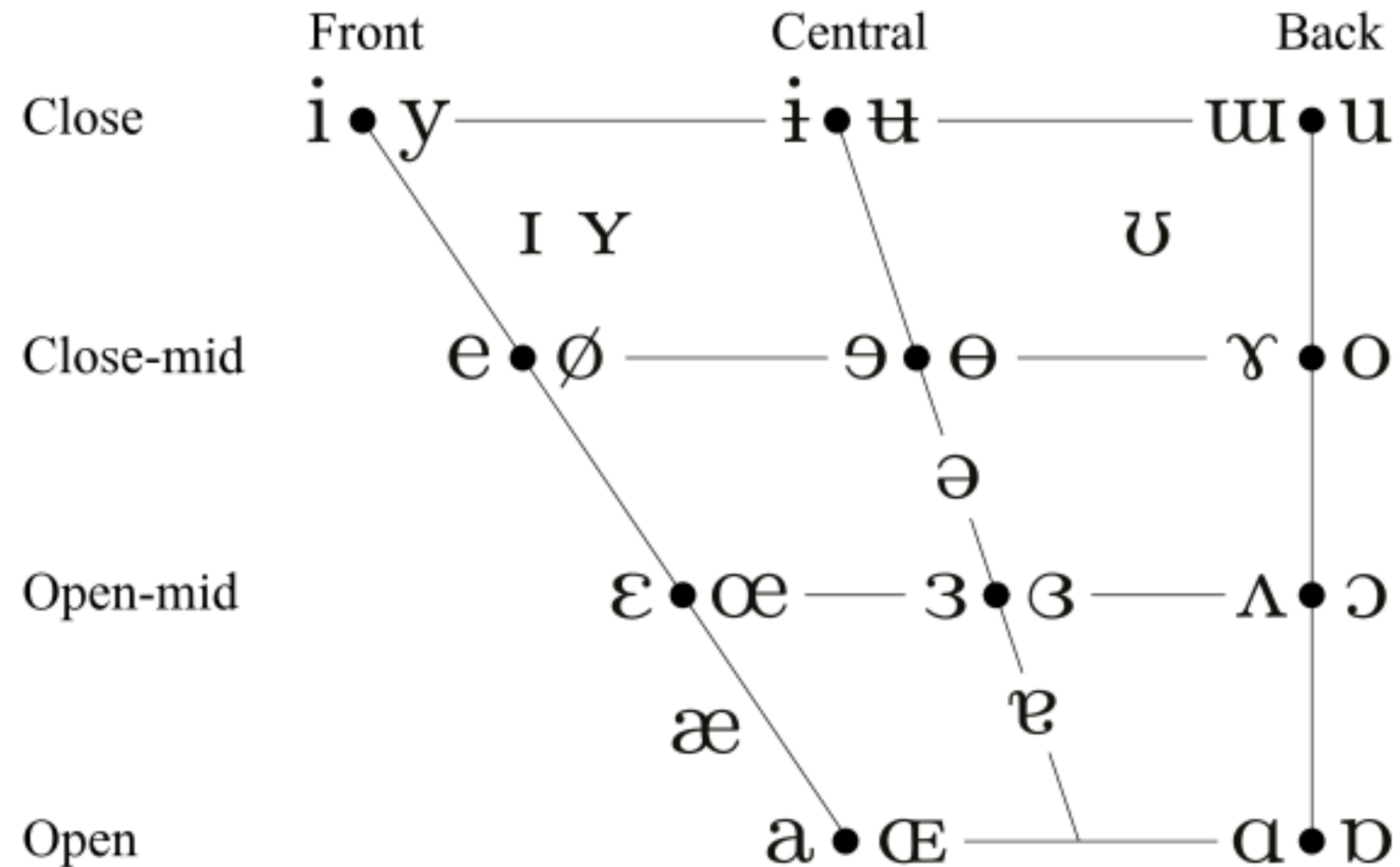
International Phonetic Alphabet

► Consonants

	Bilabial	Labiodental	Dental	Alveolar	Postalveolar	Retroflex	Palatal	Velar	Uvular	Pharyngeal	Glottal
Plosive	p b			t d		ʈ ɖ	c ɟ	k ɡ	q ɢ		ʔ
Nasal	m	ɱ		n		ɳ	ɲ	ŋ	ɴ		
Trill	ʙ			r					ʀ		
Tap or Flap				ɾ		ɽ					
Fricative	ɸ β	f v	θ ð	s z	ʃ ʒ	ʂ ʐ	ç ʝ	x ɣ	χ ʁ	ħ ʕ	h ɦ
Lateral fricative				ɬ ɮ							
Approximant		ʋ		ɹ		ɻ	j	ɰ			
Lateral approximant				l		ɭ	ʎ	ʟ			

