

# Question Tools



www.questiontools.com

## Recording speech for use in education and training materials

Recording speech for use in training and education might seem attractive. If used wisely, it can add significantly to CD-ROM and web-based materials. Here, we are concerned with recording sound for use in tests, exams and content delivered using Question Tools. However, the advice here is not really Question Tools specific. First, we will consider the reasons for and against using speech in educational and training materials, before moving on to the equipment needed, and the mechanics of how best to record speech.



The first question you should ask is, do you need speech? If you are attempting to improve something that has already been judged by its users as deadly-dull then adding speech might be part of a solution, but the main answer to such problems involves a more radical overhaul of the materials rather than the addition of extra media. If you have a manager or colleague who is convinced that speech is the panacea needed by your organization's materials then try to find some of those really dull training videos and make them sit through at least fifteen minutes. I have a promotional video on beetroot for this purpose.

## Why not get a professional voice artist?

Sometimes a professional voice artist is a good solution. However, you should remember that some 'professional' voice artists will be no better than someone you may already have in your company. In addition, if you need to record any corrections then getting the clear-speaker from down the corridor can be quicker and cheaper. If you are going to use a known actor to do the speaking then getting him or her to return to make corrections might be very expensive. One approach is to record the voice yourself, or using someone who speaks well in your organization, and then get the actor or voice artist to record the speech only when you are sure there will be no more changes.



**Voice is expensive to change.** If you are asked to change a couple of screens in your test or explanation of a topic, get a different photograph or use alternative example questions, then this can be done. However, changing speech is more difficult than it first appears. Imagine that you have been asked to change the speech in question 17. This might seem OK. However, if you re record just this sentence there is a good chance that it will not fit with your other recorded speech, even though you will have used the same person and the same recording facilities.

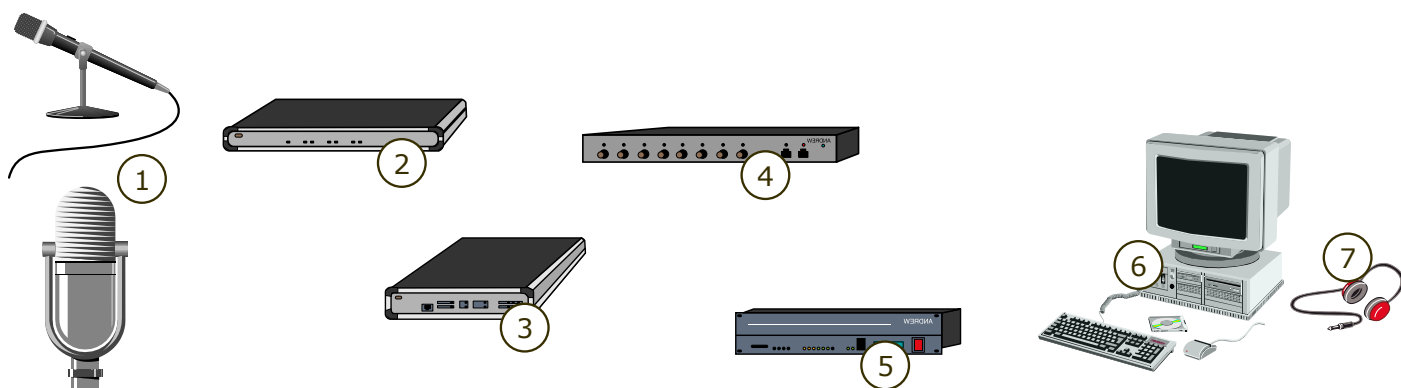
A person's voice is variable, and small variations between sessions can be easily noticed. For example,

- the pitch of the voice may be higher or lower;
- the speed of the speech can vary;
- the emphasis, and even accent, of the speaker can change;
- the speaker may be nearer or further away from the microphone;
- the acoustics in the recording booth may have changed.

To the uninitiated these may all sound like trivial differences, and they are trivial. Unfortunately, more of the human brain is allocated to processing sound than processing visual images, and we are very good at detecting these trivial variations. In short, your users will notice, and they will not like what they hear. As a consequence, if you want to change a sentence you invariably have to re record the whole section or module.

