

# Il corpus LBC tedesco: costruzione e possibili applicazioni

Sabrina Ballestracci, Claudia Buffagni, Carolina Flinz<sup>1</sup>

**Abstract (Italiano):** In questo contributo si presentano le caratteristiche principali e le possibili applicazioni del Corpus LBC tedesco, creato nell'ambito del progetto *Lessico multilingue dei Beni Culturali* (LBC). Il lavoro si compone di tre paragrafi: il primo paragrafo disamina i generi testuali, gli autori e i titoli presenti nel corpus, illustra i criteri di selezione adottati per la loro scelta e propone una valutazione dell'attuale configurazione del corpus. Successivamente (par. 2) ci si sofferma sui possibili usi del corpus, mettendo in evidenza come esso si presti ad analisi di tipo sia intralinguistico sia interlinguistico, applicabili sia in ambito didattico sia in contesti professionali. Chiudono il lavoro un riassunto e uno sguardo sulle prospettive di sviluppo futuro (par 3).

**Parole chiave:** linguistica dei corpora, didattica DaF, lessico, terminologia, arte, beni culturali.

**Abstract (English):** This essay describes the main features of the German LBC corpus, which was compiled within the research project *Lessico multilingue dei Beni Culturali* (LBC) and illustrates its possible applications. The paper is divided into three sections: Section 1 illustrates the German LBC corpus, paying special attention to the selected text genres and authors and their selection criteria; it also proposes an evaluation of the current corpus configuration. Section 2 presents various possible applications of the corpus: The proposed applications are both intralingual and interlingual. Conclusion and outlook (Section 3) conclude the work.

**Keywords:** Corpus Linguistics, German as Foreign Language Teaching, Lexicon, Terminology, Art, Cultural Heritage.

## 1. Il corpus tedesco di LBC. Costruzione

La piattaforma del corpus LBC (<<http://corpora.lessicobeniculturali.net>>) è una piattaforma multilingue specializzata composta da corpora monolingui in più lingue (tedesco, inglese, francese, italiano, russo, spagnolo), contenente

<sup>1</sup> Il presente contributo è la traduzione di un più ampio saggio in tedesco dal titolo *Das deutsche LBC-Korpus: Zusammenstellung und Anwendung*, pubblicato in Ballestracci, Buffagni, Flinz 2020. Autrice della traduzione è Sara Congregati. Come il saggio in tedesco, così anche il presente contributo è frutto della concezione comune delle tre autrici, che ne hanno discusso

sia testi originali sia traduzioni (cfr. Farina, Garzaniti 2013; Farina 2016; Billero, Martínez 2017; Billero, 2020; cfr. anche Figura 1):

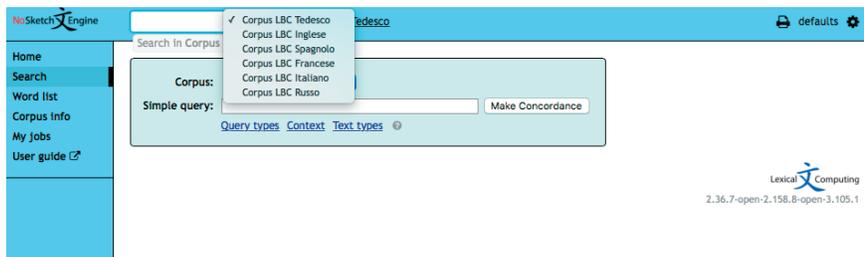


Figura 1. Screenshot dei corpora LBC nell'ambiente NoSketchEngine.

Scopo principale del progetto LBC è la redazione di un dizionario multilingue in rete, che possa anche servire da fonte per ricerche scientifiche (linguistiche, letterarie, culturali ecc.), offrendo al largo pubblico testi per una migliore conoscenza dei Beni culturali italiani. La banca dati, dunque, non solo è funzionale alla creazione di un dizionario da parte dell'unità di ricerca LBC, ma si rivolge anche a utenti esterni, come esperti del settore in storia dell'arte, italianistica ecc., traduttori professionisti, e anche insegnanti di traduzione specialistica.

È a partire da queste premesse che i singoli corpora LBC mirano a illustrare il lessico dei Beni culturali italiani. Per questa ragione si è posta fin dall'inizio particolare attenzione sulla lingua del Rinascimento e dell'arte fiorentina, nonché sulle sue opere e i suoi autori famosi (cfr. paragrafo 1.1 e paragrafo 1.2).

### 1.1 Generi testuali

Il corpus tedesco di LBC – come gli altri corpora LBC – si compone sia di testi in lingua originale tedesca sia di testi tradotti da altre lingue del progetto (italiano, francese e inglese<sup>2</sup>). Di questi testi alcuni erano disponibili in versione cartacea nelle nostre biblioteche universitarie o private, altri sono stati scaricati da archivi testuali online. Delle fonti online impiegate fanno parte rinomati archivi testuali come, per esempio, il *Deutsches Textarchiv* (<[in dettaglio le singole parti. Sabrina Ballestracci ha curato in particolare i paragrafi 1.1 e 1.2, Claudia Buffagni il paragrafo 1.3, Carolina Flinz il paragrafo 2. Alla stesura del paragrafo 3 e del cappello del paragrafo 1 hanno collaborato tutte e tre le autrici. Desideriamo qui ringraziare Katharina Müller e Anna Nissen per gli utili spunti di riflessione e i suggerimenti offerti nel corso di uno scambio proficuo.](http://www.deutsche-</a></p>
</div>
<div data-bbox=)

<sup>2</sup> Fanno parte dell'intero progetto anche altre lingue straniere, in particolare lo spagnolo e il russo; il corpus tedesco di LBC già caricato online non contiene tuttavia testi tradotti da queste lingue (cfr. anche fig. 6 nel paragrafo 1.2). Un ampliamento del corpus in tal senso sarà uno degli obiettivi delle prossime fasi di lavoro (cfr. paragrafo 1.3 e paragrafo 3).

stextarchiv.de/>), che dispone soprattutto di testi in formato PDF, e il *Projekt Gutenberg* (<<https://www.projekt-gutenberg.org/index.html>>), dove tutti i testi sono presenti in formato HTML. I testi in versione cartacea sono stati scannerizzati e dunque importati in file PDF. Tutti i file PDF – sia quelli attinti da fonti internet, sia quelli di provenienza cartacea – sono stati convertiti in formato Word tramite OCR, i testi del sito web del Progetto Gutenberg sono stati copiati direttamente in un file Word<sup>3</sup>. Ogni singolo esemplare testuale è stato sottoposto a revisione formale, ovvero linguistica.

Attualmente il corpus tedesco di LBC contiene 763 file Word (in totale 1.183.484 token)<sup>4</sup>, che sulla base del genere testuale di appartenenza possiamo suddividere in tre macrocategorie: divulgativa, tecnico-specialistica e letteraria (cfr. Tabella 1 e Figura 2).

Tabella 1. Ripartizione dei token nelle tre macrocategorie (frequenza assoluta).

Categorie	Frequenza assoluta
Divulgativa	17.450
Tecnico-specialistica	526.516
Letteraria	639.518
Totale	1.183.484

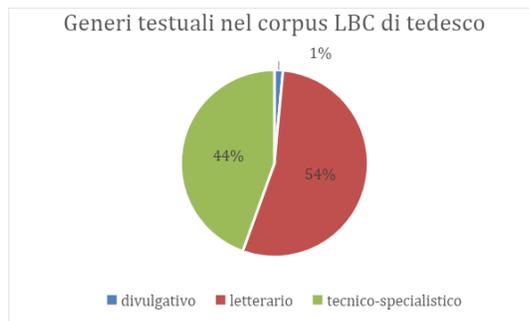


Figura 2. Ripartizione dei token nelle tre macrocategorie (percentuale).

<sup>3</sup> Per la realizzazione di questa prima fase della ricerca è stato fondamentale il lavoro di dottorandi e tirocinanti dell'Università di Firenze. Il nostro particolare ringraziamento va qui alla dott.ssa Benedetta Bronzini, alla dott.ssa Sara Congregati, a Valentina Baldi, Domenica Barone, Delia Papandrea, Maria Chiara Susini, Clara Sabatino, Giulia Giannattasio e Lisa Bandinelli.

