

Análisis y Comparación de Características Acústicas con MFCC

Walter Raul Perez Machinena

Maestría en Ciencia de Datos

Universidad Autónoma de Nuevo León

San Nicolás de los Garza, Nuevo León, México

waltermachinena@gmail.com

I. INTRODUCCIÓN

En esta actividad se llevó a cabo un análisis de tres archivos de audio utilizando técnicas de procesamiento de señales para comparar sus características acústicas. Los archivos de audio seleccionados corresponden a grabaciones de diferentes personas, y el objetivo fue comparar las características espectrales extraídas de estos audios a través de la técnica de coeficientes cepstrales en las frecuencias de Mel (MFCC).

II. HALLAZGOS

Los hallazgos principales de este análisis se centran en las diferencias y similitudes entre los coeficientes MFCC de los tres audios utilizados. A continuación se presentan algunos puntos clave:

- Los coeficientes MFCC mostraron que las grabaciones de "Lily", y "Walter2" tienen una mayor similitud en comparación con las grabaciones de "Lily", y "Walter".
- Se observaron diferencias significativas entre las grabaciones de "Walter1", y "Walter2", lo que indica que la variabilidad en la grabación afecta significativamente a las características acústicas.
- La comparación entre los MFCCs también reveló patrones de frecuencias dominantes en las grabaciones, los cuales fueron visualizados a través de espectrogramas y gráficos tridimensionales.

III. PROCEDIMIENTO

El proceso de análisis consistió en varias etapas clave, detalladas a continuación:

III-A. Carga de Archivos de Audio

Se cargaron tres archivos de audio en formato .ogg, correspondientes a tres grabaciones diferentes. Los archivos utilizados fueron:

- audio_lily.ogg
- audio_walter.ogg
- audio_walter_2.ogg

III-B. Preprocesamiento de Audio

Para asegurar la consistencia en la frecuencia de muestreo, todos los audios fueron re-muestreados a una frecuencia de 22,050 Hz. Esto permite que los coeficientes MFCC puedan ser comparados en una escala común. Los audios fueron luego

normalizados y sus características fueron extraídas utilizando la librería librosa.

III-C. Extracción de Características

Se extrajeron los coeficientes MFCC, que son una representación compacta de las características espectrales del audio. Se utilizaron 13 coeficientes MFCC para cada archivo de audio, lo que permitió realizar una comparación entre las grabaciones.

III-D. Visualización y Comparación

Las características MFCC fueron visualizadas en varias formas:

- Espectrogramas de las señales de audio.
- Gráficos de dispersión en 3D de las trayectorias de los coeficientes MFCC.
- Mapa de calor de la matriz de similitud entre los diferentes audios.
- Comparación de distancias Euclidianas entre los coeficientes MFCC.

IV. RESULTADOS

Los resultados obtenidos muestran que las grabaciones de Lily y Walter_2 presentan una similitud notable en sus características acústicas, según lo indicado por los coeficientes MFCC. Esta similitud fue visualizada tanto en el gráfico de similitud como en el análisis de la matriz de similitud.

Por otro lado, las diferencias entre Walter_1 y Walter_2 fueron marcadas, lo cual sugiere que la variabilidad en la grabación puede afectar considerablemente las características acústicas extraídas.

El análisis de las frecuencias dominantes también reveló patrones únicos en cada grabación, lo que refuerza la idea de que las características acústicas son distintivas para cada grabación, a pesar de que pueden compartir algunas similitudes.

Se generaron varias gráficas para ilustrar los resultados, incluyendo espectrogramas, gráficos de dispersión y un gráfico de radar para mostrar la comparación de similitudes entre las grabaciones. Además, la matriz de similitud proporcionó una visión clara de las distancias acústicas entre las diferentes grabaciones.

