LING 2005 Week 2 **Preliminaries to Acoustic Phonetics**

Speech has a dual identity: on the one hand as an acoustic signal that impinges on the auditory system, and on the other, as a complex orchestration of gestures produced by the articulatory and vocal system. Consequently we need to understand both aspects. Here we lay the foundation for understanding speech as an acoustic/auditory phenomenon.

1. The nature of sound.
Sounds consist of travelling waves of *pressure* fluctuations in a compressible medium - air. If we could visualize these travelling waves they would consist of regions where the air molecules are more or less densely packed, creating fronts and troughs, propagated outward from some source as in figure 1 below:

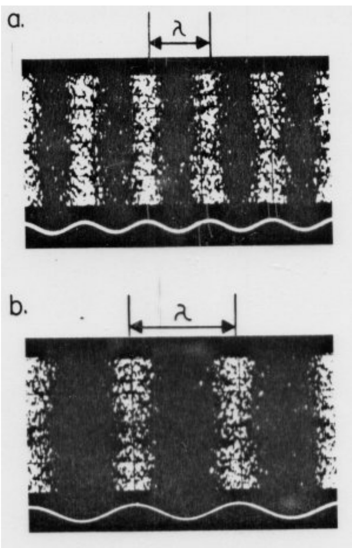


Fig.1 Pressure wave

The fronts of high pressure, where the air molecules are densely packed together show up as light regions (as if each molecule of the medium reflected a point of light under our 'molecular viewer').

The graph at the bottom of each wave shows the fluctuations in the density of the air molecules at different regions in the medium.

The term for that describes these density fluctuations in the medium (air in this case) is ___________.

These are travelling wavefronts. Lets assume they are moving in a left-right direction. The distance between each pressure front, represented by the symbol *lambda* is known as the **wavelength**. of the travelling wave. Suggest an appropriate unit for measuring the wavelength of a wave: [Answer:______________________].

We experience very slow moving wavefronts involving pressure variations in the atmosphere as variations in the _______. Forecasters report these fluctuations in what unit of measurement? [Answer: ________]. From our perspective, the only relevant difference between atmospheric pressure waves that carry the weather and sound waves is that the latter are *audible*.

Think for a bit... What are the physical differences between audible sound waves and inaudible atmospheric pressure fluctuations which we experience as 'the weather'? Make reference to properties such as *wavelength* and...

____

____

____

2. Measuring Pressure
Go to Hyper-physics. (http://hyperphysics.phy-astr.gsu.edu/hbase/hframe.html) Scroll down the Index on the right hand side of the web page until you find Pressure. Read up on *pressure* and *pressure calculation*.

Define Pressure. Give the equation for relating *pressure*, *force*, and *area*

Pressure is ____

Equation for Pressure:_______________________________

3. Why do we need a medium for sound waves?

____
____
____
____
____

Electromagnetic energy (light waves, xrays, etc.) constitute their own medium (photons, in the case of light), whereas sound (acoustic energy) relies on the movement of particles that make up the medium (air, water...).

4. Sound propagation

The vibrating molecules of air transmit sound energy by bouncing off one another - in a nice coordinated fashion, like a string of 'line dancers'. Less picturesquely, you may think of the medium as made up of a lattice of particles each of small mass, connected to one another by springs. Set one row of particles in motion and they will transmit that vibration to their neighbours.

Now go to Hyper-physics and check out *Sound Propagation*, then find *Longitudinal waves*. Find out the distinction between *longitudinal* and *transverse* waves in order to understand more precisely how sound waves are transmitted in a medium.

Explain the distinction between longitudinal and transverse wave transmission. Which is applicable in the case of sound waves?

____
____
____
____

5. Sounds and their Waveforms

Sound waves are pressure fluctuation in time and space. In fig 1. we pictured a pressure wave as a pressure fluctuation in space. But it is more common to measure and view sounds as pressure fluctuation in time. This kind of visual display is known as a *time series waveform* or *waveform* for short.

