BBN–ANG–183 Typography
Lecture 7B: Punctuation

Zoltán Kiss & Péter Szigetvári

Dept of English Linguistics, Eötvös Loránd University
outline

where punctuation begins

punctuation marks
  sentence- and clause-final punctuation
  horizontal lines
  paired symbols
    brackets
    quotation marks
  other punctuation marks

order of punctuation marks

conclusion
the first punctuation mark

in the beginning

there were only uppercase letters (not even interword spaces): *scriptio continua*, typical of ancient Greek texts.

**THISISWHATTHISKINDOFTEXTLOOKSLIKE** **ITISNOTVERYEASILYREADABLEBUTBETTERTHAN** **NOTHINGINTERWORDSPACESAREINDISPENSABLEFORFASTREADING**

the first punctuation mark: the interpunct

the Romans used raised dots (or triangles) between words (many word processors represent space this way when showing formatting characters).
sentence-final punctuation

full stop/period

- ends sentences
- ends abbreviations
- separates “words” in email addresses and URLs
- separates decimal fractions in numerals (Anglo-Saxon convention)
- separates thousands, millions etc. in numerals (continental convention)

exclamation mark/point (screamer, bang)

- derives from Latin io ‘hey’: ↓ → !
- ends sentences
- within a sentence it is usually in parentheses: She has seven(!) cats.
- also used as a negation symbol: !≡ means ‘is not equivalent to’
sentence-final punctuation

question mark (interrogation point, eroteme)

▷ derives from Latin *Quæstio* ‘question’: O → ⬇ → ?
▷ typically sentence final, but also within sentences:
  Where do we come from? where do we go? and why?

Spanish and Catalan

have sentence initial inverted exclamation and question marks:
¿Qué hora es? ¡Olé!

emphasis

may be achieved by tripling exclamation and question marks:
Excellent!!! You really mean it???
don’t use more than three exclamation or question marks, and use this
effect with care, it may annoys the reader and easily gets inflated
clause-final punctuation

comma

- ends clauses
- separates the items of lists
- separates decimal fractions in numerals (continental convention)
- separates thousands, millions etc. in numerals (Anglo-Saxon convention), e.g., 100,000.00 vs. 100.000,00 (=one hundred thousand)

serial (Oxford, or Harvard) comma

- in lists of two items there is no comma before the conjunction: X and Y; X or Y
- in lists of more than two items there are two conventions:
  - X, Y and Z
  - X, Y, and Z

the orange comma is the serial comma
clause-final punctuation

**semicolon**

- is a “stronger” type of comma, linking related sentences:
  
  I am alone; my wife left me.

- used in lists with items containing commas:
  
  I traveled to London, England; Tijuana, Mexico; and Reykjavík, Iceland.

  Lisa scored 2,845,770 points; Marcia, 2,312,860; and Jeff, 1,726,640.

  the *orange* semicolons above are “serial semicolons”

**colon**

- precedes a clarification of what has been said before

- separates hours and minutes (and seconds) in time specifications: The time now is 10:48:42.

- separates the title and the subtitle in a reference list:
  
  Star Wars Episode IV: A New Hope
spacing

general rule for spaces around sentence- and clause-final punctuation

no space before, space after (think of where the line may be broken)

departures

- it is an old-fashioned convention to leave a larger space between sentences, i.e., after a sentence-ending (not abbreviation-ending) period, exclamation mark, question mark (and possibly also after a colon); most present-day authorities discourage this practice

- it is an French convention to leave a thin space before an exclamation mark, question mark, colon, and semicolon; beware: this must be a nonbreaking space!

both of these conventions are referred to as French spacing
French spacing

**space between sentences larger than between words**


**space between sentences same as between words**

French spacing

the text below is from a Hungarian publication of 1975 (Tersányszky J. Jenő, *Misi Mókus kalandjai*, Móra), illustrating French spacing: thin space before “!”, “?”, “:”, “;”

Hát a növendékek közül Maki Mókusra rájött a csintalankodhatnák. Különben jó tanuló volt és ügyes. Ő volt a tornajátékok vezetője. Most azonban így fordítette el a tanvers végét:

*Ahány magot találsz a fán,*
*Rakd a bendődbe, mókuskám!*

Úgy ám! De a tanító rögtön meghallotta ezt.
– Mit mondtál, te haszontalan? Nem szégyelled magad? Jó tanuló voltál eddig és játékvezető, mégis ilyen rossz példával szolgálsz tár- saidnak? Kitől tanultad ezt?
French spacing

this text is from a French book: Georges Mounin, *Clefs pour la linguistique*, Seghers, 1968; the spaces before colons and question marks is even greater

quement a partir de ses usages dans la langue. Ainsi aboutissent-ils à une procédure d'extraction de la signification dont Meillet a parfaitement énoncé la théorie dans la formule suivante : « Le sens d'un mot ne se laisse définir que par une moyenne entre les emplois linguistiques d'une part et les individus et les groupes d'une même société d'autre part. » C'est la théorie contextuelle de la signification. Un mot, dit-on, n'a aucun sens hors des contextes où il apparaît : que signifie le mot « mouche » isolé ? que signifie l'énoncé « Le boulanger fait des bâ-tards », surtout si on le sépare de la chanson de Jacques Dutronc *Paris s'éveille* ? Les mots qui n'apparaissent qu'une fois dans la somme des documents dont on dispose sur un état de langue — ce sont les *kanar* — sont la plupart du temps, impossibles
the hyphen (or dash)

has two main functions: separating (orthographical) syllables and linking words

end-of-line syllabification

<table>
<thead>
<tr>
<th>a hyphen is typically applied to break words</th>
<th>an old-fashioned alternative glyph: –</th>
</tr>
</thead>
<tbody>
<tr>
<td>this is often used as an end-of-line hyphen in hyphen-linked words</td>
<td></td>
</tr>
</tbody>
</table>

linking words

- e.g., twentieth-century writers
- suspended (or hanging, or dangling) hyphen: nineteenth- and twentieth-century writers
dashes

Typewriters and ASCII had a limited range of keys/codes: many symbols are omitted or merged; as a result, many users see no difference between a hyphen, a minus sign, an en-dash, or an em-dash.

The minus sign

Is used to indicate negative numbers ($0^\circ K \approx -273^\circ C$), and subtraction ($7 - 3 = 4$); it is longer than the hyphen: exactly as large as the horizontal bar of the plus sign, “+” (HTML: &minus;)

En-dash

The en-dash typically represents ranges (pp. 13–28, 15:30–17:00); its width is 1 en (half an em), usually thinner than the minus sign (HTML: &ndash;)

Em-dash

The em-dash usually delimits parenthetical material (Looking after cats—a noble task—needs patience.); its width is 1 em (HTML: &mdash;)
dashes

some typographers do not use separate glyphs

for an en- and an em-dash, but distinguish them by surrounding the latter by spaces:
They read pages 13–28 in the book – which turned out to be important.
or They read pages 13—28 in the book — which …

spacing

only em-dashes may have a space following them

replacement

if unavailable, the en-dash is usually replaced by two hyphens (pp. 13--28), the em-dash by three (book---which)
## brackets

### types

| Parentheses | round/curved/oval brackets, parens, fingernails |
| Brackets   | box brackets, square brackets |
| Braces     | curly brackets |
| Angle brackets | chevrons |

### embedding

When brackets are embedded (that is, when they are enclosed within each other another pair of brackets [viz. square brackets \{braces within those\}]), **may** be used within round brackets, however, it is more usual to use the same parentheses (round (or curved (or oval)) brackets); whichever you choose, be consistent.

