

Typesetting Guaraní with T_EX

Javier Bezos

2004/07/00

Spanish keywords: guaraní, ortografía, fuentes, tipografía, codificación.

1 Introduction

Guaraní is the main language spoken in Paraguay and one of the two official languages of the country (with Spanish). Here are minimal tools to typeset Guaraní with T_EX, which should not be considered as a finished job.

2 Input encoding

No current encoding supports the full set of characters used in Guaraní, not even Unicode (except if we recur to composite characters). Since there is a real necessity for a practical way to enter Guaraní, a new input encoding has been devised. There are a few fonts which replace several characters in the cp 1252 encoding by new ones. In particular, all vowels with circumflex are replaced by vowels with tilde. That means some glyph are repeated, but this way, since the circumflex is available in Paraguayan keyboards, they can be typed easily. In fact, this is a fairly frequent solution when the right glyphs are not available. There is an exception, yet—since cp 1252 does not have a \hat{y} , the slot used for \tilde{y} is that of \ddot{y} .

These fonts replace the section sign (§) by \tilde{g} , which one can live with, but a few map \tilde{G} to ASCII \$, which is definitely unacceptable. There are much better candidates, whose use is very restricted as they are mathematical symbols also available in the Symbol font: \div and \times . Furthermore, this is coherent with the surrounding characters and with the fact that the uppercase variant is placed at the same offset as the rest of uppercase letters.¹ Thus, with these fonts you will see the text as it would be typeset except for \tilde{g} —I hope the designers of these fonts will reconsider the current mapping.²

Accented letters (including \acute{y}) are part of cp 1252, and therefore nothing special has to be done here.

¹There were not the only candidates, yet. The islandic thorn is a very improbable letter in this context, for example, and a few fonts places \tilde{G} replacing the uppercase thorn. The ordinal letters does not fulfil these requirements and I don't want to discard two important Spanish signs.

²This applies to vowels with tilde, too. For consistency, I would rather use the umlaut, so that the key to press in order to get the tilde would be always the same, but again \ddot{u} is important in Spanish.

The new encoding is named `win-gn`, which maps the characters with diacritical marks in the following way (also uppercase):

â	ã	á	á	÷	Ğ
ê	ẽ	é	é	×	Ğ
î	ĩ	í	í	ñ	ñ
ô	õ	ó	ó		
û	ũ	ú	ú	'	'
ÿ	ÿ	Ý	Ý		

For example:

```
\documentclass{article}

\usepackage[win-gn]{inputenc}

\begin{document}

Mayma yvypóra ou ko yvy ári iñapytl'yre ha eteïcha dignidad
ha derecho jeguerekópe; ha ikatu rupi oikuaa añetéva ha
añete'yva, iporáva ha ivaíva, tekotevê pehenguéicha oiko
oñondivekuéra.

\end{document}
```

produces:

Mayma yvypóra ou kin-o yvy ári iñapytl'yre ha eteïcha dignidad ha dere-
cho jeguerekópe; ha ikatu rupi oikuaa añetéva ha añete'yva, iporáva ha
ivaíva, tekotevê pehenguéicha oiko oñondivekuéra.

(Ğğ)

If you insist in having ã in the slot reserved to the section sign, you can write `\guaranitildegassection`, but I don't like killing a useful typographical sign not available elsewhere. If you want to map them to thorn (*both* upper and lowercase, for consistency) write `\guaranitildegasthorn` (you can combine the section and the thorn options). The uppercase form Ñ, however, will *never* be mapped to \$.

3 Babel style

A style file for Guaraní compatible with babel is provided.

It provides translations for dates [*in fact it doesn't*] but not for names, since I was unable to find reliable translations.

It provides shorthands using " for the acute accent (unfortunately, ' cannot be used because it's considered a letter) and ^ for the tilde. Note this is fine if your keyboard combines the tilde with the following vowel, provided the encoding above is used—thus, you can always use the hat as a synonymous of ~, either if combined or not.

The ' is known as *puso* and is much like the arabic ayin, in fact with a similar sound. The *puso* has been made active to adjust the spacing; otherwise it could break the visual continuity of words. The command `\setguaranipuso` allows to adjust the spacing to adapt it to the current font. It has three arguments: space to be removed before, lowering, space to be removed after. You can deactivate this feature by leaving the three arguments empty.

4 Hyphenation

Very often, a mixture of Guaraní and Spanish, known as Jopara, is spoken; therefore, these patterns are a mixture of Guaraní (mainly *mb*, *nd*, *ng*, *nt* as special cases) and Spanish. In addition, ' is a letter but note if you are using the Guaraní style then that is irrelevant, since the adjustments could kill hyphenation.