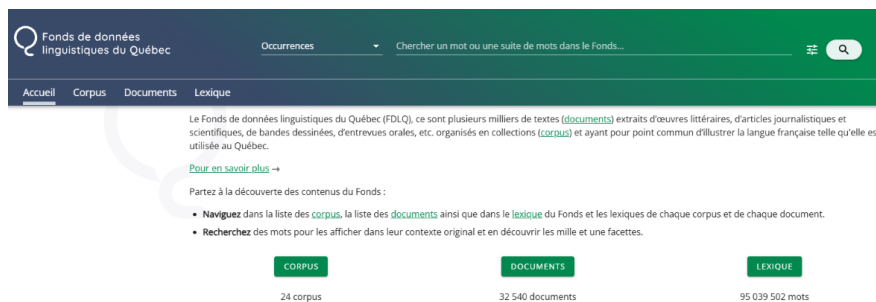


Aide en ligne du FDLQ

Bienvenue dans l'aide en ligne du Fonds de données linguistiques du Québec! Le FDLQ est une plateforme numérique qui donne accès à des documents qui rendent compte du français tel qu'il est parlé et écrit au Québec. La plateforme permet aussi d'interroger ces documents.

[Comment accéder](#) au FDLQ.

Cliquez sur les éléments de la capture d'écran qui s'affiche ci-dessous pour accéder aux rubriques de l'aide en ligne. Il existe également une [version imprimable](#) du guide.



Qui êtes-vous?

Vous faites de la recherche?



- Vous avez à votre disposition un [moteur de recherche](#) qui permet de chercher des **occurrences**¹ de mots-formes ou de suites de mots ainsi que des **cooccurrences**². Il est également possible de lancer une recherche à partir d'un **lemme**³.
- Le moteur de recherche prévoit plusieurs [options de recherche](#) et per-

¹Apparition d'une unité linguistique dans le discours; par métonymie, cette unité (définition tirée d'Usito).

²Ensemble d'unités lexicales qui apparaissent fréquemment en combinaison dans un énoncé, et qui sont liées par un rapport de proximité syntaxique et de relative dépendance (définition tirée d'Usito).

³Forme canonique d'un mot variable (au masculin singulier pour les noms et les adjectifs, à l'infinitif pour les verbes) (définition tirée d'Usito).



met entre autres d'utiliser des jokers ou des expressions régulières pour formuler des requêtes plus fines.

- Vous souhaitez analyser des contextes élargis? Consultez la section consacrée aux [options d'affichage](#) pour en savoir davantage sur l'affichage d'un [passage élargi](#).
- Dans certains cas, il est possible de consulter le [texte intégral](#) d'un document, voire le [document original](#) sous forme d'image ou d'enregistrement audio.

Vous enseignez le français?



- Nous vous invitons à prendre connaissance du fonctionnement du [moteur de recherche](#) pour repérer des exemples de mots ou de constructions que vous souhaitez intégrer dans vos exercices faits en classe. Vous pouvez [télécharger les résultats](#) de vos recherches textuelles dans un fichier *Excel*.
- Si vous enseignez le français langue seconde et que vous cherchez des données authentiques de l'oral, n'oubliez pas de consulter la rubrique consacrée aux documents qui disposent d'une [transcription avec audio](#). Plusieurs enregistrements audio versés au Fonds peuvent être écoutés dans leur intégralité.
- Les interfaces du Fonds et du [dictionnaire Usito](#) communiquent entre elles! La consultation des deux outils permet de réfléchir aux liens entre l'usage de la langue et la description qu'on en fait dans un dictionnaire.
- Vous avez des questions sur la terminologie utilisée sur la plateforme? Le [glossaire](#) comprend certaines explications qui pourront vous être utiles.

Vous vous intéressez au français québécois et avez envie d'explorer le Fonds?



- Parcourez le [lexique du Fonds](#) ou le [lexique d'un corpus](#) pour découvrir les nombreux mots qui sont attestés dans les différents documents réunis.
- Pour vous familiariser avec les contenus versés au Fonds, consultez la rubrique qui explique les principaux éléments de la plateforme, comme les [corpus](#) et les [documents](#).
- Saviez-vous qu'il est possible d'écouter certains enregistrements audio directement sur la plateforme? Notre [lecteur audio](#) permet en



même temps de suivre la transcription de ces documents.

Voir aussi

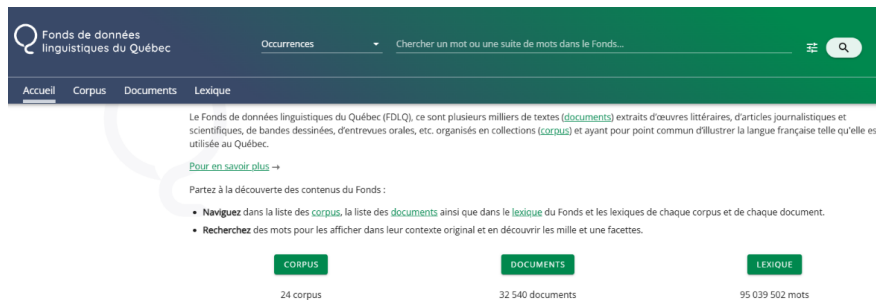
- [Corpus](#)
- [Documents](#)
- [Lexique](#)
- [Moteur de recherche](#)

Accéder au Fonds de données linguistiques du Québec

1. À partir du site du [Fonds de données linguistiques du Québec](#), cliquez sur **Accéder au Fonds**.



2. Un nouvel onglet s'affiche dans le navigateur Web, vous pouvez désormais interroger la base de données!



Voir aussi

- [Accueil](#)

Dernière modification : 21 mai, 2025

Moteur de recherche

Le moteur de recherche est un aspect central du Fonds de données linguistiques du Québec. Il permet de chercher un **mot**¹ ou une suite de mots dans les **corpus**² versés au Fonds. On peut effectuer divers types de recherche et ajuster les paramètres de recherche.

Sommaire

- [Type de recherche](#)
 - [Occurrences](#)
 - [Cooccurrences](#)
 - [Lemmes](#)
- [Champ de recherche](#)
- [Options de recherche](#)

Type de recherche

Le moteur de recherche permet d'effectuer trois types de recherche : il est possible de chercher des **occurrences**³, des **cooccurrences**⁴ et des **lemmes**⁵. Pour changer le type de recherche, cliquez sur la petite flèche qui se trouve à droite du champ et sélectionnez la valeur de votre choix.

¹Dans le contexte de la plateforme FDLQ, mot n'est pas utilisé dans le sens d'« unité lexicale » qu'il prend en linguistique, mais plutôt comme synonyme de mot-forme (mot séparé par des espaces) ou encore de token. Les mots faisant partie du lexique d'un document ou d'un corpus correspondent ainsi à l'ensemble des tokens qu'on y trouve; suivant cette logique, pomme de terre correspond à trois tokens.

²Ensemble de textes établi selon un principe de documentation exhaustive, un critère thématique ou exemplaire en vue de leur étude linguistique (définition tirée d'Usito).

³Apparition d'une unité linguistique dans le discours; par métonymie, cette unité (définition tirée d'Usito).

⁴Ensemble d'unités lexicales qui apparaissent fréquemment en combinaison dans un énoncé, et qui sont liées par un rapport de proximité syntaxique et de relative dépendance (définition tirée d'Usito).

⁵Forme canonique d'un mot variable (au masculin singulier pour les noms et les adjectifs, à l'infinitif pour les verbes) (définition tirée d'Usito).

Occurrences

Une occurrence est l'attestation d'une unité linguistique dans un corpus (définition tirée de [Usito](#)).

C'est le type de recherche qui est sélectionné par défaut, il permet d'effectuer une recherche sur toutes les occurrences du mot (ou de la suite de mots) saisi dans le [champ de recherche](#).

Cooccurrences

Une cooccurrence est l'ensemble des unités lexicales qui apparaissent fréquemment en combinaison dans un énoncé, et qui sont liées par un rapport de proximité syntaxique et de relative dépendance (définition tirée de [Usito](#)).

Le moteur de recherche permet de trouver des mots qui ne sont pas utilisés les uns après les autres, mais qui sont utilisés dans un voisinage donné.

Comment faire une recherche par cooccurrences

1. Saisissez un terme dans le premier champ de recherche. Il est à noter qu'on peut saisir une suite de mots dans le champ de recherche.
2. Indiquez le sens de la recherche. Par défaut, la flèche bidirectionnelle est choisie (la flèche bidirectionnelle indique que les deux termes peuvent apparaître dans n'importe quel ordre).
3. Saisissez le second terme à rechercher.
4. Indiquez la distance maximale entre les mots. Par défaut, la distance est réglée à 10 mots, mais on peut régler la distance à 40 mots au maximum.
5. Pressez sur la touche **Entrée** du clavier ou cliquez sur l'icône **Rechercher**.



Dans l'exemple ci-dessus, puisque la flèche bidirectionnelle est choisie, « panier » et « fruits » peuvent arriver en premier dans la séquence. Il peut y avoir au maximum 10

tokens¹, y compris la ponctuation, entre les termes de recherche. Voici des exemples de résultats :

Contexte avant	Pivot	Contexte après
un énorme	panier de fruits	avec des tas de petites
Elle offrit les	fruits de son panier	, forçant les religieuses à tendre la main pour les prendre, puis elle regagna sa place.
Tu peux bien parler, Jacquou, toi qui nous refiles des	fruits blets sous le dessus du panier	!
Un	panier de fruits	
Se ferait	panier plein de fruits	vermeils

- La flèche vers la gauche indique que le terme saisi à droite doit apparaître en premier dans les résultats;
- La flèche bidirectionnelle indique que les deux termes peuvent apparaître dans n'importe quel ordre;
- La flèche vers la droite indique que le terme saisi à gauche doit apparaître en premier dans les résultats.

