Acoustic Recording Software for Speech Production Experiments

 http://geoff-morrison.net

Release 2008-12-23	– for dual-monitor systems, different visuals are displayed on the researcher and participant monitors
Release 2008-12-20	– minor upgrade, compiled version now runs under Matlab Compiler Runtime version 7.9
Version 07 .1 4 Aug 2007	– currently a Beta version – should work but not yet fully tested, please report bugs
	– reading in languages_ini.txt file fixed so that any number of columns can be read in (previous version assumed two)
Version 07 16 July 2007	– software generalised to allow user to supply files for any language, including a right to left option; however, Matlab only supports characters in the range 0–255.
Version 06 14 May 2007	– right-channel-only on sound-playback-test now works
	– software does not provide extra-stimulus-set option when extra-stimulus-set repetitions is set to zero
	– round counter added
	– prompt counter now indicates prompt being presented rather than prompts already completed
	– position of between-rounds quit button has been moved so that it no longer hides the counters
	– sample extra-stimulus prompts added for English
Version 05 10 May 2007	– all necessary files are archived in production_experiment_v05_pkg.exe (no zip archive needed)
	– .\RecAnswers\ is created first time software runs (no longer need to include a dummy folder in distributed archive)
	– only using one window (instead of one for word display, one for good recordings and one for clipped) - this should have solved the problems with window focus in which there was no response to key presses
	– visual prompt for "Q" quit option supplied in waveform window
	– "p" and "x" responses now work when recording is clipped
	– session times and number of good recordings are saved to a file
	– Spanish instructions have been proofread
	– fixed problem with round-counting when experiment is resumed
	– timer zeroed immediately before practice experiment
	– documentation revised
Version 04 25 April 2007	– first release distributed for trial and feedback

This software is provided as-is without any guarantee that it will work. I'm willing to give some end-user support, but please read this document thoroughly and follow all the instructions before contacting me.

This software is provided free-of-charge for academic not-for-profit research. Please include appropriate acknowledgments in published papers.

Contents

Description

This software presents written prompts onscreen and records spoken responses. Prompts are presented in random order and the number of repetitions can be selected. The researcher controls the progress of the experiment and monitors the quality of the recording. For each response recording, a raw waveform is displayed onscreen (the speaker does not see the waveform if the software is setup for a dual-monitor system), and the researcher has the option of listening to the recording (the speaker does not hear the playback), accepting the recording, rejecting the recording, or quitting the experiment. If a recording is rejected, the prompt is presented again later in the experiment. This ensures that the required number of tokens are recorded with acceptable quality. Each recording is saved to a separate file using a systematic file-naming convention for easy processing in acoustic analysis software. If the experiment is interrupted no data are lost. If the experiment is resumed the speaker can either complete the missing prompts only or can start again at the beginning. Prior to the experiment, the instructions are presented in written form onscreen, and in audio form from prerecorded readings of the instructions. The speaker also completes a short practice version of the experiment.

Requirements

Matlab code version:

– installed and licenced version of 32 bit Matlab

– installed and licenced version of Data Acquisition Toolbox

– installed and licenced version of Matlab Signal Processing Toolbox

– 32 bit Windows XP

The software has been tested on the 32 bit version of Matlab R2008b running under the 32 bit version of Windows XP. The Data Acquisition Toolbox is not available on other platforms, hence the AcousticRecording software will not run on other platforms. I will not give support for Matlab releases other than R2008b.

Compiled version:

– Matlab Compiler Runtime version 7.9 (this can be downloaded from my ftp site; however, you will need to ask me for the username and password - the licencing agreement doesn't allow it to be posted to a generally accessible location on the internet).

– 32 bit Windows XP

Installing and Running

Matlab code version:

Unzip the files in the archive AcousticRecording.zip into a folder named AcousticRecording.

– To start the software, run AcousticRecording_script.m or AcousticRecording_function.m in Matlab.

Compiled version:

1. Install Matlab Compiler Runtime version 7.9
2. Place AcousticRecording_pkg.exe into a folder named AcousticRecording, and double click on AcousticRecording_pkg.exe.

– To start the software, double click on AcousticRecording.exe. (A shortcut could be made.)

Font:

You must also install the SIL Doulos IPA93 font if it is not already installed on your machine. This is available from http://scripts.sil.org/cms/scripts/page.php?site_id=nrsi&item_id=encore-ipa-download (Alternative phonetic fonts can be used, see languages_ini.txt below; however, Matlab only supports characters in the range 0–225.)

Folders and Files

In addition to a copy of this documentation and the Matlab m files for the Matlab code version or exe files for the compiled version, the folder in which you installed the software contains the a number of files and folders (some pof these will not appear until the firs time you run the software using the "Experiment" option):

The folder InitiationFiles contains a file which is used to set default parameters related to the soundcard (see *Hardware Setup* below), a file to set the default number of repetitions of the prompts (see *Running the Experiment*), and a file to supply information related to the language you are running the experiment in (see *Language Setup*).

The folder Instructions contains a series of text and wave files with English (suffixed _e.txt) and Spanish (suffixed _s.txt) instructions for the participant (see *Language Setup* for instructions on adding more languages). If necessary, you can edit and replace the instructions in the text files. The distributed

wave files were produced using a text-to-speech system. You will probably want to replace them with recordings of the instructions being read by a human with an accent appropriate to your participant population.

The folder Prompts contains the following text files:

stimuli_practice_e.txt stimuli_practice_s.txt
stimuli_core_e.txt stimuli_core_s.txt

The following text files are optional:

stimuli_extra_e.txt stimuli_extra_s.txt

These files contain the prompts which will be displayed on the screen.

