

Diagrammes de Sankey et réseaux pour comparer les éditions d'un recueil de textes

Blanche d'Aubigny¹, and Philippe Gambette² 

¹ Sorbonne Université, CNRS, Centre d'Étude de la Langue et des Littératures Françaises, CELLF, F-75006 Paris, France

² Univ Gustave Eiffel, CNRS, LIGM, F-77454 Marne-la-Vallée, France

Abstract

In this article, we focus on the interpretation of order changes in successive editions of a collection of texts: do these changes reflect a random rearrangement or do they reveal certain patterns that might lead to interpretive hypotheses? We provide a web interface that allows users to visualize these changes in order using Sankey diagrams (networks with edges of varying thickness), enhanced to display other structural characteristics of the compared editions. These visualizations can be combined to analyze more than two successive editions, and enhanced with interactive features to provide a multi-scale comparison, ranging from the order of texts within the collection, to differences in content between the texts themselves. A few examples drawn from the poetic works of Marceline Desbordes-Valmore illustrate the value of these digital exploration tools for text collection genetics.

Mots-clefs: génétique des recueils, réseaux de textes, diagrammes de Sankey, tau de Kendall

Keywords: text collection genetics, text networks, Sankey diagrams, Kendall's tau

1 Introduction

Gérard Genette considère dans *Seuils* que dans un recueil de poèmes brefs, « l'effet de séquence ou de progression est habituellement très faible, et l'ordre est le plus souvent arbitraire » [14]. Dans le cas où des recueils de poèmes, ou de textes d'autres genres littéraires (nouvelles, maximes, contes, fables, etc.) présentent des ordres différents dans leurs éditions successives, s'agit-il de réorganisations arbitraires ou existe-t-il des critères, tenant par exemple à la progression globale du recueil, ou à des contraintes structurelles plus locales, pour expliquer les différences d'ordre ?

Dans cette contribution, nous montrons de quelle manière des outils de visualisation de réseaux peuvent aider à explorer ces questions. Nous proposons des interfaces web qui permettent à la fois une observation globale des changements d'ordre entre deux versions d'un recueil de textes, et des focus sur des phénomènes plus locaux liés par exemple à la structuration du recueil en parties. Ces interfaces peuvent facilement être générées à partir des tables des matières de deux éditions d'un recueil, organisées dans un document tableur.

2 Visualisations de changements d'ordre

Les outils de visualisation et d'analyse de données tels que les plateformes web RAW-Graphs [16] et NetSage [21] proposent des visualisations en réseaux connues en anglais sous le nom de *slope graphs* pour visualiser les différences d'ordre entre deux listes contenant les mêmes éléments, ou de *bump charts* quand il s'agit de représenter l'évolution des différences d'ordre sur plus de deux listes d'éléments [23].

Dans le cas où l'on souhaite également représenter des évolutions de taille de ces éléments ordonnés, on peut avoir recours à des *diagrammes alluviaux* [16], aussi appelés *diagrammes de flux Sankey* [21]. Ces derniers ont déjà été utilisés en humanités numériques pour représenter visuellement les différences entre deux éditions de *Cloud Atlas* de David Mitchell mettant en relief des réorganisations de plusieurs segments du roman [13]. Nous proposons ici de faciliter leur construction pour des recueils de textes tout en tirant pleinement parti d'une visualisation interactive, et de la coloration, pour faciliter l'analyse des différences d'ordre et de contenu entre les recueils.

3 Corpus utilisés

Nous nous intéressons à deux corpus de textes de Marceline Desbordes-Valmore. Le premier est constitué par les éditions de 1920 et de 1927 de l'anthologie de poèmes de Marceline Desbordes-Valmore (1786-1859) traduits en allemand, en annexe de la biographie de la poète par Stefan Zweig, intitulée *Marceline Desbordes-Valmore : das Lebensbild einer Dichterin* [10 ; 11].

