

**Facultad de Ingeniería – Universidad de Buenos Aires**  
**Señales y Sistemas TB065**  
**Segundo Cuatrimestre del 2024**

*Curso 1: Rey Vega - Caldentey - Marques Rojo - Florentín*

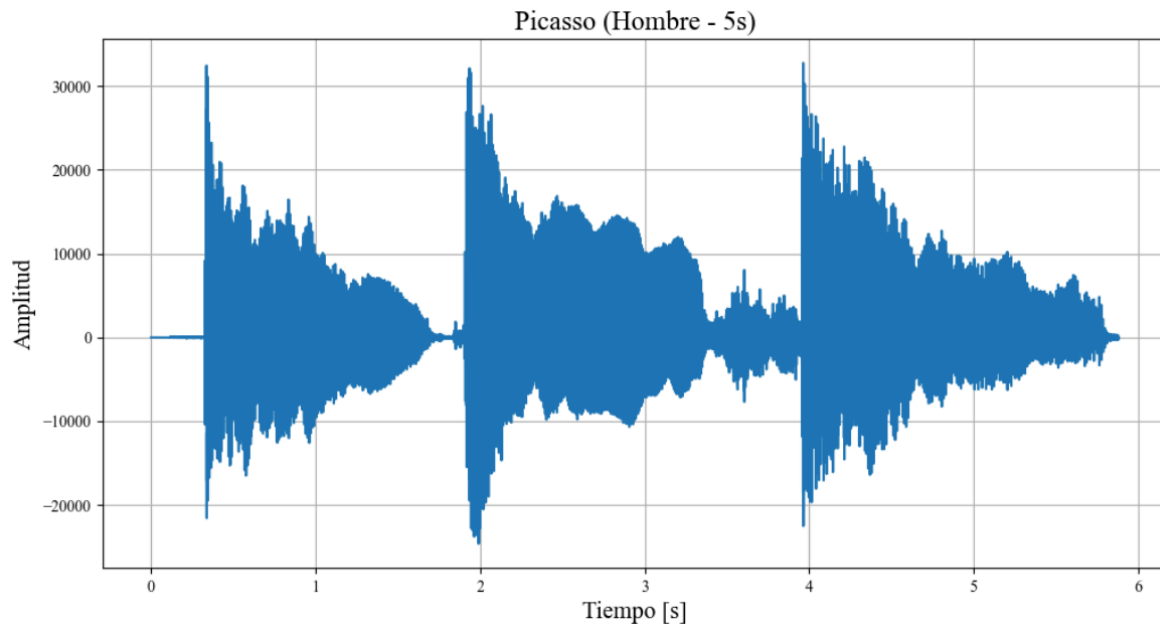
**Trabajo Práctico Especial - Parte 1**  
**Análisis de grabaciones de la palabra “Picasso”**

**Integrantes**

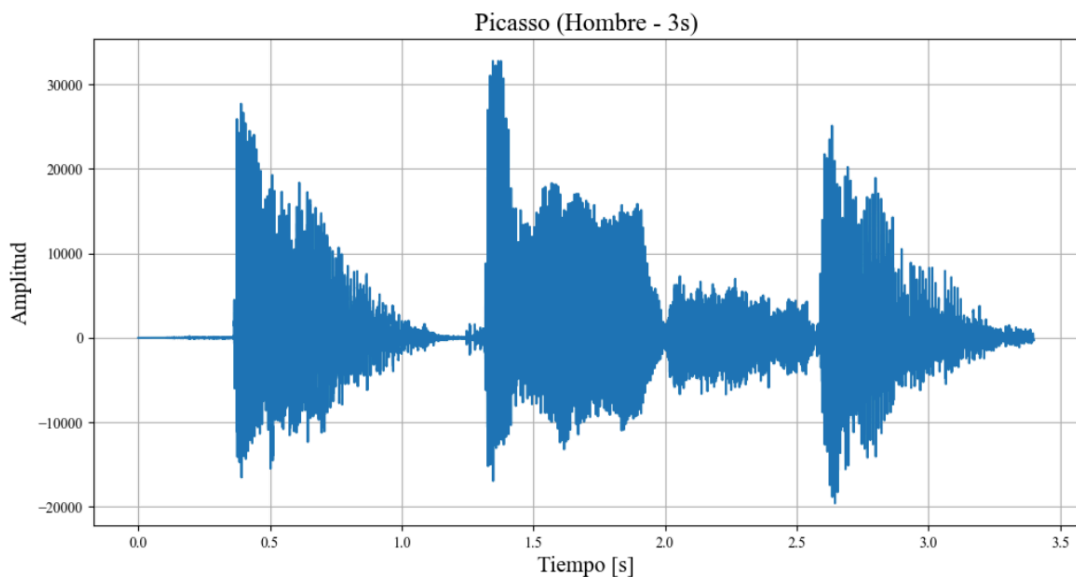
María de las Mercedes Gamberale	108834
Carolina Gonzales Peralta	110804
Ezequiel Sanchez Aliva	103799

## Grabaciones de voz iniciales

Se grabaron dos muestras de voz diciendo la palabra “Picasso” a una velocidad más lenta (marcándose esta como “Hombre - 5s”) y a una velocidad más rápida (identificable como “Hombre - 3s”) y se graficaron las señales de voz, lo que se puede apreciar en las Figuras 1 y 2 respectivamente:

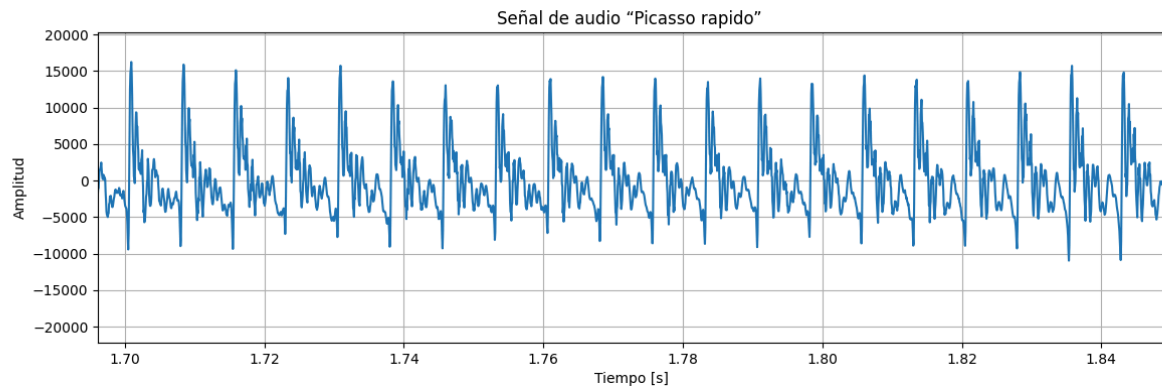


**Figura 1:** Señal temporal de una voz masculina diciendo la palabra “Picasso” a una velocidad lenta.

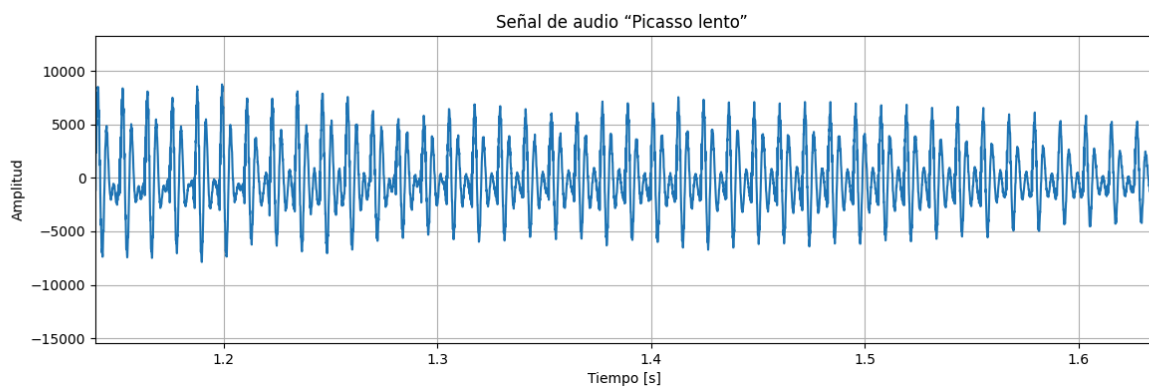


**Figura 2:** Señal temporal de una voz masculina diciendo la palabra “Picasso” a una velocidad rápida.

Luego se extrajeron fragmentos de cada muestra en las cuales se observaron señales periódicas, presentadas en las Figuras 3 y 4.



