AALTO UNIVERSITY

SCHOOL OF SCIENCE

Special assignment

Engineering physics and mathematics

Tfy-99.5111 Biophysics and Biomedical Engineering

2.8.2011

Measurement of acoustic and anatomic changes in orthognathic surgery patients

Tiina Murtola

231206

Plan for special assignment

Työn nimi Akustisten ja anatomisten muutosten mittaaminen ortognaattisen toimenpiteen läpikäyneissä potilaissa

Name of the work Measurement of acoustic and anatomic changes in orthognathic surgery patients

Background:

The motivation of this work is data acquisition for mathematical modelling of human speech production. Our main interest is in vowel production where the acoustics of the vocal tract plays the most important role. For data acquisition, an experimental arrangement has been developed at the Institute of Mathematics in cooperation with several partners (such as the phonetics laboratory of University of Helsinki and the MRI unit in TUCS) in 2007- 2010. This arrangement makes it possible to carry out high-quality sound recordings simultaneously with MR imaging of the vocal organs. The pilot data has already been obtained in June 2010, and the measurement protocols are now to be improved for clinical use.

To obtain the first clinically relevant data set, the patients undergoing an orthognathic procedure have been chosen as test subjects. Orthognathic surgery deals with the correction of abnormalities of the facial tissues. The underlying cause for abnormality may be present at birth or may be acquired during the life as the result of distorted growth. The patients undergoing orthognathic surgery are an excellent study group because we can obtain recorded speech data for acoustic analyses and MRI data of vocal tract anatomy before and after the operation. The changes in anatomy are easy to quantify since only hard tissues are involved in operations. In addition, the patients are mostly young and without any significant underlying diseases.

Prof. R.-P. Happonen (Dept. of Oral Diseases, TUCS) is responsible for the medical aspects of the project as well as the ethical permits. Dr. J. Saunavaara (TUCS, Medical Imaging Center) takes care for the questions involving MRI. Dentistry related aspects are the responsibility of Mr. Jean-Marc Luukinen (TUCS).

What should be done?

The proposed work should include experiment planning, coding, and documentation.

A signal generator software has to be created to produce a cue signal for the patient during the experiment as well as control signals for the MRI machine for timed operation. So as to the experimental arrangements, a number of phonetic aspects must be taken into consideration in cooperation with the phonetics experts at the University of Helsinki (Dr. Martti Vainio). Experimental protocols and meta data questions have to be treated so as not to mess up the experiments that may extend over two years of time and over 20 patients. Signal processing tools must be developed for fast inspection and reliable artefact detection during the patient measurements. Finally, the solutions must be tested using a healthy subject in the MRI laboratory of TUCS.

Methods and tools:

The MR imaging is takes place at TUCS. The simultaneous sound recording is carried out by the experimental arrangement developed in this project earlier. Matlab is used for signal processing and all software development. The Department of Mathematics and Systems Analysis will provide all the necessary resources for this work.

Date: 20/12/2010

Instructor: Dr Jarmo Malinen, Senior Assistant in Mathematics, Aalto University

Name of the student: Tiina Murtola *Student number:* 231206

E-Mail: tiina.murtola@aalto.fi

Previous and currently on-going special assignments (code, title and type):

Course code and topic area: Tfy-99.5111 Biophysics and Biomedical Engineering

Date:

Supervisor of the course: ARI KOSKELAINEN

Abstract

This assignment fine-tunes an existing experimental arrangement for simultaneous collection of sound and magnetic resonance imaging data. Phonetic materials are re-designed to instruct and support test subjects in sound production. A multichannel audio interface and a custom-made device allowing external triggering of the MRI machine are added, and a Matlab-based system is developed to run and control experiments. Data collection procedures, including meta data collection, are integrated to the control system.

