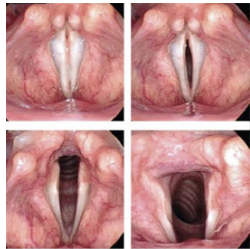LING 1010

Language and Mind

Prof. Jon Sprouse

02.01.21:   Articulatory features
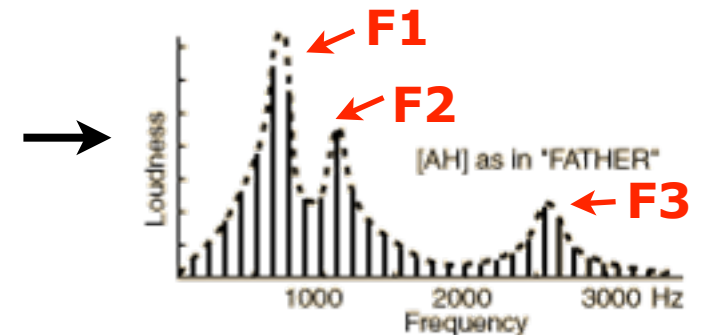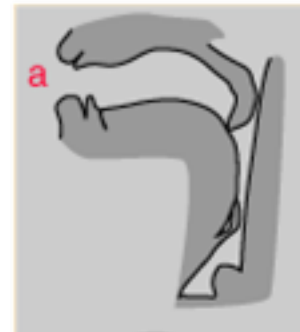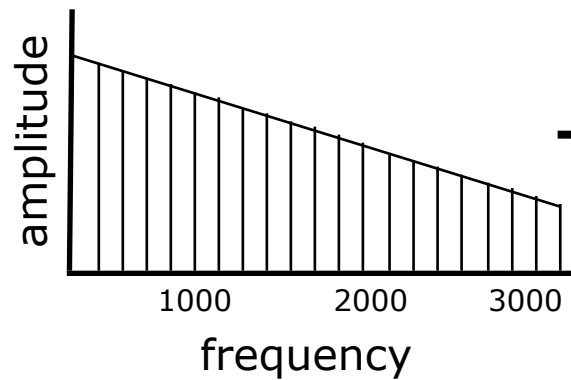
Quick Recap: Formants

source: vocal folds   **filter:** vocal tract "ah"

For the human voice, we call the highest amplitude frequencies that occur after filtering (i.e., in the EQ settings) the **formants**.

Formants are the highest amplitude peaks in the frequency spectrum created by the human vocal tract.

Much like harmonics, we label the formants in order beginning with the lowest frequency (F1, F2, etc.).

Sine Wave Speech

The importance of formants to speech can be seen by deconstructing normal human speech. Here is a typical English sentence deconstructed into the following parts:

1. Its first formant (F1)

2. Its second formant (F2)

3. Its third formant (F3)

4. The sounds that are not formants (pops, cracks, shushes)

OK, now let's put the first three formants together (without the pops and cracks), and see what it sounds like!

Could you get it? If not, here is the original recording of the sentence.

And here are the three formants again.

Sine Wave Speech

For those of you reading this as a pdf, here is a link to the sounds on the previous slide:

http://www-oedt.kfunigraz.ac.at/hlt/content/04LV4.../01-LADEFOGED-VOWELS-KONS-CD/vowels/chapter7/abirdinthehand.html

Disclaimer:

Sine wave speech isn't perfectly intelligible. For most people, it is much easier to recognize as speech if you already know what is being said. This shows that the formants aren't quite enough for normal speech perception (we will come back to this soon). But the fact that you can understand SWS does show that formants play a large role in phonemes!

The next step

We have a process for detecting phonemes:

s a t

s a d

Different words, distinct phonemes!

We have a set of symbols to represent each phonemes (IPA):

And we know that each phoneme has distinct formants created by the shape of the vocal tract:
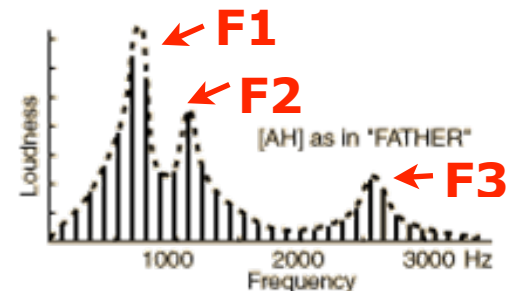
The next question we can ask is: How does the mind **represent** phonemes?

How does the mind represent phonemes?

Articulatory Features

Representations

A representation is just what it sounds like: an object that stands in a symbolic relationship with another object

let's pretend that this a real tree (and not an image in a slideshow)

a picture of a tree is a representation of the real tree

Representations

A representation is just what it sounds like: an object that stands in a symbolic relationship with another object

let's pretend that this a real tree (and not an image in a slideshow)

a drawing of a tree is also a representation of a tree

Mental Representations

A representation is just what it sounds like: an object that stands in a symbolic relationship with another object

A mental representation is also just what it sounds like: a representation made by your mind

The most frequent mental representations are the ones that you make every time you take in sensory input from the world.

When you see a tree, you are actually perceiving the mind's representation of that tree based on the workings of the visual system

of the Human Brain

Parietal Lobe

Occipital Lobe

Lobe

Temporal Lobe

Brain Stem

Cerebellum

Spinal Cord

Front

Back

Mental Representations

A representation is just what it sounds like: an object that stands in a symbolic relationship with another object

A mental representation is also just what it sounds like: a representation made by your mind

The most frequent mental representations are the ones that you make every time you take in sensory input from the world.

When you hear a sound, you are actually perceiving the mind's representation of the air vibrations based on the workings of the hearing system