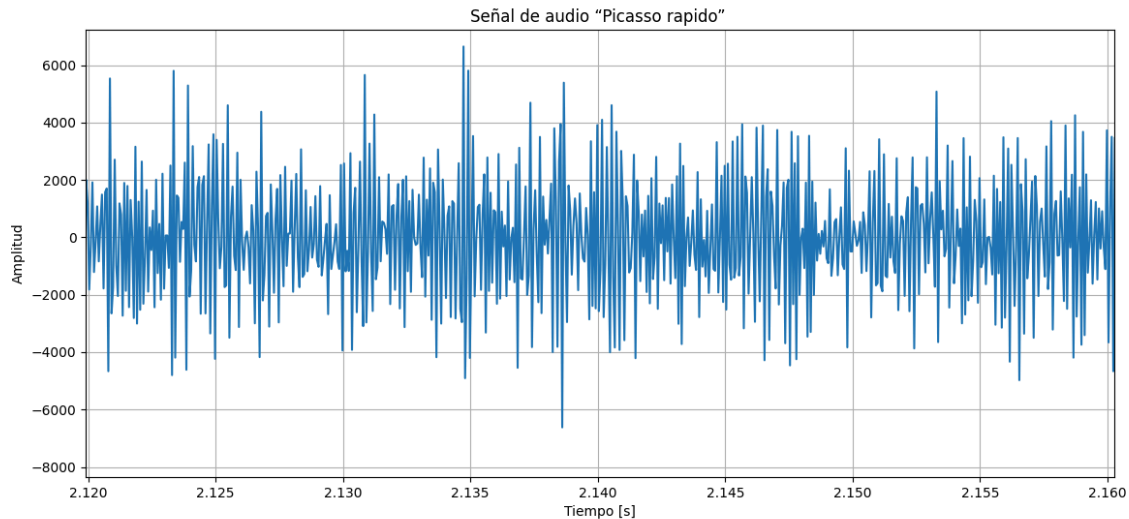
**Figura 3:** Fragmento de la señal de la voz masculina diciendo la palabra "Picasso" a una velocidad rápida en la cual se observó una sección periódica.



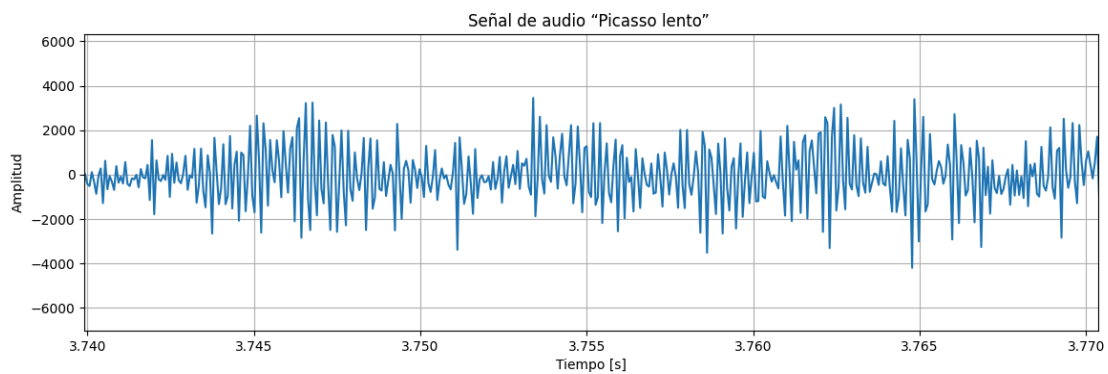
**Figura 4:** Fragmento de la señal de la voz masculina diciendo la palabra "Picasso" a una velocidad lenta en la cual se observó una sección periódica.

Ambos fragmentos se corresponden a vocales, para el fragmento de la muestra rápida se seleccionó la parte donde se escucha la vocal "a" mientras que para la muestra lenta se seleccionó el fragmento donde se escucha la vocal "i". Ambas, al ser vocales (que se producen mediante sonidos armónicos presentes en la vibración natural de la garganta al emitir dichos sonidos), se facilita la observación de la forma periódica de la señal.

Se continuó con la extracción de los fragmentos de cada muestra donde se apreciaban partes de la señal no periódicas, mostradas en las Figuras 5 y 6.



**Figura 5:** Fragmento de la señal de la voz masculina diciendo la palabra "Picasso" a una velocidad rápida en la cual se observó una sección no periódica en la señal.



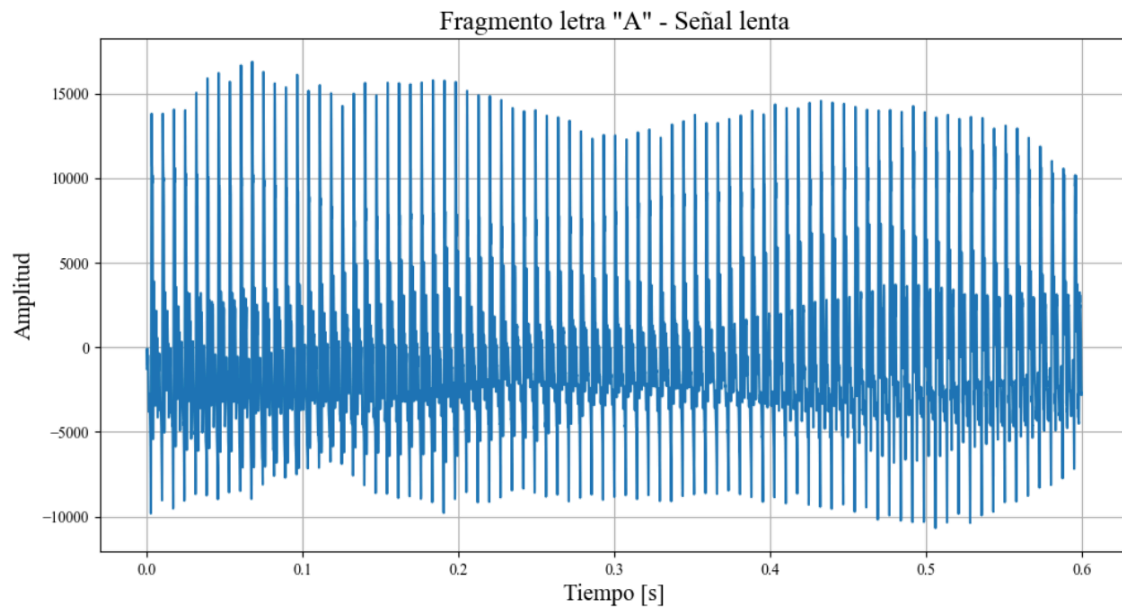
**Figura 6:** Fragmento de la señal de la voz masculina diciendo la palabra "Picasso" a una velocidad lenta en la cual se observó una sección no periódica de la señal.

Ambos fragmentos se corresponden a la parte de la muestra donde se escucha la consonante "s" (para la muestra rápida se escogió un fragmento de la "s" cercano a la vocal "a", mientras que para la muestra lenta fue un fragmento cercano a la vocal "o"). Como se observa, a diferencia de los fragmentos extraídos de vocales (Figura 3 y 4), es difícil encontrar un patrón repetitivo. Esto se debe a que es una letra sorda, al pronunciarla las cuerdas vocales no vibran entonces se producen sonidos aleatorios sin patrón definido. Estos son sonidos de percusión, no armónicos, por lo que no presentan una periodicidad.

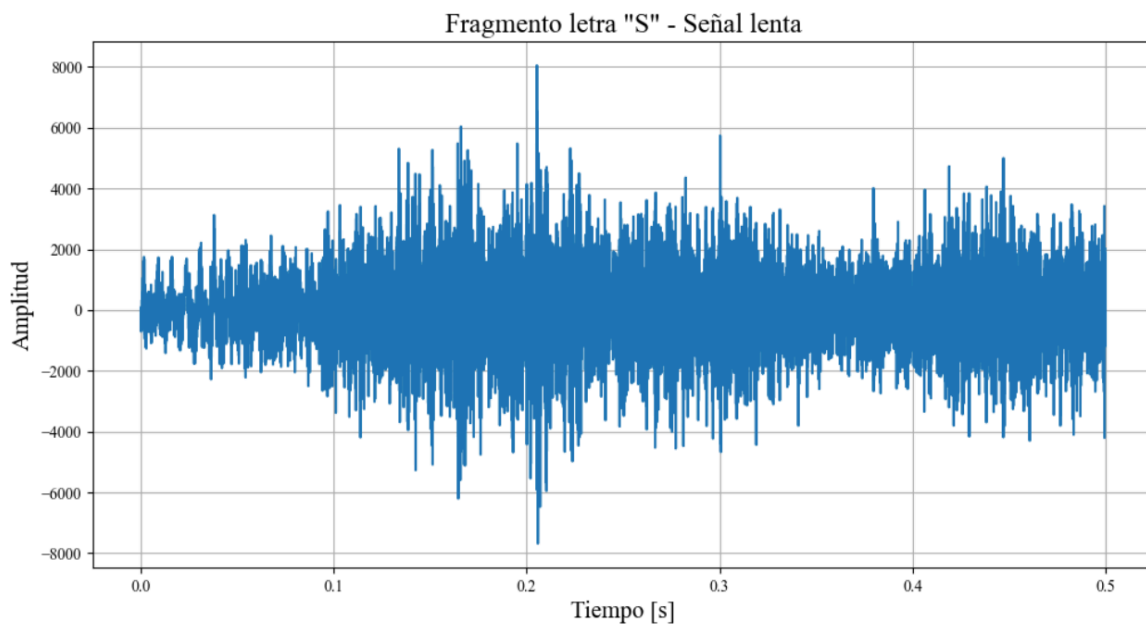
## Comparación entre los fragmentos donde se dicen las letras "a" y "s"

Se realizó la comparación utilizando la señal a una velocidad lenta. Analizando la señal grabada con detenimiento, se estima que en el intervalo de tiempo entre 2,4s y 3,0s se encuentra una letra "a" mientras que en el intervalo entre 3,4s y 3,9s se encuentra una letra "s".