Table of Contents

Abbreviations and symbols

e_d	discretisation error
f_0	target frequency for sound production
f_s	sampling frequency
MRI	magnetic resonance imaging
WAVE	Waveform Audio File Format, also known as WAV

Definitions

de-noised signal	difference between collected patient-in and noise signals produced by a differential amplifier
formant	meaningful frequency components of human speech produced by resonances in the vocal tract
key	a matrix for extracting pure samples from recorded sound
orthognathic surgery	surgery to correct conditions of the jaw and face
patient-in signal	sound signal collected from the patient
patient-out signal	sound signal to patient headphones which includes the de-noised signal, cue signal and signal from control room microphone
Playrec	a Matlab utility (MEX file) for using multichannel soundcards (www.playrec.co.uk)
pure sample	a recorded sound sample where the MRI noise is either absent or sufficiently low to allow the phonetic features of the patient sound to be analysed
session	a series of experiments (measurements) carried out with one subject/patient on one day
test phase	measurements done with one test subject to test the developed system and to resolve any remaining issues

1. Introduction

1.1. Background

A computational model of human speech production based on the resonance characteristics of the vocal tract (such as discussed for example in Hannukainen et al. (2007) and Aalto et al. (2009)) would be useful, for example, in predicting and planning the effects of surgical operations in the facial area on speech. To validate such models, a coupled data set of sound and the vocal tract anatomy which produced the sound is need.

An experimental arrangement to collect such data using *magnetic resonance imaging (MRI)* has been developed (Lukkari et al. 2007; Malinen and Palo 2009) and pilot experiments have been carried out confirming the feasibility of the arrangement (Aalto et al. 2011). The next step is to obtain a clinically relevant data set. Measurements are carried out before and after operation for 10-20 patients undergoing orthognathic surgery. Orthognathic patients are an excellent study group because the surgical operations involve only hard tissue in the facial area and because they are mostly young and healthy adults.

1.2. Specification

The aim of this project is to fine-tune the experimental arrangements. The existing methods allow data collection but are, as such, unsuitable for use with a large number of patients. Improvement are made both to the phonetic aspects of the experiments and to the hardware, software and procedures used to carry out the experiments. The MRI and sound collection arrangements remain the same as in the pilot experiments (see Aalto et al. (2011) and the next sections).

Changes in the phonetic materials address improvement suggestions based on the pilot experiments (Aalto et al. 2011). The overall the improvements aim to ensure that the articulatory task for each measurement is clear and within the ability of the patient, and that the timing and quality of the sound produced allows data collection with few repetitions.

New hardware and software are added to the measurement system to increase the number of sound samples collected in each experiment. An external trigger is used to introduce breaks in the MRI sequence which allows collecting samples with low levels of MRI noise (i.e. *a pure sample*), checking quality of articulation throughout the experiment and capturing time dependent articulatory features.

Using such a system requires timed operation of cueing, triggering of the MRI sequence and recording. Timing of the control signals must ensure that sound production is stabilised before the start of the imaging, and that pure samples are available before and after, as well as during the imaging. Furthermore, it must be possible to locate the pure samples in the record.

Other requirements on the fine-tuned measurement system are that it (1) is simple to use even without a technical background; (2) reduces human errors by minimising the number of programmes/tasks the user must handle simultaneously; (3) standardises the running of experiments to allow easier comparisons; and (4) creates and stores meta data for each measurement.

The fine-tuned arrangements will be tested with one healthy subject in a *testing phase*. Due to hardware and time considerations the testing has been left outside the scope of this assignment.

2. Methods and tools

Software development and signal processing is carried out using Matlab 7.11.0.584 (R2010b) on Linux Ubuntu 10.04 LTS. Sound collection is carried out using the experimental arrangements detailed in Lukkari et al. (2007) and Malinen and Palo (2009). Playrec, a Matlab utility (a MEX file), QjackCtl JACK Audio Connection Kit version 0.3.4, JackEQ 0.4.1 and M-Audio Delta 1010 PCI Audio Interface were used to achieve desired functionality with the audio signals.