of the Human Brain

Parietal Lobe

Occipital Lobe

Lobe

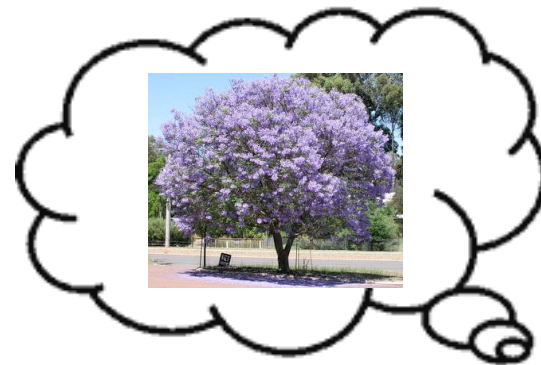Temporal Lobe

Brain Stem

Cerebellum

Spinal Cord

Front

Back

Mental Representations

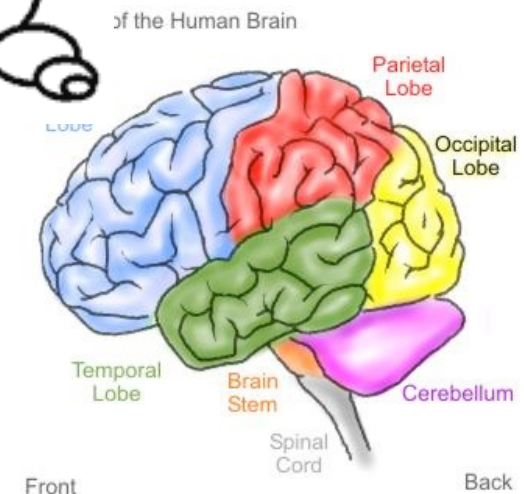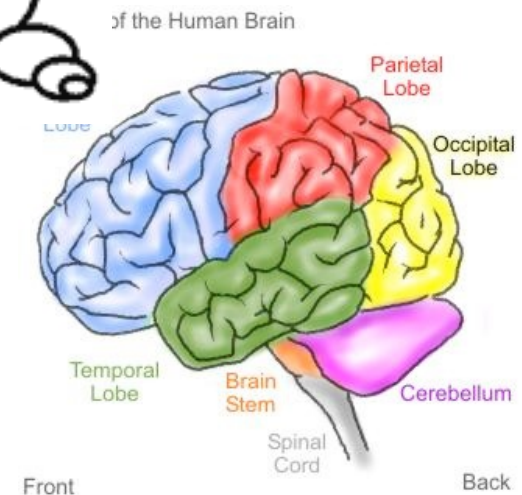A representation is just what it sounds like: an object that stands in a symbolic relationship with another object

A mental representation is also just what it sounds like: a representation made by your mind

In this way, you are already very familiar with mental representations.

Mental representations are how you interact with the world all of the time -- your mind creates representations of the world through your sensory systems

of the Human Brain

Parietal Lobe

Occipital Lobe

Temporal Lobe

Brain Stem

Cerebellum

Spinal Cord

Front

Back
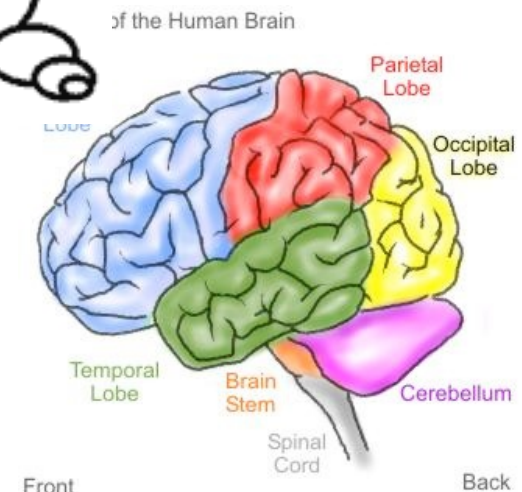
For the philosophers in the room: This means that you never actually perceive reality the way it is, but rather how your mind represents it. We just assume that it represents it correctly --- but that is not necessarily true. (Ask Rene Descartes)

How does the mind represent phonemes?

Once we have the idea of mental representations, it is natural to ask how the mind represents phonemes.

The simplest answer has probably already occurred to you. Since phonemes are made up of formants, and formants are just frequencies, perhaps the mind simply represents phonemes using the formant frequencies.

"ah"
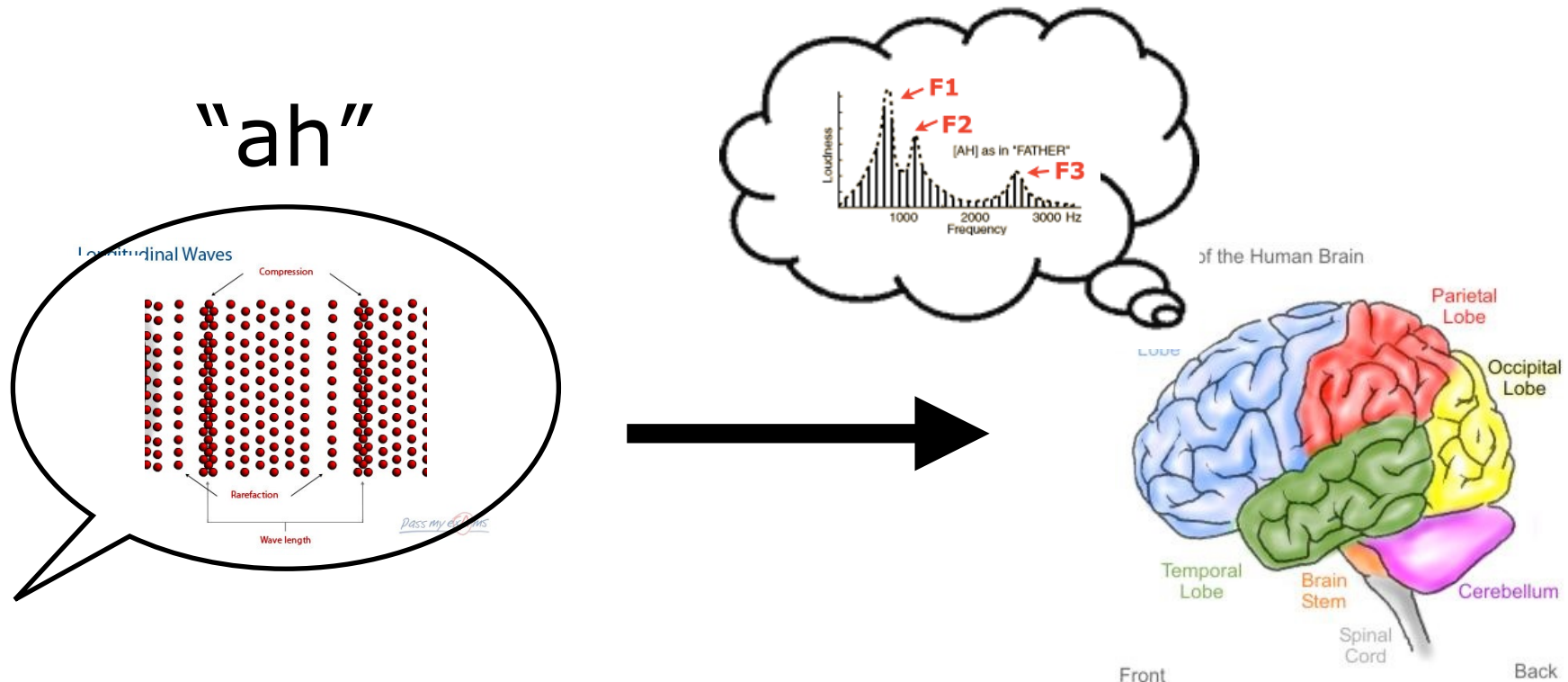
This would make lots of sense… but we have reason to believe the picture is more complicated than this!

Here is one reason

https://auditoryneuroscience.com/McGurkEffect

We call this the McGurk Effect

This is called the **McGurk Effect**.

The audio track is the syllable **BA.**

The visual track is the syllable **GA.**

The percept is either **DA** or **THA.**

The McGurk effect suggests that we fuse audio and visual information during the perception of speech.

