

## poeta\_web

2022-12-21

```
url <- "https://www.angelfire.com/in/alga/gongora.html"
lines <- readLines(url, encoding = "UTF-8", warn = FALSE)
```

```
library(stringr) #Para que deje usar str_split()
```

```
#Hay algunos casos en los que los sonetos aparecen todos en la misma linea, divididos con etiquetas <br>
for (i in (1:length(lines))) {
  lines[i] <- str_split(lines[i], "<br>|</h4>")
}
```

```
#Deshacemos las sublistas
lines <- unlist(lines)
```

```
#Quitar los títulos
titulos <- grep("[H|h]1|[H|h]2|[H|h]3|[H|h]4|title", lines)
for (i in titulos) {
  lines[i] <- ""
}
```

```
#Eliminamos líneas en blanco
texto = c()
for (i in (1:length(lines))) {
  if (lines[i] != "") {
    texto <- c(texto, lines[i])
  }
}
```

```
#Quitar las etiquetas de html
lines <- trimws(gsub("<[^>]*>", "", lines))
```

```
#Traducir elementos &Xacute
```

```
library(stringi)
traduc <- list(c("&aacute;", "á"),
               c("&Aacute;", "Á"),
               c("&eacute;", "e"),
               c("&Eacute;", "É"),
               c("&iacute;", "í"),
               c("&Iacute;", "Í"),
               c("&oacute;", "ó"),
               c("&Oacute;", "Ó"),
               c("&uacute;", "ú"),
               c("&Uacute;", "Ú"),
               c("&ntilde;", "ñ"),
               c("&Ntilde;", "Ñ"),
```

```

        c("&iquest;", "¿"),
        c("&iexcl;", "¡"),
        c("&laquo;", "«"),
        c("&raquo;", "»"),
        c("&#191;", "¿"),
        c("&#161;", "¡"),
        c("&uuml;", "ü"),
        c("&iuml;", "ï"),
        c("&Aacute;", "Á"),
        c(" ", " ")
    )
lines <- stri_replace_all_regex(lines,
                                pattern=unlist(lapply(traduc,function(x){x[1]})),
                                replacement=unlist(lapply(traduc,function(x){x[2]})),
                                vectorize=FALSE)

#Eliminar las líneas en blanco y aquellas que acaban por "{", o "}", pues es código. Tb aquellas empezadas por "
lines[grepl("\\{\\}|\\}\\|>$", lines)] <- ""

#Nos damos cuenta que en esta web todo el texto se acumula en la última línea (esto pasa en algunos casos)
lines[grepl("^if|\\++$|\\+/$", lines)] <- ""

#Eliminamos líneas en blanco
texto = c()
for (i in (1:length(lines))) {
    if (lines[i] != "") {
        texto <- c(texto, lines[i])
    }
}
lines <- texto

#Ahora hay que eliminar las líneas terminadas en ";", pues son código. Aunque también podría darse el caso de "
lineas_punto_coma <- as.integer(grep(";", lines))
for (i in (1:length(lineas_punto_coma))) {
    if (i==1|i==length(lineas_punto_coma)) {
        if (lineas_punto_coma[i]+1==length(lines)&lines[lineas_punto_coma[i]+1]!=";") {
            lines[lineas_punto_coma[i]] <- ""
        }
    } else {
        if (lineas_punto_coma[i]-1==length(lineas_punto_coma)&lines[lineas_punto_coma[i]-1]!=";") {
            lines[lineas_punto_coma[i]] <- ""
        }
    }
}

#Eliminamos líneas en blanco
texto = c()
for (i in (1:length(lines))) {
    if (lines[i] != "") {
        texto <- c(texto, lines[i])
    }
}

