

XLingPaper's use of TeX Technologies

H. Andrew Black & Hugh Paterson III

About the Presenters

Andy is the developer of XLingPaper (since 2001).

blackhandrew@gmail.com



Hugh is an XLingPaper user, bug-reporter, feature-requester, etc., since (2014).

i@hp3.me



Our Goals

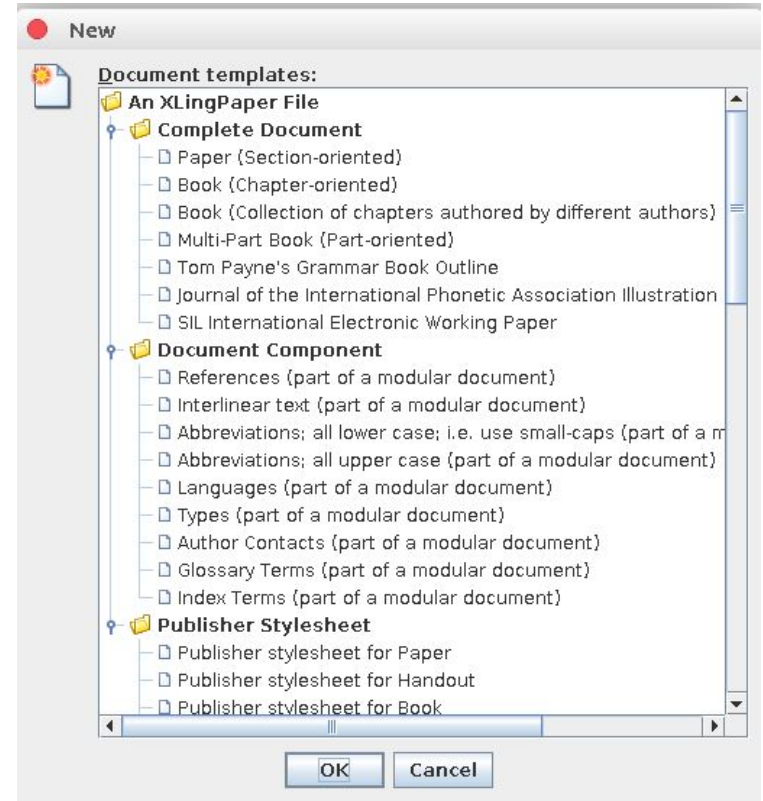
1. Say ‘thank you’ to the developers of the TeX packages we use.
2. Inform the TeX community of our project and the use of TeX technology.
3. Receive and consider any suggestions from the TeX community on the infrastructure design of XLingPaper.

XLingPaper - A Plugin to XMLMind



XMLMind Interface

- Written in JAVA
- Cross Platform (Windows, MacOS, Linux)
- Specifically designed document types
- XML Editor designed for technical document production



XMLMind Interface

The screenshot displays the XMLMind software interface. The main window shows a document titled "On rights management in anthropological and linguistic sound collections" by Hugh J. Paterson III, dated 25 February 2021, version 5 with en-dash corrections. The document is structured with a table of contents on the left and a validation panel on the right.