This raises an interesting problem for the idea that phonemes are represented directly as formant frequencies.

If the mind simply represented phonemes as formants, why would visual information change the perception of the phoneme? The formants for BA are present in the auditory signal. No other formants are there. How could visual information affect an auditory representation? That would be very surprising.

Visual information is a large component of speech perception

If you haven't seen the McGurk effect before, it may have surprised you. But in reality, visual information is a large component of speech perception.

"Can you hear me now?"

We've all experienced this when talking on the phone: sometimes there isn't enough information in the auditory signal to let us figure out the sound.
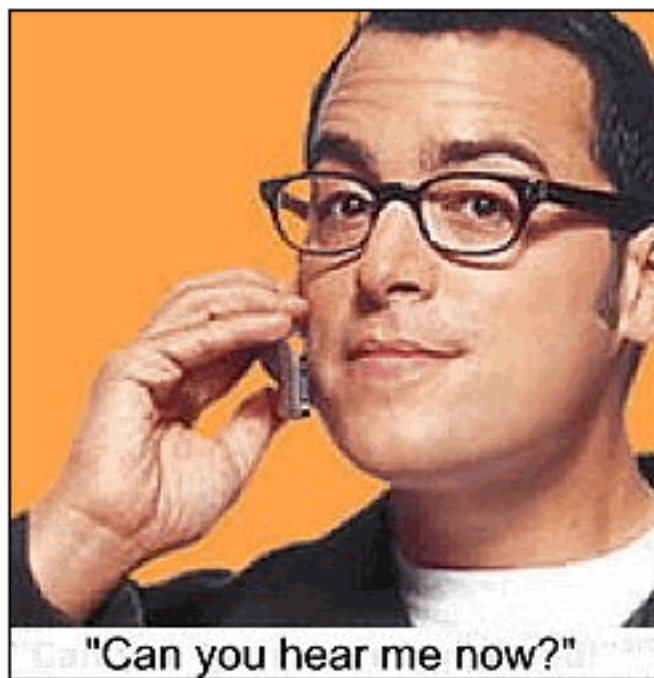
This is also happens when older people (with hearing loss) claim that they "can't hear without their glasses" - they use visual (motor) information to compensate for the degraded auditory information due to hearing loss.

Articulatory Features

The McGurk effect (and other evidence we will see soon) leads linguists to believe that the representation of phonemes is not in terms of formants at all.

Instead, linguists have proposed that phonemes are represented as a combination of **articulatory features**.

The word feature is just a formal way of saying "property". The word articulate is just a formal way of saying "produce". So an **articulatory feature** is a property of a phoneme that is related to the way it is produced.

Another way to think about articulatory features is as gross (meaning large or coarse) motor commands. Each phoneme may have several articulatory features, especially if it takes several different (simultaneous) motor commands to produce a given phoneme.

When investigating phonemes, linguists attempt to identify the smallest set of articulatory features that can be used to describe a phoneme. They then look for patterns of articulatory features among different phonemes.

Articulatory features of vowels

Linguists have been able to identify **two types** of articulatory features for vowels. Both are related to the position of the tongue during production:

1. **Height**: Is the tongue low or high

2. **Backness:** Is the tongue forward or back?

And here is an interactive chart with MRIs of a real human producing speech sounds (it also has animations)

tongue

https://www.seeingspeech.ac.uk/ipa-charts/?chart=4
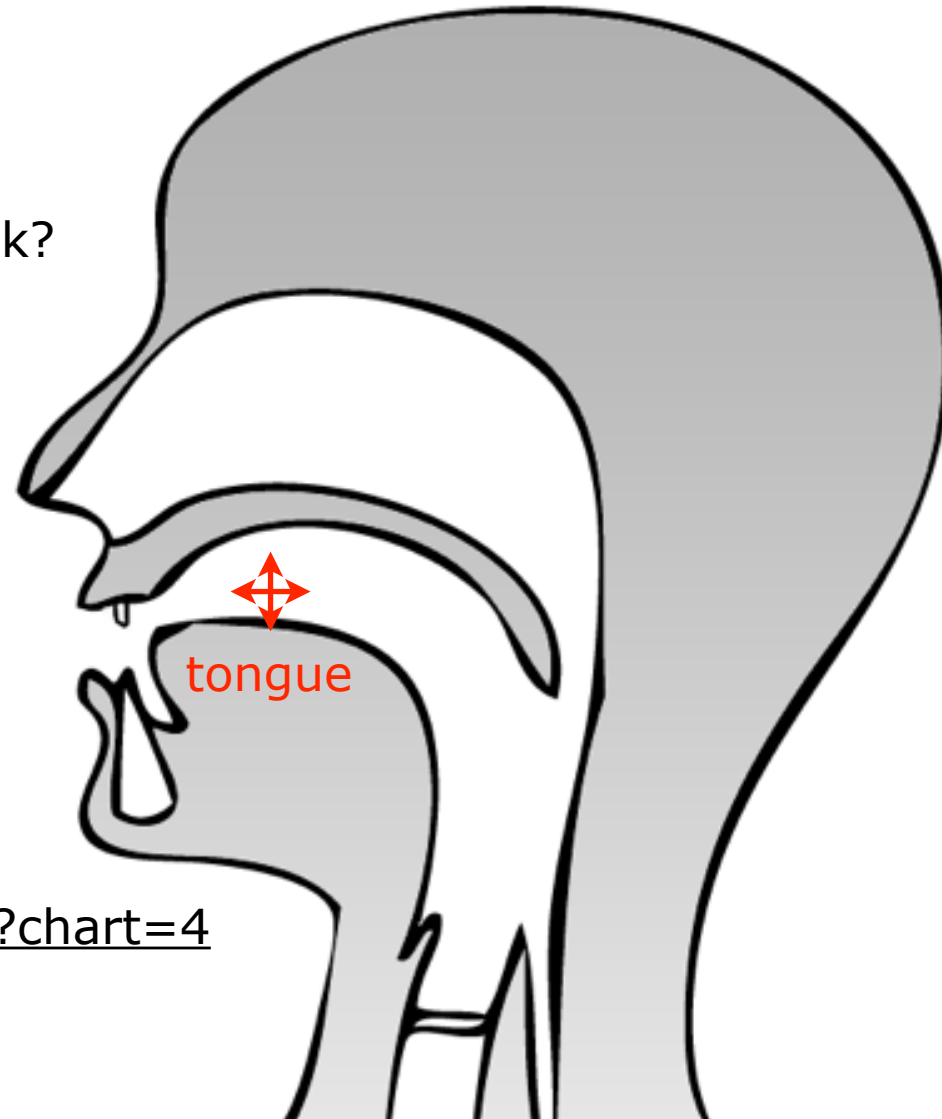
Monophthongs: fancy word for single vowels