International Phonetic Alphabet

► Vowels



Accent

- ▶ Same writing may have different pronunciation



tomato

/tə'meɪ.tou/

/tə'mɑː.tou/

Grapheme to phoneme

- ▶ Grapheme: a letter or a group of letters that represent a single phoneme
- ▶ Phoneme: the smallest unit of sound that can distinguish one word from another in a particular language
- ▶ when a child says the sound /t/ this is a phoneme, but when they write the letter 't' this is a grapheme.

Grapheme

t o m a t o

Phoneme

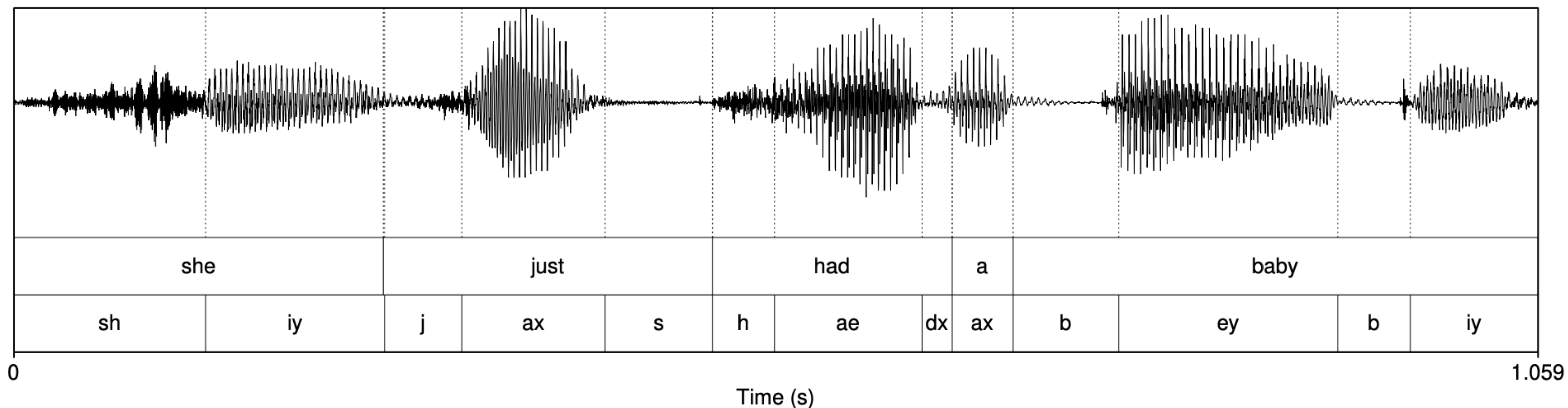
/t ə' m eɪ. t oʊ/

Grapheme to phoneme conversion

- ▶ Build a set of rules or a statistical model to convert a sequence of graphemes to phonemes

