Tutorial on Phonetics and Speech Analysis

true   true

Document compiled 12 Nov 2024 00:16

Contents

Preface

Aims

In this tutorial you will learn about **acoustics**, **phonetics**, and **speech analysis**. You will learn the core concepts in these related fields, as well as the necessary practical skills for speech analysis. The aim of this tutorial is to provide you with the phonetic insights and skills in speech analysis that you need to succefully conduct phonetic research in your own project (e.g. paper or thesis).

Under construction

This tutorial is a work in progress, resulting from an ongoing revision of an existing tutorial, and meanwhile incorporating other modules and resources.

The existing (outdated) full tutorial is still available at https://resources.lab. hum.uu.nl/resources/phonetics/index.html .

More details on the origins of this tutorial are provided below.

How to use this tutorial

You will learn the most from this tutorial if you

(1) read the explanatory texts in this tutorial,
(2) work through the questions and exercises provided,
(3) practice in applying your new knowledge hands-on, with the `Praat` computer program (detailed below), and
(4) re-read the relevant sections from this tutorial and your textbook, with the help of keywords provided per section.

Questions

Text blocks such as this one will contain questions or exercises inviting you to engage with the tutorial. You will learn most if you attempt to answer these questions (preferably in writing) *before* you proceed and *before* you take a look at the answer provided. (These questions only work in the HTML version of the tutorial; other versions will just show both the question and answer subsequently.)

Question 0.1

What is sound?

Answer 0.1

Sound is a type of energy that travels through a medium (such as air, water, or solid materials) in the form of waves. These sound waves are created by the vibration of objects, which causes the surrounding particles in the medium to move in a back-and-forth motion. This movement, or vibration, transfers energy through the medium, creating waves of high and low pressure.

Recommended software

In this tutorial you will work mostly with `Praat`[1]. This is a popular open-source program for the analysis of speech, developed by Paul Boersma and David Weenink (both at University of Amsterdam). It can be found on its own website (https://www.praat.org), where you will find a wealth of helpful documentation. `Praat` also has extensive `Help` built in. There is an online forum (https://groups.io/g/Praat-Users-List), where users share their knowledge by posting questions and providing answers.

In order to install `Praat` on your computer, go to its webpage at https://www.praat.org/, and then proceed to the download page for the operating system of your computer. Follow the installation instructions on the download page for your operating system.

Instructions for using `Praat`

Text blocks such as this one will contain instructions about how to "do" things in `Praat`.

Options in software menus, and texts of on-screen buttons, will be shown `in this way`. The notation `Main > Sub` means: first choose option `Main` from the

---

[1]The Dutch word *praat* / praːt/ means "talk".

main menu, after which a submenu will appear, then choose option `Sub` from the submenu. Commands or formulas that you have to type will be shown `in this way` too. (Commands typically need to be terminated with typing `Enter` or `Return` or   or   – which however will not be specified in the instructions.)

Structure of this tutorial

TBA

Recommended textbooks

This tutorial is intended to be used in addition to one or more textbook(s) in Phonetics, to which this tutorial will provide additional background knowledge. Some excellent textbooks in Phonetics are those by Rietveld and van Heuven (2009) (in Dutch), Johnson (2011), Ladefoged and Johnson (2015), Reetz and Jongman (2020), and Zsiga (2024).

Details

License

This work is licensed under the *GNU GPL 3* license (for details see https://www.gnu.org/licenses/gpl-3.0.en.html).

Citation

TBA

Technical details

TBA

History

This work is based on an earlier tutorial (2006-2007) titled Tutorial for self study: basics of phonetics and how to use Praat by Clizia Welker and Hugo Quené. In turn, that 2007 tutorial was partly based on older texts by Hugo Quené, Denise Bruin and Mirjam Wester (1996-2000); these older texts acknowledged valuable

comments and suggestions by Paul Boersma, Olga van Herwijnen, Kim Koppen, Eva Sittig, Joyce Vliegen and Mieke van Wijck.