Lemmes

Un lemme est la forme canonique d'un mot variable (au masculin singulier pour les noms et les adjectifs, à l'infinifit pour les verbes) (définition tirée de [Usito](#)).

Par exemple, si on inscrit "blanc" dans le champ de recherche, le système retournera les résultats suivants : blanc, blancs, blanche et blanches.

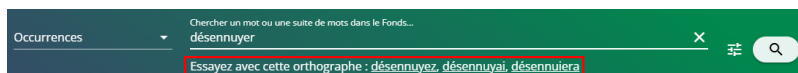
Champ de recherche

Inscrivez un mot ou une suite de mots dans le champ de recherche et cliquez sur l'icône **Rechercher** (ou pressez la touche **Entrée** du clavier). Au besoin, modifiez les [options de recherche](#).



Si vous cherchez un mot unique, le système proposera des orthographes alternatives.

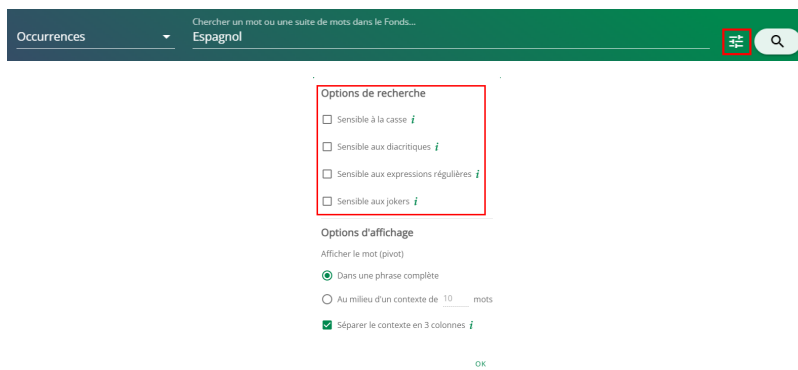
¹Unité textuelle minimale qui résulte de la segmentation automatique du texte d'un document. Il peut s'agir de mots ou de formes de mots (bleuetière, jouaient, vingt-cinq, tire-bouchon, etc.), de chiffres (500.000, 1608, etc.) ou d'abréviations (2e « deuxième », M. « monsieur », av. « avant », etc.).



Si vous cherchez une suite de mots, il n'est pas permis d'utiliser d'opérateurs logiques (+, " " (guillemets), -, *, ET, OU, etc.). La recherche s'effectue sur la suite de mots, telle que saisie. À noter que le système ne propose pas d'orthographe alternatives. Tous les signes de ponctuation sont traités comme des tokens. Ils peuvent faire l'objet d'une requête. Par exemple, la requête « bleu! » ne retourne pas les mêmes résultats que « bleu ».

Options de recherche

On accède aux options de recherche en cliquant sur l'icône qui se trouve tout juste à droite du champ de recherche. Ce faisant, une fenêtre contextuelle surgit.



Option	Description
Sensible à la casse	Si cette option est activée, le moteur de recherche tiendra compte des majuscules et des minuscules. Par défaut, cette option n'est pas activée. Par exemple, la recherche sur le mot « E spagnol » ne retournera pas les mêmes résultats que « e spagnol ».
Sensible aux diacritiques	Si cette option est activée, le moteur de recherche tiendra compte des signes diacritiques (accents aigu, grave et circonflexe, trémas et cédilles). Par défaut, cette option n'est pas activée. Par exemple, la recherche sur le mot « aigue » ne retournera pas les mêmes résultats que « aiguë » ou « aigüe ».
Sensible aux	Si cette option est activée, le moteur de recherche tiendra

Option	Description
expressions régulières	<p>compte des expressions régulières. Par défaut, cette option n'est pas activée.</p> <p>Consultez le lien suivant pour en apprendre plus sur les expressions régulières (aussi appelées <i>regex</i>).</p> <p>Par exemple, la requête « (panier casseau) de fraises » retournera les résultats « panier de fraises » et « casseau de fraises »</p>
Sensible aux jokers	<p>Si cette option est activée, le moteur de recherche tiendra compte des jokers¹ (aussi appelés <i>métacaractères</i>). Par défaut, cette option n'est pas activée.</p> <ul style="list-style-type: none"> • Le joker * (astérisque) représente un ou plusieurs caractères, voire aucun caractère. La requête « fêt* » permet de chercher <i>fête</i> et <i>fété</i>, mais aussi <i>fêtée</i>, <i>fêter</i> ou <i>fêtant</i>. • Le joker ? (point d'interrogation) remplace toujours un seul caractère. Par exemple, la requête « fêt? » permet de chercher <i>fête</i> ou <i>fété</i>.

Les options d'affichage sont définies à la section [Résultats de recherche](#).

Voir aussi

- [Résultats de recherche](#)

Dernière modification : 21 mai, 2025

¹Caractère qui permet de remplacer un ou plusieurs autres caractères dans une commande, une recherche ou une requête (définition tirée du Grand dictionnaire terminologique).

Résultats de recherche

Les résultats de la recherche qui a été effectuée à l'aide du [moteur de recherche](#) s'affichent dans la section centrale de l'écran qui correspond à la **concordance**¹. Cette liste de résultats permet de faire un tri rapide de tous les mots d'un texte (ou d'un corpus), de situer des mots pivots en contexte (KWIC – Key word in context), de compter le nombre d'occurrences, etc.

Sommaire

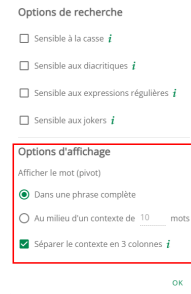
- [Options d'affichage](#)
- [Recherche sur un autre mot](#)
- [Filtres rapides](#)
 - [Corpus](#)
 - [Type de corpus](#)
 - [Année](#)
- [Tri des données](#)
- [Téléchargement des résultats](#)
- [Information à afficher](#)
- [Passage élargi](#)
- [Citer](#)
- [Nombre de lignes par page](#)

Options d'affichage

On accède aux options d'affichage en cliquant sur l'icône qui se trouve tout juste à droite du champ de recherche. Ce faisant, une fenêtre contextuelle surgit.



¹Liste des résultats d'une requête textuelle où les éléments recherchés sont affichés dans leur contexte.



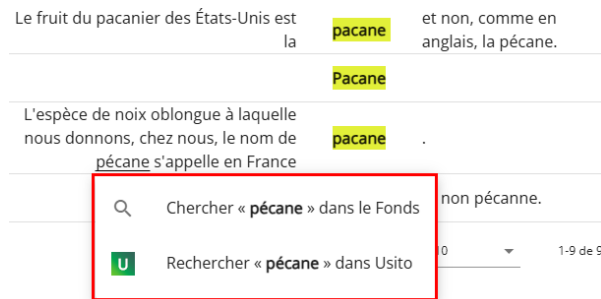
Option	Description												
Afficher le mot pivot	<p>Le mot pivot¹ est toujours surligné en jaune dans le concordancier. Sélectionnez l'une des deux options d'affichage :</p> <ul style="list-style-type: none"> • Dans une phrase complète : Le mot pivot est affiché au sein de la phrase complète auquel il appartient. Par défaut, c'est cet affichage qui est sélectionné. • Au milieu d'un contexte de _ mots : Le « contexte avant » et le « contexte après » sont composés du nombre de mots indiqué, jusqu'à concurrence de 50 mots. Par exemple, si 25 est inscrit, il y aura 25 mots dans le « contexte avant » et 25 mots dans le « contexte après ». Le nombre de mots tient compte des signes de ponctuation. À noter que pour le corpus BDTS, le contexte est toujours réduit pour des raisons de droit d'auteur. 												
Séparer le contexte en trois colonnes	<p>Par défaut, cette option est activée. Les résultats sont affichés dans 3 colonnes, Contexte avant, Pivot et Contexte après. Si on désactive cette option, les résultats seront affichés dans une phrase continue, sans surlignage.</p> <p>Option activée :</p> <table border="1" data-bbox="548 1409 1365 1707"> <thead> <tr> <th data-bbox="548 1409 971 1461">Contexte avant</th> <th data-bbox="971 1409 1133 1461">Pivot</th> <th data-bbox="1133 1409 1365 1461">Contexte après</th> </tr> </thead> <tbody> <tr> <td data-bbox="548 1461 971 1587">Après le souper, il se retire souvent dans sa chambre sous prétexte de faire une sieste et il y reçoit la bonne : « Manon vient me rejoindre et me</td> <td data-bbox="971 1461 1133 1587">désennuyer</td> <td data-bbox="1133 1461 1365 1587">... Oui, je la méprise.</td> </tr> <tr> <td data-bbox="548 1587 971 1650">Tu devrais plutôt essayer de te</td> <td data-bbox="971 1587 1133 1650">désennuyer</td> <td data-bbox="1133 1587 1365 1650">.</td> </tr> <tr> <td data-bbox="548 1650 971 1707">te</td> <td data-bbox="971 1650 1133 1707">désennuyer</td> <td data-bbox="1133 1650 1365 1707">/ oui / il part</td> </tr> </tbody> </table>	Contexte avant	Pivot	Contexte après	Après le souper, il se retire souvent dans sa chambre sous prétexte de faire une sieste et il y reçoit la bonne : « Manon vient me rejoindre et me	désennuyer	... Oui, je la méprise.	Tu devrais plutôt essayer de te	désennuyer	.	te	désennuyer	/ oui / il part
Contexte avant	Pivot	Contexte après											
Après le souper, il se retire souvent dans sa chambre sous prétexte de faire une sieste et il y reçoit la bonne : « Manon vient me rejoindre et me	désennuyer	... Oui, je la méprise.											
Tu devrais plutôt essayer de te	désennuyer	.											
te	désennuyer	/ oui / il part											

¹Dans une concordance, le mot (ou la suite de mots) choisi pour la recherche, accompagné de son contexte et mis en évidence typographiquement. L'acronyme KWIC signifie KeyWord In Context, littéralement « mot-clé dans son contexte ».

