

Tekstaro de Esperanto

Utilisation élémentaire du *serĉilo*

Introduction

Le *Tekstaro de Esperanto* est un site qui permet d'effectuer des recherches textuelles dans un corpus de textes classiques. Plus de 10 millions de mots y sont indexés.

Le champ de recherche (bouton **Serĉi**) contient évidemment le mot ou l'expression que vous cherchez, ce qu'on appelle le **motif**, mais aussi des **symboles spéciaux** qui modifient votre recherche.

La liste de ces symboles est longue et leurs emplois combinés peuvent s'avérer délicats. Ce document propose une utilisation simplifiée de l'outil.

Principaux symboles - exemples de recherche

Parmi tous les symboles, les plus utiles sont les **symboles grammaticaux**.

Mais avant tout, il est prudent de commencer le motif par le symbole `\b` qui signifie que l'on cherche un motif *séparé*. Par exemple, dans la phrase « Parmi tous les symboles, voici les plus utiles. », si mon motif de recherche est "les" je trouverais "symboles", "les" et "utiles" car tous ces résultats contiennent le motif. Mais si je désire trouver tous les mots isolés commençant par "les", je dois préfixer mon motif "`\bles`". Je ne trouverais alors que "les".

Les symboles grammaticaux

On peut affiner la recherche en précisant certaines fonctions grammaticales :

- **\FI** remplace n'importe quelle **finale**.
Par exemple le motif `\binteres\FI` trouvera *intereso*, *intereson*, *interesojn*, *interesa*, *interesaj*, *interese*, *interesi* et *interesas*, mais pas les mots composés : *seninteresa* et *interesigi*.
- Les finales -i, -o, -a et -e :
 - **\VF** remplace n'importe quelle **finale verbale**.
Par exemple le motif `\binteres\FV` trouvera *interesi*, *interesas*, *interesis*, *interesos*, *interesus* et *interesu*.
 - **\OF** remplace n'importe quelle **finale substantive**.
Par exemple le motif `\binteres\OF` trouvera *intereso*, *intereson*, *interesoj* et *interesojn*.
 - **\AF** remplace n'importe quelle **finale adjectivale**.
Par exemple le motif `\binteres\AF` trouvera *interesa*, *interesan*, *interesaj* et *interesajn*.
 - **\EF** remplace n'importe quelle **finale adverbiale**.
Par exemple le motif `\binteres\EF` trouvera *dekstre* et *dekstren*.
- **\PA** remplace n'importe quel **participe**.
Par exemple le motif `\bnask\PAa\b` (qui signifie le motif isolé avant et après *nask*, suivi du symbole

participe, suivi de la lettre a) trouvera naskanta, naskinta, naskonta, naskata, naskita kaj naskota.

- Enfin les symboles **\AP** et **\PP** à la place de **\PA**, trouveront respectivement **les participes actifs et passifs**.

La recherche

Une fois le champ **Serçi** rempli, il faut choisir l'ensemble des œuvres dans lesquelles on recherchera le motif. Par exemple, on peut choisir **Ĉiuĵ**. On lance la recherche en cliquant sur le bouton **Ekserçi**.

Trois exemples

Premier exemple

On recherche toutes les occurrences de “La nova vesto”, au singulier ou au pluriel et éventuellement à l'accusatif. On entre le motif de recherche suivant :

```
\b[Ll]a nov\AF vest\OF
```

- la balise **\b** marque une limite de mot ;
- une paire de crochets représente une occurrence quelconque des caractères qu'elle contient. Par exemple [aeiouy] représente une voyelle, Duran[dt] désigne *Durand* ou *Durant*. Ici [Ll] représente *L* ou *l* ;
- **\AF** indique toutes les variantes des finales adjectives : *a*, *an*, *aj* ou *ajn* ;
- **\OF** indique toutes les variantes des finales substantives : *o*, *on*, *oj* ou *ojn*.

Remarques :

- la balise **\b** commence pratiquement toutes les recherches ;
- la balise **\b** est automatique ajoutée. Elle n'est pas nécessaire après **\AF** et **\OF**.

La recherche trouve 20 occurrences de notre motif.

Deuxième exemple

On recherche toutes les occurrences de “mia patrino,”. On entre le motif simple :

```
\bmia patrino,
```

On obtient 53 occurrences.

Troisième exemple

On recherche toutes les occurrences de “mia patrino,” mais contenant zéro à trois mots entre “mia” et “patrino”. Alors le motif devient :

```
\bmi\AF ( \EL+ ) {0,3}patrin\OF
```

- les symboles spéciaux **\AF** et **\OF** sont déjà expliqués ;
- l'expression (\EL+) {0,3} est un sous-motif qui indique une ou plusieurs lettres en espéranto (c'est le \EL+) suivi d'un blanc. Cette sous-expression peut apparaître zéro à trois fois (ce qu'indiquent les accolades).

On obtient 67 occurrences.

Bonnes recherches !