

Frantext

Premiers pas dans la nouvelle interface

Diapo 36

Maintenant que nous disposons d'un corpus, nous allons effectuer une première recherche à l'intérieur.

Pour ça, on clique sur la mention **Rechercher**, en haut à gauche, à côté de **Corpus**.

Diapo 37

On arrive sur une page où plusieurs modes de recherche sont proposés, répartis par onglets : **Simple**, **Assistée**, **Avancée**, **Fréquences**, **Co-occurrences**, **Voisinage**. On a ensuite un onglet pour accéder aux recherches qu'on a précédemment enregistrées et un dernier pour accéder à la documentation, sur le Wiki Frantext.

Ici, nous allons nous contenter de nous intéresser à la recherche simple. Pour les choses plus complexes, je vous invite à consulter le Wiki.

On tape donc le contenu de notre recherche simple dans le champ de texte prévu à cet effet. Je veux observer les occurrences de *porte* dans mon corpus **Aymé**. J'écris donc **porte** et comme je ne veux pas de *porté* dans les résultats, je coche la case **Sensible aux diacritiques**.

Je clique ensuite sur le bouton **RECHERCHER**.

Diapo 38

Les résultats s'affichent dans la partie principale de l'écran, conformément à mes paramètres utilisateurs. On voit que l'on a récolté **456** occurrences de la forme *porte*.

Dans cette zone, on a la possibilité de changer les propriétés de données affichées et la taille du contexte, à l'aide des menus déroulants à gauche au-dessus des concordances.

On peut accéder aux métadonnées d'un texte en cliquant sur sa cote à gauche de chaque ligne de concordance.

On peut zoomer sur un résultat pour en élargir le contexte en cliquant sur la petite loupe à droite d'une ligne ou même décider de supprimer un résultat qui ne nous semble pas pertinent à l'aide de la croix située à droite de la loupe.

Je vous laisse explorer les autres possibilités par vous-mêmes.

Sur la gauche de l'écran, là où l'on avait tout à l'heure les métadonnées du corpus, on accède à présent à un ensemble d'**Actions**. Je ne vais pas présenter le détail de ces actions aujourd'hui, mais vous voyez rapidement qu'elles permettent de réorganiser les résultats (tris, mélanges), de les affiner (exclure des documents, échantillonner) ou de changer le point de vue qu'on porte sur eux (fréquences, distribution, co-occurrence).

Les métadonnées du corpus de travail restent accessibles en cliquant sur **Métadonnées**, juste à côté de **Actions**.

Diapo 39

Pour cette introduction, on va simplement s'intéresser à deux actions, qui vont nous permettre de connaître la répartition de la forme recherchée, que j'appellerai désormais pivot. Ces actions, ce sont la **Fréquence** et la **Distribution**.

Diapo 40

Commençons par la fréquence. J'y accède en cliquant sur **Fréquence** dans le menu de gauche et cela déclenche l'apparition de cette fenêtre pop-up.

On peut se servir de cette action pour observer les fréquences des contextes gauches et droit de la concordance ou du pivot lui-même. On a accès aux différentes propriétés des mots du corpus, que j'ai déjà évoqué : Forme, Catégorie grammaticale et Lemme. Il est possible de combiner des **Niveaux** pour interroger par exemple plusieurs mots différents dans les contextes en choisissant des **Offset** différents. Ces termes de niveaux et d'offset ne sont peut-être pas les termes idéaux et si vous avez des suggestions pour améliorer certains intitulés de l'interface, ici ou ailleurs, n'hésitez pas à les suggérer par l'intermédiaire des adresses mails que je vous ai fournies.

Dans le cas présent, je m'intéresse à la répartition par catégorie grammaticale de la forme *porte*. Je choisis donc d'interroger un seul niveau : je choisis la **Position Pivot**, et l'**Ensemble de données Catégorie**. J'accède à la liste des réponses possibles pour chaque critère à l'aide du petit symbole de flèche à droite, comme on a vu précédemment.