Option	Description
	<p>Option désactivée :</p> <div style="border: 1px solid #ccc; padding: 5px; margin: 5px 0;"> <p>Phrase</p> <p>Après le souper, il se retire souvent dans sa chambre sous prétexte de faire une sieste et il y reçoit la bonne : « Manon vient me rejoindre et me désennuyer... Oui, je la méprise.</p> <p>Tu devrais plutôt essayer de te désennuyer.</p> <p>te désennuyer / oui / il part</p> </div>

Recherche sur un autre mot

À tout moment, on peut effectuer une recherche sur un autre **mot**¹ affiché dans la concordance. Cliquez simplement sur un mot du contexte avant ou du contexte après et sélectionnez l'option **Chercher X dans le Fonds**.



On peut également rechercher le mot dans le dictionnaire [Usito](#).

Filtres rapides

Il est possible d'appliquer rapidement des filtres sur les résultats de recherche. En sélectionnant un filtre, les résultats sont automatiquement mis à jour (le nombre d'**occurrences**² change). On peut filtrer selon le [corpus](#) de provenance, le [type de cor-](#)

¹Dans le contexte de la plateforme FDLQ, mot n'est pas utilisé dans le sens d'« unité lexicale » qu'il prend en linguistique, mais plutôt comme synonyme de mot-forme (mot séparé par des espaces) ou encore de token. Les mots faisant partie du lexique d'un document ou d'un corpus correspondent ainsi à l'ensemble des tokens qu'on y trouve; suivant cette logique, pomme de terre correspond à trois tokens.

²Apparition d'une unité linguistique dans le discours; par métonymie, cette unité (définition tirée d'Usito).

[pus](#) et [l'année](#). Les filtres sélectionnés peuvent s'additionner et sont affichés dans le haut de la section :

Filtres

Année **est entre** 1950 et 2025 ✕

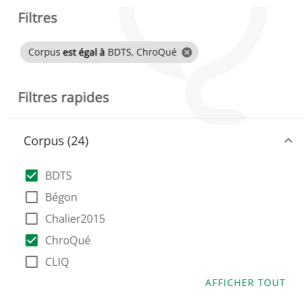
Type de corpus **est égal à** Textuel, Oral ✕

Corpus **est égal à** BDTS ✕

On peut supprimer les filtres sélectionnés en cliquant sur le petit **X** au bout de la ligne.

Corpus

Cochez les **corpus**¹ selon lesquels les résultats de recherche doivent être filtrés. La section n'affiche que les 5 premiers corpus du Fonds de données linguistiques du Québec : cliquez sur **Afficher tout** pour avoir accès à tous les autres filtres de corpus.



On peut également appliquer un filtre sur le corpus directement dans le concordancier. Pour ce faire, cliquez sur l'icône représentant 3 petites lignes en forme de triangle. La case du corpus correspondant sera automatiquement cochée dans la section **Filtres rapides**.

Corpus	Document	Auteurs	Année ↓ 1
CLIQ ☰	Bouchard 2022 Étaient ☰	Serge Bouchard, Marie-Christine Lévesque	2022
CLIQ ☰	Perrault 2022 Nord ☰	Pierre Perrault	2022

¹Ensemble de textes établi selon un principe de documentation exhaustive, un critère thématique ou exemplaire en vue de leur étude linguistique (définition tirée d'Usito).

On peut aussi appliquer ce type de filtre sur le document, mais il n'y a pas d'équivalent dans les filtres rapides.

Corpus	Document	Auteurs	Année ↓ 1
CLIQ ☰	Vigneau 2021 Naufrages ☰	Placide Vigneau	2021
CLIQ ☰	Vigneau 2021 Naufrages ☰	Placide Vigneau	2021

Type de corpus

Cochez les types de corpus selon lesquels les résultats de recherche doivent être filtrés.

Filtres

Type de corpus est égal à Textuel, Oral

Filtres rapides

Corpus (24)

Type de corpus (4)

- Dialectologique
- Oral
- Textuel
- Métalinguistique

Il y a quatre types de corpus : dialectologique (données sur la variation régionale), oral (langue parlée), textuel (langue écrite) ou métalinguistique (discours sur la langue).

Année

Sélectionnez l'intervalle selon lequel les résultats de recherche doivent être filtrés. On peut inscrire manuellement les années de départ et de fin dans les champs prévus à cet effet, mais on peut aussi utiliser la glissière.

Filtres

Année est entre 1950 et 2025

Filtres rapides

Corpus (24)

Type de corpus (4)

Année

Entre 1950 et 2025

1950 2025

Tri des données

On peut trier les résultats de recherche en cliquant sur le nom d'une colonne.

2449 occurrences

TÉLÉCHARGER LES RÉSULTATS DE CETTE PAGE

Choisir l'information à afficher
Corpus Document Auteurs -1 autres

Corpus ↑ 3	Document ↓ 4	Auteurs ↑ 2	Année ↓ 1	Contexte avant	Pivot	Contexte après
CLIQ	Perrault_2022_Nord	Pierre Perrault	2022	Entre le jaune grue, le rouge conteneur, le bleu camion, le blanc bateau, le rouge feuille d'	érable	garde-côtes, on aperçoit un peu de cap Diamant. Des miettes triviales de réalité. Mais on ne voit rien briller, scintiller,
CLIQ	Perrault_2022_Nord	Pierre Perrault	2022	-ils quitté la branche d'un brise-glace ? Se sentent-ils possédés par les emblèmes qui les entourent, les feuilles d'	érable	? Partout des images ! Des signes ! Une possession ! La raison du plus fort ! Comment, sans l'aide des images

- Au premier clic sur une colonne, le chiffre **1** et une flèche vers le haut (↑) s'affichent. Cela indique que le tri s'affiche d'abord sur cette colonne et que le tri est ascendant.
- Au second clic sur la même colonne, la flèche s'inverse (↓) : le tri est maintenant descendant.
- Au troisième clic, toujours sur la même colonne, le tri est supprimé.

On peut appliquer des tris sur plusieurs colonnes (dans l'exemple ci-dessus, il y en a 4). Toutefois, dès que plusieurs colonnes sont affichées et que l'affichage bascule à la verticale (voir section [Information à afficher](#)), il n'est plus possible de trier les résultats.

Corpus CLIQ	Document Boucharde_2021_Diesel	Auteurs Serge, 2021, Montréal (Québec), 1947, Boucharde Mark, Fortier	Année 2021	Lieu Montréal	Genre littérature	l'autre de la van. Chargé comme j'étais, j'ai dû le mettre plate comme une galette	. Personne aurait pu le rattraper. Le gars qui me suivait m'a dit que le loup s'était	... CITER
---------------------------------------	--	---	----------------------	-------------------------	-----------------------------	---	---	-----------

Téléchargement des résultats

Il est possible de télécharger les résultats de recherche en format *Excel*. Pour ce faire, cliquez sur le bouton **Télécharger les résultats de cette page**.

2449 occurrences

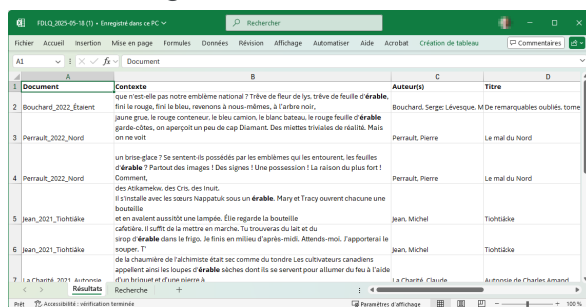
TÉLÉCHARGER LES RÉSULTATS DE CETTE PAGE

Choisir l'information à afficher

Corpus Document Auteurs +1 autres

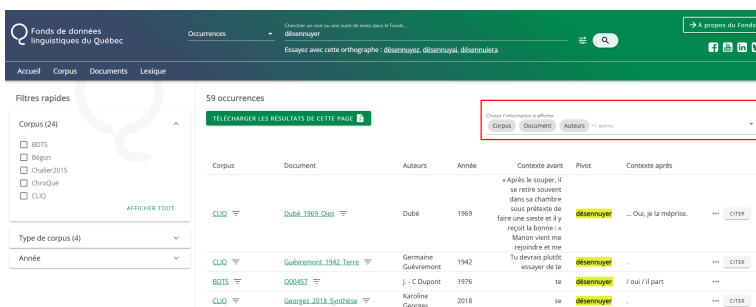
Corpus	Document	Auteurs	Année	Contexte avant	Pivot	Contexte après
CLIQ	Bouchard_2022_Étaient	Serge Bouchard, Marie-Christine Lévesque	2022	dit et le répète : que n'est-elle pas notre emblème national ? Trêve de fleur de lys, trêve de feuille d'	érable	, fini le rouge, fini le bleu, revenons à nous-mêmes, à l'arbre noir, à l'arbre à barbe
CLIQ	Perrault_2022_Nord	Pierre Perrault	2022	Entre le jaune grue, le rouge conteneur, le bleu camion, le blanc bateau, le rouge feuille d'	érable	garde-côtes, on aperçoit un peu de cap Diamant. Des miettes triviales de réalité. Mais on ne voit rien briller, scintiller,

Attention! Seuls les résultats de la page courante seront téléchargés. Au besoin, affichez un plus grand [nombre de lignes par page](#). Pour télécharger tous les résultats d'une requête, il faut donc répéter le téléchargement pour l'ensemble des pages. Dans le fichier *Excel*, le premier onglet du chiffrier est consacré aux résultats de la recherche, alors que le second onglet montre les détails de la requête.



