

LESSICO  
BENI CULTURALI

MENU

# Le Corpus LBC Français

Annick Farina (Università di Firenze), Lorella Sini (Università di Pisa)

## Introduction

Comme cela a déjà été dit dans la présentation générale de cet ouvrage, notre corpus LBC français a été conçu à partir du projet interuniversitaire *Lessico dei Beni Culturali* dans l'objectif de créer des dictionnaires monolingues et bilingues reliés entre eux par des liens hypertextuels. Ces dictionnaires, tout comme les corpus qui les alimentent, comprennent plusieurs langues : allemand, anglais, chinois, espagnol, français, italien, portugais, russe<sup>[1]</sup>. Ils devraient permettre aussi bien à un public spécialisé qu'à un public plus large d'effectuer des recherches, à la fois qualitatives et quantitatives, concernant le patrimoine artistique de la ville de Florence et de la Toscane et, au-delà, dans un second temps, du patrimoine artistique italien. Nos outils - en particulier les corpus - devraient en effet contribuer à aiguiller les rédacteurs de guides ou ceux des légendes qui accompagnent les expositions, les commentateurs ou les traducteurs dans leurs choix linguistiques. Il leur permettra d'accéder aisément à une banque de données pouvant leur fournir des informations fiables comme par exemple des équivalents attestés de certains concepts, en particulier ceux relatifs aux techniques artistiques et aux matériaux, ou les renseigner sur la fréquence de certaines collocations ou co-occurrences particulières.

Pour notre corpus français, nous avons recueilli selon les mêmes critères que pour les autres langues, un certain nombre de documents caractérisés par une typologie variée, et rédigés sur un

arc temporel étendu, depuis les célèbres écrits sur la Renaissance et le *Rinascimento* jusqu'à l'époque moderne et contemporaine. Ceci devrait faciliter, entre autres, des études de sémantique comparative inter- et intra-linguistique aussi bien en diachronie qu'en synchronie.

L'équipe française a ceci de particulier qu'elle a été en quelque sorte pionnière dans le développement du projet, bénéficiant de beaucoup de ressources dès le départ, grâce à la contribution de nombreux chercheurs et étudiants-stagiaires, qui ont collaboré à ce projet depuis 2013. Ce corpus - qui provient d'une approche '*corpus-driven*' et non '*corpus-based*' (Biber 2012) - a de fait représenté un terrain d'essai pour le projet, qui a servi inductivement de base de travail pour les autres langues, permettant de mettre en évidence les difficultés de gestion des sources et de création des métadonnées. Il intègre certains éléments destinés à tester différents aspects (principalement le bien-fondé de nos catégories lexicographiques), que nous souhaiterions développer dans les prochaines années pour l'ensemble de nos corpus. Ainsi, le corpus français contient des échantillons de textes différents de ceux que l'on retrouve dans les autres langues, à la fois au niveau de la typologie textuelle (correspondances d'artistes, par exemple mais aussi sites internet destinés aux touristes) que du point de vue de l'extension de notre principal objet d'observation - qui a jusqu'à présent guidé nos choix - puisque nous avons commencé à intégrer au patrimoine artistique toscan le patrimoine artistique matériel et immatériel italien.

Pensé comme un 'laboratoire didactique' (Farina 2016), le projet a donné lieu à de nombreux mémoires et stages d'étudiants. Des mini-corpus ont été créés dans ce cadre qui ont été rassemblés à l'intérieur du corpus principal et qui continueront de l'être dans les prochains mois et années<sup>[2]</sup>. De même, les chercheurs membres du groupe ont travaillé sur certains auteurs ou thèmes reliés à notre projet et ont transcrit pour ce faire un certain nombre de textes dont notre corpus a pu bénéficier. Nous pouvons citer par exemple le travail de Rosa Cetro (Université de Pise) sur A. Félibien, celui de Nicole Maroger (Université de Florence) concernant le *Palio de Sienna*, le projet de recherche actuellement développé par Valeria Zotti (Université de Bologne) sur les récits de voyageurs français et francophones à Bologne et celui de Dominique Garand (Université du Québec à Montréal) qui collabore actuellement au projet pour rassembler des textes sur l'Italie d'auteurs québécois. Pour ce qui concerne les rédactrices de cet article, elles ont recueilli et transcrit la plupart des documents présents aujourd'hui dans le corpus,

survolant les auteurs d'une époque à l'autre au gré de leurs découvertes<sup>[3]</sup>.

Notre travail doit beaucoup également aux collaborations inter-universitaires franco-italiennes comme celles que nous avons pu avoir avec le Laboratoire d'Informatique de l'Université de Tours, le Laboratoire UMR 5206 Triangle du CNRS, le groupe LiCoRN de l'Université de Bretagne Sud, des laboratoires qui ont aiguillé notre méthodologie tant pour la constitution du corpus que pour ses futures utilisations lexicographiques.

## 1. Contenu du corpus

La recherche de sources s'est tout d'abord effectuée sur la base de bibliographies existantes sur les récits de voyage en Italie<sup>[4]</sup> et les ouvrages de référence, relatifs à l'histoire de l'art italien et de la Renaissance en particulier. Les premiers textes intégrés à la base ont été des classiques facilement accessibles en ligne et appartenant au domaine public. En particulier, nous avons tiré une partie consistante de notre corpus dans les ressources disponibles des bibliothèques numériques Gallica, par exemple la traduction de C. Weiss des *Vies* de G. Vasari publiée en 1900, et la bibliothèque numérique archive.org, par exemple l'ouvrage *Des principes de l'architecture, de la sculpture, de la peinture...* de A. Félibien (1676), tous les deux encore en cours de transcription dans des fichiers d'édition<sup>[5]</sup>. Nous avons aussi utilisé des ouvrages présents sur Wikisource, que nous avons d'ailleurs souvent contribué à transcrire en ligne, comme les textes d'A. Dumas *La villa Palmieri* (1843) ou *Une Année à Florence* (1841).

Ceci explique que notre corpus est principalement littéraire (près de 60% du corpus). Il contient en effet tant des récits de voyage des écrivains du Grand Tour et de ceux qui ont suivi leurs traces plus tard jusqu'au début du XXe siècle, que des œuvres de fiction ayant Florence ou l'Italie comme toile de fond, sans oublier des écrits critiques ou 'amoureux' sur l'art ou la culture italienne. On sait que l'une des destinations privilégiées de ces auteurs, surtout durant la période romantique, était la ville de Florence et la Toscane, ce qui n'était pas toujours le cas des voyageurs européens qui leur ont souvent préféré Rome, Venise ou Naples. Nul ne s'étonnera donc que la plus grande partie de nos textes soit puisée dans la littérature de voyage du XIXe et de la première moitié du XXe siècle. Les nombreuses traductions qui figurent dans notre corpus, également

relatives à cette même époque, confirment l'attrait particulier des érudits français pour le patrimoine artistique italien.

Cet attrait préfigure la prédilection d'aujourd'hui pour le tourisme culturel en Italie. Même si cela représente, pour le moment, une moindre proportion dans notre corpus, des textes plus récents et plus contemporains ont été incorporés au fur et à mesure de l'avancée de nos recherches, car le voyage en Italie demeure un *topos* bien vivant dans notre littérature. Il a suscité dans le temps et continue de susciter de nombreuses publications qui vont du livre de cuisine au guide touristique, de la critique d'art au dictionnaire des peintres en passant par la correspondance et l'article de journal. Ci-dessous la répartition des auteurs figurant dans notre corpus avec indication des *tokens* [6], par ordre décroissant :

Viollet Le Duc	853,133	Lafenestre, Georges	24,527	Revel, Adèle	4,588
Fernandez, Dominique	726,437	Perrot, George	20,886	Pommier, Édouard	4,443
SA	264,073	Allais, Gustave	19,221	Schmitz, Isabelle	4,016
Vasari, Giorgio	252,132	Goncourt, J. et E.	19,194	Dufay, Henri	3,832
Stendhal	218,313	Félibien, André	18,894	Grimaldi, Edouard-F.	3,594
Dumas, Alexandre	168,272	Brosses, Charles de	18,316	Musset, Paul de	3,584
Colet, Louise	135,166	Michel, Émile	17,289	Jaucourt, Chevalier de	3,548
France, Anatole	103,614	Gautier, Théophile	15,326	Bard, Joseph	3,266
Faure, Élie	84,716	La Sizeranne, R. de	15,195	Pasquin, Antoine-C.	3,028
Lang, Jack	73,914	Renan, Ernest	14,695	Powell, Véronique G.	2,612

Bazin, René	73,301	Le Routard	14,463	Favre le Bret, Robert	2,556
Taine, Hippolyte	58,962	Montaigne, Michel de	13,057	Erdan, Andrea	2,476
Auzias, Dominique	55,897	Fréville, Valentin	12,126	Taillasson, Jean-J.	2,106
Cellini, Benvenuto	50,742	Lescure, O.	11,584	Nobecourt, Jacques	1,947
Grandgeorge, Gaston	48,737	Palustre, Léon	10,414	Prieur, Jean	1,909
Klaczko, Julian	44,386	Giono, Jean	10,135	Bertheroy, Jean	1,809
Feuillet, Michel	42,624	Méry, Joseph	9,976	Mallarmé, Camille	1,593
Meyer, Eric	40,029	Staël-Holstein, G.	9,152	Colombari, Colonel F.	1,578
Moran, Robert	34,705	Libri, Guglielmo	8,796	Creuzé de Lesser, A.	1,564
Maurel, André	32,942	Delacroix, Eugène	7,174	Beaugrand, Honoré	1,416
Rosov, Gricha	32,613	Wyzewa, Théodore de	5,299	De la Borie, G.	1,155
De Navenne, Ferdinand	30,664	Mallet, George	5,137	Gaboury, Louise	907
Chateaubriand, François-R.	29,599	Camus, Albert	5,033	Artusi, Pellegrino	388

Figure 1 - Nombre de *tokens* pour chaque auteur présent dans le corpus en ordre décroissant