Como se puede observar en las Figuras 7 y 8, existen notables diferencias entre los fonemas mencionados. En el fragmento correspondiente a la letra “a” se observa una periodicidad en la señal, producto de que el sonido de la misma sea armónico y por ende, debido a la vibración de la garganta al repetir el fonema /a/, éste presenta una periodicidad notable como se aprecia en la Figura 7. En contraste, al pronunciarse la “s” se realiza un sonido de percusión ya que, en esencia, la “s” es un silbido sordo. Es por ello que la señal presenta saltos abruptos y variados a lo largo de la pronunciación del mismo fonema de una manera más marcada que en el caso de la “a”.

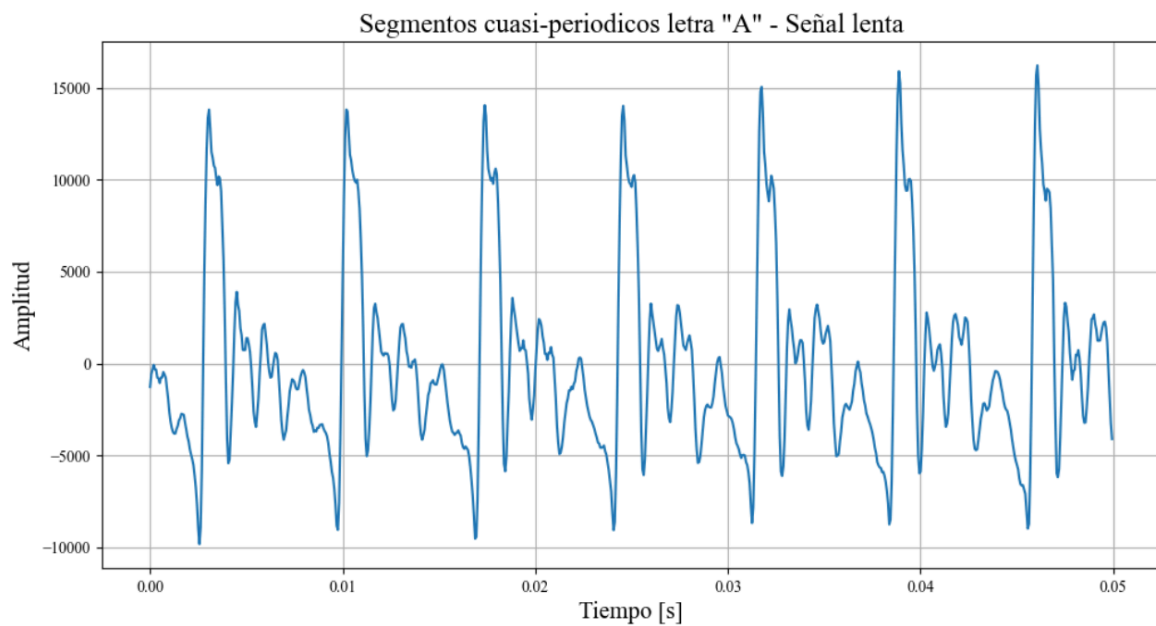


**Figura 7:** Fragmento de la señal de la voz masculina diciendo el fonema /a/ a una velocidad lenta.

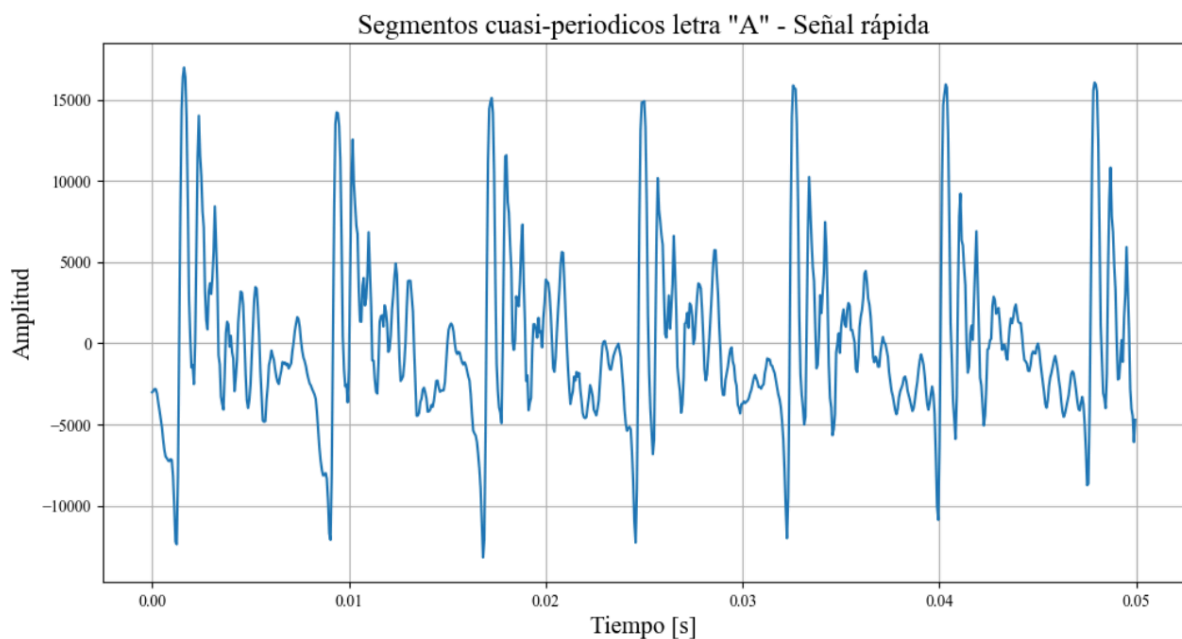


**Figura 8:** Fragmento de la señal de la voz masculina diciendo el fonema /s/ a una velocidad lenta.

En las Figuras 9 y 10 se observan segmentos cuasi-periódicos, correspondientes a la letra “a” tanto de la muestra lenta como de la rápida. A partir de los mismos es posible estimar su período (y por tanto su frecuencia).



**Figura 9:** Fragmento de la señal de la voz masculina diciendo la letra “a” a una velocidad lenta.



**Figura 10:** Fragmento de la señal de la voz masculina diciendo la letra “a” a una velocidad rápida.

En la señal lenta se estimó un periodo de  $T = 7,19 \text{ ms}$ , por lo que su frecuencia sería de  $f = 1/T = 139 \text{ Hz}$ . En la señal rápida la estimación fue de  $T = 7,88 \text{ ms}$ , por lo que su frecuencia sería de  $f = 127 \text{ Hz}$ . En ambos casos la estimación fue realizada visualmente, y es relativamente similar.

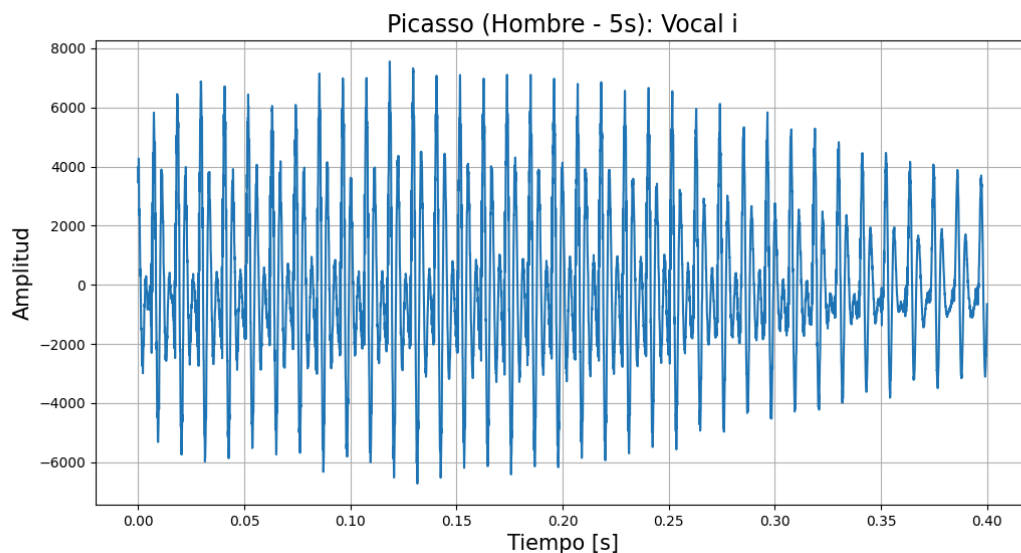
## Análisis en frecuencia de las vocales de la señal

Para el análisis en frecuencia de las vocales de la señal se realizó lo siguiente: primero se fragmentó la señal grabada a la sección en la cual se decía la vocal deseada y se realizó la FFT del fragmento completo. Luego, se buscó un solo periodo de la señal recortándola de a poco hasta visualizar un solo periodo completo, al que finalmente se le realizó la FFT. Este proceso se repitió para la señal lenta (“Hombre - 5s”) y la rápida (“Hombre - 3s”). Los resultados se mostrarán a continuación.