Information à afficher

Par défaut, les informations **Document** et **Corpus** sont affichées. Pour sélectionner d'autres éléments d'information à afficher, cliquez sur une des étiquettes ou sur la flèche du champ **Choisir les informations à afficher**.



Voici la liste des informations pouvant être sélectionnées et leur signification :

Option	Description
Accès au contenu	Le contenu textuel correspond à la transcription du document .

Option	Description
textuel	<ul style="list-style-type: none"> • Si l'accès est limité, on peut interroger le document au complet, mais on ne peut pas le lire dans son intégralité. • Si l'accès est complet, on peut non seulement interroger le document au complet, mais on peut aussi le lire dans son intégralité.
Accès au contenu binaire	<p>Le contenu binaire correspond à une version numérique du document dans son format original (il peut s'agir d'une image ou d'un enregistrement audio).</p> <ul style="list-style-type: none"> • Si l'accès est interdit sans autorisation, il n'est pas possible de consulter le document sans avoir obtenu la permission au préalable. Contactez-nous pour en savoir davantage. • Si l'accès est complet, il est possible de consulter le fichier directement sur la plateforme du Fonds.
Année	Année de publication du document.
Auteurs	Auteurs du document.
Corpus	Nom du corpus d'appartenance du document. On peut cliquer sur le nom du corpus pour ouvrir la page qui lui est consacrée.
Document	Nom du document. On peut cliquer sur le nom du document pour ouvrir la page qui lui est consacrée.
Genre	Genre du document. Par exemple, littérature, littérature jeunesse, entrevue sociolinguistique, etc.
Licence du document	Type d'entente d'utilisation du document.
Lieu	Lieu de publication du document.
Maison d'édition	Maison d'édition qui a publié le document.
Référence bibliographique	Référence bibliographique complète.
Source fichier binaire	Provenance du fichier binaire.
Source fichier textuel	Provenance de la transcription.
Sous-genre	S'il y a lieu, sous-genre littéraire du document. Par exemple, roman, contes et légendes, théâtre, etc.

Option	Description
Titre	Titre du document.
Type de corpus	Type de corpus auquel appartient le document : dialectologique, oral, textuel ou métalinguistique.

Notez qu'à partir d'un certain nombre d'informations sélectionnées, l'affichage par colonne change pour prendre moins de place. Par exemple :

Information	Contexte avant	Pivot	Contexte après
Corpus CLIQ			
Document Dubé.1969.Oies			
Auteurs 1930, Dubé			
Année 1969	» Après le souper, il se retire souvent dans sa chambre sous prétexte de faire une sieste et il y reçoit la bonne : « Manon vient me rejoindre et me		
Maison d'édition Leméac		désennuyer	... Oui, je la méprise.
Licence du document licence générale FDLQ			... CITER
Genre littérature			
Accès au contenu textuel limité			

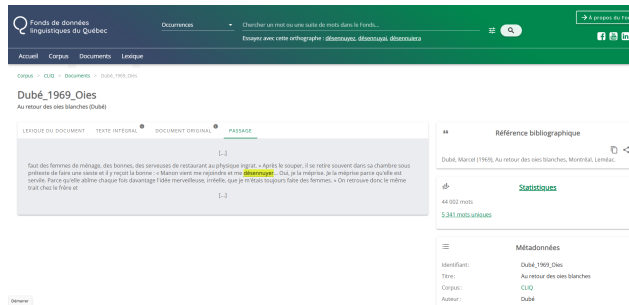
Passage élargi

En cliquant sur l'icône représentant trois points de suspension, on peut consulter le passage élargi du mot (ou de la suite de mots) recherché.

Corpus	Document	Année	Contexte avant	Pivot	Contexte après
CLIQ	Dubé.1969.Oies	1969	» Après le souper, il se retire souvent dans sa chambre sous prétexte de faire une sieste et il y reçoit la bonne : « Manon vient me rejoindre et me	désennuyer	... Oui, je la méprise.
					... CITER

Le passage élargi affiche 50 mots avant le mot **pivot**¹ et 50 mots après (sauf pour le corpus BDTs, limité à 10 mots avant et après).

¹Dans une concordance, le mot (ou la suite de mots) choisi pour la recherche, accompagné de son contexte et mis en évidence typographiquement. L'acronyme KWIC signifie KeyWord In Context, littéralement « mot-clé dans son contexte ».

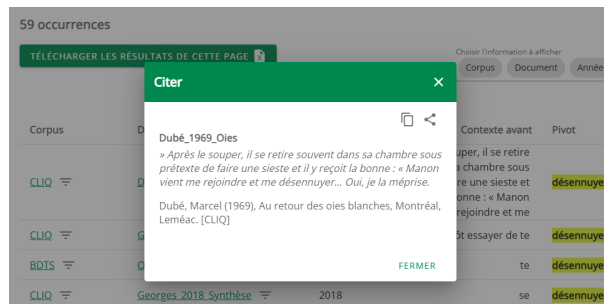


Citer

Cliquez sur le bouton **Citer** pour afficher une fenêtre contextuelle qui permet de citer le passage sélectionné.

Corpus	Document	Année	Contexte avant	Pivot	Contexte après	
CLIQ	Dubé_1969_Oies	1969	» Après le souper, il se retire souvent dans sa chambre sous prétexte de faire une sieste et il y reçoit la bonne : « Manon vient me rejoindre et me	désennuyer	... Oui, je la méprise.	... CITER

La fenêtre contextuelle affiche le nom du document, le passage en question et sa référence bibliographique. Dans le coin supérieur droit de la fenêtre, la première icône permet de copier le texte de la citation; la seconde permet plutôt de copier l'URL de la citation.



Nombre de lignes par page

Il est possible d'afficher plus ou moins de lignes de résultats par page. Pour ce faire, faites défiler la page jusqu'en bas et sélectionnez un nombre de lignes dans le champ prévu à cet effet. On peut afficher 5, 10, 25, 50 ou 100 lignes par page. Le système indique le nombre de lignes affichées, en fonction des résultats de la recherche (dans l'exemple ci-dessous, les 10 premières lignes, sur 60 lignes au total).

...tant le ...rtait un ...omplet	bleu poudre	même nuance, baissé sur les yeux. Ses souliers noirs miraient le plafond. Les	... CITER
...qu'on a ...rie de la ...amisole	bleu poudre	: une hypothèse confirmée. Avril, le printemps s'essaye, et moi aussi. Je tente une expérience hautement scientifique pour valider	... CITER

Lignes par page 10 1-10 de 60 < >

Pour naviguer d'une page à l'autre, utilisez les flèches gauche et droite.

...tant le ...rtait un ...omplet	bleu poudre	même nuance, baissé sur les yeux. Ses souliers noirs miraient le plafond. Les	... CITER
...qu'on a ...rie de la ...amisole	bleu poudre	: une hypothèse confirmée. Avril, le printemps s'essaye, et moi aussi. Je tente une expérience hautement scientifique pour valider	... CITER

Lignes par page 10 1-10 de 60 < >

Voir aussi

- [Moteur de recherche](#)
- [Corpus](#)
- [Document](#)

Dernière modification : 28 mai, 2025

Corpus

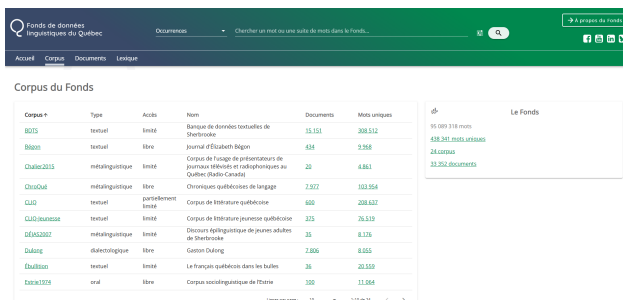
Le Fonds de données linguistiques du Québec est composé (pour l'instant!) de 24 **corpus**¹. Pour en savoir plus sur les divers corpus qui constituent le FDLQ, consultez le lien suivant : [Composition du Fonds](#).

Sommaire

- [Corpus du Fonds](#)
- [Corpus individuel](#)
 - [Référence bibliographique](#)
 - [Statistiques](#)
 - [Métadonnées](#)

Corpus du Fonds

La page **Corpus du Fonds** montre la liste de tous les corpus que le Fonds de données linguistiques du Québec a réunis, y compris quelques statistiques générales. Les principales caractéristiques de chaque corpus sont affichées dans les colonnes du tableau. Il est d'ailleurs possible d'appliquer un filtre sur les colonnes (un seul filtre est permis). Les liens de la colonne **Documents** mènent à l'onglet [Documents du corpus](#) et les liens de la colonne **Mots uniques** mènent à l'onglet [Lexique du corpus](#).