Lets access some sounds and their waveforms (http://emsah.uq.edu.au/linguistics/teaching/ling2005/waves.html) . You should be able to scroll down over the sounds 'ding', 'chord', 'sine', 'buzz', 'noise', and '   '. By clicking on each wave you should be able to hear it (provided the sound is working on your computer and/or you have headphones!). For the longer sounds, 'sine', 'buzz', 'noise' the box on the right represents an expanded (zoom-in) time scale display so that you can see the precise shape of each waveform.

Listen to each of the sounds and think about how you could describe the auditory differences between them. How might you rate them differentially on some descriptive scales that apply to the subjective experience of sounds and sound quality? Lets see what you can come up with. Focus on describing how the sounds 'sine', 'buzz' and 'noise' differ from each other :

____
____
____
____
____

We seem to lack a ready vocabulary for accurately describing subjective differences in sound quality. Musicians may have more terms for this than non-musicians. But everyday language for describing sounds seems strangely impoverished, compared for example, with colour. This is one reason why phoneticians describe speech sounds in articulatory terms. Try conveying to another phonetician that you are referring to a *voiced bilabial stop* [b] not to a *voiceless palatal fricative* [s] without the benefit of referring either to the pronunciation of the sounds or their transcribed forms. ('A short muffled sound', versus 'a longer hissy one' just does not do it.)

Nevertheless, it is possible to develop purely auditoy/acoustic feature systems for describing speech sounds. See the distinctive feature system devised by Jakobson, Fant and Halle (1952) Preliminaries to Speech Analysis MIT Press. We shall have more to say about phonetic features later. We now turn to ways of describing and classifying sounds as physical signals.

6. Types of signals (waves)

Sound waves can be regarded as signals when, as in speech, they carry information for the hearer/listener. Our auditory systems, which are fundamentally the same in all mammals, are responsive to different types of signals, so a basic and important question for acoustic phonetics is how should acoustic signals be classified? Notice that this is not the same question as asked above: How can we describe differences in subjective sound quality between different kinds of sounds? We are talking here about physical properties of acoustic signals; properties that can be measured instrumentally.

Nevertheless, any physical classification of signals should correlate with subjective perceptions of differences between sounds, provided that our classification is well founded on fundamental properties that distinguish signals in the environment and which our auditory system is well adapted to detect. There are three such basic properties or dimensions by which sounds as acoustic signals may be classified: periodicity, duration and complexity.

Periodic - Aperiodic

A fundamental distinction in the world of signals is between waves that repeat themselves regularly *(periodic*) and those which do not (*aperiodic*). We have a half-way house in this classification (*quasi-periodic*) for sounds that are not wholly repeating; sounds that have an overlay of noise, or that are almost, but not quite regular. Clearly a sound that cannot be sustained, like a 'pop' or a 'chime' cannot be regarded as periodic and there are signals that are difficult to classify - like a series of 'chimes' that occur at more or less regular intervals of time.

Listen again to the following sounds, examine their waveforms, and see why they have been classified as *periodic*, *aperiodic*, and *quasi-periodic* (http://emsah.uq.edu.au/linguistics/teaching/ling2005/periodicity.html).

Questions:

Only one of the stimuli, ‘chord’ was given a label. Supply an appropriate descriptive label for each of the other sounds.

1.___________	2. ____________	3._____________
4. __________	5. ‘chord’	

One of these sounds gives the auditory impression of being not one signal but two that have been overlaid. Which sound is that? Your labelling of the sound may reflect its composite nature. ____

Examine the waveform of the ‘composite’ sound and justify why it has been classified the way that it has.

____
____
____

An interesting, and profoundly difficult question to answer in many cases, is why did the ‘composite’ signal sounds like two signals overlaid on one another and not a single sound that has a mixture of periodic and aperiodic qualities. Why does the ear separate this stimulus into two distinct auditory objects? (Or do you not agree with this perceptual judgement?) Raise this for discussion in class. It bears on the fundamental problem of how the auditory system separates sound objects in the environment from other objects in the auditory field and both from ‘background noise’. In general terms, this is known as the figure - ground problem of perception. But let’s return to the easier problem of classifying sounds.

Continuous vs. Transient signals

Check out the contrast between continuous and transient signals (http://emsah.uq.edu.au/linguistics/teaching/ling2005/transient.html).

Explain the contrast in your own words.

