LING 1010



Language and Mind Prof. Jon Sprouse

01.27.21: Phonemes (i.e., speech sounds)

Let's start with sound and work our way up



Two big questions for today

As our first foray into the speech/sound portion of language, today we will focus on the first two logical questions that we can ask about how speech and sound works in language:



1. What are the speech sounds in a given language (e.g., English)?

2. What are the physical properties of the speech sounds in a given language (e.g., English)?

What are the speech sounds in English?

Phoneme: A technical term for speech sound

The first thing we are going to do is science-up our terminology. In this class we are going to need to be precise about the objects we are talking about, so we need technical terms that have precise meanings.

Though we all know intuitively what a speech sound is, in linguistics we use the technical term **phoneme** instead of the intuitive term "speech sound".

Because the primary function of speech sounds in a language is to construct distinct words, we can use the fact that distinct words will have at least one distinct speech sound as a technical definition for phoneme:

phoneme: The smallest segment of speech that leads to a meaningful difference between words.

The nice thing about this definition is that it comes with a built in test. Take a word, change one of the speech sounds in it. If the change leads to a change in the meaning of the word, then the original sound and the new sound are distinct phonemes.

The cognitive test for phonemes

Step 1: Pick a word in the language.	s a t	
		Different words!
Step 2: Change one sound in the word.	s a d	

If the changed sound leads to a different word, then the two sounds (the original and the new one) are **both** distinct phonemes in the language!

Here is another example of two distinct phonemes:

lake Different words! rake

So | and r are distinct phonemes in English.

And here is an example of two sounds that aren't distinct phonemes:

s a t		t ^h is a fancy way of saying "pronounce a t and blow out at the
s a <mark>t</mark> h	Same word!	same time". So t and t ^h are not distinct phonemes in English.

Let's give each phoneme a symbol

The International Phonetic Alphabet (IPA) was created so that we can easily write phonemes. Each phoneme that occurs in a language in the world is given a symbol. And each symbol is used only once so there is no confusion!

	monophthongs				diphthongs]	F
OWELS	i:	I	ប	u:	IÐ	еі		a
	sh <u>ee</u> p	sh <u>i</u> p	<u>goo</u> d	sh <u>oo</u> t	h <u>ere</u>	w <u>ai</u> t		р
	е	ə	3:	:c	បə	JI	ວບ	İa
>	b <u>e</u> d	teach <u>er</u>	b <u>ir</u> d	d <u>oor</u>	tourist	b <u>oy</u>	sh <u>ow</u>	S
	æ	۸	a:	a	eə	аі	aช	e
	c <u>a</u> t	<u>u</u> p	f <u>ar</u>	<u>o</u> n	h <u>air</u>	my	c <u>ow</u>	
	р	b	t	d	ťſ	dӡ	k	g
UTS	pea	boat	<u>t</u> ea	dog	<u>ch</u> eese	June	car	go
ONAN	f	v	θ	ð	S	Z	ſ	3
SNC	<u>f</u> ly	video	<u>th</u> ink	<u>th</u> is	see	<u>z</u> 00	<u>sh</u> all	television
S	m	n	ŋ	h	1	r	W	j
	man	now	sing	hat	love	red	wet	yes

These are the IPA symbols for English phonemes. But there are many more symbols for phonemes that occur in other languages. The IPA has a symbol for every phoneme in every language (over 300).

> You **do not have to memorize** the IPA for this class! But you should know what it is, and why it exists.

The 44 phonemes of Received Pronunciation based on the popular Adrian Underhill layout

adapted by EnglishClub.com

Why don't we just use the (Roman) alphabet?

A B C D E F G H I J K L M N O P Q R S T U V W X Y Z

The (Roman) alphabet isn't a very good system for naming phonemes. Letters often represent more than one speech sound.



For our scientific investigation, we need a one-to-one relationship between symbols (letters) and phonemes.

A classic joke...

The many-to-many relationship between letters and phonemes has led to really dorky jokes like this:

How do you pronounce this word in English?



GHOTI F I SH enough \rightarrow f women → i nation \rightarrow sh

The next step

Now that we have a process for identifying all of the phonemes in a language, and we have a set of symbols for writing each one down, we can ask the next big question:

What are the **physical properties** of each phoneme that makes them distinct from each other?

	monophthongs				diphthongs			
OWELS	i:	I	ឋ	u:	IƏ	еі		
	sh <u>ee</u> p	sh <u>i</u> p	<u>goo</u> d	sh <u>oo</u> t	h <u>ere</u>	w <u>ai</u> t		
	е	ə	3:	o:	ບອ	IC	ວບ	
>	b <u>e</u> d	teach <u>er</u>	b <u>ir</u> d	d <u>oor</u>	t <u>ou</u> rist	b <u>oy</u>	sh <u>ow</u>	
	æ	۸	a:	a	eə	аі	aช	
	c <u>a</u> t	<u>u</u> p	f <u>ar</u>	<u>o</u> n	h <u>air</u>	my	C <u>OW</u>	
	р	b	t	d	ťſ	dӡ	k	g
VTS	pea	<u>b</u> oat	tea	dog	<u>ch</u> eese	June	car	go
SONAI	f	V	θ	ð	S	Z	ſ	3
NO	fly	video	<u>th</u> ink	<u>th</u> is	see	<u>z</u> 00	<u>sh</u> all	television
CO	m	n	ŋ	h	1	r	W	j
	man	now	si <u>ng</u>	<u>h</u> at	love	red	wet	yes

The 44 phonemes of Received Pronunciation based on the popular Adrian Underhill layout

adapted by EnglishClub.com

What are the physical properties of the phonemes of English (and other languages)?

