

De L^AT_EX à X_ƎL^AT_EX

Laurent Roussarie

30 avril 2019

1 La migration

Il n'existe pas vraiment de documentation de référence de X_ƎL^AT_EX (les utilisateurs sont souvent renvoyés à celle du package `fontspec`). Normalement il n'y a pas de différence entre L^AT_EX et X_ƎL^AT_EX pour ce qui concerne le corps du document ; et donc tout se joue dans le préambule.

X_ƎL^AT_EX est un « dialecte » de L^AT_EX qui se distingue principalement par sa gestion des polices de caractères et de l'encodage UTF-8 (unicode).

Donc *normalement*, une fois que les polices sont correctement appelées, le corps du document suit la même syntaxe qu'un document L^AT_EX classique. Il faut juste compiler avec `xelatex` au lieu de `pdflatex`.

1.1 Pourquoi passer à X_ƎL^AT_EX

X_ƎL^AT_EX offre plusieurs avantages, le principal se présente lorsque l'on souhaite i) saisir du texte dans des alphabets non latins directement en unicode (par exemple au clavier) et/ou ii) rédiger dans ces alphabets en utilisant d'autres polices que la famille Computer Modern (qui est la police standard de L^AT_EX).

1.2 Packages L^AT_EX à ne pas utiliser avec X_ƎL^AT_EX

- `fontenc`
- `inputenc`
- `babel` (les nouvelles version de `babel` sont maintenant compatibles avec X_ƎL^AT_EX et fonctionnent bien, mais restent encore en rodage sur certaines langues)
- `tipa` – mais ce n'est pas si clair : selon les versions il est utile, voire nécessaire de l'appeler ; tant que ça marche, je conseille de l'appeler (si on veut écrire en API, bien sûr)
- `breakurl` (utile pour `pdflatex` en complément de `hyperref`, mais inutile et perturbant pour X_ƎL^AT_EX)

1.3 Packages à utiliser avec X_ƎL^AT_EX

1.3.1 Remplacements

- `polyglossia` en place de `babel`. Exemple d'utilisation :

```
\usepackage{polyglossia}
\setdefaultlanguage{french} %% Langue principale du document
\setotherlanguage[variant=uk]{english} %% Autres langues qui
\setotherlanguage{arabic} %% seront utilisées
```

❗ À terme, **babel** devrait complètement supplanter **polyglossia** pour $\text{X}_{\text{T}}\text{L}_{\text{A}}\text{T}_{\text{E}}\text{X}$. Et aujourd’hui, pour écrire simplement en français ou en anglais, par exemple, je conseille d’utiliser **babel** (y compris si le document comprend de l’API). Pour les autres alphabets, **polyglossia** reste encore d’utilisation plus aisée. Les exemples qui suivent (sauf indications contraires) sont sous **polyglossia**.

1.3.2 Packages nécessaires

- **fontspec**. Même si on ne sait pas trop s’en servir, mieux vaut prendre l’habitude de l’appeler. Nécessaire pour charger correctement des polices.

2 Fonctionnalités avancées

2.1 Les polices

Le package **fontspec** permet en théorie d’utiliser n’importe quelle police installée sur l’ordinateur (notamment les polices True Type). Mais en fait peu de polices ont un support unicode étendu et très peu sont suffisamment complètes pour offrir à la fois les alphabets non latins, les symboles API et les symboles mathématiques.

Voici quelques polices recommandables :

- CMU Serif : API (əf:), cyrillique (яжд), grec (φκω)
- Linux Libertine : API (əf:), cyrillique (яжд), grec (φκω)
- Gentium : API (əf:), cyrillique (яжд), grec (φκω)
- STIX : API (əf:), cyrillique (яжд), grec (φκω)
- Times New Roman : API (əf:), cyrillique (яжд), grec (φκω)
- DejaVu Serif : API (əf:), cyrillique (яжд), grec (φκω)
- Free Serif : API (əf:), cyrillique (яжд), grec (φκω)
- Charis SIL : API (əf:), cyrillique (яжд)¹

Voir cette page https://fr.overleaf.com/learn/latex/Questions/What_OTF/TTF_fonts_are_supported_via_fontspec%3F pour trouver d’autres polices plus ou moins accessibles et utilisables dans divers alphabets.

Pour savoir si une police est installée sur votre ordinateur, consultez le répertoire **Fonts** de votre système. Si certaines manquent, elles peuvent se télécharger facilement depuis le Web.

