



INSTITUTO POLITÉCNICO NACIONAL
UNIDAD PROFESIONAL INTERDISCIPLINARIA DE
INGENIERÍA Y TECNOLOGÍAS AVANZADAS



CARRERA: INGENIERÍA TELEMÁTICA

UNIDAD DE APRENDIZAJE: MULTIMEDIA

PRACTICA: 8-INTEGRACION SIT CON EL TTS (PARTE 1)



ALUMNO:

- GARNICA GONZALEZ CHRISTOPHER ALDAIR

BOLETA: 2023640373

DOCENTE: NOE SIERRA ROMERO

FECHA: 24/02/26

GRUPO: 2TV5

Resumen

Originalmente, muchas herramientas de accesibilidad o interfaces de "Texto a Voz" presentan uno de estos dos problemas:

- **Complejidad Innecesaria:** Interfaces saturadas con demasiados botones (pausar, detener, reiniciar, limpiar) que confunden al usuario final.
- **Falta de Personalización:** Herramientas que no permiten ajustar la identidad de la voz, resultando en audios monótonos o difíciles de entender por la velocidad predeterminada.
- **Rigidez:** La imposibilidad de cambiar el tono (pitch) o la rapidez (rate) para adaptarse a diferentes contextos (como una lectura rápida de noticias o una lectura lenta para aprendizaje de idiomas).

Resolucion

Se realizó una reingeniería del código base utilizando la **Web Speech API** para crear una herramienta minimalista pero potente. Las acciones clave fueron:

- **Simplificación de la UI (Interfaz de Usuario):** Se eliminaron los elementos distractores, dejando solo el área de entrada de texto y el botón de acción principal ("Play").
- **Implementación de Selectores Dinámicos:**
 - **Voz:** Se vinculó el sistema para detectar y listar automáticamente todas las voces disponibles en el navegador del usuario (Google, Microsoft, Apple).
 - **Velocidad:** Se añadió un control deslizante (slider) para ajustar el ritmo desde 0.5x hasta 2x.
 - **Tono:** Se integró un control de frecuencia para hacer la voz más aguda o más grave.
- **Optimización de Ejecución:** Se programó una función de "limpieza de cola" que cancela cualquier audio previo antes de iniciar uno nuevo, evitando que las voces se encimen si el usuario presiona "Play" varias veces.