Front **Mid** **Back**

High

heed — **i**
hid — **ɪ**
who'd — **u**
hood — **ʊ**

Mid

hay — **e**
hope — **o**
head — **ɛ**
ə
hut — **ʌ**
caught — **ɔ**

Low

had — **æ**
a
father — **ɑ**
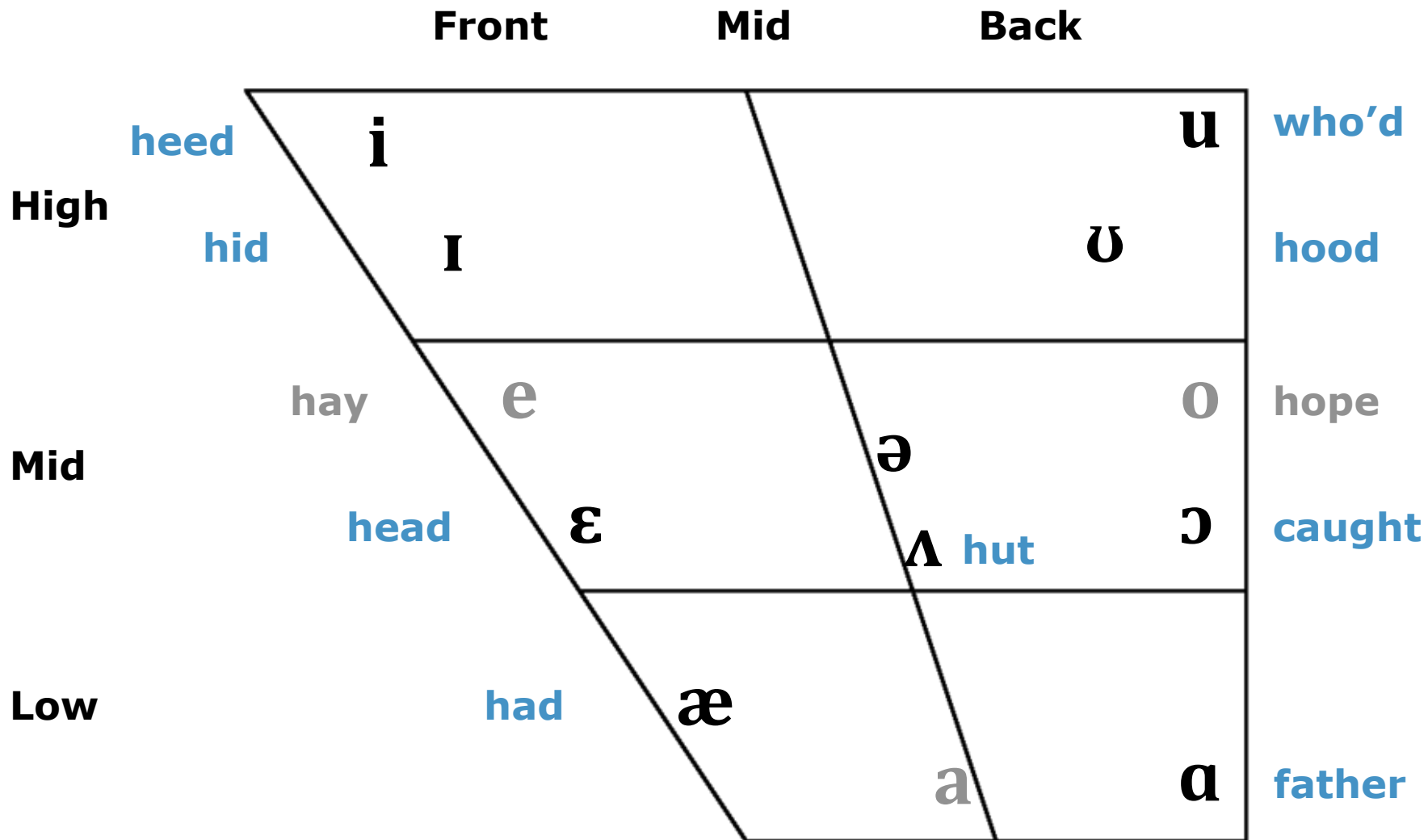
The ones in gray don't really occur in US English alone, but we use them for diphthongs so I have included them here.

https://www.seeingspeech.ac.uk/ipa-charts/?chart=4

Diphthongs: double vowels!

Front   **Mid**   **Back**

heed   **i**   **u**   who'd

High

hid   **ɪ**   **ʊ**   hood

hay   **e**   **o**   hope

Mid

head   **ɛ**   **ə**   **ɔ**   caught

ʌ hut

aʊ house

Low   had   **æ**

a   **ɑ**   father

A diphthong is a vowel that starts with the tongue in one location in the mouth, and ends with it in a second location. But crucially, the movement is done very quickly, so the entire articulation only takes the place of a single vowel! (http://www.paulmeier.com/diphthongs-and-triphthongs/)

Diphthongs: double vowels!