<sup>4</sup> Le cifre si riferiscono allo stato del corpus al 31 agosto 2018 (cfr. anche paragrafo 3.1). Da allora il corpus tedesco di LBC è stato fortemente ampliato; i nuovi testi elaborati non sono tuttavia ancora disponibili online. Raggiunta quota 1.000.000 di token, il corpus online non è stato ulteriormente incrementato per permettere di ancorare il lavoro sulle concordanze a una raccolta di testi circoscritta (cfr. Flinz, Buffagni, Ballestracci 2020). Tale decisione è estesa a tutte le lingue del progetto.

Tabella 1 e Figura 2 mostrano che il genere più rappresentato nel corpus tedesco è quello letterario (54%). Ben rappresentato è anche il genere tecnico-specialistico con una percentuale del 45%. I meno rappresentati nel corpus LBC tedesco sono i testi divulgativi con una percentuale dell'1%.

Rientrano nei testi di genere letterario biografie, opere narrative e saggi secondo la ripartizione qui di seguito illustrata (Figura 3):



Figura 3. Ripartizione dei testi letterari nelle tre categorie (percentuale).

Come possibile osservare in Figura 3, la categoria più ampia, con una percentuale del 55%, è la biografia, mentre le altre due categorie letterarie – *fiction* e saggistica – presentano una percentuale più bassa, pressoché equivalente (rispettivamente del 25% e del 20%).

Rientrano nei testi biografici sia biografie vere e proprie sia testi più vicini ad altri generi testuali, per esempio diari e resoconti di viaggio. Questo gruppo di testi accoglie in egual misura testi in traduzione – per esempio dall'italiano *Das Leben des Benvenuto Cellini, florentinischen Goldschmieds und Bildhauer. Von ihm selbst geschrieben* (2016)<sup>5</sup> di Benvenuto Cellini – oppure dal francese – *Reise in Italien (Rome, Naples et Florence en 1817)* (1922) di Stendhal – così come testi originali in lingua tedesca, ad esempio *Das Florenzer Tagebuch* (1984) di Rainer Maria Rilke.

Si annoverano fra i testi di *fiction* novelle, romanzi e altri testi di narrativa letteraria dove Firenze o la Toscana sono al centro o sullo sfondo delle vicende narrate. All'interno di questa categoria figurano soltanto testi di autori tedesco-foni. Ne sono un esempio *Florentiner Novellen* (1984) di Isolde Kurz o *Florentinische Nächte* (1969) di Heinrich Heine.

Anche i testi di saggistica includono soltanto testi in lingua originale tedesca, per esempio *Stadt des Lebens* (1905) di Isolde Kurz, oppure *Geschichte der Renaissance in Italien* (1891) di Jacob Burckhardt.

La categoria dei testi tecnico-specialistici comprende testi di varia provenienza settoriale: arte, enogastronomia e storia (cfr. Figura 4):

<sup>5</sup> Da qui in poi le indicazioni bibliografiche fanno riferimento alle edizioni dei testi contenuti nel corpus.

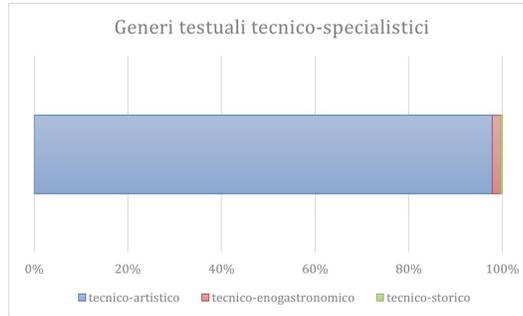


Figura 4. Ripartizione dei testi tecnico-specialistici (percentuale).

Un'alta percentuale è rappresentata dai testi tecnico-specialistici di arte (oltre 500.000 token; 98%). I testi tecnico-specialistici di enogastronomia e di storia sono invece molto poco rappresentati (in totale 11.134 token; 2%).

I testi tecnico-specialistici mostrano una varietà minore rispetto a quelli letterari: si tratta esclusivamente di traduzioni dall'italiano, come per esempio *Der Traktat von der Malerei* (1909) di Leonardo da Vinci, *Kleinere kunsttheoretische Schriften* (1877) di Leon Battista Alberti, *Von der Wissenschaft des Kochens und der Kunst des Genießens* (2000) di Pellegrino Artusi e *Geschichte von Florenz* (1934) di Niccolò Machiavelli.

Per quanto riguarda i testi divulgativi abbiamo al momento un unico esemplare: si tratta di una buona guida turistica, relativamente esaustiva (17.412 token) in lingua originale tedesca, *Florenz. Perfekte Tage in der Toskana – Metropole* (2016) di Ciccarelli, Jepson e Fisher<sup>6</sup>.

La scelta dei testi è stata spesso influenzata dalla reperibilità degli stessi; ciononostante abbiamo tentato di garantire una certa varietà per tutti i generi letterari, privilegiando quelle tipologie testuali particolarmente importanti per le finalità del progetto. Per quanto possibile, sono stati rispettati criteri quali la paternità e l'importanza dell'opera per la cultura e l'arte italiana, come pure la diffusione e la ricezione a livello internazionale. Nel caso di alcuni esemplari l'attribuzione del genere testuale di appartenenza ha presentato non poche difficoltà al gruppo di lavoro. Basti solo citare, a titolo esemplificativo, l'opera di Giorgio Vasari (1966 [1568]), che rappresenta il testo fondante del progetto, ovvero il testo da cui le singole unità di ricerca – divise per lingua – sono partite per l'acquisizione dei dati (edizione tedesca: *Leben der ausgezeichnetsten Maler, Bildhauer und Baumeister von Cimabue bis zum Jahre 1567*, 2008). Questi testi – come altri del corpus – sono caratterizzati da una forte commistione di generi testuali, il cui stile oscilla fra lo specialistico-descrittivo, il biografico-narrativo e il critico-valutativo. Il testo è stato attribuito alla categoria del genere tecni-

<sup>6</sup> In una prima fase di lavoro sono state prese in esame anche altre guide turistiche in seguito escluse di necessità dal corpus per mancanza di informazioni sulle caratteristiche della lingua impiegata (se per esempio da traduzione o da testo originale).

co-specialistico, da un lato, per il criterio delle peculiarità stilistiche prevalenti, dall'altro, per il criterio dello scopo insito nel progetto stesso. Quest'ultimo criterio ha guidato progressivamente anche la scelta dei singoli autori e delle singole opere presenti nel corpus (cfr. paragrafo 1.2).

## 1.2 Autori e opere

Nel selezionare gli autori e le opere che costituiscono il corpus LBC abbiamo scelto di privilegiare quelli particolarmente rappresentativi per l'arte e la cultura del Rinascimento. Con questo intento tutti i gruppi di lavoro LBC sono partiti all'inizio del progetto da Giorgio Vasari e dalle sue *Vite* (1550, 1568), cioè da un'opera fondamentale per l'arte e la cultura del Rinascimento, che ha contribuito alla diffusione del mito del Rinascimento italiano in gran parte dei Paesi europei – anche in ambito tedescofono (cfr. Wackenroder 1991 [1797]; Fiorillo 1798; cfr. anche Ballestracci, Buffagni 2016). Questo si è rivelato un vantaggio per l'intero progetto, anche perché l'opera di Giorgio Vasari esiste in tutte le lingue comprese nel progetto (italiano, tedesco, inglese, francese, spagnolo, russo e cinese), cosa che in futuro potrebbe anche consentirci di costruire un corpus parallelo delle *Vite* per ulteriori ricerche (cfr. paragrafo 1.3 e paragrafo 2). Fra le altre opere e gli altri autori italiani che abbiamo scelto per il nostro corpus vi sono le *Opere Volgari* (1849) di Leon Battista Alberti, il *Trattato della pittura* (1651) di Leonardo da Vinci e le *Istorie Fiorentine* (1971) di Niccolò Machiavelli. Si è cercato, per quanto possibile, di ritrovare gli stessi autori in tutte le lingue.