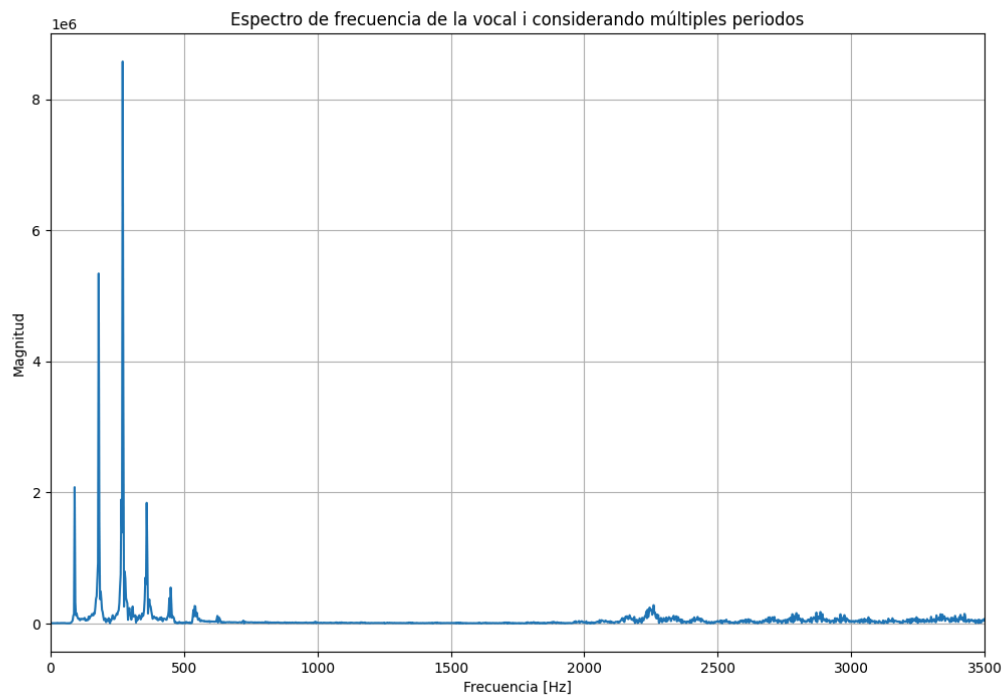
### Vocal “i”

#### *Señal lenta*

Se presenta el fragmento temporal considerando múltiples períodos en los cuales se recita la vocal “i” y la realización de la FFT del mismo en las Figuras 11 y 12 respectivamente.

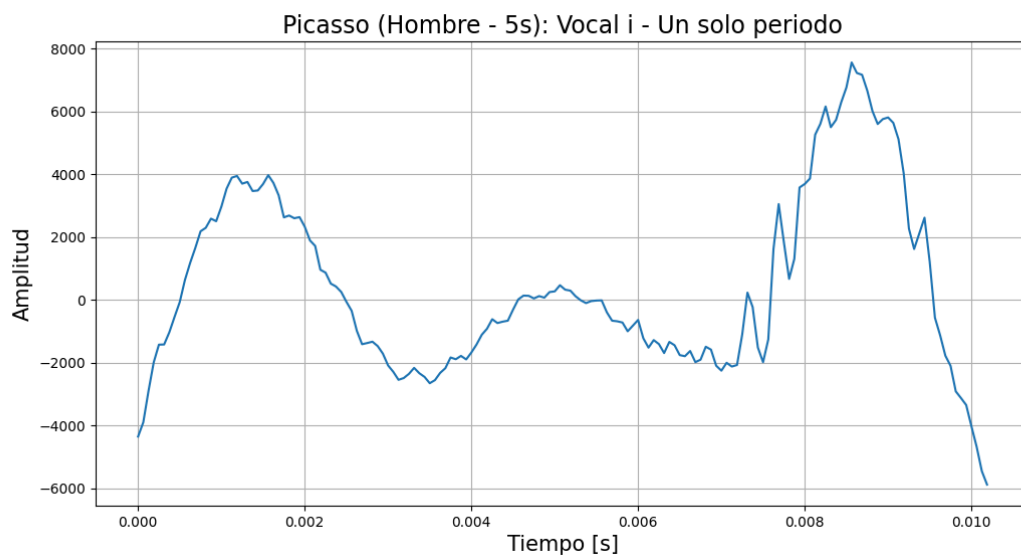


**Figura 11:** Fragmento de varios periodos de la señal lenta diciendo la letra “i.”



**Figura 12:** FFT de varios periodos de la señal diciendo la letra “i” a una velocidad lenta.

Se presenta el fragmento temporal considerando un solo periodo en los cuales se recita la vocal “i” y la realización de la FFT del mismo en las Figuras 13 y 14 respectivamente.



**Figura 13:** Fragmento de un periodo de la señal diciendo la letra “i” a una velocidad lenta.

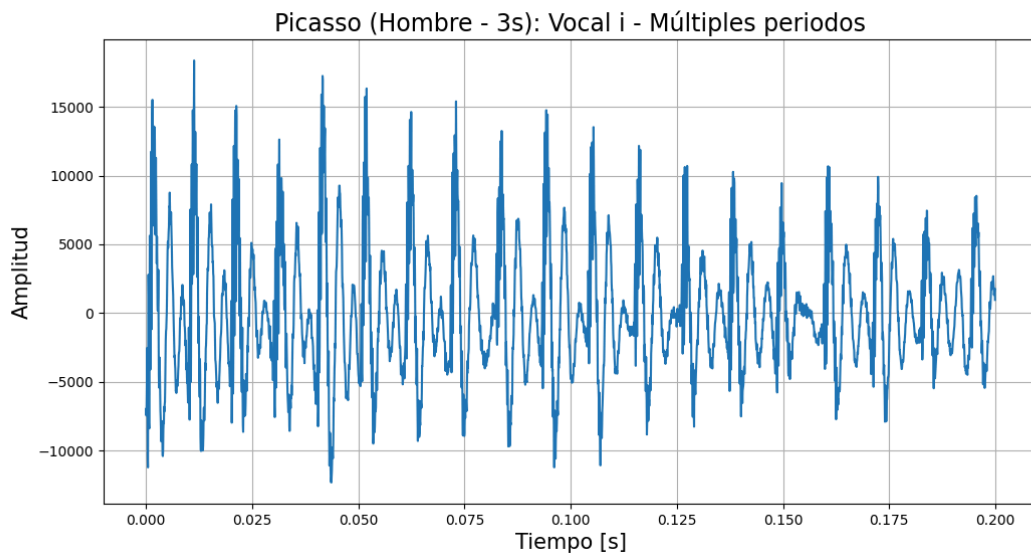




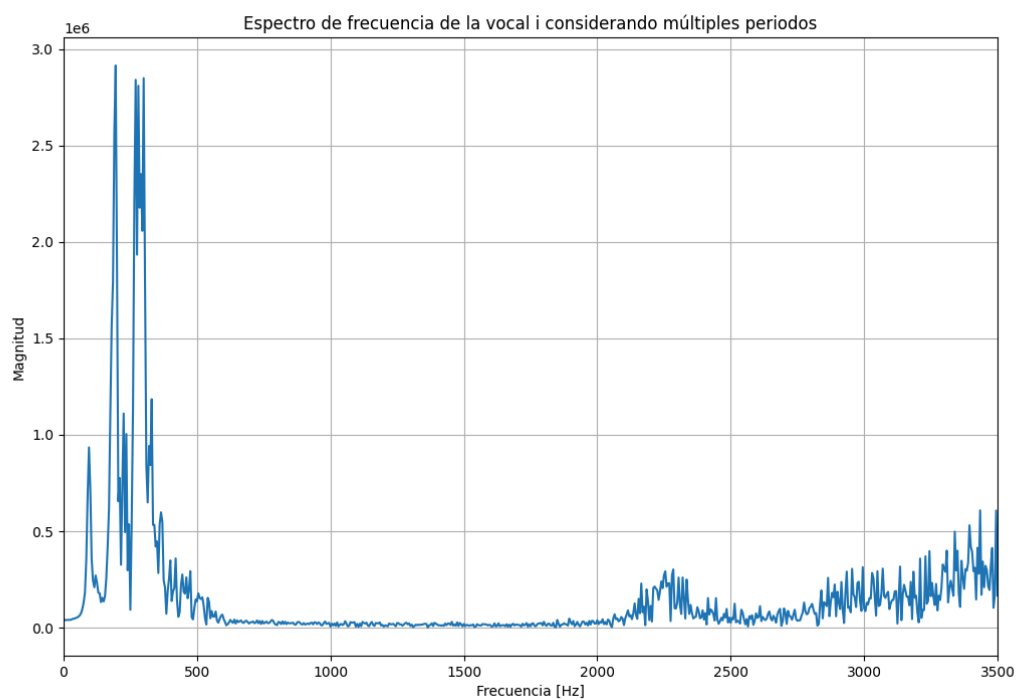
**Figura 14:** FFT de un periodo de la señal en el cual se dice la letra “i” a una velocidad lenta.

### *Señal rápida*

Se presenta el fragmento temporal considerando varios periodos en los cuales se recita la vocal “i” y se realiza de la FFT del mismo en las Figuras 15 y 16 respectivamente.

