Instrumentation for “Dummies”
Supporting Clinical Practice With Simple, Low-Cost Technology

Stephen Tasko Ph.D. CCC-SLP
Western Michigan University
Kalamazoo MI

Donald Finan Ph.D.
University of Colorado
Boulder Co

http://homepages.wmich.edu/~stasko

Copyright Tasko & Finan 2008
Topics

- Why focus this talk on high quality sound recording?
- Sound recording 101
- Available recording devices
- Alvin: PC-based recording the easy way
- Beyond just listening to your speech samples
Instrumentation and Technology

Love it or hate it?
Why focus discussion on sound recording?

- High quality digital recordings can now be made with fairly common devices at relatively low costs.

- In my view, as both a speech language pathologist and speech scientist, good sound recordings give the SLP the most “bang for the buck”.
What high quality digital sound recording offers?

- Provides a permanent archive of client behavior
- Can allow quick, easy access to samples to monitor client status
- Has face validity
- Potential for more objective baseline and follow up evaluation
- Can support and guide clinical decision making
- Can serve to streamline and standardize assessment approaches
- Allows an entrée into
  - More quantitative instrumented analysis
  - Database development for clinician directed research on outcomes
Tasko’s rating system

- Preferable
- Highly preferred
- Not recommended
- Strongly discouraged
Sound Frequency ranges

- Normal humans hearing 20 Hz – 20,000 Hz
- Human speech ~ 50 Hz – 8000 Hz
- Standard phone circuits ~ 300 - 3000 Hz
For good **audibility**, we need to be sure our recordings capture the whole frequency range for speech.
What are you using the recording for?
Components of Sound Recording

- Microphone
- Pre-amplifier (maybe)
- Cassette recorder
Components of Sound Recording

- Microphone
- Pre-amplifier (maybe)
- Or

Copyright Tasko & Finan 2008
The quality of the recording is only as good as its weakest component.
Components of Sound Recording

- Microphone
- +/- amplifier

Digital Recording Device
- Digitization of sound
- Storing of sound as file

Playback ability
- Amplifier-Speaker
- Headset
How does tape recording differ from digital recording?

Analog (tape) recording:
- Acquires a continuously varying signal

Digital recording:
- Acquires a string of discrete numbers
Analog (Continuous)

Digital (Discrete)
Digital Recording

Since sound is converted to a string of numbers, two issues have an influence on recording accuracy/fidelity:

1. How often does the device "grab" a number?
   - Sampling rate/frequency

2. How much precision is used when number recording?
   - Quantization level
Sampling Rate

High

Low
Simple Rule

- sample rate/frequency must be 2X highest frequency of interest
- For speech it must be at least 16000 Hz
Digital Devices: Common sample rates

- 5,562 Hz
- 11,025 Hz
- 22,050 Hz: ✔️ ✔️
- 44,100 Hz: ✔️
  - CD quality
  - Captures range of hearing

32 KHz
16 KHz
8 KHz
4 KHz
2 KHz
Quantization

Inches less precise

millimeters more precise
Quantization

- Expressed in “bits”
- $\uparrow$ bits $\sim$ $\uparrow$ # of marks on the “ruler” = greater precision
- $X$ bits/sample = $2^X$ possible units of measure
- For example,
  - 2 bit = 4 levels
  - 4 bit = 16 levels
  - 8 bit = 256 levels
  - 12 bit = 4096 levels
  - 16 bit = 65536 levels
Quantization

- Expressed in “bits”
- $\uparrow$ bits $\sim$ $\uparrow$ # of marks on the “ruler” = greater precision
- X bits/sample = $2^x$ possible units of measure
- For example,
  - 2 bit = 4 levels
  - 4 bit ≠ 6 levels
  - 8 bit = 256 levels
  - 12 bit = 4096 levels
  - 16 bit = 65536 levels
Checking your own computer
Sound Fidelity vs. File Size

- Digital memory is relatively cheap
- Good client records are invaluable
Components of Sound Recording

Microphone

 +/- amplifier

Digital Recording Device
- Digitization of sound
- Storing of sound as file

Playback ability
- Amplifier-Speaker
- Headset
Digital File formats

Mono vs. Stereo

Alphabet Soup

.wav .nsf .mp3 .mp4 .wma .dss .au .etc, etc, etc
Digital File formats

**Uncompressed Formats (Lossless)**
- All samples are retained
- What you record is what you get
- Most common format
  - "wav" (PCM)