Corpus #	Type	Accès	Nom	Documents	Mots uniques	
B015	textuel	libre	Banque de données textuelles de Sherbrooke	15 151	308 512	
B020	textuel	libre	Journal d'Elizabeth Bégin	534	3 368	
Chabo2015	métalinguistique	libre	Corpus de l'usage de participations de personnes bilingues et codiaphones au Québec (Québec-Canada)	20	4 861	
Chou04	métalinguistique	libre	Chroniques québécoises de langage	7 272	103 354	
CLD	textuel	partiellement libre	Corpus de littérature québécoise	600	208 637	
CLD200504	textuel	libre	Corpus de littérature jeunesse québécoise	375	26 519	
CLD2002	métalinguistique	libre	Discours épidémiologique de jeunes adultes de Sherbrooke	35	8 176	
Dubois	métalinguistique	libre	Glossaire Dubois	7 656	8 055	
États05	textuel	libre	Le français québécois dans les bulletins	36	20 559	
Estab2015	oral	libre	Corpus sociolinguistique de Thémis	100	11 564	

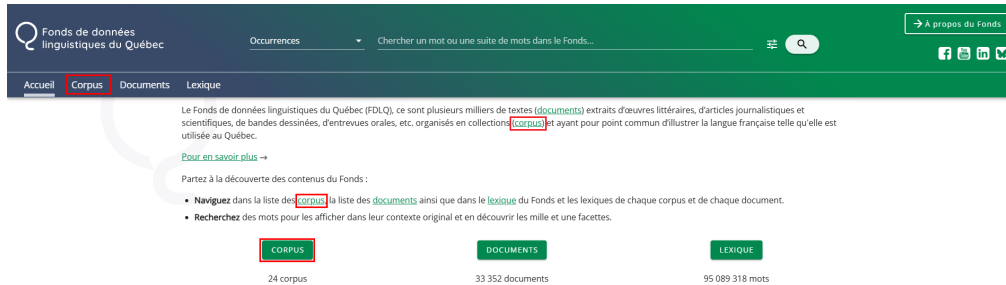
Le Fonds

- 16 000 718 mots
- 238 267 mots uniques
- 24 corpus
- 23 302 documents

On accède à la page **Corpus du Fonds** par la [page d'accueil](#) du FDLQ, en cliquant sur les liens ou le bouton vert tout à gauche de la page. On peut aussi y accéder à partir

¹Ensemble de textes établi selon un principe de documentation exhaustive, un critère thématique ou exemplaire en vue de leur étude linguistique (définition tirée d'Usito).

de n'importe quelle page du FDLQ en cliquant sur l'onglet **Corpus** qui se trouve tout juste en dessous du logo du Fonds.



Fonds de données linguistiques du Québec

Occurrences Chercher un mot ou une suite de mots dans le Fonds...

Accueil **Corpus** Documents Lexique

Le Fonds de données linguistiques du Québec (FDLQ), ce sont plusieurs milliers de **textes (documents)** extraits d'œuvres littéraires, d'articles journalistiques et scientifiques, de bandes dessinées, d'entrevues orales, etc. organisés en collections **(corpus)** et ayant pour point commun d'illustrer la langue française telle qu'elle est utilisée au Québec.

Pour en savoir plus →

Partez à la découverte des contenus du Fonds :

- Naviguez dans la liste des **corpus** la liste des **documents** ainsi que dans le **lexique** du Fonds et les lexiques de chaque corpus et de chaque document.
- Recherchez des mots pour les afficher dans leur contexte original et en découvrir les mille et une facettes.

CORPUS 24 corpus

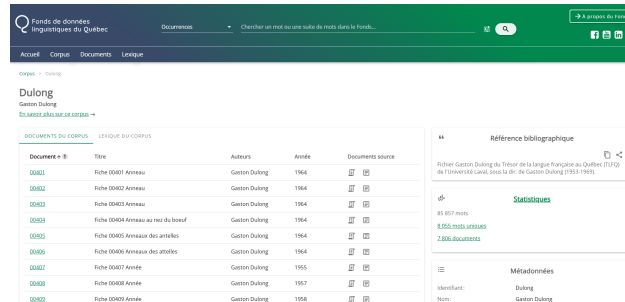
DOCUMENTS 33 352 documents

LEXIQUE 95 089 318 mots

Corpus individuels

La page d'un corpus individuel, celle du corpus *Dulong* dans l'exemple ci-dessous, montre la liste de tous les **documents** qui le constituent. Le lien **En savoir plus sur ce corpus** vous redirigera vers l'article du Fonds de données linguistiques du Québec qui décrit en détail le corpus en question.

La page individuelle des corpus s'ouvre toujours sur l'onglet **Documents du corpus**. Cliquez sur l'onglet **Lexique du corpus** pour consulter la liste de tous les mots qui composent le corpus.



Fonds de données linguistiques du Québec

Occurrences Chercher un mot ou une suite de mots dans le Fonds...

Accueil Corpus **Documents** Lexique

Corpus > Dulong

Dulong
Gaston Dulong
En savoir plus sur ce corpus →

DOCUMENTS DU CORPUS LEXIQUE DU CORPUS

Document #	Titre	Auteurs	Année	Documents source
00401	Fiche 00401 Arneau	Gaston Dulong	1964	📄 📄
00402	Fiche 00402 Arneau	Gaston Dulong	1964	📄 📄
00403	Fiche 00403 Arneau	Gaston Dulong	1964	📄 📄
00404	Fiche 00404 Arneau au nez du bœuf	Gaston Dulong	1964	📄 📄
00405	Fiche 00405 Arneau des articles	Gaston Dulong	1964	📄 📄
00406	Fiche 00406 Arneau des articles	Gaston Dulong	1964	📄 📄
00407	Fiche 00407 Arnette	Gaston Dulong	1955	📄 📄
00408	Fiche 00408 Arnette	Gaston Dulong	1957	📄 📄
00409	Fiche 00409 Arnette	Gaston Dulong	1958	📄 📄

Référence bibliographique

Fiche Gaston Dulong du Trésor de la langue française au Québec (TLFQ) de l'Université Laval, dossier de la collection Dulong 1955-1965.

Statistiques

83 807 mots
8 555 mots uniques
2 306 documents

Métadonnées

Identifiant: Dulong
Nom: Gaston Dulong

S'il y a lieu, on peut consulter le **texte intégral** d'un document (en cliquant sur l'icône 📄) et même le **document original** (en cliquant sur l'icône 📄).

La page individuelle des corpus présente toujours les sections suivantes : **Référence bibliographique**, **Statistiques** et **Métadonnées**.

Référence bibliographique

Cette section affiche la référence bibliographique complète du corpus. À droite, la première icône permet de copier le texte de la référence bibliographique, la seconde permet plutôt de copier l'URL de la référence bibliographique.

“

Référence bibliographique



Fichier Gaston Dulong du Trésor de la langue française au Québec (TLFQ) de l'Université Laval, sous la dir. de Gaston Dulong (1953-1969).

Statistiques

Cette section affiche les principales statistiques du corpus, soit le nombre de mots (et de mots uniques) qu'il contient ainsi que le nombre de documents dont il est constitué. En cliquant sur le lien **N mots uniques**, l'onglet [Lexique du corpus](#) s'affiche. En cliquant sur le lien **N documents**, l'onglet [Documents du corpus](#) s'affiche.



[Statistiques](#)

85 857 mots

[8 055 mots uniques](#)

[7 806 documents](#)

Métadonnées

Cette section affiche les **métadonnées**¹ courantes de chaque corpus : son identifiant, son nom complet, son [type](#) (dialectologique, oral, textuel ou métalinguistique), le responsable du corpus et les années de sa constitution.

¹Donnée qui renseigne sur la nature de certaines autres données dans le but d'en faciliter la compréhension et la gestion. Définition tirée du Grand dictionnaire terminologique.



Métadonnées

Identifiant:	Dulong
Nom:	Gaston Dulong
Type:	dialectologique
Responsable:	Wim Remysen
Années de constitution:	1953, 1954, 1955, 1956, 1957, 1958, 1959, 1960, 1961, 1962, 1963, 1964, 1965, 1966, 1967, 1968, 1969

Voir aussi

- [Documents](#)
- [Lexique](#)

Dernière modification : 27 mai, 2025

Documents

Le Fonds de données linguistiques du Québec est composé (pour l'instant!) de plus de 33 000 **documents**¹, répartis dans 24 **corpus**².

Sommaire

- [Documents du Fonds](#)
- [Documents du corpus](#)
- [Document individuel](#)
 - [Onglet Texte intégral](#)
 - [Onglet Document original](#)
 - [Onglet Transcription avec audio](#)
 - [Onglet Passage](#)

Documents du Fonds

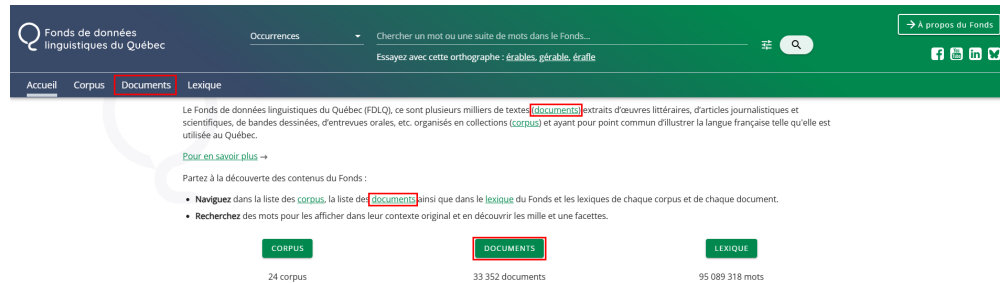
La page **Documents du Fonds** montre la liste de tous les documents que le Fonds de données linguistiques du Québec a réunis à travers ses corpus, ainsi que quelques statistiques générales. Les principales caractéristiques de chaque document sont affichées dans les colonnes du tableau.

