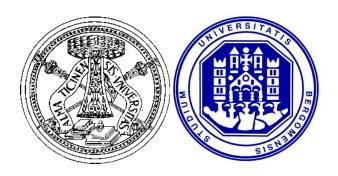
# Università degli studi di Pavia & Università degli studi di Bergamo Dottorato di ricerca in Scienze linguistiche XXXI ciclo



## COMPOSIZIONE ITALIANA IN DIACRONIA

Le parole composte dell'italiano nel quadro della Morfologia delle Costruzioni

Tesi di dottorato di Maria Silvia Micheli

> Tutor: Elisabetta Ježek

Anno accademico 2017/2018

# INDICE

Introduzi	one	7
1 LACO	MPOSIZIONE	9
	nozione di composto	
	elimitazione	
1.3 Te	ntativi di classificazione	15
	omposti dell'italiano	
	ssegna dei principali studi sull'italiano e domande aperte	
2 IL QU	ADRO TEORICO: LA MORFOLOGIA DELLE COSTRUZIONI	29
	ılla Grammatica delle Costruzioni alla Morfologia delle Costruzioni	
	Costruzione: definizione, struttura, rappresentazione	
	Costruzionario	
2.4 La	formazione delle parole	44
2.4.1	Derivazione	45
2.4.2	Composizione	46
2.4.3	Tra morfologia e sintassi: le espressioni multiparola	
2.5 Il 1	mutamento costruzionale e la formazione di affissoidi	
3. METO	DOLOGIA DI INDAGINE	50
	DIt_com: Corpus Diacronico dell'Italiano per la composizione	
3.1.1	Finalità del corpus	59 EO
3.1.1 3.1.2	Architettura del corpus	
3.1.2 3.1.3	Raccolta dei testi	
3.1.4 3.1.5	Composizione del corpusConfronto con altri corpora diacronici	
3.1.5 3.1.6	· ·	
3.1.6 3.1.7	Estrazione dei composti	
	Il dataset Compos_ita	
3.2 AI 3.2.1	tre fonti	
3.2.1 3.2.2	Corpora di italiano antico e contemporaneo	
o. <b>_</b>	rametri di analisi	
3.3.1	Parametri formali	
3.3.1 3.3.2	Parametri semantici	
3.3.2 3.3.3	Parametri quantitativi	
	ruttura dell'analisi	
	ISI	
	ılla composizione latina a quella italiana	
	lliano antico	
4.2.1	La Costruzione <sub>COMP</sub> dei composti con testa a sinistra	
4.2.2	La Costruzione <sub>COMP</sub> dei composti con testa a destra	
4.2.3	La Costruzione <sub>COMP</sub> dei composti coordinativi endocentrici	
4.2.4	La Costruzione <sub>COMP</sub> dei composti esocentrici	
4.2.5	Le Costruzioni semi-specificate	
4.2.5		
4.2.5	- 2a 355t a215 no comp [saps n]n	
4.2.6	Quid manet? Un bilancio sulla composizione dal latino all'italiano antico	
	nalisi diacronica	
4.3.1	La Costruzione <sub>COMP</sub> dei composti con testa a sinistra	
4.3.2	La Costruzione <sub>COMP</sub> dei composti con testa a destra	
4.3.3	La Costruzione <sub>com</sub> dei composti coordinativi endocentrici	201

4.3.4 La Costruzione <sub>COMP</sub> dei composti esocentrici	209
4.4 Discussione	227
5. LE COSTRUZIONI VALUTATIVE [MAL(E)-X] E [MAL(O/A)-X]	231
5.1 Introduzione	
5.2 Parametri di analisi e metodologia di indagine	
5.3 La Costruzione <sub>COMP</sub> [mal(e)-x]	
5.3.1 Dal latino all'italiano	
5.3.2 Italiano antico	236
5.3.3 Analisi diacronica	242
5.3.4 Italiano contemporaneo	252
5.3.5 Discussione	
5.4 La Costruzione [mal(o/a)-x]	256
5.4.1 Italiano antico	256
5.4.2 Analisi diacronica	259
5.4.3 Italiano contemporaneo	265
5.4.4 Discussione	272
5.4 Osservazioni conclusive	274
6. CONCLUSIONI	279
6.1 La composizione dell'italiano nel quadro della Morfologia delle Costruzioni	
6.2 Un bilancio sulla composizione dell'italiano	
Bibliografia	292
APPENDICE I: DATASET COMPOS_ITA_TLIO	306
APPENDICE II: DATASET COMPOS_ITA	316

## Ringraziamenti

Questo lavoro è il frutto di un'esperienza durata tre anni, decisivi per la mia formazione e per la mia vita: alcune persone, che di seguito ringrazio, hanno contribuito a renderli tali.