2.2 Appeler un police

Par défaut, $\text{X}_{\text{T}}\text{L}_{\text{A}}\text{T}_{\text{E}}\text{X}$ utilise la police Latin Modern ; mais celle-ci n’a pas une vraie couverture unicode, donc il va falloir appeler une autre police.

Cela se fait avec les commandes de **fontspec**.

Les polices sont identifiées soit par leur nom « officiel » (qui apparaît dans les informations données quand on double-clique sur l’icône du fichier de la police, par exemple *DejaVu Serif*), soit par le nom du fichier de la police (par exemple *DejaVuSerif.ttf*).

1. Très bonne couverture API, mais pas de grec et probablement malaisée pour les formules mathématiques.

2.2.1 Police principale

On spécifie la police principale avec `\setmainfont`. Normalement c'est une police romane (aka *serif*).

Exemples d'appel des polices mentionnées *supra* :

```
\setmainfont{CMU Serif}
```

```
\setmainfont{Linux Libertine 0}
```

```
\setmainfont[BoldFont={Gentium Basic Bold},ItalicFont={Gentium Italic}]{Gentium}
```

```
\setmainfont{STIX}
```

```
\setmainfont{Times New Roman}
```

```
\setmainfont{DejaVu Serif}
```

```
\setmainfont{FreeSerif}
```

```
\setmainfont{Charis}
```

2.2.2 Les deux autres familles

Parallèlement, on peut aussi spécifier une police particulière pour la famille linéale (aka sans serif) et pour la famille teletype avec les commandes `\setsansfont` et `\setmonofont` respectivement.

Par exemple :

```
\setsansfont{Arial}
\setmonofont{CMU Typewriter Text}
```

De la même façon, il y a la commande `\setromanfont` pour la famille romane, mais normalement celle-ci est déjà spécifiée par `\setmainfont`.

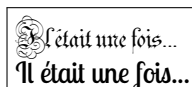
2.2.3 Appeler d'autres polices

On peut aussi appeler autant de polices supplémentaires qu'on le souhaite (contrairement à ce qui se passe typiquement avec \LaTeX). Pour ce faire on crée une commande dédiée à l'aide de `\newfontfamily`.

Exemples :

```
\newfontfamily\plume[Scale=1.5]{Parchment} %% La police Parchment est très petite, on la
redimensionne avec Scale
\newfontfamily\cursive{Lobster Two}
```

Ensuite les commandes ainsi créées (ici `\plume` et `\cursive`) deviennent des *déclarations* utilisables dans le corps de document :



```
{\large\plume Il était une fois...}
```

```
{\cursive Il était une fois...}
```

NB : comme les déclarations ne sont pas toujours très pratiques, ça vaut la peine de se définir une commande qui prend un argument. Par exemple :

```
\newcommand{\encursive}[1]{\cursive#1}
```

Ainsi on pourra saisir `\encursive{Il était une fois...}`.

- ❶ Il est souvent utile (voire nécessaire) d'ajouter l'option `Ligatures=TeX` avec `\newfontfamily`, notamment pour avoir un rendu correct des apostrophes (' au lieu de ')².

```
\newfontfamily\cursive[Ligatures=TeX]{Lobster Two}
```

2.3 Écrire dans d'autres alphabets

2.3.1 API

Si la police principale sait afficher les caractères phonétiques, on peut saisir directement de l'API en unicode dans le `.tex`, en tapant (ou collant) `əʃɪh` par exemple.

Mais je recommande d'appeler l'extension `tipa` et de saisir toutes les séquences phonétiques dans la commande `\texttipa`. Par exemple : `\texttipa{əʃɪh}`.

Mieux encore : définir sa propre commande, plus courte et éventuellement modifiable si nécessaire :

```
\newcommand{\API}[1]{\texttipa{#1}}
```

Utiliser `\texttipa` permet d'accéder aux raccourcis de `tipa` qui peuvent être très pratiques au clavier. Par exemple `\texttipa{/S@"G4A:K/}` donnera `/ʃə'ʏɥɑ:ɪ/`.

Voir la documentation de `tipa` pour toutes les fonctionnalités offertes.

- ❶ Pour combiner efficacement `tipa` et les caractères unicodes, il est nécessaire d'appeler `fontspec` avec l'option `euenc` :

```
\usepackage[euenc]{fontspec}
```

- ❶ `tipa` s'utilise aussi avec \LaTeX , mais uniquement avec ses raccourcis (l'unicode ne passe pas). Et le résultat sera forcément en Computer Modern (ce qui peut faire bizarre si la police principale du texte est différente).