Front   **Mid**   **Back**

heed   **i**   **u**   who'd

High

hid   **ɪ**   **ʊ**   hood

hay   **e**   **o**   hope

ə

Mid

head   **ɛ**   **ʌ** hut   **ɔ**   caught

buy aɪ

Low   had   **æ**

a   **ɑ**   father
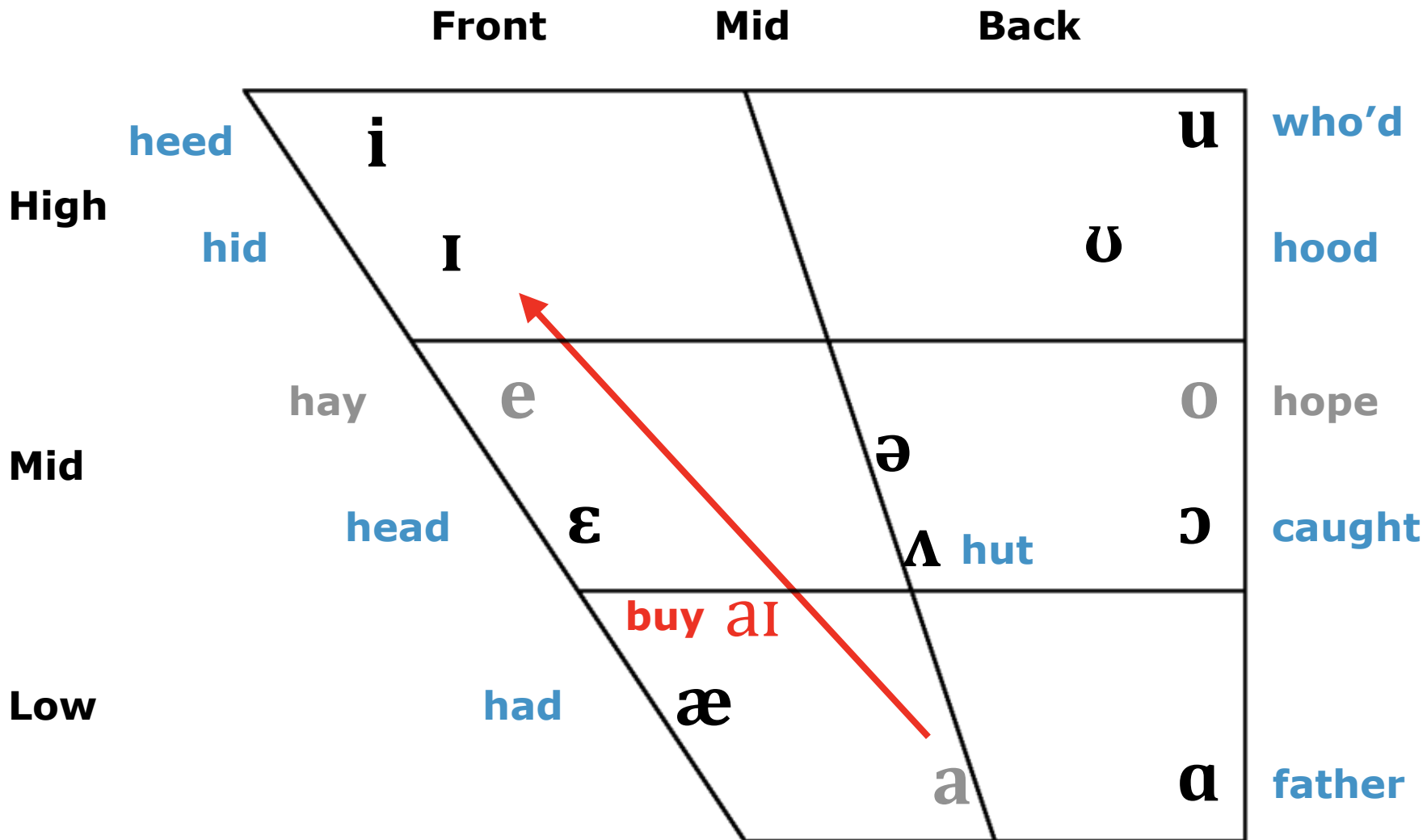
A diphthong is a vowel that starts with the tongue in one location in the mouth, and ends with it in a second location. But crucially, the movement is done very quickly, so the entire articulation only takes the place of a single vowel! (http://www.paulmeier.com/diphthongs-and-triphthongs/)

Diphthongs: double vowels!

Front   **Mid**   **Back**

High

heed   **i**   **u** who'd

hid   **ɪ**   **ʊ** hood

Mid

hay   e   ɔɪ **boy**   o hope

            ə

head   **ɛ**   **ʌ** hut   **ɔ** caught
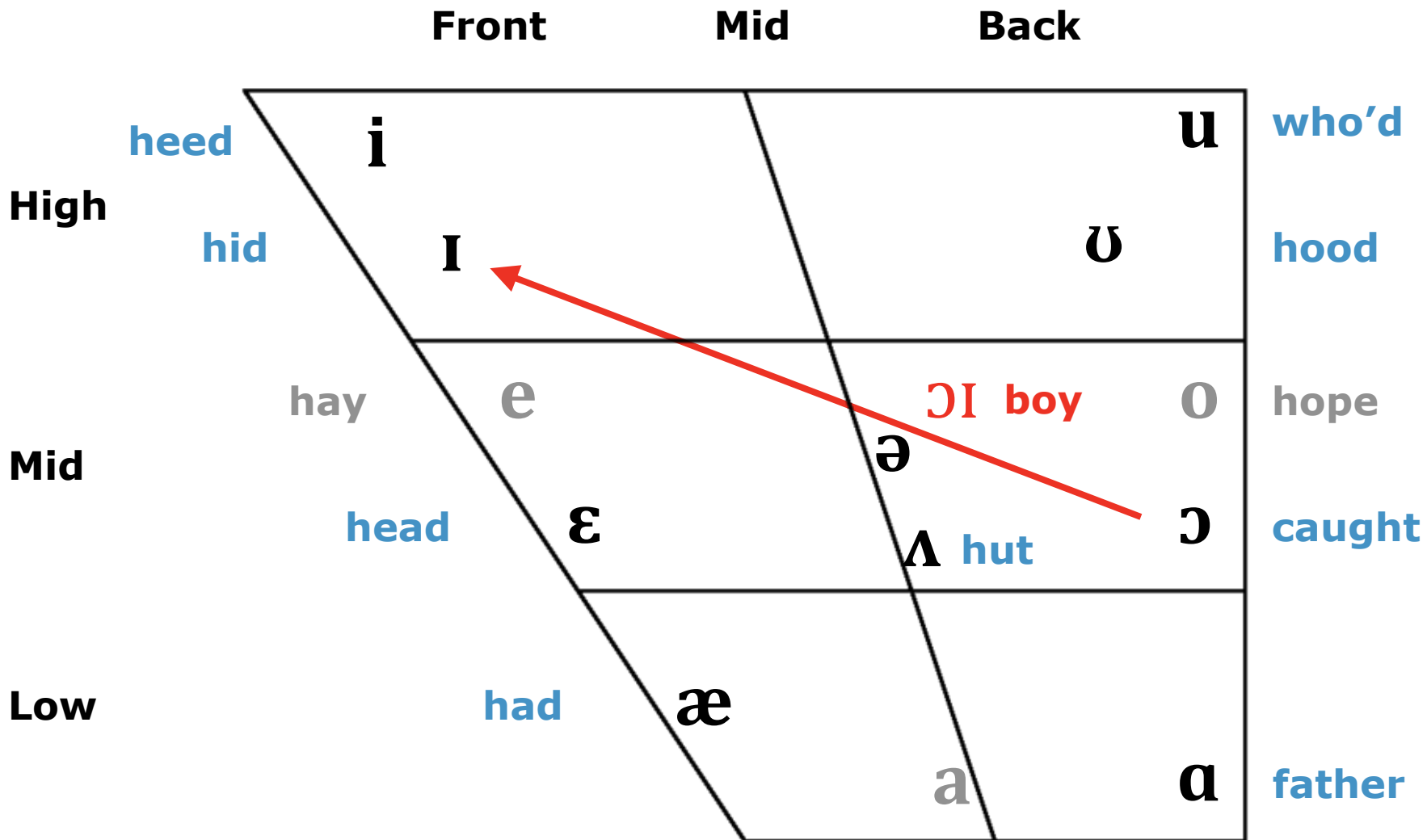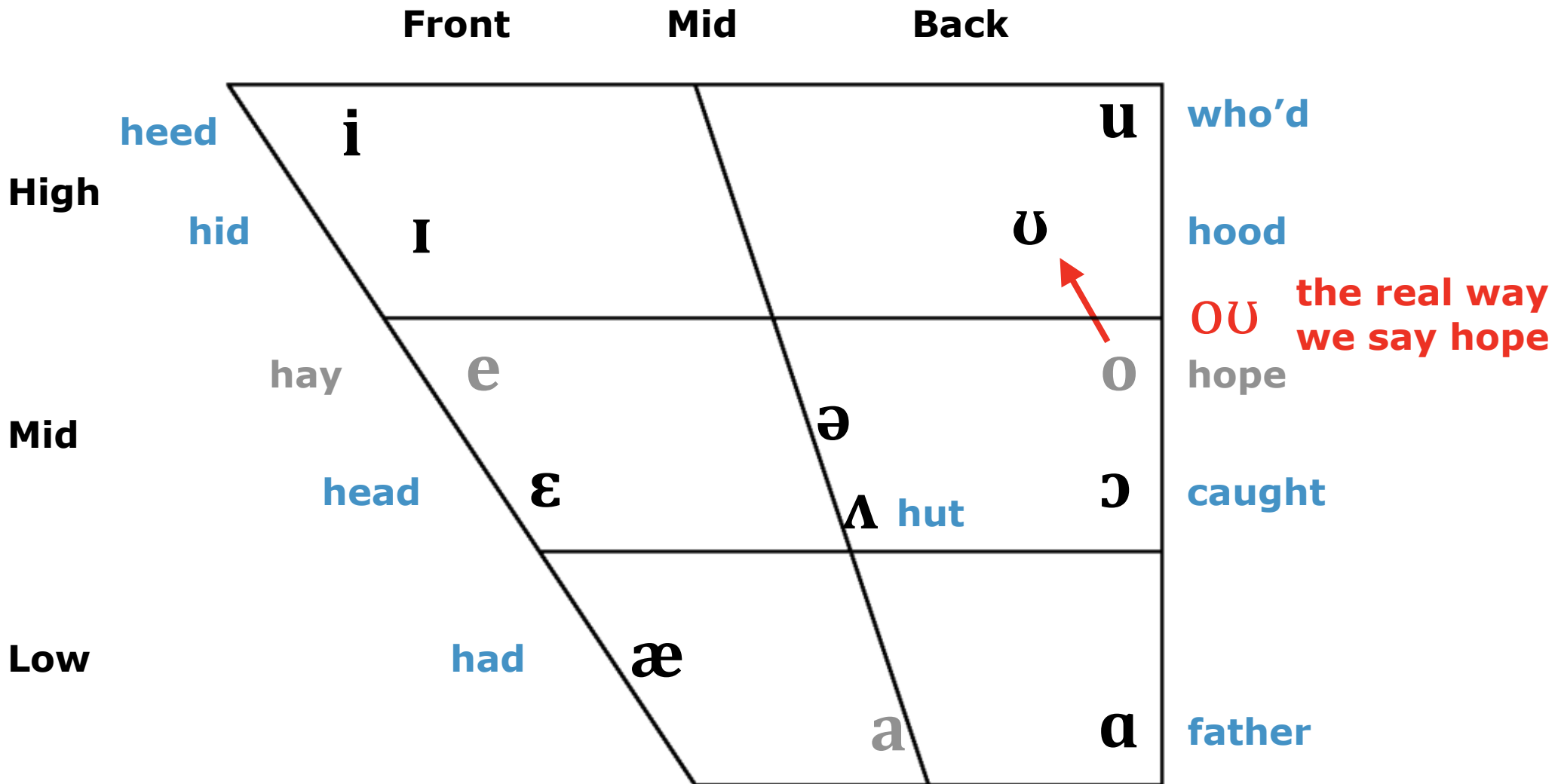
Low

