

# Cortado de señales y anotación de disfluencias en el Corpus Oral UAM

Luis Javier Rodríguez Fuentes  
Departamento de Electricidad y Electrónica  
Facultad de Ciencias. Universidad del País Vasco  
Apartado 644. 48080 Bilbao. SPAIN  
e-mail: luisja@we.lc.ehu.es

9 de mayo de 2002

## 1 Introducción

El Corpus Oral UAM (CO-UAM en lo sucesivo) está formado por diálogos entre dos o más personas, grabados en radio, en televisión, en la calle, en aulas de clase, en mesas redondas, en casa, etc. Se trata, por tanto, de un corpus de habla espontánea en castellano, que abarca diversos dominios semánticos y pragmáticos, ya que los diálogos se desarrollan en distintos contextos socio-culturales y en distintas situaciones, algunas más formales, otras más familiares. El objetivo final del trabajo es disponer de una base de datos genérica de habla espontánea, lo suficientemente grande como para obtener modelos acústicos robustos, con un vocabulario grande no restringido a una tarea. Para alcanzarlo han de llevarse a cabo dos tareas:

1. cortar la señal correspondiente a cada diálogo en tantos ficheros como turnos resulten para dicho diálogo, y
2. corregir y aumentar las anotaciones ya existentes.

No se va a procesar todo el corpus, sino sólo una parte, cuyas dimensiones se ajustan –creemos– al objetivo enunciado más arriba. CO-UAM está dividido en 17 secciones atendiendo a la temática de los diálogos. Una parte especialmente grande e importante del corpus está formada por diálogos de temática variada no encasillables en ninguna categoría. A su vez, estos diálogos se han clasificado en dos tipos: conversaciones y entrevistas. Las primeras son menos restrictivas, suelen implicar a más de dos personas e incluyen múltiples solapamientos ya que no hay un moderador que conceda turnos. Las entrevistas son conversaciones más formales, normalmente entre dos personas, una de ellas haciendo las veces de entrevistador/moderador.

En general la calidad acústica de las entrevistas es aceptable ya que en su mayor parte han sido tomadas de la radio, mientras que entre las conversaciones hay muchas grabadas en la calle, en casa, etc. con muchos ruidos, eco, etc. Se han considerado acústicamente aceptables 67 de un total de 79 entrevistas y 66 de un total de 126 conversaciones<sup>1</sup>, sumando un total de 133 diálogos. En primera instancia sólo se van a cortar y anotar estos diálogos. Ello no obsta para que en una segunda fase, si hubiera tiempo, se cortaran y anotaran también los restantes diálogos, considerados de baja calidad.

En cualquier caso, el proceso de cortado y anotación se realizará con dos criterios principales:

1. Cortar sólo señales acústicamente aceptables. Esto significa que si una señal presenta ruido o música de fondo apreciables, o contiene voces solapadas, será descartada.
2. Anotar todo lo que se escuche: ruidos de fondo, música de fondo, cortes, solapamientos, respiraciones, tics de los labios, etc. En algunos casos, la señal habrá sido descartada, es decir, no se habrá cortado y guardado en un fichero, pero seguirá siendo necesario escucharla detenidamente para anotar lo que sucede en ese tramo acústicamente "defectuoso".

---

<sup>1</sup>En realidad las conversaciones constan de 125 señales y 126 transcripciones, ya que la señal de una de ellas se perdió después de haber sido transcrita.

Para compatibilizar ambos criterios, las anotaciones deberán indicar dónde empiezan y dónde terminan los tramos "defectuosos", de manera que aún disponiendo de toda la transcripción, sea posible quedarse únicamente con los fragmentos de transcripción para los que se tenga señal. Para facilitar esta tarea, cada vez que se encuentre un tramo defectuoso, el final de éste determinará el final de un turno, de modo que el siguiente tramo acústicamente aceptable tendrá lugar ya en el turno siguiente. Ello permitirá asociar a cada señal cortada un índice de turno distinto.

## 2 Cortado de las señales

El procedimiento será manual. Se visualizará por un lado el texto de la transcripción original –en formato EHU, extensión *.dis*–, mediante un editor de texto básico, y por otro la forma de onda de la señal –en formato ESPS, a frecuencia de muestreo de 16 kHz, extensión *.esps*– mediante la herramienta *XWAVES*. Es decir, si el fichero con la transcripción se llama *aent034b.dis*, el fichero de señal correspondiente se llamará *aent034b.esps*.