____
____

The distinction is pretty obvious. But there is a question of just what ‘transient’ means. Life is ‘transient’. But a speech sound that lasts for more than half a second certainly would not be classified as ‘transient’.

Classify the sounds underlined in the following words as ‘transient’ or ‘continuous’ from a phonetic point of view. Supply the appropriate phonetic symbol.

cat smile choice mule

transient	continuous

Simple (sinusoidal) vs Complex signals

Listen to the three signals (http://emsah.uq.edu.au/linguistics/teaching/ling2005/complex.html). Try to decide why each has been classified as simple or complex. (It is not obvious from the waveform.) Supply labels for the three sounds in the box below.

Simple	Complex
1. ______________	2. _______________
	3. _______________

Those of you who know a little about musical instruments will be able to distinguish between simple and complex sounds. Complex sounds have 'overtones' or 'harmonics'. Simple tones lack harmonics. But this does not indicate why 'noise' signals are classified as complex. A complex signal (sound) is one that contains more than one *frequency component* - a notion that we will make more precise shortly. A simple signal - the building block of all complex signals - is made up of a single frequency component.

Summarizing: We have seen that there are three fundamental ways of classifying acoustic signals. These ways of classifying sounds are important because they reflect basic physical properties of signals to which all auditory systems are sensitive. These three properties may occur in combination, though not all combinations are physically possible.

Try to come up with an examples of a familiar types of sounds to illustrate as many of the combinations of these three properties as you can find. One has been done for you, to help you get started.

Example sound	periodic / aperiodic	transient / continuous	simple / complex
pitch pipe	periodic	continuous	simple

7. Properties of Waves

We now turn to consider the physical properties of waves/signals that enable us to describe them with the precision of mathematical functions. While we have no pressing need for describing speech at this level of precision, an understanding of these properties is essential for understanding the acoustic structure of the speech signal and for making effective use of the analytical methods of speech technology.

We begin by examining a revealing diagram of how the waveform of a simple signal (http://emsah.uq.edu.au/linguistics/teaching/ling2005/pic/waveform.jpg) may be generated and recorded. A mass, set in motion continues to oscillate on a near frictionless surface. Its oscillation is recorded on a moving drum by some mechanism that records deflections on the 'y-axis' over time, the trace of which is determined jointly by the rate of oscillation and the speed of the recording drum. A diagram like this serves to equally to represent Thomas Edison's original recording device as well as the PCM (pulse code modulation) representation of signals captured in .wav files on your personal computer. The main difference is that in the latter case, the wave is not a continuous (analog) function recorded on paper or on the screen of an oscilloscope, but is a series of discrete numerical values stored in a data file.

Amplitude and Period

We define the (maximum) *amplitude* of the vibration by its (maximum) displacement from the equilibrium position on the y-axis (ie. the position at which the vibration will eventually come to rest). We note that it takes .5 seconds for a complete cycle of vibration of the spring-attached weight. This quantity is known as the *period* of the vibration. It is recorded on the x-axis (time). The *period* or time for a single cycle of repetitive vibration is traditionally represented as capital T.

You may wonder where the term 'cycle' of vibration comes from. After all, the mass on the spring is oscillating up and down on a plane. The thing to appreciate however is that the mass on the spring is decelerating as it moves towards its position of maximum amplitude (A) from the equilibrium point, but it is accelerating as it moves from A towards the equilibrium point (B). The forces generated by the spring account for this behaviour, executing maximum drag when M is closest to A (or push when it is closest to C). This deceleration towards A and acceleration towards B (the equilibrium point) is precisely what one observes if uniform circular motion is projected on a plane. (Imagine the shadow of a moth circling point of light and having its shadow projected on wall by some more distant and powerful light source. The planar component (projected on the wall) of the moth's circular motion precisely describes the pattern of acceleration and deceleration of the mass on the spring towards and away from the point of maximum amplitude of vibration. Uniform circular motion is described mathematically by a sine (or cosine) function. Hence the mass-on-a-spring oscillation, when projected in moving time base can be described by a sine or cosine function. Check this out, all in nice clear diagrammatic form in hyper-physics. Scroll down the right hand scroll bar untill you find *Simple Harmonic Motion*. Explore!