Une fois mes critères sélectionnés, je lance le calcul des résultats à l'aide du bouton **CALCULER**.

Diapo 41

Dans les résultats qui s'affichent, je peux voir que 413 occurrences de *porte* ont été étiquetées comme étant des noms communs et 43 comme étant des verbes.

La liste des catégories grammaticales utilisées et de leur abréviation est disponible sur le Wiki d'aide.

Diapo 42

Si je veux regarder en détail quelles sont les 43 occurrences étiquetées verbe, je clique tout simplement sur la deuxième ligne...

Diapo 43

... et les résultats s'affichent.

J'attire ici votre attention sur le fait que l'on a désormais au-dessus des concordances une ligne composée d'une cellule bleue et de deux cellules orange. Ce sont des boutons, qui permettent de revoir quelles actions ont été réalisées, dans quel ordre et de revenir en arrière si on le souhaite. Elles se lisent de gauche à droite.

La cellule bleue correspond à notre recherche initiale, qui est également visible, bien que grisée, dans le champ recherche. Elle est écrite ici avec la syntaxe de recherche avancée, que vous apprendrez peut-être à maîtriser au fur et à mesure de vos utilisations de Frantext.

Ici : **&développer("porte"%cl).**

Les cellules orange correspondent aux actions supplémentaires. Si vous cliquez sur le petit crayon à droite à l'intérieur d'une cellule action, vous pouvez retrouver et éditer tous vos critères.

Si vous cliquez sur une cellule d'action autre que la dernière, vous annulez toutes les actions suivantes.

En cas de perte involontaire d'actions, utilisez le bouton « page précédente » de votre navigateur Web, vous n'avez pas tout perdu ! Ceci est valable pour toutes les situations où vous voulez faire un pas en arrière, y compris quand vous devez vous reconnecter parce que vous êtes restés trop longtemps inactifs.

Vous constaterez également qu'une partie **Actions en cours** est apparue dans le menu à gauche de votre écran, sous la liste des actions possibles. Il s'agit simplement des mêmes informations que celles des cellules, présentées d'une autre façon.

Diapo 44

Intéressons-nous à présent à l'action **Distribution**. Quand on clique dessus, une fenêtre s'affiche qui nous permet de choisir selon quelle métadonnée on souhaite observer la répartition. Je choisis la métadonnée **Date** et un champ apparaît me permettant de choisir la taille des périodes en années. Je choisis **10** et je lance le calcul.

Diapo 45

La répartition apparaît.

Diapo 46

J'ai la possibilité de trier cette répartition en cliquant sur l'intitulé de la colonne **Valeur**.

Ici je choisis de trier par ordre descendant des valeurs de la métadonnée que j'ai sélectionnée.

Diapo 47

Pour observer en détail les occurrences correspondant à une période précise, la méthode est la même que tout à l'heure : je clique sur la ligne qui m'intéresse, ici 1940-1949,...

Diapo 48

... et les résultats s'affichent.

Diapo 49

Admettons que je veuille à présent sauvegarder mes résultats. Je vais pouvoir le faire en cliquant sur **SAUVEGARDER**, à droite au-dessus de la concordance.

Attention, il faut bien avoir à l'esprit que ce que l'on sauvegarde, ce n'est pas la concordance elle-même, mais la succession recherche initiale + actions qui nous ont permis d'accéder à cette concordance.

Cela signifie que si on ne travaille pas sur un corpus personnel, mais sur Frantext intégral ou l'un des corpus prédéfinis, nos résultats ne seront plus identiques si des textes ont été ajoutés à la ressource. Ces ajouts, qui arrivaient environ une fois par an par le passé pourront désormais arriver plus fréquemment, parce que la procédure de mise à jour des données Frantext a été simplifiée avec le changement de moteur d'indexation de la base.