Le second est constitué par les premiers recueils de poèmes de Marceline Desbordes-Valmore, publiés de 1819 à 1830 [5 ; 6 ; 7 ; 8 ; 9]. Paru en 1819, le premier recueil au titre programmatique, *Élégies, Marie et romances*, s'organise en trois sections, autour de la nouvelle « Marie ». Celle-ci comprend huit pièces versifiées, qui disparaissent définitivement dans l'édition suivante. Le recueil des *Poésies* de 1820 conserve les élégies et les romances de 1819, à l'exception de « Médor », et leur adjoint vingt-quatre nouveaux poèmes. Les *Poésies* de 1822 retranchent deux autres romances (« Le Réveil créole » et « Jone et Sophie ») et l'édition est augmentée de seize créations. En 1825, paraissent les *Élégies et Poésies nouvelles*, où figurent cinquante-sept pièces absentes des recueils précédents : vingt élégies, quatre idylles, vingt-trois romances, six contes et trois « poésies diverses ». Dans le recueil collectif des *Poésies* de 1830, seuls onze poèmes des éditions de 1822 et 1825 ne sont pas repris et soixante-deux nouveaux apparaissent. Au cours des éditions successives, certaines « élégies » deviennent des « idylles », d'autres sont déplacées dans la catégorie « mélanges », qui est ensuite renommée « poésies diverses » et où l'on retrouve des poèmes auparavant classés parmi les « contes ».

L'ordre des poèmes dans ces recueils n'a pour le moment pas fait l'objet d'une étude approfondie, alors que, publié un an avant les *Méditations poétiques* de Lamartine, *Élégies, Marie et romances* occupe une place particulière dans le mouvement romantique français en poésie [18 ; 19]. La correspondance de Marceline Desbordes-Valmore indique qu'après avoir confié à son éditeur François Louis la tâche d'ordonner les poèmes dans ce premier recueil, dans une lettre du 26 avril 1818, elle fixe, dans une autre lettre au même destinataire, le 3 octobre 1818, l'ordre des élégies [4]. Dans sa critique dans les *Annales de la littérature et des arts*, en 1821, Jacques Ancelot déplore l'absence d'une progression dans la deuxième édition du recueil [2], ce qui a pu pousser la poète à revoir de nouveau en profondeur l'ordre des poèmes dans l'édition suivante publiée l'année suivante.

Par ailleurs, dans son étude sur le cycle des poèmes à Délie dans ces recueils [12], Edie Dixon-Young accorde une attention particulière au déplacement en première position, à partir de 1822, du poème auparavant deuxième [12]. Elle indique que pour privilégier l'un ou l'autre des deux ordres concurrents, les précédentes études de ces poèmes [1 ; 3] s'étaient surtout fondées sur des considérations biographiques, comme cela est aussi fait dans un article évoquant l'ordre des poèmes d'*Élégies, Marie et romances* [17]. Elle-même se fonde sur des considérations esthétiques pour montrer les contrastes entre les deux premiers poèmes du cycle (quel que soit leur ordre) et le troisième.

4 Méthodologie

4.1 Un outil de comparaison de deux recueils de textes par un diagramme de Sankey

Nous proposons d'une part une interface web, codée sous licence libre en HTML, CSS et Javascript¹, pour représenter deux éditions d'un même recueil de textes par un diagramme de Sankey facilitant l'analyse des différences d'ordre des textes. Les données relatives au contenu des deux éditions sont structurées dans un document tableur : elles peuvent être établies directement à partir de la table des matières, en repérant ensuite manuellement, ou semi-automatiquement à partir des titres, à quel texte de la première édition correspond éventuellement chaque texte de la seconde.

Comme dans [13], chaque recueil est alors représenté par un ensemble de rectangles verticaux, qui chacun représente un texte du recueil : la hauteur du rectangle dépend de la longueur du texte et les rectangles correspondant aux deux versions du texte dans les deux éditions du recueil sont reliés. Nous ajoutons la possibilité de répartir les textes en catégories, représentées par des couleurs dans la visualisation, qui peuvent correspondre à des parties du recueil ou à d'autres spécificités des textes.

Un formulaire permet de charger les URL de deux fichiers au format CSV, l'un pour indiquer les données relatives aux deux versions du recueil, l'autre pour les métadonnées relatives à ces deux versions. Un permalien intégrant les URL des deux fichiers CSV permet alors la diffusion des visualisations ainsi générées à partir des fichiers chargés.