Si ritengono fondamentali per il progetto anche autori non italiani che hanno avuto un ruolo di rilievo nella diffusione della cultura italiana del Rinascimento nei Paesi stranieri: ne fanno parte per esempio testi tradotti da altre lingue come *Wege zur Kunst* (1897) di John Ruskin o testi in lingua originale tedesca come *Die Geschichte der Renaissance in Italien* (1891) di Jacob Burckhardt<sup>7</sup>.

Un altro gruppo importante è rappresentato da autori o letterati famosi che hanno fatto viaggi in Italia e hanno scritto un diario o un resoconto di viaggio. La loro importanza risiede soprattutto nel fatto di aver descritto l'arte e la cultura italiane con occhi stranieri, contribuendo così alla diffusione dei Beni culturali e artistici soprattutto nel XIX secolo, ma anche fino ai giorni nostri, in realtà. Fra gli autori di testi letterari del corpus LBC tedesco si annoverano letterati famosi come, per esempio, Johann Wolfgang Goethe (*Italienische Reise*, 1997), Karl Philipp Moritz (*Reisen eines Deutschen in Italien in den Jahren 1786 bis 1788*, 2015), Fanny Lewald (*Bilderbuch*, 1992), Karl Eugen Gass (*Pisaner Tagebuch. Aufzeichnungen/Briefe*, 1961) e Stendhal (*Reise in Italien (Rom, Naples et Florence en 1817)*, 1922).

<sup>7</sup> L'opera di Burckhardt nel corpus LBC tedesco non figura ancora in versione integrale a causa delle varie impurità del file Word importato dal formato PDF, che hanno pesato non poco sulla revisione formale e linguistica del testo.

Per tutti gli autori e le loro opere sono stati scelti nelle prime fasi di lavoro solo brani nei quali spiccano Firenze, la Toscana e l'arte del Rinascimento. I criteri inizialmente adottati sono stati in seguito ampliati: da un lato, si è deciso di prendere in considerazione non solo Firenze e la Toscana ma l'Italia tutta, pur mantenendo al centro dell'attenzione l'arte e la cultura; dall'altro, sono stati presi in esame non solo l'arte e la cultura del Rinascimento ma anche l'arte e la cultura di altre epoche storiche. Per questo il corpus da noi impiegato per l'elaborazione delle concordanze (cfr. Buffagni, Flinz, Ballestracci 2020) e qui descritto è l'immagine del corpus in un dato momento, così come rappresentato in Figura 5:

<u>doc.author</u>	<u>document frequency</u>
Cellini, Benvenuto	237
Kurz, Isolde	169
Vasari, Giorgio	126
Moritz, Karl Philip	66
Artusi, Pellegrino	46
Rilke, Rainer Maria	25
Ciccarelli, Jepson, Fisher	21
Brandt, Karl	18
Heine, Heinrich	12
Burckhardt, Jacob	10
Alberti, Leon Battista	7
da Vinci, Leonardo	7
Lewald, Fanny	6
Ruskin, John	4
Stendhal	3
Gass, Karl Eugen	3
Machiavelli, Niccolò	2
Goethe, Johann Wolfgang von	1

Figura 5. Autori nel corpus LBC tedesco.

Come evidenziato in Figura 5 il corpus LBC tedesco comprende in totale 18 autori: di questi, dieci sono tedeschi (Kurz, Moritz, Ciccarelli, Jepson e Fisher, Brandt, Heine, Burckhardt, Gass e Goethe), mentre i restanti otto sono o italiani (Cellini, Vasari, Artusi, Alberti, da Vinci e Machiavelli) o di altre lingue, come Stendhal (francese) e Ruskin (inglese) (cfr. Figura 6).

<u>doc.language</u>	<u>document frequency</u>
Italiano	425
Tedesco (Germania)	331
Inglese (Regno Unito)	4
Francese (Francia)	3

Figura 6. Lingue di partenza rappresentate nel corpus LBC tedesco.

La Figura 6 mostra come dal corpus LBC tedesco restino ancora esclusi testi in traduzione da alcune lingue del progetto (es. spagnolo, russo e cinese); queste

dovranno venir prese in considerazione nelle prossime fasi di lavoro. Allo stato attuale il corpus LBC tedesco offre tuttavia testi che garantiscono una varietà diacronica relativamente ricca, essendo stati scritti o pubblicati fra il 1400 e il 2016 (cfr. Figura 7).

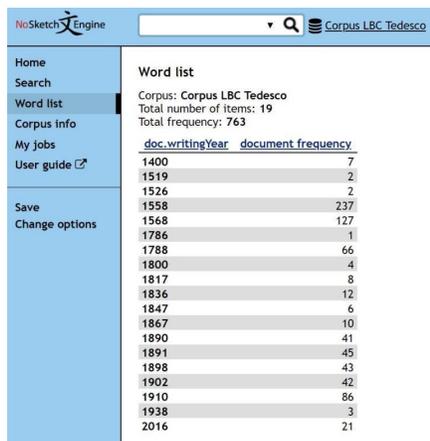


Figura 7. Lingue di partenza rappresentate nel corpus LBC tedesco.

Pur necessitando di alcuni interventi per vari aspetti, il corpus LBC tedesco evidenzia altrettanti aspetti che allo stato attuale non possono che trovare una valutazione positiva (cfr. paragrafo 1.3).

### 1.3 Lo stato attuale del corpus. Prime conclusioni

Riassumendo, la piattaforma del corpus LBC mira, unitamente alla redazione di un dizionario multilingue in rete, alla reperibilità di testi rilevanti per la conoscenza dei Beni culturali italiani, così da offrire una fonte liberamente accessibile per ricerche di tipo linguistico, letterario e culturale.

Come si evince dal paragrafo 1.1, più della metà dei testi compresi nel corpus appartiene al genere letterario, la quasi totalità dei testi rimanenti è riconducibile invece al genere tecnico-specialistico, e solo una minima parte al settore divulgativo. Tale configurazione potrebbe a prima vista apparire non del tutto equilibrata per la preponderanza dell'ambito letterario. Vanno tuttavia sottolineate innanzitutto le motivazioni anche genetiche dell'attuale configurazione della banca dati: la maggior parte dei testi letterari presenti provengono infatti da archivi online molto prestigiosi e facilmente accessibili; così non è per molti dei testi tecnico-specialistici. È sembrato dunque sensato nella prima fase di lavoro partire dalla raccolta dei testi letterari. La struttura attuale della banca dati che, come evidenziato, registra la situazione del corpus in un dato momento, corrisponde dunque in primo luogo a questo modo di procedere.

La bontà della decisione presa si è palesata anche in un'altra ottica. Dei testi letterari del corpus LBC tedesco va notato innanzitutto che si contraddistinguono per la loro varietà: includono prevalentemente biografie, la restante parte è data invece dalla somma dei generi testuali di *fiction* e saggistica. Le biografie assumono pertanto un'importanza notevole, a queste sono stati infatti aggiunti i testi di viaggio. Come nel caso delle altre lingue rappresentate nella banca dati (es. il francese), le descrizioni e i resoconti di viaggio svolgono sul piano sia quantitativo che qualitativo un ruolo particolare in relazione agli obiettivi del corpus LBC tedesco: quantitativo, poiché (cfr. paragrafo 1.2) costituiscono una parte cospicua del corpus; qualitativo, in quanto riportano la prospettiva di osservatori stranieri di fronte alle opere d'arte italiane, e fiorentine in particolare (si tratta spesso di viaggiatori del XVIII e XIX secolo, come, per esempio, il consigliere segreto Goethe o l'irrequieto soldato Stendhal). Le descrizioni che scrittori e artisti tedeschi, francesi e inglesi hanno dedicato alle opere d'arte italiane sono testimonianze dell'epoca e della vita degli autori; possiedono un elevato valore letterario e hanno contribuito alla ricezione dell'immagine dell'Italia nel mondo. Hanno infine portato alla creazione di un vero e proprio «mito dell'Italia» che avrà ampie ripercussioni sull'arte e la cultura europea ed extraeuropea. Ricordiamo fra questi anche *Wege zur Kunst* (1897, traduzione dall'inglese) di Ruskin. Le altre biografie presenti nel corpus, che comprendono le vite degli artisti del Rinascimento, si presentano da parte loro come «tratteggio di un percorso di vita» e sono ascrivibili a un'epoca più antica (per esempio le *Lebensbeschreibungen* di Vasari e l'autobiografia di Cellini). Opere architettoniche, sculture, dipinti, affreschi, stucchi e ornamenti vengono descritti e valutati rispettivamente secondo criteri indicati dagli autori. Queste opere comprendono «elementi [...] di storiografia e letteratura» (Best 1994, 71) e contengono a tal riguardo preziose informazioni storiche. Entrambe le tipologie di genere biografico possiedono un inestimabile valore storico-culturale, poiché permettono di comprendere lo sviluppo della storia dell'arte, come pure della storiografia dell'arte, fra il XV secolo e il presente.

