Recording Techniques: Wordlists/Reading lists

A good recording starts with a clear plan, appropriate population sampling, good equipment, and a good recording technique. Following these steps will help make it a smooth process.

The guidelines here are for eliciting a fairly formal, citation, style which falls on a continuum of many possible speech styles. Depending on the purposes of the study a more formal style or a less formal style, or some combination thereof, might be needed. More formal styles are elicited using techniques that induce some degree of hyperarticulation in which the speaker becomes more self-aware of words or even specific speech sounds or in which the relation of the talker to the social setting shifts. These tasks include words in isolation, minimal pairs tasks, tasks which involve some sort of request on the part of the researcher (or intermediary listener) for clarification, or tasks in which a new interlocutor from a different social group is introduced. There are often reasons for wanting to elicit a less formal style of speech as well which is marked by a decreasing self-awareness on the part of the speaker, an increase in phonetic reduction, and an increase in colloquialisms and regional or sociolectal markers (particularly stigmatized ones). In decreasing order of formalness, here are a few of the many tasks that elicit less formal speech than sentences do: reading passages, interviews, map-tasks (aka barrier tasks), conversations with a demographically matched stranger (same dialect, same age, same gender, etc.), conversations with a demographically matched friend. See Labov (1972, 1984) and Milroy (1987) for nice discussions of many issues related to social, demographic, and task factors. It is very important to choose a task which is designed to record the appropriate speech style and the appropriate words and phrases. In addition to the task it is important to control for dialect, age, and other demographic and social factors. Usually for a study involving wordlists or carrier phrases the population should be dialectally homogeneous and have an equal number of male and female speakers. Age may also be an important factor so either the speakers should be drawn from a single age range or represent the spectrum of ages in the population (balanced for gender).

1 Wordlist construction

Have a clear idea of what exactly you are going to record. Make a recording list, make sure there are no spelling mistakes, and make sure that all the words are recognizable. Use a large font and keep plenty of spacing between the lines.

In general words are embedded in a carrier phrase of the type:

“Say ___ again.”

The goal of the carrier phrase is to increase the naturalness of the pronunciation and to reduce list effects on the words of interest. It is important that you chose your carrier phrase carefully to avoid introducing confounding variables into your study. For example, if you are looking at word final voiced and voiceless stops in English, you would have to avoid the word initial unstressed vowel in “again” because it will cause /t/ and /d/ to flap. Likewise, if you wanted to measure the duration of the word final stop, you would have to avoid word initial stops like the /k/ in
“Say ___ quickly.”

because it will mask the duration of the stop closure; try word initial nasals or fricatives because they will permit you to see where the stop closure ends (that is if there is no pause between the words).

Each page of a reading list should start and end with three to five filler items (the carrier phrase with words that are not important to the study). The reason for this is that speakers generally “chunk” all the items on a list and superimpose a list intonation over the chunk. Putting fillers at the beginning and end reduces the effect of the list intonation on the words. If you want to avoid attracting your subjects’ attention to specific aspects of the words on the wordlist, you may want to include distracter words. Generally you will include the same number of distracter words as you have words of interest.

Words/sentences should appear in random order (with the exception of the filler items at the beginnings and ends of the list). You should use a different randomization per repetition (to avoid semantic interactions and list-order effects on pronunciation). Generally I prepare three randomizations which I label “A, B, C”. Each subject in the study records one of the 9 possible orders of the randomizations (ABC, ACB, BAC…). The number of repetitions is determined by the needs of the study; some researchers collect five repetitions, others even more.

2 Equipment preparation

Set up and test the equipment well ahead of time. Test it in a dry run the night before or at least an hour before to make sure that all the cables are attached etc. You should have a professional quality microphone (usually has an XLR jack which has three plugs in a triangle), a high quality amplifier (if necessary) and a high quality recorder. One missing link and the whole thing is ruined. Generally the sound cards that come stock in PCs, especially laptops, are not ideal (though still better than cassette tapes). Macs have a reasonable sound card, PCs and Macs equipped with high end sound cards with on-board DSP chips are high quality, and external (USB or Firewire) devices are usually reasonable to high quality (especially if they have XLR inputs). Never use a hand held Dictaphone, a walkman, or one of those cheap cassette recorders that they let you play with in elementary school. If you do you will end up with poor quality recordings that are useless from an instrumental point of view. (no need to worry here, you’ll be using lab equipment). Remember that you can’t fix a poor quality recording.