2.3.2 Alphabets proches de l'alphabet latin

Typiquement le grec et le cyrillique.

Là encore on peut les saisir directement en unicode dans le `.tex`, par exemple en saisissant `ευχαριστώ` ou `спасибо`.

Mais il est plus prudent de compléter cela avec le support linguistique de `polyglossia`.

Pour ce faire, d'abord déclarer ces langues avec `\setotherlanguage` :

```
\setotherlanguage{greek}
\setotherlanguage{russian}
```

À ce moment-là, `polyglossia` fournit les commandes `\textlang{...}` et l'environnement `\begin{lang} ... \end{lang}` où `lang` est le nom de la langue du texte.

```
\textgreek{ευχαριστώ}  donne :  ευχαριστώ
\textrussian{спасибо}  donne :  спасибо
```

- ❶ `polyglossia` définit aussi une commande `\langfont` qui correspond à la déclaration de la police associée à la langue `lang`. On peut redéfinir cette commande à l'aide de `\newfontfamily` (cf. *supra*) pour changer la police de cette langue. Mais, si la police principale du texte couvre déjà le grec ou le cyrillique, ce n'est pas utile de faire cet ajustement (il est même préférable de rester dans la même fonte, pour l'harmonie graphique du texte). Mais en cas de message d'erreur, faire, par exemple :

```
\newfontfamily\greekfont{STIX} %% en adaptant à la police principale
```

2. Ainsi que des ff, fl et fi au lieu de ff, fl et fi...

2.3.3 Autres alphabets

On procède comme précédemment, mais là il est généralement nécessaire de préciser la police `\langfont`.

Exemple pour l'arabe :

```
\setotherlanguage{arabic}
\newfontfamily\arabicfont[Script=Arabic,Scale=1.25]{Times New Roman}}
```

Il se trouve que Times New Roman (ainsi que Arial) ont une couverture pour l'arabe, mais ce n'est pas le cas de toutes les polices.

Voir https://fr.overleaf.com/learn/latex/Questions/What_OTF/TTF_fonts_are_supported_via_fontspe%3F pour trouver d'autres polices qui écrivent en arabe.

`\textarabic{شكرا}` donne : شكرا

Pour un paragraphe entier en arabe, mieux vaut utiliser l'environnement `\begin{Arabic} ... \end{Arabic}`³ (donnera le sens d'écriture droite→gauche) :

شكرا

Exemple pour l'hébreu :

```
\setotherlanguage{hebrew}
\newfontfamily\hebrewfont[Script=Hebrew]{Times New Roman}
```

Avec `\texthebrew{...}` on obtiendra : תודה et avec `\begin{hebrew} ... \end{hebrew}` :
תודה

2.3.4 Autres stratégies

- On peut écrire en arabe aussi sous L^AT_EX. Pour un texte qui n'insère que quelques passages en arabe, le plus simple est d'utiliser l'extension **dad**.
- Pour un document complètement en arabe ou en hébreu, mieux utiliser l'extension **arabtex** (L^AT_EX) ou **arabxetex** (X_YL^AT_EX).
- **polyglossia** ne couvre ni le chinois ni le japonais. Pour ces langues, voir les extensions **cjk** et **xecjk**, et tout autres documentations disponibles sur internet.

3 Couvertures mathématiques

Si on utilise une autre police principale que la police standard Computer Modern (ou ses clones comme CMU Serif), les formules mathématiques ne seront pas harmonisées avec le reste du texte, et ça sera moche.

Il convient donc d'appeler une extension supplémentaire (après **fontspec**) pour mettre les écritures mathématiques au diapason.

Pour Linux Libertine :

```
\setmainfont{Linux Libertine O}
\usepackage[libertine,vvarbb]{newtxmath}
```

ou :

```
\setmainfont{Linux Libertine O}
\usepackage{libertinustlmath}
```

3. La majuscule à Arabic est importante, car `arabic` est déjà défini pour autre chose dans L^AT_EX. Pour les autres langues, il n'y a pas de majuscule.

Pour Times :

```
\setmainfont{Times New Roman}
\usepackage[varvf,vvarbb]{newtxmath}
```

Pour STIX⁴ :

```
\usepackage{unicode-math}
\setmainfont{STIX}
```

Pour CMU Serif, il n’y a rien à faire, c’est du Computer Modern.

