

Hyphenation Control

Achieving Optimal Readability in Your Documents

László Németh

Csevej Bt., FSF.hu, Document Foundation



Collabora
Online

ut eos quos ex libris nouerant: corā
q; uiderēt. **S**icut pitagoras memphi-
ticos uates. sic plato egiptū. ⁊ architā
tarentinū. eandemq; oram ytalie. que
quondā magna grecia dicebat: labo-
riofissime peragrūit. et ut qui athenis
mōre erat. ⁊ potens. cuiusq; doctrinas
achademie gignasia plonabāt. fieret
pēgnus atq; discipulus. malēs aliena
uerecūde discere: qm̄ sua ipudentē ingerē.
Demq; cū lrās quasi toto orbe fugien-
tes psequit. capt⁹ a piratis ⁊ uenūda-
tus. tyrāno crudelissimo paruit. duct⁹
captiuus uind⁹ ⁊ seruus. **T**amē quia
plus maior emente se fuit. ad tyrum
lunū. lacteo eloquēcie fonte manantē.
de ultimis hispanie galliarūq; finibz.
quosdam uenisse nobiles legimus. ⁊
quos ad deumplacionē sui roma nō
traxerat: unū hoīs fama pduxit. **H**a-
buit illa etas inauditū omnibz seculis.
celebrandūq; miracūm. ut urbē tantā

Agenda

Understand the past

- ODF compliance
- OOXML compliance/DOCX interoperability

Prepare for the future

- CSS 4 features (Writer 24.2 and 24.8)
- Sidebar hyphenation control (Writer 25.2)

Make history

- LO/COOL Book Project (25.8 and 26.2)
- <http://www.numbertext.org/typography>

ODF compliance - part 1

- Exclude from hyphenation
- ODF 1.0
- Keep spell checking
- Keep font features

Hyphenation

Exclude from hyphenation

Ö →

Ö →

Die Frosch-Tunnel werden von der Öko-
bio Foundation gebaut und vom Um-
weltausschuss der Unive