What is the importance of SHM (simple harmonic motion), and do I have to learn that damn equation? SHM is the building block for understanding the composition of any complex wave. Fourier's famous theorem states that any complex wave may regarded as the sum of a number, possibly infinite, of sinusoidal (sine) waves. *Fourier analysis* is a mathematical technique for decomposing a complex wave into its component sine waves. The auditory system, as we shall see, is a *frequency* analyser. There is a computationally efficient implementation of Fourier analysis, known as the Fast Fourier Transform (FFT) which is employed in most speech analysis programs to simulate the kind of frequency analysis that we assume takes place on signals in the auditory system and which is the basis of the *digital* *spectrogram*. (More on that later.) But no, you do not *have* to learn the equation of SHM. However, it is so fundamental that it is worthwhile unpacking the components of the equation, and by the time you do that, you will probably find that it has 'stuck' in your memory anyway.

Frequency

We have previously defined the quantities Amplitude (A) and Period (T) of a wave. Just to make sure, give your own definition of each of these terms here:

Amplitude:

____

Period:

____

How is *frequency* (f) defined? Also, find out from hyper-physics the mathematical relationship between the *frequency* and the *period* of a wave.

Frequency:

____

Mathematical relationship of *f* and *T*:

____

The frequency of sound waves is measured in cycles per second, known as *Hertz* (Hz) or in thousands of cycles per second, kilo-Hertz (*kHz*). The subjective sensation of the ***pitch*** of a sound is what correlates most highly with the physical property of *frequency*.

Using hyper-physics, starting from the 'pitch' node, search out the answer to the following questions:

What is the frequency of the note 'middle C' on the piano?
______ Hz.

By traversing from the 'Pitch' node to 'Hearing concepts' and the node 'Great sensitivity' you should be able to discover the range of frequencies to which the human ear is sensitive:
____________________ Hz to kHz.

Along the way, you should discover lots of other interesting facts which will be relevant for understanding the perception of speech sounds and topics that we shall explore later in the course.

8. Summarizing and moving on

We have identified three fundamental ways of classifying signals (remind yourself: 1. _______________________, 2. ____________________, 3. ______________________) and three physical properties of signals (**A** = _________________, **T** = ________________ and **f** = _____________________), and we have looked at a special type of wave motion: **SHM** = _________________________________, the building block of complex vibrations. But how do these physical properties of sounds relate to sound as a subjective perceptual sensation and how do they relate to our perception of speech sounds in particular? Consider first the general question of sounds as subjective (auditory) sensations and their corresponding physical (acoustic) properties. It is generally agreed that there are three fundamental auditory dimensions of sound. You can by now probably suggest their acoustic correlates (corresponding physical properties), at least for one or two. You may be able to go the next step further and suggest appropriate units of measurement for one or two of the acoustic properties. However, this is something we need to work towards - ways of measuring the acoustic properties of sounds, so your answers here may be quite provisional at this stage.

Auditory dimension		Acoustic Property	Unit of measurement
Pitch	low - high		
Loudness	soft - loud		
Quality/Timbre			

We shall also consider in a provisional way how the physical classification of signals types outlined previously relates to the *perceptual* classification of speech sounds, or at least some sounds of speech - a broad sampling of the English consonants. But before we begin this exercise, a word of caution. There is no strong reason for assuming at the outset that the properties which distinguish speech sounds, so called phonetic features, are closely related to a fundamental acoustic classification of sounds, or types of signal. In fact, as we shall see, some researchers have argued that speech perception is a special case of auditory perception and that the phonetic features which characterise speech sounds simply do not apply to auditory signals in general. But this is at bottom an empirical question and one that is still the subject of much debate.

9. Mapping perceptual similarities among sounds

Suppose I select at random some triplets of consonants from the set of English consonants. I place the consonants in a [Ca] syllable frame and ask you to make the following perceptual judgement of each triplet: Which member of the triplet sounds least like the other two? (i.e.: which sound is ‘the odd man out’?).

Now further suppose that I present you with all the possible combinations of triplets that can be made by selecting 3 at a time from the set of English consonants.

Triplet	your judgement	majority judgement
pa - sa - ba	__	[s]
ta - da - ka	__	[d]

By collecting all these ‘triadic comparisons’ together from not just one but a fair sample of listeners it is possible to construct a *similarity matrix* which shows how similar each consonant is to every other consonant.