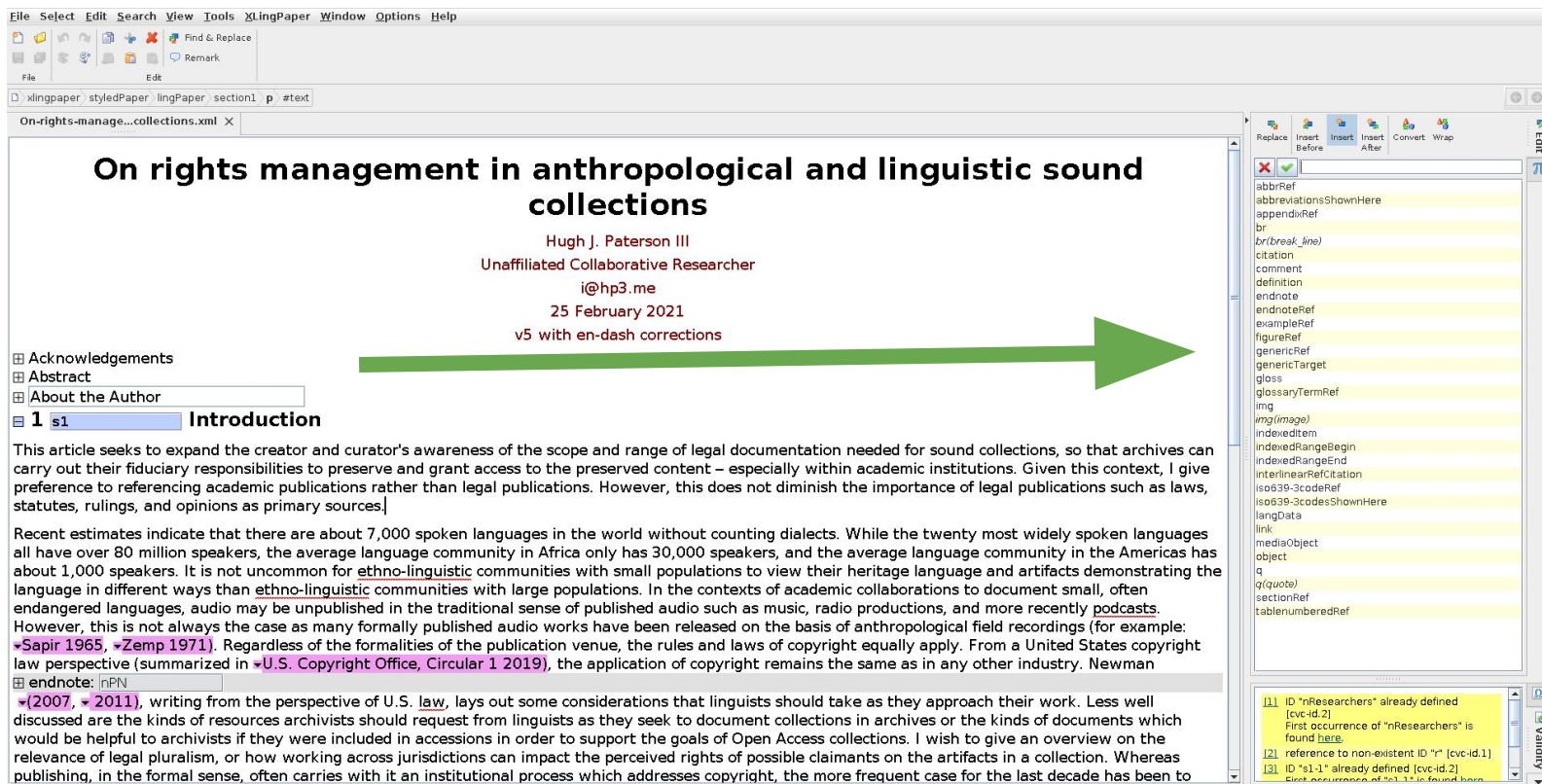
Table of Contents:

- Acknowledgements
- Abstract
- About the Author
- 1 s1 Introduction
- 2 s2-4 Preliminaries
- 3 s2 Plurality and conflicting approaches to ownership
- 4 s Considerations
- 5 s6 Conclusion
- 6 sConsultedSource Consulted Sources
- Endnotes
- References
- Languages:
- Types:
- Publishing info:
- Content control:
- Publisher Stylesheet: ARSC-Journal**

Validation Panel:

- [1] ID "nResearchers" already defined [cvc-id.2] First occurrence of "nResearchers" is found [here](#).
- [2] reference to non-existent ID "r" [cvc-id.1]
- [3] ID "s1-1" already defined [cvc-id.2] First occurrence of "s1-1" is found [here](#).
- [4] ID "s1-2" already defined [cvc-id.2] First occurrence of "s1-2" is found [here](#).