The stimuli_practice_e.txt file contains the green text below:

filename	Y	X	no_rhyme	SAMPA	IPA93
h01t	heat	meat	0	hit	hit
h02d	hid	kid	0	kId	kɪd
b03P	bapa	paper	1	bep@	bepə
_04_	"e"	met	1	E	ɛ
h05t	hat	mat	0	h{t	hæt
h06d	hud	mud	0	hVd	hʌd
b07P	boppa	hopper	1	bQp@	bɒpə
_08_	"au"	law	0	O	ɔ
h09t	hote	mote	0	hot	hot
h10d	hood	could	0	hUd	hʊd
b11P	boopa	super	1	bup@	bupə
_13_	"ai"	high	0	aI	aɪ
h14t	hout	lout	0	haUt	haʊt
h17d	hoyd	lloyd	0	hOId	hɔɪd
b18P	burpa	sherpa	0	b3~rp@	bɜːrpə
_19_	"ar"	car	0	Ar	ɑr

Column 1 contains the filename that will be used when the recording is saved (a repetition number and *.wav* extension will be appended). The sound recordings will be written to subfolders within the folder

RecAnswers. A subfolder will be created for each speaker. The subfolder will also contain a timing_info.txt file.

The prompts will be displayed in the form *X sounds like Y* (the prompt paradigm was developed by Dr. Michael Kiefte in a study of Nova Scotian English, it is designed to give an unambiguous indication of the vowel category despite the problems with English spelling).

Column 2 contains the *Y* word, and Column 3 contains the *X* word.

Column 4 indicates whether the *Y* word will appear in red instead of blue. For the stimulus files supplied, this is used in English if the *X* and *Y* words do not rhyme, and in Spanish if the *X* and *Y* words are not identical.

Column 5 provides the SAMPA transcription of the *Y* word (SAMPA is included for reference only, it is not used by the software).

Column 6 contains the transcription of the *Y* word in the SIL Doulos IPA93 font (may look like b p« etc. in a text editor).

You may edit the prompt files to include your desired words etc.. Columns must be separated by tabs. Something must be typed in each column in each row.

To facilitate editing, Excel versions are also included in the folder. Save only the necessary cells as tab delimited text files (do not include the headers when saving the text files or delete them afterwards in Notepad).

Language Setup

The file InitiationFiles\languages_ini.txt specifies language options for the software. The context of the default file are as follows:

```
e                   s
                    en
sounds like         tienes

Next                Siguiente
Play                Escuchar
Stop Playing        Grabar
Stop                Parar
Times               Times
40                  40
.1                  .1
SILDoulos IPA93     SILDoulos IPA93
L                   L
Helvetica           Helvetica
14                  14
```

The first column specifies he settings for English and the second column for Spanish. Additional languages can be added in additional columns. Columns must be separated by tabs.

The first row of each column provides a language identification code. This must be a single character. The same character must appear in the suffix of all the prompt and instruction files for that language. For example, if you add Dutch and use d as the language identifier you must add a column to languages_ini.txt and fill in all the rows in a column headed by d. You must also make Dutch versions of all the prompt and instructions files: Look for all the files in the folders Instructions and Prompts with the suffix _e.txt and _e.wav, and make Dutch versions with the suffix _d.txt and _d.wav respectively. You must make a Dutch version of every file that has an _e.txt or _e.wav suffix, otherwise AcousticRecording will crash (a Dutch version of stimuli_extra_e.txt is optional).

The second through fourth rows of each column specify parts of the carrier sentence to be used. In English the sentence will be *X sounds like Y*, and in Spanish it will be *en X tienes Y*.In general the format is: *A X B Y C* where A is given in the second row, B in the third, and C in the fourth.

The sixth through eighth rows specify text which will appear on buttons on the screen.

The ninth and tenth row specify the name of the font and the font size which will be used to display the prompt sentences.

The eleventh row specifies a margin offset for display of the prompt sentences – a larger number will move the text to the right and a smaller number will move it to the left, negative numbers are allowed, numbers outside the range -0.1 to 1.1 will probably result in some or all of the text being invisible.

The twelfth row specifies the font to be used to display phonetic symbols. If a font other than `SILDoulos IPA93` is selected, then it must be installed on the computer running the software, and Column 6 of the prompt file must supply phonetic symbols using that font.

The thirteenth row specifies whether the language is read from left-to-right L or right-to-left R. The prompt sentences will be displayed in that order and the effect of the margin offset will be reversed.

The fourteenth and fifteenth row specify the name of the font and the font size which will be used to display the instruction sentences.

Limitations on languages

Matlab only supports the display of characters in the range 0–255, it does not support unicode characters above this range; hence, at the present time it is not possible to use the software to display languages like Chinese on a non-Chinese version of Windows. Also one cannot use unicode versions of phonetic fonts unless one can do so without using characters outwith the range 0–255.

There may be some occasional encoding problems even with roman-based writing systems such as Czech. (these may be fixable under the current version of Matlab). But in general the main problem is with non-roman writing systems.

It may be possible to display languages with alphabets or syllabics (e.g., Russian, Hebrew, Cree) if one has access to or can create a font which has transposed all the characters into the range 0–255.

It is possible to display more complex writing systems if you are using a version of Windows localised to the relevant language. I have successfully tested the software on a Chinese version of Windows, and a set of initiation and prompt files for Chinese is available on my website.

If and when Matlab has been upgraded to display the full range of unicode characters, I will upgrade the software.

There may be some occasional encoding problems even with roman-based languages such as Czech. These are probably fixable under the current version of Matlab, so please contact me if you run into such problems.

Hardware Setup

The ideal recording situation is to have the participant inside a soundbooth and the researcher outside. You will need to set up two monitors, one inside the soundbooth and one outside (ideally use a laptop internal screen + external screen, or a dual head graphics card on a desktop, if neither are available use a VGA splitter). A mouse (or other pointing device) is required for both the researcher and participant (you could hook up an extra mouse via a USB hub inside the soundbooth). A keyboard is required for the researcher.

