

Welcome!

STRING MANIPULATION WITH STRINGR IN R



Charlotte Wickham

Assistant Professor at Oregon State
University

You will learn:

- stringr for manipulating strings
- Regular expressions using rebus

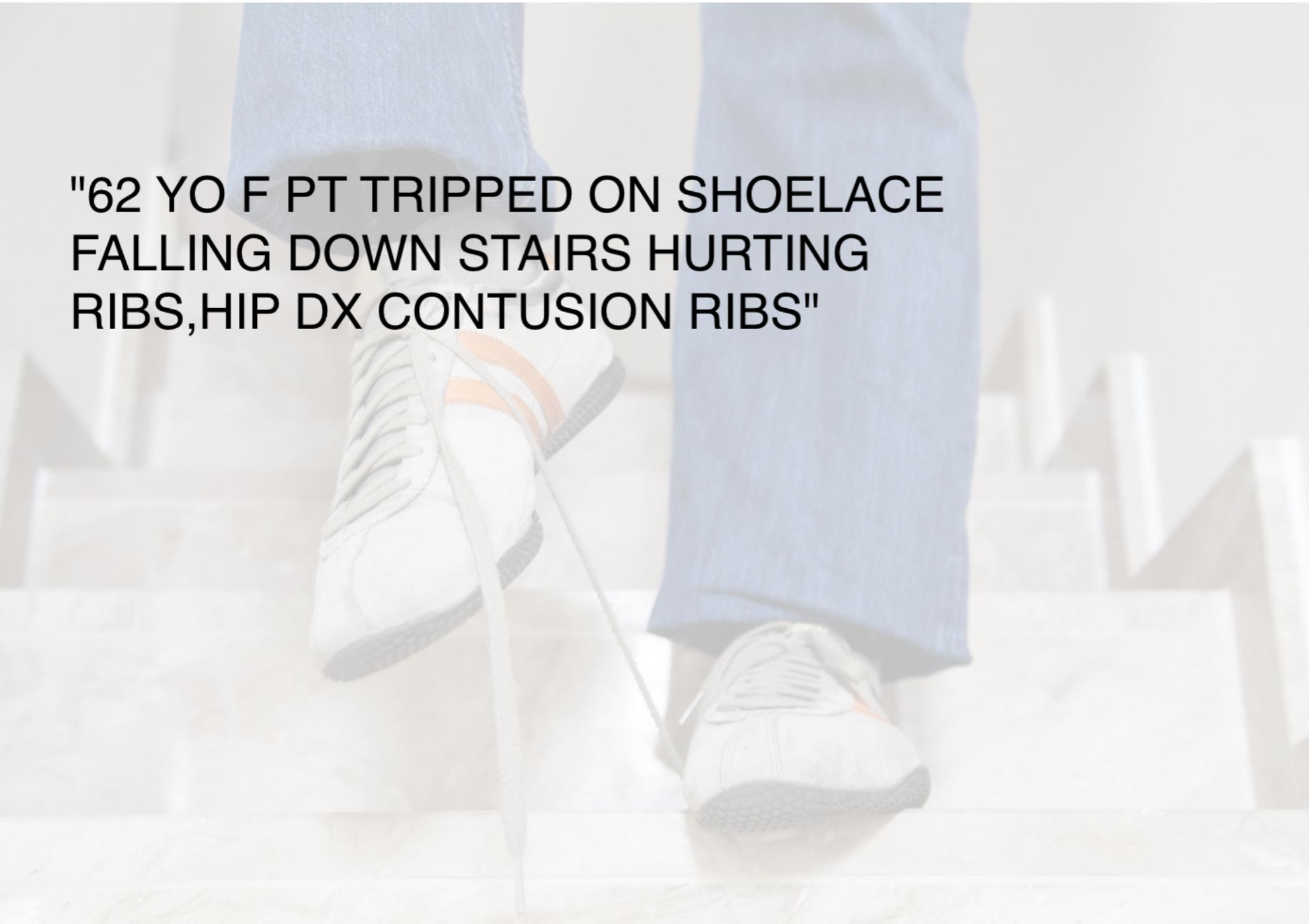


```
"TTAAGGAAACGATCGTACGCATGATAGGGTTTGCA  
GTGATATTAGTGTCTCGGTTGACTGGATCTCATCAA  
TAGTCTGGATTGGTTGATAAGTACACACTTAATAAATATGCT  
CATCAATGGATTACACATCACTTAATAAATATGCT  
GTAGTGGCCAGTGGTGAATAGGCCTCAACCACCT  
CTTCTAACGCTTCCAATTTCAGGCGGAAGGG  
TAATCTTGGCACTTTCAAGATTATGCCAATAAAG  
CAGCAAACGTCGTAACCCAGTTGTTGGGTTAAC  
GTGTACACAAGCTGCGGTAATGATCCCCTGCTGCC  
GCATCTTTCTACTCTTACATGAATAGTTCCGGGGC  
TAACAGCGAGGTTGGCTAATTAGCATAGGGT  
GTGCGTGCATTTCATTAAATGCTTCAGGATGCTG  
CGATCGAGATTATCGATCTGATAAAATTCACTCAT"
```