The 2007 version of the tutorial was subsequently revised and adapted to the current version using `R Markdown` (Xie et al., 2018) and `bookdown` (Xie, 2024) in Rstudio by Hugo Quené in 2024.

Part I: Acoustics

Chapter 1

Sound waves

1.1 Sound

Sound is a type of energy that travels through a medium (such as air, water, or solid materials) in the form of waves. These sound waves are created by the vibration of objects, which causes the surrounding particles in the medium to move in a back-and-forth (oscillatory) motion. This movement, or vibration, or oscillation, transfers energy through the medium, creating waves of high and low pressure.

1.2 Sound wave

A sound wave consists of pressure fluctuations caused by the molecules of the acoustic medium crowding together (compression) and moving apart (rarefaction). A sound wave is spread in all directions from the sound source; we could compare its propagation to that of a circular wave on the surface of a water basin. The molecules themselves move over a very short distance and do not travel along with the wave: instead, after the sound wave (the air pressure fluctuation) has passed along, they go back to their equilibrium position. Unlike wind, which is a movement of air (caused by temperature, pressure or difference in chemical composition), the air particles involved in the propagation of the sound wave do not move along with it.

There are two kinds of waves (also depending on the acoustic medium). In *longitudinal* waves (such as sound waves) the back-and-forth displacement or movement of the medium's particles is in the same direction as the propagation of the wave. In *transverse* waves (such as the waves on the surface of a pond) the back-and-forth displacement of the water particles is perpendicular to the direction of propagation of the wave.

A stadium wave provides a clear example of a transverse wave: a group of persons (the particles) starts the wave by standing up, rising their arms, sitting down, standing up again, and so on. The persons' action is directly followed by that of their neighbours on one side, who do the same and who are again followed by their next neighbours on their side, and so on, until the wave is travelling through the whole stadium. The persons' motion (up-down) is perpendicular to the propagation of the wave (left-right along the bench).

Sound propagates in all dimensions through an acoustic medium, like an expanding sphere, which is indeed the theoretical model used to describe the sound wave propagation pattern. As the sound wave moves away from its source, more particles are involved in the pressure fluctuations. As a consequence, sound waves lose energy while travelling through the medium, as some of the energy is spent in moving increasingly more particles. Finally, sound is perceived as such when the sound wave spread by the sound source and travelling through the acoustic medium finally impinges upon the eardrum of the observer.

1.3   Acoustic media

Air is only one of the media through which sound can propagate. If your head is under water (as in a bath, pool, lake or sea), the water may carry sound waves from the sound source to your eardrums, and you do hear sounds. The propagation of sound waves is faster through liquids than through gases such as air: the closer the molecules of the medium (i.e. the higher its density), the higher the speed of sound in that medium.

You can also put your ear to the ground in order to hear sounds propagated through the soil. The propagation of sound waves in solid soil is even faster than in liquids. Trying this out on dry sand on the beach, one observer noted hearing footsteps until about 25 m distant (Minnaert, 1970, §10).

1.4   The speed of sound

In air, the speed of sound (the speed of propagation of a sound wave) is about 331 m/s at 0°C (about 334 m/s at 21°C). The speed of sound also depends on the humidity: humid air holds more particles (of water), resulting in a slightly higher speed of sound.

In sea water sound travels at about 1435 m/s, in concrete 3400 m/s, in iron railway tracks about 5000 m/s.