**Compressed Formats**
- Algorithm used to reduce file size yet maintain good audibility
- Driven, in part, by music industry
- Common Formats
  - .mp3, .wma, .dss
- Effect on speech samples is not well known
Components of Sound Recording

- **Microphone**
- +/− amplifier

**Digital Recording Device**
- Digitization of sound
- Storing of sound as file

**Playback ability**
- Amplifier-Speaker
- Headset
What (I think) you need to know about microphone selection

- Type of microphone
- Underlying operation
- How it connects to other devices
- Pick up patterns
- Frequency Response
Microphones: Common Types

- Handheld
- Head worn
- Clip on/lavalier
- Built in to recorder
Underlying operation

**Dynamic**
- Very inexpensive
- Will likely work with older computers (older soundcards)
- Do not require any power to operate
  - Perform more poorly than condenser mics (less accurate)
  - Signal may be too small for some small cards

**Condenser (electret)**
- More accurate recording, better performance: “Professional”
- More sensitive and responsive: will perform better for speech analysis
  - Can be more costly
  - Can be more fragile than dynamic mics (more sensitive to rough handling).
  - Require a power source (bias or phantom power, depending on the mic).
    Most sound cards can provide bias power, some condenser mics come with a battery pack if they require phantom power.
Common Microphone Connectors

XLR (balanced): found on professional microphones

Mini XLR: often on wireless microphones

1/8” Phone Jack (mini phone jack): Required for PC sound card input
Microphone Response Pattern: Omnidirectional

- Picks up sound all around the microphone.
- Good for if the sound source is moving, and the microphone is stationary.
Good for eliminating outside noise.  Picked up from the front.

Unidirectional (Cardioid): Sound is mostly
Microphone Frequency Response

What is a frequency response curve?
- Graph that shows how the microphone responds over a range of frequencies
- Look for a relatively flat line within the frequency range for speech
Frequency Response Curve

Speech Range
Frequency Response Curve
To pre-amplify or not to pre-amplify

- Great variation in size of signal coming from mics
- May need to amplify the mic signal so it is heard by the recording device
- Often
  - Best mics put out small signals
  - Cheapest mics put out largest signals
A Tasko Seal of Approval

- **Audio-Technica ATM75 (without powerpack!! - cheaper)**

**Features**
- Condensor type Headworn
- Directional
- Acceptable frequency response
- Decent sized output signal
  - amplification not typically necessary
- Can be purchased with a variety of connectors
- **REASONABLY PRICED (~ $99)**

Copyright Tasko & Finan 2008
Audio Technica ATM75

http://www.soundprofessionals.com/cgi-bin/gold/item/ATM75-SP-NP
Recording Issues

- Keep mic-mouth distance relatively constant within and across recording sessions
- Setting your recording levels is very important
- You really want a device that monitors the recording level
Apply the “Goldilocks” Rule
That recording level is just toooooo low!

- Environmental noise becomes an issue
- Does not take advantage of the precision of the recording device
That recording level is just tooooo high!

- Signal will be “clipped” resulting in distortion
- Truncates the largest amplitude components
Components of Sound Recording

- Microphone
- Amplifier

Digital Recording Device

- Digitization of sound
- Storing of sound as file
- Playback ability
  - Amplifier-Speaker
  - Headset

/+ amplifier
A brief survey of digital recording devices

- Desktop
- Laptop
- The PC (or Mac)

- Digital equivalents to the professional-grade devices
  - iPods
  - Digital dictaphones
  - PDAs (Palms, Smartphones)

- Personal recording devices
  - Digital equivalents to the professional-grade devices

Devices developed primarily for music industry
Questions to ask about portable recording devices

Basic recording performance questions
- Are the sampling rate options appropriate for speech recording?
- Is the quantization rate appropriate for speech recording?
- Is the overall frequency response of the device acceptable for speech recording?
Questions to ask about portable recording devices

Microphone/recording level related questions

- Can an external microphone be used?
- What kind of connectors are needed?
- Will you need a mic amplifier?
- Can recording levels be adjusted?
- Can recording levels be monitored?
Questions to ask about portable recording devices

