# Guide d'utilisation de SyllabO+

Version du 4 aout 2016





## Guide d'utilisation

Avant de télécharger la base de données entière, il est recommandé d'avoir une bonne idée de l'information dont vous avez besoin, car les possibilités sont nombreuses et certaines de nos bases de données sont assez imposantes (177 MB pour la base de données des syllabes) ! *SyllabO*+ contient en effet une grande quantité de données et il est possible qu'une partie seulement soit pertinente pour vos objectifs.

### Instructions

1. Allez à <a href="http://syllabo.speechneurolab.ca/">http://syllabo.speechneurolab.ca/</a> et après avoir répondu à quelques brèves questions, vous aurez accès aux bases de données <a href="https://syllabo.speechneurolab.ca/">Syllabo</a>+. Vos réponses demeureront strictement confidentielles et ne seront utilisées qu'à des fins de rapports pour les agences de financements et autres partenaires du projet afin de quantifier l'utilisation de nos bases de données.

Accéder à SyllabO

2. Vous pourrez alors **choisir** entre les **bases de données complètes** (informations extraites à partir de tous les enregistrements audio), que ce soit pour les syllabes ou pour les phones, ou les bases de données spécifiques. Les bases de données complètes contiennent les statistiques d'utilisation globales, c'est-à-dire pour tous les contextes et tous les locuteurs (peu importe leur sexe et leur âge). Elles représentent ainsi l'usage général des locuteurs adultes du français québécois en terme de syllabes, séquences de syllabes et phones.

L'image ci-dessous illustre la base de données complète des syllabes.

## Base de données des syllabes

Lorsque vous téléchargez toute la base de données de syllabes, quatre fichiers sont inclus dans un document compressé (.ZIP): soit les syllabes, les paires, les triades et les tétrades.

Des versions alternatives de la base de données sont également disponibles ci-dessous aux format .XLSX (avec les caractères spéciaux API déjà encodés) et au format .CVS sans caractères spéciaux (transcription avec l'alphabet SAMPA – Speech Assessment Methods Phonetic Alphabet).

Télécharger la base de données complète de SYLLABES

Vous pouvez également choisir de télécharger une **base de données spécifique**, laquelle contiendra alors les statistiques d'utilisation extraites à partir d'un sous-ensemble des enregistrements audio. Les statistiques d'utilisation seront alors compilées en fonction des paramètres suivants, individuellement ou combinés: <u>contexte de communication</u> (formel, informel), <u>âge</u>, et <u>genre</u> des locuteurs. Il est ainsi possible de télécharger les statistiques d'utilisation des syllabes seulement pour les jeunes adultes, ou pour les femmes en contexte formel, etc. À noter que cette option n'est disponible que pour les syllabes pour le moment. À noter également que les résultats ne seront pas affichés en ligne mais téléchargeable. L'option de recherche en ligne devrait être offerte dans le futur (à suivre!). La recherche dans la base de données des phones sera disponible à l'automne 2016.

Pour les téléchargements de ces bases de données spécifiques, la fenêtre suivante est utilisée :

## Téléchargement selon des paramètres de recherche

vedillez identifier les parametres particuliers.					
Туре*	Syllabe				
Contexte d'élocution	Tous				
Genre des locuteurs	Tous				
Âge minimum	Optionnel				
Âge maximum	Optionnel				
Code de sécurité*	eciw				
Télécharger la base de données particulière					

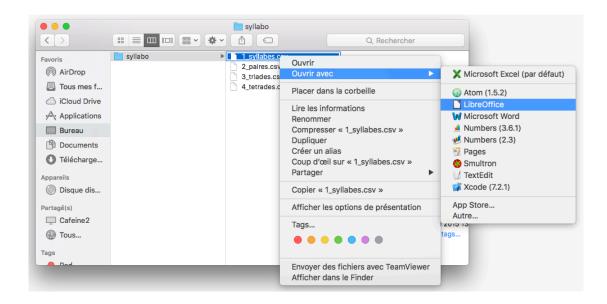
Vouillez identifier les paramètres particuliers

Cliquez pour télécharger la base de données désirée.

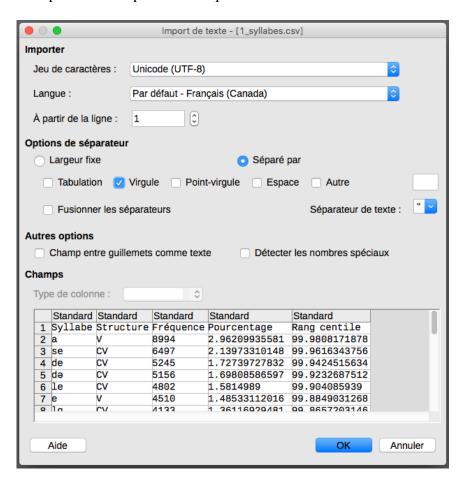
- 3. Si vous avez **téléchargé une base de données** complète, vous obtiendrez quatre fichiers différents. Pour les syllabes, vous obtiendrez un fichier pour les syllabes individuelles et leurs statistiques, un pour les paires de syllabes et leurs statistiques, un pour les groupes de 3 syllabes (triades) et leurs statistiques et un pour les groupes de 4 syllabes (tétrades) et leurs statistiques. La taille pour l'ensemble des fichiers est de 177 MB. Pour les phones, vous obtiendrez un fichier pour les phones individuels et leurs statistiques, un pour les paires de phones (diphones) et leurs statistiques, un pour les groupes de 3 phones (triades) et leurs statistiques et un pour les groupes de 4 phones (tétrades) et leurs statistiques. La taille pour l'ensemble des fichiers est de 22 MB.
- 4. Si vous avez **téléchargé une base de données spécifique**, vous obtiendrez seulement le fichier sélectionné (syllabes, paires de syllabes ou triades).

