A practical introduction to acoustic phonetics
day 2: Introduction to digital sound and PRAAT

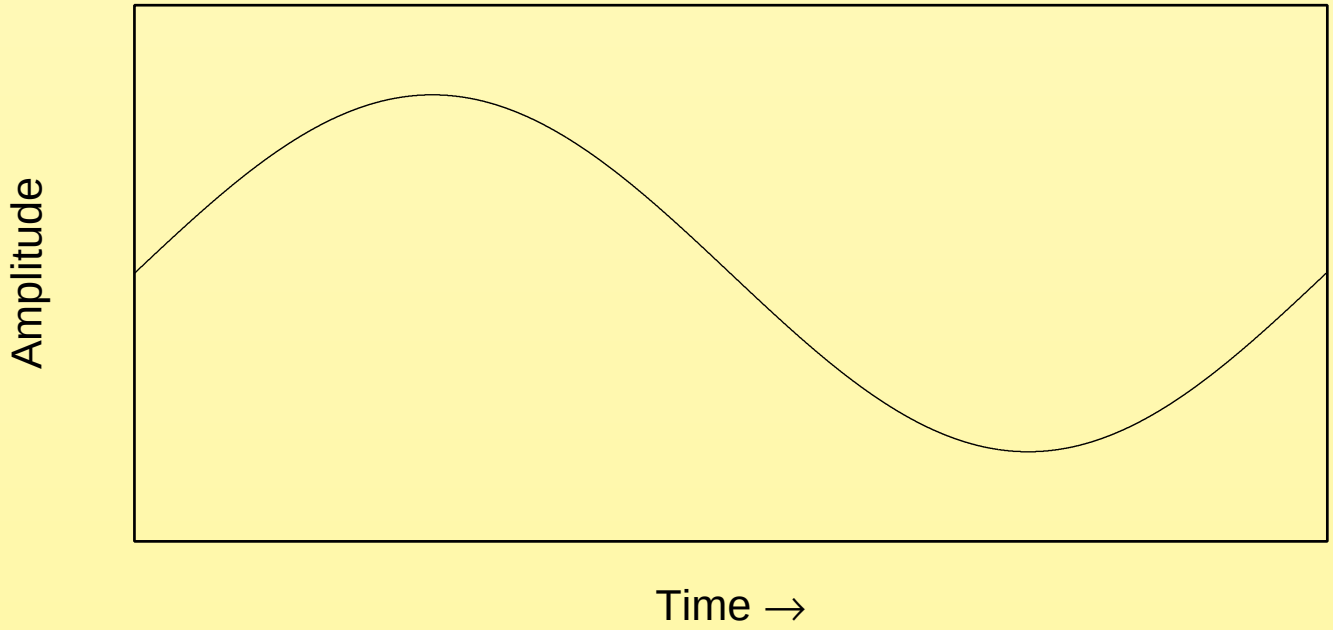
Wouter Jansen
Department of Linguistics
University College London
wouter.jansen@kuvik.net
http://wouter.jansen.kuvik.net/teaching