had   **æ**

             a   **ɑ** father

A diphthong is a vowel that starts with the tongue in one location in the mouth, and ends with it in a second location. But crucially, the movement is done very quickly, so the entire articulation only takes the place of a single vowel! (http://www.paulmeier.com/diphthongs-and-triphthongs/)

Diphthongs: double vowels!

If you ever took a romance language in school (Spanish, Italian, Portuguese), this is why your teacher told you that you were pronouncing your o's wrong. These languages have a monophthong vowel for o, and we have a diphthong.

Diphthongs: double vowels!
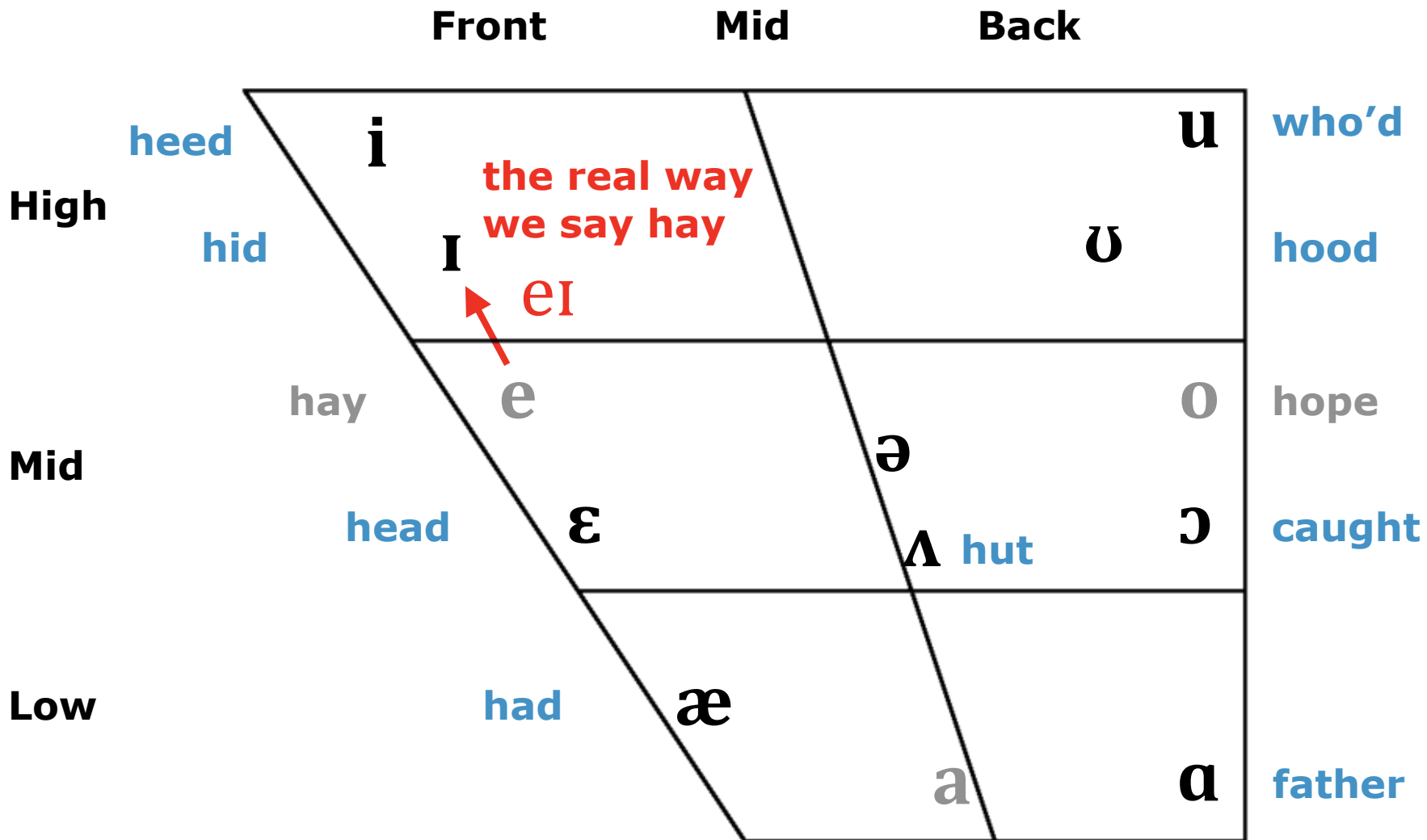
Front **Mid** **Back**