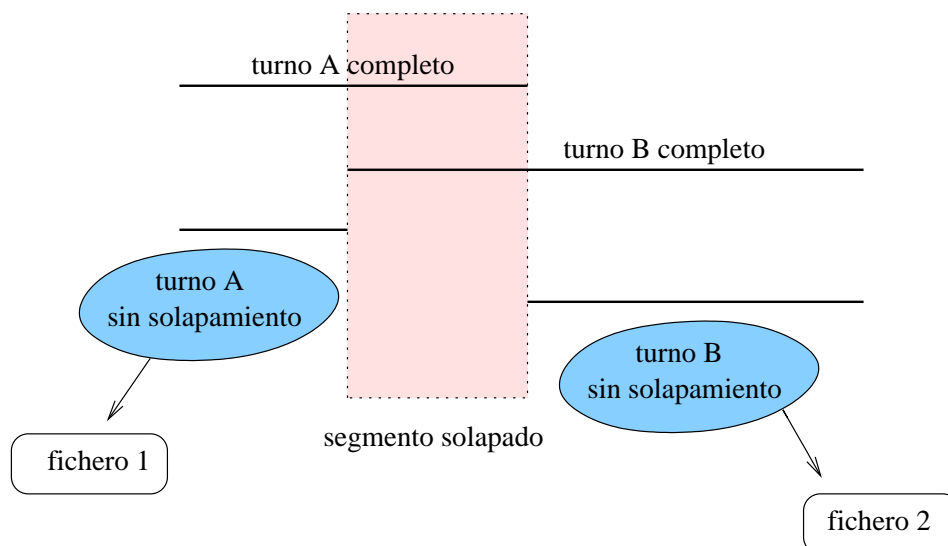
En la forma de onda se seleccionarán/cortarán los fragmentos correspondientes a los turnos o intervenciones de los hablantes. Esto se realizará de forma natural mediante el ratón:

1. se selecciona un trozo de señal,
2. se escucha (hay una opción del menú que lo permite),
3. si el segmento no se ajusta a la transcripción del turno que se pretende cortar, entonces volver al paso 1, o bien
4. si se ajusta, guardar el segmento cortado en un fichero.

El nombre del fichero se compondrá añadiendo al nombre base del diálogo dos extensiones: el índice del turno (número de tres dígitos) y el identificador del hablante. Por ejemplo, si se está cortando el cuarto turno del diálogo *aent034b.esps*, producido por el hablante *H2*, el nombre del fichero de señal sería *aent034b.004.H2.esps*. Los ficheros de señal correspondientes a los turnos de un diálogo se guardarán todos dentro del mismo directorio, que tendrá el nombre base del diálogo (en el ejemplo anterior, el directorio se llamaría *aent034b*).

Los solapamientos (dos o más interlocutores hablando a la vez) plantean dificultades a la hora de determinar dónde empiezan y dónde acaban. Como se ha comentado más arriba, se ha optado por eliminar todos los fragmentos de señal que muestren solapamientos, ya que desde el punto de vista acústico resultan de escasa o nula utilidad.

Si se produce un solapamiento entre el final de un turno y el comienzo del siguiente, el segmento solapado quedará fuera de los dos ficheros de señal que se generan. Tal como muestra la figura, en un fichero estaría la señal hasta el inicio del solapamiento y en el otro la señal a partir del fin del solapamiento. Por tanto, cada solapamiento determinará la aparición de un nuevo turno. La definición de los puntos de inicio y fin del solapamiento se hará de forma que abarquen palabras completas. Es decir, si una palabra está afectada por el solapamiento, aunque sólo sea muy al principio o muy al final, dicha palabra será incluida en el segmento solapado.



En cuanto a las anotaciones, se ha optado por incluir la transcripción de los segmentos solapados, marcando el inicio y el fin de los mismos, ya que desde el punto de vista del modelo de lenguaje no sería muy adecuado cortar las frases y dejarlas inacabadas. Así pues, los segmentos solapados, aunque descartados para las señales, deberán ser escuchados y anotados. Más adelante veremos en detalle cómo anotarlos.

Todo lo dicho en este apartado con respecto a los solapamientos de dos voces se aplicará exactamente igual a los segmentos de señal solapados con ruidos de fondo o con música: se marcan el inicio y el fin del fenómeno y se incluye la transcripción de lo que se escucha *por encima* del ruido o de la música.

Finalmente, con respecto a los silencios, todos aquéllos que sucedan en el interior de un turno se marcarán como pausas, aunque sean anormalmente largos. Sin embargo, los que sucedan al comienzo o al final de un turno serán truncados siempre que superen la duración de 0.5 segundos. Es decir, al principio y al final de un turno se tendrán silencios de como máximo 0.5 segundos. Cuando dos turnos estén separados por una pequeña pausa, inferior a un segundo, la frontera entre ambos se establecerá más o menos en el punto medio.

### 3 Anotación de los fenómenos de habla espontánea

CO-UAM se anotará en dos fases. La primera fase se realizará a la vez que el cortado de las señales en turnos, y cubrirá los fenómenos acústicos y léxicos. La segunda fase cubrirá los fenómenos sintácticos y pragmáticos.

#### 3.1 Primera fase: cortado y anotación de fenómenos acústicos y léxicos.

La primera fase se realizará de manera simultánea con el proceso de cortado descrito en la sección anterior. El anotador dispondrá de tres fuentes de información: su percepción auditiva de la señal, que podrá oír tantas veces como quiera, la forma de onda y el espectrograma de esa misma señal, que podrá recorrer y ampliar con la herramienta *X WAVES*, y la transcripción de partida en formato EHU, que podrá visualizar con el editor *gvim* debidamente configurado para mostrar las marcas en diferentes colores, lo cual facilitará las operaciones de lectura y modificación.

La transcripción de partida podrá ajustarse más o menos a la señal, dependiendo de la precisión del anotador original: algunos fenómenos estarán correctamente anotados, otros lo estarán sólo con cierta precisión y otros muchos ni siquiera habrán sido anotados.