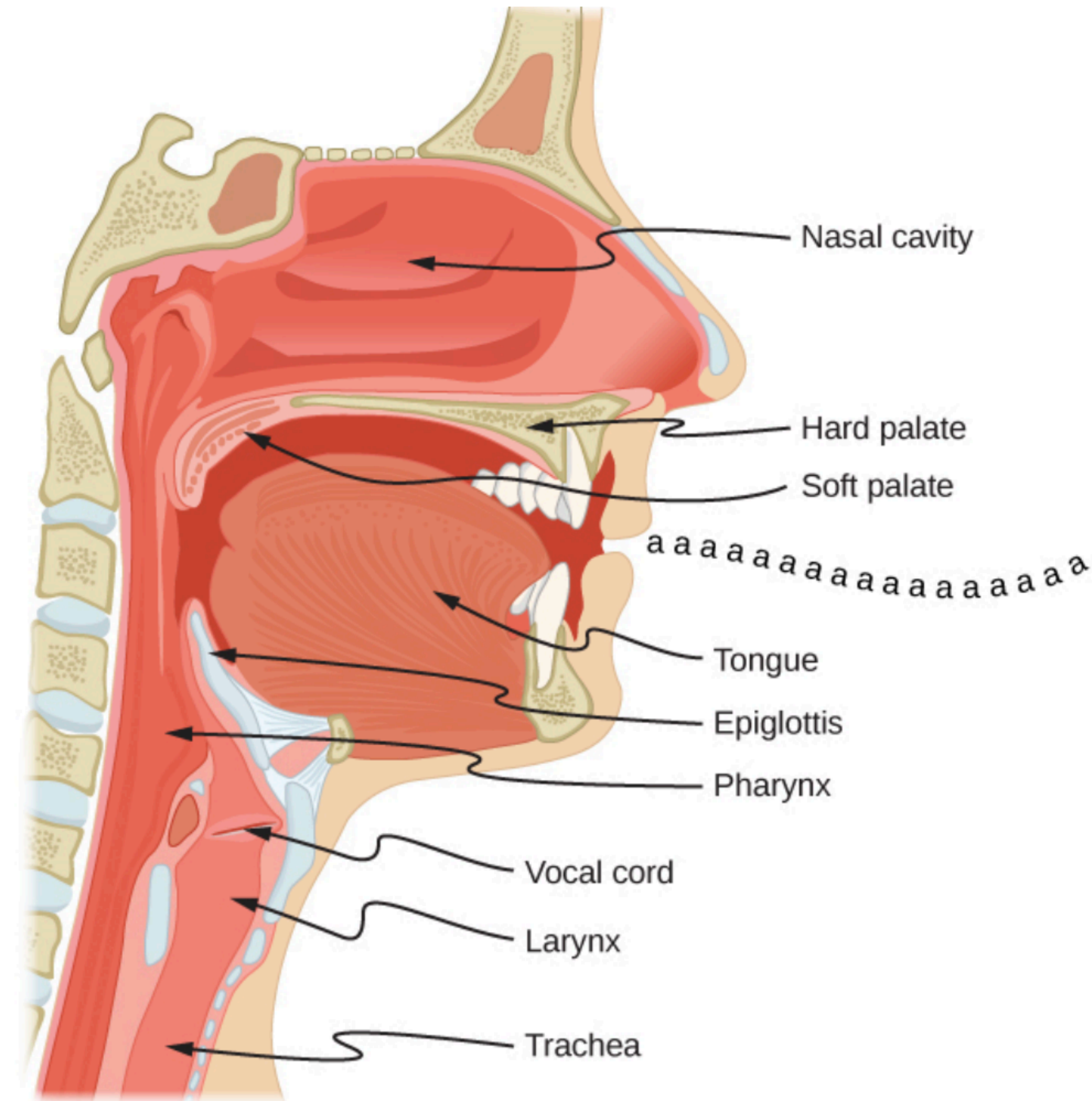
Grapheme	Phoneme
ACCENT	AH ₀ K S EH ₁ N T
ACCENTS	AE ₁ K S EH ₀ N T S
ADDICT	AH ₀ D IH ₁ K T
ADDICTS	AH ₀ D IH ₁ K T S
ADVOCATE	AE ₁ D V AH ₀ K EY ₂ T
ADVOCATES	AE ₁ D V AH ₀ K EY ₂ T S
AFFECT	AH ₀ F EH ₁ K T
AFFECTS	AH ₀ F EH ₁ K T S

Interpretation of Phones from a Waveform



Articulatory Phonetics

- ▶ Articulatory phonetics studies how phones are produced as the various organs in the mouth, throat, and nose modify the airflow from the lungs.



Vocal tract

- ▶ Vocal tract consists of **oral tract** and **nasal tract**
 - After the air leave the trachea, it can exit the body through the **mouth** or the **nose**
 - **Nasal sounds**: sounds made by air passing through the nose, they use both the oral and nasal tracts as resonating cavities
 - e.g. English [m], [n]
- ▶ Phones can be divided into two classes: vowels and consonants

Vowel and consonant

- ▶ Consonants

- Made by restriction or blocking of the airflow in some way, and can be voiced or unvoiced

