For use by the Project lecturer	Approved		Revision required	

Feedback

To be completed by the student

PROJECT PROPOSAL 2022

Project no	AO5	Revision no	0

Title	Surname	Initials	Student no	Study leader (title, initials, surname)
Mr	**Ramedies**	**I**	16023405	Mr A Oloo

Project title

A school intercom system that filters inappropriate phrases and plays emails as announcements in the voice of the sender.

Language editor name	Language editor signature
Student declaration I understand what plagiarism is and that I have to complete my project on my own.	**Study leader declaration** This is a clear and unambiguous description of what is required in this project. Approved for submission (Yes/No)
Student signature	Study leader signature and date

1. Project description

What is your project about? What does your system have to do? What is the problem to be solved?

The problem addressed in this project is the lack of physical contact between teachers and students in schools due to the recent change in schooling systems. Teachers that make announcements and students that hear them are often not able to be in the same physical location and employing an individual to collate announcements would not be an efficient solution to this problem. Thus, the problem addressed in this project is to develop a school intercom system that will make announcements on behalf of the teacher, eliminating the need for announcers or any human intervention.

The concept is to have a system that receives a text-based email from a teacher and then announces this email to the students, in the voice of the teacher that sent the email. This may be accomplished without the need for human intervention by integrating natural language processing techniques, speech synthesis, and digital signal processing into a single intercom system. The system should also account for any inappropriate phrases received from the email and should be able to play the announcement over an integrated speaker so that multiple students in a classroom are able to hear it.

2. Technical challenges in this project

Describe the technical challenges that are *beyond* those encountered up to the end of third year and in other final year modules.

2.1 Primary *design* challenges

A general system design challenge consists of the collection of data, in terms of audio, and storing this data in the system so that it is accessible by the synthesis algorithm. Another design challenge is hardware allocation of the system resources. The overall system should be designed to work in real time, so all training of artificial intelligence systems will have to be conducted prior to final product compilation. Also, selecting a correct linguistic analysis technique will have an affect on the processing time of the overall implementation.

2.2 Primary *implementation* challenges

A system implementation challenge is managing waveforms and assigning it to analyzed text correctly. Testing results of synthesized audio when compared to natural language audio might be difficult to conduct unless extra functional components are developed.

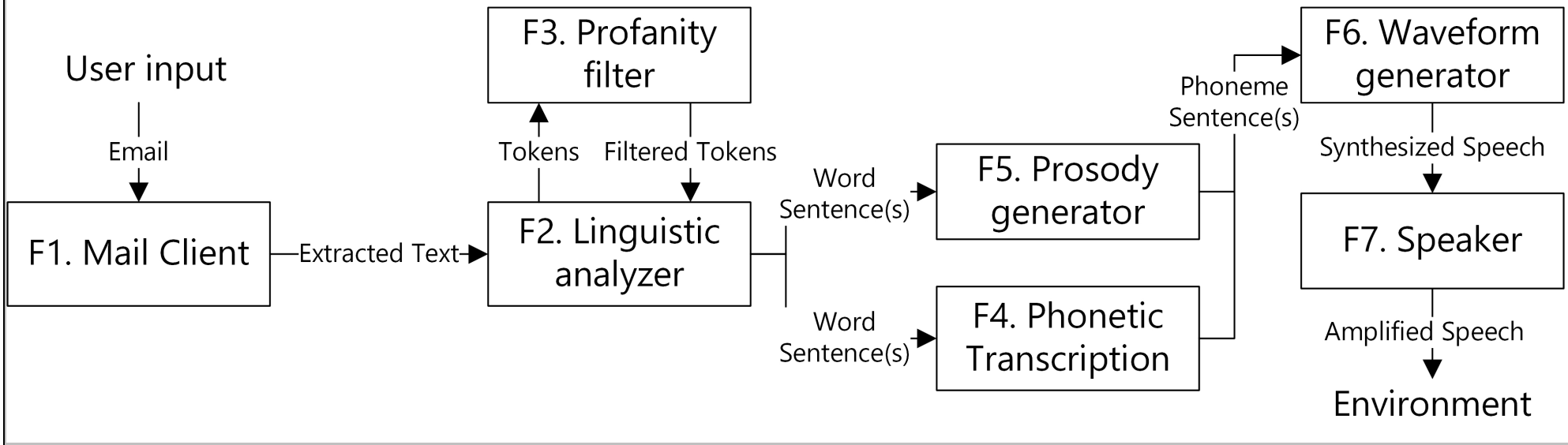
3. Functional analysis

3.1 Functional description

Describe the design in terms of system functions as shown on the functional block diagram in section 3.2. This description should be in *narrative format*.