```

*#Ahora que tenemos el texto bastante limpio, hay que guardar bloques de 14 líneas en un archivo .txt y texto*

```
## [1] "¡Oh, excelso muro! ¡Oh, torres coronadas"
## [2] "de honor, de majestad, de gallardía!"
## [3] "¡Oh, gran río, gran rey de Andalucía,"
## [4] "de arenas nobles, ya que no doradas!"
## [5] "¡Oh, fértil llano, oh, sierras levantadas,"
## [6] "que privilegia el cielo y dora el día!"
## [7] "¡Oh siempre gloriosa patria mía,"
## [8] "tanto por plumas como por espadas!"
## [9] "¡Si entre aquellas rinas y despojos"
## [10] "que enriquece Genil y Dauro baña"
## [11] "tu memoria no fue alimento mío,"
## [12] "nunca merezcan mis ausentes ojos"
## [13] "ver tu muro, tus torres y tu río,"
## [14] "tu llano y sierra, oh patria, oh flor de España!"
## [15] "YDEL DUQUE DE FERIA A NÁPOLES"
## [16] "Y AFRANCIA."
## [17] "El Conde mi señor se fue a Napoles;"
## [18] "el Duque mi señor se fue a Francia;"
## [19] "príncipes, buen vaje, que este día"
## [20] "pesadumbre dare a unos caracoles."
## [21] "Como sobran tan doctos españoles,"
## [22] "a ninguno ofrecí la Musa mía;"
## [23] "a un pobre albergue sí, de Andalucía,"
## [24] "que ha resistido a grandes, digo Soles."
## [25] "Con pocos libros libres (libres digo"
## [26] "de expurgaciones) paso y me paseo,"
## [27] "ya que el tiempo me pasa como higo."
## [28] "No espero en mi verdad lo que no creo:"
## [29] "espero en mi conciencia lo que digo,"
## [30] "mi salvación, que es lo que más deseo."
## [31] "De pura honestidad templo sagrado,"
## [32] "cuyo bello cimientó y gentil muro"
## [33] "de blanco nectar y alabastro duro"
## [34] "fue por divina mano fabricado;"
## [35] "pequeña puerta de coral preciado,"
## [36] "claras lumbreras de mirar seguro,"
## [37] "que a la esmeralda fina el verde puro"
## [38] "habeis para viriles usurpado;"
## [39] "soberbio techo, cuyas cimbrias de oro"
## [40] "al claro Sol, en cuanto el torno gira,"
## [41] "ornan de luz, coronan de belleza;"
## [42] "ídolo bello, a quien humilde adoro,"
## [43] "oye piadoso al que por ti suspira,"
## [44] "tus himnos canta, y tus virtudes reza."
## [45] "Mientras por competir con tu cabello,"
## [46] "oro bruñido, el sol relumbra en vano,"
## [47] "mientras con menosprecio en medio el llano"
## [48] "mira tu blanca frente el lilio bello;"
## [49] "mientras a cada labio, por cogello,"
## [50] "siguen más ojos que al clavel temprano,"
## [51] "y mientras triunfa con desden lozano"
```