- ▶ Vowels

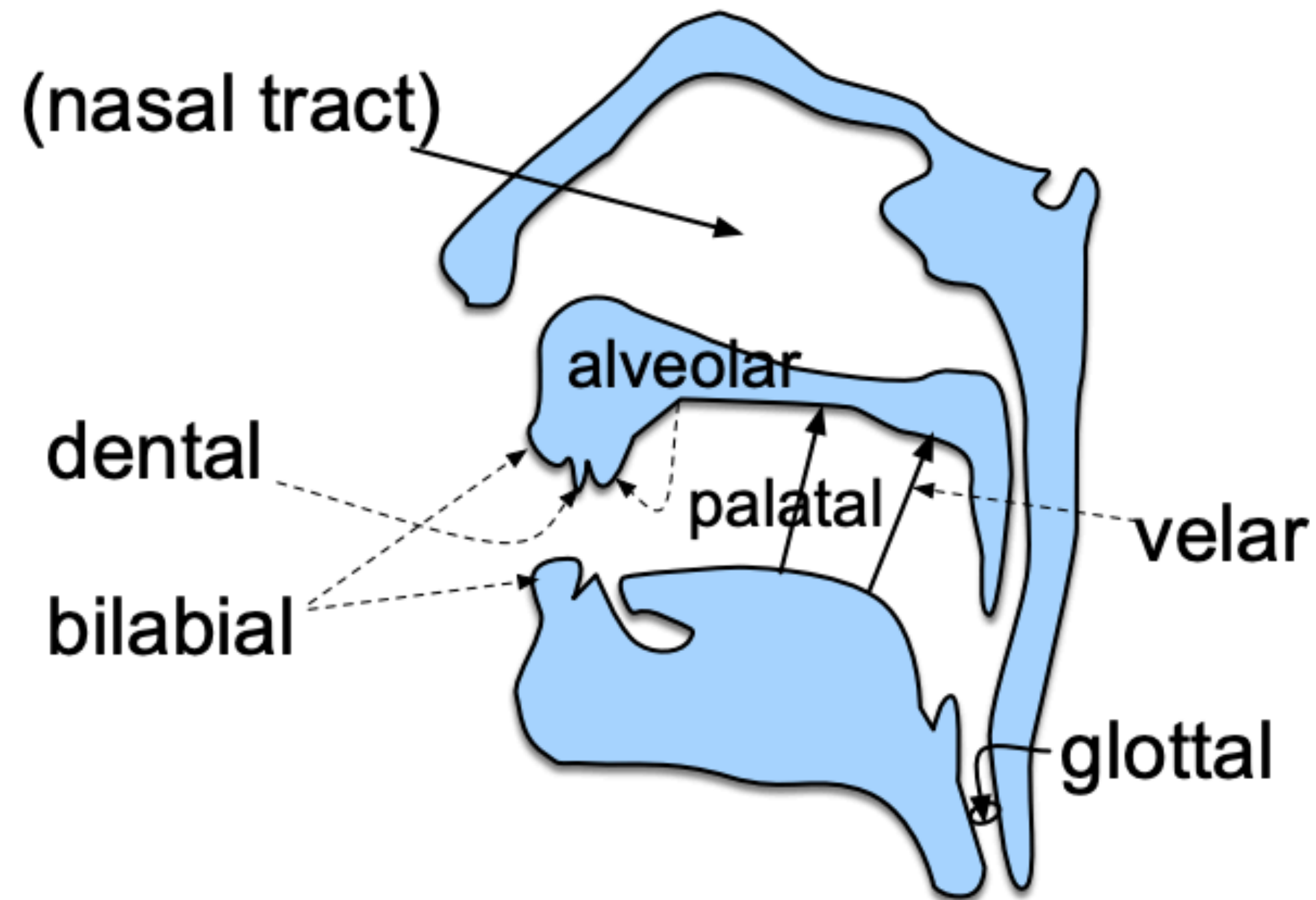
- have less obstruction, are usually voiced, and are generally louder and longer-lasting than consonants

Consonants

- ▶ A speech sound that is articulated with complete or partial closure of the vocal tract
- ▶ Place of articulations
 - where in the vocal tract the obstruction of the consonant occurs, and which speech organs are involved
- ▶ Manner of articulations
 - how air escapes from the vocal tract when the consonant sound is made

Consonants: Place of articulation

- ▶ Made by restricting airflow, and can be grouped by their place of articulation



Consonants: Place of articulation

▶ Labial

- Main restriction is formed by the two lips coming together have a bilabial place of articulation
- e.g. [p] as in pizza, [b] as in boy, [m] as in mom