Audio setup

I have tested the two soundcards listed below. Physical hookups are described for these soundcards, hopefully you can figure out what to do if you have some other soundcard.

Sound Playback Devices:	EDIROL UA-25	EDIROL FA-101 Out 1
Physical hookups:		
Soundcard: EDIROL FA-101		
participant's microphone input:	input 1/L	input 1
researcher's microphone input (intercom):	input 2/R	input 2
output to researcher only (left):	output L	output 1
output to participant only (right):	output R	output 2
switches:		
mono:*	out	out
soft control:	–	out
mon. sw:	light on, and adjust volume	–
sample rate:	44100	44100
limiter:	off	–
advance:	off	–

* Change to “in” (monaural) if you want to feed the speaker’s microphone back to their headphones, i.e., if the speakers is wearing isolating headphones which impede their ability to hear their own voice.

The **participant** will need a microphone. The participant will also need headphones to hear the instructions and warning beeps (preferable to a loudspeaker which may produce noise on the recordings).

The **researcher** will need headphones or a loudspeaker. The setup includes an intercom so that the researcher can speak to the participant. For this the researcher will need a microphone. It would probably be best to switch off the researcher's microphone when it is not being used.

In Windows Sounds and Audio Devices in the Control Panel set your default sound playback device to match the device you will be using for the experiment.

To check that your sound playback setup is correct, run the software and select the *Sound Playback Test* option at the start screen

Sounds are recorded assuming that the input is scaled to a peak amplitude of ±1. Some soundcards output lower peak amplitudes. In order for the software to detect when a recording is clipped, it needs to know what the peak range of the soundcard is. The file InitiationFiles\soundcard_ini.txt provides scaling factors for some soundcards. If you have some other soundcard, or want to check the appropriate scaling factor on your setup, run the software and select the *Soundcard Amplitude Calibration* option at the start screen. This application outputs a file InitiationFiles\ soundcard_amplitude_calibration.txt. Open soundcard_amplitude_calibration.txt in Notepad and paste the contents onto a new line in soundcard_ini.txt (make sure that the name of the soundcard and the scaling factor are separated by a tab, not spaces, and that no tabs appear elsewhere).If your system has multiple soundcards installed, and more than one appears on the list, then the software will default to the soundcard which appears first in the list. Please send me your soundcard_amplitude_calibration.txt or soundcard_ini.txt file and I'll add your devices to the next release of the software.

Monitor setup

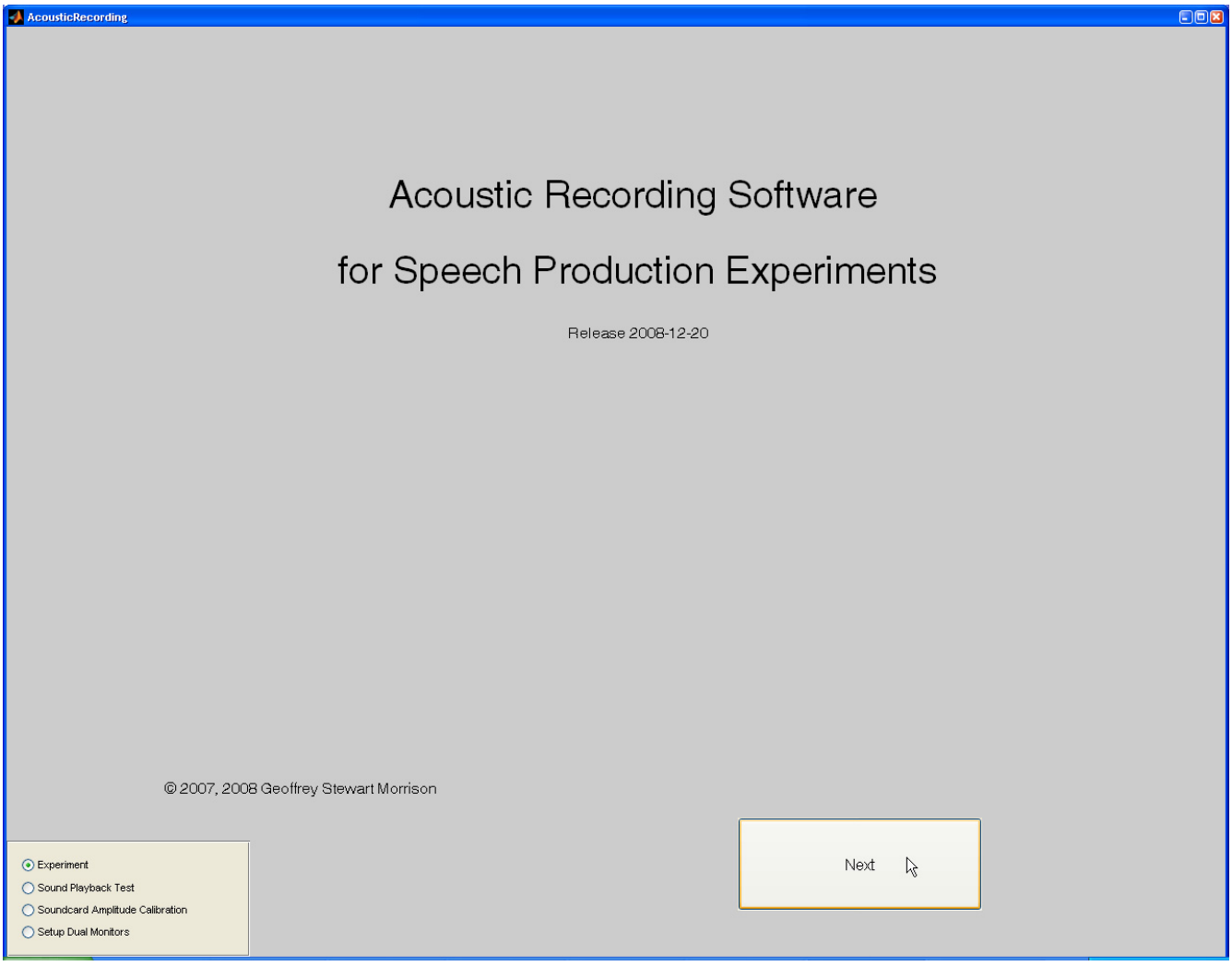
If you have a dual-monitor system, e.g., a laptop internal screen + external screen, or a dual head graphics card on a desktop, set up your system so that the Windows desktop is extended over both monitors: right click on the desktop, select properties, go to the settings tab, click on the second monitor, adjust the size and location etc., and check "extend my Windows desktop on to this monitor" (some laptops use non-Windows monitor controllers).