In the testing phase, MRI will be carried out using the arrangements detailed in Aalto et al. (2011). In particular, the 1.8 mm voxel sequence is used for which the number of triggered slices is 35 due to accelerated acquisition. External triggering of the MRI sequence is done using a custom built trigger box which is controlled with audio signals.

3. Solutions

3.1. Phonetic aspects

The phonetic aspects of the experimental arrangements can be divided into two parts: instructions and practice for the subject prior to measurements, and the cue signal for guiding sound production during measurements. Instructions are needed to guide the subject in physiological functions which may affect sound production. Practice prepares the subject for the measurement and allows checking that the task has been understood correctly. The cue signal helps the subject to find and maintain the target frequency level, f_0, and guides the subject in timing.

In the pilot experiment, no standardised instructions or practice were used as the test subject was familiar with the procedure. The cue signal consisted of 3 s of sine wave at f_0 and three count-down beeps based on gammachirp impulse response. The phonetic materials were improved according to the observations made during the pilot measurements, for example to allow time for swallowing and inhaling and to improve the audibility of the cue signal (Aalto et al. 2011). The unfamiliarity of patients or new test subjects with the measurement procedure was also taken into account.

Instructions and practice

The proposed instruction set can be divided into three parts following Aalto and Vainio (2011).

1. *Orientation:* The subject is told to breathe deeply and, if necessary, to swallow. The frequency of deglutition is 200-1000 times a day including eating and sleeping (Lear et al. 1965) and may increase with dryness of mucous membranes in the mouth, so an opportunity to swallow is necessary between the measurements.
2. *Task and practice:* The subject is told which syllable to produce, provided with a sample of $f_{0,}$ and given a few seconds to practice.
3. *Instructions for countdown:* The subject is reminded of the cue structure: how many beeps to expect, when to breathe in and when to start vowel production. The instruction for breathing is essential as good quality vowel production can only be maintained long enough (10 s or more) if the subject's lungs are full at the beginning of articulation.

Cue signal

The cue signal prompts the subject to (1) start sound production, (2) maintain sound production until all pure samples are recorded, and (3) maintain consistent sound production. Both the time and the frequency structure of the signal were changed from those used in the pilot experiments to better achieve these goals, in particular (2) and (3).

The new count-down contains five beeps with f_0 starting on the fourth pulse. The subject is instructed to begin vowel production on the fourth, allowing more time to inhale. Because sound recording is done using a separate channel, the cue signal can be played to the subject throughout imaging, helping to maintain f_0 and telling the subject when sound production can be ended. This leaves 2 s for stabilisation and sample collection before MR imaging begins.

In the pilot experiment, two f_0 (110 and 137.5 Hz) were used in order to study the vocal tract shape with different larynx positions (Aalto et al. 2011). MRI noise interfered with the higher frequency, however, and so f_0 for males were changed 100 and 125 Hz, and females to 160 and 200 Hz.

A pure sine wave was found to be problematic as it is difficult to hear over the MRI noise at low frequencies (Aalto et al. 2011), and interpreting and mimicing the intended f_0 from it may be difficult. Possible alternatives include triangular waves, neutral vowels, and synthesised target vowels. In the case of synthesised target vowels, it must be checked that there is no tendency of the subject mimicing the formants in the cue signal.

3.2. Measurement system

Figure 1 illustrates the experimental arrangements, in particular the audio connections. The control system, consisting of a signal generator and a recorder (shown separately in Figure 1), operates in Matlab sending and receiving audio signals via the mixer and the audio interface.

The sound collection arrangements have been described in Lukkari et al. (2007) and Malinen and Palo (2009). A differential amplifier performs a subtraction of the collected patient-in and noise signals producing a *de-noised signal*, the primary data signal. All of three signals are recorded and available on the audio interface output lines for listening in the control room.

The mixer performs a weighted addition to obtain the patient-out signal and controls the volumes of the patient-out and MRI trigger signals. The patient hears their own de-noised voice added to the cue signal with a delay of about 90 ms (see below). The MRI trigger signal is made into square pulses in the trigger box before passing to the MRI machine.