## Required equipment

Most of the items shown here are optional. You can, at worst, get by with a microphone plugged into the sound card in your computer.



1. Microphone, surrounded by sound proofing, usually with a pop shield
2. Preamplifier (optional, can use direct connection to sound card instead)
3. De-esser (optional)
4. Compressor/limiter (optional)
5. External digitizer (optional—using sound card inside computer to perform digitization more common)
6. Sound card
7. Headphones

**Microphone.** A good microphone is a must. There is nothing wrong with experimenting using a £30 (\$50) microphone. However, if you want a good voice microphone then you should expect to pay £200 (\$350) or more. A large diaphragm, professional microphone, such as a Neumann, is going to cost at least £500 (\$750). However, there is an additional cost to using most professional microphones—they require power (often called phantom power). This means that the preamplifier needs to supply this, and only the more expensive preamplifiers do this.

**Pop shield.** When we speak certain sounds release a rush of air. Typically, these are the 'b' and 'p' sounds. Try holding your hand just in front of your mouth as you speak, and try a few sentences with words like 'probability' and 'bipolar' in them. Each time you make a 'p' or 'b' sound you should feel a rush of air onto your hand. These sounds, and the movements of air they create, can cause microphones to 'pop'. On the recording it sounds like a thump. A pop shield is something that sits in front of a microphone to dissipate the energy from this rush of air, and prevent the microphone diaphragm from being pushed back by the force of this small blast. They can be bought for £30 (\$50). Some people even make their own, using bits of wire and stockings or tights. A pop shield should be positioned about 12 cm in front of a microphone.

**Sound proofing.** It is a common misconception that sound proofing in a recording booth is to stop sounds from outside leaking in. Sound proofing, except when it is installed as a piece of heavy engineering with sand-filled walls, has little effect on loud noises from outside. The main aim of sound proofing is to stop the sound of the voice from bouncing off walls behind, above, below and to the sides of the microphone. In this respect you can save money. Professional sound proofing is not much more effective than some bits of old carpet and a few cheap bed covers or quilts. The aim in the recording is to have a 'dead' or 'dry' sound, with little or no ambience, and you do not need to spend a lot of money achieving this. The setup used for recording the speech on the Question Tools CD involves a very expensive microphone, an expensive preamplifier that is part of a mixer, a very expensive sound card and top-of-the-range software—and yet the sound proofing comprises a number of cheap quilts bought from a supermarket.

**Preamplifier.** Microphones produce very small signals. So small, in fact, that you usually need to amplify them before they reach the computer. A preamplifier is really just an amplifier dedicated to amplifying very small signals. A cheap preamplifier should cost around £30 to £50 (\$50 to \$80), while a reasonable preamplifier is likely to cost more than £200 (\$300). If you have opted for a high quality microphone then you will probably need a preamplifier that supplies power to the microphone (phantom power). If you have used a cheaper microphone then you may be able to miss out a preamplifier all together and plug straight into the microphone socket of your computer's sound card.

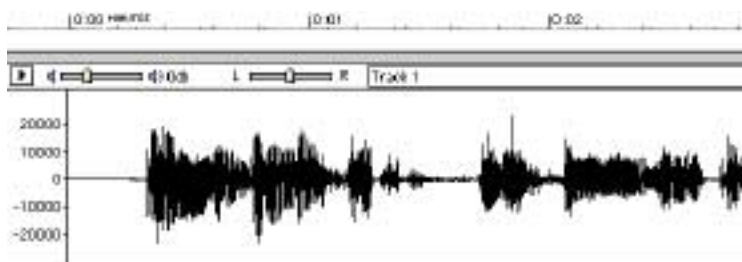
**De-essor.** Sibilance is the professional name given to 'sss' sounds. Sometimes these sounds can be over-emphasized on a recording. Providing this effect is not too great you can remove it automatically with a de-essor. De-essors are usually employed only by professionals for speakers who are susceptible to the problem of producing excessive sibilance.