Cela signifie également que lorsqu'on va sauvegarder notre travail, il faudra bien noter sur quel corpus on a effectué cette recherche. Notamment si on a utilisé la possibilité de supprimer des résultats qui ne nous semblaient pas pertinents.

Diapo 50

En cliquant sur **SAUVEGARDER**, on fait apparaître une nouvelle fenêtre.

On peut alors donner un nom à notre sauvegarde et en faire une courte description. Je vous invite à y noter sur quel corpus vous avez travaillé.

Pour valider la sauvegarde, on clique sur **CRÉER**.

Diapo 51

Pour accéder aux recherches qu'on a sauvegardées, on se rend dans l'onglet **Mes recherches**, entre **Voisinage** et **Documentation**.

Diapo 52

On y retrouve toutes les recherches qu'on a sauvegardées, avec un code couleur qui nous permet de voir rapidement s'il s'agit de concordances, de fréquences ou de co-occurrences.

Pour reprendre une recherche, on clique sur le bouton **CHARGER**.

Si on a modifié son corpus de travail courant entre-temps, on n'oublie pas de remettre le bon corpus, à l'aide du menu déroulant en haut à droite que l'on vu au tout début de l'atelier.

Diapo 53

Cette histoire d'indépendance entre recherche et corpus présente un intérêt non négligeable pour la comparaison d'époques, d'auteurs, de genres ou de n'importe quoi que vous aurez décidé de ranger dans des corpus différents.

Si on reprend la succession d'actions précédentes, qui nous a amenés à observer les occurrences de *porte* dans les romans et nouvelles de Marcel Aymé écrits dans les années 40. Il nous suffit de basculer vers le corpus **Moderne** pour étendre notre recherche à l'ensemble des textes répertoriés comme tels dans la base.

Diapo 52

Vous constatez ici que le contenu grisé du champ recherche et la succession d'actions n'ont pas changé. En revanche, on a désormais 5801 résultats parmi 630002 occurrences de *porte*.

Diapo 55-56

À vous

Diapo 57

Passons maintenant à l'export de résultats. Pour l'instant, cette fonctionnalité n'est disponible que pour les concordances.

À partir d'une concordance, vous devez simplement cliquer sur **Exporter**, à droite en dessous de **SAUVEGARDER**.

Diapo 58

Une fenêtre apparaît, qui contient un formulaire, dans lequel vous allez pouvoir paramétrer votre export.

Vous choisirez dans un premier temps le format (csv, xml ou texte brut).

Ensuite, vous êtes invités à donner des informations sur les résultats eux-mêmes. Si vous ne souhaitez pas l'ensemble des résultats, vous pouvez exporter seulement une sous-partie. Dans ce cas, vous aurez sans doute préalablement trié vos résultats ou, au contraire, les aurez mélangés. Vous pouvez exporter uniquement les formes, leur catégorie grammaticale, leur lemme ou un panaché des trois. Enfin, vous choisissez la taille du contexte.

Dans un dernier temps, vous décidez des métadonnées du texte et de l'extrait de texte que vous souhaitez voir apparaître.

Diapo 59

Pour ma part, j'ai choisi un format csv, que je vais pouvoir exploiter sous Excel. Dans ce cas, chaque métadonnée incluse va donner lieu à la création d'une nouvelle colonne.

Mon pivot étant composé d'une seule forme, l'option **séparer les mots de la cible**, qui a pour but de segmenter en mots les pivots composés n'est pas pertinente.

Je choisis d'inclure une colonne avec les numéros de résultats.

Une fois tout cela paramétré, je clique sur **EXPORTER**.

Diapo 60

J'obtiens le fichier CSV attendu, que j'importe dans Excel.

Attention : pour l'instant, les colonnes ne disposent pas d'intitulés. Je vous conseille donc de noter l'ordre des métadonnées que vous avez demandé, pour vous y retrouver.

Diapo 61

J'avais prévu cet exemple à la fin, pour vous montrer ce qui s'affiche quand on clique sur les crayons dans chaque cellule orange.