XMLMind Interface



File Select Edit Search View Tools XMLPaper Window Options Help

Find & Replace
Remark

File Edit

D:\xmlpaper\styledPaper\lingPaper\section1 p #text

On-rights-manage...collections.xml X

On rights management in anthropological and linguistic sound collections

Hugh J. Paterson III
Unaffiliated Collaborative Researcher
i@hp3.me
25 February 2021
v5 with en-dash corrections

Acknowledgements
Abstract
About the Author

1 s1 Introduction

This article seeks to expand the creator and curator's awareness of the scope and range of legal documentation needed for sound collections, so that archives can carry out their fiduciary responsibilities to preserve and grant access to the preserved content – especially within academic institutions. Given this context, I give preference to referencing academic publications rather than legal publications. However, this does not diminish the importance of legal publications such as laws, statutes, rulings, and opinions as primary sources.

Recent estimates indicate that there are about 7,000 spoken languages in the world without counting dialects. While the twenty most widely spoken languages all have over 80 million speakers, the average language community in Africa only has 30,000 speakers, and the average language community in the Americas has about 1,000 speakers. It is not uncommon for ethno-linguistic communities with small populations to view their heritage language and artifacts demonstrating the language in different ways than ethno-linguistic communities with large populations. In the contexts of academic collaborations to document small, often endangered languages, audio may be unpublished in the traditional sense of published audio such as music, radio productions, and more recently podcasts. However, this is not always the case as many formally published audio works have been released on the basis of anthropological field recordings (for example: Sapir 1965, Zemp 1971). Regardless of the formalities of the publication venue, the rules and laws of copyright equally apply. From a United States copyright law perspective (summarized in U.S. Copyright Office, Circular 1 2019), the application of copyright remains the same as in any other industry. Newman

endnote: hPN

(2007, 2011), writing from the perspective of U.S. law, lays out some considerations that linguists should take as they approach their work. Less well discussed are the kinds of resources archivists should request from linguists as they seek to document collections in archives or the kinds of documents which would be helpful to archivists if they were included in accessions in order to support the goals of Open Access collections. I wish to give an overview on the relevance of legal pluralism, or how working across jurisdictions can impact the perceived rights of possible claimants on the artifacts in a collection. Whereas publishing, in the formal sense, often carries with it an institutional process which addresses copyright, the more frequent case for the last decade has been to

Replace Insert Before Insert After Convert Wrap

abbrRef
abbreviationsShownHere
appendixRef
br
br(break_line)
citation
comment
definition
endnote
endnoteRef
exampleRef
figureRef
genericRef
genericTarget
gloss
glossaryTermRef
img
img(image)
indexedItem
indexedRangeBegin
indexedRangeEnd
interlinearRefCitation
iso639-3codeRef
iso639-3codesShownHere
langData
link
mediaObject
object
q
q(quote)
sectionRef
tablenumberedRef