Un primo ringraziamento va a Pierluigi Cuzzolin, come coordinatore del dottorato, e a Elisabetta Ježek, Nicola Grandi, Caterina Mauri e Andrea Sansò, come tutor e lettori, per avermi permesso di portare avanti questo progetto: ringrazio, in particolare, Elisabetta Ježek per aver discusso con me l'intero lavoro e avermi aiutata a migliorarlo.

Gran parte delle idee che ho maturato sulla morfologia e sulla composizione la devo al confronto con Fabio Montermini e Francesca Masini: ringrazio il primo per avermi accolta al CNRS di Tolosa e per i numerosi consigli e spunti che mi hanno aiutato a dar forma a questo lavoro; la seconda per avermi introdotta con entusiasmo e metodo al quadro costruzionista, all'interno del quale ho cercato di ritagliare uno spazio per la composizione dell'italiano.

Negli ultimi due anni è stata per me fondamentale l'esperienza all'Università di Milano – Bicocca: per il supporto costante e le piacevoli ore di lavoro passate insieme, ringrazio Davide Colussi e Paolo Zublena; devo un ringraziamento particolare a Giorgio F. Arcodia per il tempo che in questi anni mi ha dedicato, aiutandomi e consigliandomi con pazienza e disponibilità.

Ho imparato molto anche dall'esperienza di lavoro all'Università Cattolica di Milano: sono grata a Eleonora Litta e Marco Passarotti per avermi accolto al CIRCSE e avermi dato questa possibilità. A Luca Serianni e Isabella Chiari, che per primi mi hanno introdotto allo studio della Linguistica italiana e generale, devo l'esser arrivata fin qua: li ringrazio per la formazione ricevuta alla Sapienza e per il supporto mai venuto meno in questi anni.

Questo percorso non sarebbe stato lo stesso se non l'avessi condiviso con i miei colleghi, a cui va un grande ringraziamento per i momenti passati insieme. Ringrazio in particolare Silvia Ballarè e Guglielmo Inglese perché hanno condiviso e alleggerito, da amici prima che da colleghi, i non rari momenti di sconforto e noia di questo percorso. Per le tante chiacchierate, accademiche e (soprattutto) non, e il supporto costante, ringrazio Marco Budassi, Anna Feltracco, Shanshan Huang e Matteo Pellegrini.

Ad Alessio riservo il ringraziamento più personale, per la cura con cui mi è stato accanto e mi ha sostenuto e per aver condiviso con me i momenti più felici di questi anni.

Dedico questo lavoro a mio padre.

## Introduzione

Questo lavoro si colloca nell'ambito degli studi dedicati alla formazione delle parole dell'italiano: in particolare, esso si concentra sulla composizione, indagata finora soltanto in chiave sincronica e in relazione a particolari tipologie di composto. Nonostante si tratti di un fenomeno ampiamente diffuso nelle lingue, la composizione sfugge a una definizione universalmente valida, manifestandosi con caratteristiche in parte diverse da lingua a lingua: scopo di questo lavoro è tracciare un quadro sistematico della composizione dell'italiano. Analizzando in prospettiva diacronica le parole composte che l'italiano ha creato dalla fase antica a quella contemporanea, si intende delineare un profilo della composizione italiana, volto a metterne in luce le caratteristiche formali e semantiche, la produttività e i mutamenti subiti nel corso dei secoli; tale studio intende inoltre fornire gli strumenti per interpretare la composizione nella fase contemporanea.

La prospettiva che si è scelto di adottare è quella proposta dalla Morfologia delle Costruzioni (cfr. Booij 2010b), teoria che riprende e applica al piano morfologico i principi fondamentali della Grammatica delle Costruzioni (cfr. Goldberg 1995, 2013): le parole composte dell'italiano verranno quindi analizzate come Costruzioni, ossia associazioni convenzionalizzate di una forma e una funzione. Tale quadro teorico è stato finora prevalentemente adottato per descrivere la composizione di lingue germaniche (cfr. Booij 2015b, 2010c): questo lavoro intende colmare una lacuna nella letteratura costruzionista, applicando gli strumenti della Morfologia delle Costruzioni alla formazione delle parole composte di una lingua romanza come l'italiano.