Sound is a distortion in air pressure

Sound is a wave that travels through air. This means that a sound wave is a disturbance in **air pressure**, or how closely packed the air molecules are.



Physics note: The compression-rarefaction cycle occurs because gas molecules try to fill the volume that they are in. When you push them together, they then try to spread apart:

<u> https://</u>

www.livescience.com/ 53304-gases.html

Sound travels in waves

Everybody knows that sound is a wave, but what exactly does that mean? The first thing to realize is that there are two types of waves:



The pdf version of this slide won't show the animation, so you can use this link to see the motion of transverse and longitudinal waves:

http://gifsoup.com/view/3529701/longitudinal-waves.html

Properties of waves

Waves have several properties. Here are two that have an impact on the way we experience sound, so you might think that they are relevant to phonemes:

- Amplitude is a measure of the force applied to an area of air during compression. The perceptual effect of amplitude is a change in loudness. High amplitude sounds are perceived to be louder, low amplitude sounds are less loud.
- 2. **Frequency** is a measure of the **number of compression cycles** that a wave completes in a given unit of time. The perceptual effect of frequency is a change in pitch (or tone). High frequency sounds have high pitches, low frequency sounds have low pitches.

Physics note about frequency: The sounds that we care about here involve multiple compression cycles (not just one) because speech sounds have a fairly long duration (so they need multiple cycles).

Can you think of a way to test whether amplitude and/or frequency are important to phonemes?

Is amplitude important to phonemes?

Here is a simple experiment to determine if **amplitude** is critical to the difference between phonemes.

Step 1: say "ah"

Step 2: say "ah" with high amplitude

Step 3: say "ah" with low amplitude

Remember, amplitude is a measure of the size of the distortion -- it is the force applied to the air to cause the disturbance

Question: Did varying the amplitude result in a different phoneme? (e.g., 'ee')

Alternative experiment: say `ah' and `ee' with the same amplitude...

Conclusion: Varying the amplitude does not result in changes in the phonemes, only changes in **loudness**, so amplitude is not critical to the difference between phonemes.

Is frequency important to phonemes?

Here is a simple experiment to determine if **frequency** is critical to the difference between phonemes.

Step 1: say "ah"

Step 2: say "ah" with high frequency

Step 3: say "ah" with low frequency

Remember, frequency is a measure of the number of cycles the wave completes in a given time. You might know it as the pitch of the sound.

Question: Did varying the frequency result in a different phoneme? (e.g., 'ee')

Alternative experiment: say 'ah' and 'ee' with the same frequency...

Conclusion: Varying the frequency does not result in changes in the phonemes, only changes in **pitch**, so frequency is not critical to the difference between phonemes.

Properties of your voice

OK, so that was a bust. But it turns out that there are more complicated properties of your voice that do seem to matter for phonemes. But to see them, we must look closer at how your voice works.

There are two components of your voice: your vocal folds and your vocal tract.

Your vocal tract acts as a **filter** to the sound created by your vocal folds. The shape of your oral cavity and pharynx directly affect the properties of the sound.

The **source** of your sound comes from your vocal folds.

Here is a video of the vocal folds of four people singing!

http://www.youtube.com/watch?v=-XGds2GAvGQ



The frequencies of your voice

1. The first important property is that your vocal folds create a large set of frequencies simultaneously (thanks to the complex physics of vibrating objects, which we won't go into here).



We can represent this with a graph like this: frequencies are on the x-axis and the amplitude of the frequencies is on the y-axis.

FundamentalThe lowest frequency generated by a sound source. WeFrequency:use the abbreviation F0 for the fundamental frequency.For your voice, this is the basic pitch that you hear.

Harmonics: The additional frequencies that are created by the source. There is one harmonic at each integer multiple of the F0. We label them H2, H3, H4, etc. Each higher harmonic is weaker in amplitude, hence the decreasing amplitude in the graph above.

The frequencies of your voice

1. The first important property is that your vocal folds create a large set of frequencies simultaneously (thanks to the complex physics of vibrating objects, which we won't go into here).



We can represent this with a graph like this: frequencies are on the x-axis and the amplitude of the frequencies is on the y-axis.

These graphs of frequencies might look strange, but you have seen them before in **music equalizers**. Equalizers let you set the amplitude of different frequencies, so you can get "more bass" or "more treble". In this case, your vocal folds come with a default EQ setting like the one above.



Quick physics aside (not on the exam)

For this class, you just need to know that your vocal folds create a fundamental frequency and a series of harmonics.