High

heed **i**

hid **ɪ**

the real way
we say hay

eɪ

hay **e**

u who'd

ʊ hood

o hope

Mid

head **ɛ**

ə

ʌ hut

ɔ caught
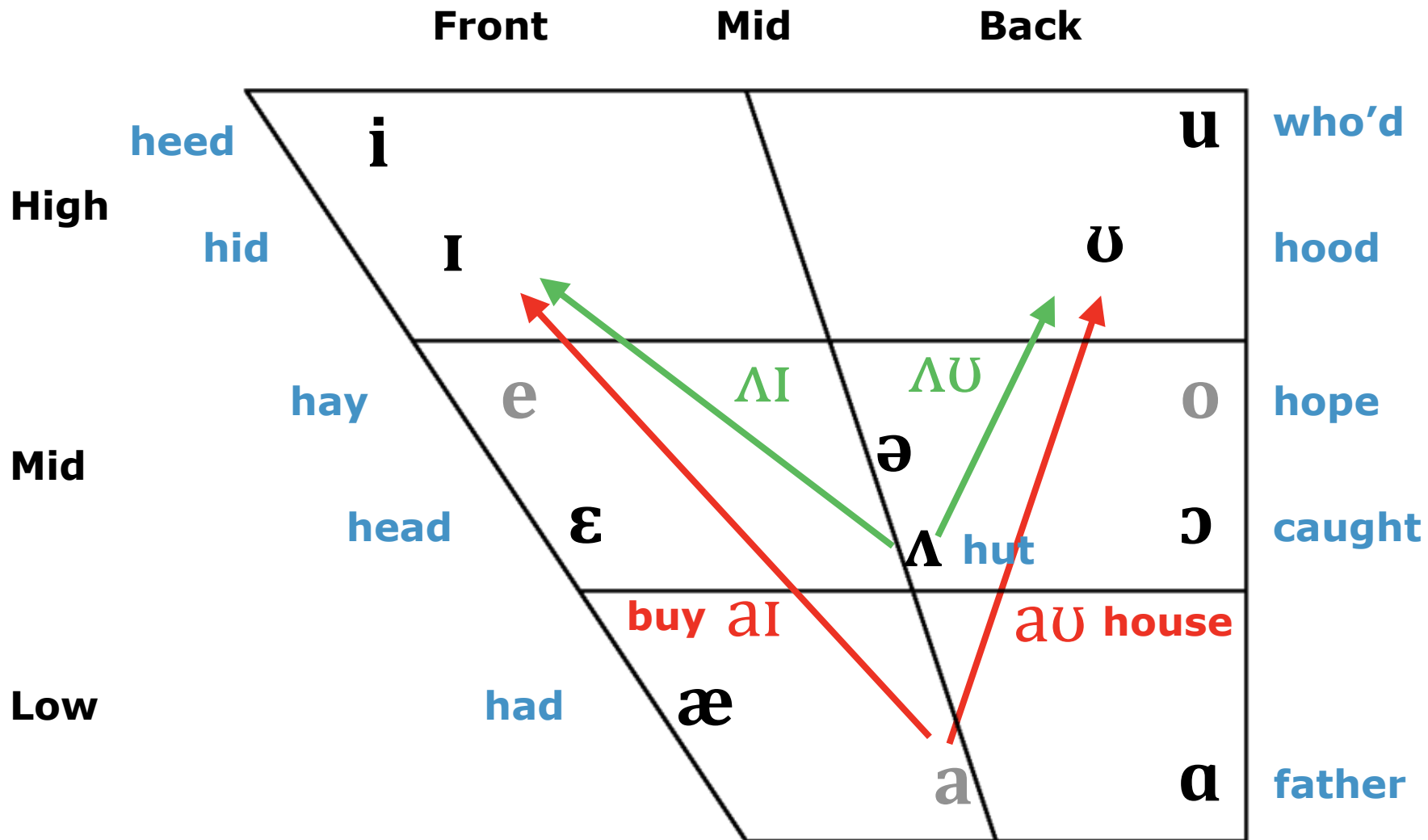
Low

had **æ**

a **ɑ** father

If you ever took a romance language in school (Spanish, Italian, Portuguese),
this is why your teacher told you that you were pronouncing your e's wrong.
These languages have a monophthong vowel for e, and we have a diphthong.

Why do Canadians sound different?

Front   **Mid**   **Back**

High

Mid

Low

heed — **i**
hid — **ɪ**
hay — **e**
head — **ɛ**
had — **æ**

who'd — **u**
hood — **ʊ**
hope — **o**
caught — **ɔ**
father — **ɑ**

ə
ʌ hut
a

ʌɪ   **ʌʊ**

buy aɪ   **aʊ house**

For two common diphthongs in English, Canadians begin the diphthong with their tongues a bit higher. They end in the same place, but start higher. We call this Canadian Raising.

Articulatory features of consonants

Linguists have been able to identify **three types** of articulatory features for consonants.

1. **Place of Articulation**:

   Where in the vocal tract is the airflow being obstructed?