El anotador deberá explorar cada turno contrastando lo que oye con la anotación que tiene, añadiendo, borrando o modificando marcas a medida que avanza. El inventario de fenómenos incluye ruidos externos, ruidos producidos por el locutor, pausas de silencio, pausas habladas, alargamientos de sonidos (principalmente vocales), palabras cortadas, palabras mal pronunciadas, sonidos guturales que pueden indicar afirmación o negación, siglas, palabras extranjeras y solapamientos.

Casi todos estos fenómenos han sido descritos con anterioridad en el manual de anotación de la base de datos INFOTREN, formada por diálogos hombre-máquina. Sin embargo, los tres últimos han sido incluidos por primera vez en esta base de datos, y por ello pasamos a describirlos con detalle a continuación.

### 3.1.1 Siglas.

Las siglas o acrónimos no tienen una forma regular de pronunciación. Algunas veces se deletrean y otras se pronuncian como palabras normales. Por ejemplo, FBI se deletrea mientras que CIA se lee como una palabra. Ello no impide que los hablantes elijan una u otra forma de pronunciación según su costumbre. Precisamente para señalar aquellas ocasiones en que unas siglas se deletrean se ha creado un tipo especial de disfluencias léxicas, llamado *acronym*. Por ejemplo:

```
<l type="acronym"> FBI </l>,
```

que en formato simplificado se anotaría:

```
(ls FBI).
```

La tendencia a minimizar el esfuerzo hace que muchas siglas se pronuncien como palabras, lo cual es mucho más eficiente –y más breve– que deletrearlas. Tal suele ser el caso de CAMPSA, CIA, REPSOL, etc. Estos casos simplemente *no se marcan*. Por otro lado, no siempre es posible pronunciar las siglas como palabras, debido a combinaciones impronunciables de consonantes. Estos casos suelen ser tratados por los hablantes de forma mixta, con una parte que se deletrea y otra que se lee, como en CSIC, leído *ze-sik* en lugar de *ze-ese-i-ze*. ¿Cómo anotar estos casos? No es fácil, ya que debemos mantener la transcripción original y al mismo tiempo hacer constar su pronunciación. La solución que se propone cumple ambos propósitos a costa de incrementar ligeramente el coste de anotación. Se trata de añadir un nuevo atributo, llamado *pronunciation*, que da cuenta de la pronunciación. Por ejemplo:

```
<l type="acronym" pronunciation="C-SIC"> CSIC </l>,
```

que en formato simplificado se anotaría:

```
(ls CSIC C-SIC).
```

En realidad, el atributo *pronunciation* es opcional y sólo se añade cuando resulta necesario. Cuando no aparece, se toma por defecto una pronunciación completamente deletreada. Es decir,

```
(ls FBI)
```

sería equivalente a:

```
(ls FBI F-B-I).
```

### 3.1.2 Palabras extranjeras.

Nos referimos a palabras tomadas de otra lengua cuya pronunciación no se ha castellanizado, es decir, palabras que se escriben y pronuncian en su lengua original, como *software*, *light*, etc. Suelen ser palabras en inglés, sobre todo nombres propios, términos técnicos y palabras en argot. Para estos casos se ha creado un nuevo tipo de disfluencia léxica, llamado *foreign*, que hará uso del atributo *pronunciation* para dar cuenta de la transcripción en castellano de la palabra extranjera. Por ejemplo:

```
<l type="foreign" pronunciation="sofwar"> software </l>,
```

que en formato simplificado se anotaría:

```
(lx software sofwar).
```

O también:

```
<l type="foreign" pronunciation="lait"> light </l>
```

que en formato simplificado se anotaría:

```
(lx light lait).
```

### 3.1.3 Solapamientos.

En los diálogos que se van a anotar aparecen dos tipos de solapamientos:

1. los que marcan la transición entre dos turnos:
  - (a) suelen ser bastante largos,
  - (b) en ambos casos la señal solapada forma parte (es final o inicio) de un turno,
  - (c) el turno entrante (el del hablante que interrumpe) continua después de manera normal, es decir, sin solapamiento, y
  - (d) normalmente estos solapamientos están ya anotados en la transcripción original de los diálogos;
2. los que suceden en el interior del turno de un hablante:
  - (a) suelen ser cortos,
  - (b) tienen una función fática, el hablante que interrumpe trata de mostrar acuerdo, confirmar una idea, etc. mediante una palabra, una breve expresión o incluso un sonido gutural,
  - (c) dada su corta duración, estas intervenciones pueden incluso suceder entre dos palabras de la intervención principal, pero también pueden superponerse,
  - (d) no forman parte de ningún turno, y
  - (e) no suelen estar anotados en la transcripción original de las entrevistas.

Ambos tipos de solapamientos van a recibir el mismo tratamiento, que será coherente con los criterios de cortado de los turnos con solapamiento. En todos los casos la parte solapada del primer hablante marcará el final de un turno, y la parte solapada del segundo hablante marcará el principio del turno siguiente. Las partes solapadas se marcarán de forma que después puedan ser o bien descartadas o bien procesadas. Utilizaremos para ello la marca XML *o* (de *overlap* en inglés):

H1[15]: no es normal que los jóvenes se reúnan en un parque con unas bolsas del supermercado y se pongan a beber como cosacos . Los padres <o> deberían hacer algo </o> .

H2[16]: <o> y la educación en la escuela </o> . la escuela es importante .

H1[17]: pero los padres tienen mayor responsabilidad .