User interface and documentation

The control system is constructed as shown in Figure 2. The user will be required to deal with one Matlab file containing the parameters to define the experiment, such as the target vowel and f_0. In addition, during the run of an experiment, dialogue windows will ask the user for input and provide the means to abort the experiment if necessary.

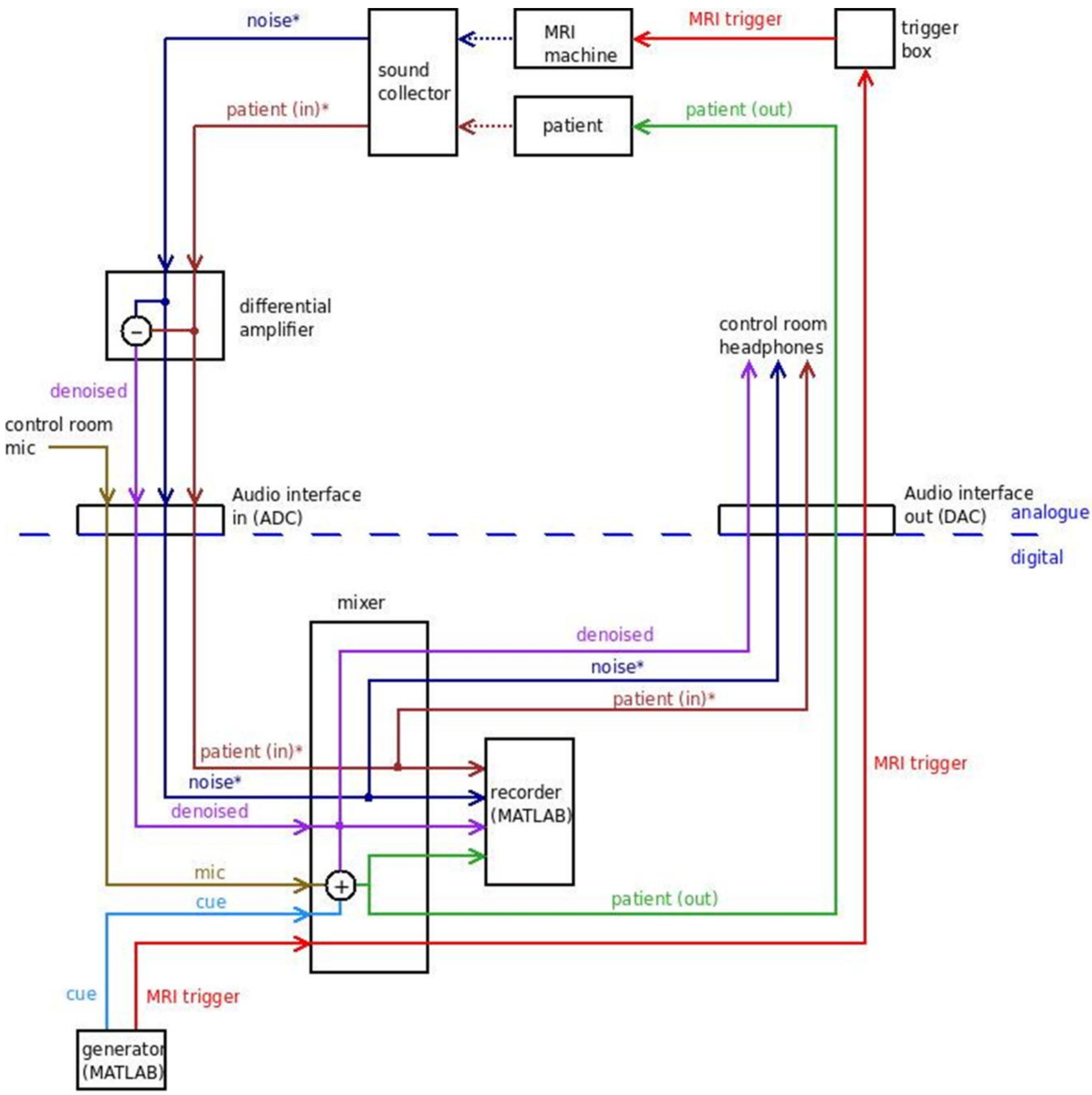

Figure 1: Measurement system - audio connections

*Patient (in) /noise channel includes two empty lines from the sound collector

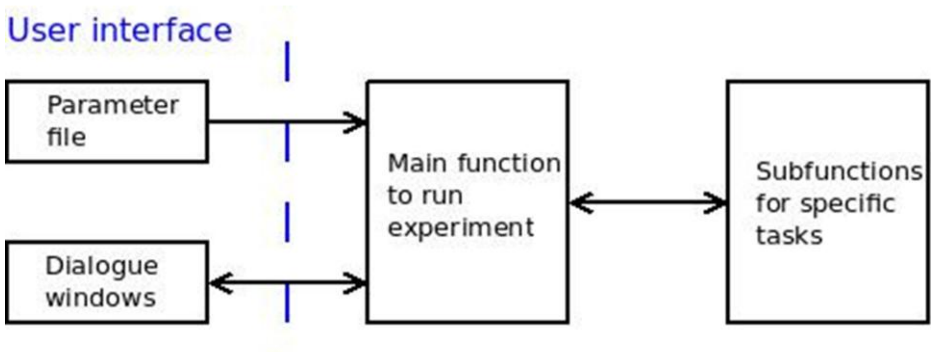

Figure 2: Control system structure in Matlab

In addition to the assignment of parameter values, the parameter file contains comments and a call to the main function. In the testing phase, the comments divide the parameters into groups as shown in Table 1. In patient measurements, all parameters except f_0 and vowel are constant and will be removed from this file. The user only needs to adjust these two assignments and to save and run the file which removes the need for previous experience with Matlab.

Parameter group	Parameter	Parameter type
cue	f_0	integer
	vowel	string
	cue type	string
trigger sequence	number of slices before each break	row vector
	duration of breaks	row vector
	duration of default pause	floating point number
sample extraction	echo time in MRI room	floating point number
	sample duration	floating point number
	uncertainty[1]	floating point number
	delay	floating point number

Table 1: Experiment parameters in the parameter file

The main function checks the validity of the input parameters either explicitly (for example, if breaks are long enough to encompass echo time, sample time and uncertainties) or implicitly (for example, reading the cue file determined by the cue parameters). Human errors, such as incorrect choice of the number of MRI slices and breaks, are checked for by asking the user for confirmation.

In addition to Matlab, the user must also deal with the mixer to determine output volumes. Four different sliders may need to be adjusted during an experiment (MRI trigger and patient-out volumes and the weights of the de-noised and cue signals in patient-out). Sufficient MRI trigger volume can be determined by checking lights on the trigger box. The balance and volume of the patient-out signal must be done based on post-experimental patient feedback and hence may prove to be challenging. If more than one experiment is needed to set the volumes or if fluctuation in the volume of the patient sound production necessitates continuous adjustments, the patient-out signal can be copied to another output in the audio interface, so that adjustments can be done during

1 Time (ms) added to each break so that regardless of uncertainties, there's always a sample available during a break

experiments.

Correct settings, including audio connections, must be determined at start-up in both JACK and Matlab. In Matlab, these are done with a start-up file which Matlab runs automatically during its starting operations. This file sets up necessary paths, changes the current folder, initialises Playrec, opens the parameter file, and initialises session data. In JACK, start-up creates default connections which are unnecessary for the experiments but do not affect the system performance unless currently unused ports are taken into use. The essential connections are made automatically using an active Patchbay profile. An illustrated guide has been written to help the user if the default connections need to be removed. The guide also shows how to clear all connections and build the connections from scratch if setting up goes wrong.