Die Frosch-Tunnel w
Ökobio Foundation geba
weltausschuss der Unive

Die Frosch-Tunnel w
Ökobio Foundation geba
weltausschuss der Unive

Character

Font Font Effects **Position** Asian Layout Hyperlin

Position

Normal Raise/lower by: 1% - +

Superscript Relative font size: 100% - +

Subscript

Rotation / Scaling

0 degrees Scale width: 100% - +

90 degrees

270 degrees

Spacing

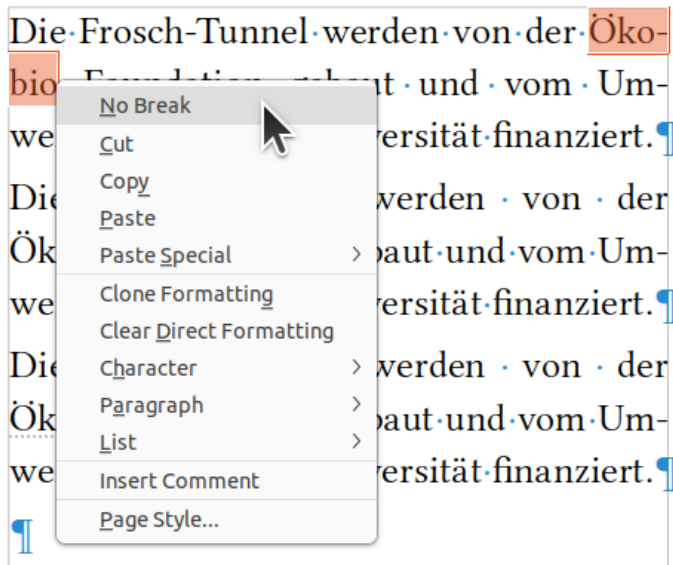
Character spacing: 0,0 pt - +

Hyphenation

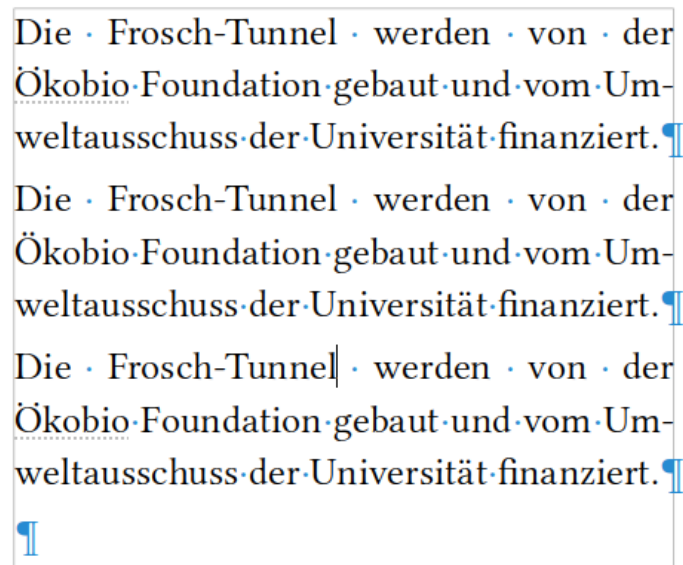
Exclude from hyphenation

Exclude words from hyphenation

Context menu



Visualization



ODF compliance – part 2

(Do not) hyphenate across page

- ODF 1.0
- CSS 4
- New Hart Rules (English orthography: visible page pair: spread)

The screenshot shows a text editor window with a paragraph of text: "Hyphenate last word of the page The Earth is no different to any other celestial body out there in space. It merely moves along in space inertially. Even just one inch above the surface of the Earth is space, except that it has an at- no different to any other". A cursor is positioned at the end of the paragraph. A "Paragraph" dialog box is open, showing the "Text Flow" tab. The "Hyphenation" section is expanded, showing the following settings:

- Automatically
- 2 Characters at line end
- 2 Characters at line begin
- 2 Compound characters at line end
- 0 Maximum consecutive hyphenated lines
- 5 Minimum word length in characters
- Hyphenation zone: 0,00 cm
- Hyphenate words in CAPS
- Hyphenate last word
- Hyphenation across**
 - Last full line of paragraph
 - Column
 - Page
 - Spread

The "Breaks" section shows:

- Insert
- Type: Page
- Position: Before
- With page style:
- Page number: 1

The "Split Options" section shows:

- Keep with next paragraph
- Allow to split paragraph
- No split at beginning of paragraph (2 lines)
- No split at end of paragraph (2 lines)

Buttons at the bottom of the dialog include "Help", "Reset", "Cancel", and "OK".

DOCX interoperability – part 1

Beware of the ladder

- Add missing DOCX import/export

According to a study by a German research institute, it is not possible to reproduce the control of consecutive hyphenated paragraph lines in the OpenDocument document format, as it doesn't support it.

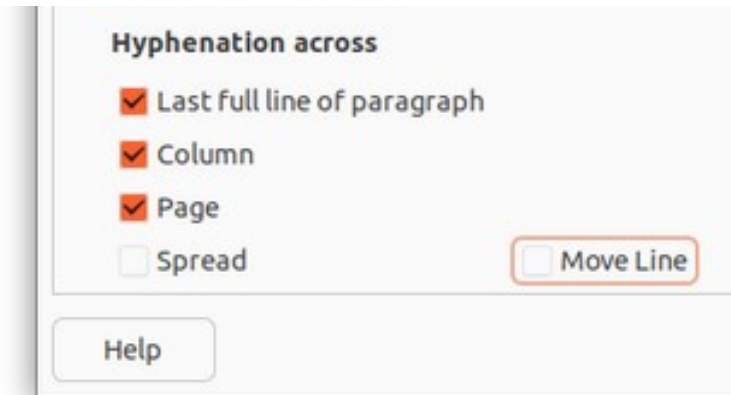


According to a study by a German research institute, it is not possible to reproduce the control of consecutive hyphenated paragraph lines in the OpenDocument document format, as it doesn't support it.