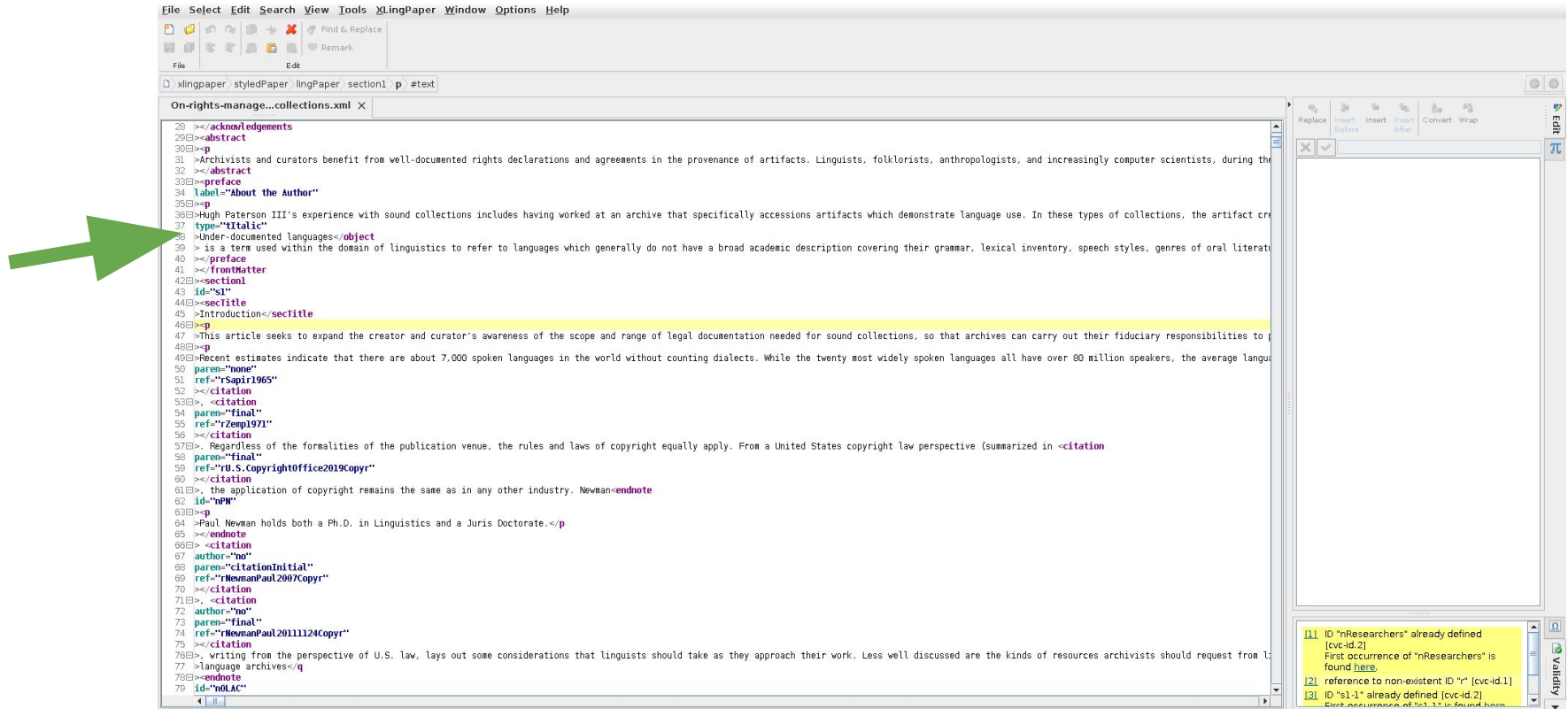
Validity

[1] ID "nResearchers" already defined [cvc-id.2]
First occurrence of "nResearchers" is found here.

[2] reference to non-existent ID "r" [cvc-id.1]

[3] ID "s1-1" already defined [cvc-id.2]
First occurrence of "s1-1" is found here.