Run the software and select the *Setup Dual Monitors* option at the start screen. Position the Participant Window on the participant's monitor, and the Researcher Window on the researcher's window. In both cases you probably want to maximise the windows to fill the whole screen. When both windows have been positioned, click Next. The application saves a file InitiationFiles\WindowPositions_ini.mat. If you change your monitor setup, delete this file and run the *Setup Dual Monitors* again.

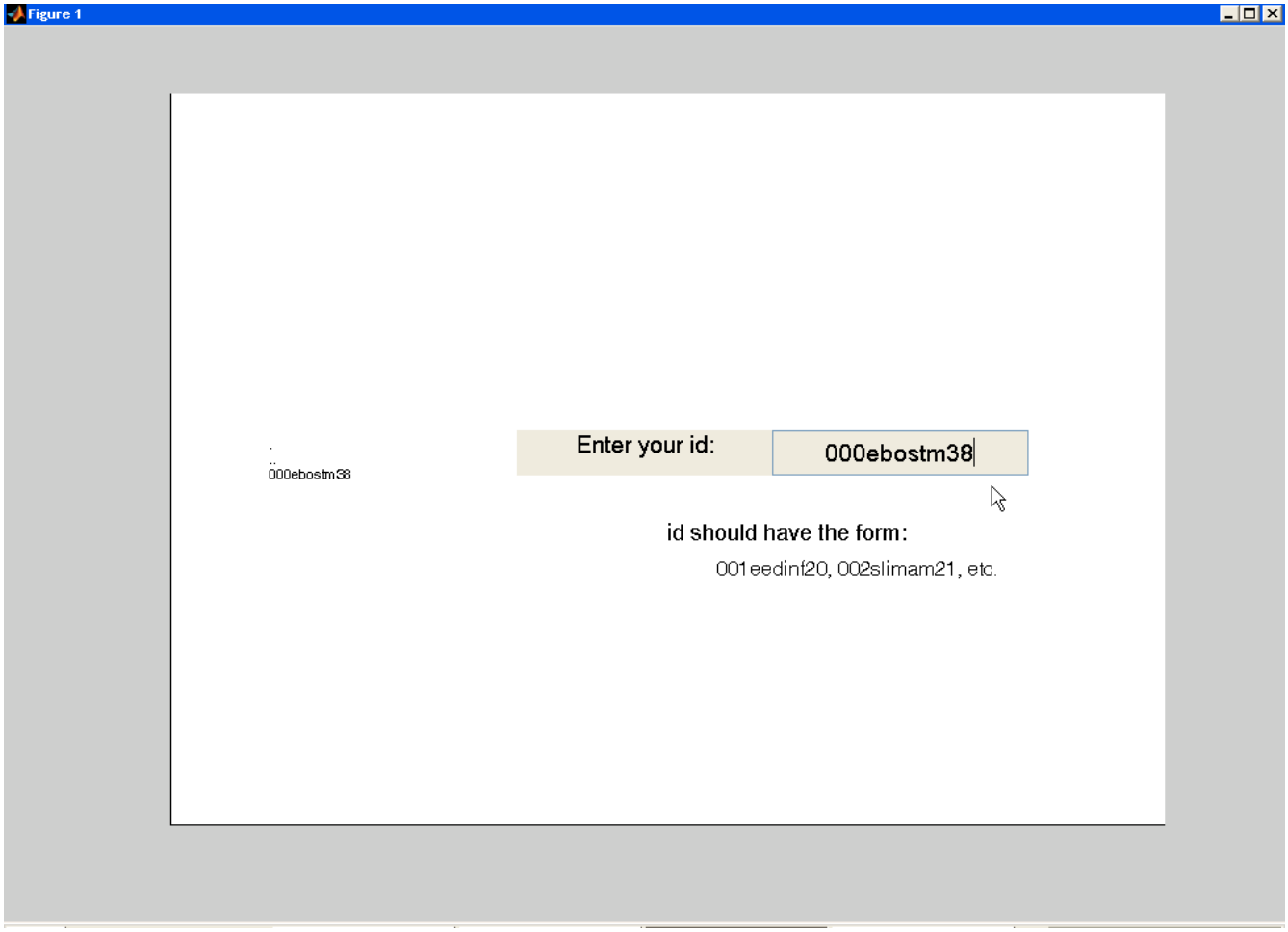
The next time you start the software, separate Researcher and Participant windows will appear on their respective monitors. Entering participant id and experiment options will only appear on the Researcher window. The sound test control buttons will only appear on the Researcher window. The Next buttons during the instructions will only appear on the Participant window. During the experiment, the counters, waveforms, and IPA symbols will only appear on the Researcher window.

Running the Experiment

When you run the software, you should see the following screen (it may take several seconds for the initial screen to appear). Choose the "Experiment" option and click Next.



User id



Click in the input box, input the user id and press Enter. The user id string is checked to ensure that it is legal. If the string does not follow the convention above then nothing will happen. If the string is legal, then the software will proceed.