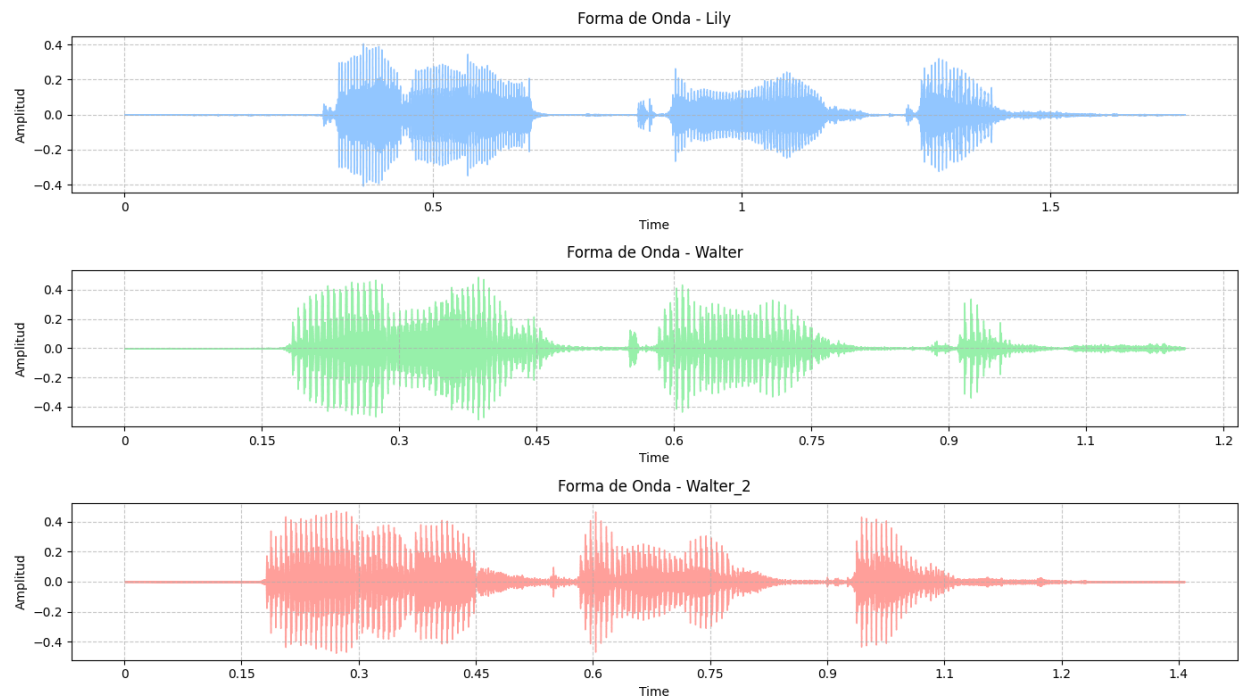


Figura 1: Comparación de formas de onda de los audios de Lily, Walter y Walter2.

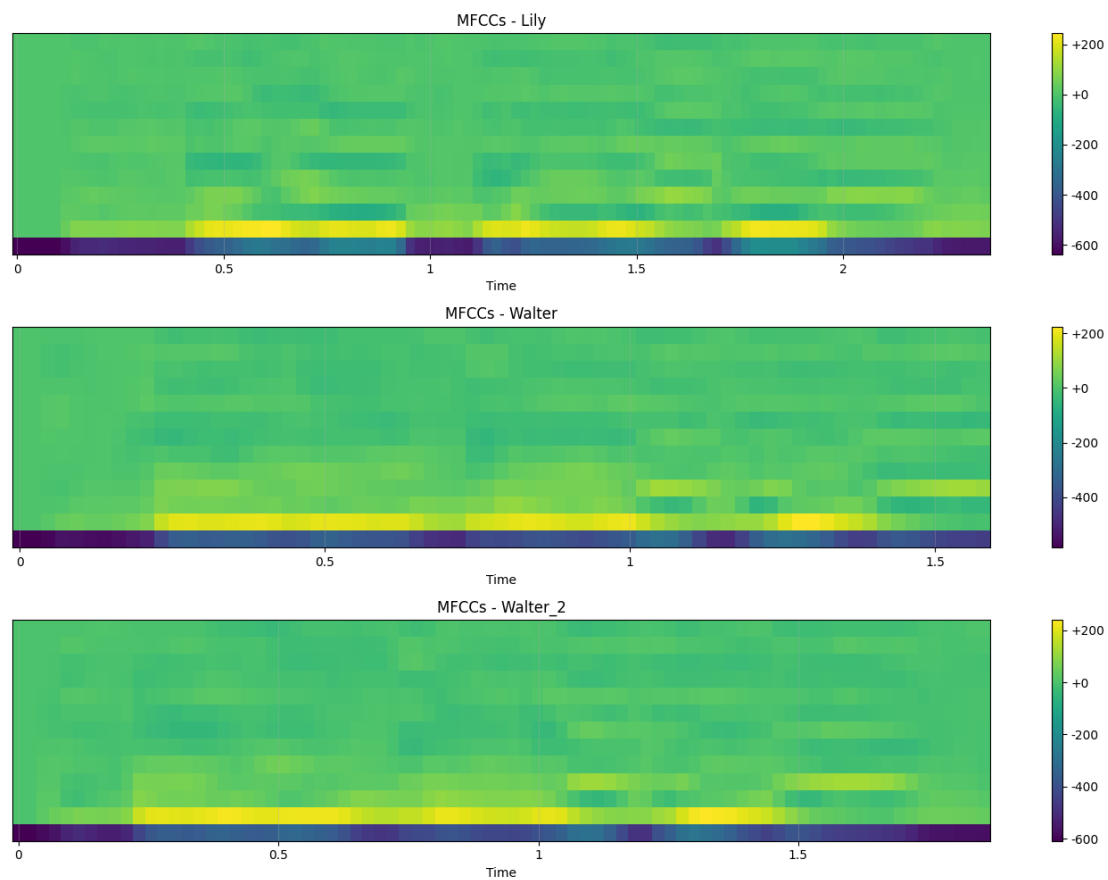


Figura 2: Representación gráfica de los coeficientes MFCC extraídos de los audios de Lily, Walter y Walter2.

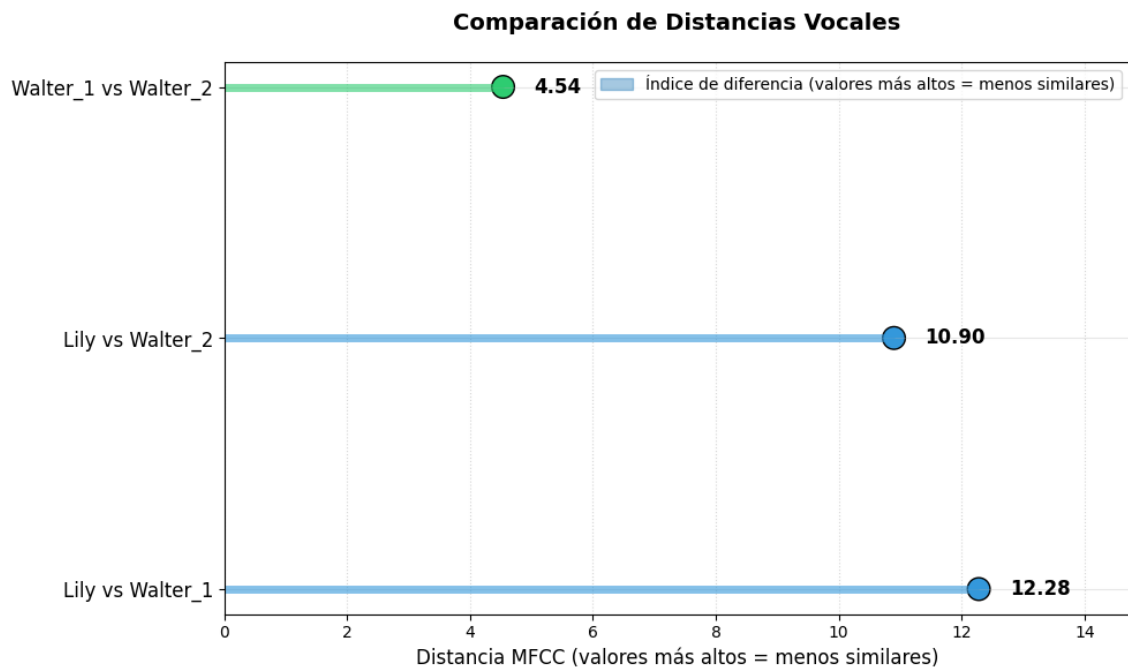


Figura 3: Gráfico comparativo de distancias entre los audios de Lily y Walter1, Lily y Walter2, y Walter1 y Walter2.