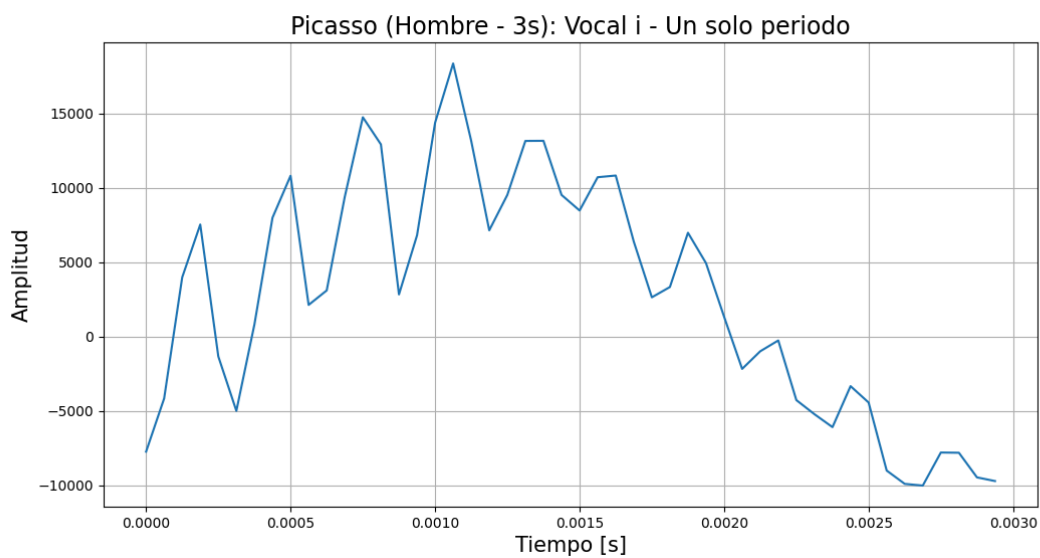


**Figura 15:** Fragmento de varios periodos de la señal diciendo la letra “i” a una velocidad rápida.



**Figura 16:** FFT de varios periodos de la señal diciendo la letra “i” a una velocidad rápida.

Se presenta el fragmento temporal considerando un solo periodo en los cuales se recita la vocal “i” y se realiza la FFT del mismo en las Figuras 17 y 18 respectivamente.



**Figura 17:** Fragmento de un periodo de la señal diciendo la letra “i” a una velocidad rápida.



**Figura 18:** FFT de un periodo de la señal diciendo la letra “i” a una velocidad rápida.

A partir de los gráficos mostrados, se pueden apreciar varias características en las FFTs de señales que incluyen múltiples periodos de la señal que contrastan drásticamente con las presentes en las FFTs en las que se consideró un solo periodo: por un lado, en las de múltiples periodos se ve que la magnitud o amplitud es significativamente mayor. Por ejemplo, en la Figura 12 el mayor pico se da en los 275Hz y tiene una magnitud de 8,61M; mientras que en la Figura 14 el mayor pico se da en los 290Hz con una magnitud de 0,26M. Esto es debido a que en la FFT de señales de múltiples periodos al haber más muestras para considerar hay mayor energía de la señal presente que en el caso de un solo periodo.

Otra diferencia significativa se encuentra en que en la FFT en la que se consideran múltiples periodos hay una variación en amplitud entre frecuencias próximas mucho más marcada que en la de un solo periodo, donde las variaciones son más paulatinas. Esto se evidencia en la Figura 16, donde a 184Hz la amplitud de la FFT es de 2,92M mientras que a 248Hz la amplitud es de 0,1M, lo que muestra una gran variabilidad en un rango corto de frecuencias; en contraste, en la Figura 18 se observa que a 327Hz la amplitud es de 0,24M mientras que a 455Hz la amplitud es de 0,16M, viéndose gráficamente un descenso suave de la amplitud de la FFT. Esta suavidad de la FFT de un solo periodo como la de la Figura 18 probablemente se deba a la menor cantidad de muestras, ya que el número de puntos considerado es significativamente menor y se ven las frecuencias de forma más abstracta, sin los detalles finos que sí se aprecian en las FFTs con mayor número de puntos como la de la Figura 16.

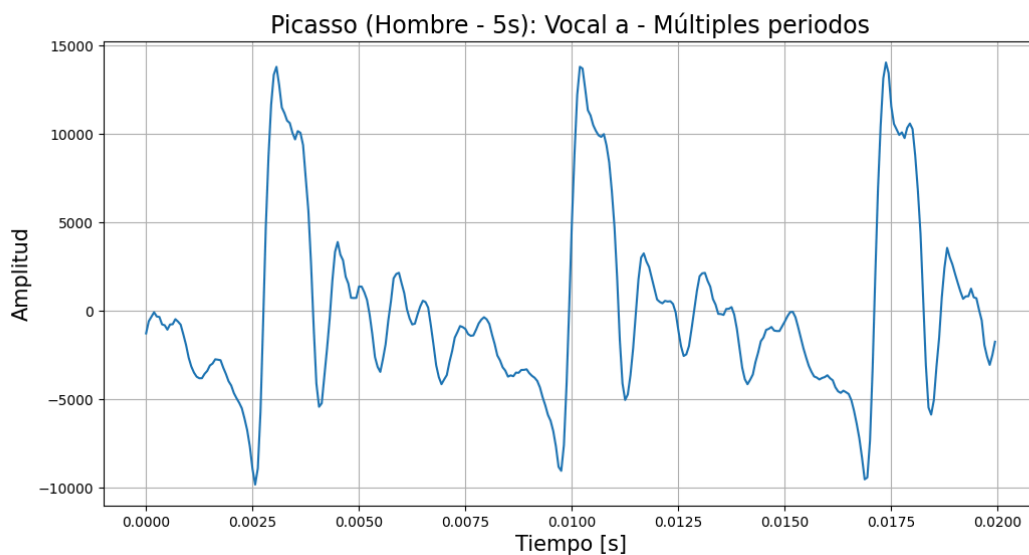
Finalmente, otro punto a resaltar es el “ruido,” ya que en las FFTs de varios periodos se observan pequeñas oscilaciones constantes en la FFT, mientras que en las FFTs de un solo periodo a frecuencias cercanas se observan amplitudes bastante similares. Esto se visualiza de manera clara en el rango de

1KHz a 2KHz en las Figuras 16 y 18; mientras que en la Figura 16 se observan pequeñas fluctuaciones en amplitud, en la Figura 18 esto se ve como una sección continua. Esto puede deberse a ruido producido por las variaciones ligeras en tono y timbre presentes en la voz al hablar, ya que naturalmente varía ligeramente incluso al pronunciar el mismo sonido: es por ello que los sonidos armónicos producidos por la voz humana son cuasi-periódicos y no perfectamente periódicos.

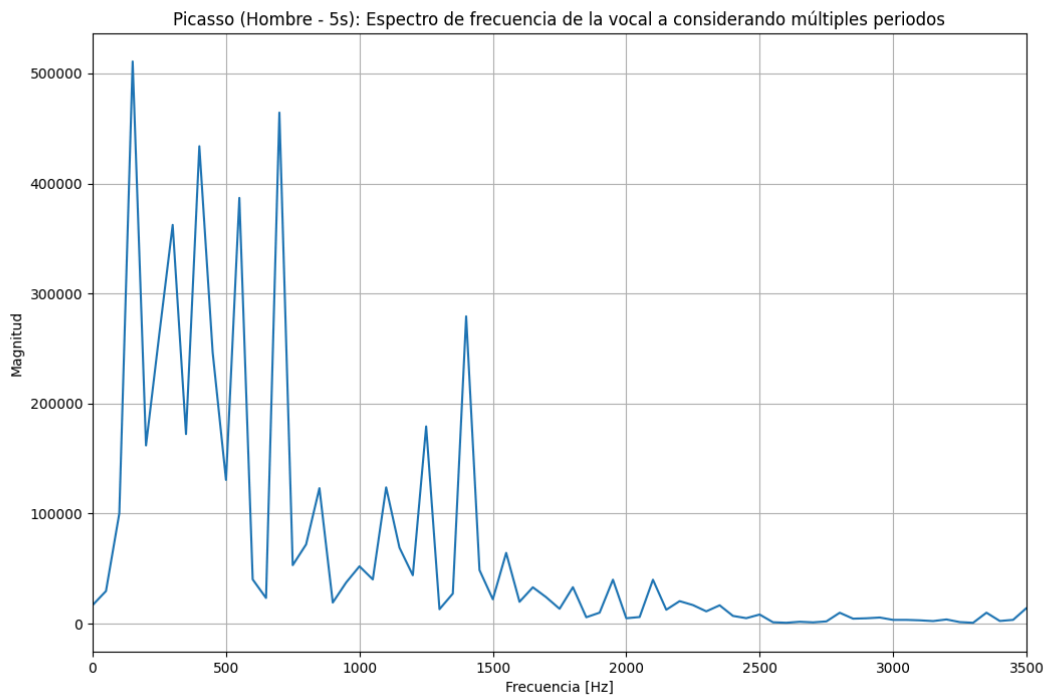
## Vocal “a”

### *Señal lenta*

Se presenta el fragmento temporal considerando sendos periodos en los cuales se recita la vocal “a” y se presenta la FFT del mismo en las Figuras 19 y 20 respectivamente.