**Limiters and compressors.** Sometimes it can be useful to have a limiter or compressor to take the signal after the preamplifier and process it before it reaches the computer. Limiters and compressors work in slightly different ways to achieve similar effects. A limiter watches the volume of the sound. If it exceeds a certain threshold then it starts to restrain the increase. A limiter can be a good way of restricting the volume of a sound to make sure that it does not peak with loud utterances from your speaker. Compressors work in a slightly different way. They amplify the signal, such that small sounds are amplified more than large sounds. However, this can also bring out noise and breathing on the recording.

**Digitizer.** In order for the sounds the microphone detects to be used on a computer they must be converted from analog waves into numbers the computer can understand. If you plug your microphone directly into your sound card then the sound card is going to be performing this digitization. Digitization is a key determinant of sound quality, but computer sound cards, such as the SoundBlaster series, are not such a bad starting point. There are three options:

1. Connect the microphone directly to whatever sound card has been installed in the computer. This approach is cheap and simple. It will probably not produce the best results, but it could be acceptable. In any event, it is a good starting point.
2. Install a professional sound card, such as an AudioMedia III. This will produce better results. However, you will not be able to directly connect your microphone. Instead, you will have to send the signal via a preamplifier instead. In addition, many specialist sound cards are designed to accept multiple channels at any one time, and yet you will typically only use one channel when recording speech—hence you will be paying for facilities you will probably not use.
3. Use an external digitizer. The problem with any sound card in a computer is that it will be subject to a certain amount of electromagnetic radiation from the other components in the computer, and this can reduce sound quality, even with an expensive sound card. The best option is usually a digitizer that is outside the computer, and these are more specialist. Moreover, you will need to ensure that once digitized the signal can get into your computer. USB connectors are a possibility, although specialist external digitizers will use music industry connections and standards, and your computer will usually need a further plug-in card to accept these.

**Recorder/Editor.** You will need some software on your computer to record and edit the speech. SoundEdit is a popular choice, although there are also one or two very good free editors available. Personally, I use Cubase Audio to record, edit and apply sound compression, and Sound Edit to chop the files up. Your recorder and editor needs to do five things:



1. Record the incoming speech as a sound file.
2. Allow you to select parts of the sound and quieten it.
3. Edit the speech, to remove mistakes and obvious breathing sounds.
4. Apply a sound compressor (different from a file compressor).
5. Copy or export sections of your speech into a number of smaller sound files.

**File compressor.** Once you have your speech recorded and edited you can use the sound files just as they are. However, they will tend to be large. You can make these files much smaller by using a file compressor. Some specialist compression software is rather expensive, but some MP3 compression software is very good, as well as cheap, and in some cases free. Although the MP3 compression format was designed for music, it does a good job of compressing speech. Most importantly, Question Tools can handle this file format, as can most versions of Windows. Try to find a compressor (also called a 'ripper' by some) that will process more than one file at once (batch process). If you have 100 files to compress then setting the compressor going and making a cup of tea is much more acceptable than compressing each one individually.

## Script

A well-written piece of text does not necessarily make a good script. For one thing, a good script does not necessarily follow the rules of grammar. When we speak we typically use extra words and phrases. Words that look out of place when written down do not sound out of place when heard because of the emphasis the speaker uses. Consider these two examples, both of which deal with house building.

### **Example 1**

As building methods and practice have evolved houses have been sited in ever-smaller plots. The main cause is economic, with the price of land dictating building priorities.

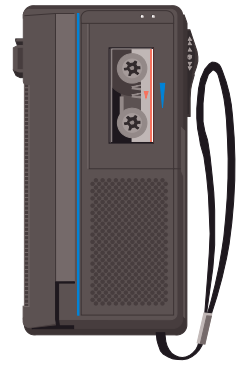
### **Example 2**

Over the years house-building has changed. Houses are crammed into smaller and smaller plots of land. And the cause of this is the price of the land itself. Land is expensive, and the more houses builders can squeeze into the available land, the more profit they will make.

Example 2 uses everyday language, and is longer as a result. It might not be good writing (starting a sentence with 'and' is rarely attractive), but it sounds better when spoken, and this is the real test. If a colleague or customer passes you a script make sure they appreciate the need to rework some aspects of the script. If you accept a formal essay to be used as a script, in the belief that a lively and professional speaker can make it work, you are probably making a mistake.

