**Chapter 4: Implementation and Discussion of Results**

* 1. **Introduction**

System implementation evaluation involves systematically assessing the newly implemented system to ensure it meets the specified requirements and performs effectively in a real-world environment. After system analysis, what follows is the design. System design is the development of the actual configuration of the hardware and software. This chapter covers the implementation details of the PDF to Audio Converter System. This chapter includes the results of the testing conducted on the application, showcasing the functionality of the system and its performance in converting PDF documents to speech.

* 1. **Development Environment and Tools**

The development of the Application involved both client-side and server-side technologies. Below is an overview of the environment and tools used during the implementation:

* + 1. **Backend (Server-Side) Development**
* **Programming Language**: PHP
* **Text Extraction Library**: PDF2Text, a PHP-based library for extracting text from PDF files.
* **Database**: MySQL for storing metadata related to uploaded PDF files.
* **Web Server**: Apache (XAMPP) for local development and testing.
  + 1. **Frontend (Client-Side) Development**
* **HTML/CSS**: Used for building the user interface and ensuring responsiveness using Bootstrap.
* **JavaScript**: For handling speech synthesis using the Web Speech API.
* **Web Speech API**: A browser-native API used to convert extracted text into speech, providing speech synthesis functionality.
  + 1. **Development Tools**
* **Integrated Development Environment (IDE)**: Visual Studio Code.
* **Browser**: Google Chrome and Mozilla Firefox for testing Web Speech API compatibility.
  1. **Implementation**
     1. **System Implementation**

1. **File Upload Functionality** The backend provides a file upload handler that accepts PDF files from the user via an HTML form. The uploaded file is processed by PHP, which checks the file’s type and size to ensure it is a valid PDF. The file is then stored in a designated directory on the server.
2. **Text Extraction using PDF2Text** Once the file is uploaded, the **PDF2Text** library is used to extract the text from the PDF file. This library parses the PDF and retrieves any text it can find, converting it into a plain-text format that can be processed for speech synthesis.
3. **Database Integration** MySQL is used to store metadata about the uploaded PDF files, such as the file name, file path, and upload timestamp. This allows for record-keeping and potential future use of the stored PDF files.
   * 1. **User Interface Implementation**
4. **User Interface** The frontend is built using HTML and Bootstrap to provide a clean and responsive interface. The user can upload PDF files using a simple form, view the extracted text, and interact with the playback controls (play, pause, resume).

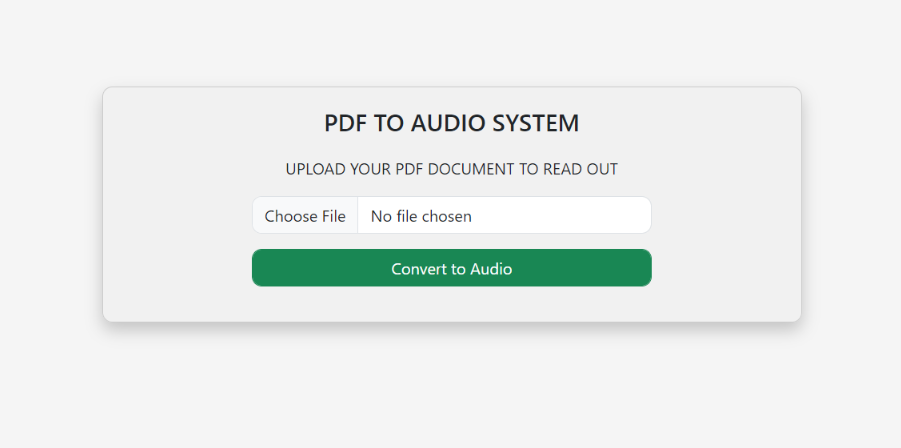


Figure. 4.1 PDF File Upload Interface

1. **Speech Synthesis using Web Speech API** Once the text is extracted from the PDF, it is displayed on the frontend. JavaScript is used to interact with the Web Speech API, which handles the conversion of text to speech. Users can play, pause, and resume the speech using the following code:
2. **Customization of Speech:** Users are given the ability to adjust the rate and pitch of the speech using sliders. These controls are linked to the Web Speech API's rate and pitch properties.



Figure 4.2 Speech Customization Controls

1. **Playback Controls:** The Web Speech API provides several methods to control the playback of text-to-speech (TTS), allowing users to interact with the spoken content more flexibly. The key playback controls include read-aloud (start), pause, resume, and stop. These controls provide users with the ability to manage speech synthesis dynamically, improving the overall experience.