The constructed system has been documented both in comments in all the Matlab files and functions and in a User's guide. The User's guide contains the aforementioned audio set-up guide, guide to using the control system in Matlab, and a troubleshooting guide.

Timed operations

The requirements set for the timing of tasks would be met, ideally, by synchronisation the three audio signals used in the experiment: the cue signal, the MRI trigger signal, and the incoming sound signal. However, perfect synchronisation cannot be achieved when the signals are subject to different delays and uncertainties. Hence the sequence of tasks was constructed in such a way that moderate delays between the onsets of the different signals can be tolerated. The key feature in this implementation is the ability to estimate an upper bound for all possible delays and uncertainties.

Figure 3 represents the timing of the audio signals in Matlab. To keep the audio matrices a manageable size, the instructions and practise signal is played separately to the patient. The MRI trigger signal is buffered with blocks of zeros to time it correctly with the cue signal. Recording is started simultaneously with the beginning of the cue and trigger signals and contains equal number of samples. The output signals and the corresponding inputs will be delayed by different amounts by the time they are recorded (Figure 3b).

Delays

Estimates of the delays between the reference timing (Figure 3a) and the timing of the recorded signals (Figure 3b) are needed to ensure sufficiently long recording and to locate the beginning of

MRI noise in the recorded signals.

The delays in the system in Figure 1 were estimated in two parts: (1) the digital part of the system and the audio interface, and (2) the audio wave guides in the upper half of the system. Any delays caused by the patient, MRI machine, trigger box, or differential amplifier are assumed to be insignificant. The first estimate was done by comparing recorded signals with the input signals using two mixer settings: (1) connection via on slider (MRI trigger) and (2) connection via two sliders, i.e. weighted addition (patient-out). The second part was estimated using speed of sound in room temperature.

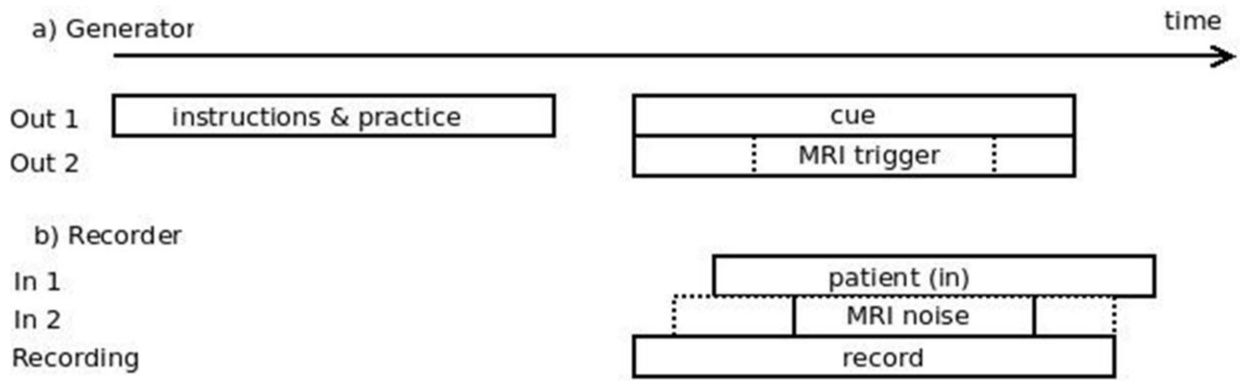

Figure 3: Signal timing (a) as seen in the signal generator and (b) as seen in the recorder

Given the results in Table 2, MRI noise is recorded with a delay of 59±1 ms and patient-in signal with 91±1 ms. To account for these delays 100 ms was added to the duration of recording by increasing cue signal duration. It should also be noted, that moderate unanticipated delays (less than ~450 ms) will, at worst, reduce the number of pure samples per experiment by one.