January 18, 2005

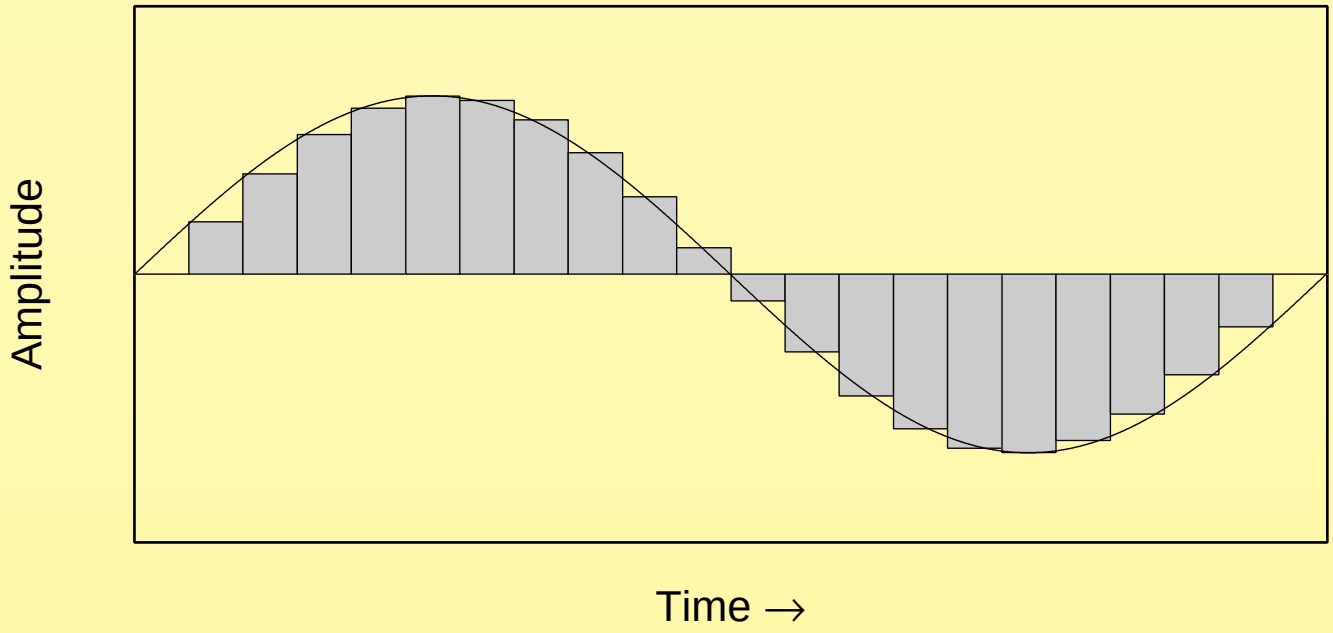
Analog vs. digital signals

- Devices such as taperecorders and grammophone records store sound signals in an analog way: for every possible time t you can find an amplitude D
- Devices such as CDs, DAT tapes, Minidiscs, and computer hard drives store sound digitally: the signal is represented as a sequence of discrete time slices, each of which is assigned a single amplitude value from some finite range of possible values