4.2 Une visualisation en réseau des premiers recueils de poèmes de Desbordes-Valmore

Pour représenter les recueils de poèmes de Desbordes-Valmore publiés de 1819 à 1830, nous choisissons de généraliser le principe du diagramme de Sankey aux ensembles de recueils successifs. Mais nous le simplifions en uniformisant épaisseur des liens et taille des points représentant les textes², pour plus de lisibilité. Le recueil *Élégies, Marie et romances* (1819) apparaît donc comme un ensemble horizontal de points qui représentent les poèmes qu'il contient. Les éditions suivantes de ce recueil, les *Poésies* de 1820 et de 1822, adoptent la même forme, en dessous, tout comme les *Élégies et poésies nouvelles* de 1825, qui ne contiennent que des poèmes inédits et les *Poésies* de 1830 en deux tomes réunissant le tout, avec de nouveau des poèmes inédits réunis pour l'essentiel à la fin du deuxième tome, donc en bas à droite de la visualisation.

Nous proposons une visualisation interactive, en permettant, par le survol avec la souris d'un lien reliant deux éditions successives d'un poème, de mettre en gras tous les liens situés entre éditions successives de ce poème. Les titres de ces deux versions liées par le lien survolé sont alors affichés en haut de la visualisation. Passer la souris sur un point coloré correspondant à un poème permet d'en lire le titre et d'épaissir les liens qui le relie à ses éditions dans les autres recueils. Cliquer sur ce point permet d'accéder à une version en ligne du texte du poème, et cliquer sur un lien ouvre une comparaison entre les deux versions du poème (utilisant parfois le texte obtenu par reconnaissance automatique de caractères, non relu), sous forme d'un alignement automatique construit avec le logiciel MEDITE [22], appelé de façon répétée grâce à un script Python³.

1. <https://philippegambette.github.io/txtCompare/sankeyCompare>

2. <https://philippegambette.github.io/txtCompare/recueilsMDV/>

3. <https://github.com/PhilippeGambette/txtCompare/tree/master/pairwiseMeditate>.

5 Résultats et discussion

Dans la visualisation en diagramme de Sankey du premier corpus, montrée en figure 2⁴, la couleur cyan indique les titres des poèmes de l’anthologie en annexe de la biographie de Marceline Desbordes-Valmore (1786-1859) publiée en 1921 par Stefan Zweig, qui ne sont pas repris dans la réédition de 1927. De même, sont colorées en bleu deux nouvelles traductions, dues à Friderike Zweig. La couleur des liens est automatiquement assombrie si le titre diffère, ce qui permet de détecter une seule variante de titre, l’ajout de la précision « (vor seiner Reise in das Pensionat) » au poème « An meinen Sohn ». La visualisation montre un nombre conséquent de différences d’ordre, qui sont analysées par Ina Schabert comme une volonté de suivre davantage la thématique biographique du texte de Zweig, au détriment d’éléments plus variés de l’œuvre de Desbordes-Valmore qu’avait choisies la traductrice Etzel-Kühn, décédée peu après la publication de la première édition [20].

Nous avons choisi dans la capture d’écran de la figure 1 de mettre en relief l’évolution de la place du poème « Le Sommeil de Julien » pour montrer qu’il est d’abord publié comme une romance située plutôt en fin de recueil, avant de devenir une des dernières élégies en 1820, puis de devenir la première romance des recueils à la fois de 1822 et de 1830. L’utilisation des couleurs pour indiquer les sections génériques de ces divers recueils (qui n’apparaîtront plus dans les recueils suivants de la poète) permet donc de mettre en évidence la fluidité générique de certaines élégies de 1819 et 1820 qui deviennent idylles à partir de 1822 et de trois idylles de 1825 intégrées dans les romances en 1830.

De même, les cinq contes en vers repris du recueil de 1825 ne sont plus distingués des autres poésies diverses de 1830, même s’ils restent placés de façon contigüe, dans le même ordre. Cela peut conduire à lire comme des contes les poèmes placés juste avant, tirés des poésies diverses de l’édition de 1822, « Conte imité de l’arabe », « La Mouche bleue », « L’Écolier » et « Conte d’enfant », ou de l’édition de 1825, « Le Billet d’une amie » et « L’Autruche et le Pélican ».