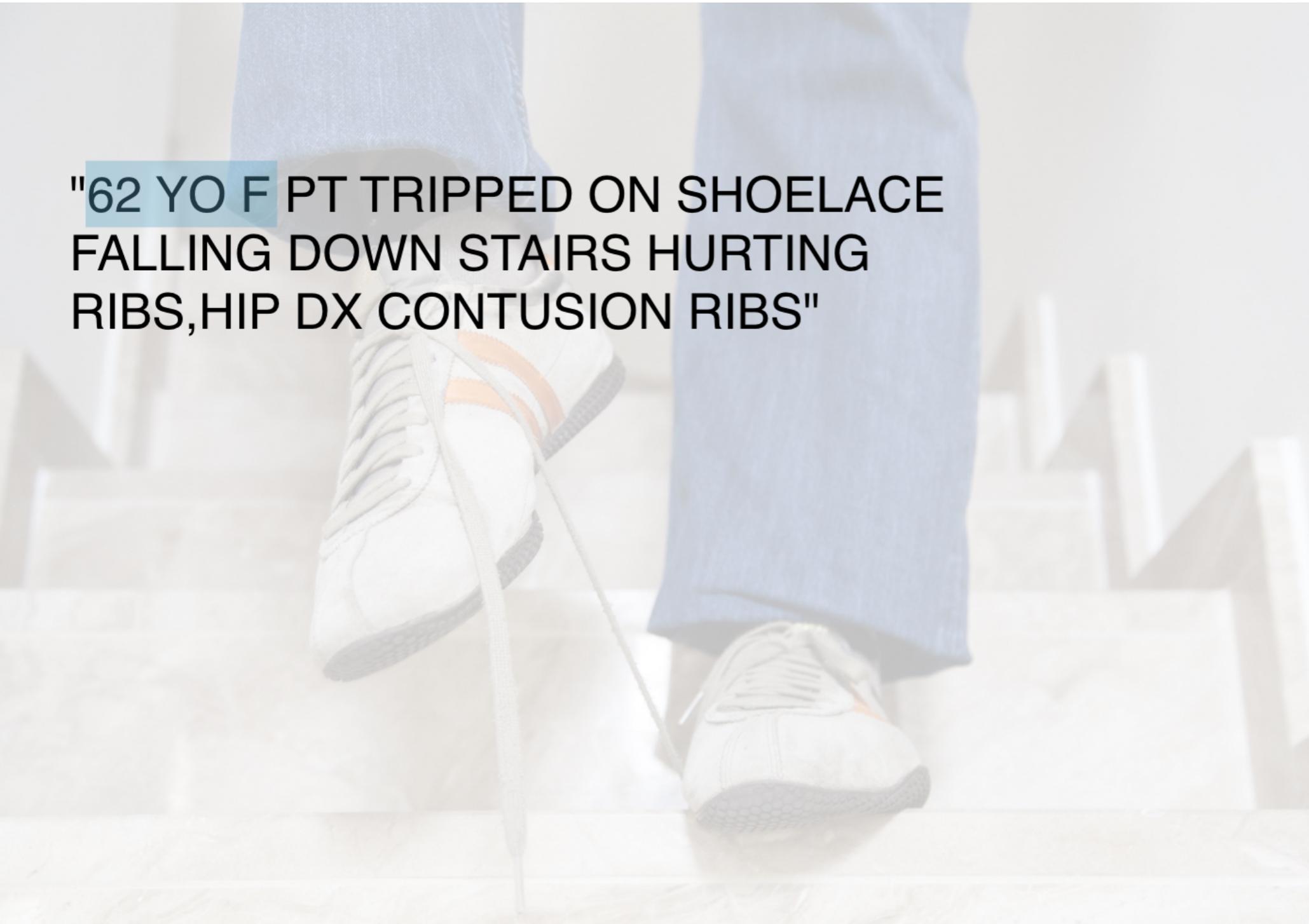
```
"TTAAGGAACGATCGTACGCATGATAGGGTTTGCA  
GTGATATTAGTGTCTCGGTTGACTGGATCTCATCAA  
TAGTCTGGATTGTTGATAAGTACCTGCTGCAATG  
CATCAATGGATTACACATCACTTAATAAAATATGCT  
GTAGTGGCCAGTGGTGTAAATAGGCCTCAACCACCT  
CTTCTAACGCTTCCAATTTCAGGCGGAAGGG  
TAATCTTGGCACTTTCAAGATTATGCCAATAAAG  
CAGCAAACGTCGTAACCCAGTTGTTGGGTTAAC  
GTGTACACAAGCTGCGGTAATGATCCCCTGCTGCC  
GCATCTTTCTACTCTTACATGAATAGTTCCGGGGC  
TAACAGCGAGGTTTGGCTAATTCAAGCATAGGGT  
GTGCGTGCATTTCATTAAATGCTTCAGGATGCTG  
CGATCGAGATTATCGATCTGATAAAATTCACTCAT"
```

```
"TTAAGGAAACGATCGTACGCATGATAGGGTTTGCA  
GTGATATTAGTGTCTCGGTTGACTGGATCTCATCAA  
TAGTCTGGATTGGTTGATAAGTACACACTTAATAAATATGCT  
CATCAATGGATTACACATCACTTAATAAATATGCT  
GTAGTGGCCAGTGGTGAATAGGCCTCAACCACCT  
CTTCTAACGCTTCCAATTTCAGGCGGAAGGG  
TAATCTTGGCACTTTCAAGATTATGCCAATAAAG  
CAGCAAACGTCGTAACCCAGTTGTTGGGTTAAC  
GTGTACACAAGCTGCGGTAATGATCCCCTGCTGCC  
GCATCTTTCTACTCTTACATGAATAGTTCCGGGGC  
TAACAGCGAGGTTGGCTAATTAGCATAGGGT  
GTGCGTGCATTTCATTAAATGCTTCAGGATGCTG  
CGATCGAGATTATCGATCTGATAAAATTCACTCAT"
```

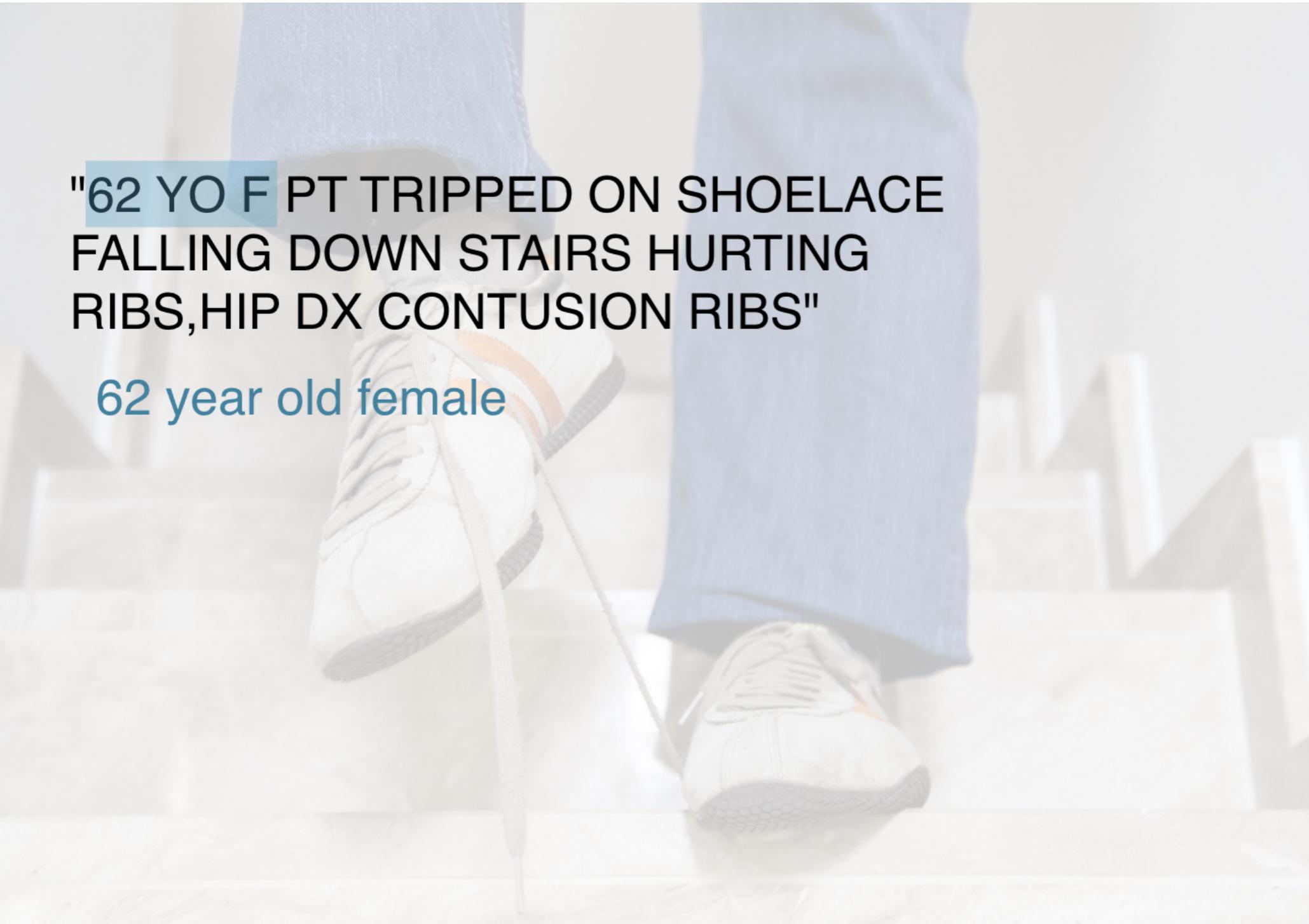
```
"TTAAGGAAACGATCGTACGCATGATAGGGTTTGCA  
GTGATATTAGTGTCTCGGTTGACTGGATCTCATCAA  
TAGTCTGGATTGGTTGATAAGTACACACTTAATAAATATGCT  
CATCAATGGATTACACATCACTTAATAAATATGCT  
GTAGTGGCCAGTGGTGTAAATAGGCCTCAACCACCT  
CTTCTAACGCTTCCAATTAAAAAGGGCGGAAGGG  
TAATCTTGGCACTTTCAAGATTATGCCAATAAAG  
CAGCAAACGTCGTAACCCAGTTGTTGGGTTAAC  
GTGTACACAAGCTGCGGTAATGATCCCCTGCTGCC  
GCATCTTTCTACTCTTACATGAATAGTTCCGGGGC  
TAACAGCGAGGTTGGCTAATTAGCATAGGGT  
GTGCGTGCATTTCATTAAATGCTTCAGGATGCTG  
CGATCGAGATTATCGATCTGATAAAATTCACTCAT"
```



"62 YO F PT TRIPPED ON SHOELACE
FALLING DOWN STAIRS HURTING
RIBS,HIP DX CONTUSION RIBS"



"62 YO F PT TRIPPED ON SHOELACE
FALLING DOWN STAIRS HURTING
RIBS,HIP DX CONTUSION RIBS"



"62 YO F PT TRIPPED ON SHOELACE
FALLING DOWN STAIRS HURTING
RIBS,HIP DX CONTUSION RIBS"

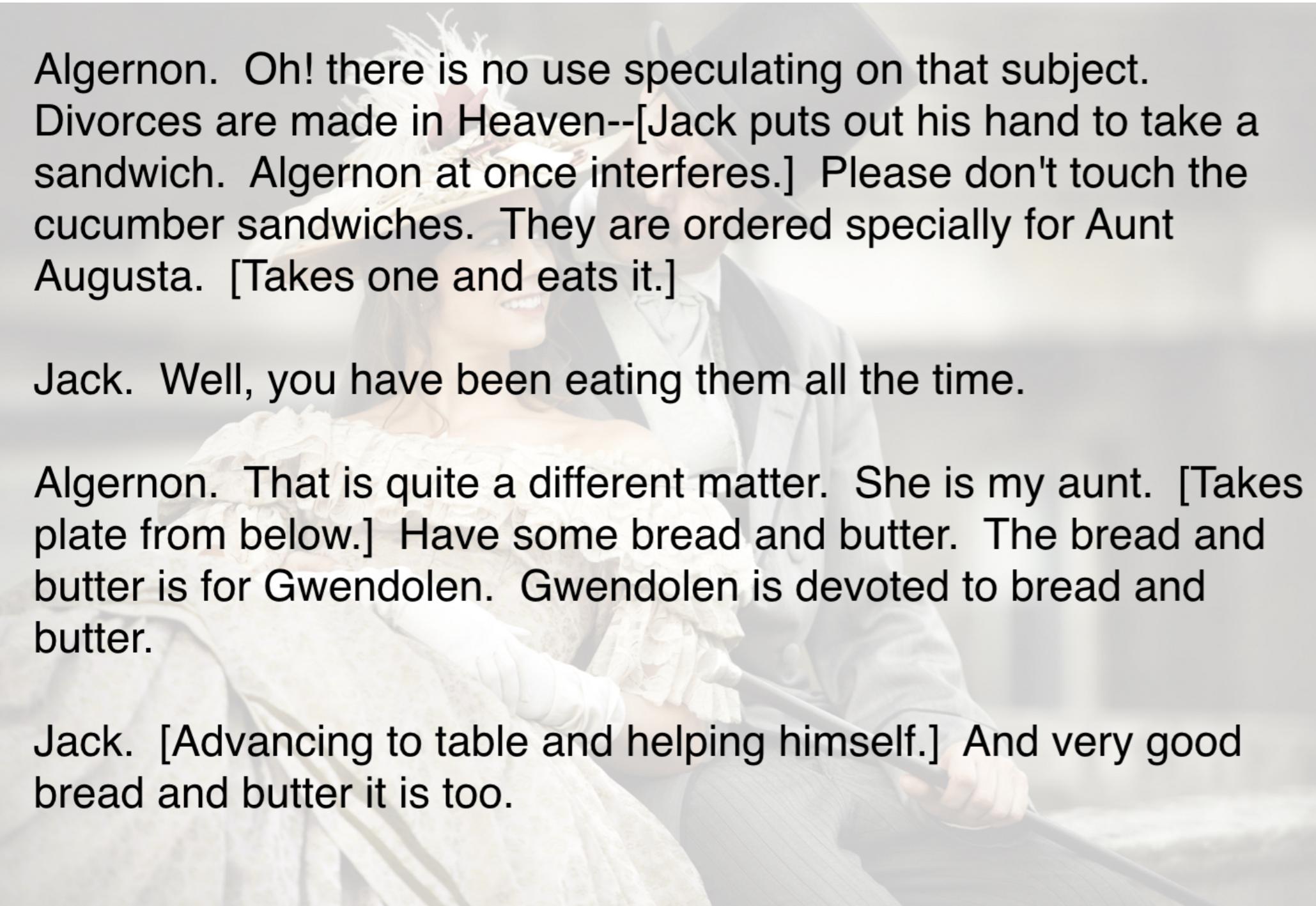
62 year old female

```
[1] "Call me at 555-555-0191"  
[2] "123 Main St"  
[3] "(555) 555 0191"  
[4] "Phone: 555.555.0191 Mobile: 555.555.0192"
```

```
[1] "Call me at 555-555-0191"  
[2] "123 Main St"  
[3] "(555) 555 0191"  
[4] "Phone: 555.555.0191 Mobile: 555.555.0192"
```



```
[1] "Call me at XXX-XXX-XXXX"  
[2] "XXX Main St"  
[3] "(XXX) XXX XXXX"  
[4] "Phone: XXX.XXX.XXXX Mobile: XXX.XXX.XXXX"
```

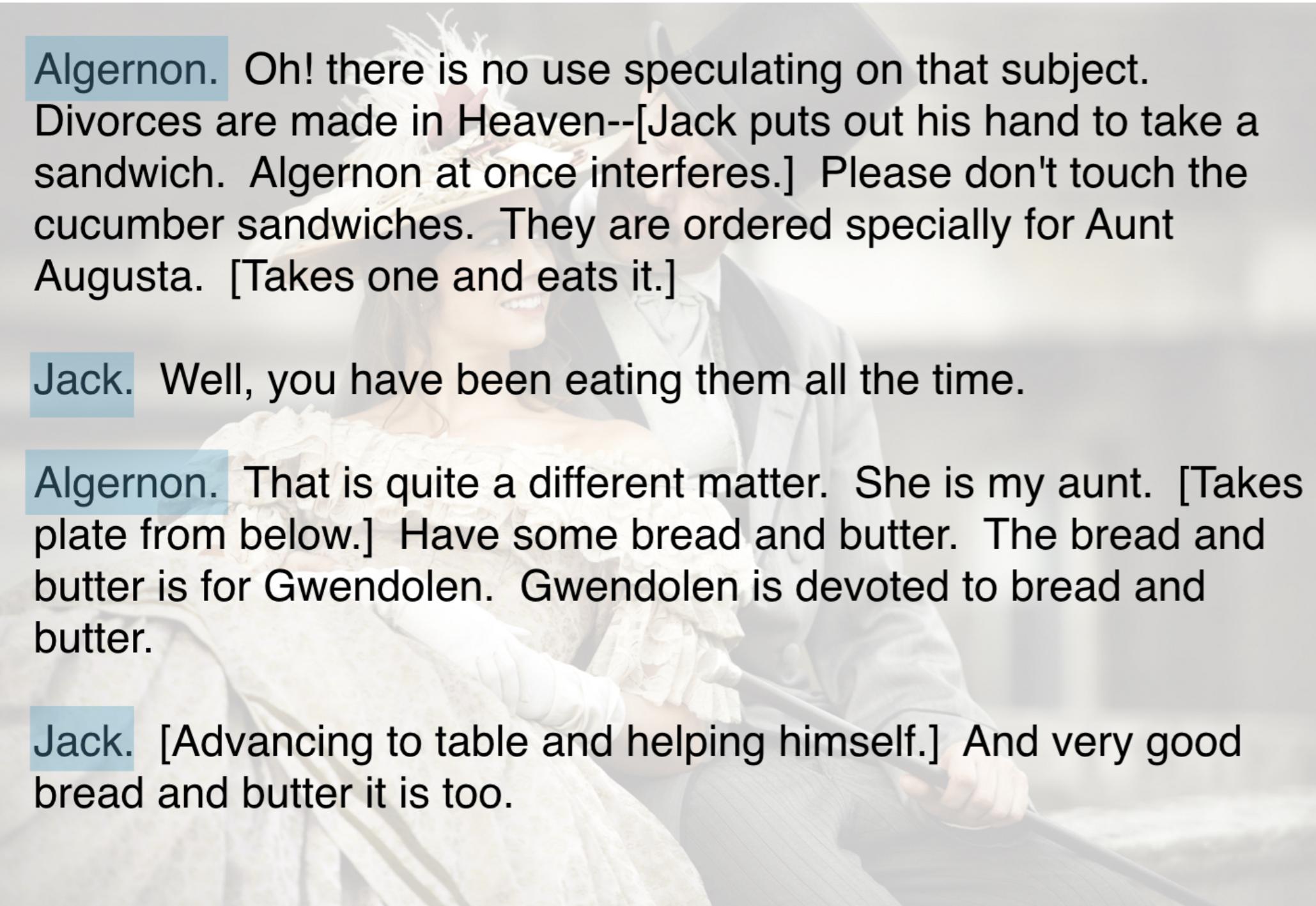


Algernon. Oh! there is no use speculating on that subject. Divorces are made in Heaven--[Jack puts out his hand to take a sandwich. Algernon at once interferes.] Please don't touch the cucumber sandwiches. They are ordered specially for Aunt Augusta. [Takes one and eats it.]

Jack. Well, you have been eating them all the time.

Algernon. That is quite a different matter. She is my aunt. [Takes plate from below.] Have some bread and butter. The bread and butter is for Gwendolen. Gwendolen is devoted to bread and butter.

Jack. [Advancing to table and helping himself.] And very good bread and butter it is too.



Algernon. Oh! there is no use speculating on that subject. Divorces are made in Heaven--[Jack puts out his hand to take a sandwich. Algernon at once interferes.] Please don't touch the cucumber sandwiches. They are ordered specially for Aunt Augusta. [Takes one and eats it.]

Jack. Well, you have been eating them all the time.

Algernon. That is quite a different matter. She is my aunt. [Takes plate from below.] Have some bread and butter. The bread and butter is for Gwendolen. Gwendolen is devoted to bread and butter.

Jack. [Advancing to table and helping himself.] And very good bread and butter it is too.

Chapter 1

- How to enter strings
- How to control numbers as strings
- Combine strings into sentences and tables

Entering strings

```
"hi!"