DOCX interoperability – part 2

Hyphenation across

- Add missing import/export and UX settings of 2 DOCX compatibility options
- Move last hyphenated word (default in MSO 2016 and newer) or move last hyphenated line (MSO 2013)



DOCX interoperability – part 3

Smart justify

- Very basic layout interoperability with MSO 2013 and newer

Integer sodales tincidunt tristique
viverra odio. Donec auctor
Cras sodales sed nisi mattis
pharetra, vestibulum, magna
fauces, faucibus, magna nec, a

DOCX interoperability – part 4

Inline heading

- Different paragraph styles in the same line
- Table of Content
- PDF navigation
- Select the first word(s) of the paragraph, and choose a style

*Word style separators are perhaps the key feature lacking in LibreOffice that prevents widespread Law office/legislator usage of LO Writer.
(Bruce Hatfield)*

Hyphenation of compound words

Computerpro-
gramm

Computer-
programm

- For headings and (future) Hyphenation across

Hyphenation

Automatically

2	-	+	Characters at line end
2	-	+	Characters at line begin
3	-	+	Compound characters at line end

Future ODF standardization

ODF/LibreOffice (white background: new in LibreOffice, "loext:" LibreOffice extensions for OpenDocument)	XSL-FO	CSS 4	OOXML/MS Word DOCX extension Document-level setting of hyphenation, but optional disable in paragraphs (suppressAutoHyphens).
hyphenate	hyphenate	hyphens	autoHyphenation
hyphenate (in character formatting)	hyphenate	hyphens	–
hyphenation-remain-char-count	hyphenation-remain-character-count	hyphenate-limit-chars	–
hyphenation-push-char-count	hyphenation-push-character-count		–
loext: hyphenation-word-char-count	–		–
loext:hyphenation-compound-remain-char-count	–	–	–
hyphenation-ladder-count	hyphenation-ladder-count	hyphenate-limit-lines	consecutiveHyphenLimit
hyphenation-keep (auto and page)	hyphenation-keep (auto, page and column)	hyphenate-limit-last (none, spread, page, column and always)	allowHyphenationAtTrackBottom and useWord2013TrackBottomHyphenation (default: column/page)
loext:hyphenation-no-caps	–	–	doNotHyphenateCaps
loext:hyphenation-no-last-word	–	–	–
loext:hyphenation-zone	–	hyphenate-limit-zone	hyphenationZone (default: 18 pt)

DTP capabilities – part 1

End zones

- Fine control in last hyphenated lines, like in desktop publishing

Hyphenation across

<input checked="" type="checkbox"/> Last full line of paragraph	Paragraph end zone:	<input type="text"/>	-	+
<input checked="" type="checkbox"/> Column	Column end zone:	27 pt	-	+
<input checked="" type="checkbox"/> Page	Page end zone:	<input type="text"/>	-	+
<input checked="" type="checkbox"/> Spread	Spread end zone:	48 pt	-	+

Move Line

Help

DTP capabilities – part 2

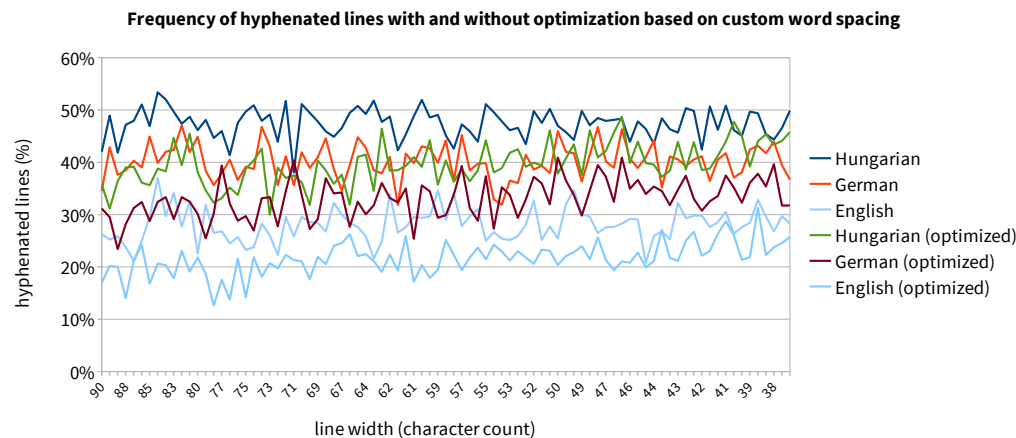
Custom word spacing

- On Paragraph settings → Adjustment pane
- Justification → minimum, desired and maximum word spacing
- E.g. for normal line width: 80%, 100%, 133% (default in DTP)
- Avoid of hyphenation within this range