Cuando el solapamiento no esté marcado en la transcripción original, será necesario dividir el turno donde sucede en dos partes, una anterior y otra posterior a la interrupción, lo cual implica además crear un nuevo turno para el hablante que interrumpe. De hecho, muchos de los solapamientos marcados en las anotaciones UAM originales corresponden en realidad a pequeñas interrupciones elevadas a la categoría de turnos. En el siguiente ejemplo H2 aprueba lo que dice H1 diciendo "claro" de forma superpuesta con "el esfuerzo". Sin embargo, en la transcripción original no hay ninguna referencia a esa intervención de H2:

H1[6]: quiero decir que no merece la pena el esfuerzo . tardaríamos demasiado .

Aplicando lo dicho anteriormente nos quedaría lo siguiente:

H1[6]: quiero decir que no merece la pena <o> el esfuerzo </o> .

H2[7]: <o> claro </o> .

H1[8]: tardaríamos demasiado .

Nótese la diferencia entre los dos ejemplos anteriores. En el primer ejemplo, el hablante H1 es interrumpido pero puede terminar su intervención y pasa a escuchar lo que H2 ha empezado a decirle. Por su parte, H2 empieza su turno hablando a la vez que H1 pero luego continúa su intervención en solitario. Cuando H1 vuelve a intervenir, lo hace para replicar a lo que acaba de decir H2 en el turno 16, pero el turno 17 no es continuación del turno 15. Por el contrario,

en el segundo ejemplo H1 no llega a ser interrumpido por H2, sino que continúa su intervención como si éste último no hubiera dicho nada. Es decir, el turno 8 es continuación del turno 6. Esta circunstancia ha de ser anotada porque desde el punto de vista del modelo de lenguaje la intervención completa de H1 estaría formada por la concatenación de los turnos 6 y 8. Para ello definimos una nueva marca, llamada *continued*, que no tendrá atributos. Su sola presencia indicará que el turno actual será continuado en un turno posterior. Así el segundo ejemplo quedaría como sigue:

H1[6]: quiero decir que no merece la pena <o> el esfuerzo </o> <continued> .

H2[7]: <o> claro </o> .

H1[8]: tardaríamos demasiado .

En formato simplificado se utilizará la marca (*o TEXTO*) para los solapamientos, y la marca [*CONTINUA*] para marcar la continuación del turno actual en un turno posterior:

H1[6]: quiero decir que no merece la pena (o el esfuerzo) [CONTINUA] .

H2[7]: (o claro) .

H1[8]: tardaríamos demasiado .

### 3.1.4 Ruidos, segmentos no transcritos y cortes en la grabación.

La división del turno natural de un hablante en varios pequeños turnos artificiales se producirá también cuando la intervención del hablante se vea solapada con ruido ambiente (risas, aplausos, música, etc.) y cuando aparezca alguna de las marcas [*NO TRANSCRITO*] o [*CORTE*]. En tales casos se utilizará también la marca [*CONTINUA*] (en formato simplificado) para indicar la continuación del turno. Esta regla surge, una vez más, para mantener la coherencia de las anotaciones con las reglas de cortado: cada señal ha de corresponder a un turno distinto y ha de tenerse la transcripción completa de todas las señales.

Resulta difícil justificar la marca de continuación cuando se encuentra un segmento no transcrito o un corte en la grabación. Estrictamente no sabemos qué relación puede unir a los segmentos anterior y posterior a un segmento no transcrito o a un corte en la grabación. No obstante, si en la transcripción original formaban parte del mismo turno, mantendremos esa relación mediante la marca de continuación.

Por ejemplo, el siguiente turno:

H3[34]: pues hoy estábamos revisando algunas (nw de las canciones de Serrat) cuando hemos recibido varias llamadas telefónicas felicitándonos .

se dividirá en los dos siguientes:

H3[34]: pues hoy estábamos revisando algunas (nw de las canciones de Serrat) [CONTINUA] .

H3[35]: cuando hemos recibido varias llamadas telefónicas felicitándonos .

lo cual significa que tendremos dos señales, una para el turno 34 –que no incluirá el segmento contaminado con ruido– y otra para el turno 35. Obsérvese la utilidad de la marca [*CONTINUA*] para enlazar ambos turnos, puesto que efectivamente uno es continuación del otro.

Análogamente, el siguiente turno:

H1[15]: lo cierto es que mi madre acostumbraba a [NO TRANSCRITO] pueblo que no me gustaba nada, aunque pasaba allí todos los veranos .

se desdoblaría en los dos siguientes:

H1[15]: lo cierto es que mi madre acostumbraba a [NO TRANSCRITO] [CONTINUA] .

H1[16]: pueblo que no me gustaba nada, aunque pasaba allí todos los veranos .

Nuevamente generaremos dos ficheros de señal distintos, uno para el turno 15 –correspondiente únicamente a la parte transcrita– y otro para el turno 16. Nótese cómo a pesar de que mantenemos la relación de continuidad entre los turnos, no hay relación aparente entre los bloques de texto anterior y posterior al segmento no transcrito.

### **3.2 Segunda fase: anotación de fenómenos sintácticos y pragmáticos.**

En la segunda fase las señales ya estarán cortadas y las anotaciones acústicas y léxicas corregidas y aumentadas, de modo que el anotador podrá concentrarse únicamente en la detección y caracterización de los fenómenos sintácticos (reformulaciones) y pragmáticos (marcadores de discurso). Esta segunda fase se efectuará siguiendo estrictamente el mismo esquema: procedimiento, marcas, etc. aplicado en INFOTREN, al cual nos remitimos.