Data Storage/Transfer Questions

- What file format is used to save the data?
- Can it record in a lossless format?
- What is the recording capacity of the system with the preferred recording settings?
- Does it have only built-in memory or can you use standard memory media (e.g., SD cards)?
- How easy is it to move the recordings off the device and onto a PC for archiving (i.e., USB)?
Questions to ask about portable recording devices

Human Factors Questions

- Can you easily play back recordings?
- What is quality of device playback?
- Is it easy to use?
- Is it too big or small?
- …
Personal Recording Devices

- Digital dictaphones
- iPods
- PDAs (Palms, Smartphones)

- Many have good specifications
- File format options can be quite limited
- Often have serious limitations regarding microphone/recording level flexibility

Copyright Tasko & Finan 2008
Dictaphones

- Frequency Range (Super High Quality Recording): 100-17000Hz
- Recording time: At least 1 hour at super high quality.
Other options: iPod
Personal Recording Devices

Not Recommended:

• Voice recorders using DSS (Digital Speech Standard) recording format.
  – (Some recorders use DSS format only for their lower quality recordings)

• Voice recorders with a frequency range of less than 100-8000Hz
  – (many have a range of ~300-7000Hz)
“Professional” Grade Devices

- Greater flexibility in all areas
- Minimize concerns about digitization specs
- Typically use a .wav (uncompressed) format
- Larger and sometimes bulky
- Can be quite expensive
- Good values if you look at devices designed for the music industry
Professional Duty

Marantz PMD660
~ $500
Professional Duty

Boss Micro BR
~ $230

Don Finan’s Seal Of Approval
Professional Duty

M-Audio Microtrack < $300
A Significant Disadvantage of Recording with Portable Devices

- **Housekeeping duties**:
  - Sound recording file management
  - Naming files so you know what is in them
  - Moving files from device or portable media to a more permanent archive (PC, CD)
  - Finding software to replay the files you’ve acquired
Alvin: Digital Recording the Easy Way
What is “Alvin”?

- Freely available software that exploits recording and play capabilities of modern personal computers
- Originally designed to as a control and user interface for psychophysical experiments
- Offers enough flexibility that it has potential to ease clinical sample collection
Why is it called “Alvin”?

- Named in honor of Alvin Liberman
- Prominent speech researcher at Haskins Laboratories
- 1917-2000
“Alvin” vs. “Alvin Clinical”

- Identical software “under the hood”
- Alvin Clinical is a streamlined version of Alvin that focuses on features useful for the SLP
It is really just a more sophisticated recording and playback device. It can be customized to ease collection of standard speech/voice/language samples and removes much of the “housekeeping” associated with sound recording, allowing for easy recording and playback. It takes care of many of the aforementioned concerns.
Alvin Clinical

- Alvin can be customized to specific user needs
- I’ve attempted to make a set of templates that would fit most user needs.
- Find a template that matches your general need, then customize the content that goes in it
- Not as difficult as it sounds
Prolong the vowel 'ah' for 3-5 seconds

Trial# 1 / 16
However...

- You do need to get familiar with how to navigate through folders and files in Windows (using Windows Explorer for example)
- You will need to use a basic text editor like Wordpad to create your custom protocol
- It helps to be able to make shortcuts to the Desktop
Recording templates

- Text presentation only
  - Text recording_nfb
  - Text sound recording_nfb

- Image presentation only
  - Pic recording_nfb
  - Pic sound recording_nfb

- No prompting
  - Basic recording_nfb
  - Basic recording_fb

As above but with automated playback of recorded samples
Three key components

- Alvin software
- Protocol “script”
  - Tells Alvin software what to do
  - Uses the Lua scripting language
  - Has a “.luax” extension
  - I’ve designed template scripts to avoid having people mess with this
- Stimulus file
  - Provides the specific details of what speech samples you want to acquire
  - Has a “.stim” extension
  - You will have to edit this initially
Stimulus files (.stim)

- A simple text file
- Can be created and edited with Wordpad
- Contains information about
  - prompts you want to use
  - names you want to give the recordings

Unless you want to modify the protocol in substantial ways, it is the only file you should mess with!
Sample stimulus (.stim) file for text presentation format

- "Prolong the vowel 'ah' for 3-5 seconds" ahh.wav
- "Prolong the vowel 'ee' for 3-5 seconds" eee.wav
- "The blue spot is on the key again" -Repeat 3 times" key.wav
- "Buy Bobby a puppy" -Repeat 3 times" bob.wav
- "Susie sees the sky" -Repeat 3 times" sus.wav
- "Mama made lemon jam" -Repeat 3 times" mam.wav
- "Early one morning a man and a woman were ambling along a one-mile lane ..."
- "He saw half a shape mystically cross fifty or sixty steps in front of him ..."
- "When the sunlight strikes raindrops in the air, they act as a prism and ..."
- "Briefly describe the picture provided" pic.wav
- "Pitch glide" pgl.wav
- "Sing 'Happy Birthday'' bir.wav
- "Sharp cough" cgh.wav
- "Additional Sample 1" vs1.wav
- "Additional Sample 2" vs2.wav
- "Additional Sample 3" vs3.wav

Text you want displayed

Name of sound file (must be .wav)
Sample stimulus (.stim) file for picture presentation format

```
ape.bmp  ape.wav
B.bmp   b.wav
bale.bmp bale.wav
```

- Name of image (must be .bmp)
- Name of sound file (must be .wav)
Note

- To avoid having to modify the instruction file, stick to images with a fixed size of 300 by 300 pixels
Adjusting sound recording and playback in Windows
More Advanced Activities
What if I want to rename the template

Steps

- Copy and paste the whole template folder into Alvin Clinical Folder
- Rename the template folder
- Rename the .stim file
- Rename the .luax file
- Open the .luax file with Wordpad
- On line 17 or so there is a line that starts “controlFile=filename.stim”
- Change this to the renamed .stim file, save and exit
---[ ]

Experiment: Prompted recording
Presentation: an image
Response: A recording of the utterance
Feedback: None
Stimulus: <prompt> <filename>
Output: <trial index> <filename> <response> <reaction time> <replay count>

---]

defineExperiment {
    helpFile="instructions.txt",
    shuffle=0,
    controlFile="pic_recording_fb.stim",
    stimulusDir="stimuli"
}

function exp.layout()
    canvas.backcolor("forest green")
    bitmap.new{name="bitmap", path=exp.getStimulusFile("blank.bmp"), x=40, y=90, w="300px", h="300px"}
Adding the new protocol to the Menu

- Go into Lua folder
- Open menus.lua with Wordpad
- Add a path line to the file using the same syntax as the already existing menu options
- Save and exit
menu.addMenu("Sample protocols",
{
"Basic recording without feedback", "basic_recording_nfb/basic_recording_nfb.lua"},
{"Basic recording with feedback", "basic_recording_fb/basic_recording_fb.lua"},
{"Text prompted recording without feedback", "text_recording_nfb/text_recording_nfb.lua"},
{"Text prompted recording with feedback", "text_recording_fb/text_recording_fb.lua"},
{"Text and sound prompted recording without feedback", "text_sound_recording_nfb/text_sound_recording_nfb.lua"},
{"Text and sound prompted recording with feedback", "text_sound_recording_fb/text_sound_recording_fb.lua"},
{"Image prompted recording without feedback", "pic_recording_nfb/pic_recording_nfb.lua"},
{"Image prompted recording with feedback", "pic_recording_fb/pic_recording_fb.lua"},
{"Image and sound prompted recording without feedback", "pic_sound_recording_nfb/pic_sound_recording_nfb.lua"},
{"Image and sound prompted recording with feedback", "pic_sound_recording_fb/pic_sound_recording_fb.lua"},
{"Sample Playback", "play_samples/play_samples.lua"},
})
Easiest way to access Alvin Clinical Protocols

- You don’t need to use the Menu itself
- Make a shortcut on the desktop to the specific .luax file
- When you click on this, it will automatically load
Issues with PC recording

- Not all PC sound cards are identical
- Vast majority work perfectly well
- Occasional problems include
  - Noise or static
  - Shifts in recording baseline (does not affect how it sound)
Beyond Listening

Freely available programs for basic acoustic analysis
Freely available

Can be used to view and analyze acoustic characteristics of sounds you have recorded

Developed by P. Milenkovic (UW Madison)

http://www.medsch.wisc.edu/~milenkvc/tools.html