Nel corpus si trovano anche, classificati come generi letterari – nella sottocategoria dei testi di saggistica – classici moderni della storiografia dell'arte come la *Geschichte der Renaissance in Italien* (1891) di Burckhardt. L'opera di Burckhardt, che ha avuto un'influenza notevole su altri autori presenti nel corpus (es. Isolde Kurz), si distingue per le neoacquisite conoscenze scientifiche dell'epoca e offre un'immagine dell'arte del Rinascimento ancor più ricca di sfaccettature. Sul piano stilistico mostra un lessico variegato, duttile, che accanto ai termini tecnici presenta anche peculiarità estetiche.

Sono poi da menzionare come generi di *fiction* con valore prototipico in ambito letterario i testi narrativi presenti nel corpus, fra cui il racconto *Florentinische Nächte* di Heinrich Heine (1969) e il ciclo di novelle *Florentiner Novellen* (1984) di Isolde Kurz. La città di Firenze con gli usi e costumi della sua popolazione fanno spesso da sfondo alle storie narrate, il cui ampio spettro di varianti si esprime, fra l'altro, nella possibilità di venir ambientate in epoche passate (come nelle novelle di Kurz) o nel presente (come nel racconto di Heine).

Fra i generi testuali tecnico-specialistici, tutti in traduzione, spiccano in particolare i *Kleinere kunsttheoretische Schriften* (1877) di Leon Battista Alberti e *Traktat von der Malerei* (1909) di Leonardo da Vinci: il primo segna una svolta importante nella storiografia dell'arte dell'epoca, avendo introdotto un lessico tecnico innovativo (architettura, scultura e pittura); l'altro rappresenta una riflessione approfondita sull'essenza della pittura che proprio all'epoca vuole affermarsi accanto alla poesia come arte di pari dignità (Gombrich 2000). La *Geschichte von Florenz* (1934) di Niccolò Machiavelli costituisce l'unico testo tecnico di contenuto storico, mentre *Von der Wissenschaft des Kochens und der Kunst des Genießens* (2000) di Pellegrino Artusi è il solo testo dedicato alla gastronomia. Si tratta in entrambi i casi di opere di grande valore culturale. A questo proposito si rivelano rappresentativi per un lungo lasso di tempo (dal 1400 al 1891) anche i generi testuali tecnico-specialistici che occupano quasi la metà del corpus, mentre si annovera fra i testi divulgativi una guida turistica contemporanea in lingua tedesca, *Florenz. Perfekte Tage in der Toskana – Metropole* (2016) di Ciccarelli, Jepson e Fisher.

Come emerge chiaramente da quanto argomentato sopra, la stragrande maggioranza dei testi presenti nel corpus è accomunata da un elevato valore stilistico che contraddistingue la banca dati multilingue e il corpus tedesco. Proprio lo stile alto possiede un valore intrinseco ed è rilevante soprattutto per la storia dell'arte e la storiografia artistica. Non a caso Reiners contrappone allo «stile enciclopedico privo di visione» il «testo creato da un buon scrittore di prosa», per giungere alla conclusione: «Le opere di storia dell'arte contrappongono immagini di uno stile [dello stile vivace, plastico, n.d.a.] all'altro [lo stile enciclopedico, n.d.a.] per metterne in risalto le peculiarità. Anche nel regno della parola la bellezza della grande prosa risalta con maggior chiarezza sullo sfondo dello stile medio<sup>8</sup>» (Reiners 1991, 219 s.).

E quanto più diversi tra loro si presentano i generi, tanto più ricca si rivela la stragrande maggioranza dei testi all'interno del corpus. Da un punto di vista qualitativo il corpus è da considerarsi adeguato agli obiettivi perseguiti. Come annunciato nel paragrafo 1.2, nelle prossime fasi di lavoro la prospettiva verrà ulteriormente ampliata sul versante dello spazio (a tutta l'Italia) e su quello del tempo (all'arte e alla cultura di altre epoche); ci si concentrerà, anche in futuro, sulla qualità stilistica e sul valore storico-culturale dei testi.

Va inoltre aggiunto che, soprattutto in virtù della propria costituzione testuale, il corpus rappresenta una miniera linguistica inesauribile: le prime ricerche effettuate con approcci differenti hanno dimostrato di condurre all'acquisizione di conoscenze fruttuose su numerosi fenomeni linguistici (es. lessicali, sintattici, testuali, traduttivi, cfr. paragrafo 2).<sup>9</sup>

<sup>8</sup> Con «stile medio» Ludwig Reiners intende il tipico stile enciclopedico che egli critica come astratto e «privo di visione». Agli antipodi troviamo la grande prosa. Secondo l'autore il «buon scrittore di prosa» porta a compimento un lavoro significativo che richiede molta forza d'immaginazione e conoscenza del mondo per riprodurre complessi processi mentali e renderli vivi (Reiners 1991: 219).

<sup>9</sup> Vanno qui menzionate anche numerose tesi triennali redatte a Firenze e a Siena e incentrate sul testo di Vasari o su altri testi del corpus LBC.

## 2. Possibilità applicative del corpus

La piattaforma LBC (<<http://corpora.lessicobeniculturali.net/de/>>) mette a disposizione i vari corpora del progetto (tedesco, inglese, francese, italiano, russo, spagnolo) permettendo la loro consultazione tramite *NoSketchEngine* (versione ridotta di *SketchEngine*<sup>10</sup>). Gli utenti possono analizzare il corpus<sup>11</sup> fruendo di diverse funzionalità, come *Wordlist* con la quale si possono estrarre liste di parole sulla base della frequenza assoluta (parole, lemmi, tags ecc.), *Concordance* con la quale si possono analizzare singole/i parole/lemmi nel loro contesto (*Keywords in Context*) e *Collocations* con la quale si possono identificare possibili collocazioni di un lessema (cfr. Kilgarriff *et al.* 2004; Kilgarriff *et al.* 2014). Si possono poi valutare i risultati ottenuti confrontandoli fra loro in ottica intralinguistica o interlinguistica.

Dato il grande potenziale scientifico e pedagogico dei corpora (cfr. Flinz 2020; Flinz, Katelhön 2019), entrambe le prospettive trovano possibilità applicative in diversi ambiti scientifici e didattici: per la prassi traduttiva dei traduttori professionisti, per la formazione dei futuri traduttori e anche per la didattica del tedesco come lingua straniera (*DaF – Deutsch als Fremdsprache*) in generale<sup>12</sup>. L'uso dei corpora LBC si presta inoltre non solo alla ricerca e all'insegnamento del lessico specialistico ma anche all'indagine di fenomeni grammaticali. Illustreremo in seguito alcuni di questi usi.

### 2.1. Analisi intralinguistiche

Il corpus tedesco risulta composto da 1.183.484 token (cfr. Figura 8). Trattandosi tuttavia di un *corpus monitor* (Lemnitzer, Zinsmeister 2015: 140), è soggetto a integrazioni e ampliamenti costanti.

Corpus LBC Tedesco ©

Corpus LBC Tedesco

Counts		General Info		Lexicon sizes	
Tokens	1,183,484	Corpus description	<a href="#">Document</a>	word	66,972
Words	1,018,253	Language	German	tag	54
Paragraphs	10,935	Encoding	UTF-8	lemma	47,026
Documents	763	Compiled	11/12/2019 11:01:29	lc ©	62,797
		Tagset	<a href="#">Description</a>	lemma_lc ©	45,498

Structures and attributes

doc 763 ▾

Figura 8. Informazioni sul corpus LBC tedesco.