1.5   Pressure

Pressure is the amount of force on a surface. In physics, *force* is defined as an influence causing an object to accelerate. It is expressed in Newton units; a Newton is the amount of force that increases the velocity of a 1-kilogram object by one meter per second (m/s). *Pressure*, in turn, is defined as force per unit of area. It is measured in Pascal units, which correspond to Newton (N) per square meter ($1\ Pa = 1\ N/m^2$). Under normal conditions, atmospheric air pressure is centered at about 1013 hPa (101300 Pa, an average value calculated on a medium latitude at sea level, 0°C), with normal meteorological fluctuations of about $\pm$ 5000 Pa. Sound wave fluctuations in air pressure are far smaller, ranging from about $\pm$20 µPa (mikropascal, or $\pm$0.00002 Pa) at the lower threshold of hearing to about $\pm$20 Pascal at the upper threshold of hearing. Even louder sounds, with variations in air pressure exceeding about $\pm$20 Pascal, are painful and cause hearing damage.

1.6   Questions

Question 1.1

Explain why a sound wave loses energy the further it is spread from the oscillation source.

Answer 1.1

The more a sound *wave* moves away from the source, the more particles of the medium (e.g. air) are involved. The amount of initial energy (spread with the source oscillation) is spread over a larger surface, of an expanding imaginary sphere, and consequently the sound wave displaces more particles. The overall amount of energy remains the same. Therefore, the energy on a single medium particle or on a single portion of the sound wave is smaller. Thus, the sound wave fades as the distance to the sound source increases.

Remember that the sound *wave* travels through the medium, but the particles in the medium remain more or less in place.

Question 1.2

What parameters influence the speed of a sound wave?

Answer 1.2

The propagation speed of a sound wave depends on the acoustic medium: different media have a different characteristic sound propagation speed. In gases (like air) the sound speed is lower than in liquids or solids, as the particles are

less close to one another (the medium is less dense). Moreover, factors such as temperature and pressure influence the density of the acoustic medium, and thus these factors indirectly influence the speed of sound in the medium. Heat, for instance, makes the particles less close to one another, and thus slows down the sound speed. For gases such as air, sound speed is determined by their chemical composition.

1.7   How to work with `Praat`

`Praat` is a computer program designed to process, analyze and visualize speech sounds.

After starting the program, `Praat` opens two windows: an *Objects* window (typically on the left) and a *Picture* window (typically on the right).

Objects

In `Praat`, signals and derived representations are all seen as objects. Objects may have different types, e.g. Sound, Spectrum, Pitch, etc[1].
Each type of object comes with pre-defined operations that are possible. If you select an object of a different type, then the buttons (operations) change with the object type. As an analogy, consider the various types of objects in your room: clothing items and human bodies may be washed, food may be cooked but human bodies may not be cooked, furniture and food items may be opened but human bodies may not be opened, clothing can be inside furniture, etc. Moreover, relations between object types are also specified: for example, a plant can become a food item (by means of cooking), but a human body may not.

Before working with an object, you need to select that object, by clicking on it in the list of objects displayed in the Praat Objects window.

Objects of any type may be saved and opened using the `Save` and `Open` options in the top menu of the Objects window. This is a great way to save the objects resulting from your phonetic analyses, i.e., to save your results.

The buttons at the bottom of the Objects window are *always* available for objects of any type: `Rename...`, `Copy...`, `Inspect` (to take a deeper look), `Info`, and `Remove`.

---

[1]By convention, object types are written with a capital; this helps to distinguish physical properties (e.g. the intensity of a sound) from the `Praat` representations of those properties (e.g. the Intensity object computed from a Sound object, both within `Praat`).

We will often work with objects of the *Sound* type. Such a Sound object is a digital sound (sampled audio), which you can Play or Scale or Convert or Combine, etc. You may analyze a Sound, which will typically result in an object of a different type (e.g. Pitch). Sound objects can be opened from disk, and saved as audio files in a wide variety of audio formats.

Picture window

`Praat` will draw its visualisations (figures) in its Praat Picture window. The figure will be scaled to the area with the pink boundary, the so-called viewport. By changing the viewport after drawing a part of a figure, you may obtain multiple visualizations in a single figure, as will be illustrated in this tutorial.

The combined figure in the viewport may be saved (`File > Save`) or printed (`File > Print`) in the top menu of the Praat Picture window.