▶ Dental

- Made by placing the tongue against the teeth are dentals
- e.g. [th] as in thing, [dh] as in though

▶ More classes

- <https://web.stanford.edu/~jurafsky/slp3/28.pdf>

Consonants: Manner of articulation

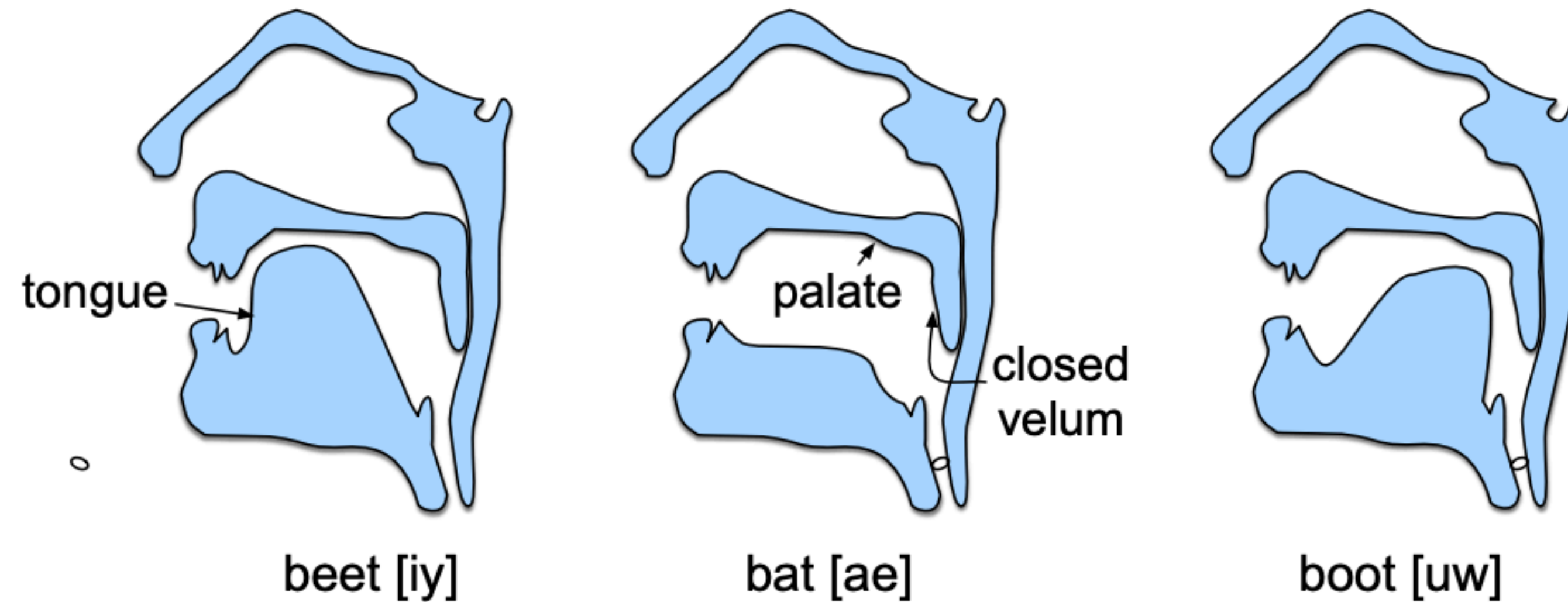
- ▶ Consonants are also distinguished by *how* the restriction in airflow is made, for example, by a complete stoppage of air or by a partial blockage
- ▶ Stop
 - airflow is completely blocked for a short time
 - Voiced: [b] [d] [g]
 - Unvoiced: [p] [t] [k]
- ▶ Fricatives
 - airflow is constricted but not cut off completely. e.g. [f] [v]

Vowel

- ▶ Vowels can be characterized by the position of the articulators as they are made
 - Heights
 - the vertical position of the tongue relative to either the roof of the mouth
 - Backness
 - the position of the tongue during the articulation of a vowel relative to the back of the mouth
 - Roundedness
 - the amount of rounding in the lips during the articulation of a vowel