Le corpus LBC français comprend 3,164,995 mots, correspondant à 252 textes (livres, articles ou chapitres de livre) dont 129 textes en langue originale française et 123 textes traduits de l'italien vers le français [8] qui sont essentiellement des fragments pour permettre la future création de bases parallèles, selon les proportions suivantes divisées par siècle :

siècles	Tokens	dates de rédaction	Tokens	pourcentages de tokens <sup>[9]</sup>
XVIe	13 057	1580	13 057	0,34%
XVIIe	18 894	1676	18 894	0,55%
XVIIIe	21 864	1739	18 316	0,63%
		1751	3 548	
XIXe	1 796 410	1801	1 564	52,17%
		1807	11 258	
		1815	5 137	
		1817	212 493	
		1825	582	
		1827	29 599	
		1829	2 691	
		1830	7 174	
		1832	879	
		1841	4 953	
		1842	3 028	
		1843	122 326	
		1844	3 594	
		1849	14 695	
		1850	1 212	
		1853	13 808	
		1854	857 977	
		1855	19 194	
		1862	135 166	
		1863	4 588	
1864	58 962			
1865	15 326			
1866	11 584			
1867	247			

siècles	Tokens	dates de rédaction	Tokens	pourcentages de tokens <sup>[9]</sup>
		1870	5 631	
		1878	20 886	
		1880	44 386	
		1888	141	
		1892	10 414	
		1894	176 915	
XXe	554 153	1901	17 289	16,09%
		1903	30 664	
		1905	5 299	
		1906	18 896	
		1907	48 737	
		1909	19 221	
		1910	15 195	
		1913	34 751	
		1923	84 716	
		1924	1 593	
		1934	1 909	
		1936	5 033	
		1938	2 556	
		1953	10 135	
		1969	1 947	
		1982	221 507	
1994	34 705			
XXIe	1 038 979	2002	249 289	30,17%
		2007	2 612	
		2008	297 931	
		2012	113 943	
		2013	33 768	

siècles	Tokens	dates de rédaction	Tokens	pourcentages de tokens <sup>[9]</sup>
		2015	907	
		2016	321 722	
		2017	18 473	

Figure 2 - Distribution des *tokens* par année/siècle de publication

À partir de ce tableau, on comprend clairement que c'est la langue française des XIXe et XXIe siècles qui est la mieux représentée dans le corpus, périodes qui présentent une plus grande accessibilité des sources, les textes du XIXe étant désormais libres de droits et les textes contemporains du XXIe étant facilement accessibles en ligne. Il est cependant nécessaire de noter que toutes les catégories de textes insérés dans le corpus ne sont pas distribuées de manière homogène sur les différentes périodes attestées, comme nous le verrons ci-dessous (figure 4). Il s'agit là d'un aspect dont il faudra tenir compte lors de recherches ciblées sur des genres textuels (cfr. partie 2.)

Les textes ont été répartis selon quatre catégories ou genres : les textes de vulgarisation (DIV pour '*divulgativo*'), les dictionnaires (DIZ pour '*dizionario*'), les textes techniques destinés aux spécialistes (TEC pour '*tecnico*'), les textes littéraires (LET pour '*letterario*'), selon la répartition suivante :

	Pourcentage sur mots corpus	Sous-catégorie	Nombre de documents	Nombre de mots	Pourcentage total mots corpus
DIV (vulgarisation)	10,92%	BLG (blogs)	1	27 029	0,8%
		RIV (revues et journaux)	8	41 940	1,3%
		GUI (guides touristiques)	4	277 179	8,7%
DIZ (dictionnaire)	22,26%	MON (monolingue)	7	704 622	22,26%
TEC (technique)	8,42%	ARC (architecture)	13	24 290	0,7%



	Pourcentage sur mots corpus	Sous-catégorie	Nombre de documents	Nombre de mots	Pourcentage total mots corpus
		ART (arts en général)	67	243 398	7,6%
		GAS (œno-gastronomie)	1	321	0,0%
LET (littéraire)	58,35%	BIO (biographies)	124	733 487	23,1%
		FIC (fiction)	10	524 493	16,5%
		SAG (essais)	17	589 228	18,6%

Figure 3 - Distribution des documents selon les catégories (et sous-catégories) du corpus

1) Sous la classification *[DIV]*, nous avons compté les textes de vulgarisation destinés au grand public, puisés essentiellement dans des documents récents, comme nous pouvons le constater sur le tableau ci-dessous (figure 4). Ces textes se répartissent en trois sous-catégories : 1 texte sous forme de blog participatif d'amateur de voyage conçu comme une nomenclature (*[DIV\_BLG]*), 8 textes (XXe et XXIe) tirés de revues spécialisées et de la presse généraliste *[DIV\_RIV]*, 4 guides touristiques dont 1 de 1829 et les autres postérieurs aux années 2000 (*[DIV\_GUI]*).

2) Les dictionnaires (*[DIZ]*) comptent 7 textes dont 2 du XVIIIe et 5 du XIXe. Cette catégorie a été construite dans l'objectif d'établir une base de travail pour la composition de futurs dictionnaires monolingues spécialisés comprenant aussi bien des définitions que des citations et des références sourcées. Comme nous pouvons le constater, cette catégorie est constituée quasi exclusivement du *Dictionnaire raisonné de l'architecture française* de Viollet-le-Duc (par ailleurs non entièrement reporté dans notre base qui comprend 5 tomes sur 11). Ce texte important et incontournable pour la culture artistique française, occupant quasiment un quart de nos données, peut paraître surdimensionné au regard des autres textes recueillis et des autres langues qui n'ont pas bénéficié d'une ressource documentaire similaire. C'est pourquoi nous avons introduit la possibilité d'exclure des recherches ce sous-corpus afin d'éviter de biaiser les résultats. En outre, 2 articles de *l'Encyclopédie* : l'article « Florence » et l'article « École florentine » y ont été adjoints ;

3) Certains textes destinés à un public spécialisé dans l'étude des arts, ont été classés comme 'techniques' ([TEC]). Ils sont composés de 81 documents, subdivisés en trois catégories ou sous-corpus : des textes techniques spécialisés dans l'architecture ([TEC\_ARC]) ; des textes relatifs aux autres techniques artistiques, telles que la sculpture ou la peinture ([TEC\_ART]) et enfin des textes concernant le domaine de l'œno-gastronomie ([TEC\_GAS]), une spécialité dont nous supposons qu'elle tiendra un rôle important dans la constitution du patrimoine gastronomique italien. Les célèbres œuvres de Artusi ont été recueillies à cet effet.

4) Des textes littéraires ([LET]) : ont été classés sous cette typologie les œuvres ayant l'Italie et plus particulièrement la Toscane et la ville de Florence en toile de fond. Il s'agit de 151 documents qui constituent plus de la moitié de notre corpus. Cette catégorie a été subdivisée en trois sous-corpus :

a) Les œuvres ayant un caractère biographique ([LET\_BIO] : 23% du corpus). La majorité est constituée par les récits de voyages en Italie des grands écrivains du XIXe : du *Voyage en Italie* de Chateaubriand (1827), à *l'Italie des Voyages* d'Ernest Renan (1849), en passant par *Une année à Florence* de Dumas (1841). De même, nous avons intégré les traductions en français de la vie de Benvenuto Cellini par lui-même (pour le moment, les livres I et II, traduction de 1847 de Léopold Leclanché). Nous disposons déjà de la plupart des lettres de Michel-Ange (traduction de Boyer d'Agen datant de 1911) mais nous n'avons pas encore décidé de leur future classification et en particulier de la possibilité d'intégrer des fragments d'un même ouvrage dans des catégories différentes avant de les rendre disponibles dans le corpus (cfr. conclusion). Nous souhaiterions, dans un avenir proche, ajouter des traductions des carnets de Léonard de Vinci, en commençant par son *Traité de la peinture* et des fragments de *l'Histoire de Florence* de Machiavel sur lesquelles nous travaillons actuellement.

b) Les œuvres relevant de la fiction ([LETT\_FIC] : 16,5% du corpus) : une grande partie de ce sous-corpus est constitué d'œuvres du XIXe dont le célèbre récit de Madame de Staël, *Corinne ou l'Italie* (1807), en particulier *Le séjour à Florence* du livre XVIII.

c) Les essais ([LETT\_SAG] : 18,6% du corpus) sont constitués de 17 textes répartis sur un peu plus de deux siècles, du XVIIIe jusqu'à nos jours. Il s'agit, entre autres, de *l'Histoire de la peinture en Italie* de Stendhal (1817), de certains articles parus dans la *Revue des deux mondes* (1832, 1878, 1901, 1905, 1906). Nous avons également

numérisé puis océrisé le *Que-sais-je sur l'Art italien* (Michel Feuillet, 2008) mais aussi des *Dictionnaires amoureux de l'Italie* de Dominique Fernandez (1997 et 2008). Cette sous-catégorie est plus délicate à circonscrire que les autres car le pacte de lecture a pu changer au fil des siècles. En effet, ce qui était destiné à un public de spécialistes des arts et des techniques artistiques au XIXe siècle, s'avère être aujourd'hui un témoignage littéraire légué par d'illustres critiques d'art de l'époque.

	<b>DIV (vulgarisation)</b>	<b>DIZ (dictionnaire)</b>	<b>TEC (technique)</b>	<b>LET (littéraire)</b>
XVIe	0	0	0	0,9%
XVIIe	0	0	80%	0
XVIIIe	0	28,6%	0	3,6%
XIXe	8,3%	71,4%	13,3%	61,6%
XXe	16,7%	0	6,7%	28,6%
XXIe	75%	0	0	5,3%

Figure 4 - Pourcentage de documents de chaque catégorie selon les siècles.

## 2. Exemple de recherches à l'intérieur du corpus

Grâce à la participation active d'un informaticien à notre projet (Riccardo Billero) qui a obtenu des contrats de recherche postdoctoraux, nous avons pu mettre au point tant un modèle de gestion des sources qu'un formulaire de recherche *ad hoc*<sup>[11]</sup> à l'intérieur des corpus. Nous illustrerons ici quelques exemples de recherches à l'intérieur du corpus avec différents filtres et options<sup>[12]</sup>.