Annexe

A Quelques raccourcis TIPA

Ces raccourcis sont activés dans l’argument de la commande `\textipa`, après la déclaration `\tipaencoding` et dans l’environnement IPA, fournis par l’extension `tipa`.

```
\textipa{...}
{\tipaencoding ...}
\begin{IPA} ... \end{IPA}
```

Bien sûr, ces raccourcis sont surtout utiles si on ne dispose pas d’un moyen simple de saisir les caractères phonétiques directement en unicode au clavier.

TAB. 1: Raccourcis TIPA avec CMU Serif

ɐ	5	ɑ	A	ɒ	6	ʌ	2	ɓ	\!b	β	B ⁵
ç	C	ɔ	\!d	ɔ̃	\:d	ð	D	ə	@	ə	9
ε	E	ɜ	3	g	g	ɠ	\!G	ɣ	G	ɣ	7
ɦ	H	ɥ	4	ɨ	1	ɪ	I	ʝ	J	f	\!j
ɬ	\ ~l	ɭ	\:l	ɱ	M	ɰ	W	ɳ	N	ɳ	\:n
ʊ	\!o	ø	8	ø	\o	œ	\oe	œ	\OE	ɔ	0
ɸ	\th	ϕ	F	ɾ	R	ɽ	\:r	ɽ	*r	ɽ	\:R
ɸ	K	ɶ	\:s	ʃ	S	ʈ	\:t	ɸ	*t	θ	T
ɸ	θ	υ	U	υ	V	ɰ	*w	χ	X	ɸ	L
Y	Y	z	\:z	ʒ	Z	ʔ	P	ʔ	Q		
		'	"	,	""	:	:	'	;	˘	\t*{}

NB : les caractères API qui se présentent sous forme de petites capitales s’obtiennent, pour la plupart, avec le préfixe `\`; . Par exemple `\textipa{\ ;A\ ;B\ ;G\ ;H\ ;L}` donne `ABGHL`.

4. Sous réserve; je n’ai pas fait l’essai.

5. Apparemment ne marche pas avec Xe_{La}TeX (ça produit β); utiliser à la place `\textbeta` ou directement le caractère β (beta).

TAB. 2: Raccourcis TIPA avec Linux Libertine

\mathfrak{e}	5	\mathfrak{a}	A	\mathfrak{v}	6	\mathfrak{A}	2	\mathfrak{b}	\!b	$\mathfrak{\beta}$	B
\mathfrak{c}	C	\mathfrak{d}	\!d	\mathfrak{d}	\:d	\mathfrak{D}	D	\mathfrak{e}	@	\mathfrak{e}	9
\mathfrak{e}	E	\mathfrak{z}	3	\mathfrak{g}	g	\mathfrak{G}	\!G	\mathfrak{Y}	G	\mathfrak{Y}	7
\mathfrak{h}	H	\mathfrak{u}	4	\mathfrak{i}	l	\mathfrak{I}	I	\mathfrak{j}	J	\mathfrak{f}	\!j
\mathfrak{t}	\ ~l	\mathfrak{l}	\:l	\mathfrak{m}	M	\mathfrak{w}	W	\mathfrak{n}	N	\mathfrak{n}	\:n
\mathfrak{O}	\!o	\mathfrak{e}	8	\mathfrak{o}	\o	\mathfrak{oe}	\oe	\mathfrak{OE}	\OE	\mathfrak{o}	0
\mathfrak{p}	\th	\mathfrak{f}	F	\mathfrak{r}	R	\mathfrak{r}	\:r	\mathfrak{r}	*r	\mathfrak{r}	\:R
\mathfrak{k}	K	\mathfrak{s}	\:s	\mathfrak{S}	S	\mathfrak{t}	\:t	\mathfrak{t}	*t	\mathfrak{t}	T
\mathfrak{u}	0	\mathfrak{u}	U	\mathfrak{v}	V	\mathfrak{w}	*w	\mathfrak{x}	X	\mathfrak{x}	L
\mathfrak{Y}	Y	\mathfrak{z}	\:z	\mathfrak{Z}	Z	\mathfrak{P}	P	\mathfrak{Q}	Q	\mathfrak{P}	
$\mathfrak{ }$		$\mathfrak{'}$	"	$\mathfrak{,}$	""	$\mathfrak{:}$:	$\mathfrak{;}$;	$\mathfrak{\textasciitilde}}$	\t*{}