-CODIGO:

```
<!DOCTYPE html>
<html lang="es">
<head>
<meta charset="UTF-8">
<meta name="viewport" content="width=device-width, initial-scale=1.0">
<title>Texto a Voz Simplificado</title>
<style>
body {
    font-family: 'Segoe UI', Tahoma, Geneva, Verdana, sans-serif;
    max-width: 500px;
    margin: 50px auto;
    padding: 20px;
    background: #f4f7f6;
    border-radius: 12px;
    box-shadow: 0 4px 15px rgba(0,0,0,0.1);
}
h2 { text-align: center; color: #2c3e50; }

textarea {
    width: 100%;
    height: 120px;
    padding: 12px;
    border: 2px solid #ddd;
    border-radius: 8px;
    resize: none;
    box-sizing: border-box;
    font-size: 16px;
}

.controls {
    margin-top: 20px;
    display: flex;
    flex-direction: column;
    gap: 15px;
}

.control-group {
    display: flex;
    flex-direction: column;
    gap: 5px;
}

label { font-weight: bold; color: #555; font-size: 14px; }

select, input[type="range"] {
    width: 100%;
    cursor: pointer;
}

#btn-speak {
    width: 100%;
    margin-top: 25px;
    padding: 15px;
    background: #4CAF50;
    color: white;
    border: none;
    border-radius: 8px;
    font-size: 18px;
    font-weight: bold;
    cursor: pointer;
    transition: background 0.3s;
}

#btn-speak:hover { background: #45a049; }

#status {
    margin-top: 15px;
    text-align: center;
```

```

        font-size: 14px;
        color: #888;
    }
</style>
</head>
<body>

<h2> 📺 Lector de Texto</h2>

<textarea id="input-area" placeholder="Escribe tu mensaje aquí..."></textarea>

<div class="controls">
    <div class="control-group">
        <label>Seleccionar Voz:</label>
        <select id="voice-select"></select>
    </div>

    <div class="control-group">
        <label>Velocidad: <span id="rate-val">1</span>x</label>
        <input type="range" id="rate" min="0.5" max="2" step="0.1" value="1">
    </div>

    <div class="control-group">
        <label>Tono (Pitch): <span id="pitch-val">1</span></label>
        <input type="range" id="pitch" min="0" max="2" step="0.1" value="1">
    </div>
</div>

<button id="btn-speak" onclick="speak()">▶ PLAY</button>

<div id="status">Listo para reproducir</div>

<script>
    const synth = window.speechSynthesis;
    const voiceSelect = document.getElementById('voice-select');
    const rateInput = document.getElementById('rate');
    const pitchInput = document.getElementById('pitch');
    const status = document.getElementById('status');
    let voices = [];

    function loadVoices() {
        voices = synth.getVoices();
        voiceSelect.innerHTML = "";
        voices.forEach((voice, i) => {
            const option = document.createElement('option');
            option.value = i;
            option.textContent = `${voice.name} (${voice.lang})`;
            voiceSelect.appendChild(option);
        });
    }

    if (speechSynthesis.onvoiceschanged !== undefined) {
        speechSynthesis.onvoiceschanged = loadVoices;
    }
    loadVoices();

    // Actualizar visualmente los valores de los sliders
    rateInput.oninput = () => document.getElementById('rate-val').innerText = rateInput.value;
    pitchInput.oninput = () => document.getElementById('pitch-val').innerText = pitchInput.value;

    function speak() {
        const text = document.getElementById('input-area').value;
        if (!text) return;

        // Cancelar cualquier audio que se esté reproduciendo actualmente
        synth.cancel();

        const utterance = new SpeechSynthesisUtterance(text);

        // Aplicar configuraciones elegidas

```

```
utterance.voice = voices[voiceSelect.value];
utterance.rate = rateInput.value;
utterance.pitch = pitchInput.value;

utterance.onstart = () => status.innerText = "Reproducendo...";
utterance.onend = () => status.innerText = "Finalizado.";

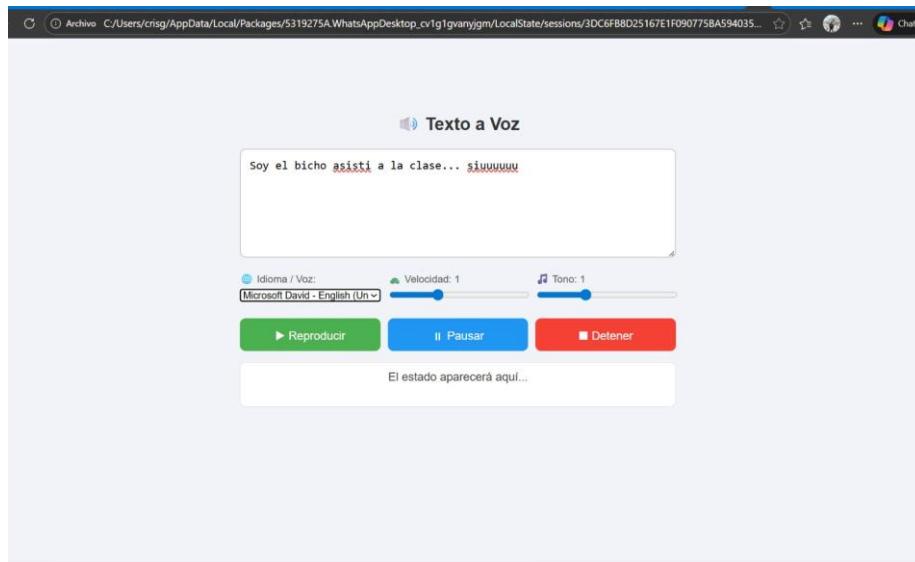
synth.speak(utterance);
}

</script>

</body>
</html>
```

-EVIDENCIA:

Conversor de texto inicial.



Conversor de texto modificado.