Dans le cadre d'une comparaison entre le corpus français et allemand (Farina, Flinz 2020), nous avons entrepris une recherche sur les différentes dénominations du principal monument de Florence, la cathédrale Santa Maria del Fiore, dans les deux langues selon les types de textes et leur époque de rédaction. Pour vérifier le nombre d'occurrences de la dénomination *Sainte-Marie-des-Fleurs* et les contextes dans lesquels on la trouve, nous avons d'abord fait la recherche dans les seuls textes 'en langue originale' et dans le sous-

corpus des textes appartenant à la catégorie 'Littéraire' en utilisant le formulaire de recherche disponible sous 'Search'/ 'Text types' (fig. 5) :

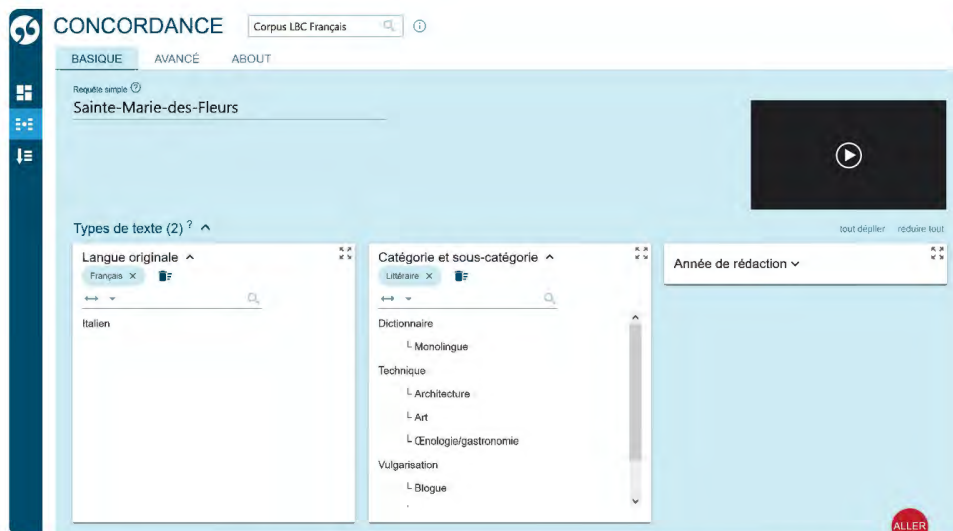


Figure 5 – Recherche de concordances contenant *Sainte-Marie-des-Fleurs* dans les textes en langue originale de la catégorie 'Littéraire'.



Figure 6 – Les onze occurrences de *Sainte-Marie-des-Fleurs* dans les sources littéraires du corpus.

Cette recherche, comparée à celles de la même forme dans les autres catégories textuelles, puis des autres formes (*Santa Maria Novella*, *Dôme* et *Duomo*) dans les différentes catégories nous a permis de remarquer que, alors que le nom français du monument *Sainte-Marie-des-Fleurs* n'est jamais utilisé dans nos textes de divulgation (dans les guides touristiques) ni ne se trouve dans les textes en traduction, il apparaît dans une bonne proportion par rapport aux autres noms dans les ouvrages littéraires. Un croisement de ces données avec les datations des occurrences (visualisation des dates de rédaction obtenue en indiquant l'option

d'affichage des dates de rédaction en sélectionnant les « Options d'affichage du détail des lignes » disponibles en cliquant sur l'icône d'information à côté de l'information bibliographique de n'importe quelle référence, comme montré dans les figures 7 et 8) nous a permis de mieux comprendre la distribution des choix de dénomination du monument.

The screenshot shows the CONCORDANCE search interface. The search term is 'Sainte-Marie-des-Fleurs' with 11 results. A dropdown menu is open over the search results, showing 'Options d'affichage du détail des lignes'. The menu items include: 'Année de rédaction' (checked), 'Année de rédaction traduction', 'Auteur', 'Catégorie et sous-catégorie', 'Décennie de rédaction', 'Décennie rédaction traduction', 'Langue originale', 'Siècle de rédaction', 'Siècle rédaction traduction', 'Titre original', 'Titre traduction', 'Traducteur', 'Année de publication', 'Année de publication traduction', 'Fragment', and 'Fragment traduction'. The 'Année de rédaction' option is selected, and the value '1841' is visible next to it.

Figure 7 – Modification des options d'affichage du détail des lignes.

The screenshot shows the 'Options d'affichage du détail des lignes' menu in detail. The 'Année de rédaction' option is selected, and the value '1841' is displayed next to it. Other options include 'Année de rédaction traduction' (value: ===NONE===), 'Auteur' (value: Dumas, Alexandre), 'Catégorie et sous-catégorie' (value: Littéraire, Littéraire::Biographique), 'Décennie de rédaction' (value: 1840 - 1849), 'Décennie rédaction traduction' (value: ===NONE===), 'Langue originale' (value: Français), 'Siècle de rédaction' (value: 1800 - 1899), 'Siècle rédaction traduction' (value: ===NONE===), 'Titre original' (value: Une année à Florence), 'Titre traduction' (value: ===NONE===), 'Traducteur' (value: ===NONE===), 'Année de publication', 'Année de publication traduction', 'Fragment', and 'Fragment traduction'. The 'FERMER' and 'SAUVEGARDER' buttons are visible at the bottom right of the menu.

Figure 8 – Sélection de l'affichage de la date seulement dans les options d'affichage du détail des lignes.

The screenshot shows the 'CONCORDANCE' interface for the 'Corpus LBC Français'. The search term is 'Santa Maria del Fiore' with 34 results. The results are sorted by 'writingYear' (date of writing). The interface includes a search bar, a list of results with checkboxes, and a 'Détails' button. The results list shows the following entries:

Year	Text Snippet
1815	cathédrale appelée autrefois l' église de la Reparata aujourd' hui <b>Santa Maria del fiore</b> , que fut massacré Julien de Médicis : la famille des Pazzi, riche
1817	s'engage envers la confrérie de marchands qui se réunissaient à <b>Santa Maria del Fiore</b> , à tirer une statue haute d'environ neuf brasses (cinq mètres vir
1817	ie de Saint Matthieu, qui se voit encore dans la première cour de <b>Santa Maria del Fiore</b> , et qu'il abandonna peut-être comme ayant une position trop co
1817	lais de naissance, celui qui fut enterré avec tant de pompe dans <b>Santa Maria del Fiore</b> , la cathédrale de Florence, et sur le tombeau duquel on voit un
1817	dans la plaine, et j'ai aperçu de loin, comme une masse sombre, <b>Santa Maria del Fiore</b> et sa fameuse coupole, chef-d' œuvre de Brunelleschi. C'est là q
1817	Florence : ce sont des forteresses. Je regardais tout à l'entour de <b>Santa Maria del Fiore</b> (bâtie en 1293), ces arcades légèrement gothiques, dont la poin
1817	s arcades qui entouraient la place immense au milieu de laquelle <b>Santa Maria del Fiore</b> s' élève isolée. Je me sentais heureux de ne connaître personne
1817	de fous s'ils se présentaient aujourd' hui au grand café à côté de <b>Santa Maria del Fiore</b> ; mais comparés aux personnages de Paul Véronèse et du Tint
1862	Florence et ses tours se déploient sous mes yeux ; la coupole de <b>Santa Maria del Fiore</b> , le palais Vecchio et le Bargello dominent les groupes des mon
1862	ence montent en relief ; les coupoles des églises Santa Croce et <b>Santa Maria del Fiore</b> scintillent au soleil. A ma droite, les parterres du jardin dei Semp
1862	superbe, un chant à travers lequel je distingue ces paroles : Bella <b>santa Maria del Fiore</b> Dei Toscani sei l'amore ! Bellissima torre del Giotto Tu sei la gioi
1862	mmortels, parrains des fils de la Toscane. Le même jour, je visite <b>Santa Maria del Fiore</b> , le Baptistère et le Campanile de Giotto. Que dire de ces trois r
1862	ne d'un navire autour d'un grand mât. Je fais le tour du Dôme ou <b>Santa Maria del Fiore</b> , nom suave que l' église emprunte au nom même de Florence €
1862	mmentaient Dante au quinzième siècle, dans cette église même de <b>Santa Maria del Fiore</b> , comme il eût fait d'un texte sacré. Nos deux amoureux marche
1862	encement du quinzième par le pape Eugène IV. Comme celle de <b>Santa Maria del Fiore</b> , la façade est restée inachevée : elle est en pierres brutes. On t
1862	ntent l' Église militante et l' Église triomphante. La cathédrale de <b>Santa Maria del Fiore</b> , reproduite dans une de ces fresques, symbolise l' Église militai
1862	hons, nous voyons poindre au-dessus des toitures la coupole de <b>Santa Maria del fiore</b> de Brunelleschi, puis le campanile de Giotto, puis quelques dôm
1864	ls de Laurent, institue une discussion sur la véritable amitié dans <b>Santa Maria del Fiore</b> et propose en prix au vainqueur une couronne d'argent. On voit
1906	ux œuvres de beauté publique, portes du Baptistère, Coupole de <b>Santa Maria del Fiore</b> , etc. Fêtes, vêtements et jeux, gardaient encore un caractère de
1906	e dame, admiratrice de Savonarole, le décide à l'accompagner à <b>Santa Maria del Fiore</b> . Il s' y sent mal à l'aise, entouré d' œillades défiantes, au milieu

Figure 9 - Visualisation des occurrences de *Santa Maria del Fiore* dans la catégorie 'Littéraire' du corpus, classées par date de rédaction (classement en cliquant en haut à gauche sur 'Détails').

De fait, 90% des occurrences de *Sainte-Marie-des-Fleurs* sont du XIXe siècle et près de 80% de *Santa Maria del Fiore* sont des XXe et XXIe siècles. Si la différence de genres textuels nous a ainsi semblé un élément pertinent pour expliquer les différences de désignation. La désignation en langue originale italienne se manifeste en particulier dans une plus grande fréquence dans les guides touristiques qui optent pour le maintien de la forme originale des noms propres en général, et qui se retrouve aussi dans les textes en traduction. Cependant, la période de rédaction des documents n'en reste pas moins importante. La datation des occurrences nous permet ainsi de pallier le manque d'homogénéité dans les proportions données à chaque période historique dans les différentes catégories textuelles de notre corpus que nous avons déjà illustrées dans la partie 1.