Parmi ces six poèmes, c’est peut-être pour « Le Billet d’une amie » que ce genre peut sembler le plus surprenant, en apportant justement une autre expérience de lecture. Les différences entre le texte des deux éditions du poème, visibles en cliquant sur le lien entre ces deux éditions, montrent un changement dans le dernier vers : « Je le déteste... eh quoi ! l’aimais-je encor ? » devient « Je hais l’amour... Eh quoi ! l’aimais-je encor ? ». Alors que le pronom « le » de l’édition de 1825 pouvait être ambigu à la lecture des vers précédents (désigne-t-il l’amour, ou l’amant qui a trahi ?), la version de 1830 écarte la seconde hypothèse. Elle introduit en revanche une ambiguïté sur le pronom « l’ » final, autorisant une lecture plus générale du dernier vers (« aimais-je encore l’amour ? ») qui sonne alors comme une morale à ce conte.

6 Conclusion et perspectives

L’exemple complexe des premiers recueils de poèmes de Marceline Desbordes-Valmore montre comment les outils de visualisation présentés dans cette contribution combinent plusieurs possibilités de lecture, guidée par les couleurs ou par l’exploration à la souris des fonctionnalités d’interactivité et de retour au texte, pour des analyses approfondies des variations d’ordre. Il montre aussi de façon plus globale et percutante, par les segments parallèles, comment l’ordre des poèmes des éditions de 1822 et 1825 est stabilisé en 1830 : après les hésitations, parfois empreintes de modestie, sur l’ordre des premiers recueils de 1819 et 1820, Marceline Desbordes-Valmore n’est plus une actrice qui écrit des vers, mais une poète qui confirme ses choix de structuration en recueils.

4. Version interactive sur <https://philippegambette.github.io/txtCompare/sankeyCompare/?id=4>.

Il pourrait être intéressant de déterminer à quel point les perturbations d'ordre que l'on peut constater sur les visualisations conservent l'ordre initial des textes du recueil, ou au contraire, tendent tout particulièrement à le renverser. Des fonctionnalités de calcul statistique pourraient alors être ajoutées dans l'interface, si l'on se limite à analyser les textes du premier recueil qui ont été repris, de façon unique, dans le second recueil : ceci permet de modéliser la différence entre les deux ordres par une permutation, dont il est alors possible de compter le nombre d'inversions, c'est-à-dire le nombre de paires de textes qui ont vu leur ordre s'inverser. Ce nombre pourra alors être comparé avec la distribution des nombres d'inversions de permutations aléatoires pour déterminer si l'ordre initial, ou l'ordre inverse, est plus conservé dans le second ordre des textes que ce qu'on pourrait attendre si le nouvel ordre avait été tiré au hasard. Le *tau* de Kendall [15], qui fournit un score compris entre -1 et 1, proche de 1 en cas de respect de l'ordre original, et proche de -1 en cas de proximité avec l'inverse de l'ordre original, pourrait aussi, plus simplement, être calculé.

Références

- [1] AMBRIÈRE, Francis. *Le Siècle des Valmore: Marceline Desbordes-Valmore et les siens*. T. 1. Seuil, 1987, p. 190-192.
- [2] ANCELOT, Jacques. « Poésies de M^{me} Desbordes-Valmore ». In : *Annales de la littérature et des arts* 2, no. 19 (1821). <https://gallica.bnf.fr/ark:/12148/bpt6k9631888h/f224.item>, p. 208-209.
- [3] BERTRAND, Marc. *Les Œuvres poétiques de Marceline Desbordes-Valmore*. T. 1. Presses Universitaires de Grenoble, 1973, p. 266-267.
- [4] CAVALLUCCI, Giacomo. « Poésies de 1819 ». In : *Bibliographie critique de Marceline Desbordes-Valmore, d'après des documents inédits*. T. I. Pironti, 1936, p. 8-10.
- [5] DESBORDES, Marceline. *Élégies, Marie et romances*. Paris : François Louis, 1819.
- [6] DESBORDES-VALMORE, Marceline. *Élégies et poésies nouvelles*. Paris : Ladvocat, 1825.
- [7] DESBORDES-VALMORE, Marceline. *Poésies*. Paris : François Louis, 1820.
- [8] DESBORDES-VALMORE, Marceline. *Poésies*. Paris : Théophile Grandin, 1822.
- [9] DESBORDES-VALMORE, Marceline. *Poésies*. Paris : Boulland, 1830.
- [10] DESBORDES-VALMORE, Marceline et ETZEL-KÜHN, Gisela (trad.) « Gedichte ». In : ZWEIG, Stefan. *Marceline Desbordes-Valmore: das Lebensbild einer Dichterin*. Insel-Verlag, 1920, p. 69-136.
- [11] DESBORDES-VALMORE, Marceline, ETZEL-KÜHN, Gisela (trad.) et ZWEIG, Friderike Maria (trad.) « Zweiter teil, Gedichte ». In : ZWEIG, Stefan. *Marceline Desbordes-Valmore: das Lebensbild einer Dichterin*. Insel-Verlag, 1927, p. 83-130.
- [12] DIXON-YOUNG, Edie. « The Other Woman: Female friendship and homoeroticism in the poetry of Marceline Desbordes-Valmore ». Mémoire de master, Oxford University, p. 18-23. 2025.
- [13] EVE, Martin Paul. « “You Have to Keep Track of Your Changes”: The Version Variants and Publishing History of David Mitchell’s *Cloud Atlas* ». In : *Open Library of Humanities* 2, no. 2 (2016). <https://olh.openlibhums.org/article/id/4426/>, e1.
- [14] GENETTE, Gérard. *Seuils*. Poétique. Paris : Éditions du Seuil, 1987, p. 287-288.