- 5. Tous les fichiers sont au **format CSV** en API (Comma Separated Values), qui est une manière simple de stocker des données tabulaires en fichiers textes bruts. À noter que deux autres formats sont maintenant disponibles (avec les caractères spéciaux API déjà encodés) et au format .CVS sans caractères spéciaux (transcription avec l'alphabet SAMPA Speech Assessment Methods Phonetic Alphabet). Ces formats ne sont toutefois disponibles que pour le téléchargement des bases de données complètes.
- 6. Si vous téléchargez les fichiers .CVS API, le plus simple est de les ouvrir avec le logiciel LibreOffice (tableur). Nous recommandons l'utilisation de ce logiciel parce qu'il permet de traiter les caractères phonétiques (API) correctement. Microsoft Excel ne traite pas ces caractères spéciaux correctement. À noter: LibreOffice est un logiciel gratuit et Open source, disponible pour téléchargement à <a href="http://fr.libreoffice.org/">http://fr.libreoffice.org/</a>. Pour la version en français, téléchargez l'installateur principal, puis l'interface en français. (Allez à <a href="http://www.libreoffice.org/">http://www.libreoffice.org/</a> pour la version en anglais).

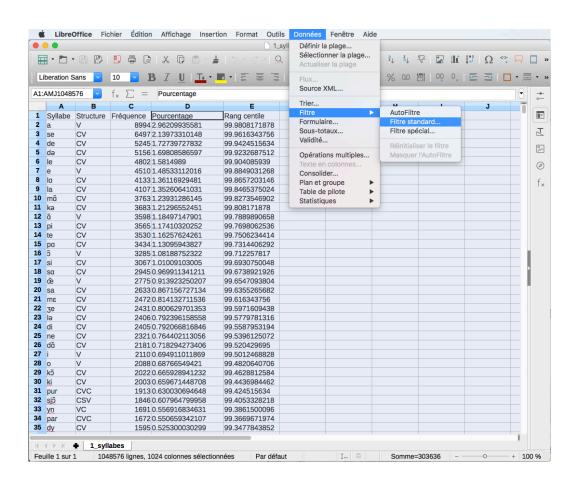




7. Sélectionnez **Unicode (UTF-8)** comme jeu de caractères. Sélectionnez **Virgule** comme option de séparateur. Cliquez sur **OK** pour ouvrir.



- Une fois le fichier ouvert, vous pouvez sélectionner les colonnes pertinentes pour votre projet.
  Masquez ou supprimez les colonnes non désirées.
- 9. Filtrez les données selon vos besoins : sélectionnez toutes les données, allez au menu, puis cliquez sur *Données Filtre Filtre Standard*.



10. Vous pouvez alors utiliser le(s) filtre(s) de votre choix pour obtenir l'information désirée.Indiquez quelle colonne doit être filtrée selon quel paramètre. Par exemple :

		Filtre standard			
Critères de filt	tre				
Opérateur	Nom de champ	Condition		Valeur	
	Structure	=	•	CV	<u> </u>
ET 💲	Fréquence	>	<b>\$</b>	150	_
•	- aucun(e)-	=	•		~
•	- aucun(e)-	=	0		~
▶ Options					
Aide				ОК	Annuler

### 11. Vous obtenez vos **résultats**!

### Exemple d'élaboration de stimuli

Voici un exemple qui illustre comment des stimuli spécifiques peuvent être construits à partir de *SyllabO*. À partir du tableau des groupes de 3 syllabes, ces colonnes ont été sélectionnées :

- Triade (transcription en API)
- Structure 1<sup>ère</sup> syllabe
- Structure 2<sup>e</sup> syllabe
- Structure 3<sup>e</sup> syllabe
- Rang centile 1<sup>ère</sup> syllabe
- Rang centile 2<sup>e</sup> syllabe
- Rang centile 3<sup>e</sup> syllabe

Le filtre « = CV » a été appliqué à toutes les colonnes « structure » et le filtre « >= 95 » a été appliqué à toutes les colonnes « rang centile ». Le tableau résultant de ces filtres (*Tableau 1*) est une sélection de triades de syllabes, où chaque syllabe possède une structure simple « consonne – voyelle » et possède une fréquence élevée puisqu'elle est située entre le 95<sup>e</sup> et le 100<sup>e</sup> rang centile. Les groupes de 3 syllabes issus de ce tableau peuvent être utilisés comme stimuli dans une étude expérimentale.

Triade	Structure 1	Rang centile 1	Structure 2	Rang centile 2	Structure 3	Rang centile 3
kə mã se	CV	99.0216765778	CV	99.8273546902	CV	99.9616343756
vu za ve	CV	98.849031268	CV	98.6188375216	CV	98.273546902
mã do ne	CV	99.8273546902	CV	97.602148475	CV	99.5396125072
le zã fã	CV	99.904085939	CV	98.4270093996	CV	97.0650297334
di fe rã	CV	99.5587953194	CV	96.451179743	CV	98.0625359678
•••						

Voilà! Bonnes recherches, amusez-vous bien, et restez branchés pour les développement futurs (base de phones, recherches en ligne améliorées et base de données lexicales). Merci de citer notre article dans vos publications et rapports découlant de l'utilisation de SyllabO+: (Bédard et al., Révisions soumises):

Bédard, P., Audet, A.-M., Drouin, P., Roy, J. P., Rivard, J., & Tremblay, P. (Révisions soumises). SyllabO+: a new tool to investigate sublexical phenomena in Québec spoken French. *Behavioral Research Methods*.