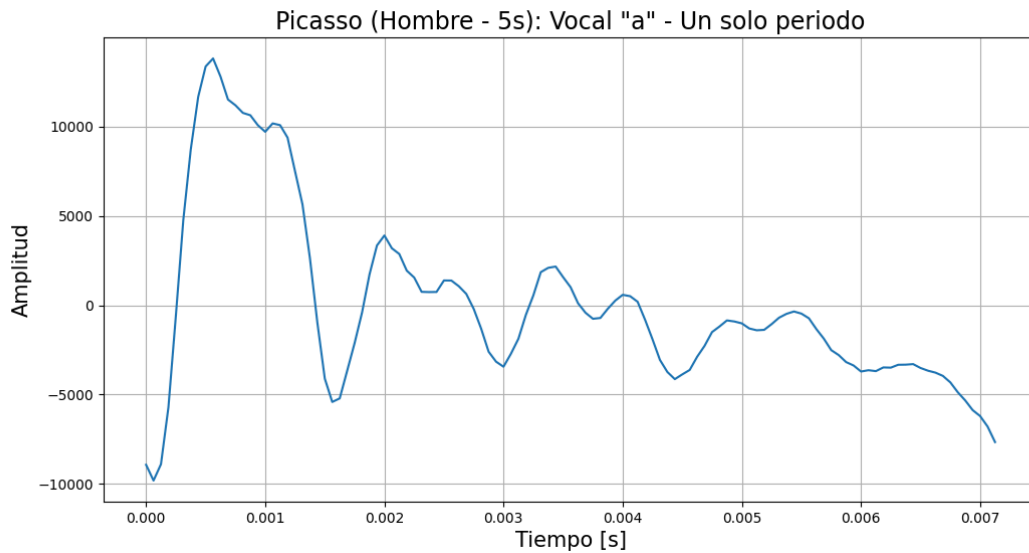


**Figura 19:** Fragmento de varios periodos de la señal en el cual se dice la letra “a” a una velocidad lenta.



**Figura 20:** FFT de varios periodos de la señal en el cual se dice la letra “a” a una velocidad lenta.

Se presenta el fragmento temporal considerando un solo periodo en el cual se recita la vocal “a” y se presenta la FFT del mismo en las Figuras 21 y 22 respectivamente.



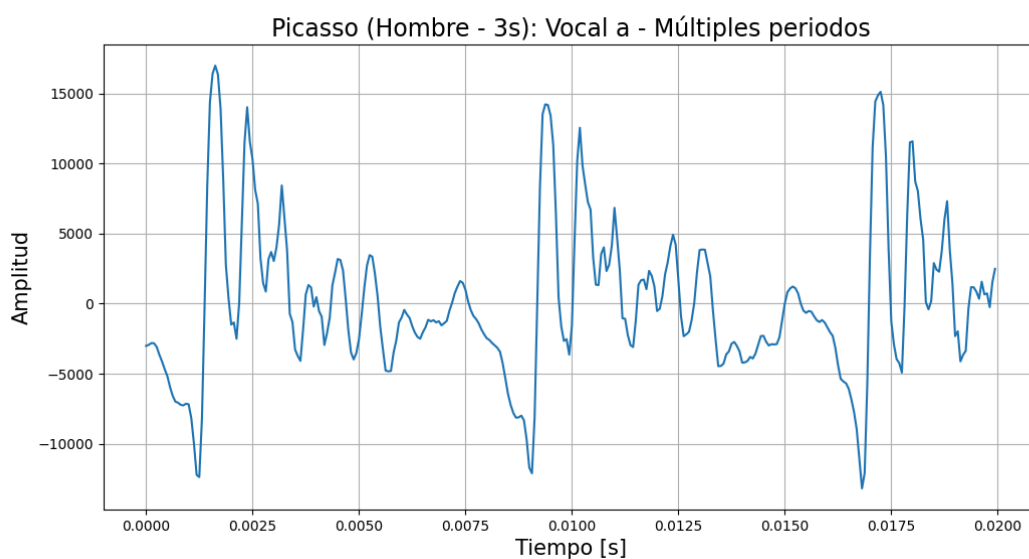
**Figura 21:** Fragmento de un periodo de la señal en el cual se dice la letra “a” a una velocidad lenta.



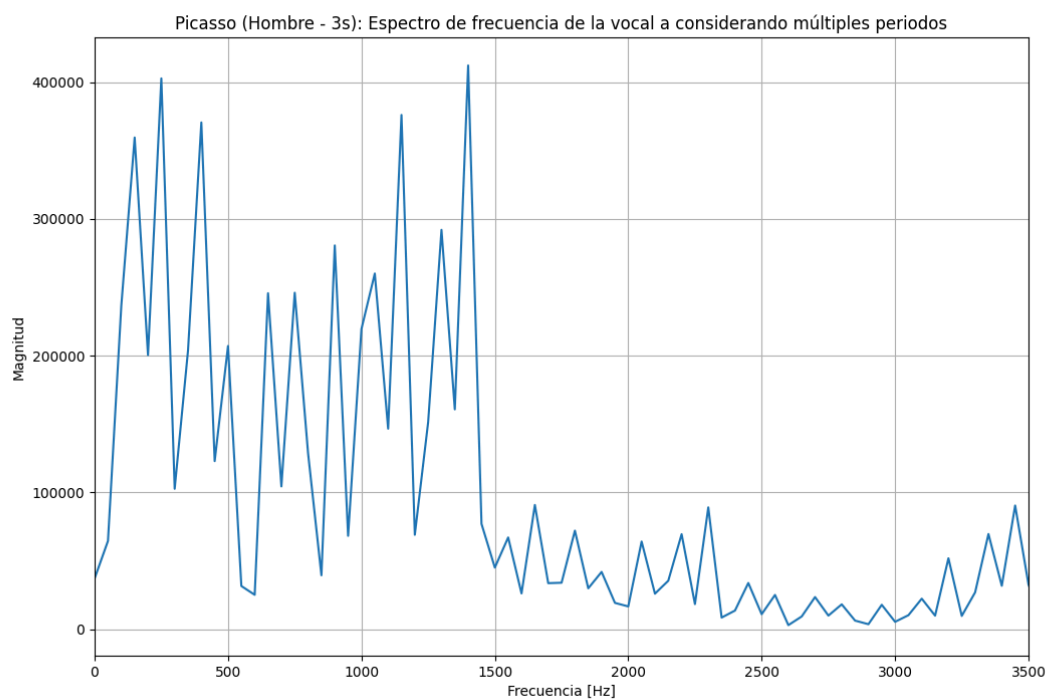
**Figura 22:** FFT de un periodo de la señal en el cual se dice la letra “a” a una velocidad lenta.

### *Señal rápida*

Se presenta el fragmento temporal considerando varios periodos en los cuales se recita la vocal “a” y se presenta la FFT del mismo en las Figuras 23 y 24 respectivamente.

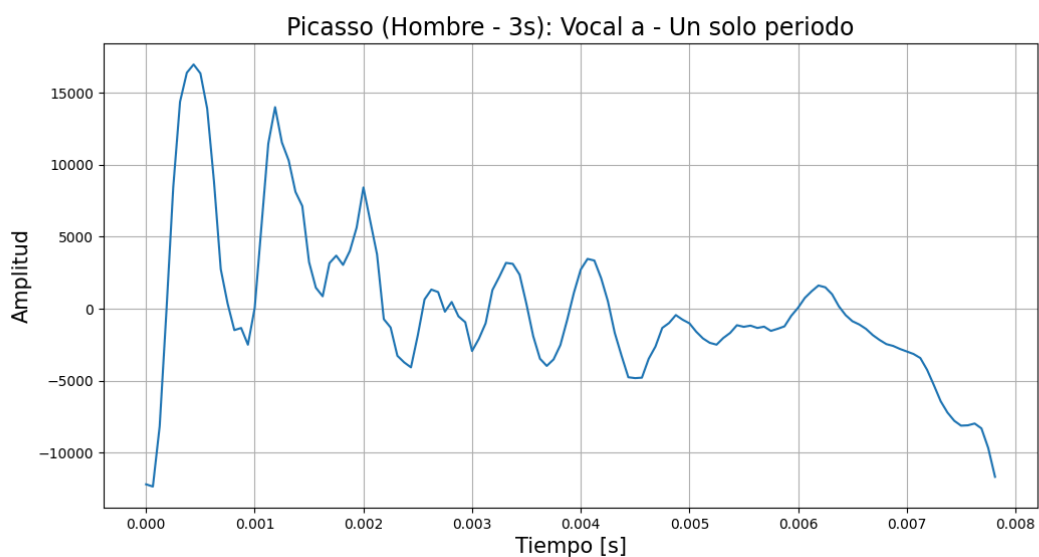


**Figura 23:** Fragmento de varios periodos de la señal en el cual se dice la letra “a” a una velocidad rápida.

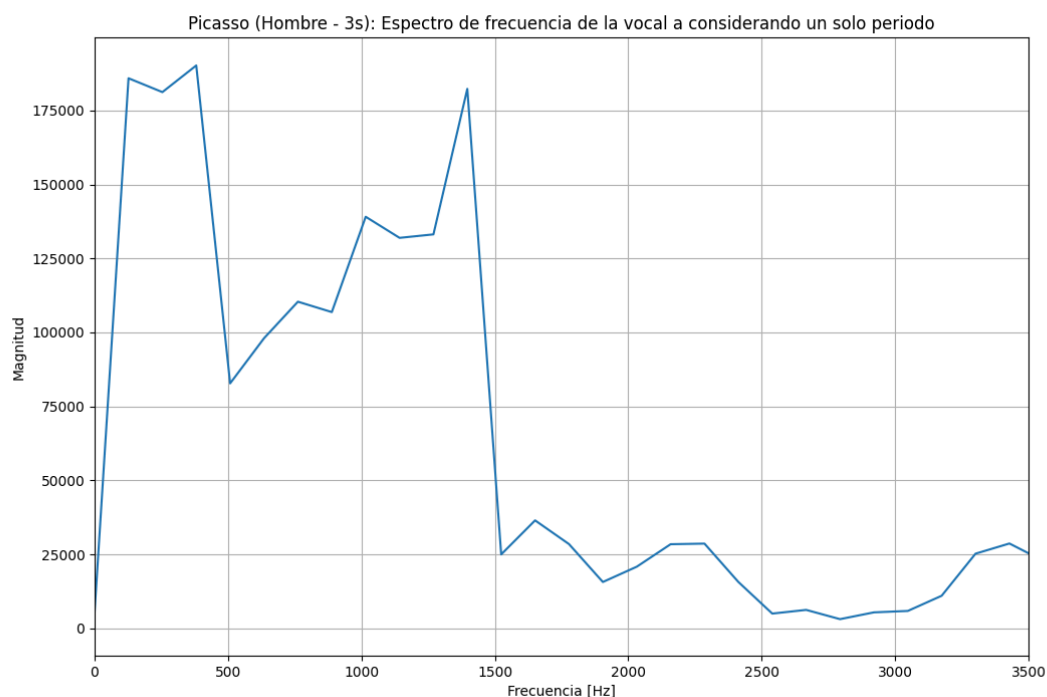


**Figura 24:** FFT de varios periodos de la señal en el cual se dice la letra “a” a una velocidad rápida.

Se presenta el fragmento temporal considerando un único periodo en el que se recita la vocal “a” y se presenta la FFT del mismo en las Figuras 25 y 26 respectivamente.