The following similarity matrix was collected for 12 English consonants. Only a representative subset of the 16 consonant phonemes of English was used in this experiment due to the large number of triplets required to systematically map the similarity of each consonant with every other consonant. However, there are enough sounds included here to broadly capture the ‘auditory space’ of consonant sound contrasts.

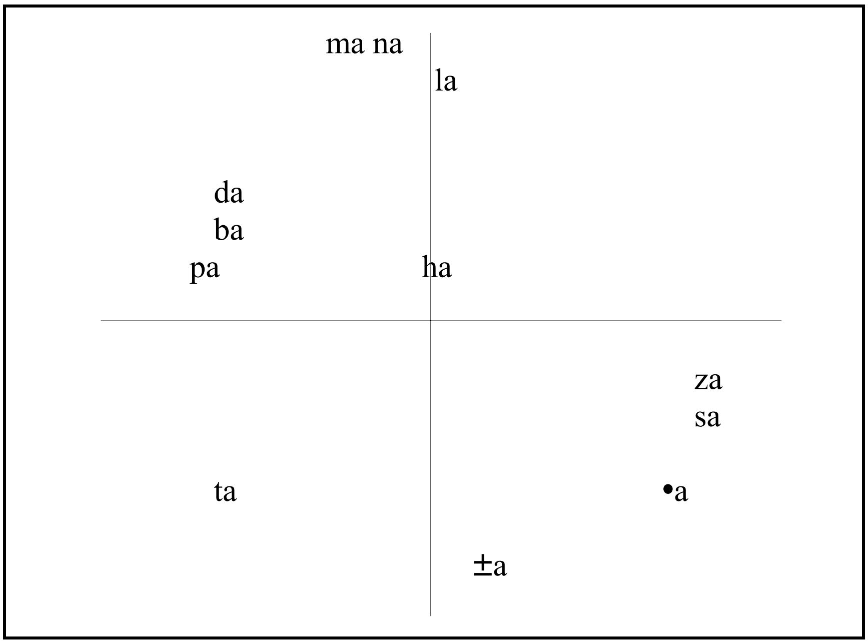
Similarity matrix: English consonants												
	p	m	±	b	•	t	n	z	l	h	s	d
p												
m	22											
±	35	43										
b	07	15	33									
•	43	46	11	40								
t	19	35	20	20	30							
n	25	03	41	17	42	35						
z	42	40	19	37	09	34	38					
l	27	09	39	19	40	35	04	32				
h	17	20	25	13	28	24	15	24	14			
s	41	43	16	38	05	32	40	04	36	27		
d	08	11	36	03	41	24	13	40	39	15	40	

The similarity matrix contains a good deal of information about how perceptually alike or different certain consonants are to each other.

However it is difficult to visualize the basis for listener’s similarity judgements by examining the matrix and its similarity values.

So, we employed a mapping technique called ‘multidimensional scaling’ which treats the similarity ratings as if they were (systematically related to) distances in a perceptual space.

It turned out (the details need not detain us) that a good mapping of the similarity judgements could be obtained by treating the sounds as points in a two dimensional space, where the distance between any two sounds in that space reflects their similarity to each other. Consider how the sounds are distributed in this two dimensional space.



From the way that the consonant sounds distribute themselves in ‘perceptual space’ can you make any observations about the acoustic properties that listeners used in making their similarity judgements? Look at how the sounds line up on each of the two axes, which we might try to interpret as ‘auditory perceptual dimensions’. From your knowledge of how the consonants themselves sound and what your intuitions tell you about their acoustic properties, can you discern any relationship with the fundamental acoustic classification of signals that we developed in section 6.?

If you find that you can interpret the results of this perceptual mapping of English consonants in terms of the basic acoustic classification of signals given earlier, what implications does this have for the question that we raised earlier - that the perception of speech sounds is in some ways ‘special’ and different from the way that sounds in general are perceived?

Reference:
Ingram J.C. L Perceptual dimensions of phonemic recognition. In Prideaux G. et al. (eds.) *Experimental Linguistics* E.Story-Scientia :Ghent, 1980, 275-291.