Corpus	Document	Titre	Auteurs	Année	Documents source
ESQ16334	ESQ16334_236	Entrée 236	Normand Beaulac, Pierre Martel, Michel Thériault	1974	
ESQ16334	ESQ16334_112	Entrée 112	Normand Beaulac, Pierre Martel, Michel Thériault	1974	
ESQ16334	ESQ16334_153	Entrée 153	Normand Beaulac, Pierre Martel, Michel Thériault	1974	
ESQ16334	ESQ16334_211	Entrée 211	Normand Beaulac, Pierre Martel, Michel Thériault	1974	
ESQ16334	ESQ16334_180	Entrée 180	Normand Beaulac, Pierre Martel, Michel Thériault	1974	
ESQ16334	ESQ16334_138	Entrée 138	Normand Beaulac, Pierre Martel, Michel Thériault	1974	
ESQ16334	ESQ16334_104	Entrée 104	Normand Beaulac, Pierre Martel, Michel Thériault	1974	
ESQ16334	ESQ16334_102	Entrée 102	Normand Beaulac, Pierre Martel, Michel Thériault	1974	
ESQ16334	ESQ16334_160	Entrée 160	Normand Beaulac, Pierre Martel, Michel Thériault	1974	
ESQ16334	ESQ16334_144	Entrée 144	Normand Beaulac, Pierre Martel, Michel Thériault	1974	

¹Unité de base d'un corpus (qui se définit comme une collection de documents), par exemple une œuvre littéraire, une chanson, une bande dessinée, un article de journal, une fiche dialectologique ou encore une entrevue. Dans le cadre du Fonds, une distinction est faite entre le contenu textuel du document (transcription) et le contenu binaire (version numérisée de l'original : image ou enregistrement audio).

²Ensemble de textes établi selon un principe de documentation exhaustive, un critère thématique ou exemplaire en vue de leur étude linguistique (définition tirée d'Usito).

On accède à la page **Documents du Fonds** par la [page d'accueil](#) du FDLQ, en cliquant sur les liens ou le bouton vert au centre de la page. On peut aussi y accéder à partir de n'importe quelle page du FDLQ en cliquant sur l'onglet **Documents** qui se trouve tout juste en dessous du logo du Fonds.



Documents du corpus

L'onglet **Documents du corpus** affiche tous les documents qui constituent un corpus. Un lien permet d'accéder à chaque document individuellement : en cliquant sur ce lien, c'est la page individuelle de chaque document qui s'affiche. L'onglet [Lexique du corpus](#) permet de consulter le lexique spécifique du corpus.

Corpus > DÉJAS2007

DÉJAS2007
Discours épilinguistique de jeunes adultes de Sherbrooke
[En savoir plus sur ce corpus](#) →

DOCUMENTS DU CORPUS		LEXIQUE DU CORPUS		
Document ↕ 1	Titre	Auteurs	Année	Documents source
DÉJAS2007_Adèle	Entrevue Adèle	Amélie-Hélène Rheault	2007	
DÉJAS2007_Alexandra	Entrevue Alexandra	Amélie-Hélène Rheault	2007	
DÉJAS2007_Andrée-Anne	Entrevue Andrée-Anne	Amélie-Hélène Rheault	2007	
DÉJAS2007_Annie	Entrevue Annie	Amélie-Hélène Rheault	2007	
DÉJAS2007_Ariane	Entrevue Ariane	Amélie-Hélène Rheault	2007	

Document individuel

Cette page comprend une série d'onglets. Par défaut, c'est le [lexique du document](#) qui s'affiche. Le choix des autres onglets qu'on peut ouvrir varie selon la nature du document (texte intégral et document intégral pour les documents écrits, transcription avec audio pour les enregistrements sonores) et selon les autorisations obtenues pour diffuser certains contenus (il arrive par exemple que nous n'ayons pas le droit de diffuser une transcription complète). L'identifiant des documents est spécifique à chaque corpus.

Corpus > Témoins/Actes > Documents > Aleyrac_1935 [1755], Aventures

Aleyrac_1935 [1755] Aventures

Aventures militaires au XVIIIe siècle d'après les mémoires de Jean-Baptiste d'Aleyrac (Jean-Baptiste d'Aleyrac)

LEXIQUE DU DOCUMENT TEXTE INTÉGRAL DOCUMENT ORIGINAL

Tout en haut
Lexique alphabétique

A B C D E F G H I J K L M N O P Q R S T U V W X Y Z

< a = aventures >

a

a(1)	ajoutent(1)	après(1)	au(5)	aventures(1)
adjectif(1)	aleyrac(2)	après-midi(1)	autre(1)	
agit(1)	amarcer(1)	attacher(1)	aux(1)	

Référence bibliographique

Aleyrac, Jean-Baptiste d', *Aventures militaires au XVIIIe siècle d'après les mémoires de Jean-Baptiste d'Aleyrac*, publi. par Charles Coche, Paris, Ed. Berger-Levrault, 1935, [Témoignage de 1755]

Statistiques

245 mots
164 mots uniques

Métadonnées

Identifiant: Aleyrac_1935 [1755] Aventures
Titre: Aventures militaires au XVIIIe siècle d'après les mémoires de Jean-Baptiste d'Aleyrac

Texte intégral

L'onglet **Texte intégral** affiche le texte intégral du document converti en caractères exploitables par un système informatique. Par exemple, ce texte intégral tiré du document *Bégon-1748-001* du corpus *Bégon*.

Corpus > Bégon > Documents > Bégon-1748-001

Bégon-1748-001

Lettre d'Élisabeth Bégon (Élisabeth Bégon)

LEXIQUE DU DOCUMENT **TEXTE INTÉGRAL** DOCUMENT ORIGINAL

Le 12 novembre 1748

a [preñ mon cher fils que je me vois debarrair de tans d'escrits qui m'ont beaucoup costé, je pourai avec la meme satisfaction que j'ay toujours eu a m'entretenir avec toy le faire tous les jours, et te repeter tant fois que ces tous ce (ou) me reste de confessions, tu Escy cher fils combien ton absence m'est dure a supporter. j'y ay eu de la peine a te voir partir jeug quelle dot entre celles ou j'Escy a preñ un couple les trois quard du jours avec ma chere petite-fille, je l'ay mandé que maitre et lilly ont été a quebec avec me de lilly le 15 juillet, elle ne font revenue qu'il y a quelque jours tres entoufflé de des [pailles de] quebec ou m' l' [intendant] fait dantir a toute main, les habit ne font pas a priori d'aller a toute ces feste, mais elle y ont manger et yeu certe belles agremie qui font le belle en des confessions si la mode me lanodere y est la brillante tous le reste n'est rien en comparaison il se fassent et nous honore de les yeux comme tu Escy de la bonne fation, notre general est

Référence bibliographique

Lettre d'Élisabeth Bégon, 12 novembre 1748, BANQ Québec, Fonds Famille Bégon (P2.P1), Transcription et établissement du texte par France Martneau et Yves Charles Morin (2009).

Statistiques

247 mots
181 mots uniques

Métadonnées

Identifiant: Bégon-1748-001
Titre: Lettre d'Élisabeth Bégon
Corpus: Bégon

Document original

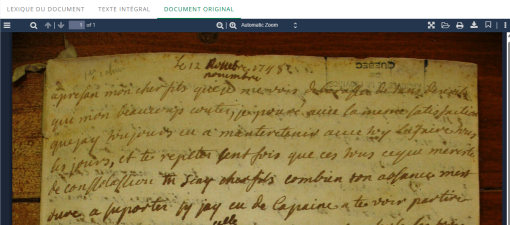
L'onglet **Document original** affiche le document numérisé original. Par exemple, ce document original tiré du document *Bégon-1748-001* du corpus *Bégon*.

Corpus > Bégon > Documents > Bégon-1748-001

Bégon-1748-001

Lettre d'Élisabeth Bégon (Élisabeth Bégon)

LEXIQUE DU DOCUMENT TEXTE INTÉGRAL **DOCUMENT ORIGINAL**



Référence bibliographique

Lettre d'Élisabeth Bégon, 12 novembre 1748, BANQ Québec, Fonds Famille Bégon (P2.P1), Transcription et établissement du texte par France Martneau et Yves Charles Morin (2009).

Statistiques

247 mots
181 mots uniques

Métadonnées

Identifiant: Bégon-1748-001
Titre: Lettre d'Élisabeth Bégon
Corpus: Bégon

Transcription avec audio

L'onglet **Transcription avec audio** permet de consulter la transcription du document, avec son fichier audio. Par exemple, cette transcription avec audio tirée du document *PFC-Québec_Grande-Rivière_cqhd11* du corpus *PFC-Québec*.

Lorsqu'on fait jouer l'enregistrement, la transcription est automatiquement synchronisée et un surlignage gris indique ce qui est prononcé. Cliquez n'importe où dans la transcription pour jouer l'extrait en question.