### embedding in mathematical formulas

\[
\left( \left( 5 - \left( 13 \div 4 \right) \right) \times 7 - 2 \right) \div 3
\]
brackets

use

besides embedded parentheticals

▶ square brackets are used for
  ▶ explanatory or missing material (especially in quotes) ("I [...] don’t see it [the cat].")
  ▶ narrow, phonetic transcription
  ▶ deleted letters in classical philology ("cum Cæsar[e] venit")

▶ braces are used for
  ▶ sets ("\(\mathbb{N} = \{0, 1, 2, 3, \ldots\}\)
  ▶ disjunctive choices ("r \to \emptyset / \{C, \#\}")

▶ angle brackets are used for
  ▶ spelt forms (the word bε: may be spelt \langle bare\rangle or \langle bear\rangle)
  ▶ inserted letters in classical philology (cum C⟨æ⟩sare venit)

▶ primary school teachers have a bad habit of encouraging pupils to enclose letters/words to be ignored (deleted) in parentheses; such text should rather simply be rubbed out crossed out
brackets

replacement

- avoid replacing parentheses by slashes /like this/ (only for broad, phonemic transcriptions!)
- angle brackets are sometimes rendered as less-than and greater-than signs <like this>, a symptom of poor typography; ⟨this is right!⟩

spacing

there is a space before an opening and after a closing bracket (unless the latter is followed by some punctuation mark), there is no space after an opening and before a closing bracket; this is especially important when they are replaced /e.g., like here/
undirected and directed quotes

typewriters and ASCII have one symbol for opening and closing quotation marks: these are called undirected (or dumb) quotes (ASCII 96, the backtick, is not the mirrored version of the single quote)

```
"X" 'X' “X” ‘X’
```

undirected double quotes    undirected single quotes    directed double quotes    directed single quotes

many word processors automatically change undirected quotes to directed quotes (this feature is called “smart quotes”, and can usually be set somewhere like [Tools > AutoCorrect > Custom Quotes])

since the “right” direction is calculated from the previous character, wrong results may emerge, e.g., in the ’80s → in the ‘80s

⇒ be careful with smart quotes!
quotation mark shapes

the glyphs of quotation marks are the most language-specific of all punctuation marks: here are a few examples

<table>
<thead>
<tr>
<th>Language</th>
<th>Opening</th>
<th>Closing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anglo-Saxon</td>
<td>“one”</td>
<td>‘two’</td>
</tr>
<tr>
<td>Hungarian</td>
<td>„egy”</td>
<td>»kettő«</td>
</tr>
<tr>
<td>German</td>
<td>„eins“</td>
<td>,zwei‘</td>
</tr>
<tr>
<td>French</td>
<td>« un »</td>
<td>« deux » (called guillemets)</td>
</tr>
<tr>
<td>Dutch</td>
<td>«een»</td>
<td>&lt;twe&gt;</td>
</tr>
<tr>
<td>Finnish/Swedish</td>
<td>”yksi”</td>
<td>»kaksi»</td>
</tr>
</tbody>
</table>

- French leaves a thin space after an opening and before a closing guillemet
- in Hungarian text use the Hungarian glyphs, even when quoting an English phrase: vigyázzunk a „smart quote”-ok használatával!
embedding quotation marks

quotations within quotations are normally enclosed by a different set of quotation marks, e.g.,

- "HAL said, 'Good morning, Dave,'" recalled Frank.
- 'HAL said, "Good morning, Dave,"' recalled Frank.
- „»Jó reggelt, Dave« mondta Hal” emlékezett Frank.

the use of double quotes as primary quotation marks is a better choice, because they are more readily distinguishable from other punctuation marks than single quotes
other punctuation marks

apostrophe (/-/fi:/)

- used for genitives (Bob’s), omission (can’t, in the ’60s)
- not to be confused with the prime mark: ’ ≠ ’
- the glyph is usually the same as that of the closing single quote mark

ellipsis

in good typesetting systems the three dots of ellipsis are a special glyph: these dots are further apart from each other than three full stops (... ≠ ...) (except in a monospace font!)

slash (or oblique, virgule, slant, diagonal separatrix)

- indicates a disjunction (“and/or”)
- features in abbreviations (w/o = without, c/o = care of)
- indicates a joint between words (“the Gimson/Wells transcription”)
principles

- principle of logic: nest punctuation properly (i.e., close first what has been opened most recently)
- principle of esthetics: do not leave the baseline empty

conflicts

- quoting part of a sentence: both orders occur
  - logical order: They talked about “new data”.
  - esthetic order: They talked about “new data.”
- footnote references: only the esthetic order is common
  They talked about new data.\textsuperscript{12}
there is a plethora of punctuation conventions, 
the rule of thumb is: BE CONSISTENT!