

All this words should be splitted when using "pdftotext -bbox test_wordbreaks.pdf"