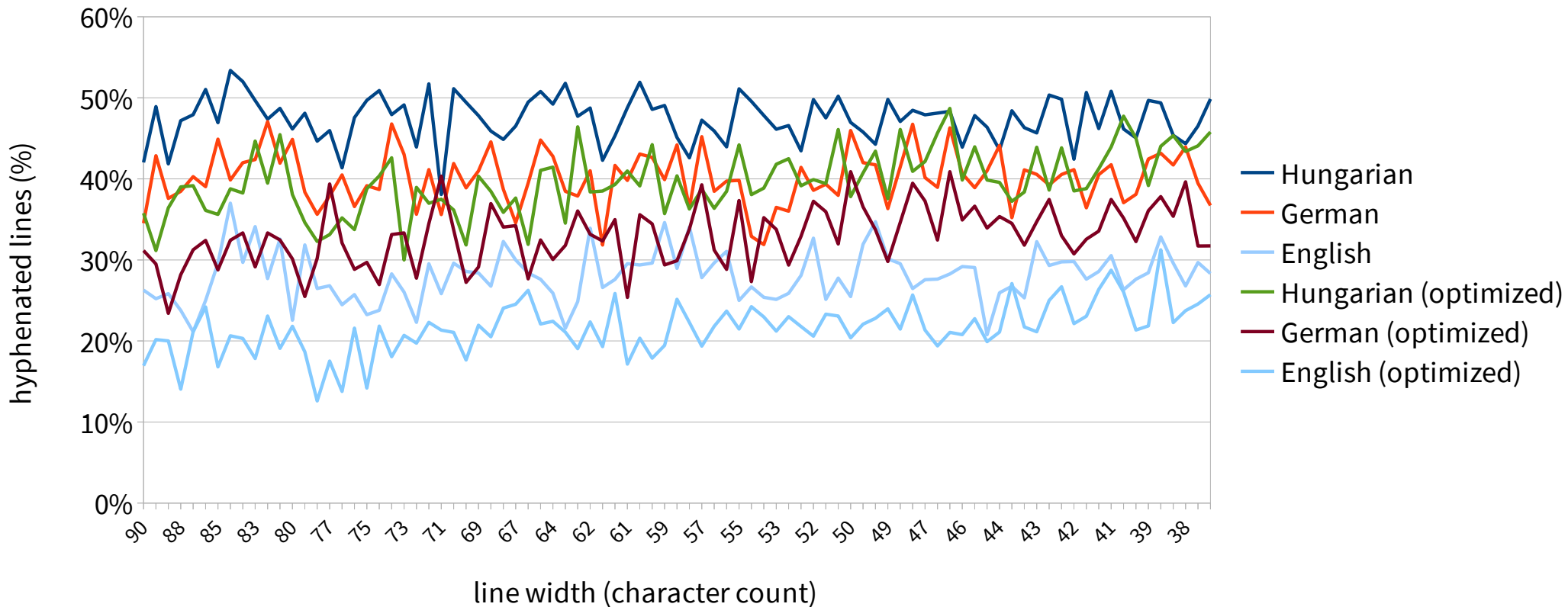
Research

- Automatic hyphenation of compound words at constituent boundaries in the open source LibreOffice Writer word processor
- Custom word spacing optimized for hyphenation (80%, 100%, 133%)
- same text, different line width: large fluctuation
- big average optimization:
 - English: 28% → 22%
 - German: 40% → 33%
 - Hungarian: 47% → 40%



Achieving optimal readability

Frequency of hyphenated lines with and without optimization based on custom word spacing



Smart justify - part 2

Shrinking is not enough

- At lot of spaces, 1-letter words
- Full algorithm revealed: minimum & maximum word spacing