Il lavoro si articola come segue. Nel primo capitolo viene introdotta e delimitata la categoria di composto, di cui vengono discussi i tentativi di classificazione e gli studi dedicati alla composizione dell'italiano. Il secondo capitolo introduce il quadro teorico entro il quale si colloca il presente lavoro: i principali assunti della Morfologia delle Costruzioni vengono presentati e discussi, con particolare riferimento alla formazione delle parole. La descrizione delle risorse utilizzate e dei parametri di analisi adottati è oggetto del terzo capitolo: in particolare, si presentano le fonti e la metodologia con cui sono stati estratti i dati. Il capitolo 4 rappresenta il nucleo centrale del lavoro, in cui vengono analizzati i composti dell'italiano in ottica costruzionista: nella prima parte del

capitolo, si focalizza l'attenzione sui composti dell'italiano antico; nella seconda, ciascuna tipologia di composto viene analizzata dal punto di vista diacronico. Il quinto capitolo è dedicato a due studi di caso in cui vengono analizzate due Costruzioni di cui si è osservata la rilevanza nel corso dell'analisi diacronica, ossia le Costruzioni peggiorative [mal(e)-x] e [mal(a/o)-x]. Infine, il capitolo 6 riepiloga e discute i principali risultati di questo lavoro, mettendone in luce i limiti e tracciando un bilancio della composizione italiana.

#### 1 LA COMPOSIZIONE

In questo capitolo si introducono le nozioni teoriche su cui si basa il presente lavoro. Nella prima parte (§ 1.1-2), si discute la nozione di parola composta e se ne delimitano i confini; nella seconda (§ 1.3), dopo aver illustrato la classificazione dei composti proposta da Bisetto & Scalise (2005), di cui verranno ripresi alcuni principi fondamentali, si fornisce una panoramica di altri tentativi di classificazione e si presenta quella adottata in questo lavoro. Il capitolo si chiude con un quadro delle tipologie di composto attestate in italiano contemporaneo (§ 1.4) e una rassegna dei principali studi dedicati ai composti dell'italiano (§ 1.5).

## 1.1 La nozione di composto

Una parola composta in italiano si può intuitivamente definire come una parola nata dall'unione di due parole; meno intuitivamente si può aggiungere che in italiano le due parti del composto sono legate da una relazione che non è esplicitata da nessuna marca formale. Nonostante si tratti di una nozione disponibile tra i parlanti a livello intuitivo, la ricerca di una definizione di composto sufficientemente stringente e soprattutto valida interlinguisticamente ha una lunga storia nella letteratura:1 la definizione standard di composto come 'parola formata dall'unione di due (o più) parole',2 infatti, porta con sé almeno due problemi: 1) non chiarisce che cosa si debba intendere per "parola", se entità lessicale autonoma, lessema, radice, etc.; 2) include nella categoria di composto qualsiasi elemento lessicale costituito da più parti, come parole sintagmatiche, espressioni idiomatiche, etc. Il primo problema mette in luce che tale definizione non ha validità interlinguistica: lingue come l'inglese, l'italiano o il latino, infatti, si differenziano rispetto al tipo di unità che prendono parte al processo di composizione. In particolare, in inglese le parole sono generalmente monomorfemiche: un composto come blackbird 'merlo' è infatti costituito da due entità lessicali autonome, black<sub>A</sub> e bird<sub>N</sub>, entrambe monomorfemiche. In italiano, lingua in cui le parole hanno generalmente almeno due morfemi, i costituenti dei composti sono parole autonome e conservano il proprio

<sup>&</sup>lt;sup>1</sup> Per una sintesi delle definizioni di composto adottate si veda Bauer (2017: 3-28).

