Procédure de transformation

# Préparation

Ouvrir le fichier rtf dans Libreoffice Writer, et enregistrer au fomat odf

Installer le plugin de remplacement avancé dans Writer (permet de remplacer un saut de paragraphe suivi d’un caractère) : <http://extensions.libreoffice.org/extension-center/alternative-dialog-find-replace-for-writer>

# Traitement dans Writer

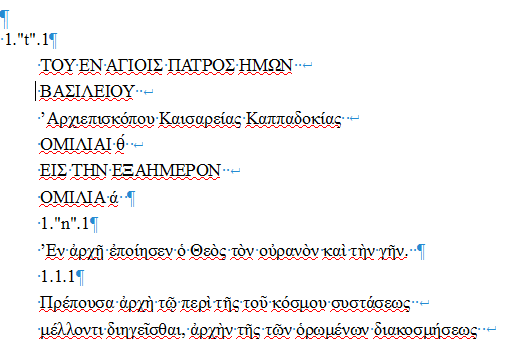
**Le fichier a 2 types de contenu :**

* numérotation -> style snsg\_cita
* texte en grec -> style snsg\_greek

à la fin de chaque ligne de texte grec : 2 espaces, et un caractère retour chariot (Shift+Entrée)

à la fin d’un bloc de texte grec, ou d’une ligne de numérotation: parfois 2 espaces, et un saut de paragraphe (Entrée)

au début de chaque ligne, un espace



**Structuration :**

* 1er niveau : identification de chaque Homélie

Structure (avec « X » le numéro de l’Homélie) :

*X."t".1*

*Numéro de l’Homélie (en grec). Ex :* ΟΜΙΛΙΑ δ́.

*X."n".1*

*Titre de l’Homélie (en grec) Ex :*  Περὶ συναγωγῆς τῶν ὑδάτων

*X.1.1*

* 2e niveau : section ( ?), ne fait pas l’objet d’une numérotation explicite, mais présent implicitement dans le 3e niveau
* 3e niveau : lignes (Homélie. Section. Ligne)

Structure (avec X le numéro de l’Homélie, Y le numéro de section, L la ligne)

*X.Y.L*

**Choix faits :**

* On peut concaténer le numéro de l’homélie et son titre. Ex : ΟΜΙΛΙΑ δ́. Περὶ συναγωγῆς τῶν ὑδάτων
* Il est sans doute inutile de conserver les numéros de ligne (si on voulait le faire, il faudrait utiliser des balises milestone ou pb)

**Règles de transformation :**

* Supprimer tous les espaces en début de ligne

Remplacer (avec option expression régulières) "^ " par ""

Remplacer (avec option expression régulières) "(\n) " par "$1"

* Supprimer les doubles espaces

Remplacer (avec option expression régulières) " \*" par " "

* Passer la numérotation de 1er niveau en style Titre 1
* Supprimer les coupures de mots du texte grec

Rechercher et remplacer : "- (\n)" par "$1"

Rechercher et remplacer : "- $" par ""

* Générer une numérotation de 2e niveau, en style Titre 2
  + Si une numérotation de 3e niveau est du type X.Y.1 (1re ligne de la section Y de l’homélie X), la transformer en X.Y

Rechercher (tout chercher avec option expression régulières) "^([0-9]\*\.[0-9]\*)\.1$"

Appliquer le style « Titre 2 » à tous les passages sélectionnés

Rechercher et remplacer : "^([0-9]\*\.[0-9]\*)\.1$" par "$1"

* Supprimer la numérotation de 3e niveau

Rechercher et remplacer : "^[0-9]\*\.[0-9]\*\.[0-9]\*$" par ""

* Supprimer les sauts de ligne du texte grec

Rechercher et remplacer : "\n" par ""

* Supprimer les lignes vides

Rechercher et remplacer : "^$" par ""

**Regrouper les éléments de la structure de 1er niveau (titre d’Homélie) :**

Avec l’éditeur spécial :

Rechercher et remplacer : "\p([0-9]\*)\."n"\.[0-9]\*\p" par "."

Rechercher et remplacer : "([0-9]\*\."t"\.[0-9]\*)\p" par "\1"

Avec l’éditeur normal :

Sélectionner tous les titres d’Homélie

Rechercher tout "^[0-9]\*\."t".\*$". Changer le style de la sélection en Titre 1

Remplacer la numérotation des Homélies

Rechercher et remplacer  "([0-9]\*).\"t"\.[0-9]\*" par "@n=$1@@type=homily@"

Rechercher et remplacer  "^[0-9]\*\.([0-9]\*)$" par "@n=$1@type=section@"

Avec l’éditeur spécial

Rechercher et remplacer  " \p[^@]" par " "

Supprimer tous les sauts de ligne au milieu du texte grec :

Ne garder que les sauts de ligne précédés ou suivis d’une arrobase

Avec l’éditeur spécial :

Rechercher et remplacer : "([^@])\p([^@])" par \1\2

Isoler le titre de l’œuvre (actuellement inclus dans le 1."t".1 avant ΟΜΙΛΙΑ ά

# Transformation du fichier odt

Envoi de l’ODT dans Odette <http://bookmeka.huma-num.fr/Odette>

Post-traitement à faire pour interpréter @n=$1@type=section@