# Proceso de evaluación de resultados

1. Consiste en comprobar el número de palabras coincidentes entre diferentes ficheros, teniendo en cuenta...
   1. **Observaciones:**

Todos los ficheros a procesar están contenidos en **dos tipos de carpetas**:

* + - * La carpeta denominada **“Manual”**: compuesta por documentos que contienen una lista de palabras (que pueden ser compuestas) separadas entre sí por un retorno de carro.
      * Las carpetas **“Frec”**, **“FrecSem”** y **“Sem”**: estas carpetas están formadas por documentos que contienen una lista de palabras (que pueden ser compuestas) y un peso que indica su importancia. Palabras y pesos están separadas entre sí por “:”. La ordenación de estas listas de palabras se ha hecho en función del peso (cuanto mayor es el peso mejor posición ocupa la palabra).
  1. **Proceso**:
     1. Comparar los documentos de la carpeta “**Frec”**con los documentos de la carpeta **“Manual”**. Se compararán entre sí aquellos ficheros con el mismo nombre. Como resultado de este proceso se producirán tres salidas de cada comparación:

a) resultado de la comparación de las primeras 5 palabras de los documentos;

b) resultados de la comparación de las 10 primeras palabras; y

c) resultados de la comparación de las 15 primeras palabras.

* + 1. Este proceso de comparación se repite para todas las combinaciones posibles, que son las siguiente:

1. Comparación **“Manual”** vs. **“Frec”**

2. Comparación **“Manual”** vs. **“FrecSem”**

3. Comparación **“Manual”** vs. **“Sem”**

4. Comparación **“Frec”** vs. **“FrecSem”**

5. Comparación **“Frec”** vs. **“Sem”**

6. Comparación **“FrecSem”** vs. **“Sem”**

iii. Los resultados obtenidos de las 6 comparaciones se incluirán en diferentes ficheros en formato txt o excel. Concretamente, se obtendrá un fichero por cada combinación (seis en total) de las 5 primeras palabras, otro por cada combinación (seis en total) de las 10 primeras palabras, y por último, otro por cada combinación (seis en total) de las 15 primeras palabras. Es decir, en total se obtendrán 18 ficheros txt (o excel) (6 combinaciones x 3 grupos de palabras) que recogerán los siguientes resultados:

1. Comparación **“Manual”** vs. **“Frec”**: para las...

a)5 primeras palabras

b) 10 primeras palabras

c) 15 primeras palabras

2. Comparación **“Manual”** vs. **“FrecSem”**: para las...

a)5 primeras palabras

b) 10 primeras palabras

c) 15 primeras palabras

3. Comparación **“Manual”** vs. **“Sem”**: para las...

a)5 primeras palabras

b) 10 primeras palabras

c) 15 primeras palabras

4. Comparación **“Frec”** vs. **“FrecSem”**: para las...

a)5 primeras palabras

b) 10 primeras palabras

c) 15 primeras palabras

5. Comparación **“Frec”** vs. **“Sem”:** para las...

a)5 primeras palabras

b) 10 primeras palabras

c) 15 primeras palabras

6. Comparación **“FrecSem”** vs. **“Sem”**: para las...

a)5 primeras palabras

b) 10 primeras palabras

c) 15 primeras palabras

iv) Cada uno de los ficheros txt (o excel) tendrá la siguiente información: NombreDocumento; Num. palabras coincidentes; Pal1Coincidente; Peso Pal1Coincidente; … ; PalNCoincidente; Peso PalNCoincidente

La siguiente imagen muestra un posible formato de salida:

