# Regular expressions

("regexps," "regexp," "regex")

## Links to helpful pages

http://www.emacswiki.org/emacs/RegularExpression
https://www.gnu.org/software/emacs/manual/html_node/elisp/Replacing-Match.html

---

## Recipes

### Replace single spaces after end of sentence with two spaces in Emacs

<2015.6.24>

Type the following:

```
[M-x] query-replace-regexp [RET]
\. \([A-Z]\) [RET]
.  \1 [RET]
```

---

### Emacs replacement trick ADC used when writing SPM table report tool

(from here)

In output file ("taldaemon_i_plus.td.txt" e.g.), AC used emacs regexp (regular expression) to reformat BA's, and just copied a rectangle to get region descriptions:

```
[M-x] replace-regexp [RET]
^.*Brodmann area.\([0-9]+\).* [RET]
BA\1 [RET]
```

## Emacs regexp for removing newline characters from text copied from Adobe PDF file

[Originally from Anthony's `science.txt` file, entry dated 2015-09-03]

To remove all Windows newlines:

```
M-x replace-regexp
C-q C-j RET
RET
```

To add a newline before heading names, e.g. “I.” or “A.”

```
M-x replace-regexp
.\. RET
C-q C-j \& RET
```

---

## Sort article lists by year using awk and sed

[Originally from Anthony's `science.txt` file, entry dated 2015-07-18.]

Exported bibliography to clipboard in Zotero, pasted into Emacs to write new file:

**…**`/VNLab/studies/ipsNumMeta/sources_number.txt`

Try using awk to print column with year before whole rest of line

```
gawk '{match($0,"\\([0-9]*[a-z]?\\)",a)} {print a[0], $0}' sources_number.txt
> sources_number_yearCol.txt
```

Remove parens from FIRST (year) on a line:

```
sed -r 's/[()([0-9]*[a-z]?)()]/\1/' sources_number_yearCol.txt >
sources_number_yearCol_noParens.txt
```

Used rectangle register copy trick in Emacs to copy only first five chars (catches both 2012 and 2012b) and pasted ultimately to Excel column.

From:
https://wiki.anthonycate.org/ - **Visual Cognitive Neuroscience**

Permanent link:
**https://wiki.anthonycate.org/doku.php?id=resources:emacs_resources&rev=1453163437**

Last update: **2019/05/22 20:08**