<sup>10</sup> Ai tempi iniziali del progetto LBC, *SketchEngine* (<<https://www.sketchengine.eu/>>) era a pagamento, pertanto l'Unità di Ricerca ha optato per la versione ridotta *open source* *NoSketchEngine*. Successivamente grazie ad un progetto europeo *SketchEngine* è diventato gratuito per la comunità scientifica e accademica e rimarrà tale fino al 2022.

<sup>11</sup> I corpora vengono lemmatizzati e annotati a livello morfo-sintattico in maniera automatica.

<sup>12</sup> I corpora LBC possono venire impiegati anche con altre finalità, che si estendono per esempio agli studi lessicografici e all'analisi del discorso.

Una possibilità di impiego consiste nel generare liste di frequenza per lemma (Figura 9) che danno informazioni sulla frequenza dei lemmi registrati nel corpus.

**Word list**

Corpus: Corpus LBC Tedesco  
 Total number of Items: 10,324  
 Total frequency: 963,888

Page 1 (Go) Next >

lemma	frequency
die	100,366
und	37,338
sein	28,572
in	24,357
eine	22,406
er	21,792
zu	17,789
ich	16,055
von	14,196
mit	10,164
haben	10,113
sich	8,468
sie	8,343
werden	7,985
dies	7,617
nicht	6,994
auf	6,964
so	6,816
als	6,620
dass	6,551
an	6,309
es	6,256
man	5,800
wie	5,308
welche	5,231
ihr	5,181
nach	4,915
alle	4,079
auch	3,840
aus	3,719
andere	3,558
aber	3,549
für	3,524
groß	3,321
noch	3,290
bei	3,251
sehen	3,131
mein	3,058
können	3,050
um	2,855
viele	2,852
durch	2,830
wir	2,746
sehr	2,731
sagen	2,617
nur	2,550
da	2,462

Figura 9. Screenshot della lista dei lemmi estratta.

Da qui si possono per esempio ricavare i sostantivi, i verbi, gli aggettivi più frequenti:

- Sostantivi: *Werk* (2.183), *Zeit* (1.864), *Bild* (1.617), *Florenz* (1.466), *Leben* (1.414), *Kunst* (1.369), *Hand* (1.346), *Kirche* (1.307), *Rom* (1.150), *Haus* (1.145), *Papst* (1.131), *Herr* (1.120) ecc.
- Verbi: *sein* (28.572)<sup>13</sup>, *haben* (10.113), *werden* (7.985), *sehen* (3.131), *können* (3.050), *sagen* (2.617), *lassen* (2.273), *wollen* (2.205), *machen* (1.859), *geben*

<sup>13</sup> Abbiamo notato che per *sein* si trovano sia le occorrenze dell'ausiliare *sein* sia quelle dell'aggettivo possessivo *sein*, il che può essere dovuto a errori nel *tagging* automatico. Questo pro-

- (1.808), *kommen* (1.800), *sollen* (1.714), *malen* (1.677), *gehen* (1.677), *finden* (1.373), *stehen* (1.307), *arbeiten* (1.274), *müssen* (1.260) ecc.
- Aggettivi: *andere*<sup>14</sup> (3.558), *groß* (3.321), *viele* (2.852), *schön* (2.424), *gut* (2.213), *alt* (1.234) ecc.

Restringendo il campo di osservazione per esempio ai soli sostantivi, si nota subito che a ricorrere più di frequente sono termini tecnici dell'arte come *Werk*, *Bild*, *Kunst* e *Kirche*, e anche nomi di città come *Florenz* e *Rom*. Fra i verbi, come prevedibile, ricorrono più frequentemente gli ausiliari e i modali, ma anche verbi tipici del linguaggio tecnico, come per esempio *malen*. Questo uso può incentivare lo sviluppo di una sensibilità per le collocazioni<sup>15</sup>, aspetto di una lingua straniera assai difficile e quindi soggetto a errori (su questo punto cfr. anche il paragrafo 2.2).

Si possono anche effettuare ricerche sull'uso di una parola nel linguaggio tecnico e nella lingua standard, come pure sul suo sviluppo diacronico, specificando nel campo di ricerca rispettivamente tipologia testuale, autore ed epoca storica (Figura 10).

Oltre alle liste di parole, nel corpus si possono ricercare anche parole isolate. Poiché *Kirche* è una parola con un'alta frequenza, è interessante riflettere sui suoi composti. Con l'opzione di ricerca *\*kirche* possono venir estrapolati tutti i composti della parola:

*Peterskirche* (51), *Hauptkirche* (18), *Servitenkirche* (14), *Domkirche* (9), *Paulskirche* (7), *Taufkirche* (5), *Andreaskirche* (3), *Weltkirche* (3), *Bettelordenskirche* (2), *Bischofskirche* (2), *Familienkirche* (2), *Franziskanerkirche* (2), *Kathedralkirche* (2), *Klosterkirche* (2), *Porkirche* (2), *Vorgängerkirche* (2), *Augustinerkirche* (1), *Barfüßerkirche* (1), *Benediktinerkirche* (1), *Dorfkirche* (1), *Franziskaner-Kirche* (1), *Gesetzeskirche* (1), *Gewölbekirche* (1), *Hauptporkirche* (1), *Hauskirche* (1), *Kapuzinerkirche* (1), *Karmeliterkirche* (1), *Kartheuserkirche* (1), *Kollegialkirche* (1), *Kuppelkirche* (1), *Madonnenkirche* (1), *Marienkirche* (1), *Markuskirche* (1), *Ordenskirche* (1), *Pfarrkirche* (1), *Unterkirche* (1).

A lezione è possibile riflettere sia sul contenuto (il significato dei composti) sia sulla forma dei composti (grafia con o senza elementi di raccordo), integrando le ricerche eventualmente con dati estratti da altri corpora online. Ciò è possibile anche con apprendenti di livello linguistico diverso.

blema, tuttavia, può anche rivelarsi una risorsa a lezione, laddove si solleciti la sensibilità degli apprendenti verso la polifunzionalità di alcuni lessemi.

<sup>14</sup> La lemmatizzazione automatica di alcune parole (*anders*, *viel* ecc.) avviene erroneamente con il femminile.

<sup>15</sup> Il termine «collocazione» trova varie definizioni in letteratura: una empirica (fra gli altri, Firth 1957; Evert 2009; Belica, Perkuhn 2015), il cui criterio principale è la frequenza, e una teorica (fra gli altri, Hausmann 1984, Burger 2015), che definisce le collocazioni come unità lessicali in relazione sintattica (Hausmann 1984) o come unità lessicali con un significato unitario. La tesi sostenuta in questo contributo è che nel *DaF* entrambe le definizioni possano trovare applicazione. La scelta verrà determinata dagli obiettivi dell'unità di apprendimento.

The screenshot shows a search interface titled "Text types". At the top, there is a "Subcorpus" dropdown menu set to "None (whole corpus)" and a link "info create new". Below this are several filter sections:

- LINGUA ORIGINALE**: Includes checkboxes for "Francese (Francia)", "Inglese (Regno Unito)", "Italiano", and "Tedesco (Germania)", with a "Select All" button.
- LINGUA DI TRADUZIONE**: Includes a checkbox for "Tedesco (Germania)" and a "Select All" button.
- CATEGORIA E SOTTOCATEGORIA**: Includes checkboxes for "Divulgativo", "Letterario", and "Tecnico", with a "Select All" button.
- AUTORE**: A text input field.
- TITOLO**: A text input field.
- FRAMMENTO**: A text input field.
- ANNO DI REDAZIONE**: A text input field.

Figura10. Screenshot del campo di ricerca.