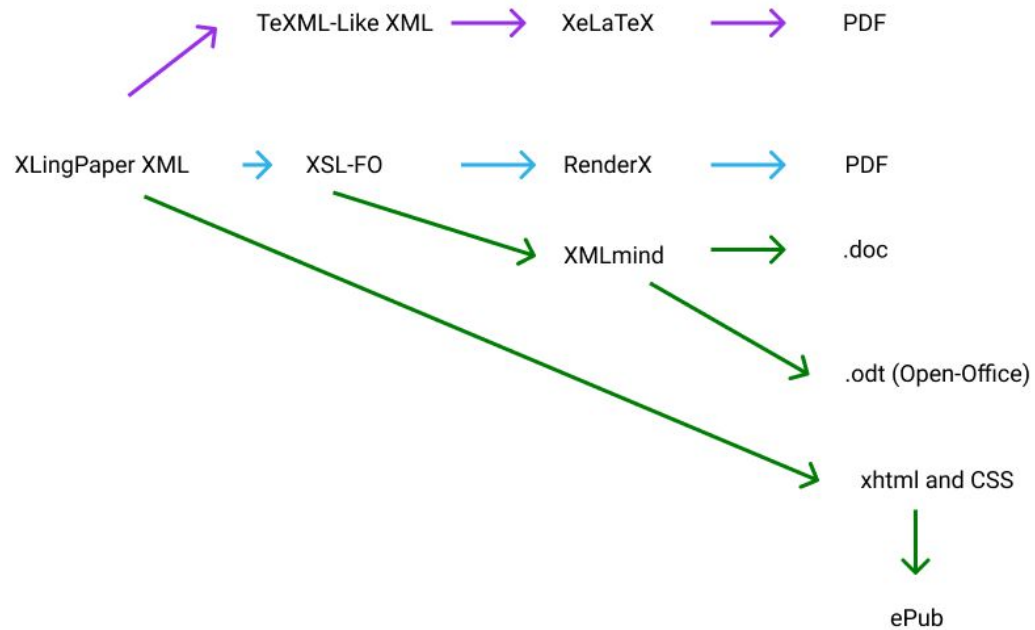
XMLMind Interface



XLingPaper Design Goals

- As free as possible.
- Content separated from styles with publishers having unique publishing style sheets expressed in XML.
- Low bandwidth target user group
 - No TeXLive
 - Pre-selected packages

PDF Creation



TeXML

- TeXML requires Python and the design requirements did not want to dictate that XLingPaper users to need to install a version of Python for TeXML when that version may conflict with other versions of Python they might already have installed. Furthermore, without adding Python the installation package would remain much smaller.
- XLingPaper also needed some extensions for formatting white space (more finely as far as the developers could determine).

Ling-TeX

- Contains packages that provide solutions for interlinear texts, but do not allow for the larger number of capabilities XLingPaper already dealt with. So the developer rolled his own.

(Remember XLingPaper started with RenderX for PDF processing.)

TeX Packages Used

- attachfile2
- booktabs
- calc
- color
- colortbl
- etoolbox
- fancyhdr
- fontspec
- footmisc
- hyperref
- lineno
- longtable
- lscape
- mdframed
- multirow
- normalem
- polyglossia
- setspace
- tabularx
- xltextra

Custom TeX Commands

Command for	Purpose
Table of contents	Store and retrieve page numbers; format the contents.
Lists	Numbered and bulleted lists with control over indents, etc.
Examples	Example number and example content, where the content can be a line, a list of lines, a set of words, a list of a set of words, interlinear, a list of interlinears, etc.
Indexes	Handle keeping track of XLingPaper's indexing capability, including page numbers.
Interlinears	Handle lines in an interlinear text or example, including dealing with an ISO 639-3 code in an interlinear example.
Block quotes	Handle special cases needed for block quotes.
Table headers	Attempt to calculate a column's width via its contents.

Typesetting Tasks XLPaper Users Often Encounter

- Common linguistic numbered example layouts:
 - List of words and their glosses
 - Interlinear clauses
 - Headings for an example
- Automatically wrapping interlinear texts
- Abbreviations used in glosses and interlinears
- Hyperlinks between:
 - Example number references and examples
 - Abbreviations and a list of abbreviation definitions

Interlinear examples (1/4)

In the XMLMind Editor (part 1/2):

2.1.1 sCuant La frase cuantificadora

Una frase cuantificadora puede acompañar al sustantivo (véanse [Los Cuantificadores](#) y [Los Números Cardinales](#)). Cuando se presenta esta frase, siempre va delante del núcleo de la frase nominal, como en los ejemplos en (▼xQN).

(xQN)
(xSmajiin6) (T-ID0EFGAC-6)
(xFC5p1) (T-ID0EMKAE-5.1)

El cuantificador puede presentarse en construcciones donde no hay sustantivo expreso, como se explica en [Los Cuantificadores](#). Un ejemplo se incluye aquí.