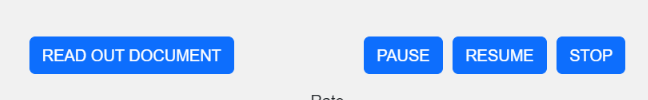


Figure 4.3 Playback Control Buttons

1. **Speech Voices**: A voice in speech synthesis is essentially a profile or a set of parameters that defines how the text will be spoken. Each voice may vary in the following aspects:

* Language: Voices are typically linked to specific languages, allowing for the pronunciation of text in a language-appropriate manner (e.g., English, French, Chinese).
* Accent/Region: Voices may also represent different regional accents or dialects within the same language (e.g., American English vs. British English).
* Gender: Some voices are designed to mimic male or female speakers.

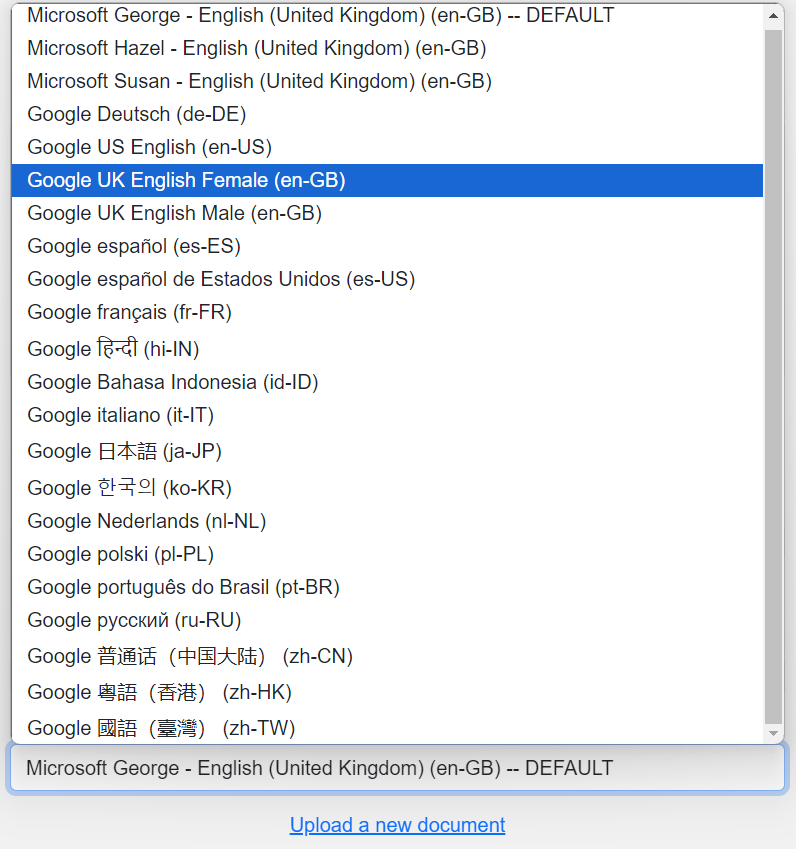


Figure 4.4 Speech Voices/Accents Selection

* 1. **Testing and Results**
     1. **Testing Procedure**

The system was tested on various browsers and with different types of PDF documents to ensure its functionality and performance. The primary objectives of the testing were:

1. **File Upload and Validation**: Ensuring that only valid PDF files could be uploaded and processed.

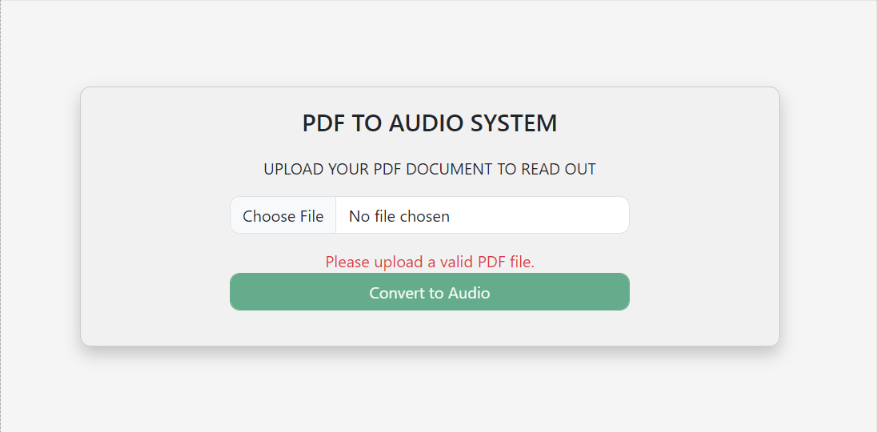


Figure 4.5 File Upload Validation

1. **Text Extraction**: Verifying that the system correctly extracts text from a variety of PDFs, including those with simple text, formatted text, and non-text elements like images.
2. **Speech Synthesis**: Testing the speech synthesis functionality, including the play, pause, and resume features, as well as customization of speech rate and pitch.
3. **Cross-Browser Compatibility**: Ensuring that the Web Speech API works on different browsers such as Google Chrome, Mozilla Firefox, and Microsoft Edge.
   * 1. **Test Results**