Vowel

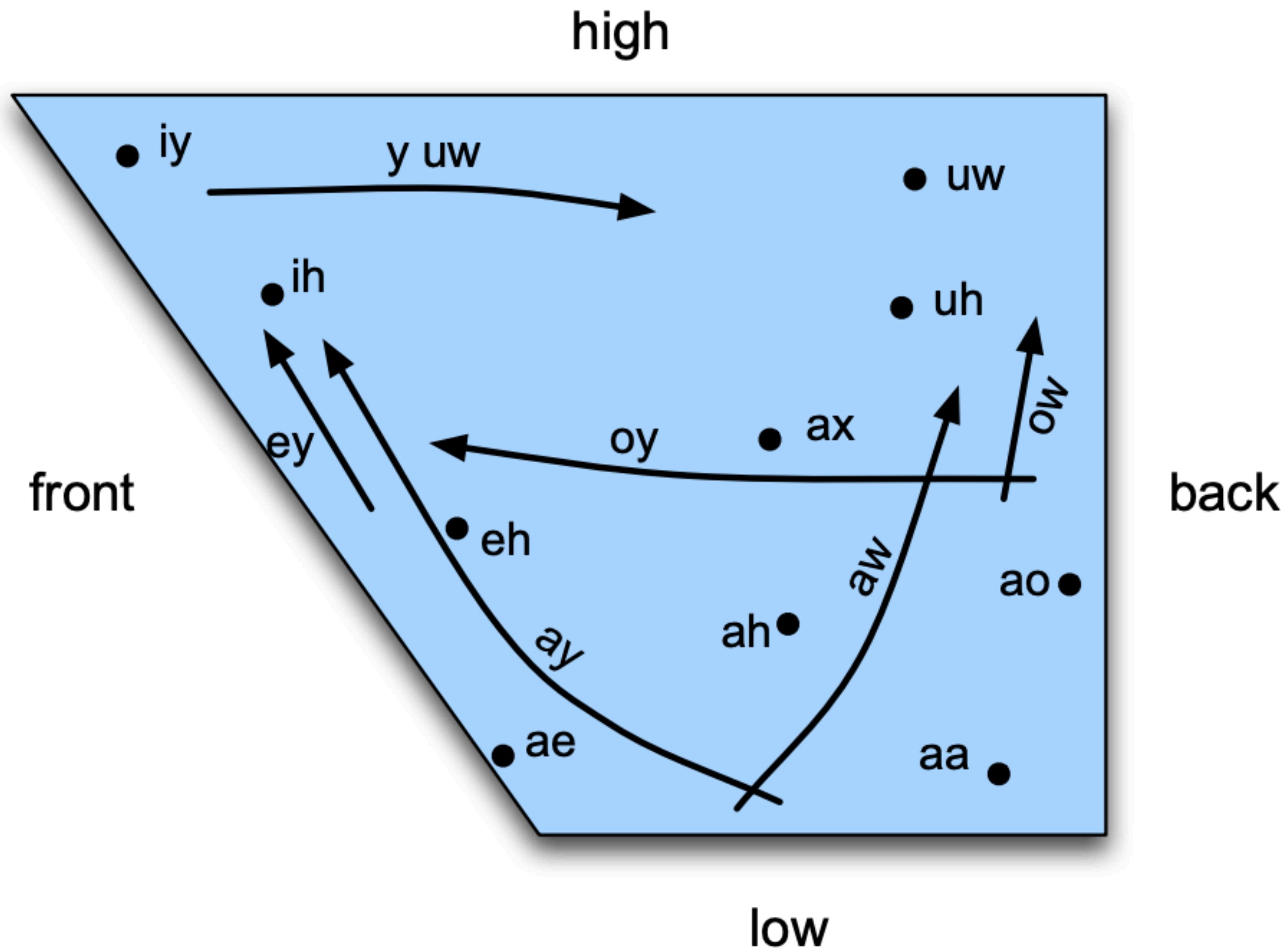
- ▶ Vowels can be characterized by the position of the articulators as they are made
 - Tongue positions for English high front [iy], low front [ae] and high back [uw].