Integer sodales tincidunt tristique. Sed a metus posuere, adipiscing nunc et, viverra odio. Donec auctor molestie sem, sit amet tristique lectus hendrerit sed. Cras sodales nisi sed orci mattis laculis. Nunc eget dolor accumsan, pharetra risus a, vestibulum mauris. Nunc vulputate lobortis mollis. Vivamus nec tellus faucibus, tempor magna nec, facilisis felis. Donec commodo enim a vehicula pellentesque. Nullam vehicula vestibulum est vel ultricies. Aliquam velit massa, laoreet vel leo nec, volutpat facilisis eros. Donec consequat arcu ut diam tempor luctus. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Praesent vitae lacus vel leo sodales pharetra a a nibh. Vestibulum ante ipsum primis in faucibus orci luctus et ultrices posuere cubilia Curae; Nam luctus tempus nibh, fringilla dictum augue consectetur eget. Curabitur at ante sit amet tortor pharetra molestie eu nec ante. Mauris tincidunt, nibh eu sollicitudin molestie, dolor sapien congue tortor, a pulvinar sapien turpis sed ante. Donec nec elit elementum, euismod nulla in, mollis nunc. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Vestibulum consequat mi quis pretium semper. Proin luctus orci ac neque venenatis, quis commodo dolor posuere. Curabitur dignissim sapien quis cursus egestas. Donec blandit auctor arcu, nec pellentesque eros molestie eget. In consectetur aliquam hendrerit. Sed cursus mauris vitae ligula pellentesque, non pellentesque urna aliquet. Fusce placerat mauris enim, nec rutrum purus semper vel. Praesent tincidunt neque eu pellentesque pharetra. Fusce pellentesque est orci. Integer sodales. Tincidunt tristique. Sed a metus posuere, adipiscing nunc et, viverra odio. Donec auctor molestie sem, sit amet tristique lectus hendrerit sed. Cras sodales nisi sed orci mattis laculis. Nunc eget dolor accumsan, pharetra risus a, vestibulum mauris. Nunc vulputate lobortis mollis. Vivamus nec tellus faucibus, tempor magna nec, facilisis felis. Donec commodo enim a vehicula pellentesque. Nullam vehicula vestibulum est vel ultricies. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Vestibulum consequat mi quis pretium semper. Proin luctus orci ac neque venenatis, quis commodo dolor posuere. Curabitur dignissim sapien quis cursus egestas. Donec blandit auctor arcu, nec pellentesque eros molestie eget. In consectetur aliquam hendrerit. Sed cursus mauris vitae ligula pellentesque, non pellentesque urna aliquet. Fusce placerat mauris enim, nec rutrum purus semper vel. Praesent tincidunt neque eu pellentesque pharetra. Fusce pellentesque est orci. Aliquam velit massa, laoreet vel leo nec, volutpat facilisis eros. Donec consequat arcu ut diam tempor luctus. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Praesent vitae lacus vel leo sodales pharetra a a nibh. Vestibulum ante ipsum primis in faucibus orci luctus et ultrices posuere cubilia Curae; Nam luctus tempus nibh, fringilla dictum augue consectetur eget. Curabitur at ante sit amet tortor pharetra molestie eu nec ante. Mauris tincidunt, nibh eu sollicitudin molestie, dolor sapien congue tortor, a pulvinar sapien turpis sed ante. Donec nec elit elementum, euismod nulla in, mollis nunc. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Vestibulum consequat mi quis pretium semper. Proin luctus orci ac neque venenatis, quis commodo dolor posuere. Curabitur dignissim sapien quis cursus egestas. Donec blandit auctor arcu, nec pellentesque eros molestie eget. In consectetur aliquam hendrerit. Sed cursus mauris vitae ligula pellentesque, non pellentesque urna aliquet.