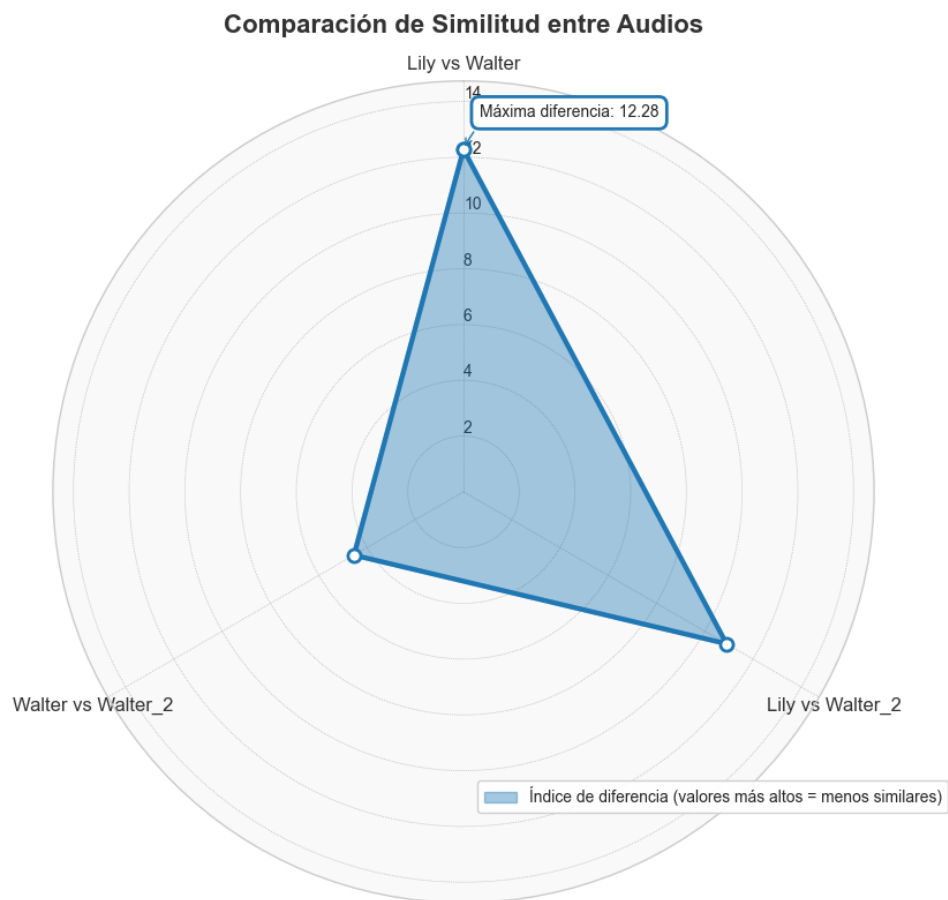


Figura 4: Gráfico radar para comparar la similitud entre los audios de Lily, Walter1 y Walter2.

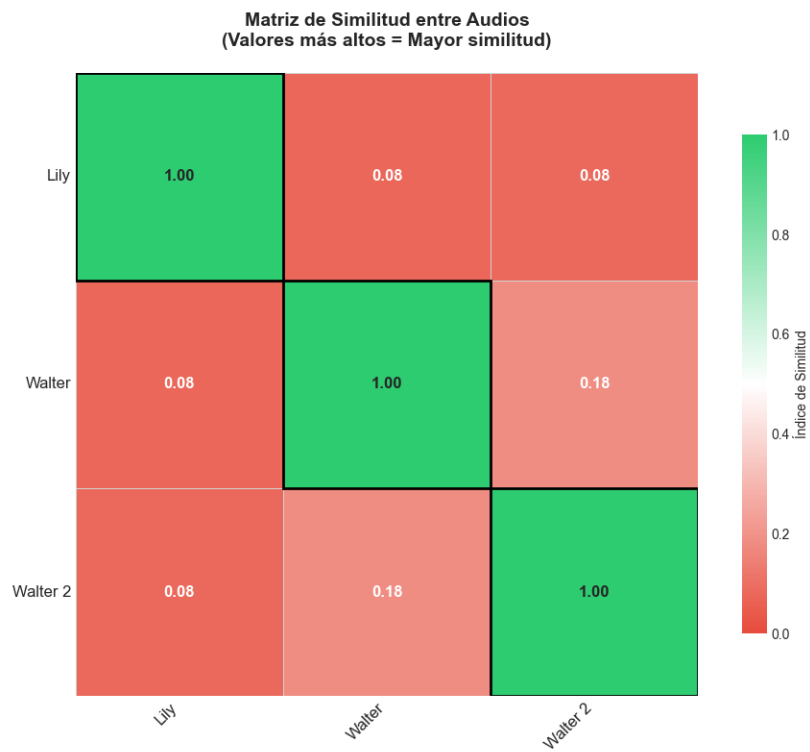


Figura 5: Matriz de similitud entre los audios de Lily, Walter1 y Walter2.