The full user id should look something like this: 023eedinm38, 045sgastf21

The convention is as follows:

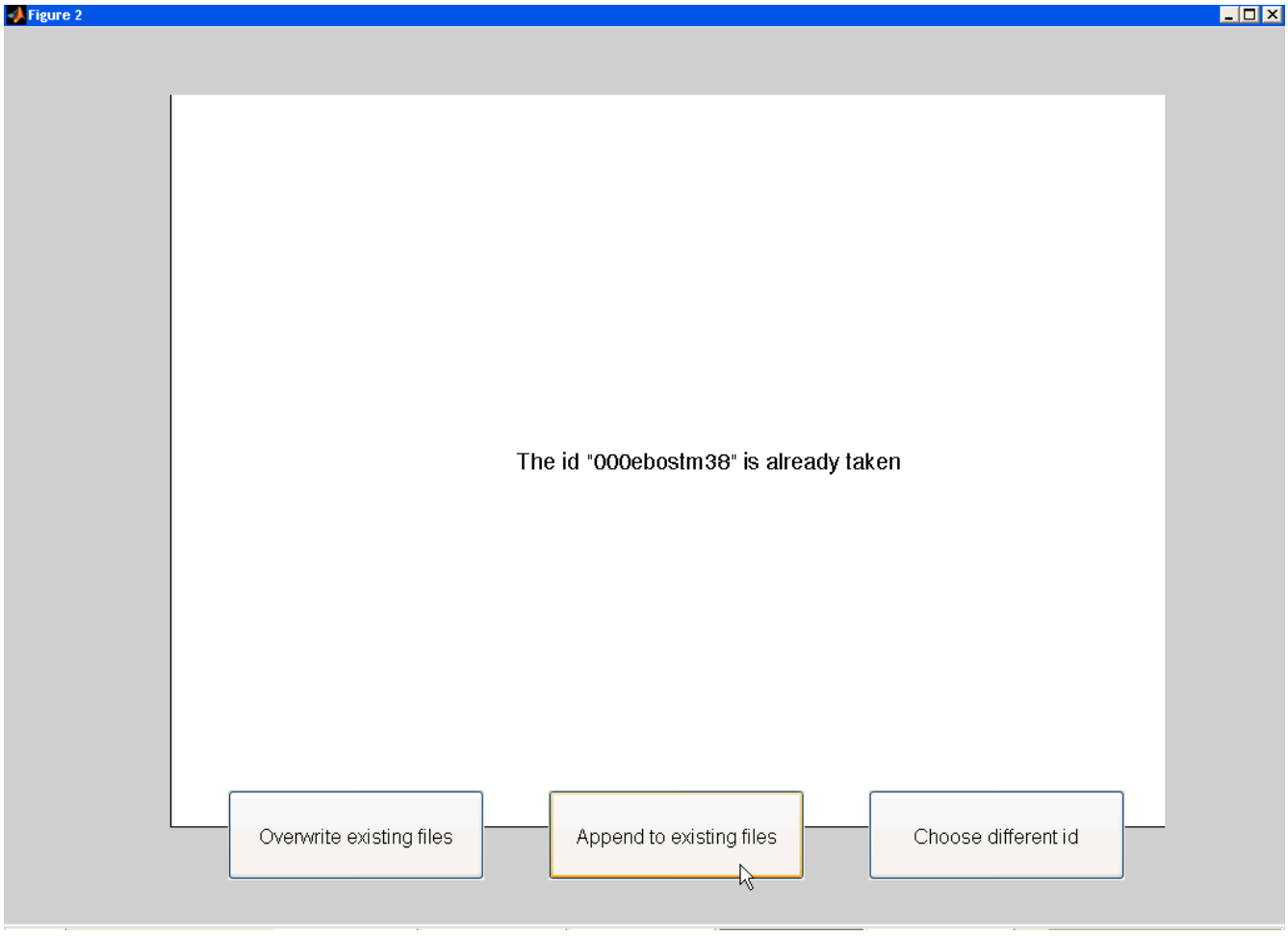
3 digit participant id number:	001, 002, ... 999
1 letter speaker's language:	e = English s = Spanish
4 letter location/dialect:	e.g., canb, edin, edmo, esse, gast, lima, madr, mexi, newc
1 letter speaker's gender:	f = female m = male
2 digit speaker's age:	18, 19, ... 89, 99