But for those of you who want to learn about the physics behind harmonics, you will want to look into the phenomenon of standing waves (particularly in the way that objects like strings vibrate):

<u>http://www.physicsclassroom.com/class/waves/Lesson-4/Harmonics-and-</u> <u>Patterns</u>

http://www.physicsclassroom.com/class/sound/Lesson-4/Fundamental-Frequency-and-Harmonics

But you don't need to know that here. This is not a physics class. We just need to know that your vocal folds produce both a fundamental frequency and a series of harmonic frequencies (in decreasing amplitude) so that we can figure out how speech sounds work!

The filtering by your vocal tract

2. The second important fact is that **the shape of your vocal tract** changes the amplitude of the frequencies created by your vocal folds. Some frequencies are increased, others decreased.



In essence, we can change the EQ settings by changing the shape of our vocal tracts. When we shape our vocal tracts to make an "ah" sound, we get a very specify EQ setting (shown above)



Quick physics aside (not on the exam)

For this class, you just need to know that your vocal tract increases the amplitude of some frequencies, and decreases the amplitude of others. This process is called **filtering**.

But for those of you who want to learn about the physics behind filtering, you will want to look into the phenomena of constructive and destructive interference (particularly for tubes).

This link from Khan academy covers both standing waves on strings and constructive/destructive interference:

https://www.khanacademy.org/science/physics/mechanical-waves-and-sound/ standing-waves/v/wave-interference-pulses

And here is the full unit on wave physics:

https://www.khanacademy.org/science/physics/mechanical-waves-and-sound

But you don't need to know that here. This is not a physics class. We just need to know that your vocal tract filters the frequencies created by your vocal folds.

So to create an "ah", we do this

source: vocal folds

filter: vocal tract "ah"



When we shape our oral cavity and pharynx to make an "ah" sound, we get a very specify EQ setting (shown above)



And to create an "oo", we do this

And if we shape our oral cavity and pharynx differently to make an "oo" sound, we get a different EQ setting (shown below)!



source: vocal folds

filter: vocal tract "oo"



The difference between phonemes

The idea is that the difference between phonemes is a difference in the pattern of frequencies that are created by the filtering properties of the vocal tract (the different "EQ patterns"):



source: vocal folds

filter: vocal tract "oo"



And here is a test to prove it!

In this demonstration, the frequencies created by the source stay the same each time (the same duck call). But the source is placed inside of different filters (the plastic tubes), which changes the frequency pattern. Listen to the result!





eh





source: duck call

http://www.exploratorium.edu/ exhibits/vocal_vowels/ vocal_vowels.html



OK, let's learn some new terms

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.).

Here are the formants for two phonemes

source: vocal folds

filter: vocal tract "ah"



source: vocal folds

filter: vocal tract "oo"



This is a great example of "structure" in the mind!

When you hear "ah", you think it is a single sound. In cognitive science, we say that you **perceive it as a single percept** (where percept just means "thing that is perceived").

filter: vocal tract "ah"



But physically, that phoneme is really a combination of three frequencies, the three formants, put together.

The really cool thing is that **you can't hear the three frequencies**, no matter what you try. You just can't. That is because the mind is structured to perceive speech as a single percept! filter: vocal tract "oo"



But you can hear formants in music!

Even though we can't hear formants in speech, it is possible to hear them in music. They are called overtones in music. Overtones are always present in music - they are the reason two instruments sound different!

Jaw Harps are a metal reed that you pluck in between your teeth in order to make a tone. Although the fundamental frequency never changes (the reed is always the same length), its vibrations resonate in your mouth giving rise to different formants based on the shape of your mouth:

www.youtube.com/watch?v=yx0nnZZVnd8

www.youtube.com/watch?v=VDnio2axqNI

Overtone signing (overtones are what musicians call formants) is a singing technique in which performers change the shape of their vocal tract to create different formants. In effect this creates two notes at once: the low fundamental frequency and a higher frequency formant. In many traditions, the low frequency fundamental is kept relatively constant, such that the melody is actually carried by the higher frequency formant! Listen for the high-pitched melody in these clips:

www.youtube.com/watch?v=7zZainT9v6Qd

And you can see that speech is just tones

This person took songs that include lyrics, and fed them into a midi-synthesizer that synthesizes piano sounds. What this means is that everything is being played on the piano, even the singing part.

The fact that speech is just a series of tones means that even after converting it to piano music, you can still (mostly) understand the lyrics (and in particular the vowels):

https://www.youtube.com/watch?v=ZY6h3pKqYI0

Some Conclusions

The Language Faculty is a mental system that converts a physical signal (sound or sign) into meaning, and vice versa.

The Language Faculty makes several intermediate step between sound and meaning.

Phonemes are the smallest segments of speech that lead to a meaningful difference between words is.

The properties of the vocal folds (the source), such as frequency and amplitude, are not critical to speech sounds.

It is the properties of the vocal tract (the filter) that are critical.

Formants are the highest amplitude peaks in the frequency spectrum created by the human vocal tract. Different formants lead to different speech sounds.

Formants are great example of **structure in the human mind**. During speech, you can't hear the distinct formants. Instead, you perceive speech sounds as a single object (a single percept).