Corpus > PFC-Québec > Documents > PFC-Québec_Grande-Rivière_cqhd11

PFC-Québec_Grande-Rivière_cqhd11
cqhd11(Marie-Hélène Côté)

LEXIQUE DU DOCUMENT TRANSCRIPTION AVEC AUDIO

00:00:00 SB (0000)

00:00:00 DB Tu dois trouver qu'on parle vite hein?

00:00:00 SB La plupart du monde là ici.

00:00:00 EQ

00:00:00 DB Mmh.

00:00:00 SB Mmh.

00:00:00 EQ

PFC-Québec_Grande-Rivière_cqhd11

Référence bibliographique

Côté, Marie-Hélène (2012). « Conversation libre cqhd11, Grande-Rivière ». Cité dans Phonologie du français contemporain : corpus Québec (PFC-Québec).

Statistiques

11 930 mots
1 225 mots uniques

Métadonnées

Identifiant: PFC-Québec_Grande-Rivière_cqhd11
Titre: cqhd11
Corpus: PFC-Québec

Au besoin, consultez le document d'aide consacré au [lecteur audio](#).

Passage

L'onglet **Passage** ne s'affiche que si on clique sur l'icône représentant trois points de suspension à partir de la **concordance**¹, après avoir effectué une recherche. L'onglet **Passage** affiche 50 **mots**² avant le mot **pivot**³ et 50 mots après (sauf pour le corpus BDTs, limité à 10 mots avant et après).

¹Liste des résultats d'une requête textuelle où les éléments recherchés sont affichés dans leur contexte.

²Dans le contexte de la plateforme FDLQ, mot n'est pas utilisé dans le sens d'« unité lexicale » qu'il prend en linguistique, mais plutôt comme synonyme de mot-forme (mot séparé par des espaces) ou encore de token. Les mots faisant partie du lexique d'un document ou d'un corpus correspondent ainsi à l'ensemble des tokens qu'on y trouve; suivant cette logique, pomme de terre correspond à trois tokens.

³Dans une concordance, le mot (ou la suite de mots) choisi pour la recherche, accompagné de son contexte et mis en évidence typographiquement. L'acronyme KWIC signifie KeyWord In Context, littéralement « mot-clé dans son contexte ».

ACFBullB-020

[Bulletin de linguistique] (Académie canadienne-française)

LEXIQUE DU DOCUMENT

TEXTE INTÉGRAL

DOCUMENT ORIGINAL

PASSAGE

[...]

Le fruit de l'amandier est l'amande. C'est à tort que nous l'appelons noix piquée. Le fruit du pacanier des États-Unis est la **pacane**, et non, comme en anglais, la pécanne. Le noyer d'Amérique produit une noix appelée noix du noyer. Pour quelle raison est-elle devenue ici noix longue ? Le mot fait image, mais il n'est pas correct. Que dire maintenant du nom d'ortel de nègre donné à la noix du Brésil ? Il s'agit, cette fois, d'un américanisme.

[...]

Voir aussi

- [Corpus](#)
- [Lexique](#)
- [Lecteur audio](#)

Dernière modification : 26 mai, 2025

Lecteur audio

Le lecteur audio n'apparaît que dans l'onglet **Transcription avec audio** des **documents**¹ pour lesquels une transcription est disponible (**corpus**² oraux).

The screenshot shows a web interface for a document titled 'PFC-Québec_Grande-Rivière_cqhc2l' by Marie-Hélène Côté. The interface is split into two main panels. The left panel, titled 'TRANSCRIPTION AVEC AUDIO', displays a transcript with speaker labels (TG, JL, JG, JG, JG, JG, JG, JG) and time markers. The right panel contains a 'Référence bibliographique' section with the citation: 'Côté, Marie-Hélène (2012). « Conversation libre cqhc2l, Grande-Rivière ». Cité dans Phonologie du français contemporain : corpus Québec (PFC-Québec)'. Below this is a 'Statistiques' section showing '12 458 mots' and '1 492 mots uniques'. At the bottom right is a 'Métadonnées' section with fields for 'Identifiant', 'Titre', and 'Corpus'.



Les exemples présentés ici proviennent du fichier [PFC-Québec_Grande-Rivière_cqhdb1](#).

Sommaire

- [Lecteur audio par défaut](#)
- [Lecteur audio avancé](#)

Lecteur audio par défaut

Par défaut, le lecteur audio est minimisé pour laisser plus de place à la transcription.

Cliquez sur **Lecture**  pour écouter l'extrait audio et sur **Pause**  pour mettre

¹Unité de base d'un corpus (qui se définit comme une collection de documents), par exemple une œuvre littéraire, une chanson, une bande dessinée, un article de journal, une fiche dialectologique ou encore une entrevue. Dans le cadre du Fonds, une distinction est faite entre le contenu textuel du document (transcription) et le contenu binaire (version numérisée de l'original : image ou enregistrement audio).

²Ensemble de textes établi selon un principe de documentation exhaustive, un critère thématique ou exemplaire en vue de leur étude linguistique (définition tirée d'Usito).

sur pause l'extrait audio. Dans le bas du lecteur, la barre verte indique la progression de la lecture du fichier audio.

00:18:58 JL
des petits patins de gars. Fait qu'il y avait pas de coche.

00:19:01 JC2
Ça t'intéresse-tu toi les, les cours de danse qu'ils donnent?

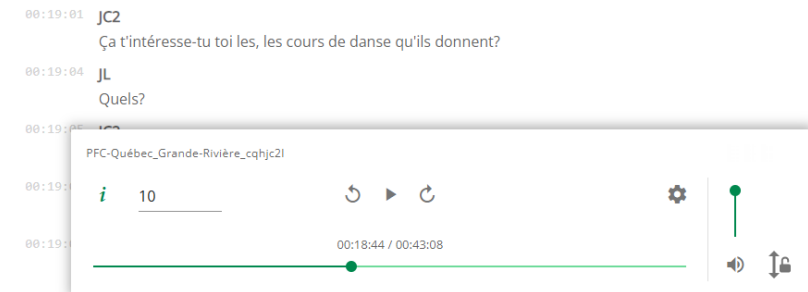
00:19: PFC-Québec_Grande-Rivière_cqhjc2l







Cliquez sur le chevron vers le haut ^ pour activer l'affichage avancé du lecteur audio.

Lecteur audio avancé

Le lecteur audio avancé offre plus de contrôle sur la lecture du fichier audio.



Options	Description
Nombre de secondes recul/ avance rapide	Par défaut, défini à 10 secondes. On peut modifier cette valeur pour réduire ou augmenter le recul rapide et l'avance rapide.
Recul rapide	Permet de revenir en arrière d'une certain nombre de secondes (par défaut 10 sec.).
Avance rapide	Permet d'avancer d'une certain nombre de secondes (par défaut 10 sec.).
Volume (après avoir cliqué sur )	Permet de réduire ou d'augmenter le volume de lecture du fichier audio. En cliquant sur l'icône  , on rend muette la lecture; en cliquant à nouveau, le son revient.

Options	Description
Verrouillage du défilement automatique (après avoir cliqué sur )	Cliquer sur l'icône  permet de figer le défilement de la page. Si on active cette option, la page ne défilera plus même si la lecture du fichier audio se poursuit. Cliquer à nouveau sur l'icône permet de réactiver le défilement automatique.
Barre de progression	Permet de se déplacer sur la ligne de temps du fichier audio.

Voir aussi

- [Documents \(onglet Transcription avec audio\)](#)

Dernière modification : 27 mai, 2025

Lexique

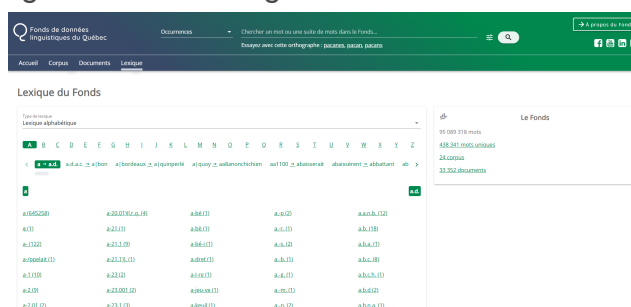
Le Fonds de données linguistiques du Québec est composé (pour l'instant!) de plus de 95 millions de **mots**¹, dont plus de 435 000 mots uniques, répartis dans plus de 33 000 **documents**² provenant de 24 **corpus**³.

Sommaire

- [Lexique du Fonds](#)
- [Lexique du corpus](#)
- [Lexique du document](#)

Lexique du Fonds

La page **Lexique du Fonds** montre la liste de tous les mots du Fonds de données linguistiques du Québec, tels qu'attestés dans les documents de tous les corpus. Quelques statistiques générales sont également affichées.



¹Dans le contexte de la plateforme FDLQ, mot n'est pas utilisé dans le sens d'« unité lexicale » qu'il prend en linguistique, mais plutôt comme synonyme de mot-forme (mot séparé par des espaces) ou encore de token. Les mots faisant partie du lexique d'un document ou d'un corpus correspondent ainsi à l'ensemble des tokens qu'on y trouve; suivant cette logique, pomme de terre correspond à trois tokens.

²Unité de base d'un corpus (qui se définit comme une collection de documents), par exemple une œuvre littéraire, une chanson, une bande dessinée, un article de journal, une fiche dialectologique ou encore une entrevue. Dans le cadre du Fonds, une distinction est faite entre le contenu textuel du document (transcription) et le contenu binaire (version numérisée de l'original : image ou enregistrement audio).

³Ensemble de textes établi selon un principe de documentation exhaustive, un critère thématique ou exemplaire en vue de leur étude linguistique (définition tirée d'Usito).

attestations du mot en question dans le document (un filtre rapide est appliqué sur le document).