Vowels sounds

- ▶ Heights
 - the vertical position of the tongue relative to either the roof of the mouth
- ▶ Backness
 - the position of the tongue during the articulation of a vowel relative to the back of the mouth
- ▶ Roundedness
 - the amount of rounding in the lips during the articulation of a vowel

Vowel space



Vowels sounds

- ▶ Two types of vowel sounds

monophthongs

One vowel

Examples

Me, that, this, work

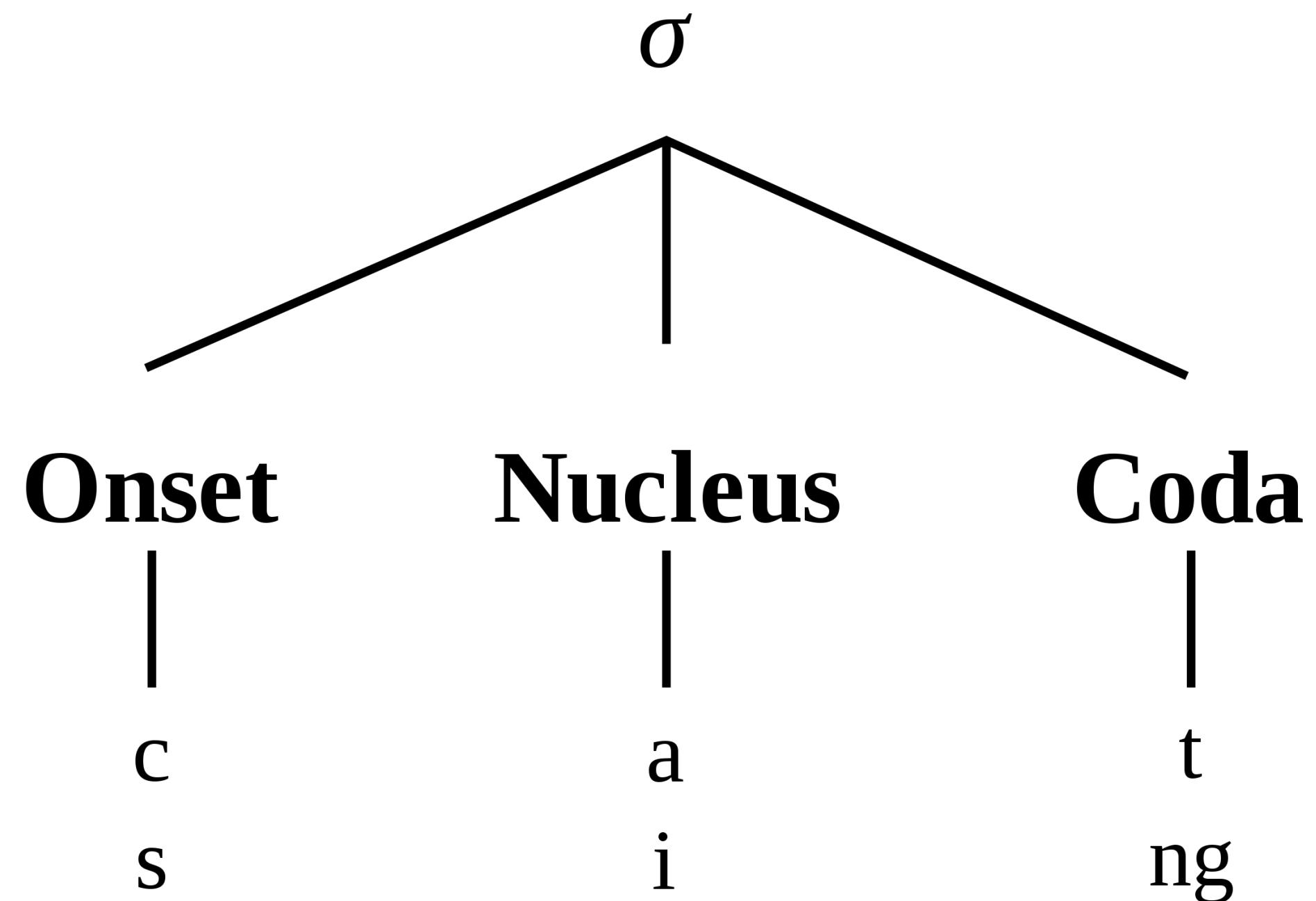
diphthongs

Two vowels

Play, town, slow, toy

Syllable

- ▶ a unit of organization for a sequence of speech sounds
 - typically made up of a syllable nucleus (most often a vowel) with optional initial and final margins (typically, consonants).
- ▶ Syllables are often considered the phonological "building blocks" of words.