TAB. 3: Raccourcis TIPA avec Gentium

\mathfrak{e}	5	\mathfrak{a}	A	\mathfrak{v}	6	\mathfrak{A}	2	\mathfrak{b}	\!b	$\mathfrak{\beta}$	B
\mathfrak{c}	C	\mathfrak{d}	\!d	\mathfrak{d}	\:d	\mathfrak{D}	D	\mathfrak{e}	@	\mathfrak{e}	9
\mathfrak{e}	E	\mathfrak{z}	3	\mathfrak{g}	g	\mathfrak{G}	\!G	\mathfrak{Y}	G	\mathfrak{Y}	7
\mathfrak{h}	H	\mathfrak{u}	4	\mathfrak{i}	l	\mathfrak{I}	I	\mathfrak{j}	J	\mathfrak{f}	\!j
\mathfrak{t}	\ ~l	\mathfrak{l}	\:l	\mathfrak{m}	M	\mathfrak{w}	W	\mathfrak{n}	N	\mathfrak{n}	\:n
\mathfrak{O}	\!o	\mathfrak{e}	8	\mathfrak{o}	\o	\mathfrak{oe}	\oe	\mathfrak{OE}	\OE	\mathfrak{o}	0
\mathfrak{p}	\th	\mathfrak{f}	F	\mathfrak{r}	R	\mathfrak{r}	\:r	\mathfrak{r}	*r	\mathfrak{r}	\:R
\mathfrak{k}	K	\mathfrak{s}	\:s	\mathfrak{S}	S	\mathfrak{t}	\:t	\mathfrak{t}	*t	\mathfrak{t}	T
\mathfrak{u}	0	\mathfrak{u}	U	\mathfrak{v}	V	\mathfrak{w}	*w	\mathfrak{x}	X	\mathfrak{x}	L
\mathfrak{Y}	Y	\mathfrak{z}	\:z	\mathfrak{Z}	Z	\mathfrak{P}	P	\mathfrak{Q}	Q	\mathfrak{P}	
$\mathfrak{ }$		$\mathfrak{'}$	"	$\mathfrak{,}$	""	$\mathfrak{:}$:	$\mathfrak{;}$;	$\mathfrak{\textasciitilde}}$	\t*{}

TAB. 4: Raccourcis TIPA avec Times New Roman

\mathfrak{e}	5	\mathfrak{a}	A	\mathfrak{v}	6	\mathfrak{A}	2	\mathfrak{b}	\!b	$\mathfrak{\beta}$	B
\mathfrak{c}	C	\mathfrak{d}	\!d	\mathfrak{d}	\:d	\mathfrak{D}	D	\mathfrak{e}	@	\mathfrak{e}	9
\mathfrak{e}	E	\mathfrak{z}	3	\mathfrak{g}	g	\mathfrak{G}	\!G	\mathfrak{Y}	G	\mathfrak{Y}	7
\mathfrak{h}	H	\mathfrak{u}	4	\mathfrak{i}	l	\mathfrak{I}	I	\mathfrak{j}	J	\mathfrak{f}	\!j
\mathfrak{t}	\ ~l	\mathfrak{l}	\:l	\mathfrak{m}	M	\mathfrak{w}	W	\mathfrak{n}	N	\mathfrak{n}	\:n
\mathfrak{O}	\!o	\mathfrak{e}	8	\mathfrak{o}	\o	\mathfrak{oe}	\oe	\mathfrak{OE}	\OE	\mathfrak{o}	0
\mathfrak{p}	\th	\mathfrak{f}	F	\mathfrak{r}	R	\mathfrak{r}	\:r	\mathfrak{r}	*r	\mathfrak{r}	\:R
\mathfrak{k}	K	\mathfrak{s}	\:s	\mathfrak{S}	S	\mathfrak{t}	\:t	\mathfrak{t}	*t	\mathfrak{t}	T
\mathfrak{u}	0	\mathfrak{u}	U	\mathfrak{v}	V	\mathfrak{w}	*w	\mathfrak{x}	X	\mathfrak{x}	L
\mathfrak{Y}	Y	\mathfrak{z}	\:z	\mathfrak{Z}	Z	\mathfrak{P}	P	\mathfrak{Q}	Q	\mathfrak{P}	
$\mathfrak{ }$		$\mathfrak{'}$	"	$\mathfrak{,}$	""	$\mathfrak{:}$:	$\mathfrak{;}$;	$\mathfrak{\textasciitilde}}$	\t*{}