El anotador deberá escuchar las señales de nuevo para asegurarse de la presencia o no presencia de un cierto fenómeno. Sin embargo, la atención que tendrá que poner en esas nuevas escuchas será radicalmente inferior. De hecho los fenómenos sintácticos y pragmáticos podrían anotarse casi exclusivamente sobre el texto, siempre que las anotaciones acústica y léxica fueran suficientemente precisas.

## Apéndice 1. Guía práctica para el anotador.

Como ya hemos comentado más arriba, el anotador utilizará como únicos recursos el editor de texto genérico *gvim* y la aplicación de visualización de señales de voz *XWAVES*, en entorno UNIX. En esta sección trataremos de describir la forma de interaccionar con el entorno de trabajo (ventanas, shell UNIX, etc.) y con las dos herramientas mencionadas, ya que el anotador deberá manejar todo ello con soltura.

### El entorno de trabajo.

#### Para empezar.

El *jack* de los auriculares (el negro) deberá conectarse en la parte lateral izquierda de la máquina, en el conector que ocupa la posición inferior derecha. Después se enciende el monitor, se teclea el nombre de usuario y la *password* e inmediatamente se pone en marcha una nueva sesión UNIX que nos muestra un escritorio con varios iconos.

#### Abriendo y cerrando ventanas.

En general, el botón izquierdo del ratón sirve para poner en marcha una tarea y el botón derecho para desplegar un menú. Por ejemplo, para abrir cualquiera de los iconos del escritorio basta colocarse encima del icono y hacer click con el botón izquierdo del ratón. Si apretamos el botón derecho del ratón sobre el fondo del escritorio, aparecerá el menú general, que por ejemplo incluye la opción *Log Out* para salir de la sesión UNIX.

En la parte inferior derecha del escritorio aparecerán el reloj y el control de audio. Este último nos permitirá aumentar o disminuir el volumen de reproducción. Para comprobar el funcionamiento del sistema de ventanas abriremos el control de audio. Si queremos que la ventana ocupe toda la pantalla, hacemos click con el botón izquierdo del ratón sobre el cuadradito que aparece en la parte superior derecha de la ventana. Para devolver la ventana a su tamaño normal volvemos a hacer click en el mismo cuadradito. Para devolver la ventana a la forma de icono basta apretar con el botón izquierdo del ratón sobre el puntito que aparece junto al cuadradito en la parte superior derecha de la ventana.

En la ventana distinguimos tres zonas importantes: el borde –una franja muy fina que la envuelve por completo–, la barra del título –que ocupa la parte superior y tiene una anchura aproximada de un centímetro– y el espacio interior. El borde y la barra del título se manejan siempre igual, mientras que el espacio interior tendrá un comportamiento distinto dependiendo de lo que muestre la ventana.

Para mover la ventana se mantiene apretado el botón izquierdo del ratón sobre la barra del título y se desplaza el ratón hasta la posición que deseemos.

Para redimensionar la ventana colocamos el ratón en una de las esquinas del borde, mantenemos apretado el botón izquierdo del ratón y movemos éste hasta conseguir el tamaño deseado. También podemos hacer lo mismo sobre alguno de los bordes rectos, pero entonces la ventana crece o decrece sólo en una dimensión.

Para *matar* una ventana colocamos el ratón en la barra del título y hacemos click con el botón derecho, lo cual despliega un menú. Con el botón derecho apretado nos desplazamos hasta la última de las opciones que muestra dicho menú (*Quit*) y lo soltamos ahí. La ventana desaparecerá.

### El *shell* de UNIX.

En la parte inferior izquierda del escritorio aparece un icono con el nombre de *bash*. Si lo abrimos nos encontramos con una ventana que presenta un procesador de comandos UNIX, que también recibe el nombre de *shell*. En la ventana aparece un pequeño texto, llamado *prompt*, con el siguiente aspecto:

```
luisja@fonos[~]$
```

El *prompt* indica que el sistema está preparado para recibir cualquier comando UNIX que deseemos introducir. Habitualmente consta de los siguientes elementos:

- el nombre del usuario,



- el símbolo '@',
- el nombre de la máquina,
- entre corchetes, el *path* (camino) completo del directorio de trabajo, y
- el símbolo '\$', tradicional en los *shell* de UNIX.

El símbolo ~ representa el directorio raíz del usuario. A continuación mostramos una tabla con los comandos UNIX más comunes:

Comando	Descripción	Opciones
ls	Lista el contenido del directorio de trabajo. Si tiene argumentos, lista sus argumentos. Si uno de los argumentos es un directorio, lista el contenido del directorio.	-l Listado largo (alias: ll). -a Lista ficheros ocultos. -d Lista sólo el nombre de los directorios, no su contenido.
cd	Cambia al directorio que se le suministre como argumento. El argumento puede ser un camino relativo (por ejemplo, "cd aent001a") o absoluto (por ejemplo, "cd /usr/local"). Si no tiene argumentos, nos lleva al directorio raíz. El comando "cd .." nos lleva al directorio padre.	
mkdir	Crea un nuevo directorio; por ejemplo, "mkdir carpeta1".	
cp	Copia un fichero en otro. Puede utilizarse para crear un nuevo fichero igual al original en el mismo directorio: "cp fich1 fich2". También puede utilizarse para crear un nuevo fichero con el mismo nombre que el original pero en otro directorio; por ejemplo, "cp fich1 carpeta1" copiaría el fichero "fich1" (con el mismo nombre) en el directorio "carpeta1".	-r Se utiliza para copiar recursivamente todo el contenido de un directorio en otro. Por ejemplo, "cp -r carpeta1 carpeta2" crearía el directorio "carpeta2" con el mismo contenido que "carpeta1". Si el directorio "carpeta2" ya existe entonces el efecto del comando anterior es crear "carpeta1" dentro de "carpeta2".
mv	Cambia el nombre de un fichero, o bien cambia la ubicación del mismo. Por ejemplo, "mv fich1 f1" cambia el nombre de "fich1" a "f1". Si el segundo argumento del comando es un directorio, entonces el efecto sería <i>mover dentro</i> del directorio el fichero pasado como primer argumento.	
rm	Borra un fichero. No borra directorios. ES IRREVERSIBLE: una vez borrado, el fichero no se puede recuperar.	-r Borra recursivamente un directorio y todo su contenido. MUY PELIGROSO.
more	Lista en pantalla el contenido de un fichero de texto. Para avanzar por el listado, teclear RETURN o SPACE. Para finalizar la operación, teclear q.	
man	Muestra información sobre el uso y la funcionalidad de un comando. La información se muestra de la misma forma que lo hace el comando more. Hágase por ejemplo "man ls".	
sgplay	Permite escuchar un fichero de señal.	

Inicialmente el procesador de comandos se encuentra en el directorio raíz del usuario. El directorio raíz de un usuario XXX suele ubicarse en /home/XXX. Para cambiar de ubicación se utilizará el comando cd y para ver el contenido de un directorio el comando ls o ls -l (que abreviadamente es ll). Para saber en qué lugar del árbol de ficheros nos encontramos en un momento dado basta con consultar el *prompt*. Por ejemplo:

```
luisja@fonos[/usr/local]$
```

dice que nos encontramos en el directorio `/usr/local`. Pero también podemos usar el comando `pwd`:

```
luisja@fonos[/usr/local]$ pwd
/usr/local
luisja@fonos[/usr/local]$
```

El comando `rm` (para borrar ficheros y/o directorios) es especialmente peligroso ya que no es reversible: una vez que borramos un fichero es imposible recuperarlo. Por tanto, conviene utilizar este comando con precaución y sólo cuando estamos absolutamente seguros de que queremos borrar un fichero.

### Algunos trucos del *bash*.

1. Cuando estamos escribiendo el nombre de un fichero o de un directorio, mediante la **tecla del tabulador** podemos completarlo automáticamente, o al menos llegar tan lejos como sea posible sin que haya ambigüedad. Es decir, supongamos que en mi directorio tengo tres ficheros: `pepe`, `pepa` y `maria`. Si ahora tecleo `gvim m<TAB>`, el nombre `maria` se completa automáticamente, ya que no existe ningún otro fichero cuyo nombre empiece por `m`. Sin embargo, si tecleo `gvim p<TAB>`, sólo se expande la cadena `pep`, ya que hay dos ficheros: `pepe` y `pepa`, que comparten esa raíz común. Obviamente, el resto del nombre deberá completarlo yo, porque la máquina no puede saber a cuál de los dos me refiero.
2. Mediante los **cursores** `↑` y `↓` es posible recorrer la historia de los comandos tecleados hasta la fecha en el *bash*. Incluso podemos alcanzar y editar comandos tecleados en una sesión anterior. Mediante la tecla `↑` retrocedemos en la historia y mediante `↓` avanzamos. Una vez alcanzado un cierto comando, para ejecutarlo basta teclear `RETURN`. Si no queremos ejecutarlo tal cual, sino cambiando una parte del comando, con los cursores `←` y `→` podemos retroceder y avanzar por el texto del comando para borrar y/o añadir caracteres.
3. Con el botón izquierdo del ratón podemos seleccionar y copiar cualquier texto, y con el botón central pegarlo en la línea de comando. Esto es útil, por ejemplo, si listamos el contenido de un directorio, después seleccionamos el nombre del fichero que queremos procesar y finalmente lo pegamos en la línea de comando.
4. El asterisco (\*) tiene un significado muy especial en el *bash*. Sirve para representar cualquier cadena de caracteres (de longitud 0 o mayor). Por ejemplo, si queremos listar sólo los ficheros que empiezan por `p`, ejecutaríamos el comando `ls p*`. Para listar los ficheros que terminan en `.esps` ejecutaríamos `ls *.esps`. Un último ejemplo: ficheros que contienen la cadena `con` y terminan en `.dis` se representarían mediante `*con*.dis`.

### El editor *gvim*.

Se trata de una herramienta de edición basada en el antiguo editor de pantalla *vi*, que fue desarrollado originalmente como parte del sistema operativo UNIX. En realidad, *gvim* es completamente compatible con *vi*, pero además incorpora notables mejoras: el uso de los cursores y de las teclas de borrado, el coloreado *inteligente* del texto, el interfaz gráfico con menús, etc. A diferencia de otros editores –tipo *Microsoft WORD*– los ficheros generados con *gvim* son ficheros de texto, es decir, son completamente portables, legibles en cualquier plataforma y mediante cualquier editor.