Interlinear examples (2/4)

In the XMLMind Editor (part 2/2):

(T-ID0EFGAC-6)

Náa majñuu nákhū iduu iya'

nāā māhjūùⁿ nákhù īdūū ījā?

▼LOC entre ▼TOT.cuatro ojo.▼3SG agua

'De entre los cuatro manantiales

⊕ endnote: nManantial

'

(T-ID0EMKAE-5.1)

Gí'doo witsu rakhóó mikhúdú

▼EST.tener.▼3SG cinco nariz.▼3SG (▼EST).picud@

'Tiene cinco esquinas picudas'

Interlinear examples (3/4)

In the XeLaTeX PDF output:

2.1 Los elementos que van antes del sustantivo

Hay una posición para un elemento antes del núcleo de la frase nominal, y puede ser ocupada por una frase cuantificadora, un artículo indefinido o un pronombre para un uso muy especializado.

2.1.1 La frase cuantificadora

Una frase cuantificadora puede acompañar al sustantivo (véanse [Los Cuantificadores](#) y [Los Números Cardinales](#)). Cuando se presenta esta frase, siempre va delante del núcleo de la frase nominal, como en los ejemplos en (2).

- (2) a. [tcf- Naa majñuu nákhü iduu iya'
Zila] náā mähjüüⁿ nákù idūū ījā?
LOC entre TOT.cuatro ojo.3SG agua
'De entre los cuatro manantiales⁵'

[Smajiin:6]

- b. [tpl- Gí'doo witsu rakhóó mikhúdú
Tlac] EST.tener.3SG cinco nariz.3SG (EST).picud@
'Tiene cinco esquinas picudas'

[FC:5.1]

Source Document:

<https://mexico.sil.org/sites/mexico/files/mephaafrasenominal.pdf>

El cuantificador puede presentarse en construcciones donde no hay sustantivo expreso, como se explica en [Los Cuantificadores](#). Un ejemplo se incluye aquí.

Interlinear examples (4/4)

Abbreviations are hyperlinks to their definition.

XeLaTeX-produced PDF after clicking on the LOC abbreviations:

~~irrealis, IRRSBD = irrealis subordinado, LOC = locativo, MD = medial, NEG = negativo, PSV = pasivo,~~
~~PFV = perfectivo, PL = plural, PROG = progresivo, RAZÓN = razón, RR = reflexivo/recíproco, SBD:ANIM~~
~~= subordinador animado, SBD:INAN = subordinador inanimado, TOT = totalidad.~~

Automatic wrapping in interlinear text (1/2)

Text as in XLingPaper:

1 stexto Rikha
endnote: nrikha
(FC) Interlinear text: FC
shortened title: FC
(T-ID0EMKAE-1)
Rikha rígi' najmaā náā yúoo' ra'kha ká',
endnote: nMax
ra'kha suan'
endnote: nespina
khamí náā yúoo' ra'kha' májin'.
endnote: nchilacayote1
flor.de.calabaza ▼INAN:PROX ▼IMPF.producirse ▼LOC guía.▼3SG calabaza.especie
calabaza.especie y ▼LOC guía.▼3SG chilacayote
'La flor de calabaza se da en la guía de la calabaza de Castilla, de la "calabaza
espina" y del chilacayote.'

(T-ID0EMKAE-2)
Rí rikhoo ra'kha suan', nagí'duu namídi rí gun' agóstó.
▼SBD:INAN flor.de.calabaza.▼3SG calabaza.especie ▼IMPF.empezar.▼3SG.▼FM ±
▼IMPF.florear ▼SBD:INAN luna agosto*
'La flor de la "calabaza espina" empieza a abrir en el mes de agosto.'

(T-ID0EMKAE-3)
Mba'ju, mujmu' rí'jiuu.
(▼EST).grande:▼PL (▼EST).amarill@ flor.▼3SG
'Sus flores son grandes y amarillas.'