Le concordanze<sup>16</sup> della parola *Kirche* possono inoltre fornire anche informazioni sul suo uso, come per esempio:

- in combinazione con aggettivi in funzione attributiva: *achteckige Kirche, alte Kirche, bischöfliche Kirche, christliche Kirche, große Kirche, kleine Kirche, neue Kirche, römische Kirche, schöne Kirche, untere Kirche*;
- in combinazione con altri sostantivi: *Bild der Kirche, Bronzetür der Kirche, Chor der Kirche, Dach der Kirche, Eingang der Kirche, Fassade der Kirche, Grundriss der Kirche, Hauptaltar der Kirche, Hauptkapelle der Kirche, Haupttür der Kirche, Kapelle der Kirche, Kreuz der Kirche, Mittelschiff der Kirche, Querschiff der Kirche, Schiff der Kirche, Teil der Kirche, Tür der Kirche, Vorderwand der Kirche, Wand der Kirche, ecc.*;
- in combinazione con preposizioni: *in der Kirche, inmitten der Kirche, innerhalb der Kirche, nach der Kirche, neben der Kirche, unterhalb der Kirche, von der Kirche, vor der Kirche, für die Kirche, in die Kirche, ecc.*;
- in combinazione con verbi (filtro a destra): *eine Kirche bauen, eine Kirche malen, eine Kirche vollenden, eine Kirche verfertigen, ecc.*;

<sup>16</sup> I KWICs possono venir ordinati tramite l'opzione *sort Left, sort Right* ecc.. Questa opzione offre la possibilità di richiamare l'attenzione su determinate caratteristiche della lingua analizzata.

- come nome proprio: *Kirche del Carmine, Kirche della Pace, Kirche der Annunziata, Kirche der Madonna, Kirche des Heiligen Andreas, Kirche des S. Agostino, Kirche Ognissanti, Kirche San Giovanni, Kirche St. Andrea, Kirche von S. Agostino, ecc.*

In questo modo è possibile favorire negli apprendenti del tedesco sia una maggiore consapevolezza delle collocazioni di una determinata parola sia una maggior sensibilità per determinati fenomeni grammaticali.

## 2.2 Analisi interlinguistiche

I corpora compresi nella piattaforma LBC hanno tutti più di un milione di token (Tabella 2):

Tabella 2. Numero dei token per lingua.

	Tedesco	Inglese	Italiano	Francese	Russo	Spagnolo
Token	1.183.484	1.784.990	1.647.595	3.818.747	1.647.595	1.162.849

Per ogni corpus sono presenti informazioni diverse, così da poter ricavare, fra l'altro, il numero dei rispettivi testi per anno di pubblicazione, autore, genere testuale (Figura 11):

<u>doc.printingYear</u>	<u>document frequency</u>
1966	162
1973	125
1970	66
1971	10
1550	2
1910	1

Figura 11. Distribuzione dei documenti nel corpus italiano in base all'anno di pubblicazione.

Il sistema permette anche di costruire subcorpora sulla base della tipologia testuale, così da poter anche effettuare studi interlinguistici orientati su questo aspetto.

Possono venir utilizzate per confronti interlinguistici anche le liste di frequenza estratte dai corpora, che possono essere utili per l'identificazione di possibili equivalenti (cfr. Tabella 3):

Dalla Tabella 3 è possibile infatti identificare i seguenti equivalenti: *cosa/Ding, opera/Werk, mano/Hand, tempo/Zeit, figura/Figur, anno/Jahr* ecc. Il significato dei lessemi rimanenti, per esempio ‚Papst‘, può poi venire eventualmente desunto tramite contesto (Figura 12):

La parola cercata *Papst* può tuttavia venir aperta anche in un'altra finestra al di fuori della rappresentazione KWIC, per esempio come testo (Figura 13):

Tabella 3. Liste di parole basate sulla frequenza (frequenza assoluta): sostantivi nel corpus tedesco e in quello italiano.

Sostantivi (D)	<->	Sostantivi (I)
Werk (2.183)		cosa (4671)
Zeit (1.864)		opera (3.615)
Bild (1.617)		mano (2.143)
Florenz (1.466)		tempo (2.093)
Leben (1.414)		figura (1.731)
Kunst (1.369)		modo (1.693)
Hand (1.346)		anno (1.693)
Kirche (1.307)		uomo (1.510)
Jahr (1.276)		parte (1.510)
Rom (1.150)		città (1.462)
Haus (1.145)		disegno (1.414)
Papst (1.131)		luogo (1.377)
Herr (1.120)		parere (1.319)
Tag (1.118)		Roma (1.309)
Stadt (1.096)		storia (1.304)
Meister (972)		tavola (1.284)
Figur (945)		chiesa (1.278)
Mann (913)		messer (1.216)
Künstler (903)		maniera (1.206)
Ding (899)		casa (1.194)

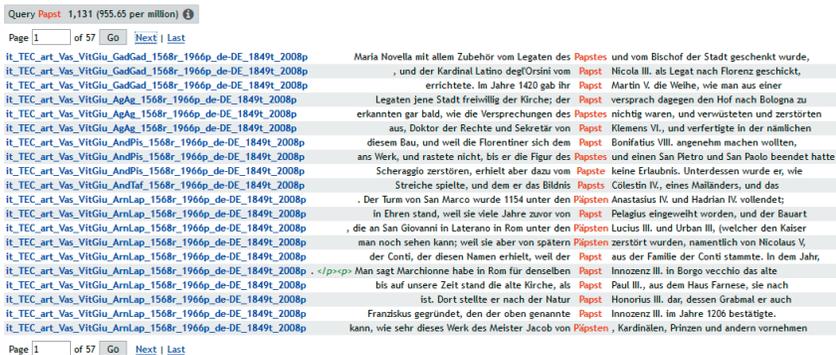


Figura 12. KWICs della parola cercata, 'Papst'.

Un altro uso particolarmente rilevante per la didattica *DaF* è il confronto interlinguistico delle collocazioni. Prendendo sempre a esempio *Kirche* e in questo caso la traduzione italiana (*chiesa*)<sup>17</sup>, possiamo riflettere prima sulle possibili collocazioni delle due parole (per esempio partendo dal tipo AGGETTIVO + SOSTANTIVO) e poi individuare possibili relazioni di equivalenza (Figura 14):

- *Kirche*: *genannte Kirche, bischöfliche Kirche, alte Kirche, christliche Kirche, untere Kirche, römische Kirche* ecc.
- *Chiesa*: *detta Chiesa, nuova chiesa*<sup>18</sup>, *chiesa principale, chiesa vecchia, chiesa romana, chiesa piccola* ecc.

< [previous](#) vergrößern. Sie hatten eine bedeutende Summe Geldes zusammen gebracht, und viele in der Stadt versprochen allen Beistand zu leisten. Deshalb ward der Bau der neuen Kirche am Sankt-Lucas-Tag 1278 begonnen, und der Cardinal Latino degli Orsini vom **Papst** Nicola III. als Legat nach Florenz geschickt, legte feierlich den Grundstein. Baumeister jener Kirche waren: der Bruder Giovanni, ein Florentiner, und der Bruder Ristoro da Campi, beide Bekenner eines Ordens, welche auch [next](#) >

Figura 13. Frammento di testo per la parola cercata ‚Papst‘.

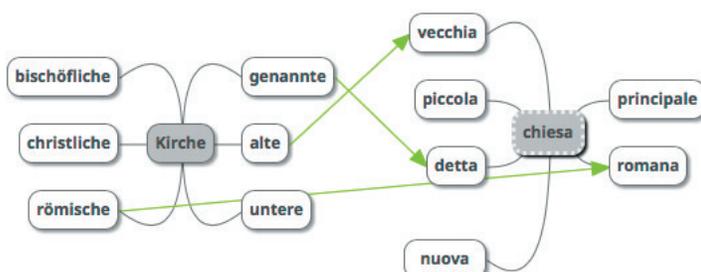


Figura 14. Rapporti di equivalenza fra le collocazioni aggettivali (*Kirche/chiesa*).