You may also save the figure in a different way, as a "recipe" set of instructions to re-create the figure (`File > Save as Praat picture file`), for later reuse.

Chapter 2

Converting sound to bytes

2.1 Overview

In order to process sounds by means of a computer program, or telephone, we first need to convert that sound, the variations in air pressure, to numbers that are then further processed by a computer or by a telephone device. This is a two-step process, involving at least two key components in order:

(1) the **microphone**: this device transforms variations in air pressure into matching variations in an electrical signal. The microphone has a thin membrane, and deviations of the membrane (caused by the sound wave hitting the membrane) are transformed into proportional fluctuations in electric current (Ampere), electric voltage (Volt) or electric resistance (Ohm). For more details about how to handle a microphone, see the text box in §2.2) below. The analog electrical signal is then passed on from the microphone to...

(2) the **analog-to-digital-converter** (ADC): this device converts a continuous, analog electrical signal into a stream of discrete, digital numbers. The ADC measures the input signal, and reports the digital output value of that input signal. This process is also called 'sampling'. Sampling a signal is done with a certain 'sampling frequency' (number of measurements per second) and with a certain precision of measurement (known as 'amplitude resolution'), both explained below. The result is an output stream of digital numbers (in bytes), to be handled further by computer software (e.g. to be displayed, compressed, transmitted, stored, played back, etc.)[1] [2]

---

[1]The input signal to be sampled often comes from a microphone, but other signals may also be sampled, e.g. the signal coming from an electro-encephalogram (EEG) electrode.

[2]In a speaker's telephone, the stream of numbers (output from the ADC) constitutes the

Very soon, whenever you want to hear sound from a computer or from a telephone connection, you will also need

(3) a **digital-to-analog-converter** (DAC): this device converts a stream of discrete, digital numbers into a continuous analog electrical signal, with a pre-specified conversion frequency and amplitude precision. The result is an output analog electrical signal, to be handled further by audio hardware (e.g. to be amplified, sent to a loudspeaker, etc.)

2.2   How to handle a microphone

- A good microphone is a very sensitive and very expensive device. Treat it with great care. Never blow into a microphone (it's far better to just say `test` or `check` or to say anything else). Do not tap on its surface.

- Do not plug or unplug the microphone into/from a "hot" port (first set the port's input/output volume to zero, then plug/unplug).

- Do not speak *into* the microphone, but just over it or alongside. The microphone should measure sounds, but *not* the flow of air coming out of a speaker's mouth and nose. If the microphone comes with a foam cap to dampen airflow, then use it.

- Do not touch the microphone while it is working; this will result in undesired (and often loud) contact sounds in the output signal.

---

input for subsequent processing and data compression, even before speech data are transmitted to the receiving phone.

Bibliography

Johnson, K. (2011). *Acoustic and Auditory Phonetics*. Blackwell, Malden, MA, 3rd edition.

Ladefoged, P. and Johnson, K. (2015). *A Course in Phonetics*. 7th edition.

Minnaert, M. (1970). *De natuurkunde van 't vrije veld*, volume 2. Thieme, Zutphen.

Reetz, H. and Jongman, A. (2020). *Phonetics: Transcription, Production, Acoustics, and Perception*. Wiley-Blackwell, Chichester, 2nd edition.

Rietveld, A. and van Heuven, V. (2009). *Algemene Fonetiek*. Coutinho, Bussum, 3rd edition.

Xie, Y. (2024). *bookdown: Authoring Books and Technical Documents with R Markdown*. R package version 0.41.

Xie, Y., Allaire, J., and Grolemund, G. (2018). *R Markdown: The Definitive Guide*. Chapman and Hall/CRC, Boca Raton, Florida.

Zsiga, E. C. (2024). *The sounds of language: An introduction to phonetics and phonology*. Wiley-Blackwell, Chichester, 2nd edition.