Integer sodales tincidunt tristique. Sed a metus posuere, adipiscing nunc et, viverra odio. Donec auctor molestie sem, sit amet tristique lectus hendrerit sed. Cras sodales nisi sed orci mattis laculis. Nunc eget dolor accumsan, pharetra risus a, vestibulum mauris. Nunc vulputate lobortis mollis. Vivamus nec tellus faucibus, tempor magna nec, facilisis felis. Donec commodo enim a vehicula pellentesque. Nullam vehicula vestibulum est vel ultricies. Aliquam velit massa, laoreet vel leo nec, volutpat facilisis eros. Donec consequat arcu ut diam tempor luctus. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Praesent vitae lacus vel leo sodales pharetra a a nibh. Vestibulum ante ipsum primis in faucibus orci luctus et ultrices posuere cubilia Curae; Nam luctus tempus nibh, fringilla dictum augue consectetur eget. Curabitur at ante sit amet tortor pharetra molestie eu nec ante. Mauris tincidunt, nibh eu sollicitudin molestie, dolor sapien congue tortor, a pulvinar sapien turpis sed ante. Donec nec elit elementum, euismod nulla in, mollis nunc. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Vestibulum consequat mi quis pretium semper. Proin luctus orci ac neque venenatis, quis commodo dolor posuere. Curabitur dignissim sapien quis cursus egestas. Donec blandit auctor arcu, nec pellentesque eros molestie eget. In consectetur aliquam hendrerit. Sed cursus mauris vitae ligula pellentesque, non pellentesque urna aliquet. Fusce placerat mauris enim, nec rutrum purus semper vel. Praesent tincidunt neque eu pellentesque pharetra. Fusce pellentesque est orci. Integer sodales. Tincidunt tristique. Sed a metus posuere, adipiscing nunc et, viverra odio. Donec auctor molestie sem, sit amet tristique lectus hendrerit sed. Cras sodales nisi sed orci mattis laculis. Nunc eget dolor accumsan, pharetra risus a, vestibulum mauris. Nunc vulputate lobortis mollis. Vivamus nec tellus faucibus, tempor magna nec, facilisis felis. Donec commodo enim a vehicula pellentesque. Nullam vehicula vestibulum est vel ultricies. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Vestibulum consequat mi quis pretium semper. Proin luctus orci ac neque venenatis, quis commodo dolor posuere. Curabitur dignissim sapien quis cursus egestas. Donec blandit auctor arcu, nec pellentesque eros molestie eget. In consectetur aliquam hendrerit. Sed cursus mauris vitae ligula pellentesque, non pellentesque urna aliquet. Fusce placerat mauris enim, nec rutrum purus semper vel. Praesent tincidunt neque eu pellentesque pharetra. Fusce pellentesque est orci. Aliquam velit massa, laoreet vel leo nec, volutpat facilisis eros. Donec consequat arcu ut diam tempor luctus. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Praesent vitae lacus vel leo sodales pharetra a a nibh. Vestibulum ante ipsum primis in faucibus orci luctus et ultrices posuere cubilia Curae; Nam luctus tempus nibh, fringilla dictum augue consectetur eget. Curabitur at ante sit amet tortor pharetra molestie eu nec ante. Mauris tincidunt, nibh eu sollicitudin molestie, dolor sapien congue tortor, a pulvinar sapien turpis sed ante. Donec nec elit elementum, euismod nulla in, mollis nunc. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Vestibulum consequat mi quis pretium semper. Proin luctus orci ac neque venenatis, quis commodo dolor posuere. Curabitur dignissim sapien quis cursus egestas. Donec blandit auctor arcu, nec pellentesque eros molestie eget. In consectetur aliquam hendrerit. Sed cursus mauris vitae ligula pellentesque, non pellentesque urna aliquet.

pellentesque eros molestie eget. In consectetur aliquam hendrerit. Sed cursus mauris vitae ligula pellentesque, non pellentesque urna aliquet.

Acknowledgements

- LibreOffice/Collabora Online Typography project was funded through the NGIO Entrust Fund, a fund established by NLnet Foundation with financial support from the European Commission's Next Generation Internet programme.
- Compound word hyphenation support was funded by FSF.hu Foundation, Hungary
- Justin Luth (Collabora), Michiel Leenaars (NLnet), Gellért Gyuris and many others from the free software community

LO/COOL Book project

The project enhances readability of text documents by adding highly customizable paragraph-level line breaking and microtypography to Writer.

It creates a new type of software, with print quality of proprietary DTP programs and with productivity of word processors.



Collabora
Online

Thank you!

László Németh
nemeth@numbertext.org



@CollaboraOffice
hello@collaboraoffice.com
www.collaboraoffice.com