## Speaking

If you intend to record your own voice then spend some time practising. We all drop sounds from words. Most of us are lazy speakers, because the people we are speaking to can usually work out what we meant despite our mumbles and errors. However, on a recording these dropped sounds and careless pronunciations sound dreadful. There are several things you should do:



- Practice, record yourself and listen to how you sound.
- Do not be afraid of any regional accent you might have. A clear, warm voice in almost any accent is appreciated by listeners, while someone trying to disguise an accent will come across as false.
- Try pronouncing the beginning and ends of words more distinctly.
- If you gabble, or you are a lazy reader (i.e. you read what you think is written and not what is actually written in the script) then slow down.
- Try to introduce more variation in the pitch of your voice. Most of us sound flat when recorded.
- As you get better start thinking about who you are when you play this pretend game of speaking into the microphone—you are the reassuring and knowledgeable friend of the listener speaking directly into his or her ear. Try to adopt this tone if you can.

Listen to some of the people on the radio, particularly spoken programs if you want an example of how it can be done well. My personal preferences are the BBC Radio 4 and World Service broadcasts. You will feel foolish as you experiment—it is embarrassing at first. You may even feel humiliated when a recording of your voice is played in front of other people. Your own reaction to your voice is probably the least important—it is how others hear your voice that matters. Once you have practised your technique you should ask several people for an opinion—preferably people you can trust to be frank.

## Getting ready to record

Setting up to record for the first time should be something to which you allocate several hours. You can plug in and get going in a matter of minutes. However, if you do this you will probably find, some time later, that there is a way of using your equipment to get a much better result. You will then have the difficult decision of whether to re-record the speech you have already recorded, edited and compressed.

Your recording software may record at a fixed sample rate and depth. However, if you can adjust this then CD-quality is good choice (44KHz, 16bit). If you can also select a file format for your recorded sound then Wave or AIFF (.wav, .aif) are good choices. Do not select any formats that require compression, such as MACE or MP3. These will slow down your computer and degrade the quality of your sound. As a general rule file compression should only ever be applied once—to the finished sound.

A better approach is to just record a few sentences. Vary the settings, and try again. The approach is to vary things (and take notes) to see which approach gives you the best quality of recorded sound. The ideal recording is one where the recorded speech is just loud enough to reach the maximum volume at which speech can be recorded. Too quiet and you will find the background hiss or general noise will be louder than it should be. Too loud and you will get 'clipping'. This is where the volume of your speech causes distortion.

**Use headphones.** When judging sound quality always make the judgement using headphones with the volume setting somewhere around medium. Loudspeakers are convenient, but reverberation in the room will make judgements difficult.

**Use a poor-quality computer.** Using headphones to make judgements about sound is a sensible choice. Nevertheless, after all of your experiments it is best to check your finished output through a poor quality computer. Even the top music recording studios do this—they have equipment to simulate poor quality radios and CD players.

**Sound proofing.** You may want to start by varying the carpets, quilts and whatever else you are using to create that dry sound. A dry sound is one where there is no reverberation (sound reflections) from the room. If you listen to the recorded speech through some headphones, and it sounds as if the voice is right next to your ear, or almost in your head, then you have probably got a good, dry sound.

**Preamp volume.** Adjust this volume so it is as loud as it can be, without any distortion. If you get some distortion then just turn down the volume a bit and try again. Once you have a setting you like, then it might want to turn down the volume on the preamplifier just a little more, to provide enough headroom for any loud speech.

**Input volume.** The software you are using to record your speech will usually have some way of adjusting the recording level. Try altering this so that it is as loud as possible without introducing distortion.

**Autoleveling.** Some cheap tape recorders have autolevelers. These are very aggressive and crude sound compressors. Some sound cards, and some sound recording software, also have these dreadful devices. Turn any autoleveling off. If you want compression during recording then it is best use a proper compressor.

**Limiters and compressors.** There is no correct answer as to how to use and set these. Experts do not always agree on the best settings. Personally, I have tried limiting and compression during recording, but always got a better result with a completely clean signal.