```

```
"hi!"
```

```
# I said "hi!"  
"I said "hi!""
```

```
Error: unexpected symbol in ""I say "hi"
```

Entering strings

```
'I said "hi!"'      # A single quote
```

```
"I said \"hi!\""
```

Entering strings

```
'I said "hi!"'      # A single quote
```

```
"I said \"hi!\""  
# An escape sequence
```

```
"I said \"hi!\""  
# # An escaped double quote
```

```
"I said \"hi!\""
```

When to use \" vs. '

```
"hi!"      # No quotes in the string, use double quotes
```

```
"hi!"
```

```
'I said "hi!"'      # Doubles quotes in the string,  
# use single quotes
```

```
"I said \"hi!\""
```

```
"I'd say \"hi!\""      # Doubles and single quotes in the string,  
# use double quotes
```

```
"I'd say \"hi!\""
```

Let's practice!

STRING MANIPULATION WITH STRINGR IN R

Turning numbers into strings

STRING MANIPULATION WITH STRINGR IN R



Charlotte Wickham

Assistant Professor at Oregon State
University

Turning a number into a string

```
estimate <- 1.34019029100
```

\$1.34

```
as.character(estimate)
```

"1.340190291"

```
format(estimate, digits = 3)
```

"1.34"

```
formatC(estimate,  
        format = "f", digits = 2)
```

"1.34"

Fixed and scientific formats

- Fixed: decimal point between ones and tenths
- Scientific: decimal point after first digit

6371 km

Fixed and scientific formats

- Fixed: decimal point between ones and tenths
- Scientific: decimal point after first digit