- [15] KENDALL, Maurice G. « A New Measure of Rank Correlation ». In : *Biometrika* 30, no. 1/2 (1938), p. 81-93. DOI : 10.1093/biomet/30.1-2.81.
- [16] MAURI, Michele, ELLI, Tommaso, CAVIGLIA, Giorgio, UBOLDI, Giorgio et AZZI, Matteo. « RAWGraphs: a visualisation platform to create open outputs ». In : *Proceedings of the 12th biannual conference on Italian SIGCHI chapter*. 2017, p. 1-5.
- [17] MICHEL, Eugène. « Un chef-d’œuvre dispersé : Les *Élégies* 1818 de Marceline Desbordes ». In : *Lieux d’être* 50 (2010), p. 134-136.
- [18] NEWIRK, Elaine Francine Charlotte. *La Renommée littéraire de Marceline Desbordes-Valmore (1820-1950)*. Thèse de doct. University of Illinois at Urbana-Champaign, 1954, p. 5.
- [19] PLANTÉ, Christine. « Une poète, cette chose si rare ». In : *Barbey d’Aurevilly critique des femmes poètes, Barbey d’Aurevilly en tous genres*, sous la dir. de Brigitte DIAZ. Presses Universitaires de Caen, 2011, p. 161.
- [20] SCHABERT, Ina. « Marceline Desbordes-Valmore en traductions allemandes ». In : *J’écris pourtant* 6 (2025), p. 225-242.
- [21] SCHOPF, Jennifer M, TURNER, Katrina, DOYLE, Dan, LAKE, Andrew, LEIGH, Jason et TIERNEY, Brian L. « The NetSage measurement and analysis framework in practice ». In : *Cluster Computing* 25, no. 4 (2022), p. 2967-2990.
- [22] SELLAMI, Zied, GANASCIA, Jean-Gabriel et BOUKHALED, Mohamed Amine. « MEDITE: logiciel d’alignement de textes pour l’étude de la génétique textuelle ». In : *Actes de la 22^e conférence sur le Traitement Automatique des Langues Naturelles. Démonstrations*. 2015, p. 1-2.
- [23] TUFTE, Edward. « Slopegraphs for comparing gradients: Slopegraph theory and practice ». *The Work of Edward Tufte and Graphics Press*, <https://www.edwardtufte.com/notebook/slopegraphs-for-comparing-gradients-slopegraph-theory-and-practice/>. 2011.

