

Exercice de balisage, solutions

1. *Conversion de la balise locuteur au format lexico :*

Rechercher une suite de trois « &&& » suivie d'un passage à la ligne et remplacer par une balise fermante suivie d'un passage à la ligne. (>^p). Soit : Rechercher &&&^p et remplacer par >^p

Dans ce cas de figure il n'est pas nécessaire d'utiliser les caractères génériques.

Puis rechercher le motif &&& et le remplacer par <locuteur=

2. *Compter les balises locuteur :*

Une astuce consiste à saisir le motif <locuteur= dans la fenêtre de recherche et à le copier dans la fenêtre de remplacement. Appeler ensuite la commande « remplacer tout ».

3. *Insertion de la balise date :*

Chaque intervention est notée ainsi : source, entretien du jj/mm/aaaa

On remplacera par exemple le motif « entretien du » par la suite de caractères « date=

Plus concrètement on recherche : espace, entretien, espace, du, espace. Le plus simple est de copier l'expression dans le texte et de la coller dans la fenêtre de recherche. On remplace par l'expression suivante : ^p<date=(C'est-à-dire l'expression <date= précédée d'un passage à la ligne).

Il reste encore à insérer la balise fermante. Sachant que tous les textes sont notés 2002, on peut rechercher le motif /2002^p et le remplacer par /2002>^p (Remplacer tout). 27 remplacements doivent être effectués.

4. *Modifier la mise en forme des balises locuteur et date*

4.1. *Balise locuteur :*

Il s'agit de modifier la balise locuteur afin qu'elle apparaisse en caractères non gras. On recherche l'expression suivante (en utilisant les caractères génériques). \<locuteur= ?*\> L'anti slash est employé ici pour déspecialiser les caractères <et>. L'expression « ?* » signifie que la recherche porte sur un ou plusieurs caractères quelconques. Cliquer ensuite dans la fenêtre de remplacement puis dans la barre d'outils de Word désactiver l'option gras (G).

4.2 *Balise date :*

On choisira ici une variante de l'exemple qui précède au moyen de l'expression suivante :

```
\<date=[0-9]{2;}/[0-9]{2;}/[0-9]{4;}\>
```

5. *Suppression des commentaires et méta info*

Ces indications étant notées en gras dans le texte, il suffit de rechercher tous les caractères en gras : ?*, attribut Gras et de remplacer par un ensemble vide. Cliquer dans la fenêtre de remplacement puis sur la commande « sans attribut ».

6. *Suppression des interventions des journalistes :*

Même formule que précédemment en choisissant l'attribut « italique ».

Le document ainsi balisé et nettoyé peut désormais être traité sous lexico, après avoir été sauvegardé en texte seul.