<sup>&</sup>lt;sup>2</sup> Cfr. Fabb (1999: 66): «A compound is a word which consists of two or more words».

morfema flessivo, es. *cap-o-stazion-e* (pl. *cap-i-stazion-e*).<sup>3</sup> In latino, le parole composte sono costituite da entità lessicali non autonome, generalmente legate da un elemento di raccordo (*linking element*), come nel caso di *pisciceps* 'pescatore' (*pisc-* 'pesce' + -*i-* + - *ceps* da *capio* 'raccoglitore'). Inoltre, anche all'interno di una stessa lingua, possono presentarsi situazioni diverse: è il caso dell'italiano, in cui nei composti V-N il primo elemento è interpretato da alcuni studiosi come un tema (o una radice morfomica, *morphomic stem*, nel senso di Aronoff 1994).<sup>4</sup>

Una diffusa e largamente accettata definizione di composto, che cerca di rendere conto della diversità con cui la composizione si presenta nelle lingue del mondo, è quella di Bauer (2003), secondo cui la composizione consiste nella formazione di un nuovo lessema attraverso l'unione di due o più lessemi. Il termine *lessema* è generalmente usato in lessicografia e lessicologia per indicare «l'unità del lessico assunta come forma base alla quale sono ricondotte le forme flesse» (Jezek 2011: 34). Analogamente, in morfologia, il lessema rappresenta un'entità astratta a cui vengono ricondotte le concrete realizzazioni (ad es. le forme flesse) con cui una forma può occorrere nell'uso.<sup>5</sup> L'adozione della nozione di lessema per riferirsi alle unità di input della composizione ha il merito di essere allo stesso tempo abbastanza stringente da escludere gli affissi e sufficientemente larga da includere radici e parole libere di uno o più morfemi: essa permette quindi di superare i limiti alla validità interlinguistica della nozione intuitiva di composto. L'approccio basato su lessemi, proposto da Aronoff (1976, 1994) e ripreso da Scalise (1984), sembra quindi il più adatto a rendere conto della composizione, in cui le unità di input sono lessemi che assumono una determinata forma all'interno di un composto (ad es., nei composti del latino, facio può occorrere come -facio, -ficus, -ficium, -fex, -fico): come osservato da Montermini (2010: 87), infatti, assumendo la nozione di lessema, non è più fondamentale distinguere tra radice, tema, parola autonoma, dal momento che «[i]n this perspective, roots and stems are simply concrete manifestation of a lexeme in a subset of its syntactic and morphological usage».

-

<sup>&</sup>lt;sup>3</sup> Fanno eccezione i composti in cui il primo costituente occorre nella forma apocopata senza la vocale o la sillaba finali (es. *gentiluomo* invece che *gentileuomo*, *granduca* invece che *grandeduca*).

<sup>&</sup>lt;sup>4</sup> Le ipotesi formulate per spiegare la natura del primo costituente dei composti V-N verranno discusse più diffusamente al § 4.3.4.

<sup>&</sup>lt;sup>5</sup> La nozione di lessema in morfologia è generalmente associata al lavoro di Matthews (1972); i più recenti contributi dedicati a tale nozione sono raccolti nel volume curato da Bonami *et al.* (2018).

Il secondo problema legato alla definizione di composto come 'parola di più parole' riguarda i confini della categoria di composto rispetto a quelli di espressione multiparola (*multiword expression*),<sup>6</sup> categoria che include una varietà piuttosto ampia di fenomeni (espressioni idiomatiche, parole sintagmatiche, espressioni cristallizzate, etc.). La letteratura dedicata alle espressioni multiparola è decisamente ricca<sup>7</sup> e numerosi sono i criteri proposti per delimitarla rispetto ai composti e ai sintagmi, a seconda dell'approccio adottato: questo lavoro assume come quadro teorico la Morfologia delle Costruzioni (Booij 2009b), i cui assunti principali saranno oggetto del capitolo 2; si rimanda quindi al § 2.4.3 per la discussione sui confini tra fenomeni multiparola e composti all'interno di questo *framework* teorico e con particolare riferimento all'italiano.

Un'ulteriore osservazione sulla nozione di composto riguarda la collocazione delle parole composte tra il piano della morfologia e quello della sintassi: in questo lavoro la composizione è considerata una categoria morfologica e, seguendo Grandi (2006: 36), «"a doppio accesso", in cui, cioè, trovano collocazione elementi lessicali che hanno origine produttivamente nella morfologia ed elementi lessicali che invece traggono origine nel componente sintattico. Questi ultimi subiscono processi di lessicalizzazione e fossilizzazione che possono di fatto rendere possibile un loro trattamento in termini morfologici». Questo presupposto teorico consente di considerare tra le parole composte dell'italiano anche entità nate dalla cristallizzazione e successiva univerbazione di espressioni originariamente nate sul piano della sintassi ma che hanno assunto caratteristiche formali assimilabili a quelle dei composti: è il caso, ad esempio, di molti composti N-A (es. camposanto, terracotta, etc) o di forme che hanno origine dall'univerbazione di espressioni latine (es. palafitta da pala fixa, mappamondo da mappa mundi, etc.). Tali elementi lessicali, pur avendo un'origine diversa, presentano caratteristiche formali del tutto assimilabili a quelle dei composti nativi e possono costituire il modello per la creazione di nuove forme attraverso un procedimento, questa volta, puramente morfologico.