* **File Upload**: The file upload functionality worked as expected, with proper validation of file type (only PDFs) and file size. Invalid files (such as images or text documents) were rejected with appropriate error messages.
* **Text Extraction**: The text extraction worked well for PDFs containing simple text. However, for complex PDFs with images, tables, or multiple columns, the extraction was not always perfect, as the PDF2Text library struggled with layout and formatting.
* **Speech Synthesis**: The speech synthesis was successful across all tested browsers. The play, pause, and resume functionality worked as intended, and users were able to adjust the speech rate and pitch using the provided sliders. The audio output was clear, and customization options allowed users to tailor the experience to their preferences.
* **Cross-Browser Compatibility**: The Web Speech API was fully functional on Google Chrome and Mozilla Firefox. However, support for the Web Speech API on other browsers, such as Safari and Internet Explorer, was limited.
  + 1. **Challenges Encountered**
* **Complex PDFs**: Extracting text from PDFs with complex layouts (e.g., multi-column text, images) was a challenge, as the text extraction was not always accurate. Future improvements may involve integrating more advanced PDF parsing libraries that can handle complex layouts better.
* **Browser Compatibility**: While the Web Speech API worked well on modern browsers like Chrome and Firefox, there were limitations on other browsers like Safari and Internet Explorer, where the API either performed poorly or was unsupported.
  1. **Results**

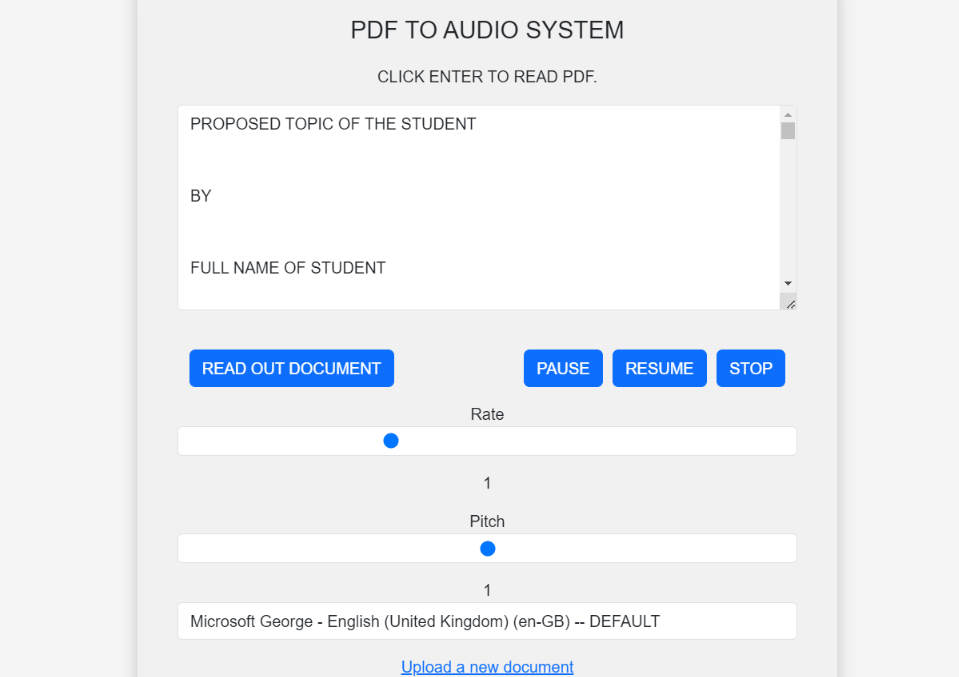
****

Figure 4.6 Successful File Upload and Controls

The PDF to Audio Converter Application performed well in most cases, meeting the primary objectives of converting PDF text into audio. The file upload and text extraction process were efficient, and the speech synthesis provided a seamless experience for users. The ability to pause, resume, and customize speech playback enhanced the user experience.

However, there are still areas for improvement. The PDF text extraction module could be further enhanced to handle more complex PDF layouts, including those with images and tables. Additionally, addressing browser compatibility issues would make the application more universally accessible.