## [52] "del luciente cristal tu gentil cuello;"  
 ## [53] "goza cuello, cabello, labio y frente,"  
 ## [54] "antes que lo que fue en tu edad dorada"  
 ## [55] "oro, lilio, clavel, cristal luciente,"  
 ## [56] "no sólo en plata o vola troncada"  
 ## [57] "se vuelva, más tú y ello juntamente"  
 ## [58] "en tierra, en humo, en polvo, en sombra, en nada."  
 ## [59] "Cual parece al romper de la mañana"  
 ## [60] "aljófar blanco sobre frescas rosas,"  
 ## [61] "o cual por manos hecha, artificiosas,"  
 ## [62] "bordadura de perlas sobre grana,"  
 ## [63] "tales de mi pastora soberana"  
 ## [64] "parecían las lágrimas hermosas"  
 ## [65] "sobre las dos mejillas milagrosas,"  
 ## [66] "de quien mezcladas leche y sangre mana."  
 ## [67] "Lanzando a vueltas de su tierno llanto"  
 ## [68] "un ardiente suspiro de su pecho,"  
 ## [69] "tal que el más duro canto enterneciera,"  
 ## [70] "si enternecer bastara un duro canto,"  
 ## [71] "mirad que habrá con un corazón hecho,"  
 ## [72] "que al llanto y al suspiro fue de cera."  
 ## [73] "De chinches y de mulas voy comido,"  
 ## [74] "las unas culpa de una cama vieja,"  
 ## [75] "las otras de un Señor que me las deja"  
 ## [76] "veinte días y más, y se ha partido."  
 ## [77] "De vos, madera anciana, me despido,"  
 ## [78] "miembros de algún navío de vendeja,"  
 ## [79] "patria común de la nación bermeja,"  
 ## [80] "que un mes sin deudo de mi sangre ha sido."  
 ## [81] "Venid, mulas, con cuyos pies me ha dado"  
 ## [82] "tal cox el que quizá tendrá mancilla"  
 ## [83] "de ver que me comeis el otro lado."  
 ## [84] "A Dios, Corte envainada en una villa,"  
 ## [85] "a Dios, toril de los que has sido prado,"  
 ## [86] "que en mi rincón me espera una morcilla."  
 ## [87] "-¿De dónde bueno, Juan, con pedorreras?"  
 ## [88] "-Señora tía, de Cagalarache."  
 ## [89] "-Sobrino, ¿y cuántos fuistes a Alfarache?"  
 ## [90] "-Treinta soldados en tres mil galeras."  
 ## [91] "-¿Tanta gente? -Tomámoslo de veras."  
 ## [92] "-¿Desembarcastes, Juan? -¡Tarde piache!,"  
 ## [93] "que al dar un Santiago de azabache,"  
 ## [94] "dio la playa más moros que veneras."  
 ## [95] "-Luego, ¿es de moros? -Sí, señora tía;"  
 ## [96] "mucha algazara, pero poca ropa."  
 ## [97] "-¿Hicieron os los perros algún daño?"  
 ## [98] "-No, que en ladrando con su artillería,"  
 ## [99] "a todos nos dio cámaras de popa."  
 ## [100] "-¡Salud serían para todo el año!"  
 ## [101] "Ayer naciste, y morirás mañana."  
 ## [102] "Para tan breve ser, ¿quien te dio vida?"  
 ## [103] "¿Para vivir tan poco estás lucida?"  
 ## [104] "Y, ¿para no ser nada estás lozana?"  
 ## [105] "Si te engañó su hermosura vana,"

```
## [106] "bien presto la verás desvanecida,"
## [107] "porque en tu hermosura está escondida"
## [108] "la ocasión de morir muerte temprana."
## [109] "Cuando te corte la robusta mano,"
## [110] "ley de la agricultura permitida,"
## [111] "grosero aliento acabará tu suerte."
## [112] "No salgas, que te aguarda algún tirano;"
## [113] "dilata tu nacer para la vida,"
## [114] "que anticipas tu ser para tu muerte."
## [115] "Ya besando unas manos cristalinas,"
## [116] "ya anudándose a un blanco y liso cuello,"
## [117] "ya esparciendo por el aquel cabello"
## [118] "que Amor sacó entre el oro de sus minas,"
## [119] "ya quebrando en aquellas perlas finas"
## [120] "palabras dulces mil sin merecello,"
## [121] "ya cogiendo de cada labio bello"
## [122] "purpúreas rosas sin temor de espinas,"
## [123] "estaba, oh, claro sol invidioso,"
## [124] "cuando tu luz, hiriendome los ojos,"
## [125] "mató mi gloria y acabó mi suerte."
## [126] "Si el cielo ya no es menos poderoso,"
## [127] "porque no den los suyos más enojos,"
## [128] "rayos, como a tu hijo, te den muerte."
## [129] "La Palestra de Euterpe."
```

```
con2 <- file( description = "soneto_web.txt", open = "wt")
writeLines(texto, con2, sep="\n")
close(con2)
```