## Recording for real

Generally, you should aim to record in sections, usually with no more than thirty sentences in a section. However, these should be natural breaks in the script. After you finish a sentence, or a couple of sentences that go together, leave a pause of three seconds or more before you move onto the next sentence—this is helpful during editing. If you make a mistake during a sentence then pause and try the sentence again. Do not try to leave it until later. If you repeatedly make mistakes, particularly the same mistake, then slow your speech. Most people who record for the first time speak too quickly.

Recording can feel very pressuring. You have to:

- avoid bumping the microphone stand or rustling the script;
- read the script accurately;
- read with the appropriate tone and variation;
- sound enthusiastic.



By the time you have managed to speak the script without fault you will probably sound flat and dull. In short, you will need several takes to get it just right. Make sure someone is listening to you as you record. They will be able to let you know when you sound too flat or dull. They may also spot mistakes that you miss, and they should not be afraid to interrupt and ask you to try a sentence again.

## Editing

When you have a good recording, then your first step is to find where your sound recording software has placed the file and make a copy of it. When editing it is easy to make a mistake, and backups are quick, easy and often needed.



1. A false start/mistake—to be silenced
2. A breath before the sentence—to be silenced
3. After the sentence, some light breathing and general noise—to be silenced
4. A small breath taken between words—better to quieten than silence



Start at the end of the recording and work backwards. The last take for any sentence is usually the best. In addition, sound file editors can often be quicker if you start from the back of the file. You should aim to silence the portions between sentences, as well as any mistakes. To do this select the portion of the sound in your editor. There will then be a Silence command, or an Amplify command that allows you to set the volume to 0%.

Your next step should be to look for sounds within the sentences you want to remove. Typically, these will be loud breaths, or gasps, as the speaker attempts to grab some more air between parts of the sentence. You could silence these, although a better effect can sometimes be found by reducing these sounds (by selecting the gasp and setting Amplification to 20%, or using a Quieten command if there is one present). This can be time-consuming. However, it is one of those tricks that produces a better quality sound. People will tell you it sounds better, but they will rarely know why.

**Compression.** Once you have your sound edited, the next step is to apply any compression and amplification. This means sound compression, not file compression. Compression, as mentioned earlier, tends to increase the volume of quiet sounds and decrease the volume of loud sounds. Mild compression can increase the clarity of the 's' and 't' sounds, as well as other word endings.

**Amplification.** You can, if you wish, increase the volume of the recorded sound. If you increase the sound level too much you will get clipping, where the sound peaks reach the maximum values. This can degrade the sound. Strangely enough, however, this does not introduce distortion (unlike clipping during the recording phase). If you do amplify the speech, make sure you apply this amplification to all of the speech in the section, not just individual sentences. If you apply amplification (or normalization) to individual sentences you can get very noticeable and unnatural changes in volume.

**Cutting the speech up.** You then need to cut your speech up into individual sentences. This is because the multimedia or e-learning software will not normally play one long sound file, but will usually play one file for roughly one sentence. In most editors you can achieve this by copying and pasting into a new file. At this stage you can choose to compress the sound files, but it is better to stick to AIFF or Wave files at this point if you can.

**File compression.** As mentioned earlier, I have found that MP3 to be a good choice for sound file compression, which is surprising given that it was designed and optimized for music rather than speech. Setting the compression software to 24KHz, 40Kbits per second (mono) produces good results in terms of acceptable speech quality and small file sizes. It is wise not to delete the original files, as you may have to return to them to re compress the files using slightly different settings.

## Summary

Recording speech to a professional standard is difficult. Now you have read this article you will probably begin to appreciate why. You might think that recording musical instruments is even more difficult. As it happens, the reverse is true. Most experienced studio engineers will tell you that recording a brass section, a string ensemble or a guitar played through an amplifier at full volume can be much easier than recording speech.

The experiments you will need to get really good results can very time consuming. Nevertheless, if you persist, you can achieve professional results. You will be able to record speech when you want at little extra cost, while others are suffering the delays and costs of contracting the work out to professional voice artists in recording students.

A cheap microphone plugged directly into a computer's built-in sound card is a good place to start. You might consider aiming for speech quality that is acceptable rather than professional. As your experience grows you can improve your equipment, technique and your results.