Digital sampling rates:

I generally use a standard sampling rate to ensure that my sound files will be compatible with any computer. There are 3 standard rates (44.1, 22.05, 11.025 kHz) and each is appropriate for a different task. Most music is recorded at 44.1 kHz (which has a Nyquist frequency of 22.05 kHz) because music contains energy in frequencies up to the maximum sensitivity of the average young person’s hearing range of around 20kHz. Speech is optimized for transmission of information under less than ideal hearing conditions (including the elderly with
typically limited hearing above 10 kHz) and therefore contains little information that is crucial above 8 kHz. Vowels and sonorants contain little information above 5 kHz while many obstruents rely on information that is distributed in the signal up to 8 kHz. Therefore, many researchers use a sampling rate of 22.05 kHz (Nyquist at 11.025) for speech in general and for research on obstruents and some limit the sampling rate to 11.025 kHz (Nyquist at 5.5125 kHz) for vowels and sonorants, or for pitch related work. Since you can’t recover frequencies that haven’t been sampled, I generally err on the side of caution and choose 22.05 kHz as my sampling rate, and down sample to a lower frequency if desired for specific aspects of the project. Any modern analog to digital algorithm (or downsampling algorithm) will anti-alias filter the input before sampling; however if you have home grown code or are using antique tools you should always low-pass filter at the Nyquist before sampling to avoid aliasing (artifact frequencies created by the sampling process).

3 Informant preparation

Once your informant has arrived make them comfortable and go over the recording list once more. Relieve them of any food, drinks, gum, candy, crinkly wrappers and other inadvertent noise makers. Make sure you get all the background information before you begin recording. They should sign a Human Subjects Consent form before you collect any data. Use good mic-side manner: be polite, patient, and accommodating but firm (especially about the gum).

4 Microphone technique

Good microphone technique is everything! Different mics are made for recordings at different angles. For off-axis mics, like the Shure headmounted, the mic should be 1/2 inch away from and TO THE SIDE of the mouth. For other mics, like Electrovoice in the recording booth, you want the mic straight in front of the speaker about 3 inches away. You want to pick as much of the speech and as little background as possible, but you must avoid bursts of air. Also check to see if the orientation of the informant’s nostril is blowing air onto the mic. Test the setup by having the person say words with /p/, /s/, and nasals and laterals in a variety of words.

5 Set the levels.

On an analog device the indicators should be hovering between the middle of the dial and the beginning of the red area during speech. On a digital device you have more dynamic range and it’s important to NOT overload; therefore, I usually don’t let the indicators get into the red at all, and only just into the yellow at the very loudest peaks. Adjust the input levels on the tape and amplifier, as well as adjusting the relationship of the microphone to the speaker’s mouth. While setting the levels have the informant read continuously from the reading list. Most people start out loud and get quieter as the session goes on so monitor the levels during the recording. Set the peak levels with words that contain open vowels like /a/ since they are the loudest.

6 Begin the recording.

Put the reading list out of reach of the informant (paper makes a ton of noise), encourage them to get comfortable, recheck the microphone, begin recording. The
first thing on the recording should be your name, the date, the session number (if multiple sessions), the place (UW Linguistics Phonetics Lab), and an identifier for the informant. Have your informant read the list as naturally as possible. If they start off sounding very careful in their pronunciation you should let them get used to the recording setup by making one dry run through the wordlist.

7 Monitor the recording conditions.
Monitor continuously; keep track of the condition of your participant, the levels of the recording and the recording list. If the participant makes a mistake, don’t panic, don’t start over, just interrupt and have them begin reading again from a few lines above where the mistake was made. Better yet build in enough repetitions that a single mistake doesn’t matter to the overall experiment. If an informant wants to stop recording, stop. It’s their right to leave at any time. Moreover, if they are uncomfortable it can ruin the recording even if they stick it out.

8 Thank the informant.
Thank the informant. Fill them in on any details that they didn’t get before hand. Make sure that they have all of the information that they need to go away feeling completely comfortable with you using the recordings. If they indicate that they are uncomfortable with you keeping the recording, tell them that you will destroy the recording and all of the information related to it.

References