Une autre option extrêmement utile pour nos recherches lexicographiques est celle qui nous permet d'obtenir des informations sur les collocations et co-occurrences sur les différents mots présents dans le corpus. Il est possible de faire cette recherche de deux manières différentes :

a) En utilisant l'option 'Collocation' du logiciel qui permet d'obtenir une liste des formes présentes à une distance à définir dans le menu de recherche de X mots (en cliquant sur l'icône présent dans le menu en haut à droite de l'écran). Par exemple dans la figure 10, une

recherche des mots que l'on trouve à une distance de un ou deux mots de *arc* sur la droite :

CONCORDANCE Corpus LBC Français

simple *arc* • 1 788  
456,13 fréquence/occurrence • 0,046%

COLLOCATIONS

BASIQUE AVANCÉ ABOUT

Atrous ? Fourchette ?

mot -5 -4 -3 -2 -1 KWIC 1 2 3 4 5

ALLER

Détails Contexte de gauche KWIC Contexte de droite

1	<input type="checkbox"/> Brosses, Lettre...	s, le jeune Néron, le Marc-Aurèle enfant, et l'Amour qui tire de l'	<i>arc</i>	; parmi celles de pierres précieuses, le Lysimachus de calcédois
2	<input type="checkbox"/> Taine, Voyage e...	ments délicats, de feuillages et de figurines peintes, couvrent les	<i>arcs</i>	et les plafonds de leur multitude harmonieuse ; le regard s'en re
3	<input type="checkbox"/> Taine, Voyage e...	juste de cette forme et de toute forme, que, copiant à Rome les	<i>arcs</i>	de triomphe et les amphithéâtres, il les dessinait à l'œil aussi si
4	<input type="checkbox"/> Feuillet, L'art...	les lignes de force se dessinent. La plaine du Pô, délimitée par l'	<i>arc</i>	alpin et l'Apennin, conserve des édifices imposants : on parle d
5	<input type="checkbox"/> Feuillet, L'art...	rmande à des éléments de l'architecture islamique, tels que des	<i>arcs</i>	outrépassés, des arcatures entrecroisées, des plafonds en sta
6	<input type="checkbox"/> Feuillet, L'art...	re moitié du XIIIe siècle et qui ont formé des maîtres toscans. L'	<i>arc</i>	trionphal et la conque absidiale de la basilique San Clemente d
7	<input type="checkbox"/> Feuillet, L'art...	hiques du centre-ouest de l'Europe, le recours systématique à l'	<i>arc</i>	ogival permet de lancer les travées jusqu' à des hauteurs vertig
8	<input type="checkbox"/> Feuillet, L'art...	n architecture mêle harmonieusement la pure tradition romane (	<i>arcs</i>	en plein cintre) et les références à l'antique (pilastres cannelés,
9	<input type="checkbox"/> Feuillet, L'art...	475-1564) est le prénom francisé de Michelangelo Buonarroti. L'	<i>arc</i>	de temps que dessine sa longue vie – quatre-vingt-neuf ans – v
10	<input type="checkbox"/> Feuillet, L'art...	nt-des-Théâtres et de la chapelle du Saint-Suaire, constituées d'	<i>arcs</i>	superposés et entrecroisés laissant filtrer la lumière. Le sculpte
11	<input type="checkbox"/> Feuillet, L'art...	), Non loin de là, la Piazza della Repubblica, ses arcades et son	<i>arc</i>	de triomphe rappellent le statut de capitale dont Florence a bén
12	<input type="checkbox"/> Chateaubriand, ...	a Paix, les ruines du palais doré de Néron, celles du Colisée, les	<i>arcs</i>	de triomphe de Titus, de Septime Sévère, de Constantin ; vaste

Figure 10 – Recherche d'une liste de mots à distance de 1 ou 2 mots à droite du lemme *arc* dans le menu 'Collocation'.

CONCORDANCE Corpus LBC Français

simple *arc* • 1 788  
456,13 fréquence/occurrence • 0,046%

Collocations MODIFIER LES CRITÈRES RETOUR AU CONCORDANCIER

Mot	Cooccurrences ?	Candidats ?	T-score	MI	LogDice ±
<input type="checkbox"/> doubleaux	189	194	13,74	11,05	11,61 ...
<input type="checkbox"/> ogives	180	247	13,41	10,64	11,50 ...
<input type="checkbox"/> doubleau	129	129	11,35	11,09	11,11 ...
<input type="checkbox"/> décharge	60	72	7,74	10,83	10,05 ...
<input type="checkbox"/> ogive	44	102	6,63	9,68	9,57 ...
<input type="checkbox"/> tiers-point	34	77	5,82	9,91	9,22 ...
<input type="checkbox"/> plein	43	626	6,51	7,23	9,19 ...
<input type="checkbox"/> triomphe	31	152	5,56	8,80	9,03 ...
<input type="checkbox"/> brisé	28	78	5,28	9,61	8,94 ...
<input type="checkbox"/> cercle	29	244	5,36	8,02	8,87 ...
<input type="checkbox"/> cintre	26	153	5,09	8,54	8,78 ...
<input type="checkbox"/> diagonaux	23	38	4,79	10,37	8,69 ...
<input type="checkbox"/> brisés	21	43	4,58	10,06	8,55 ...
<input type="checkbox"/> voûtes	24	995	4,81	5,72	8,14 ...
<input type="checkbox"/> formeret	12	72	3,45	8,51	7,72 ...
<input type="checkbox"/> E	12	375	3,41	6,13	7,51 ...
<input type="checkbox"/> croisés	10	47	3,16	8,86	7,48 ...
<input type="checkbox"/> sont	50	8891	6,50	3,62	7,26 ...
<input type="checkbox"/> B	9	624	2,90	4,98	6,93 ...
<input type="checkbox"/> au-dessus	11	1 170	3,16	4,36	6,93 ...

Lignes par page : 20 1-20 de 138 1 > >1

Figure 11 – Résultat de la recherche de la figure 10 : liste de mots à distance de 1 ou 2 mots du lemme *arc*.

Cette recherche nous permet d'obtenir tant les collocations ayant une structure en *arc* + adjectif ou nom, que celles qui ont une structure *arc* + préposition + nom : *arc ogive*, *arcs doubleaux*, *arc formeret*, etc. et *arc en ogive*, *arc en tiers-point*, *arc de voûte*, etc.

b) Lorsque l'on fait une recherche de concordances sur un mot, en utilisant l'option d'affichage 'Trier' de visualisation 'Sort/Left' ou

'Sort/Right', en triant par exemple sur la droite comme sur la figure 12, on peut aussi observer les différentes collocations en se déplaçant dans les différentes pages de résultats, comme présenté sur la figure 13 (toujours dans le cadre d'une recherche de collocations avec le mot *arc*).

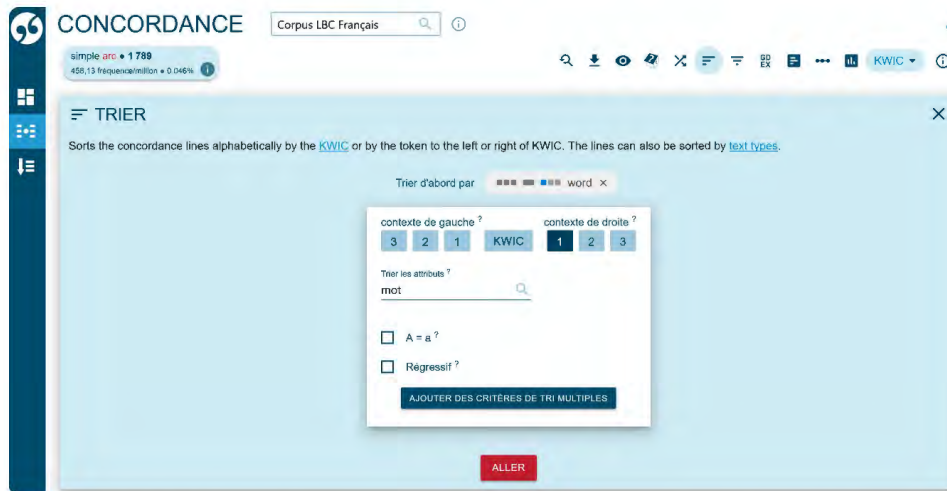


Figure 12 – Recherche de résultats à l'aide de l'option 'Trier' sur les collocations contenant le mot *arc* (ordonnées selon le contexte à gauche du mot).