A Annexes

A.1 Extraits de la correspondance de Marceline Desbordes-Valmore et d’une critique à propos de l’ordre des poèmes dans ses premiers recueils

Extrait d’une lettre du 26 avril 1818 de Marceline Desbordes-Valmore à son éditeur François Louis : « Je m’en remets à vous seul du soin de classer par ordre chaque pièce de vers – elles sont copiées pêle-mêle – je n’ai pas eu la patience de choisir leur place. Donnez-leur le rang qu’il vous plaira, et le titre qui leur convient car l’auteur de tout cela est un peu comme Monsieur Jourdain qui fait de la prose sans le savoir. » [4]

Extrait d’une lettre du 3 octobre 1818 de Marceline Desbordes-Valmore à son éditeur François Louis : « [...] j’ai cru qu’il fallait mettre les élégies dans l’ordre que je vous envoie par numéros ; » [4]

Extrait de la critique de Jacques Ancelot dans les *Annales de la littérature et des arts*, en 1821 : « Un recueil d’élégies est ordinairement un petit roman qui a son exposition, son action, son dénouement : cette marche plaît au lecteur qui aime à rencontrer un fil qui le guide parmi ces morceaux détachés ; elle est surtout favorable à *l’unité*, si précieuse à conserver en quelque chose qu’on écrit : M^{me} Desbordes n’a pas suivi cette méthode. Ses élégies, qui paraissent presque toutes avoir été inspirées par un même sentiment, n’ont cependant aucune suite entre elles ; un raccommodement précède une rupture, et à des plaintes sur l’inconstance d’Olivier, succède un éloge de sa fidélité. Dans une des élégies

adressées à *Délie*, M^{me} Desbordes lui adresse les plus vifs reproches ; dans la suivante, elle l'engage à venir pleurer au tombeau de son fils. Peut-être des motifs que nous ignorons ont pu engager l'auteur à mettre dans son ouvrage cette apparente confusion, mais elle peut être certaine que ce désordre ne paraît pas au lecteur *un effet de l'art.* » [2]