6371 km

6371

Fixed and scientific formats

- Fixed: decimal point between ones and tenths
- Scientific: decimal point after first digit

6371 km

6371

6.371×10³

Fixed and scientific formats

- Fixed: decimal point between ones and tenths
- Scientific: decimal point after first digit

6371 km

6371.0

6.371×10^3

Fixed and scientific formats

19890000000000000000000000000000 kg

Fixed and scientific formats

19890000000000000000000000000000 kg → 1.989×10^{30}

Fixed and scientific formats

198900000000000000000000000000 kg → 1.989×10^{30}

0.00000000008 kg → 8×10^{-12} kg

198900000000000000000000000000

1.989e+30

0.0000000008

8e-12

format() and formatC()

```
x <- c(1989000000000000000000000000, 0.0000000008)
```

```
format(x, scientific = TRUE)
```

```
"1.989e+30" "8.000e-12"
```

```
format(x, scientific = FALSE)
```

```
"198899999999999901909255192576.000000000000"  
"0.0000000008"
```

format() and formatC()

```
formatC(x, format = "f")
```

```
"198899999999999901909255192576.0000"  
"0.0000"
```

```
formatC(x, format = "e")
```

```
"1.9890e+30" "8.0000e-12"
```

```
formatC(x, format = "g")
```

```
"1.989e+30" "8e-12"
```

Let's practice!

STRING MANIPULATION WITH STRINGR IN R

Putting strings together

STRING MANIPULATION WITH STRINGR IN R



Charlotte Wickham

Assistant Professor at Oregon State
University

paste()

```
paste("E", "I", "E", "I", "O")
```

```
"E I E I O"
```

```
paste("E", "I", "E", "I", "O", sep = "-")
```

```
"E-I-E-I-O"
```

```
paste(c("Here", "There", "Everywhere"), "a")
```

```
"Here a"      "There a"      "Everywhere a"
```

```
paste( c("Here", "a")  
      "There",           "a"  
      "Everywhere")     "a"  → "Here a"  
                           "a"  → "There a"  
                           "a"  → "Everywhere a"
```



paste()

```
animal_goes <- "moo"  
paste(c("Here", "There", "Everywhere"), "a", animal_goes)
```

```
"Here a moo"      "There a moo"      "Everywhere a moo"
```

```
paste(c("Here", "There", "Everywhere"), "a", animal_goes,  
      collapse = ", ")
```

```
"Here a moo, There a moo, Everywhere a moo"
```

```
paste(c("Here", "There", "Everywhere"), "a",  
      c(animal_goes, animal_goes,  
          paste(rep(animal_goes, 2), collapse = "-")),  
      collapse = ", ")
```

```
"Here a moo, There a moo, Everywhere a moo-moo"
```

paste()

```
old_mac <- function(animal, animal_goes){  
  eieio <- paste("E", "I", "E", "I", "O", sep = "-")  
  old_mac <- "Old MacDonald had a farm"  
  writeLines(c(  
    old_mac,  
    eieio,  
    paste("And on his farm he had a", animal),  
    eieio,  
    paste(c("Here", "There", "Everywhere"), "a",  
      c(animal_goes, animal_goes,  
        paste(rep(animal_goes, 2), collapse = "-")),  
        collapse = ", "),  
    old_mac,  
    eieio))  
}
```

```
old_mac("cow", "moo")
```

```
Old MacDonald had a farm  
E-I-E-I-O  
And on his farm he had a cow  
E-I-E-I-O  
Here a moo, There a moo, Everywhere a moo-moo  
Old MacDonald had a farm  
E-I-E-I-O
```

```
old_mac("dog", "woof")
```

```
Old MacDonald had a farm  
E-I-E-I-O  
And on his farm he had a dog  
E-I-E-I-O  
Here a woof, There a woof, Everywhere a woof-woof  
Old MacDonald had a farm  
E-I-E-I-O
```

Let's practice!

STRING MANIPULATION WITH STRINGR IN R