351	<input type="checkbox"/>	<input type="radio"/>	Viollet Le Duc,...	nt leur offrir. Les constructeurs gothiques n'avaient point trouvé l'	arc	brisé ; il existait, ainsi que nous l'avons vu plus haut, dans des c
352	<input type="checkbox"/>	<input type="radio"/>	Viollet Le Duc,...	anchement roman. Mais les architectes gothiques appliquèrent l'	arc	brisé à un système de construction dont ils sont bien les seuls e
353	<input type="checkbox"/>	<input type="radio"/>	Viollet Le Duc,...	nté quelque chose chez nous avant le XVIe siècle. Il en est de l'	arc	brisé comme de toutes les inventions de ce monde qui sont à l'
354	<input type="checkbox"/>	<input type="radio"/>	Viollet Le Duc,...	industrie et à notre activité. Il est donc puéril de nous dire que l'	arc	brisé étant de tous les temps, les constructeurs du XIIe siècle n'
356	<input type="checkbox"/>	<input type="radio"/>	Viollet Le Duc,...	fait un pas en arrière, puisque leurs devanciers avaient adopté l'	arc	brisé, à la suite de fâcheuses expériences, comme poussant m
356	<input type="checkbox"/>	<input type="radio"/>	Viollet Le Duc,...	-cercle environ. Mais si les arcs doubleaux sont tracés suivant l'	arc	brisé EMI, le remplissage en moellons du triangle COD ne char
357	<input type="checkbox"/>	<input type="radio"/>	Viollet Le Duc,...	prendre moins de place sur la voie publique. Peu importe que l'	arc	brisé choque les yeux des parisiens exclusifs de l'antiquité, si o
358	<input type="checkbox"/>	<input type="radio"/>	Viollet Le Duc,...	noins d'importance que les arcs sont plus aigus. Ainsi (1) soit un	arc	brisé ABC, ayant les centres des deux courbes en A et B : c'est
359	<input type="checkbox"/>	<input type="radio"/>	Viollet Le Duc,...	t pas entièrement les traditions romanes (16). Le plein cintre et l'	arc	brisé se méient, et l' école nouvelle ne se montre que dans la fc
360	<input type="checkbox"/>	<input type="radio"/>	Viollet Le Duc,...	des moulures. Ici même, le plein cintre apparaît au-dessus de l'	arc	brisé ; ce qui prouve encore combien, pendant l' époque de trar
361	<input type="checkbox"/>	<input type="radio"/>	Viollet Le Duc,...	soins de la construction. La fenêtre inférieure est fermée par un	arc	brisé, parce que cette fenêtre est plus large que l'autre, et que l'
362	<input type="checkbox"/>	<input type="radio"/>	Viollet Le Duc,...	sur les claveaux d'un plein cintre entre les sommiers et la clef ; l'	arc	brisé n'est qu'un moyen de parer au danger d'une rupture vers l
363	<input type="checkbox"/>	<input type="radio"/>	Viollet Le Duc,...	doutaient sans cesse de voir se reproduire ces fâcheux effets. L'	arc	brisé leur paraissait un moyen excellent d' éviter des désastres
364	<input type="checkbox"/>	<input type="radio"/>	Viollet Le Duc,...	notre hypothèse, c'est le relevé exact d'un grand nombre de ces	arcs	brisés primitifs qui donnent exactement une flèche plus longue c
365	<input type="checkbox"/>	<input type="radio"/>	Viollet Le Duc,...	claveaux faible. Or ici, dans les premières voûtes possédant des	arcs	brisés, l'acuité de ces arcs est à peine sensible, tandis que dans
366	<input type="checkbox"/>	<input type="radio"/>	Viollet Le Duc,...	ige, en adoptant l'arc en tiers-point, n'ont rien inventé : il y a des	arcs	brisés dans les monuments les plus anciens de Grèce et d' Étru
367	<input type="checkbox"/>	<input type="radio"/>	Viollet Le Duc,...	acs ou les Romains auraient fait des voûtes engendrées par des	arcs	brisés, qu'est -ce que cela ferait, si le principe général de la con
368	<input type="checkbox"/>	<input type="radio"/>	Viollet Le Duc,...	is diagonales : c'est le plein cintre qui commande la hauteur des	arcs	brisés. Les arcs ogives, soit dit en passant (c'est ainsi qu'on noi
369	<input type="checkbox"/>	<input type="radio"/>	Viollet Le Duc,...	dont ils sont bien les seuls et les véritables inventeurs. Il y a des	arcs	brisés, au XIIe siècle, par toute l' Europe occidentale. Il n' y a dr
370	<input type="checkbox"/>	<input type="radio"/>	Viollet Le Duc,...	ue nous voyons, dans les constructions gothiques primitives, les	arcs	brisés uniquement adoptés pour les arcs doubleaux et les archi

Figure 13 – Résultat de la recherche de concordances contenant le mot *arc*, avec classement par ordre alphabétique sur le mot qui suit *arc*.

Conclusion : Analyse critique et développements futurs



La recherche d'un équilibre formel et d'une homogénéisation des données entre autant de langues différentes est une gageure. Ainsi le choix des textes traduits a dû s'opérer en concertation avec les autres équipes des autres langues. C'est à partir des textes italiens et d'une bibliographie de référence que celles-ci ont procédé à une recherche des traductions des œuvres préalablement répertoriées. Ainsi se sont imposées, pour les textes français, les traductions de Stendhal et de Taine vers les autres langues, pour les textes allemands de Bukhardt et de Goethe, de Ruskin pour l'anglais, etc. Mais ce sont surtout les nombreuses traductions de l'italien Giorgio Vasari (*le Vite de' più eccellenti pittori, scultori, architettori*, 1550) vers les différentes langues qui vont permettre de construire une base de données parallèles essentielle à partir de nos différents corpus (cfr. Zotti 2017). Pour le français, il s'agit pour le moment de 65 vies parmi *Les vies des plus excellents peintres, sculpteurs et architectes*, selon les traductions de L. Leclanché et de C. Weiss, que nous pourrions aussi comparer à celles de A. Chastel. Ces textes, qui ont donné lieu à des traductions et re-traductions de la part de spécialistes d'histoire de l'art dans les différentes langues, ne manqueront pas de constituer un outil précieux pour le recueil des équivalents appartenant à un vocabulaire spécialisé en histoire de l'art. Elles constituent également une documentation riche d'enseignements sur les stratégies traductives adoptées et leur évolution dans le temps (cfr. Ballestracci 2018).

L'observation des textes en traductions que nous souhaitons intégrer à notre corpus a également permis d'identifier certaines approximations qui remettaient en question leur fiabilité. Si nous avons, en effet, considéré les textes traduits comme ayant une valeur et un intérêt comparable aux textes-sources (Frawley 1984), certains types de textes pertinents pour notre recherche (sites internet, guides touristiques) nécessitaient un contrôle de fiabilité linguistique avant d'être intégrés à plein titre dans le corpus. Aussi avons-nous effectué une nécessaire et préalable sélection méthodique, un véritable 'contrôle de qualité', tout au moins dans la mise en place d'une première version publiée du corpus français. Ce dernier a consisté en une comparaison d'un premier sous-corpus des traductions avec le corpus en 'langue naturelle', nous permettant de mettre en évidence des anomalies au niveau des statistiques des occurrences comparées et d'éliminer certains textes (Farina, Billero 2019).

Le fait que, contrairement aux autres langues, le corpus français contienne une part importante de données sous forme de

dictionnaires (alors que la catégorie [DIZ] est encore absente dans les autres langues), ne devrait pas surprendre le public dans la mesure où leur présence reflète une tradition académique nationale dont nous avons tiré avantage. Nous avons également envisagé l'introduction de dictionnaires bilingues lors de la création de corpus parallèles, qui pourraient nous aider dans la recherche d'équivalents traductifs.

Comme nous l'avons dit, notre objectif de départ était de réunir une typologie textuelle étendue aussi bien en synchronie qu'en diachronie. Cette importante variété de typologie qui embrasse plusieurs siècles a impliqué certaines hésitations concernant le classement des sous-corpus. Faut-il distinguer les textes techniques appartenant au domaine de l'architecture, des textes techniques artistiques ? En particulier le dictionnaire déjà mentionné de Viollet-le-Duc et celui de Félibien *Des principes de l'architecture, de la sculpture, de la peinture*, (1676) - ce dernier étant en cours de « nettoyage » -, pourraient plutôt être classés dans la sous-catégorie [TEC\_ART] ou [DIZ\_TEC] (et non pas en [DIZ\_MON] pour le premier et en [TEC\_ARC]) pour le second), c'est-à-dire dans la catégorie des textes de spécialités relevant du domaine artistique en général et/ou qui ont une forme de dictionnaire. Ceci nous conduira sans doute prochainement à fusionner certains sous-corpus.

En outre, le pacte de lecture change en fonction de l'évolution des arts et des techniques et en fonction des moyens de communication ou du médium utilisé. Les textes de Vasari étaient à l'origine des textes destinés aux spécialistes de l'époque et leur diffusion en Europe a permis aussi le développement d'un lexique artistique spécialisé dans toutes les langues et en particulier dans un français de spécialité qui s'est constitué à partir de l'italien depuis le XVI<sup>e</sup> siècle pour se poursuivre au cours des siècles suivants. Aujourd'hui, ils sont parfois plutôt considérés comme des recueils d'anecdotes sur les artistes et leurs œuvres. De même, les guides touristiques du XIX<sup>e</sup> siècle seraient plutôt à classer aujourd'hui, par leur qualité littéraire, comme des récits de voyage.

Par ailleurs, la limite que nous nous étions imposée de prendre en considération prioritairement la ville de Florence et la Toscane, évolue vers un dépassement de celle-ci afin d'intégrer d'autres villes italiennes qui ont suscité l'intérêt des voyageurs étrangers au cours des siècles. Nous avons ainsi commencé à introduire des œuvres, des chapitres ou des fragments relatifs à d'autres villes, en particulier chez Edmond et Jules de Goncourt (*L'Italie d'hier*, 1855), Venise,

Parme, Modène, Bologne, Vérone etc. Cet élargissement nous a permis de tester la possibilité d'une recherche par ville ou région.

Toutes ces remarques illustrent la particularité des corpus que nous publions aujourd'hui, que nous estimons être un premier résultat important après un parcours de recherches que le groupe LBC effectue déjà depuis près d'une décennie et qui porte l'empreinte du *Work in progress*. Il est évident que des ajouts et des ajustements les transformeront dans les prochaines années au vu de la complexité de notre travail en équipes différenciées par langue. La variété des utilisations que nous en ferons entraînera des modifications qui seront mises à l'épreuve de nos recherches futures.

## Bibliographie

Ballestracci Sabrina (2018), Tra norma e variazione: allora, ora e poi in *Vite de' più eccellenti pittori scultori e architettori* di Giorgio Vasari. Un'analisi in ottica contrastiva con il tedesco. In Marco Meli (ed.), *Le norme stabilite e infrante: saggi italo-tedeschi in prospettiva linguistica, letteraria e interculturale*, Firenze University Press, Florence, pp. 21-40.

