

# Corriger des fichiers T<sub>E</sub>X contenant des caractères huit bits depuis la ligne de commande à l'aide de Francais-GUTenberg

Christophe Pythoud

26 juin 1998

## Table des matières

<b>1</b>	<b>Le problème</b>	<b>1</b>
<b>2</b>	<b>La solution</b>	<b>2</b>

## 1 Le problème

Lorsque ISPELL est invoqué depuis la ligne de commande, comme dans :

```
ispell -d francais document.tex
```

le programme passe en mode T<sub>E</sub>X, à cause de l'extension `.tex` du fichier<sup>1</sup>. Cela a deux conséquences :

- tout mot précédé d'une barre oblique inversée est considéré comme une commande T<sub>E</sub>X et n'est donc pas soumis au correcteur ; d'autre part, ISPELL reconnaît certaines commandes de mise en page et n'en analyse pas les arguments ;
- les caractères accentués « à la T<sub>E</sub>X » sont reconnus (`\'e` à la place de `é`).

La première propriété est indispensable. La seconde peut rendre service, mais elle présente un caractère exclusif ennuyeux : toutes les équivalences sept bits définies dans `francais.aff` (vers la ligne 55) sont reconnues, par contre les caractères huit bits ne le sont plus. Il devient donc impossible de corriger les mots contenant des caractères huit bits. (Le problème ne se pose pas si ISPELL est invoqué depuis (X)Emacs.)

---

1. On peut également forcer ISPELL à passer en mode T<sub>E</sub>X en lui donnant l'option `-t` ou `-T TeX`.

## 2 La solution

On peut envisager deux solutions peu satisfaisantes :

- débrayer le mode `TeX` pour effectuer la correction de ce type de fichier<sup>2</sup> : le problème est alors que les commandes `TeX` ne sont plus reconnues comme telles et vérifiées (bref, `ISPELL` devient inutilisable) ;
- convertir le texte en sept bits avant d'en vérifier l'orthographe (et éventuellement le reconvertir en huit bits après) : c'est envisageable pour un long texte, mais très pénible pour une série de petits fichiers ; d'autre part, il est nettement plus agréable de voir s'afficher des caractères accentués à l'écran plutôt que les codes `TeX`.

Heureusement, le problème peut être réglé dans `francais.aff` en supprimant purement et simplement la partie définissant les équivalences `TeX`. C'est ce qui est fait dans le fichier `francais-TeX8b.aff`. Comment procéder ? Il est possible de créer un dictionnaire `francais-TeX8b.hash` en répondant oui à la question suivante (posée par `makehash.pl`) :

Souhaitez-vous également produire un dictionnaire permettant de corriger des fichiers TeX huit bits depuis la ligne de commande (pas nécessaire si vous utilisez EMACS -- voir docs/tex8b.pdf pour plus de détails) (o/n) ?

Copiez ensuite `francais-TeX8b.aff` et `francais-TeX8b.hash` à l'endroit où `ISPELL` va chercher ses dictionnaires et vous pourrez dès lors corriger les fichiers `TeX` en huit bits depuis la ligne de commande :

```
ispell -d francais-TeX8b document.tex
```

et pour les documents en sept bits :

```
ispell -d francais document.tex
```

---

2. `ispell -n -d francais document.tex`