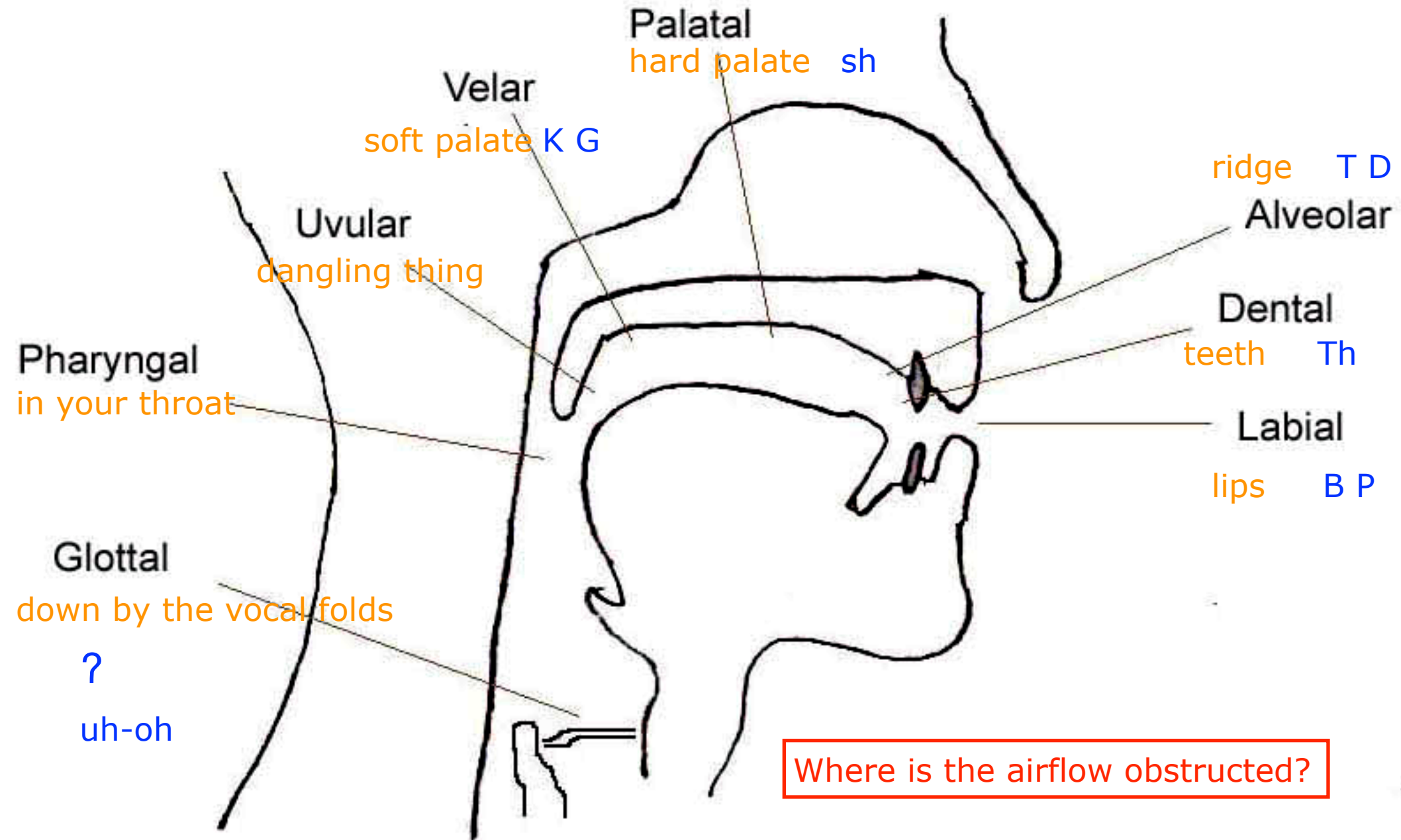2. **Manner of Articulation**:

   How is the airflow being obstructed?

3. **Voicing**:

   Are the vocal folds vibrating during this obstruction or not?

And here is an interactive chart with MRIs of a real human producing speech sounds (it also has animations)

https://www.seeingspeech.ac.uk/ipa-charts/?chart=1

Place of Articulation

Palatal
hard palate   sh

Velar
soft palate K G

ridge   T D
Alveolar

Uvular
dangling thing

Dental
teeth   Th

Pharyngal
in your throat

Labial
lips   B P

Glottal
down by the vocal folds

?

uh-oh

Where is the airflow obstructed?

Manners of Articulation

Stops (Plosives) - the airflow is completely obstructed

B P T D G K

Nasals - the airflow is diverted to the nasal cavity

N M

Fricatives - the airflow is disturbed, but not completely stopped

F V Th Sh

Affricates - a stop + fricative

Ch J

Laterals - the tongue blocks the air, but air escapes around the sides

L

Approximants - not much obstruction, very similar to vowels
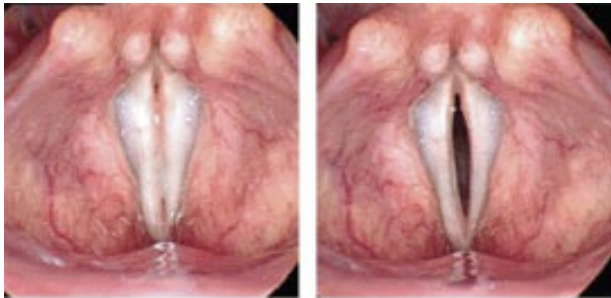
R Y

How is the airflow obstructed?

Voicing

The vocal folds must vibrate to create the frequencies necessary for vowels.

But when it comes to consonants, which are just obstructions to the airflow, there are two options.

Voiced: Let the vocal folds vibrate during the obstruction

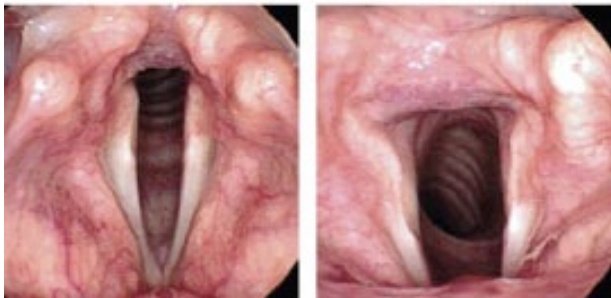**Voiceless**: Stop the vocal folds from vibrating during the obstruction

Stroboscopical imaging of the vocal fold
movement using the LED laryngoscope

Voiced Stops

B D G

Voiceless Stops

P T K

Are the vocal folds vibrating during this obstruction?

A feature-based consonant chart

Don't worry, you **don't have to memorize this**. It is just a useful chart.

	Bilabial	Labiodental	Interdental	Alveolar	Palatal	Velar	Glottal
Stop	p b			t d		k g	?
Nasal	m			n		ŋ	
Fricative		f v	θ ð	s z	ʃ ʒ		
Affricate					tʃ dʒ		
Glide					j	w	h
Liquid				l	r		

Places of Articulation form the column labels

Manners of Articulation form the row labels

Voicing is delineated by the pairs: the one on the right is voiced

And here is a rapper putting it all together

https://www.youtube.com/watch?v=_LGkbvkCS3I

Articulatory features are structure!

Once again, we see an amazing bit of structure in the mind. Our minds convert physical speech sounds into bundles of articulatory features.

This isn't something you were explicitly taught, nor is it something you can decide not to do. It simply seems to be the way the human mind represents speech sounds!

Vowels:

Height

Backness

Consonants:

Place of Articulation

Manner of articulation

Voicing

Some conclusions

Formants are the **physical properties** of phonemes that appear to be most critical to speech perception

Sine wave speech demonstrates the importance of formants. However, it is not perfectly intelligible on its own.

We <u>don't represent phonemes using formants directly.</u>

The McGurk effect suggests there is more to phoneme representation than acoustic information.

Instead, we represent phonemes using **articulatory features**, which can be thought of as formal abstractions of gross motor commands.

      Consonants are each made up of three features: place of articulation, manner of articulation, and voicing.

      Vowels are each made up of two features: tongue height and tongue backness