Biber Douglas (2012), Corpus-based and Corpus-driven Analyses of Language Variation and Use. In *The Oxford Handbook of Linguistic Analysis* Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780199544004.013.0008>

*Bibliographie italico-française universelle, ou catalogue méthodique de tous les imprimés en langue française sur l'Italie ancienne et moderne, depuis l'origine de l'imprimerie (1475-1885)* (1886), Joseph Blanc, Paris.

Bibliothèque nationale de Florence [consultation nov. 2020], *Indici*, site sur le Grand Tour <https://grandtour.bncf.firenze.sbn.it/indici>

Billero Riccardo, Nicolás Martínez Carlota (2017), Nuove risorse per la ricerca del lessico del patrimonio culturale: corpora multilingue LBC. In *CHIMERA Romance Corpora and Linguistic Studies*, Madrid, UAM, 4.2, pp. 203-216.

Billero Riccardo (2020), Cultural Heritage Lexicon: A Case Study. In Ana Pano Alamán, Valeria Zotti, *The language of art and culture heritage: a plurilingual and digital perspective*, Cambridge Scholars Publishing, pp. 86-103.

Castiglione Minischetti Vito, Dotoli Giovanni, Musnik Roger (2002), *Bibliographie du voyage français en Italie du moyen âge à 1914*, Presses

de L'Université de Paris Sorbonne/Schena editore, Paris/Fasano.

Centre de Recherche sur la littérature des voyages, bibliographie des récits de voyages sur : <http://www.crlv.org/>

Farina Annick (2016), Le portail lexicographique du Lessico plurilingue dei Beni Culturali, outil pour le professionnel, instrument de divulgation du savoir patrimonial et atelier didactique. In *PUBLIF@RUM*, vol. 24. <https://www.publiforum.farum.it/index.php/publiforum/article/view/564>

Farina Annick, Billero Riccardo, (2018), Comparaison de corpus de langue « naturelle » et de langue « de traduction » : les bases de données textuelles LBC, un outil essentiel pour la création de fiches lexicographiques bilingues. In *JADT 2018 - International Conference on Statistical Analysis of Textual Data, Roma, 12-15 giugno 2018*, UniversItalia, pp. 108-116.

Farina Annick (2019), *Florence en V.O.*, Atlande, Paris.

Farina Annick, Carolina Flinz (2020), Analisi comparativa dei corpora LBC. La visione francese e tedesca del patrimonio fiorentino: l'esempio del Duomo di Firenze. In Fernando Funari et Annick Farina (eds.), *Il passato nel presente: la lingua dei beni culturali*, Firenze University Press, Florence (in stampa).

Frawley William (1984), Prolegomenon to a theory of translation. In William Frawley (ed.), *Translation: Literary, Linguistic and Philosophical Perspectives*, Univ. of Delaware Press, Newark, pp. 159-175.

Gallica (Bibliothèque nationale de France), Dossier thématique sur les Voyages en Italie <https://gallica.bnf.fr/dossiers/html/dossiers/VoyagesEnItalie/>

Zotti Valeria (2017), L'integrazione di corpora paralleli di traduzione alla descrizione lessicografica della lingua dell'arte: l'esempio delle traduzioni francesi delle Vite di Vasari. In Valeria Zotti, Pano Ana (eds.), *Informatica Umanistica. Risorse e strumenti per lo studio del lessico dei beni culturali*. Firenze University Press, Florence, pp. 105-134.

Sources du Corpus<sup>[13]</sup>

Textes originaux français

Allais, Gustave (1909), *Lamartine en Toscane et les harmonies poétiques et religieuses (d'après la correspondance et les manuscrits)*, Société française d'imprimerie et de librairie, Paris.

Auzias, Dominique, Labourdette, Jean-Paul (2016), *Guide Petit Futé, Carnets de Voyage, Toscane*, Le Petit Futé, Paris.

Bard, Joseph (1854), *Turin, Gênes, Florence, Rome en 1854*, Timon frères, Vienne : 17-24

Bazin, René (1894), *Les Italiens d'aujourd'hui*, Calmann Lévy, Paris.

Beaugrand, Honoré (1889), « Quinzième lettre », *Lettres de voyages*, Presses de La Patrie, Montréal, 1889 : 150-156.

Bertheroy, Jean (1913), « La couronne d'épines », *Le Gaulois*, 26 septembre 1913, n. 13131: 3.

Brosses, Charles de [1739] (1858), *Lettres familières écrites d'Italie à quelques amis, en 1739 et 1740, avec une étude littéraire et des notes par Hippolyte Babou*, vol.1, Poulet-Malassis et de Broise, Paris :

- "Lettre XXIV À M. De Quentin. Mémoire sur Florence": 184-203.
- "Lettre XXIII À M. De Blancey. Route de Bologne à Florence": 204-210
- "Lettre XXVI À M. De Blancey. Route de Florence à Livourne": 210-217.
- "Lettre XXV À M. De Neuilly. Suite du séjour à Florence": 229-236.

Camus, Albert (1938), « Le désert », *Noces*, Gallimard, Paris.

Chateaubriand, François-René de [1827] (1833), *Voyage en Italie* in *Œuvres*, T. XIII, Paris, Lefèvre Libraire.

Colet, Louise (1862), *L'Italie des Italiens*, tome 2, E. Dentu, Paris:

- « Chapitre II » : 13-26.
- « Chapitre III » : 26-36.
- « Chapitre IV » : 36-102.
- « Chapitre V » : 102-136.
- « Chapitre VI » : 136-144
- « Chapitre VII » : 144-153.
- « Chapitre VIII » : 153-189.
- « Chapitre IX » : 190-213
- « Chapitre X » : 213-246.
- « Chapitre XI » : 246-255.
- « Chapitre XII » : 255-267.

- « Chapitre XIII » : 268-281.
- « Chapitre XIV » : 282-292.

Colombari, Colonel F. (1854), « Fête du 16 août à Sienne », *L'illustration, Journal Universel*, n° 602, vol. XXIV, 9 septembre 1854 : 1-2.

Creuzé de Lesser, Augustin Francois [1801] (1806), "La Toscane", in *Voyage en Italie et en Sicilie*, Imprimerie de P. Didot l'aîné, Paris, 1806 : 60-65.

De la Borie, Guillemette (2013), « Sienne vibre pour sa course de chevaux, le Palio », *la-croix.com*, 11 août 2013.

Delacroix, Eugène (1830), « Michel-Ange », *Revue de Paris*, vol. XV, 1830, pp. 41-58.

De Navenne, Ferdinand (1903), *Entre le Tibre et l'Arno*, Paris, Plon :

- « Aux sources de l'Arno et du Tibre » : 1-28
- « À travers l'Apennin toscan » : 29-116.

Dufay, Henri [1853] (1854), « Florence » *Souvenirs de vacances. Naples, Rome, Florence en 1853*, Dubiez, Senlis : 52-63.

Dumas, Alexandre [1841] (1861), *Une année à Florence*, Michel Lévy, Paris:

- « Florence » : 168-176
- « La Pergola » : 177-186.
- « Sainte-Marie-des-Fleurs » : 186-220.
- « Le Palais Riccardi » : 220-235.
- « Le Palais Vieux » : 235-258.
- « La place du grand-duc » : 259-278.

Dumas, Alexandre [1843] (1865), *La Villa Palmieri*, Michel Lévy, Paris.

Erdan, Andrea (1867), « Lettres d'Italie », *Le Temps*, Paris, 20 août 1867, n. 2292 : 4.

Faure, Élie [1923] (1924), *Histoire de l'art III : l'art renaissant (Nouvelle édition, revue et augmentée)*, G. Crès et Cie, Paris.

Favre le Bret, Robert (1938), « Le 'Palio' de Sienne », *Le Temps*, Paris, 1<sup>er</sup> juillet 1938, n° 28052 : 5.

Félibien, André (1676), *Des principes de l'architecture, de la sculpture, de la peinture, et des autres arts qui en dépendent. Avec un Dictionnaire*

*des termes propres à chacun de ces arts*, Paris, J.-B. Coignard :

- « Livre I – Chap. I » : 1-8.
- « Livre I – Chap. II » : 9.
- « Livre I – Chap. III » : 9-13.
- « Livre I – Chap. IV » : 14-17.
- « Livre I – Chap. V » : 18-21.
- « Livre I – Chap. VI » : 22-23
- « Livre I – Chap. VII » : 26-28.
- « Livre I – Chap. VIII » : 30-31.
- « Livre I – Chap. IX » : 33-35.
- « Livre I – Chap. X » : 37-39.
- « Livre I – Chap. XI » : 41-43.
- « Livre I – Chap. XII » : 44-76.

Fernandez, Dominique (1982), *Dans la main de l'ange*, Grasset, Paris.

Fernandez, Dominique (2008), *Dictionnaire amoureux de l'Italie de N à Z*, Plon, Paris.

Fernandez, Dominique (2002), *La course à l'abîme*, Grasset, Fasquelle, Paris.

Fernandez, Dominique (1997), *Le voyage d'Italie. Dictionnaire amoureux*, Paris, Plon :

- « François d'Assise » : 158-160.
- « Jardins » : 192-194.
- « Laurent le Magnifique » : 197-199.
- « Littérature » : 224-237.
- « Sienne 1 » : 576-582.
- « Sienne 2 » : 576-582.

Feuillet, Michel (2009), *L'art italien*, Que sais-je ?, PUF

« Florence » [2017], guide du routard en ligne <[www.routard.com](http://www.routard.com)> [consultation juin 2017]

France, Anatole (1894), *Le lys rouge*, Calmann-Lévy, Paris.

Fréville, Valentin<sup>[14]</sup>, (1872) « VI. A M. ET MADAME BAR, A PARIS », *Les Apennins et la mer Adriatique*, Limoges, Barbou Frères, 1872 : 129-151.

Gaboury, Louise (2015), « Italie, Le Palio : ils sont fous, ces Siennois ! », 11 avril 2015, [ledevoir.com](http://ledevoir.com), [21 février 2017].

Gautier, Théophile (1865), *Quand on voyage*, Michel Lévy frères, Paris :

- « Florence » : 177-238.
- « Venise » : 161-176.