I singoli raggruppamenti possono venir considerati anche nei loro contesti e confrontati fra loro (Figura 15):

Query **chiesa** 1,429 • Positive fitter (excluding KWIC) **vecchia** 6 (5.10 per million) ⓘ

doc#5 - In Florenza, che dipinse nella **vecchia chiesa** del Tempio, che era dove è oggi la cittadella

doc#11 Vigna, dove stanno Frati del Zoccolo, nella **chiesa vecchia** era in un quadro un Cristo morto, tanto

doc#187 quali andarono per terra quando fu rovinata la **chiesa vecchia** di San Piero per rifar la nuova. <-p>> Fu

doc#202 - Il quale, andato a Perugia, fece nella **chiesa vecchio** di S. Domenico dei Frati Predicatori

doc#228 ; il quale per lo interesse che egli aveva nella **chiesa vecchio** -che la capella e l' altare maggiore era

doc#332 deputati, ordinarono che si facesse in mezzo la **chiesa vecchia** una cappella da Antonio da Sangallo,

Query **Kirche** 1,307 • Positive fitter (excluding KWIC) **alten** 13 (10.98 per million) ⓘ

doc#8 dorthin kommen, und er arbeitete in der **alten Kirche** San Domenico, welche den Prädikanten-Mönchen

doc#11 Katharina darin darstellte. In der **alten Kirche** S. Domenico malte er auf einer Wand, wiederum

doc#12 versetzte er in Florenza eine Kapelle in der **alten Kirche** S. Spirito derselben Stadt, welche beim Brand

doc#37 und endlich die Einweihung jener sehr **alten Kirche** , welche von Papst Paschalis II. vollzogen

doc#57 - sagt man, habe er sehr Jung noch in der **alten Kirche** , die an dem Platz stand, wo nunmehr die alte

doc#65 Frati del Zoccolo weihen, sah man in der **alten Kirche** einen toten Christus auf Herolden von ihm

doc#114 fügte hinzu: man müsse es nicht nach der **alten Kirche** von St. Peter bringen, wo kein Platz mehr dafür

doc#137 von Antonio da Sangallo in der Mitte der **alten Kirche** eine Kapelle zum Teil aus Künemern (spiegel)

doc#191 und zu Ravenna der Marmor aus drei **alten Kirchen** (S. Severo, S. Apollinare in Classe und Galla

doc#193 blieb beharrlich, warf die Hälfte der **alten Kirche** nieder und legte die Fundamente der neuen (18.

doc#336 diesen köhnen Gedanken vorliegt. Von der **alten Kirche** , welche der Kaiser Konstantin auf diesem Fleck

doc#336 Hinter grunde oder der Tribune, wo aus der **alten Kirche** Konstantin, nach des größte Heiligtum,

doc#750 II. machte reine Bahn durch Abbruch der **alten Kirche** - Mit den größten Meilen sollte der schönste

Figura 15. Collocazioni: *chiesa vecchia* e *alte Kirche*.

<sup>17</sup> È stato scelto a titolo esemplificativo il parametro *range* (-1 to 1). L'opzione (-5 to 5) è imposta per default.

<sup>18</sup> Sono state elencate solo le combinazioni più frequenti. Combinazioni come *chiesa nuova* e *principale chiesa* compaiono anch'esse nel corpus, presentano tuttavia una bassa frequenza.

Come mostrano gli esempi sopra riportati, con l'aiuto degli strumenti messi a disposizione dalla piattaforma LBC gli apprendenti possono fare ricerche all'interno dei vari corpora sviluppando la *corpus literacy* (Mukherjee 2002: 179-180) necessaria per diventare sempre più autonomi.

### 3. Conclusioni e sviluppi futuri

In questo contributo si sono descritte le caratteristiche più importanti del corpus LBC tedesco unitamente ad alcune delle sue possibilità d'impiego. Tale descrizione, illustrando il corpus tedesco con determinate proprietà in un dato momento, vuole essere l'istantanea di un *work in progress*. Il corpus necessita infatti di ampliamenti testuali continui, qualitativamente elevati, che costituiscano un valido modello linguistico per tutti i gruppi di utenti a cui esso si rivolge: tutte le tipologie testuali presenti verranno arricchite di nuovi esemplari testuali, cui andranno ad aggiungersi altre tipologie testuali; si dovrà spostare l'attenzione su altre epoche storiche di riferimento (non solo i Beni culturali del Rinascimento, ma anche movimenti artistici successivi) e dal punto di vista diatopico si dovranno prendere in considerazione altri territori italiani oltre alla Toscana. Anche il lavoro scientifico e lessicografico è *in progress*, tanto che vecchie e nuove fasi del processo lessicografico vengono a sovrapporsi, come è tipico delle risorse online (Klosa 2016: 29).

Nonostante il carattere *in fieri* del lavoro i dati fin qui esposti dimostrano che già allo stato attuale il corpus tedesco – come gli altri corpora LBC – può venir utilizzato per numerosi scopi di ricerca e anche per attività didattiche. Come illustrato nel paragrafo 1, sono stati impiegati generi e tipologie testuali molto diversi in ottica sia sincronica che diacronica. Questa configurazione offre agli utenti – che si tratti di esperti o di dilettanti – una vasta gamma di applicazioni (cfr. paragrafo 3), come pure riflessioni sulle possibilità traduttive. In sintesi, il corpus tedesco si presta all'analisi di svariati fenomeni lessicali, morfologici, sintattici, morfosintattici e stilistici da prospettive diverse come per esempio:

- in una prospettiva orientata alla tipologia testuale;
- in prospettiva intralinguistica e interlinguistica;
- in prospettiva traduttiva (cfr. fra l'altro il focus sulla preponderanza di italianismi nelle traduzioni raccolte, Farina, Billero 2018);
- in prospettiva contrastiva (italiano-tedesco/tedesco-italiano);
- in prospettiva diacronica e sincronica.

Speriamo in generale con il nostro corpus di aver messo a disposizione di tutti gli interessati che fanno ricerca e lavorano nel campo dei Beni culturali un valido strumento dal quale poter ricavare utili indicazioni sulle proprietà linguistiche delle tipologie testuali inerenti all'arte e alla cultura. A questo proposito il corpus tedesco di LBC è già stato impiegato in concreto come fonte primaria per l'elaborazione del lemmario provvisorio del progettato dizionario LBC (cfr. Farina, Flinz 2020) e per la selezione delle concordanze (cfr. Buffagni, Flinz,

Ballestracci 2020) che fotografano l'attuale stato dell'arte dell'unità di ricerca LBC per il tedesco.

## Bibliografia

- Alberti L.B. 1849, *Opere Volgari*, Tipografia Galilei, Firenze.  
 — 1877, *Kleinere kunsttheoretische Schriften*, Braumüller, Wien.
- Artusi P. 2000, *Von der Wissenschaft des Kochens und der Kunst des Genießens*, Hahn Verlag, Berlin.
- Ballestracci S., Buffagni C. 2016, *La traduzione dei 'gerundi italiani' in tedesco: un'analisi semantica esemplare*, in Saraçgil A., Vezzosi L. (eds.), *Lingue, letterature e culture migranti*, FUP, Firenze: 11-35.
- Ballestracci S., Buffagni C., Flinz C. 2020, *Das deutsche LBC-Korpus: Zusammenstellung und Anwendung*, in Id., *Das deutsche LBC-Korpus*, FUP, Firenze. In pubblicazione, <<http://corpora.lessicobeniculturali.net/de/>>.
- Belica C., Perkuhn R. 2015, *Feste Wortgruppen/Phraseologie I: Kollokationen und syntagmatische Muster*, in Haß U., Storjohann P. (eds.), *Handbuch Wort und Wortschatz*, de Gruyter, Berlin-Boston: 201-225.
- Best O.F. 1994, *Handbuch literarischer Fachbegriffe. Definitionen und Beispiele*, Fischer, Frankfurt.
- Billero R., Nicolás Martínez M.C. 2017, *Nuove risorse per la ricerca del lessico del patrimonio culturale: corpora multilingue LBC*, «CHIMERA Romance Corpora and Linguistic Studies», 4(2): 203-216.
- Billero R. 2020, *Cultural Heritage Lexicon: A Case Study*, in Pano Alamán A., Zotti V. (eds.), *The language of art and cultural heritage: a plurilingual and digital perspective*, Cambridge Scholars Publishing, Newcastle upon Tyne: 86-103.
- Brandi K. 1927, *Die Renaissance in Florenz und Rom. Acht Vorträge von Karl Brandi*, Springer Fachmedien, Wiesbaden.
- Buffagni C., Flinz C., Ballestracci S. 2020, *Das deutsche LBC-Korpus: provisorische Stichwortliste und Konkordanzen*, in Flinz C., Buffagni C., Ballestracci S., Billero R., Farina A., Nicolás Martínez M.C., *Deutsche Lexik der Kunst auf der Basis des Korpus LBC*, FUP, Firenze. In pubblicazione, <<http://lexicon.lessicobeniculturali.net/de/>>.
- Burger H. 2015, *Phraseologie. Eine Einführung am Beispiel des Deutschen*, Erich Schmidt Verlag, Berlin.
- Burckhardt J. 1891, *Die Geschichte der Renaissance in Italien*, Ebner & Seubert, Stuttgart.
- Cellini B. 2016, *Das Leben des Benvenuto Cellini, florentinischen Goldschmieds und Bildhauer. Von ihm selbst geschrieben*, Karl-Maria Guth, Berlin.
- Ciccarelli Roming C., Jepson T., Fisher T. 2016, *Florenz. Perfekte Tage in der Toskana – Metropole*, Taschenbuch Verlag, München.
- da Vinci L. 1651, *Trattato della pittura*, Jacques Langlois, Parigi.  
 — 1909, *Der Traktat von der Malerei*, Eugen Diederichs, Jena.  
*Deutsches Textarchiv*, <<http://www.deutschestextarchiv.de/>>, (10.2020).
- Evert S. 2009, *Corpora and Collocations*, in Lüdeling A., Kytö A. (eds.), *Corpus Linguistics*, de Gruyter, Berlin-New York: 1212-1248.
- Farina A. 2016, *Le portail lexicographique du Lessico plurilingue dei Beni Culturali, outil pour le professionnel, instrument de divulgation du savoir patrimonial et atelier didactique*, «Publif@rum», 24, <[http://www.farum.it/publifarum/ezine\\_articles.php?art\\_id=335](http://www.farum.it/publifarum/ezine_articles.php?art_id=335)>.