The user input to the system will be an email text announcement that is sent to the system via the internet. F1 is a mail client that will receive emails. The text is extracted from the email and put through F2, a linguistics analysis function. The linguistics analyzer will be used to detect sentences, and create tokens from those sentences. The tokens are run through a profanity filter F3. The profanity filter will remove all inappropriate phrases contained within a set of tokens and will return the output to the linguistic analyzer. F2 will then go through a lemmatization and cleaning process to ensure the sentence is correctly formatted, and then tag each token with its part of speech. Each tag is then run through a linguistics analysis technique that will determine what they mean. The words are then sent to F4, a phonetic transcription function, that will append symbols to every part of the word representing what it will sound like. The symbols correspond to certain sounds as well decided by F5, a prosody generation function that appends these sounds. The phoneme sentences obtained from the combination of F4 and F5 are then used in F6, a waveform generator or synthesizer, to map the voice of the user to the words. Synthesized speech is then amplified by F7, a speaker that will relay the announcement that was received by the user. The output is heard in the classroom environment by students.

3.2 Functional block diagram



4. System requirements and specifications

These are the core requirements of the system or product (the mission-critical requirements) in table format IN ORDER OF IMPORTANCE. Requirement 1 is the most fundamental requirement.

	Requirement 1: the fundamental functional and performance requirement of your project	**Requirement 2**	**Requirement 3**
1. Core mission requirements of the system or product. Focus on requirements that are core to solving the engineering problem. These will reflect the solution to the problem.	The system should be able to convert an arbitrary text input to a speech output in natural language.	The text input will have to be processed so that it is correctly perceived by the students, i.e. it should be checked for correctness in punctuation and spelling.	The announcement has to be broadcast in the same fundamental frequency range of the person that is sharing the announcement so that it sounds like that person is speaking.
2. What is the target specification (in *measurable* terms) to be met in order to achieve this requirement?	A linguistics analysis module produces word error rates (WER) of 0% after processing text. Pitch and gain of output audio is 100% accurately aligned with the output sequence of the linguistics analyzer.	The accuracy of a spell checker should be above 95%, and the punctuation checker should account for all the beginnings and endings of sentences.	Generated waveforms of synthesized speech should be in the same fundamental frequency range of the persons voice Males typically have fundamental frequency ranging from 85-155 Hz and females are 165-255 Hz. Variations of these are included.
3. Motivation: *how* or *why* will meeting the specification given in point 2 above *solve the problem*? (Motivate the *specific* target specification selected)	A WER of 0% ensures that the text received from the user is correctly filtered and the sequence of phonemes produced by the analyzer is corect. The synthesis module should correspondingly assign correct sounds according to the generated sequence.	The probability of a word being spelt correctly has to be accurate enough so that the correct pronunciation with the correct phonemes are produced. Punctuation should split sentences up so that there is no confusion in the announcement sentences.	The fundamental frequency of the persons voice delivers an impression of a persons voice. By narrowing this parameter down, a better sound could be obtained for a synthesized speech.
4. How will you demonstrate at the examination that this requirement (point 1 above) and specification (point 2 above) has been met?	Arbitrary text data sets will be loaded to a simulated linguistics analysis module and each analysis output will be displayed for comparison. The corresponding output sequences should display no changes in the structure to confirm a 0% WER.	A data set of sentences will be input to a simulated spell and punctuation checker. The output of the data set sentences should be displayed along with the results of accuracy.	Recordings of multiple announcers voices will be played in conjunction with the synthesized speech for auditory comparison. Waveforms of the speech will also be displayed and fundamental frequency ranges will be compared.
5. Your own design contribution: what are the aspects that *you will design and implement yourself* to meet the requirement in point 2? If none, *remove this requirement.*	A text-to-speech synthesizer software with a linguistics analysis module will be implemented.	A software implementation of a spell checker and a basic punctuation checker will be developed and implemented.	The back-end of a text-to-speech synthesizer consisting of prosody generation and phonetic transcription will be designed and implemented.
6. What are the aspects to be taken off the shelf to meet this requirement? If none, indicate "none"	None.	A data set of words to be used for probability theory calculations.	A microcontroller board capable of digital signal processing will be purchased. A storage medium for audio clips will be purchased.