*Geoguide Toscane Ombrie*(2016), Gallimard Loisirs, Paris,

Giono, Jean (1953), « Florence », *Voyage en Italie*, Folio, Gallimard, Paris, 1953.

Goncourt, Jules et Edmond de [1855] (1894), *L'Italie d'hier*, Charpentier & Fasquelle, Paris :

- « Domodossola » : 1-3.
- « Milan » : 5-11.
- « Brescia » : 13-14.
- « Vérone » : 15-18.
- « Venise » : 19-50.
- « Padoue » : 53.
- « Mantoue » : 55-56.
- « Parme » : 57-59.
- « Modène » : 61-62.
- « Bologne » : 63-70.
- « Pistoja » : 71.
- « Florence » : 73-155.

Grandgeorge, Gaston [1907] (1909), *Toscane et Ombrie*, Librairie Plon, Paris :

- « Chapitre I : Pise » : 1-29
- « Chapitre II : Florence » : 30-177.
- « Chapitre IV : Florence » : 242-248.
- « Chapitre V – Sienne » : 249-285.
- « Chapitre VI : le retour » : 286-289.

Grimaldi, Edouard-F. [1844] (1850), *Dix jours à Gênes et huit jours à Florence*, E. Thunot, Paris.

Jaucourt, Chevalier de [1751-1765], « Ecole Florentine », in Diderot, Denis, Alembert, d', *Encyclopédie ou Dictionnaire raisonné des sciences, des arts et des métiers*, vol. 5, Briasson, Paris : 317-318.

Jaucourt, Chevalier de [1751-1765], « Florence », in Diderot, Denis, Alembert, d', *Encyclopédie ou Dictionnaire raisonné des sciences, des arts et des métiers* vol. 6, Briasson, Paris : 876-877.

Klaczko, Julian 1880, *Causeries florentines : Dante et Michel-Ange. Béatrice et la poésie amoureuse*, E. Plon et Cie, Paris :



- « Dante et Michel-Ange » : 1-79.
- « Béatrice et la poésie amoureuse » : 81-155.

La Sizeranne, Robert de (1910), « Les Masques et les Visages-Portraits de Florentine, le long de la Seine et de l'Arno I. XVe siècle. », *Revue des Deux Mondes*, t.60, Paris : 160-190.

Lafenestre, Georges [1870] (1882), « Chapitre 1 La sculpture italienne aux XIIIe et XIVE siècles », in *Maitres anciens : études d'histoire et d'art*, H. Loones, Paris : 1-19.

Lafenestre, Georges (1906), « La Crise de la beauté à Florence au XV<sup>e</sup> siècle », *Revue des Deux Mondes*, 5e période, tome 36 : 528-565.

Lang, Jack et Colin, Lemoine, *Michel-Ange*, Fayard, Paris, 2012.

Lescure, O. [1866] (1906), « Milan, 5 septembre 1866 », *Lettres familières sur Milan, Venise, Florence, etc.*, F. Blétit, Paris : 3-29.

Libri, Guglielmo (1832), « Toscane – Modène – Parme », *Revue des Deux Mondes*, tome 7 : 347-367.

Mallarmé, Camille (1924), « Lettre d'Italie. Le 'Palio' siennois », *Le Gaulois*, Paris, 18 juillet 1924 : 3-4.

Mallet, George [1815] (1817), « Chapitre III – Florence », *Voyage en Italie dans l'année 1815*, J.J. Paschoud, Paris : 47-67.

Maurel, André 1913, *Paysages d'Italie*, tome 1, Hachette, Paris :

- « I. La fenêtre entr'ouverte : Volterra » : 1-23.
- « II. Les fiancés de Sainte Catherine : Sienne » : 24-56.
- « III. Près du sang : Montepulciano » : 57-75.
- « IV. La leçon sur la place : Pienza » : 76-89.
- « V. La fenêtre ouverte : Chiusi » : 90-103.
- « VI. La fenêtre refermée : Corneto » : 104-129.

Méry, Joseph [1853] (1868), *Les Nuits italiennes: contes nocturnes*, Michel Lévy Frère, Paris :

- « Livourne : La vallée de l'Arno » : 17-23.
- « Un dimanche à Florence - La villa Catalani - L'album d'une reine » : 22-34.
- « Les exilés à Florence » : 35-37.
- « L'Atelier de Bartolini » : 45- 56.
- « Arrivée à Florence. Sur l'album de Madame Marini » : 127- 128.
- « Florence. A mon ami Adolphe Sturler, peintre d'histoire » : 129-137.

Meyer, Eric (ed.) (2012), « La Renaissance italienne », *Geo Art*, Prisma media, Paris.

Michel, Émile (1901), « Le Dessin chez Léonard de Vinci », *Revue des Deux Mondes*, 5e période, tome 1 : 342-375.

Montaigne, Michel de [1580-1581] (1774), « de Pescia a Viterbe » in *Journal du voyage de Michel de Montaigne en Italie, par la Suisse et l'Allemagne en 1580 et 1581, avec des notes par M. de Querlon*, Le Jay, Rome/Paris.

Moran, Robert (1994), *Secrets de peintres : apprêts, marouflage, médiums, pigments, glacis, vélatures, dorure, vernis*, Fleurus, Paris.

Musset, Paul de [1843] (1866), « Florence-Pise », *Voyage en Italie et en Sicile, cinquième édition revue et corrigée*, Charpentier, Paris : 282-292.

Nobecourt, Jacques (1969), « Le Palio, fête de la survivance », *Le Monde*, 8 octobre 1969.

Palustre, Léon (1892), « Chapitre III », *L'Architecture de la Renaissance*, Librairies-imprimeries réunies : 93-134.

Perrot, George (1878), « Florence depuis l'annexion au royaume d'Italie », *Revue des Deux Mondes*, 3e période, tome 27 : 111-149.

Pommier, Édouard (2016), « Notes sur la couleur dans la littérature artistique de la renaissance en Italie », in Jackie Pigeaud (dir.), *La couleur, les couleurs. Xles Entretiens de La Garenne-Lemot*, Presses Universitaires de Rennes, Rennes : 81-90

Powell, Véronique Gerard (2007), « Vie d'un autre artiste », in Vasari, Giorgio, *Vies des artistes* (traduction à partir de L. Leclanché de C. Weiss), Grasset, Les Cahiers Rouges, Paris : 7-13.

Prieur, Jean (1934), « La Fête du Palio à Sienne », *L'Auvergne Littéraire, Artistique & Historique*, Clermont-Ferrand, n° 75 : 37-40.

Renan, Ernest [1849] (1875), « Italie », *Voyages, Italie (1849) — Norvège (1870)*, Éditions Montaigne, Paris : 13-119.

Revel, Adèle (1863), « Florence », *Naples, Rome et Florence, impressions de voyage*, Imprimerie de l'union typographique-éditrice, Turin : 55-69.

Rosov, Gricha (2013), *Petit dictionnaire des Beaux-Arts et de la Peinture à l'usage des artistes autodidactes /// version 1.0* Index terminologique [consulté en 2017].

[sa] (1829), « Il.I. Route de Gênes à Lucques par la Rivière du Levant », *Nouveau guide du voyageur en Italie*, Sonzogno, Milan, vol. 1 : 236-244.

Schmitz, Isabelle (2017), « Un été en Toscane : Florence, berceau du foot », *Le Figaro.fr*, 6 juin 2017.

Schmitz, Isabelle (2017), « Un été en Toscane : le Palio de Sienne », *Le Figaro.fr*, 4 juin 2017.

Schmitz, Isabelle (2017), « Un été en Toscane : les contrade, un modèle de société », *Le Figaro.fr*, 5 juin 2017.

Staël-Holstein, Germaine de (1807) « Livre XVIII. Le séjour à Florence » in *Corinne, ou L'Italie*, Vol. 2, H. Nicolle, Paris : 370-409.

Stendhal [1817] (1854), *Histoire de la peinture en Italie*, Michel Lévy Frères, Paris.

Stendhal [1825] (1935), « Lettre de Rome sur l'état actuel de la littérature italienne », in *Courrier anglais*, 4, le Divan, Paris : 232-257.

Stendhal [1817] (1854), *Rome, Naples et Florence*, tome 2, Michel Lévy frères, Paris. (édition revue)

Taillasson, Jean-Joseph (1807), « Léonard de Vinci », *Observations sur quelques grands peintres*, Duminil-Lesueur, Paris : 1-9.

Taine, Hippolyte [1864] (1990), *Voyage en Italie II. D'Assise à Florence*, Éd. Complexe, Bruxelles.

Viollet Le Duc (1854-1868), *Dictionnaire raisonné de l'architecture française du XIe au XVIe siècle*, Bance éditeur, Paris :

- Tome 1 : « Préface » V-XX.
- Tome 1 : « abaque-aronde ».
- Tome 2 : « arts-chapiteaux ».
- Tome 3 : « charnier-console ».
- Tome 4 : « construction-cyborium »
- Tome 5 : « dais-fût ».

Wyzewa, Théodore de (1905), « À propos d'une nouvelle biographie de Giotto », *Revue des Deux Mondes*, 5e période, tome 29, Paris : 459-468.