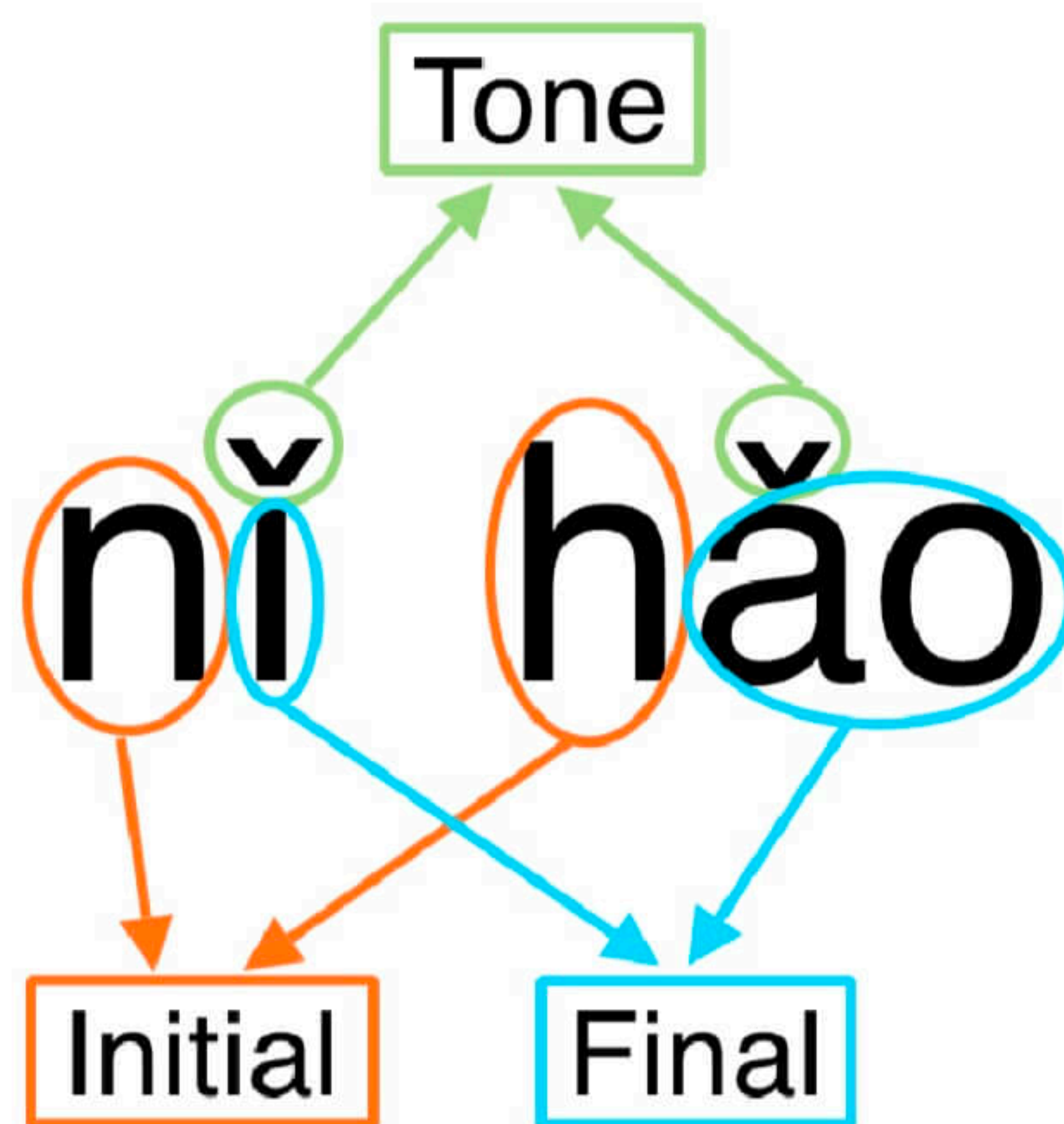
Syllable

► Examples

- Congratulation 5 syllables: con-grat-u-la-tion
- International 5 syllables: in-ter-na-tio-nal
- Water 2 syllables: wa-ter
- Group 1 syllable
- Categorization 6 syllables: cat-e-go-ri-za-tion

Syllable: Mandarin Chinese

- ▶ There are about 1300 syllables
 - Each syllable consists of an initial, a final and a tone



Summary

- ▶ International Phonetic Alphabet, and Grapheme-to-phoneme conversion
- ▶ Articulatory phonetics: vowels and consonants

Reading

- ▶ Chapter 28: Phonetics
 - <https://web.stanford.edu/~jurafsky/slp3/28.pdf>