**Figura 25:** Fragmento de un periodo de la señal en el cual se dice la letra “a” a una velocidad rápida.



**Figura 26:** FFT de un periodo de la señal en el cual se dice la letra “a” a una velocidad rápida.

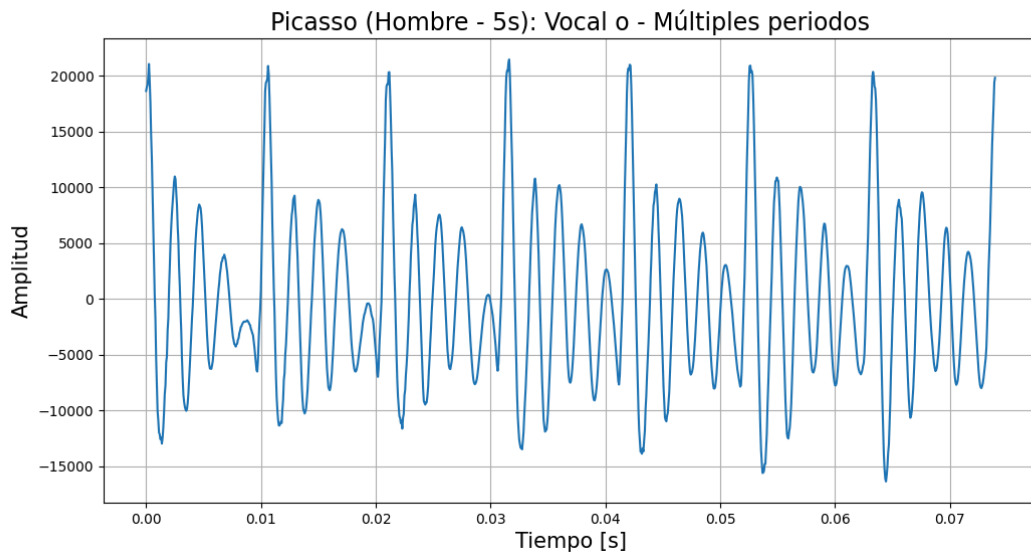
Resulta interesante comparar las frecuencias con mayor energía del fonema /a/ e /i/, ya que en la letra “a” se ven secciones con gran energía a frecuencias ligeramente más bajas que la letra “i”. Comparando en la muestra lentas con múltiples periodos, la frecuencia con mayor energía de la letra “i”, como se ve en la Figura 12, se encuentra a los 271Hz; en contraste, la frecuencia con mayor energía de la letra “a” se encuentra a los 139Hz, como se exhibe en la Figura 20.

## Vocal “o”

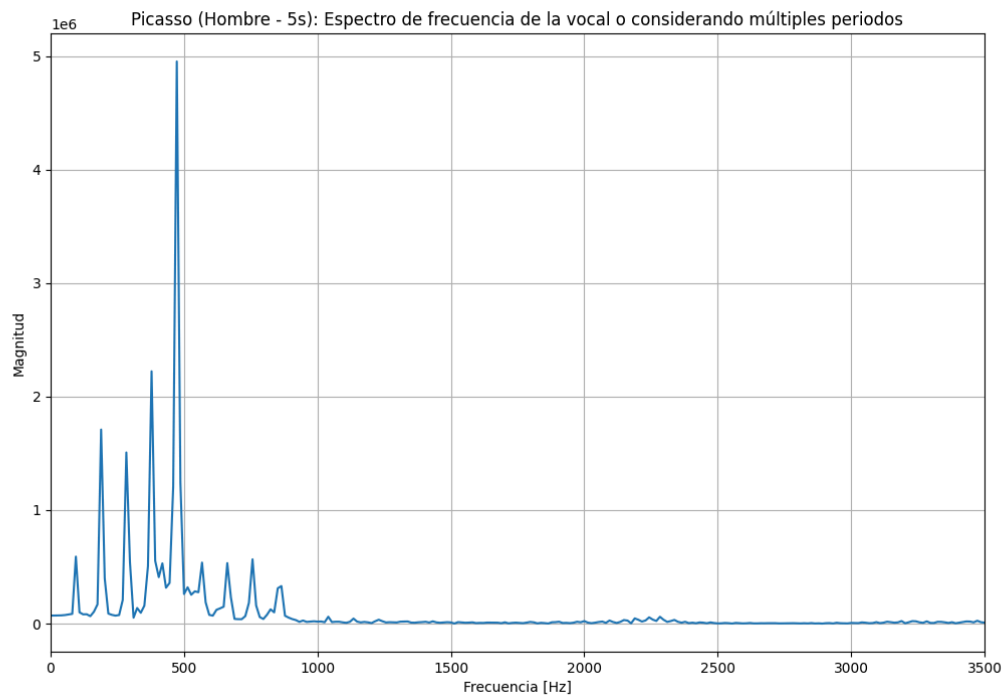
### *Señal lenta*

Se presenta el fragmento temporal considerando varios periodos en los cuales se recita la vocal “o” y se presenta la FFT del mismo en las Figuras 27 y 28 respectivamente.