Automatic wrapping in interlinear text (2/2)

XeTeX-produced PDF:

Rikha²

FC:1

Rikha rígi' najmaā náā yúoo' rā'khā ká',³ rā'khā suan'⁴
flor.de.calabaza INAN:PROX IMPF.producirse LOC guía.3SG calabaza.especie calabaza.especie

khamí náā yúoo' rā'khā májin'.⁵
y LOC guía.3SG chilacayote

'La flor de calabaza se da en la guía de la calabaza de Castilla, de la "calabaza espina" y del chilacayote.'

FC:2

Rí rikhoo rā'khā suan', nagí'dyū namīdī rí
SBD:INAN flor.de.calabaza.3SG calabaza.especie IMPF.empezar.3SG.FM ± IMPF.florear SBD:INAN

gūn' agóstó.
luna agosto*

'La flor de la "calabaza espina" empieza a abrir en el mes de agosto.'

FC:3

Mba'ju, mujmu' rī'jiyū.
(EST).grande:PL (EST).amarill@ flor.3SG

'Sus flores son grandes y amarillas.'

FC:4

Typesetting Tasks XLPaper Users Often Encounter

- Tables
- Font rendering
- Hyphenation for non-English languages
- Author contact information
- Vertical fill
- Blank page

XeLaTeX font rendering (1/2)

pattern is exemplified in the Mbo data in (▼xATR-zmw):

In XLingPaper:

```
(xATR-zmw )
(xi-zmw ) [ex[ko-sis-o]pr] [ex[tPitch 444.111.111]pr] ‘move
forward’
(xia-zmw ) [ex[kɔ-kij-a]pr] [ex[tPitch 444.111.111]pr] ‘act’
(xuo-zmw ) [ex[ko-bund-o]pr] [ex[tPitch 444.111.111]pr] ‘break’
(xɯa-zmw ) [ex[kɔ-bɯt-a]pr] [ex[tPitch 444.111.111]pr] ‘become
long’
(xeo-zmw ) [ex[ko-beɲ-o]pr] [ex[tPitch 444.111.111]pr] ‘wink’
(xɛa-zmw ) [ex[kɔ-kek-a]pr] [ex[tPitch 444.111.111]pr] ‘decorate’
(xoo-zmw ) [ex[ko-sok-o]pr] [ex[tPitch 444.111.111]pr] ‘cackle’
(xɔa-zmw ) [ex[kɔ-mvɔɖ-a]pr] [ex[tPitch 444.111.111]pr] ‘suck’
(xaa-zmw ) [ex[kɔ-bab-a]pr] [ex[tPitch 111.444.444]pr] ‘carry’
```


XeLaTeX font rendering (2/2)

XeLaTeX-produced (using Graphite with Charis SIL font):

Bantu D30 canonical infinitive verb pattern is exemplified in the Mbo data in (11):

- (11)
- | | | | |
|----|-------------------|-----------------|--------------|
| a. | [ex[ko-sis-o]ex] | [ex[[- - -]]ex] | move forward |
| b. | [ex[kɔ-kij-a]ex] | [ex[[- - -]]ex] | act |
| c. | [ex[ko-ɓund-o]ex] | [ex[[- - -]]ex] | break |
| d. | [ex[kɔ-ɓut-a]ex] | [ex[[- - -]]ex] | become long |
| e. | [ex[ko-ɓɛɲ-o]ex] | [ex[[- - -]]ex] | wink |
| f. | [ex[kɔ-kɛk-a]ex] | [ex[[- - -]]ex] | decorate |
| g. | [ex[ko-sok-o]ex] | [ex[[- - -]]ex] | cackle |
| h. | [ex[kɔ-mvɔɗ-a]ex] | [ex[[- - -]]ex] | suck |
| i. | [ex[kɔ-bab-a]ex] | [ex[[- - <]]ex] | carry |

XLingPaper website

<https://software.sil.org/xlingpaper/>