Para editar un fichero –digamos que `aent001a.dis`– con el *gvim* basta hacer:

```
luisja@fonos[~]$ gvim aent001a.dis
```

Aparecerá una nueva ventana mostrando una barra de menú, una barra de comandos y el texto del fichero (o nada si el fichero es nuevo). De la barra de menú sólo utilizaremos los dos primeros menús: *Fichero*, para guardar y salir, y *Edición*, para cortar, copiar, pegar y buscar texto. La barra de comandos presenta una serie de pequeños iconos que permiten acceder de forma rápida a algunas de las funciones de los menús: guardar, deshacer cambios, rehacer cambios, cortar, copiar, pegar, buscar, etc. Si ponemos el ratón sobre uno de esos iconos, al cabo de un instante aparece un mensaje indicando su funcionalidad.

Para editar es necesario colocar el puntero del ratón en el interior de la ventana de edición. El editor se inicia en modo *comando*, es decir, podemos ejecutar comandos pero no escribir directamente texto en el fichero. Para pasar a modo *edición*, lo más sencillo es teclear el comando *i*. Al hacerlo, en la parte inferior izquierda de la pantalla aparecerá el mensaje – *INSERT* – indicando que podemos escribir y borrar caracteres en el fichero. En modo *edición* podemos movernos a través del fichero con los cursores, pero también mediante el ratón o utilizando la barra de *scroll* que aparece a la derecha de la ventana. Para volver a modo *comando* basta pulsar la tecla *ESC*.

Estando en modo edición, podemos borrar caracteres utilizando las teclas de retroceso (*backspace*) y borrado (*supr*). También podemos copiar o cortar y luego pegar trozos de texto seleccionándolos primero con el botón izquierdo del ratón, y utilizando después los elementos de menú correspondientes. En concreto, para pegar también podemos utilizar el botón central del ratón, previo posicionamiento del cursor en el punto donde queremos pegar el texto.

## La herramienta de visualización y cortado *XWAVES*.

El *XWAVES* es una potente herramienta de visualización, análisis y procesamiento de señales de voz. En este caso se va a utilizar únicamente para escuchar y cortar turnos en un conjunto de diálogos. En primer lugar, es necesario preparar el sistema de ventanas mediante el siguiente comando:

```
luisja@fonos[~]$ mkf
```

A continuación, suponiendo que existe un fichero llamado *aent001a.esps* con la señal correspondiente a un diálogo, para visualizarla haremos:

```
luisja@fonos[~]$ xwaves aent001a.esps &
```

El símbolo *&* se añade para evitar que se bloquee el procesador de comandos *bash* mientras se ejecuta el *XWAVES*. De esta forma podemos seguir utilizando el mismo *bash* para ejecutar otros comandos. De lo contrario, el *bash* quedaría bloqueado a la espera de que finalizase el *XWAVES*.

Aparecerán tres ventanas nuevas, una de ellas muy alargada –que llamaremos *ventana de señal*– con la forma de onda de la señal; otra, arriba a la izquierda, que recibe el nombre de *panel de control*; y una tercera, la más pequeña, arriba a la derecha, que permite configurar distintos elementos de la herramienta, si bien el anotador no tendrá que utilizarla. A continuación pasamos a describir con más detalle la ventana de señal y el panel de control.

### La ventana de señal.

Mediante el ratón será posible recorrer la señal y seleccionar trozos de la misma, y mediante un menú desplegable podrán efectuarse distintas operaciones, entre ellas varios tipos de *zoom*, avanzar y retroceder, reproducir toda la señal o una parte de la misma, guardar un trozo de señal en un fichero, etc. Este menú aparece pulsando el botón derecho del ratón sobre la señal.

Por otro lado, si pulsamos el botón izquierdo y lo mantenemos apretado mientras movemos el ratón, el efecto es seleccionar un trozo de señal, que aparecerá marcado en cuanto soltemos el botón izquierdo. Una vez marcado un trozo de señal, para escucharlo basta pulsar el botón central del ratón.

En la parte superior de la ventana aparece información sincronizada con el cursor que recorre la señal, concretamente el instante (en segundos) en el que se encuentra dicho cursor. En general –salvo que hagamos un *zoom* inverso completo– sólo veremos un fragmento de la señal, por lo que los instantes reportados corresponderán únicamente a ese fragmento.

El lugar que ocupa el fragmento dentro de la señal aparece marcado sobre una barra horizontal estrecha en la parte superior de la ventana. Dicha barra representa toda la señal y si colocamos el puntero del ratón dentro de ella, el indicador de tiempo se sincroniza con ella. Si además presionamos el botón central del ratón sobre un punto de dicha barra, el efecto es mostrar un segmento de señal alrededor del punto marcado. Así, por ejemplo, si deseamos ver el instante 125 pero la ventana muestra sólo la señal comprendida entre los instantes 23 y 45, haremos lo siguiente:

1. colocar el puntero del ratón dentro de la barra horizontal,
2. mover el puntero sobre la barra hasta que el indicador de tiempo marque 125 segundos, y

3. presionar el botón central del ratón.

Con el puntero del ratón en la barra horizontal, el botón izquierdo del ratón permite avanzar en la señal, y el botón derecho retroceder.