System part	**Arrangement**	**Total delay (ms)**	
		Mixer setting (1)	Mixer setting (2)
Matlab – mixer – audio interface – mixer – Matlab	output of the audio interface connected to its input	50[2]	73.3
Audio wave guides	3 m both to patient and from sound collector	9 (each direction)	

Table 2: Delays in the measurement arrangements

2 Uncertainty a few tenths of milliseconds

Pure sample extraction with triggered imaging

Ideally, all recorded audio data could be used by removing the MRI noise. A suitable noise cancellation technique is yet to be found, however, so breaks are inserted in the MRI sequence to increase the number of usable samples. When the breaks are long enough for noise in the MRI room to fade, the recorded sound will be pure enough to allow formant extraction and other data analysis.

The MRI sequence is triggered using an external, custom made device. 1 kHz sine signal input to the trigger box produces output at _TTL_ level one. When the output is divided into pulses lasting at least 10 ms (MRI machine specification) and separated by pauses of at least 225 ms (in the sequence used), each pulse can be used to trigger a MRI slice.

The pure samples need to be extracted from the recorded sound. This is done by calculating a *key* which contains the beginning and end times of the samples. Due to uncertainty in the timing of the measurements, the key does not extract all available pure sound from the record. Instead, the breaks in the imaging sequence are made long enough that the samples specified by the key always fall within a break.

Delays and uncertainties

The delays in the control of the MRI machine affect the key calculation whereas the delays in the patient-in signal do not. This is assuming that sound production starts approximately 2 s before MR imaging so that delays of a few hundred ms are negligible in comparison and will not affect sound quality around the last 100 ms before imaging. In addition to delays, the system also introduces uncertainty, i.e. variable shifts in either direction, in the timing of the signals – both in the delays and in the trigger signal. From the discussion above, 1 ms is enough to capture the uncertainty in the delays.

The output of the trigger box at sampling frequency (f_s) 44.1 kHz was measured, and the pulse duration was found to be 11.6±0.4 ms and the pause duration for 240 ms specified was 237.4±1.4 ms. Furthermore, the pulse and pause durations varied in both directions within every sequence so that accumulation of the errors is not likely. For a 35 slice sequence an extra 70 ms in each break allows the key to capture the samples regardless of the uncertainty. When the location of the breaks has been finalised in the test phase, this estimate can be recalculated for each break separately which will reduce the total measurement duration.

Discretisation error, e_d, in the trigger signal is accounted for by the aforementioned measurements. This error depends on f_s (Equation 1) and so the duration measurements should be repeated if it is changed. The impact of incorrect estimation of e_d is only a few ms but average pulse and pause durations may be affected by changes in f_s and errors in these accumulate.

$|e_d| < (2m + 1)/f_s$, where m is the number of slices. (Equation 1)

Data collection

The aim of the experiments is to collect a set of MR images and sound. In addition, meta data needs to be created in order to place the primary data in the correct experimental context. The control system deals with the sound and meta data collection. Acquiring and storing the image data is done by the MRI system.

Sound data is saved automatically as in Waveform Audio File Format (WAVE) with an inbuilt Matlab function. When an experiment is completed, the name and location of the WAVE file is returned to the user. This allows quick and easy access to the sound data for fast inspection using the same functions which are used for full data analysis. The need to re-read the data could be eliminated, and hence faster access achieved, by passing the data matrix to the user as such but this would require re-writing the analysis functions to recognise matrices as input.

All experimental parameters are saved automatically in a text meta data file. This is faster and less prone to errors than having the user type in or copy-paste the parameters into the file. The user is able to add “Other information” in the file during the process and to open the file immediately for viewing and editing. No restrictions are placed on editing the file. This creates an opportunity for accidental deletion of the file or its contents but it is necessary at least in the testing phase when unanticipated situations may arise. The user must be made aware of this risk.

If the option to edit the meta data file is not removed for the patient measurements, paper forms complementing the electronic records would eliminate the hazard of accidental deletion of the meta data files or their content. Furthermore, in the patient measurement phase, only the fields shown in Table 3 need to be filled in. Extra time per patient needed is small and mostly overlapping with other activities (e.g. measurements). Nevertheless, the records should be kept primarily electronic to keep them as accessible as the primary data.