Textes traduits de l'italien

Cellini, Benvenuto [1558] [trad. 1847] (1881), *La vie de Benvenuto Cellini écrite par lui-même* (traduction de Léopold Leclanché), A. Quantin imprimeur-éditeur, Paris, 188 :

- « Livre I, I » : 3-4.
- « Livre I, II » : 4-6.
- « Livre I, III » : 7-10.
- « Livre I, IV » : 10-11.
- « Livre I, V » : 12-14.
- « Livre I, VI » : 14-16.
- « Livre I, VII » : 16-18.
- « Livre I, VIII » : 18-19.
- « Livre I, IX » : 19-23.
- « Livre I, X » : 23-25.
- « Livre I, XI » : 25-27.
- « Livre I, XII » : 27-29.
- « Livre I, XIII » : 30-34.
- « Livre I, XIV » : 34-36.
- « Livre I, XV » : 36-38.
- « Livre I, XVI » : 38-40.
- « Livre I, XVII » : 40-42.
- « Livre I, XVIII » : 42-44.
- « Livre I, XIX » : 44-48.
- « Livre I, XX » : 48-50.
- « Livre I, XXI » : 50-52.
- « Livre I, XXII » : 52-53.
- « Livre I, XXIII » : 54-56.
- « Livre I, XXIV » : 57-61.
- « Livre I, XXV » : 61-63.
- « Livre I, XXVI » : 63-67.
- « Livre I, XXVII » : 67-70.
- « Livre I, XXVIII » : 70-72.
- « Livre I, XXIX » : 72-76.
- « Livre I, XXX » : 76-82.
- « Livre I, XXXI » : 82-85.
- « Livre I, XXXII » : 85-88.
- « Livre I, XXXIII » : 89-94.
- « Livre II,XXXIV » : 95-99.
- « Livre II,XXXV » : 99-101.
- « Livre II,XXXVI » : 101-104.
- « Livre II,XXXVII » : 104-107.
- « Livre II,XXXVIII » : 107-110.
- « Livre II,XXXIX » : 110-112.
- « Livre II,XL » : 112-116.

- « Livre II, XLI » : 116-119.
- « Livre II, XLII » : 119-121.
- « Livre II, XLIII » : 121-124.
- « Livre II, XLIV » : 125-128.
- « Livre II, XLV » : 129-131.
- « Livre II, XLVI » : 132-135.
- « Livre II, XLVII » : 135-136.
- « Livre II, XLVIII » : 137-138.
- « Livre II, XLIX » : 139-141.
- « Livre II, L » : 141-142.
- « Livre II, LI » : 143-145.
- « Livre II, LII » : 145-149.
- « Livre II, LIII » : 149-151.
- « Livre II, LIV » : 151-152.
- « Livre II, LV » : 152-154.
- « Livre II, LVI » : 154-157.
- « Livre II, LVII » : 158-161.

Vasari, Giorgio [1568] (trad. 1900), *Les vies des plus excellents peintres, sculpteurs, et architectes* (traduction de Weiss, Charles), Dorbon-Aine, 1900, Paris

- « Alesso Baldovinetti » : 389-391.
- « Andrea dal Castagno di Mugello et Domenico Viniziano » : 402-407.
- « Antonio et Bernardo Rossellino » : 418-421.
- « Antonio Filarete et Simone » : 364-367.
- « Arnolfo di Lapo » : 130-138.
- « Benozzo Gozzoli » : 413-416.
- « Berna » : 224-226.
- « Buonamico Buffalmacco » : 186-190.
- « Cosimo Rosselli » : 437-439.
- « Dello » : 277-279.
- « Desiderio da Settignano » : 422-423.
- « Don Bartolommeo, Abbé de San Clemente et Gherardo » : 439-443.
- « Don Lorenzo Monaco » : 247-249.
- « Donato » : 347-358.
- « Fra Giovanni da Fiesole » : 375-381.
- « Francesco di Giorgio et Lorenzo Vecchietto » : 416-418.
- « Gentile da Fabriano et Vittore Pisanello de Vérone » : 407-410.
- « Gherardo Starnina » : 243-245.
- « Giovanni Cimabue » : 123-129.
- « Giuliano da Maiano » : 367-370.

- « Jacopo della Quercia » : 267-274.
- « Jacopo, Giovanni et Gentile Bellini » : 430-436.
- « Leon-Batista Alberti » : 381-384.
- « Lippo » : 245-246.
- « Lorenzo Costa » : 427-430.
- « Lorenzo di Bicci » : 251-256.
- « Luca della Robbia » : 281-288.
- « Margaritone » : 154-157.
- « Masaccio » : 311-317.
- « Mino da Fiesole » : 424-427.
- « Niccola et Giovanni » : 138-149.
- « Pesello et Francesco Peselli » : 411-412.
- « Piero della Francesca » : 370-374.
- « Pietro Cavallini » : 193-195.
- « Pietro Laurati » : 179-181.
- « Simone et Lippo Memmi » : 195-202.
- « Spinello Aretino » : 235-243.
- « Stefano et Ugolino » : 176-179.
- « Taddeo Bartoli » : 249-251.
- « Vellano de Padoue » : 391-393.

Vasari, Giorgio [1568] (trad. 1900), *Vies des artistes* (traduction à partir de L. Leclanché de C. Weiss), Grasset, Les Cahiers Rouges, Paris, 2007

- « Andrea Pisano » : 38-42.
- « Andrea Mantegna » : 161-167.
- « Antonello da Messina » : 129-133.
- « Antonio et Piero Pollaiolo »
- « Cimabue » : 15-22.
- « Domenico Ghirlandaio » : 134-147.
- « Filippo Brunelleschi » : 75-109.
- « Fra Filippo Lippi » : 393-401.
- « Fra Giovanni da Fiesole (Fra Angelico) » : 122-128.
- « Ghiberti » : 50-66.
- « Giorgione da Castelfranco » : 194-198.
- « Giotto » : 23-37.
- « Giovannantonio da Verzelli » : 122-128.
- « Jules Romain » : 322-335.
- « Le Rosso » : 122-128.
- « Le Titien » : 439-459.
- « Léonard de Vinci » : 178-193.
- « Marc Antonio » : 295-321.
- « Masaccio » : 67-74.

- « Michel-Ange » : 346-438.
- « Paolo Uccello » : 43-49.
- « Pietro Perugino » : 168-177.
- « Raphaël d'Urbino » : 220-254.
- « Sandro Botticelli » : 155-160.

Vasari, Giorgio [1568] (trad. 2002), « Giotto », in *Vies d'artistes : édition bilingue français-italien* (traduction de Gérard Luciani), Gallimard, Paris, 2002 : 77-93.

Vasari, Giorgio [1568] (trad. 1839), « Giotto », in *Vies des peintres, sculpteurs et architectes* (traduction Léopold Leclanché), Jeanron, Philippe-Auguste, Paris, 1839-42 : 201-225.

## Notes

[1] Les corpus chinois et portugais n'ayant pas encore atteint le million de mots que nous avons posé comme nécessaire pour la mise en ligne, ils ne seront publiés que dans une phase successive de notre projet.

[2] Plusieurs étudiantes de Valeria Zotti et d'Annick Farina ont travaillé par exemple sur les traductions de G. Vasari en français. Une étudiante suivie par Annick Farina pour le français et par Sabrina Ballestracci pour l'allemand a travaillé récemment à la comparaison des traductions de la correspondance de Michel-Ange dans ces deux langues et ses transcriptions seront bientôt intégrées. Nous avons aussi suivi plusieurs mémoires sur les textes de voyageurs français tant à Florence (A. Farina) qu'à Bologne (V. Zotti).

[3] Ces découvertes ont servi de base à la rédaction d'un ouvrage sur Florence et ses mots-clés à travers les auteurs qui l'ont décrite (Farina 2019).

[4] Citons en particulier la *Bibliographie italico-française...*(1886), la bibliographie du Grand Tour disponible sur le site de la Bibliothèque nationale de Florence, celle du site Centre de Recherche sur la littérature des voyages, le dossier thématique sur les Voyages en Italie du site Gallica de la Bibliothèque nationale de France et Castiglione Minischetti, Dotoli, Musnik (2002).

[5] Le site archive.org a l'avantage de donner accès à des ressources présentes dans plusieurs bibliothèques, nous avons ainsi pu choisir

l'édition avec laquelle l'océrisation était la meilleure lorsque ce choix de ressources se présentait.

[6] Dans SketchEngine, qui est le logiciel utilisé pour notre corpus, les données statistiques sont données en nombre de 'tokens' et/ou de 'mots' selon les types de recherche. Un 'token' y est défini comme "the smallest unit that a corpus consists of. A token normally refers to: a word form: going, trees, Mary, twenty-five... / punctuation: comma, dot, question mark, quotes.../ digit: 50,000.../ abbreviations, product names: 3M, i600, XP, FB...anything else between spaces / There are two types of tokens: words and nonwords. Corpora contain more tokens than words. Spaces are not tokens. A text is divided into tokens by a tool called atokenizer which is often specific for each language"  
[https://www.sketchengine.eu/my\\_keywords/token/](https://www.sketchengine.eu/my_keywords/token/) [consultation nov. 2020]

[7] sans auteur

[8] Il s'agit des traductions des *Vies* de G. Vasari (65 vies d'artiste ont été intégrées, le travail de transcription est encore en cours), de B. Cellini, *La vie de Benvenuto Cellini* (Livres I et II), et une recette de P. Artusi. Ces textes représentent 303 262 mots donc environ 10% du corpus.

[9] Pourcentage calculé sur le nombre total de textes hors traductions présent dans le corpus (3 443 357)

[10] Seuls les textes en langue 'originale' ont été pris en considération : avec les textes traduits, le problème se pose de l'époque de rédaction du texte original qui ne coïncide pas avec celle de rédaction de la traduction et donc des caractéristiques linguistiques de la langue utilisée qui varie selon les dates. Lors de la consultation des données il est possible de filtrer ces textes soit sur la base de l'année de rédaction de l'original soit sur celle de sa traduction.

[11] L'interface de consultation du corpus est actuellement en italien mais elle sera disponible sous peu dans la langue de chaque corpus.

[12] Nous renvoyons aux articles de Riccardo Billero sur le choix du logiciel NosketchEngine et sur la gestion des données et métadonnées de nos corpus (Billero, Nicolás Martínez 2017 et Billero 2020).

[13] Il s'agit de la bibliographie des textes disponibles en novembre 2020, certains sont encore en train d'être encodés avant publication



et ne seront consultables que dans les prochains mois. Un sous-corpus contenant un grand nombre de textes sur Bologne, préparés par Valeria Zotti, seront intégrés pendant l'année 2021.

[14] Pseudonyme de Alfred Driou.



---

Farina, Annick. Corpus LBC Français

© 2024 - Author(s) | Published by Firenze University Press

e-ISBN: 979-12-215-0309-8 | DOI: 10.36253/979-12-215-0309-8

Content license: CC BY-SA 4.0 International | Metadata license: CC0 1.0 Universal