



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 2)

ALUMNO:

- GARNICA GONZALEZ CHRISTOPHER ALDAIR

BOLETA: 2023640373

DOCENTE: NOE SIERRA ROMERO

FECHA: 24/02/26

GRUPO: 2TV5

Introducción

Este proyecto consiste en el desarrollo de una interfaz web de comunicación bidireccional utilizando las APIs nativas del navegador. El objetivo principal fue transformar un conversor de Texto a Voz (TTS) estándar en una herramienta minimalista, directa y altamente accesible. Para lograrlo, se buscó eliminar la complejidad visual para el usuario final, ocultando las configuraciones técnicas y sumando la capacidad de reconocimiento de voz (STT) para crear un ciclo completo de dictado y lectura.

Desarrollo (Los pasos realizados)

Para cumplir con los requerimientos del proyecto, el desarrollo se dividió en las siguientes fases:

Definición de parámetros (Punto 1): Se realizaron pruebas para determinar y elegir los valores óptimos de voz (idioma español), tono (pitch) y velocidad (rate) que garantizaran una reproducción clara y natural.

Refactorización del Código y UI (Puntos 2 y 3): Se modificó el código fuente para eliminar todos los controles visuales (sliders y menús desplegables). Los parámetros de voz, tono y velocidad elegidos en el paso anterior se establecieron como valores por omisión (hardcodeados) dentro del JavaScript. Así, la interfaz gráfica cumplió el objetivo de reducirse exclusivamente a una caja de Texto y un botón de Play.

Integración de STT con TTS (Punto 4): Se incorporó la API SpeechRecognition (Speech-to-Text). Se agregó un botón de "Dictar" que activa el micrófono, captura la voz del usuario y la transcribe automáticamente a la caja de texto. Posteriormente, esta transcripción puede ser leída por el sistema utilizando la función TTS previamente configurada, logrando la integración exitosa de ambas tecnologías.

-CODIGO:

```
<!DOCTYPE html>
<html lang="es">
<head>
    <meta charset="UTF-8">
    <meta name="viewport" content="width=device-width, initial-scale=1.0">
    <title>STT y TTS Integrados</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: 150px;
            padding: 12px;
            border: 2px solid #ddd;
            border-radius: 8px;
            resize: none;
            box-sizing: border-box;
            font-size: 16px;
            margin-bottom: 15px;
        }

        .buttons {
            display: flex;
            gap: 10px;
        }

        button {
            flex: 1;
            padding: 15px;
            color: white;
            border: none;
            border-radius: 8px;
            font-size: 16px;
            font-weight: bold;
            cursor: pointer;
            transition: opacity 0.3s;
        }

        button:hover { opacity: 0.85; }

        #btn-dictate { background: #2196F3; } /* Azul para STT */
        #btn-speak { background: #4CAF50; } /* Verde para TTS */

        #status {
            margin-top: 15px;
            text-align: center;
            font-size: 14px;
            color: #666;
            font-weight: bold;
        }
    </style>
</head>
<body>

    <h2> Dictado y Lectura</h2>
```

```

<textarea id="input-area" placeholder="Presiona 'Dictar' y habla, o escribe tu mensaje aquí..."></textarea>

<div class="buttons">
  <button id="btn-dictate" onclick="startDictation()">🎤 Dictar (STT)</button>
  <button id="btn-speak" onclick="speak()">▶ Play (TTS)</button>
</div>

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

<script>
// =====
// 1. CONFIGURACIÓN TTS (Text-to-Speech)
// =====
const synth = window.speechSynthesis;
const statusDiv = document.getElementById('status');
const inputArea = document.getElementById('input-area');

let defaultVoice = null;

// Cargar voces y fijar una por omisión (Punto 2 y 3 del pizarrón)
function setupVoice() {
  const voices = synth.getVoices();
  // Buscar una voz en español por defecto, si no, usa la primera disponible
  defaultVoice = voices.find(v => v.lang.includes('es')) || voices[0];
}

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

function speak() {
  const text = inputArea.value.trim();
  if (!text) {
    statusDiv.innerText = "⚠️ No hay texto para leer.";
    return;
  }
  synth.cancel(); // Limpiar lecturas previas

  const utterance = new SpeechSynthesisUtterance(text);

  // Valores por omisión definidos en código (sin selectores en pantalla)
  if (defaultVoice) utterance.voice = defaultVoice;
  utterance.rate = 1.0; // Velocidad normal
  utterance.pitch = 1.0; // Tono normal

  utterance.onstart = () => statusDiv.innerText = "🔊 Reproduciendo...";
  utterance.onend = () => statusDiv.innerText = "✅ Finalizado.";
  utterance.onerror = () => statusDiv.innerText = "✖️ Error en reproducción.";

  synth.speak(utterance);
}

// =====
// 2. CONFIGURACIÓN STT (Speech-to-Text)
// =====
const SpeechRecognition = window.SpeechRecognition || window.webkitSpeechRecognition;
let recognition;

if (SpeechRecognition) {
  recognition = new SpeechRecognition();
  recognition.lang = 'es-ES'; // Idioma español para escuchar
  recognition.interimResults = false; // Solo devuelve el texto final

  recognition.onstart = function() {
    statusDiv.innerText = "🎙️ Escuchando... habla ahora.";
  }
}

```

```
};

recognition.onresult = function(event) {
    const transcript = event.results[0][0].transcript;
    // Agrega el texto dictado al área de texto
    inputArea.value += (inputArea.value ? " " : "") + transcript;
    statusDiv.innerText = "✅ Texto capturado.";
};

recognition.onerror = function(event) {
    statusDiv.innerText = "❌ Error al escuchar: " + event.error;
};
} else {
    statusDiv.innerText = "⚠ Tu navegador no soporta el dictado por voz (STT).";
}

function startDictation() {
    if (recognition) {
        recognition.start();
    } else {
        alert("El dictado por voz no es compatible con este navegador. Intenta usar Google Chrome.");
    }
}
</script>

</body>
</html>
```

Resumen / Conclusión

El resultado final es una aplicación web limpia y funcional que integra de manera invisible tanto el reconocimiento de voz (STT) como la síntesis de voz (TTS). Al asignar los valores de voz por defecto directamente en el código y dejar una interfaz conformada únicamente por el área de texto y los botones de acción ("Dictar" y "Play"), se resolvió la problemática de saturación visual. Se entregó un prototipo eficiente, listo para usarse sin requerir configuración previa por parte del usuario.