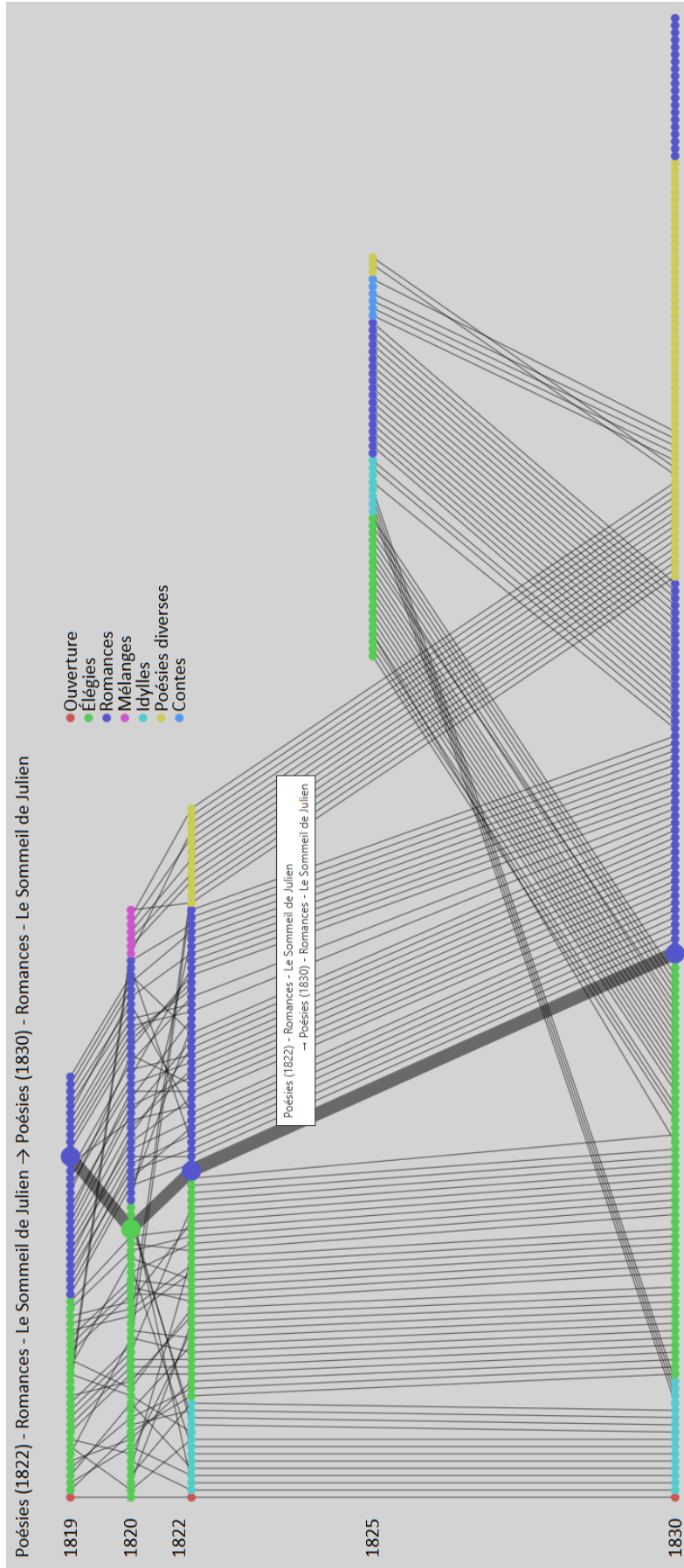


Figure 1 – Visualisation des recueils de Marceline Desbordes-Valmore publiés entre 1819 et 1830.

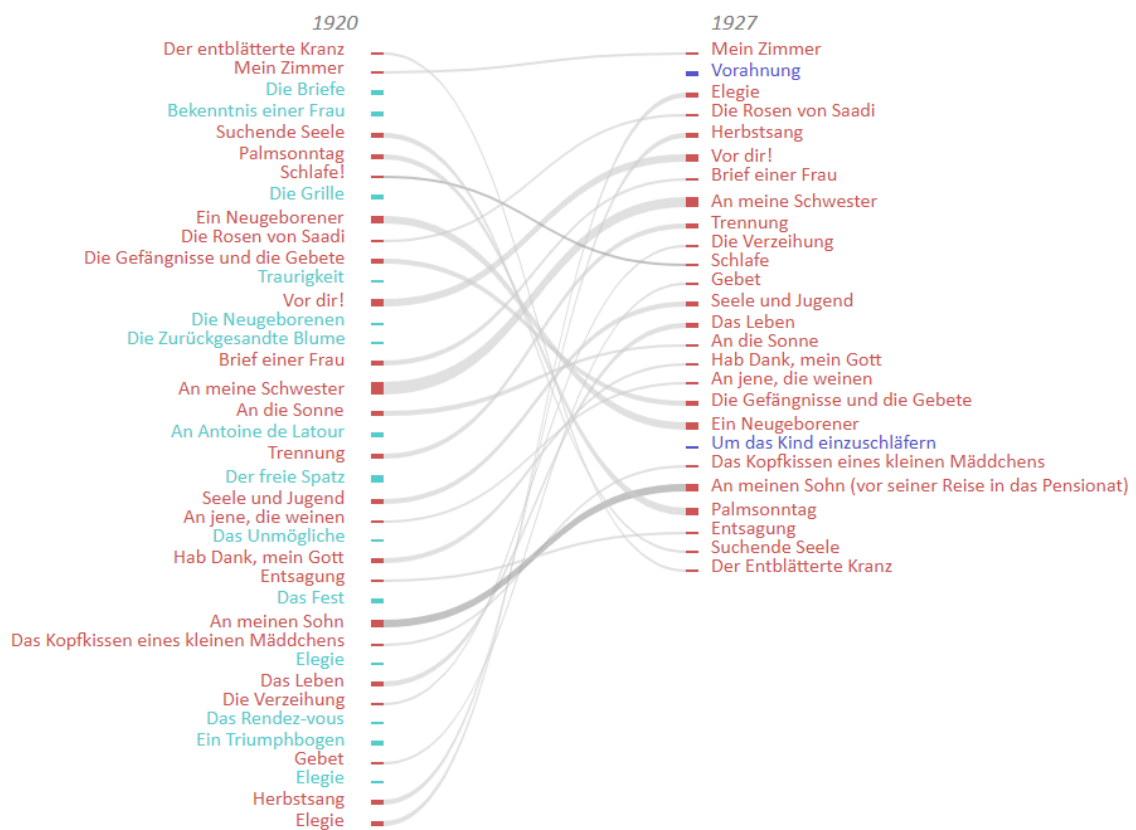


Figure 2 – Visualisation en diagramme de Sankey des éditions de 1920 et de 1927 de l’anthologie de poèmes en annexe de l’ouvrage de Stefan Zweig *Marceline Desbordes-Valmore : das Lebensbild einer Dichterin*.