- Farina A., Billero R. 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 Lezzi D.F., Celardo L., Misuraca M. (eds.), *JADT 2018 – International Conference on Statistical Analysis of Textual Data, Roma, 12-15 giugno 2018*, UniversItalia, Roma: 108-116.
- Farina A., Flinz C. 2020, *LBC-Dictionary: a Multilingual Cultural Heritage Dictionary. Data collection and data preparation*, in Gavriilidou Z., Mitsiaki M., Fliatouras A. (eds.), *Proceedings of XIX EURALEX Congress: Lexicography for Inclusion – Volume 1*, Ljubljana University Press, Ljubljana: 371-379.
- Farina A., Garzaniti M. 2013, *Un portale per la comunicazione e la divulgazione del patrimonio culturale: progettare un lessico multilingue dei beni culturali on-line*, in Filipovic A., Troiano W. (eds.), *Strategie e Programmazione della Conservazione e Trasmissibilità del Patrimonio Culturale*, Edizioni scientifiche Fidei Signa, Roma: 500-509.
- Fiorillo J. D. 1798, *Die Geschichte der römischen und florentinischen Schule*, Rosenbusch, Göttingen.
- Firth J.R. 1957, *Modes of Meaning*, in *Papers in Linguistics 1934-1951*, Oxford University Press, London: 190-215.
- Flinz C. 2020, *Vergleichbare Spezialkorpora für den Tourismus: eine Chance für den Fachsprachenunterricht*, in Hepp M., Salzmann K. (eds.), *Sprachvergleich in der mehrsprachig orientierten DaF-Didaktik: Theorie und Praxis*, Istituto Italiano di Studi Germanici, Roma: 133-151.
- Flinz C., Buffagni C., Ballestracci S., Billero R., Farina A., Nicolás Martínez M.C. 2020, *Deutsche Lexik der Kunst auf der Basis des Korpus LBC*, FUP, Firenze. In pubblicazione, <<http://lexicon.lessicobeniculturali.net/de>>.
- Flinz C., Katelhön P. 2019, *Korpora nella didattica del tedesco come lingua straniera. Proposte per l'insegnamento del linguaggio specialistico del turismo*, «EL.LE», 8.2: 319-344.
- Gass K. E. 1961, *Pisaner Tagebuch. Aufzeichnungen / Briefe*, Lambert Schneider Verlag, Heidelberg.
- Goethe J.W. 1997, *Italienische Reise*, Deutscher Taschenbuch Verlag, München.
- Gombrich E.H. 2000, *La storia dell'arte raccontata da Ernst H. Gombrich*, Leonardo Arte, Milano.
- Gutenberg-Projekt*, <<https://www.projekt-gutenberg.org/index.html>>, (10.2020).
- Hausmann F.J. 1984, *Wortschatzlernen ist Kollokationslernen. Zum Lehren und Lernen französischer Wortverbindungen*, «Praxis des neusprachlichen Unterrichts», 31: 395-406.
- Heine H. 1969, *Florentinische Nächte*, Artemis & Winkler Verlag, München.
- Institut für Deutsche Sprache 2017, *Deutsches Referenzkorpus / Archiv der Korpora geschriebener Gegenwartssprache 2017-I (Release vom 08.03.2017)*, Institut für Deutsche Sprache, Mannheim. PID: 10932/00-0373-23CD-C58F-FF01-3.
- Kilgariff A., Rychlý P., Smrz P., Tugwell D. 2004, *The Sketch Engine*, in Geoffrey W., Vessier S. (eds.), *Proceedings of the Eleventh EURALEX International Congress, Lorient, France July 6-10, 2004*, Faculté des Lettres et des Sciences Humaines, Université de Bretagne, Lorient: 99-104.
- Kilgariff A., Baisa V., Bušta J., Jakubiček M., Kovář V., Michelfeit J., Rychlý P., Suchomel V. 2014, *The Sketch Engine: ten years on*, «Lexicography», 1: 7-36.
- Klosa A. 2016, *Der lexikographische Prozess im Projekt elexiko*, in Hildenbrandt V., Klosa A. (eds.), *OPAL Online publizierte Arbeiten zur Linguistik, 1/2016, Lexikographische Prozesse bei Internetwörterbüchern*: 29-38.

- Kurz I. 1905, *Stadt des Lebens*, Hermann Seemann Nachfolger, Leipzig.  
 — 1984, *Florentiner Novellen*, Phaidon, Essen.  
 LBC-Korpusplattform, <<http://corpora.lessicobeniculturali.net/de/>>, (10.2020).  
 Lemnitzer L., Zinsmeister H. 2015, *Korpuslinguistik. Eine Einführung*, Narr, Tübingen.  
 Lessico dei Beni Culturali, <<http://corpora.lessicobeniculturali.net/>>, (10.2020).  
 Lewald F. 1992, *Bilderbuch*, Urike Helmer Verlag, Frankfurt am Main.  
 Machiavelli N. 1934, *Geschichte von Florenz*, Phaidon, Essen.  
 — 1971, *Istorie Fiorentine*, Sansoni, Firenze.  
 Moritz K.P. 2015, *Reisen eines Deutschen in Italien in den Jahren 1786 bis 1788*, e-artnow, o.O.  
 Mukherjee J. 2002, *Korpuslinguistik und Englischunterricht. Eine Einführung*, Peter Lang, Berlin u.a.  
 Reiners L. 1991, *Stilkunst. Ein Lehrbuch deutscher Prosa*, Beck, München.  
 Rilke R.M. 1984, *Das Florenzer Tagebuch*, Suhrkamp, Frankfurt am Main.  
 Ruskin J. 1897, *Wege zur Kunst. Eine Gedankenlese aus den Werken von John Ruskin*, Heitz und Mündel, Straßburg.  
 Stendhal 1922, *Reise in Italien (Rome, Naples et Florence)*, Propyläen Verlag, Berlin.  
 Vasari G. 1550, *Le vite de' più eccellenti architetti, pittori, et scultori italiani: da Cimabue in sino a tempi nostri*, Torrentino, Firenze. <<http://archiviovasari.beniculturali.it/index.php/1550-la-prima-edizione-de-le-vite/>>.  
 — 1966, *Le Vite de' più eccellenti Pittori, Scultori et Architettori*, S.P.E.S/Sansoni, Firenze.  
 — 2008, *Leben der ausgezeichnetsten Maler, Bildhauer und Baumeister von Cimabue bis zum Jahre 1567*, Directmedia Publishing, Berlin.  
 Wackenroder W.H. 1991, *Herzensergiessungen eines kunstliebenden Klosterbruders*, in Vietta S., Littlejohns R. (eds.), *Sämtliche Werke und Briefe, Historisch-kritische Ausgabe*, Winter Universitätsverlag, Heidelberg.