The screenshot shows a digital library interface for the 'Fonds de données linguistiques du Québec'. The search results for 'Charlevoix_1744 [1720]_Histoire' are displayed. The interface includes a search bar, a list of search results, and a detailed view of the document with a word frequency table and a reference section.

LENDRE DU DOCUMENT TEXTE INTEGRAL DOCUMENT ORIGINAL

Titre de l'œuvre
Lecteur de l'abulique

A B C D E F G H I J K L M N O P Q R S T U V W X Y Z

← active → active

active	active	active	active	active
active(1)	active(1)	active(1)	active(1)	active(1)
active(1)	active(1)	active(1)	active(1)	active(1)
active(1)	active(1)	active(1)	active(1)	active(1)
active(1)	active(1)	active(1)	active(1)	active(1)

Référence bibliographique

Charlevoix, Pierre-François-Xavier de. Histoire et description générale de la Nouvelle-France, avec le journal historique d'un voyage fait par ordre du Roi dans l'Amérique septentrionale. Paris, Pierre-Thomas Collet, 1744. (Édition de 1720)

Statistiques

220 mots
227 mots-clés

Métadonnées

Identifiant: Charlevoix_1744 [1720]_Histoire
Titre: Histoire et description générale de la Nouvelle-France, avec le journal historique

Voir aussi

- [Corpus](#)
- [Documents](#)

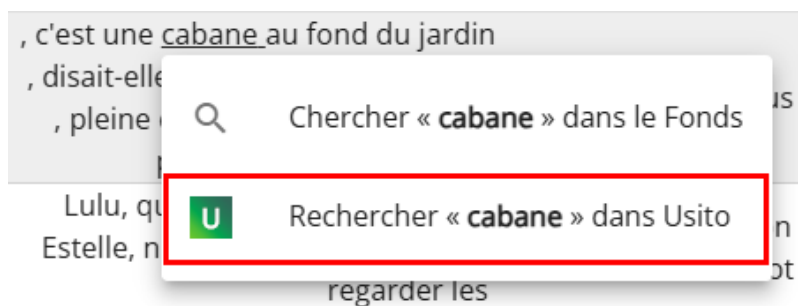
Dernière modification : 27 mai, 2025

Intégration avec *Usito*

Il existe une intégration entre *Usito* et le Fonds de données linguistiques du Québec. [Usito](#) est un dictionnaire conçu au Québec.

Du FDLQ vers le dictionnaire *Usito*

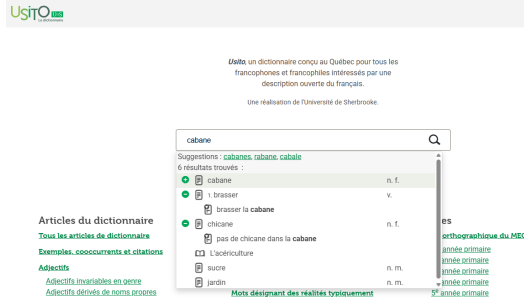
1. À partir de la **concordance**¹ du FDLQ, dans les résultats de recherche, passez avec la souris sur les **mots**² : ils deviennent soulignés.
2. Cliquez sur n'importe quel mot souligné, dans le contexte avant ou dans le contexte après, ou sur le mot pivot lui-même.
3. Dans la fenêtre contextuelle qui apparaît, sélectionnez l'option **Rechercher « x » dans Usito**.



4. Le dictionnaire *Usito* s'ouvre automatiquement dans un nouvel onglet de votre navigateur Web. Le champ de recherche est déjà saisi du mot en question.
5. Pressez sur la touche **Entrée** ou cliquez sur une entrée du dictionnaire proposée.

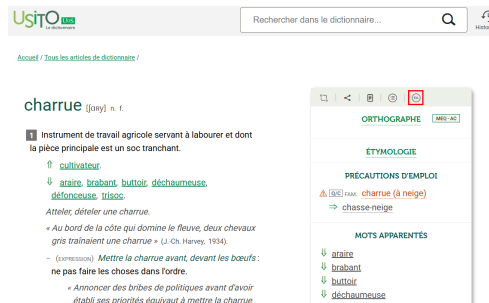
¹Liste des résultats d'une requête textuelle où les éléments recherchés sont affichés dans leur contexte.

²Dans le contexte de la plateforme FDLQ, mot n'est pas utilisé dans le sens d'« unité lexicale » qu'il prend en linguistique, mais plutôt comme synonyme de mot-forme (mot séparé par des espaces) ou encore de token. Les mots faisant partie du lexique d'un document ou d'un corpus correspondent ainsi à l'ensemble des tokens qu'on y trouve; suivant cette logique, pomme de terre correspond à trois tokens.



Du dictionnaire *Usito* vers le FDLQ

1. À partir du site du dictionnaire [Usito](https://www.usito.ca/), effectuez une recherche.
2. Dans l'article du mot recherché, cliquez sur l'icône **Exemplification** dans le car-touche qui se trouve à droite de l'écran.



3. La page **Exemplification de l'article** s'affiche. Repérez la section **Recherches associées**, au bas de la page.



4. Le premier champ permet de déterminer si le système doit **Chercher toutes les formes du mot** ou s'il doit **Chercher exactement le mot**.

5. Le second champ contient le **mot**¹ à rechercher :
 - Si l'option **Chercher toutes les formes du mot** a été sélectionnée, on ne peut rien changer ici;
 - Si l'option **Chercher exactement le mot** a été sélectionnée, choisissez sur quelle forme du mot la recherche doit se faire dans le FDLQ. Par exemple, pour le mot *charrue*, on peut chercher spécifiquement la forme *charrue* ou la forme *charrues*.
6. Indiquez dans quel **corpus**² d'œuvres effectuer la recherche. Les choix disponibles sont :
 - Littérature québécoise (littérature générale et jeunesse);
 - Littérature générale québécoise;
 - Littérature jeunesse québécoise.
7. Cliquez sur le bouton **Lancer la recherche**.
8. Automatiquement, un nouvel onglet du navigateur est ouvert et les résultats de la recherche s'affichent dans la concordance.

Voir aussi

- [Résultats de recherche](#)
- [Moteur de recherche](#)

Dernière modification : 28 mai, 2025

¹Dans le contexte de la plateforme FDLQ, mot n'est pas utilisé dans le sens d'« unité lexicale » qu'il prend en linguistique, mais plutôt comme synonyme de mot-forme (mot séparé par des espaces) ou encore de token. Les mots faisant partie du lexique d'un document ou d'un corpus correspondent ainsi à l'ensemble des tokens qu'on y trouve; suivant cette logique, pomme de terre correspond à trois tokens.

²Ensemble de textes établi selon un principe de documentation exhaustive, un critère thématique ou exemplaire en vue de leur étude linguistique (définition tirée d'Usito).

Glossaire

Concordance

Liste des résultats d'une requête textuelle où les éléments recherchés sont affichés dans leur contexte.

Cooccurrence

Ensemble d'unités lexicales qui apparaissent fréquemment en combinaison dans un énoncé, et qui sont liées par un rapport de proximité syntaxique et de relative dépendance. Définition tirée d'[Usito](#).

Corpus

Ensemble de textes établi selon un principe de documentation exhaustive, un critère thématique ou exemplaire en vue de leur étude linguistique. Définition tirée d'[Usito](#).

Document

Unité de base d'un corpus (qui se définit comme une collection de documents), par exemple une œuvre littéraire, une chanson, une bande dessinée, un article de journal, une fiche dialectologique ou encore une entrevue. Dans le cadre du Fonds, une distinction est faite entre le contenu textuel du document (transcription) et le contenu binaire (version numérisée de l'original : image ou enregistrement audio).

Joker (métacaractère)

Caractère qui permet de remplacer un ou plusieurs autres caractères dans une commande, une recherche ou une requête. Définition tirée du [Grand dictionnaire terminologique](#).

Lemme

Forme canonique d'un mot variable (au masculin singulier pour les noms et les adjectifs, à l'infinitif pour les verbes). Définition tirée d'[Usito](#).

Métadonnée

Donnée qui renseigne sur la nature de certaines autres données dans le but d'en faciliter la compréhension et la gestion. Définition tirée du [Grand dictionnaire terminologique](#).

Mot

Dans le contexte de la plateforme FDLQ, *mot* n'est pas utilisé dans le sens d'« unité lexicale » qu'il prend en linguistique, mais plutôt comme synonyme de *mot-forme* (mot séparé par des espaces) ou encore de *token*. Les mots faisant partie du lexique d'un document ou d'un corpus correspondent ainsi à l'ensemble des tokens qu'on y trouve; suivant cette logique, *pomme de terre* correspond à trois tokens.

Occurrence

Apparition d'une unité linguistique dans le discours; par métonymie, cette unité. Définition tirée d'[Usito](#).

Pivot (mot pivot, KWIC)

Dans une concordance, le mot (ou la suite de mots) choisi pour la recherche, accompagné de son contexte et mis en évidence typographiquement. L'acronyme KWIC signifie *KeyWord In Context*, littéralement « mot-clé dans son contexte ».

Token

Unité textuelle minimale qui résulte de la segmentation automatique du texte d'un document. Il peut s'agir de mots ou de formes de mots (*bleuetière, jouaient, vingt-*

cinq, tire-bouchon, etc.), de chiffres (*500.000, 1608*, etc.) ou d'abréviations (2^e « deuxième », *M.* « monsieur », *av.* « avant », etc.).

Dernière modification : 27 mai, 2025