Análisis 3D de Características Vocales

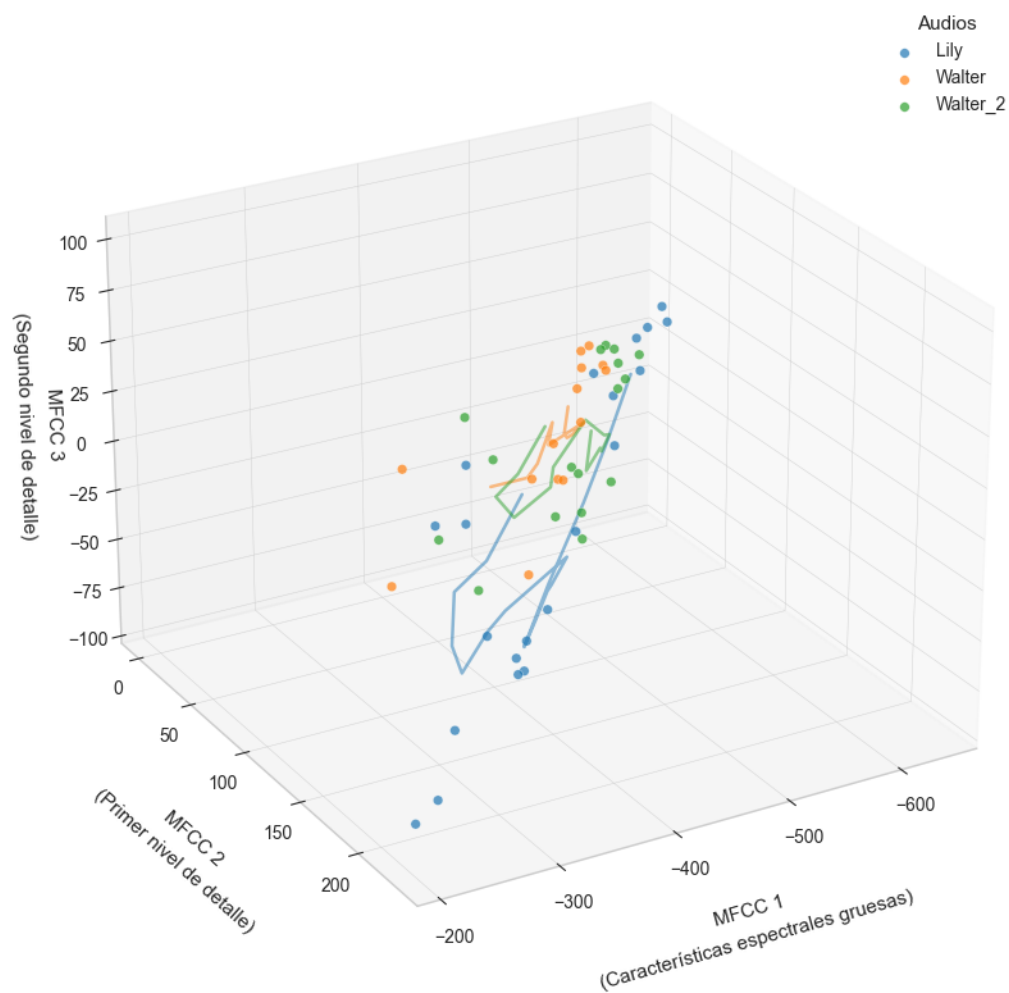
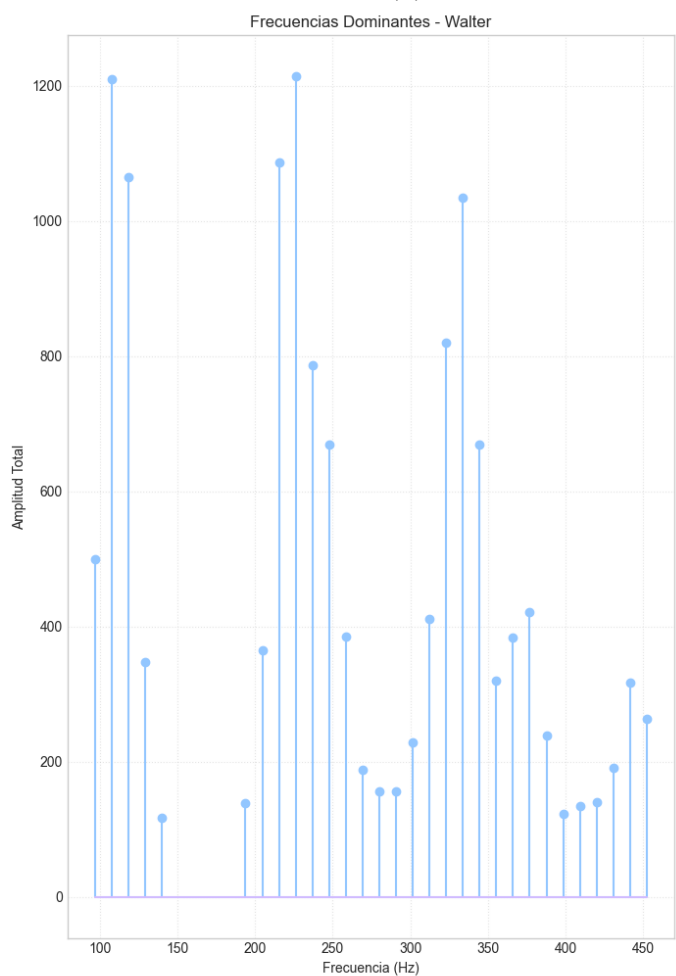
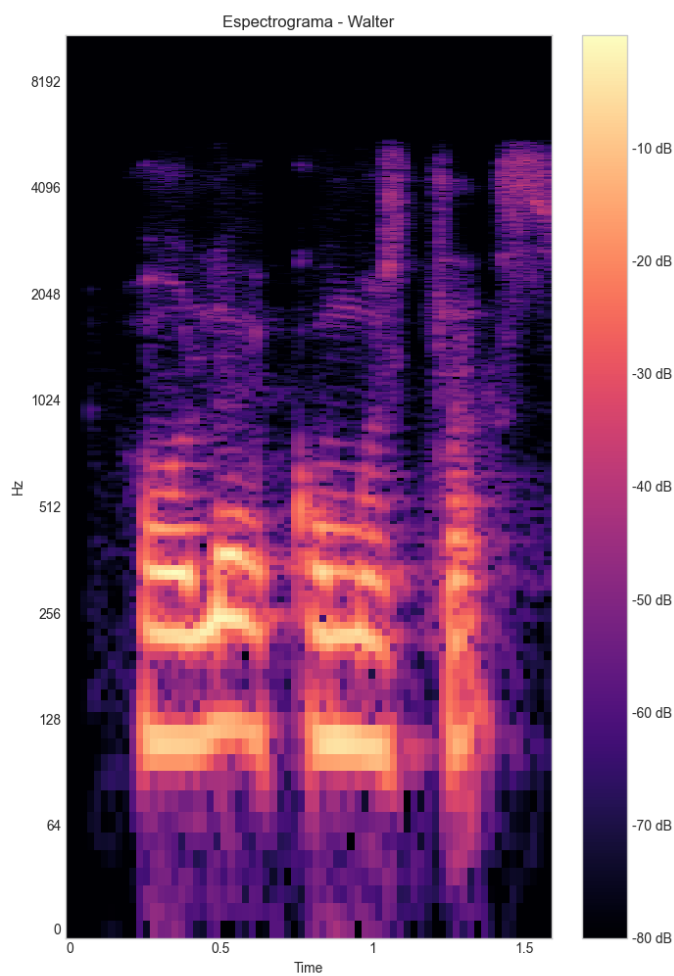
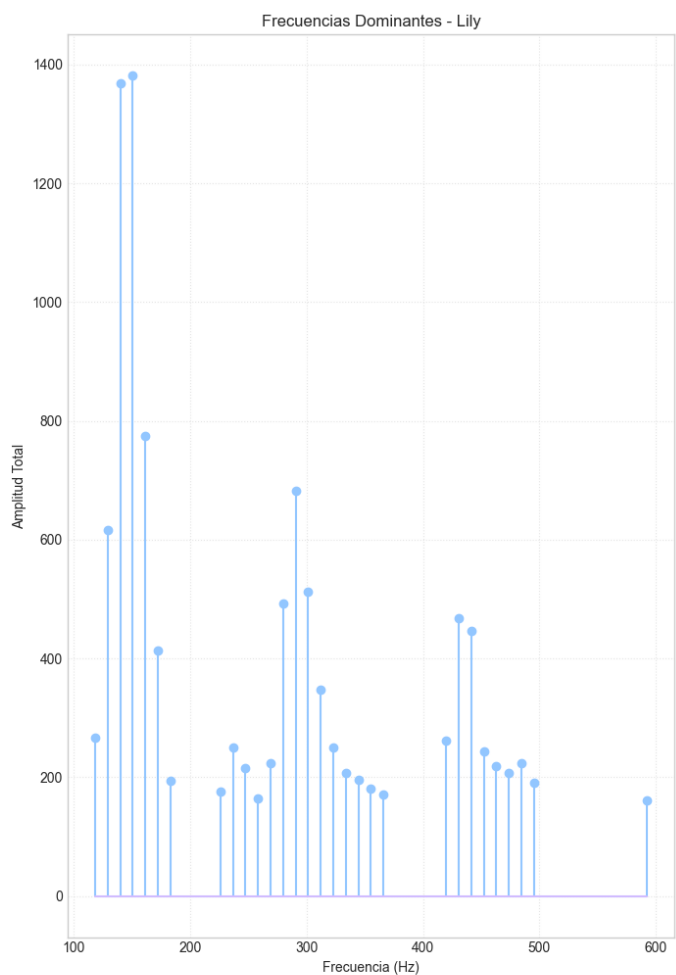
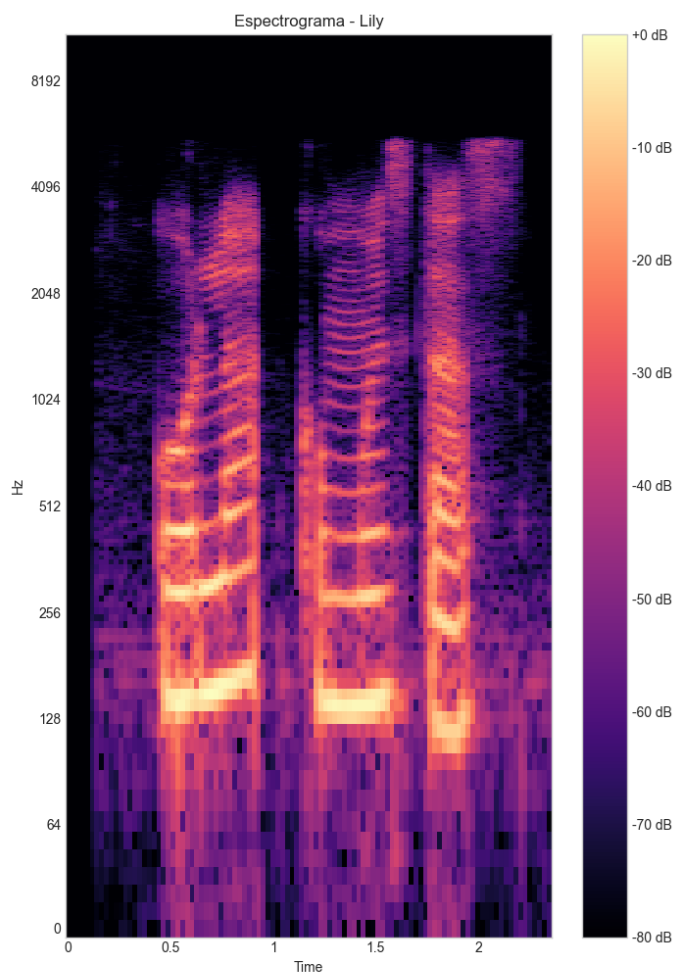


Figura 6: Grafica de 3 dimensiones características vocales



Espectrograma - Walter 2

Frecuencias Dominantes - Walter 2

V. CONCLUSIÓN

El análisis realizado permitió comparar las características acústicas de tres grabaciones de audio mediante la extracción de coeficientes MFCC. A través de la visualización de las distancias entre los diferentes audios, las similitudes y las diferencias entre las grabaciones fueron claramente evidentes. Este análisis resalta la importancia de los MFCC como herramienta para el análisis de audio y su utilidad en tareas como la clasificación de sonidos y la identificación de patrones acústicos.