I recommend using a unique 3 digit id number for each speaker,

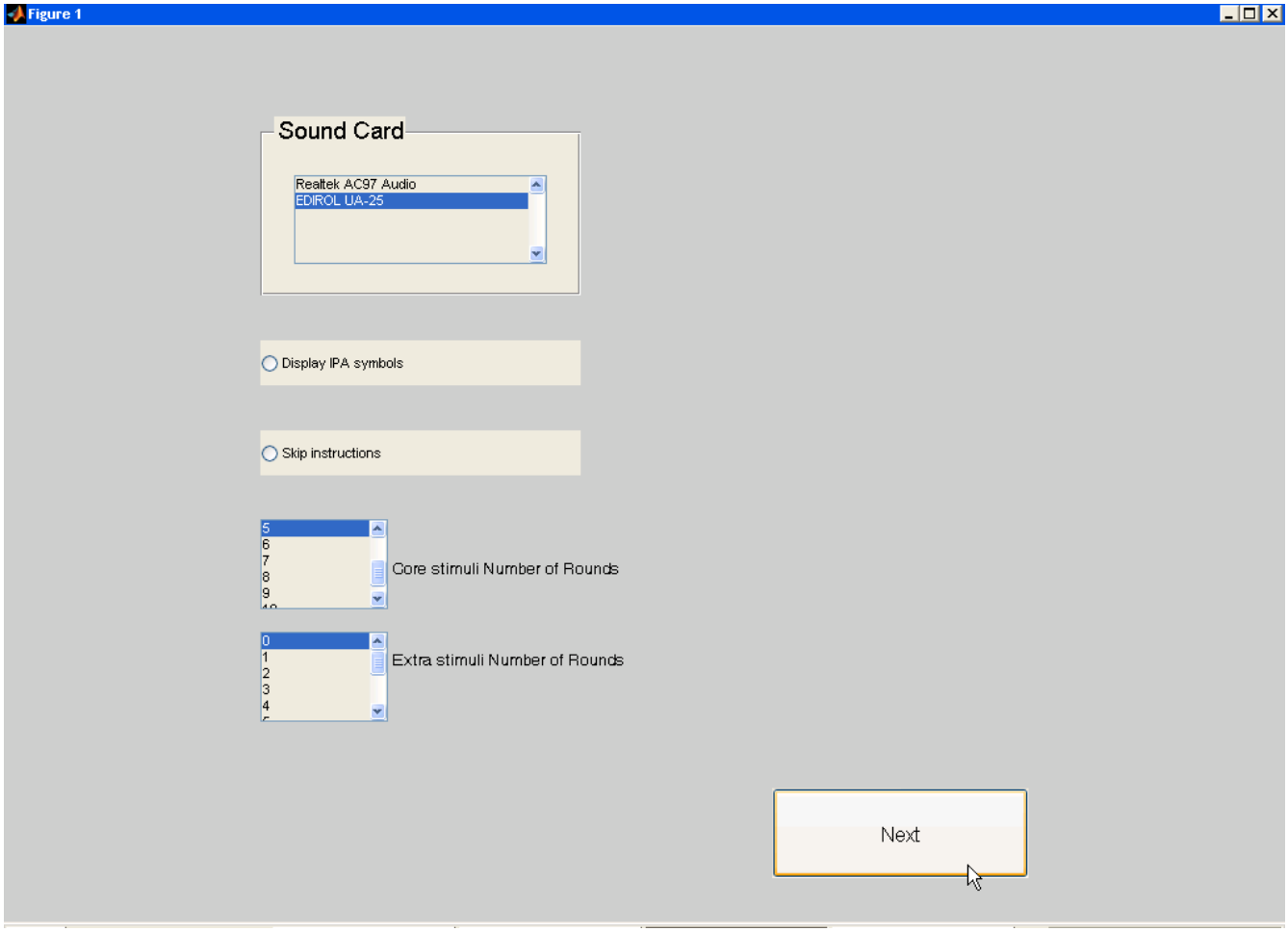
i.e.: 001ecanbm19 002slimaf22

rather than: 001ecanbm19 001slimaf22

If the string has been used before, you will be given the options of entering a different user id, deleting the exiting recordings for that user, or resuming the experiment by appending new response recordings to the existing recordings. Use the latter option if you want to complete an experiment which was previously deliberately or accidentally interrupted.



Experiment options



– select a soundcard (only soundcards installed on your machine are shown, the highest listed available soundcard in InitiationFiles\soundcard_ini.txt is the default)

– display IPA symbols for the speaker (default no)

– skip instructions and practice experiment (default no)

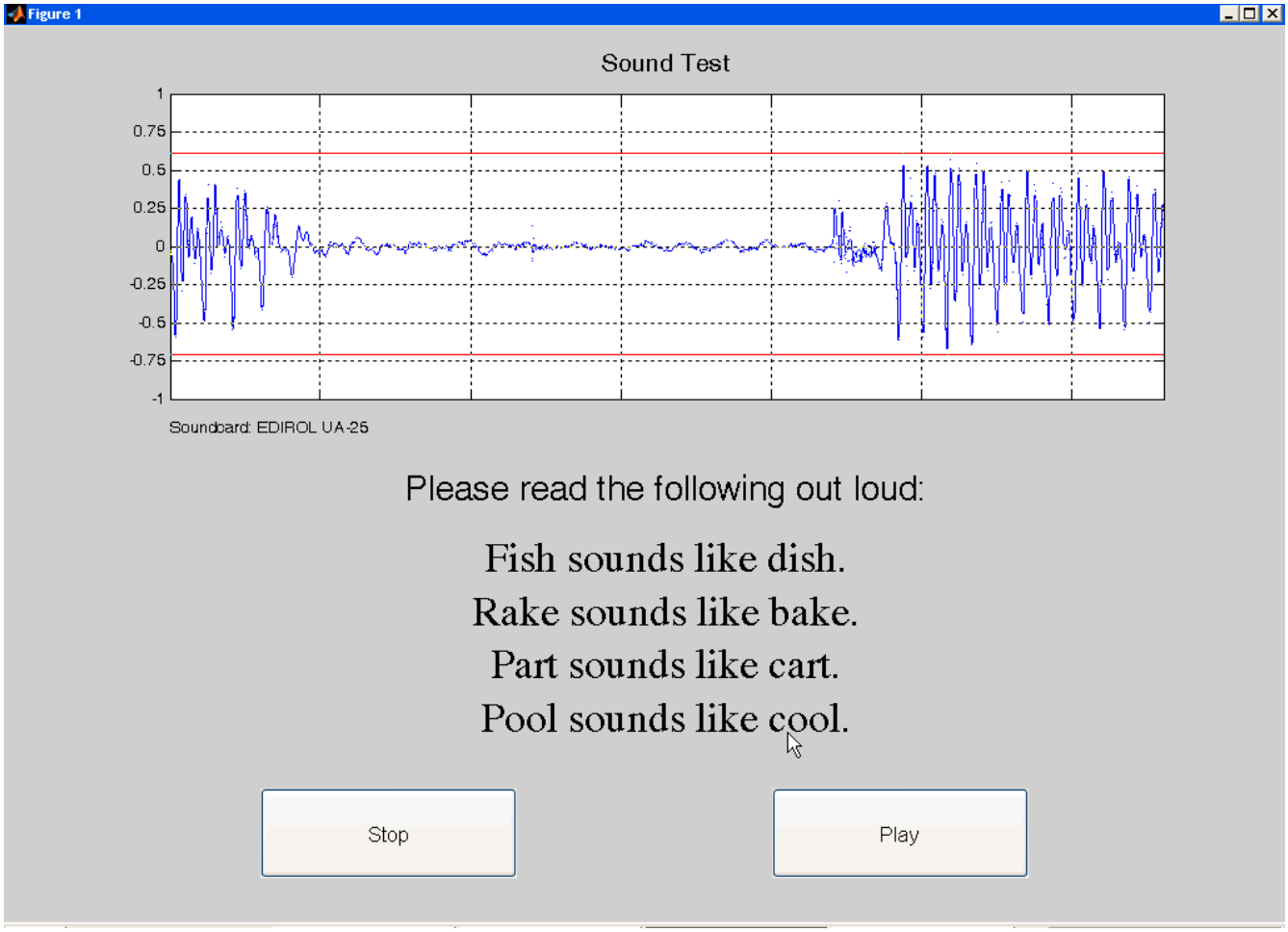
– choose the number of repetitions for the core stimulus set