Data type	Variable	Frequency	Field type[3]
session data[4]	patient name	once per patient	open
	date		open
	names/initials of people present		open/multiple choice
	other information		open
experiment parameter	f0	once per experiment	multiple choice
	vowel		multiple choice
	other information		open

Table 3: Record fields in patient measurements

4. Conclusions

The testing of the system was removed from the scope of this assignment due a delay in the production of suitable hardware. The testing phase is needed for assessing the overall functioning and practicality of the system and for addressing the following questions.

1. *Cue type:* Possible cue types are triangular waves, neutral vowels and the target vowel. The chosen cue type must be audible over the MRI noise, convey the correct f_0 to the subject but not cause mimicing of formants.
2. *Number, location and duration of breaks in the imaging sequence:* The number of breaks and their duration must be balanced against an increase in the duration of sound production. In addition to the sample time, other parameters affecting break durations, namely those to do with echo time and uncertainties, must be confirmed.
3. *Other parameters:* The values of the parameters affecting key calculation should be confirmed, especially delays. The accuracy of the key should be tested and adjustments made until it finds the pure samples reliably.

After the testing phase has been completed, all parameters which have become fixed need to be removed from the parameter file. In patient measurements, all parameters except for f_0 and vowel are in a separate file which need not be open during experiments but where they can be easily

3 Open: space for free writing; or multiple choice: tick the correct choice; or a mixture of both

4 Refers to data which does not change for every experiment

accessed.

The experimental arrangements may be further automatized for patient measurements with a start-up script that launches all software needed for the experiments upon login. If the parameter file proves to be confusing for the user even after removal of fixed parameters, the choice of f_0 and vowel may be moved to a dialogue window. The user would then only need to know how to run the parameter file in Matlab.

For this project, it is assumed that the location of sound data, keys, and meta data files are sufficient to carry out fast inspection of samples and, later, data analysis. The details of what is done with the data are outside the scope of this project. However, it is acknowledged that these details might now or in the future impose requirements that are not met by the current output of the system. The documentation, especially comments in the Matlab codes, is aimed to make later changes to the system as easy as possible.

This assignment has fine-tuned the experimental arrangements developed and used by Lukkari et al. (2007), Malinen and Palo (2009) and Aalto et al. (2011) and developed a software interface for running the experiments. The overall aim has been to produce a measurement arrangement that maximises the quality of data by removing as may sources of error as possible. Test subjects and patients are the one source of error that cannot be completely eliminated, however. Practice and guidance improve the situation but the system relies on sample inspection and repetition of experiments if necessary.

5. References

Aalto, A., Alku, P. and Malinen, J. (2009). A LF-pulse from a simple glottal flow model. *MAVEBA 2009* (pp. 199-202). Florence, Italy.

Aalto, D., Malinen, J., Palo, P., Aaltonen, O., Vainio, M., Happonen, R.-P., Parkkola, R. and Saunavaara, J. (2011). Recording speech sound and articulation in MRI. *Biodevices 2011* (pp. 168-173).

Aalto, D. and Vainio, M. (2011). Cue-signaali. Personal communication to T. Murtola. 25 Jan 2011.

Hannukainen, A., Lukkari, T., Malinen, J. and Palo, P. (2007). Vowel formants from the wave equation. *Journal of the Acoustical Society of America Express Letters*, *122*(1), EL1-EL-7.

Lear, C. S. C., Flanagan Jr, J. B. and Moorrees, C. F. A. (1965). The frequency of deglutition in man. *Archives of Oral Biology*, *10*(1), 83-99, IN13-IN15.

Lukkari, T., Malinen, J. and Palo, P. (2007). Recording speech during magnetic resonance imaging. *MAVEBA 2007* (pp. 163-166). Florence, Italy.

Malinen, J. and Palo, P. (2009). Recording speech during MRI: Part II. *MAVEBA 2009* (pp. 211-214). Florence, Italy.