System requirements and specifications page 2

	Requirement 4	Requirement 5	Requirement 6
1. Core mission requirements of the system or product. Focus on requirements that are core to solving the engineering problem. These will reflect the solution to the problem.	The message that is announced should be filtered of all inappropriate phrases.	A mail system should receive emails from a user.	
2. What is the target specification (in *measurable* terms) to be met in order to achieve this requirement?	The accuracy, recall and precision of a large sample set, should be above 94% each to ensure that the profanity filter is accurate enough.	The text data that is sent by the user in the email should be correctly reflected by the receiving end of the system with no changes made to the input.	
3. Motivation: *how* or *why* will meeting the specification given in point 2 above *solve the problem*? (Motivate the *specific* target specification selected)	There are often phrases that seem inappropriate but are not. High accuracy, recall and precision ensures that False Positives are identified from the True positives in the model as well as the model being a reliable one.	To ensure that the mail system receives emails correctly, it can not be allowed for the receiving end to make any changes to the users input.	
4. How will you demonstrate at the examination that this requirement (point 1 above) and specification (point 2 above) has been met?	A large data set of phrases containing profanity and False Positives will be input into a simulated profanity filter. The outcome of accuracy, recall and precision will be measured by the models filtered output in comparison to the expected output.	A display screen should indicate that an email is received upon receiving an email. The output of the text should be displayed as well and confirmed with the user.	
5. Your own design contribution: what are the aspects that *you will design and implement yourself* to meet the requirement in point 2? If none, *remove this requirement.*	A text-based profanity filter will be developed and implemented.	Implementation of a email text display feature.	
6. What are the aspects to be taken off the shelf to meet this requirement? If none, indicate "none"	None.	A Wi-Fi module capable of Internet connection and an LCD screen for debugging and display purposes will be purchased. An email library will be used for the implementation of the mail client.	

System requirements and specifications page 3

	Requirement 7	**Requirement 8**	**Requirement 9**
1. <u>Core mission requirements of the system or product.</u> Focus on requirements that are core to solving the engineering problem. These will reflect the solution to the problem.			
2. What is the <u>target specification</u> (in *measurable* terms) to be met in order to achieve this requirement?			
3. <u>Motivation</u>: *how* or *why* will meeting the specification given in point 2 above *solve the problem*? (Motivate the *specific* target specification selected)			
4. How will you <u>demonstrate</u> at the examination that this requirement (point 1 above) and specification (point 2 above) has been met?			
5. <u>Your own design contribution:</u> what are the aspects that *you will design and implement yourself* to meet the requirement in point 2? If none, *remove this requirement.*			
6. What are the aspects <u>to be</u> taken off the shelf to meet this requirement? If none, indicate "none"			

System requirements and specifications page 4

	Requirement 10	Requirement 11	Requirement 12
1. Core mission requirements of the system or product. Focus on requirements that are core to solving the engineering problem. These will reflect the solution to the problem.			
2. What is the target specification (in *measurable* terms) to be met in order to achieve this requirement?			
3. Motivation: *how* or *why* will meeting the specification given in point 2 above *solve the problem*? (Motivate the *specific* target specification selected)			
4. How will you demonstrate at the examination that this requirement (point 1 above) and specification (point 2 above) has been met?			
5. Your own design contribution: what are the aspects that *you will design and implement yourself* to meet the requirement in point 2? If none, *remove this requirement.*			
6. What are the aspects to be taken off the shelf to meet this requirement? If none, indicate "none"			

5. Field conditions

These are the REAL WORLD CONDITIONS under which your project has to work and has to be demonstrated.

	Field condition 1	Field condition 2	Field condition 3
Field condition requirement. In which field conditions does the system have to operate? Indicate the one, two or three most important field conditions.	The intercom system has to operate correctly upon receiving emails with incorrect punctuation and spelling.		
Field condition specification. What is the specification (in measurable terms) for this field condition?	The system should be able to correct, with 90% accuracy, the contents of the email input to the system and announce the corrected output.		

6. Student tasks

6.1 Design and implementation tasks

List your primary design and implementation tasks in bullet list format (5-10 bullets). These are *not* product requirements, but *your* tasks.

The intercom system must be designed and implemented.
An appropriate linguistics analysis technique has to be selected.
Audio data from multiple people will have to be collected.
Large sentence text data sets with profanity have to be collected and sorted.
A physical cover for the system will have to be designed and built.
The system should be integrated onto a PCB board.
Simulated versions of the functions should be developed for debugging and demonstration purposes.

6.2 New knowledge to be acquired

Describe what the theoretical foundation to the project is, and which new knowledge you will acquire (*beyond* that covered in any other undergraduate modules).

The student will be required to master theoretical background knowledge on Natural Language Processing and Speech Synthesis.
The student will be required to acquire knowledge on Linguistics Analysis techniques.
The student will need to learn how to work with audio data sets.
The student will be required to learn about rulings of speech in the English language.