– choose the number of repetitions for the extra stimulus set

(default repetition values are set in the InitiationFiles\repetitions_ini.txt file)

Choose your desired options then click Next.

Sound test



Get the speaker to read the sentences while you adjust the microphone gain on your soundcard. Try to set peak volume around ±0.75. The Play button allows you to replay the last few seconds of sound. Click on it again to return to live mode. Click on the Stop button when you are happy with the sound level.

Instructions and practice

– Let the speaker go through the instructions and practice experiment.

– Be available to answer questions and monitor the practice experiment.

During the experiment

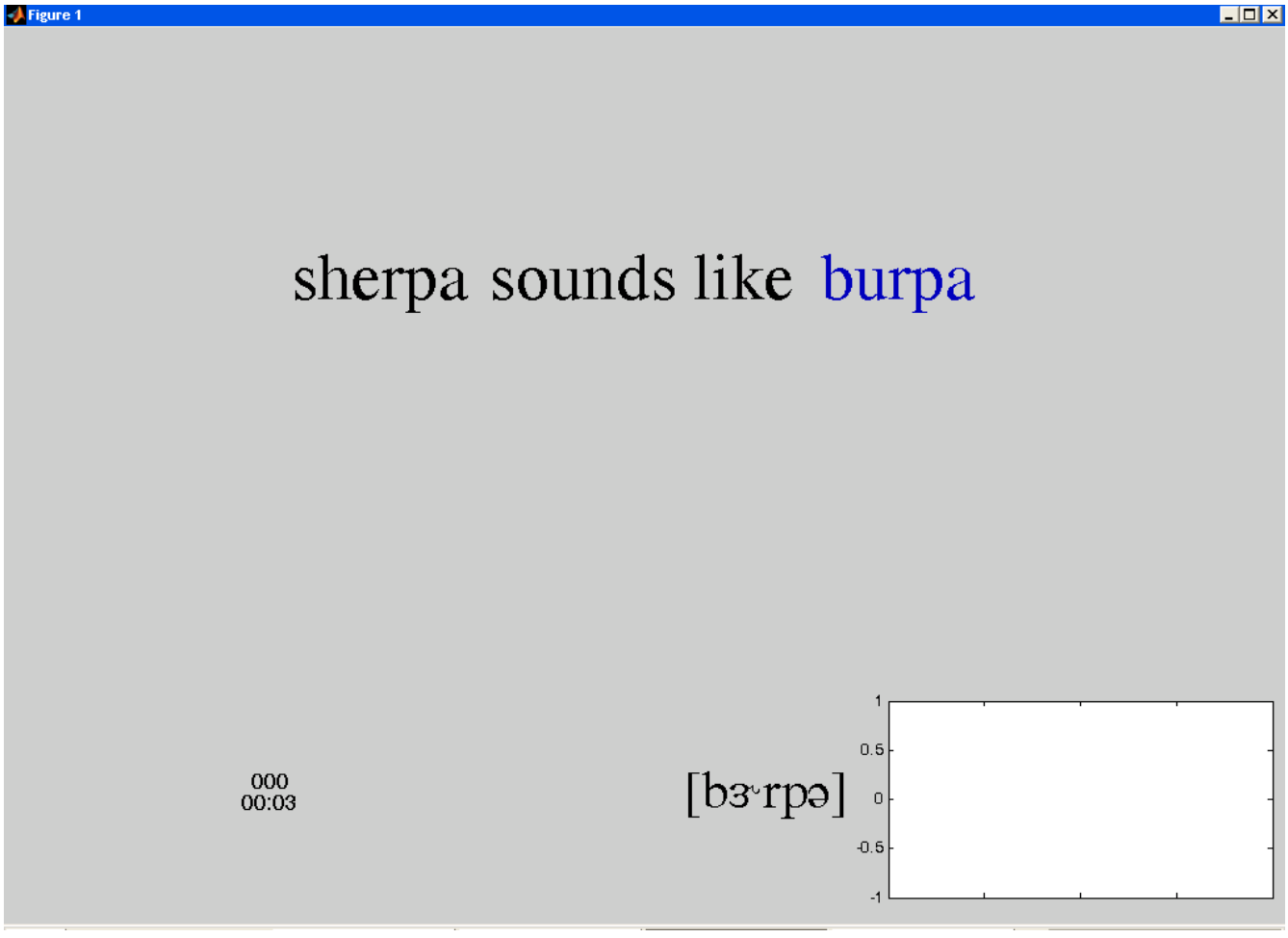
Proceed with the experiment proper and monitor the recordings.

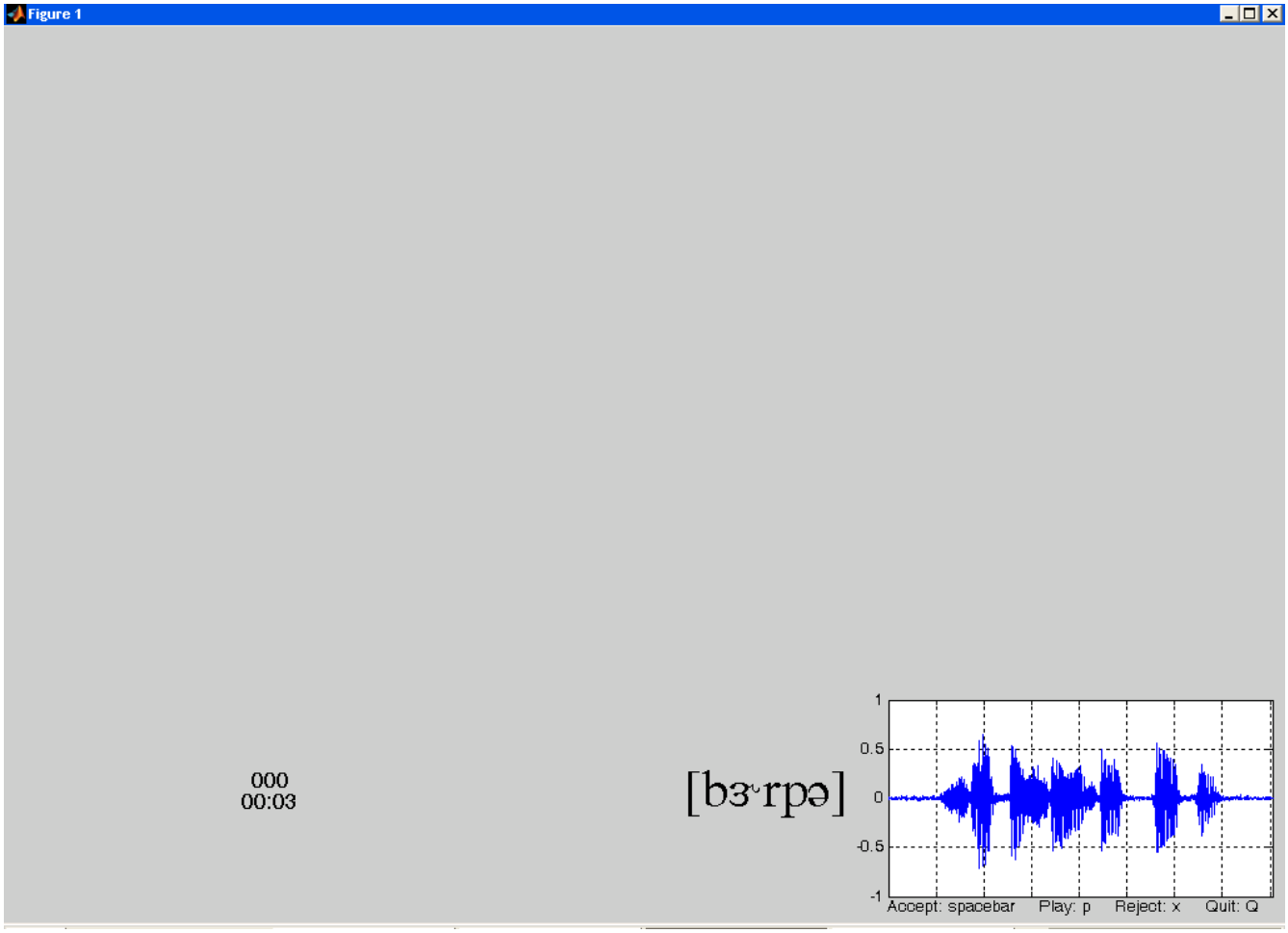
– Both the researcher and speaker will hear a beep.

– Both will see a prompt sentence displayed on the screen, and the software will start recording sound from the speaker.

– The speaker will read the sentence out loud (the researcher will hear the speaker)

– The researcher will press the spacebar to stop the recording (there is also a 20 second timeout).





– A raw waveform will be displayed in the bottom-right of the screen. If you have dual monitors set up, only the researcher will see this.

– The researcher has the following options:

– to play the recording press p (only the researcher will hear the recording)

– to accept the recording press a or the spacebar

– to reject the recording press x

– to quit the experiment press Q (must be capital)

– If the recording is clipped or timed out, then a message to that effect appears over the raw waveform

– pressing the spacebar rejects the recording rather than accepting it

– After a short pause, there will be a beep and a new prompt will be displayed.

The software is designed so that in most instances the researcher presses only the spacebar.

Possible reasons for rejecting a recording:

- clipping during target word
- background noise, coughs etc.
- stutters or other disfluencies
- incorrect reading of target word (the *X* sounds like *Y* procedure is not foolproof)

The speaker is encouraged to take a rest between recording rounds (repetition sets). They may take a rest at any other time. If the researcher does not press a button when the raw waveform is displayed, then the software will wait. For long experiments, it may be a good idea to provide the speaker with a bottle of water.

Troubleshooting

Make sure Caps-Lock is off.

If you need to quit in the middle of an experiment, it is recommended that you wait until the waveform window appears then press "Q", or press the "Quit" button between rounds. This allows the software to run some clean-up activities.

The experiment has been tested on screens with a size of 1024 × 768 pixels. If you are using another screen size and are having problems with the display, try setting the screen size in Windows to 1024 × 768 pixels.

If there are errors in Column 1 of the prompt file, e.g., the filename has 3 instead of 4 digits, then the software will not behave as expected when an experiment is resumed.