Para terminar este apartado, enumeramos y describimos algunas de las operaciones que pueden realizarse mediante el menú desplegable:

Opción del menú	Descripción
play between marks	Reproduce el trozo de señal marcado en la ventana.
play window contents	Reproduce todo el trozo de señal presentado en la ventana.
play entire file	Reproduce toda la señal.
play to end of file	Reproduce desde la posición actual del cursor hasta el final del fichero.
page ahead	Avanza un cierto espacio en la señal (como media ventana).
page back	Retrocede un cierto espacio en la señal (como media ventana),
window ahead	Avanza una ventana hacia adelante. Es decir, presenta una ventana de la misma longitud que la actual, a partir del final de la actual.
window back	Retrocede una ventana.
bracket markers	Presenta en la ventana únicamente el segmento marcado, ampliando y redibujando la señal. Es una forma de <i>penetrar</i> directamente en el segmento de señal que nos interesa, para verlo con más detalle.
zoom in	Amplia la señal alrededor del cursor, es decir, vemos un segmento más pequeño pero con mayor detalle.
zoom out	Realiza un zoom inverso alrededor del cursor, es decir, vemos un segmento de señal más grande pero con menor detalle.
zoom full out	Realiza un zoom inverso completo, de forma que vemos toda la señal.
save segment in file	Almacena el segmento marcado en un nuevo fichero, dando a éste el nombre que se indique en el campo <i>OUTPUTfile</i> del panel de control.

### El panel de control.

El panel de control permite especificar el nombre de la señal de entrada (campo *INPUTfile*), también **muy importante** el nombre que queremos dar al fichero donde guardar un trozo de señal (campo *OUTPUTfile*), y ejecutar un comando sobre el fichero de señal abierto por el *XWAVES* (campo *COMMAND*). También presenta varios botones, entre ellos uno con la leyenda *QUIT!* que sirve para finalizar la ejecución del *XWAVES*, y otro con la leyenda *xwaves MANUAL*, que permite acceder al manual de uso del *XWAVES*. Por último, el campo *Attach* presenta varios botones que permiten ejecutar aplicaciones secundarias como el *xlabel*, el *xchart* o el *xspectrum*. No obstante, de todo lo anterior el anotador tan sólo manejará los campos *INPUTfile* y *OUTPUTfile*, y el botón *QUIT!*.

## Apéndice 2. Lista de etiquetas EHU actualizada para CO-UAM.

Tabla 1: Ruidos.

(nw)	Ruido exterior aislado
(nw <b>TEXTO</b> )	Ruido exterior solapado con voz
(na)	Ruido de aspiración o espiración producido por el locutor
(nl)	Ruido de labios producido por el locutor
(nt)	Ruido de tos

Tabla 2: Disfluencias acústicas.

(p)	Pausa de silencio
(fa)	Pausa hablada, vocalizada /a/
(fe)	Pausa hablada, vocalizada /e/
(fm)	Pausa hablada, nasalizada /m/
(fb)	Pausa hablada, inclasificable
(a X)	Alargamiento de la letra X

Tabla 3: Disfluencias léxicas.

(lm <b>P</b> Bien <b>P</b> Mal)	Palabra mal pronunciada
(lu <b>P</b> Entera <b>P</b> Cortada)	Palabra cortada
(lg sí)	Afirmación gutural
(lg no)	Negación gutural
(ls <b>CSIC C-SIC</b> )	Siglas (deletreadas total o parcialmente)
(lx light lait)	Palabra extranjera

Tabla 4: Disfluencias sintácticas.

(rr <b>BLOQUE</b> )	Repetición
(rs <b>BLOQUE</b> )	Sustitución
(ri <b>BLOQUE</b> )	Reformulación con inserción de elementos
(rd <b>BLOQUE</b> )	Reformulación con borrado de elementos
	<b>BLOQUE = (m TEXTO) (s TEXTO) (c TEXTO)</b>
	(m <b>TEXTO</b> ) Reparandum: lo que se repite, sustituye o reformula
	(s <b>TEXTO</b> ) Señal: marca de edición (opcional)
	(c <b>TEXTO</b> ) Corrección: la parte que sustituye al reparandum
(b <b>TEXTO</b> )	Frase abandonada

Tabla 5: Marcadores de discurso.

<b>(do hola)</b>	De apertura
<b>(dc adios)</b>	De cierre
<b>(da bueno)</b>	De aceptación
<b>(dr nada más)</b>	De rechazo
<b>(de o sea)</b>	De explicación
<b>(dq por favor)</b>	De petición
<b>(df entonces)</b>	De relleno
<b>(dx hala)</b>	De exclamación

Tabla 6: Solapamientos.

<b>(o TEXTO)</b>	Fragmento en el que aparecen solapadas dos o más voces
------------------	--

Tabla 7: Otros.

<b>[NO TRANSCRITO]</b>	Fragmento no transcrito (no se oye o no se entiende)
<b>[CORTE]</b>	Corte en la grabación. Pérdida de continuidad debida a un problema en la grabación, como por ejemplo el fin de la cinta, o en la transferencia de la señal a formato digital, como por ejemplo la inserción accidental de un trozo de señal perteneciente a otro diálogo.
<b>[CONTINUA]</b>	El turno actual continua en uno posterior (corte debido a ruidos o voces solapados)