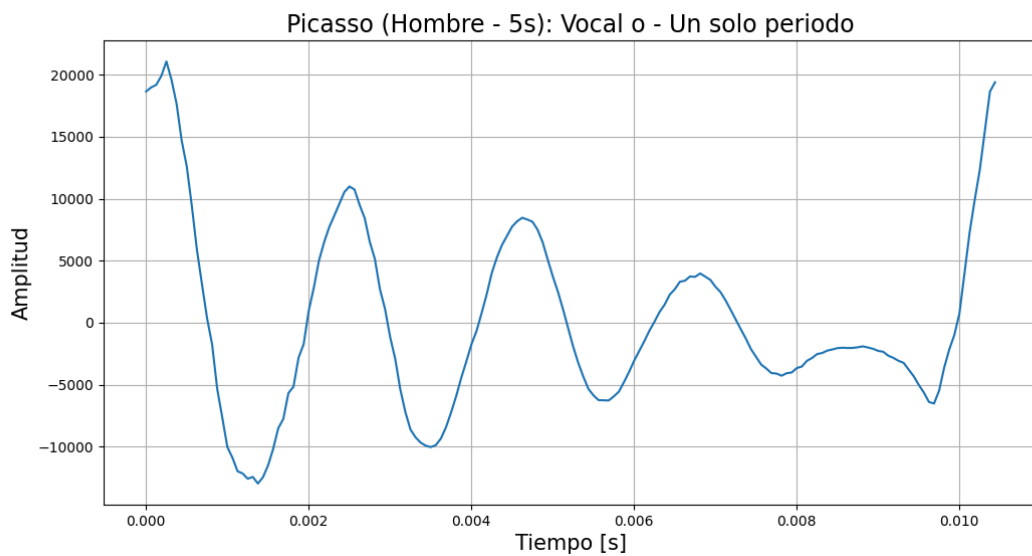


**Figura 27:** Fragmento de varios periodos de la señal en el cual se dice la letra “o” a una velocidad lenta.

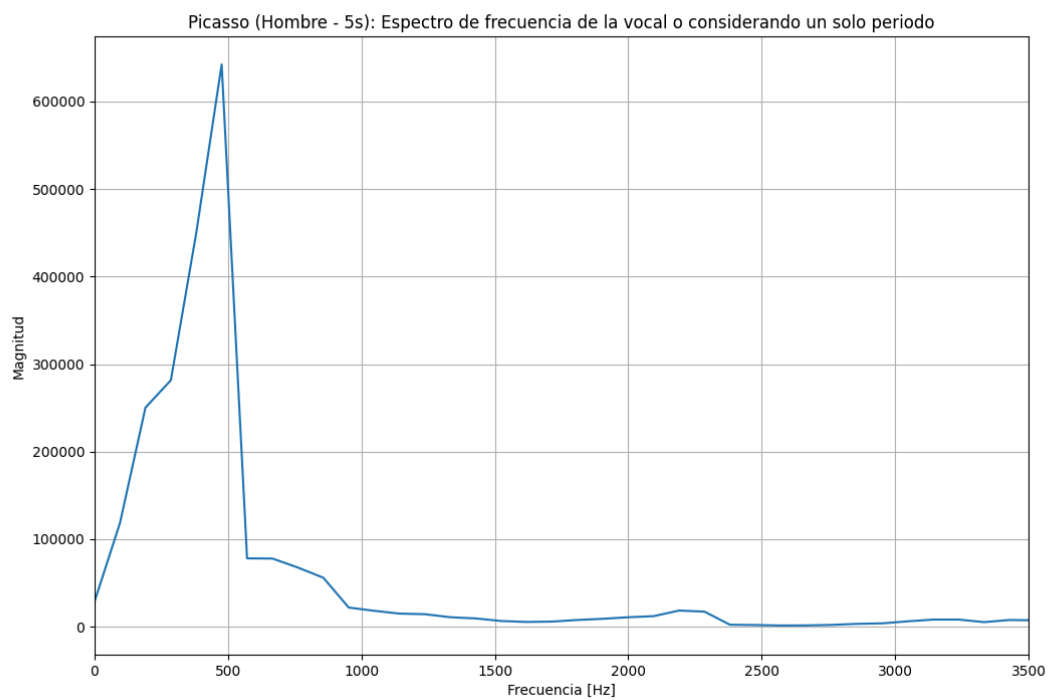


**Figura 28:** FFT de varios periodos de la señal en el cual se dice la letra “o” a una velocidad lenta.

Se presenta el fragmento temporal considerando un solo periodo en el que se recita la vocal “o” y se presenta la FFT del mismo en las Figuras 29 y 30 respectivamente.



**Figura 29:** Fragmento de un periodo de la señal en el cual se dice la letra “o” a una velocidad lenta.

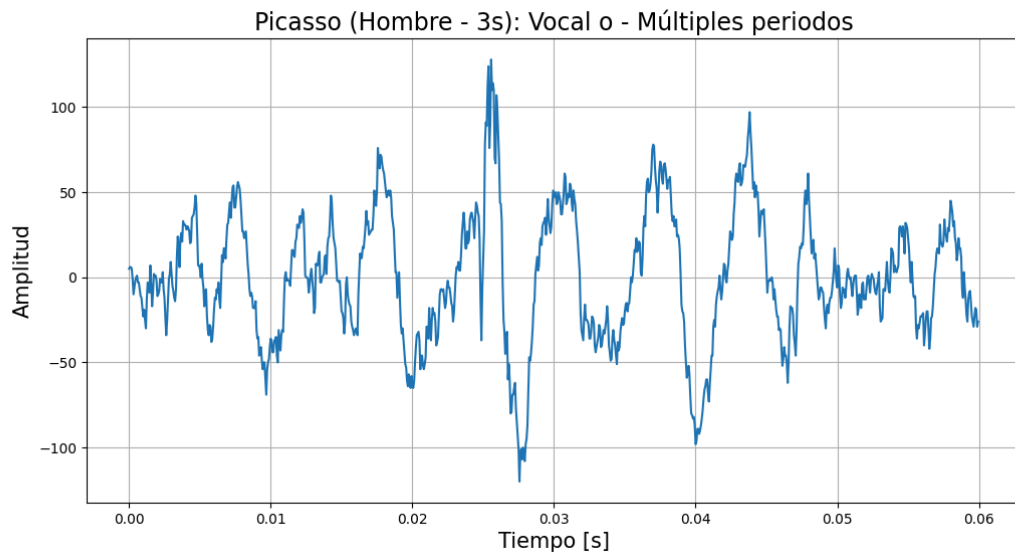


**Figura 30:** FFT de un periodo de la señal en el cual se dice la letra “o” a una velocidad lenta.

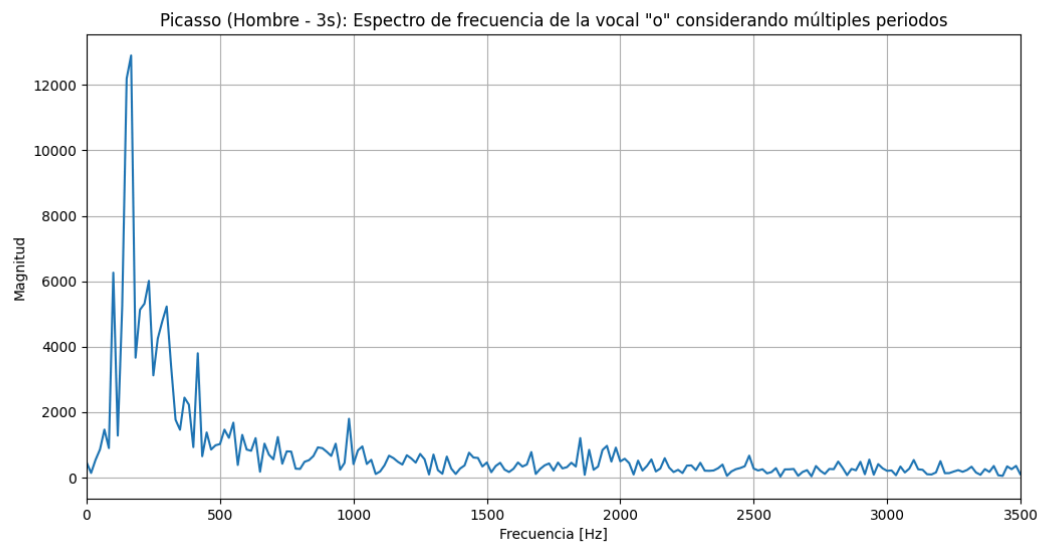
En la Figura 30 se observa que, como ocurría previamente, al ser la “o” una vocal abierta su energía máxima en la FFT la alcanza a una frecuencia menor que en el caso de la “i”, a unos 128Hz (viendo la Figura 30) en comparación con 271Hz que se observaban en la Figura 12.

### *Señal rápida*

Se presenta el fragmento temporal considerando sendos periodos en los cuales se recita la vocal “o” y se presenta la FFT del mismo en las Figuras 31 y 32 respectivamente.

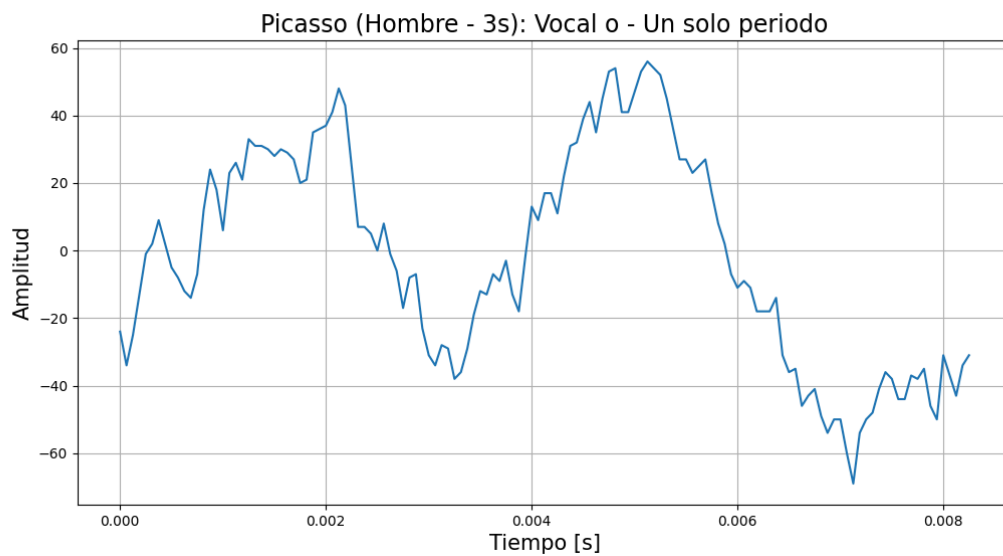


**Figura 31:** Fragmento de varios periodos de la señal en el cual se dice la letra “o” a una velocidad rápida.



**Figura 32:** FFT de varios periodos de la señal en el cual se dice la letra “o” a una velocidad rápida.

Se presenta el fragmento temporal considerando un único periodo en el cual se recita la vocal “o” y se realiza la FFT del mismo en las Figuras 33 y 34 respectivamente.



**Figura 33:** Fragmento de un periodo de la señal en el cual se dice la letra “o” a una velocidad rápida.



**Figura 34:** FFT de un periodo de la señal en el cual se dice la letra “o” a una velocidad rápida.

## Conclusiones

A lo largo de este Trabajo Práctico se observó la diferencia entre los sonidos producidos por fonemas armónicos como los de las vocales y su contraste con los fonemas de percusión como el producido por la “s.” También se vio el efecto en la amplitud de la FFT al considerarse distintas cantidades de puntos, obteniéndose energías mayores a la misma frecuencia para señales en las que se consideraba una mayor cantidad de puntos; y se vio que este aumento a su vez impactaba el nivel de detalle que se

podía observar en la FFT, ya que a menor cantidad de muestras se perdía precisión significativamente en las frecuencias presentes en la señal ingresada. En adición a esto, se observaron las diferencias entre las vocales abiertas y cerradas. A través de este Trabajo Práctico se evidenció la utilidad de realizar el pasaje de tiempo a frecuencia mediante la FFT, ya que en el dominio frecuencial se pudieron notar detalles de la señal que no hubiesen sido fáciles de ver en el dominio temporal.