Analog



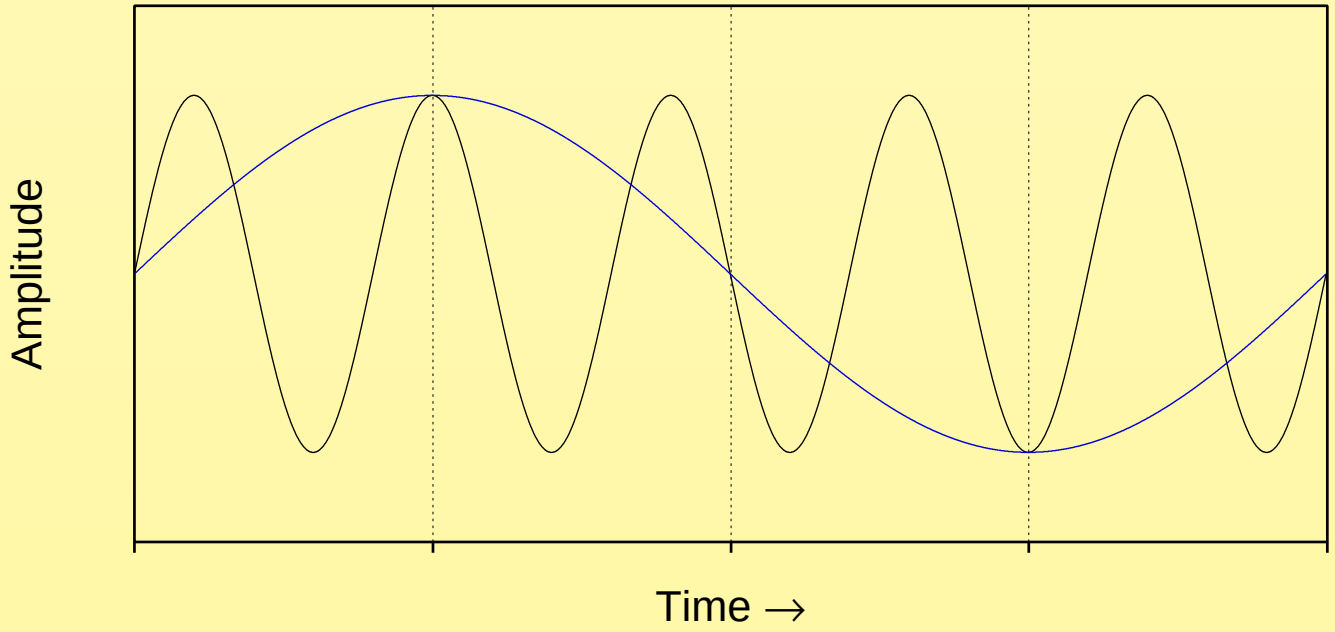
Digital



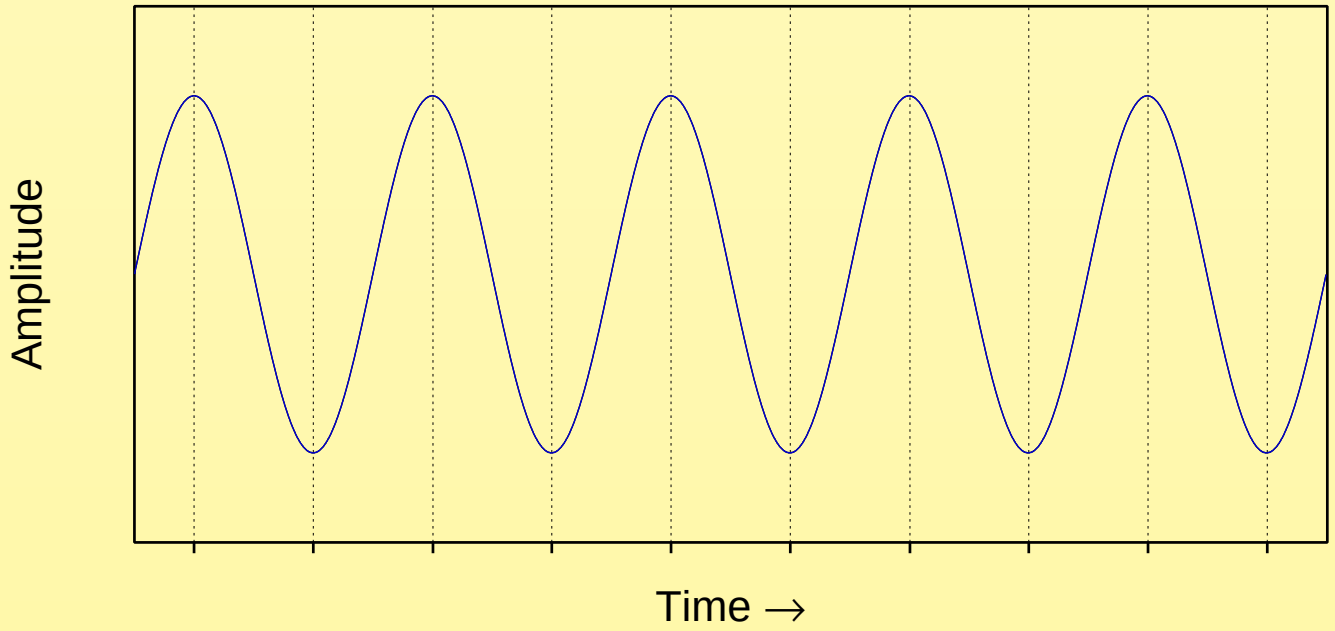
Sampling

- Sampling frequency denotes the number of time slices per unit time in the digital signal
- This quantity is usually expressed in time slices per second (Hz)
- The highest signal frequency that can be reliably encoded is half the sampling frequency
- This frequency cut-off point is commonly known as the Nyquist frequency

Frequency components $>$ Nyquist: aliasing



Frequency components = Nyquist



Quantisation

- Quantisation refers to the number of amplitude slices or 'bins' that are available
- This number is often expressed as the bit rate of the digital signal, which denotes the number of bits available for encoding amplitude
- Thus an 8-bit encoding gives you 2^8 = 256 values, 16-bit gives 16^2 = 65,536, whilst 32-bit yields 32^2 = 4,294,967,296 values

Sampling, Quantisation and storage

- Higher sampling and quantisation (bit) rates give you greater accuracy
- . . . but also increase your storage requirements: doubling either sampling frequency or bit-rate doubles your file size per unit time of recorded material (uncompressed files)
- 22,500 Hz/ 16 bit encodings are generally appropriate for speech material

Info about recording and digitisation

- Ladefoged (2003); Plichta (2001)
- LINGUIST List 14.707, 14.97, 14-2229
- Audio File Formats FAQ http://www.cnpbagwell.com/audio.html
- Your local AV technicians (!)

Storing your digitised speech

- Better to avoid (highly) lossy compression formats such as MP3, as they tend to suppress information that can be important in acoustic analysis (use gz, zip instead)
- RIFF-WAVE, aka .wav files are used a lot with PCs running MS-Windows
- NeXT/Sun (.au; .snd), traditionally used on Apple, Unix
- AIFF/AIFC (.aif; .aiff; .aifc), Apple, some unixes
- PRAAT binary format (.Sound; .Collection)

Software

- PRAAT:http://www.praat.org
- Wavesurfer: http://www.speech.kth.se/wavesurfer/
- Speech Filing System (SFS): http://www.phon.ucl.ac.uk/resource/sfs/ (program created and maintained at UCL)
- EMU: http://emu.sourceforge.net/
- Audacity http://audacity.sourceforge.net: free, open source, general audio editor (cf. CoolEditPro; Adobe offerings)

Today's practical

- Digitising/ storing sound
- The PRAAT interface: Objects vs. files; file formats, etc.
- Interacting with sounds: editors, queries
- Annotation and TextGrids
- Analysing and transcribing pitch

References

Johnson, K. (1997) *Acoustic and auditory phonetics.* London: Blackwell.

Ladefoged, P. (2003) *Phonetic Data Analysis. An Introduction to Fieldwork and Instrumental Techniques.* Oxford: Blackwell.

Plichta, B. (2001) Best Practices in the Acquisition, Processing, and Analysis of Acoustic Speech Signals. Penn State Working Papers in Linguistics [. . . ?]