<sup>&</sup>lt;sup>6</sup> Nel corso del lavoro, si utilizzeranno i termini *espressione multiparola* e *polirematica* come sinonimi (cfr. § 2.4.3).

<sup>&</sup>lt;sup>7</sup> Si rimanda a Masini (2012) per un quadro sulle espressioni multiparola, con particolare riferimento all'italiano.

#### 1.2 Delimitazione

La categoria di composto è stata spesso discussa in relazione ai confini che la separano da altre categorie: di seguito si discute la delimitazione tra ciò che in questo lavoro si considera composizione nativa, e che sarà oggetto di analisi in questo lavoro, e altri due meccanismi di formazione di parola, ossia la derivazione (in particolare, la prefissazione) e la composizione neoclassica.<sup>8</sup>

Le parole derivate si distinguono da quelle composte in quanto costituite da uno o più affissi e una parola autonoma, laddove i composti sono definiti dall'unione di due parole autonome: questa distinzione, pur permettendo di tracciare una linea di demarcazione piuttosto netta tra i due meccanismi, non fornisce una soluzione ad alcuni casi dubbi, la quasi totalità dei quali coinvolge prefissi invece che suffissi. Vale quindi la pena richiamare in questa sede le proprietà che sono state individuate per identificare i prefissi dall'italiano (cfr. Iacobini 2004b: 105-108):

- 1. sono affissi privi di categoria sintattica, che si premettono a una base lessicale con lo scopo di modificarne la semantica;
- 2. formano parole nuove premettendosi a parole;
- 3. non possono costituire base di derivazione né hanno una flessione;
- 4. costituiscono un inventario tendenzialmente chiuso;
- 5. possono occorrere soltanto in posizione iniziale di parola;
- 6. presentano restrizioni circa la lunghezza;
- 7. non influiscono sull'accento primario;
- 8. sono in rapporto di subordinazione con la base lessicale rispetto a cui svolgono la funzione di determinante:
- 9. selezionano la base secondo criteri legati solo parzialmente alla categoria sintattica della base, ma spesso di natura semantica;
- 10. esprimono significati di tipo funzionale-relazionale, non lessicale né flessivo;
- 11. formano parole di norma endocentriche in cui non rappresentano mai l'elemento testa:

<sup>&</sup>lt;sup>8</sup> Ci si concentra in particolare sui confini con prefissazione e composizione neoclassica perché, nel corso dell'analisi diacronica, assumeranno rilievo i rapporti che la composizione nativa dell'italiano intrattiene con questi meccanismi di formazione di parola.

<sup>&</sup>lt;sup>9</sup> Un caso discusso in letteratura è quello degli avverbi in *-mente*, considerati da Scalise (1990) derivati e da Bosque (1987) e Zagona (1990) composti.

Tra le proprietà identificate da Iacobini, le più cruciali per distinguere i prefissi dalle parole e ascriverli alla categoria degli affissi derivazionali sono la posizione fissa (5), l'impossibilità di essere derivati o flessi (3) e la relazione determinativa con la base (8). Un caso dubbio tra prefissazione e composizione è quello delle unità che possono comparire anche autonomamente in sintassi come le preposizioni, ad es. sotto- (ma anche contro-, dopo-, oltre-, sopra-, senza-) in parole come sotto-commissione, sotto-scala, sotto-bicchiere, etc.<sup>10</sup> Questo tipo di forme è stato ascritto alla composizione da Scalise (1994: 136) e alla prefissazione da Montermini (2008: 137-148). In particolare, Scalise (1994) propone di distinguere tra derivati con il prefisso sotto-, come sotto-commissione, e composti esocentrici P-N del tipo sottobicchiere, sottoscala: su questa linea si pongono anche Iacobini (1992: 110-112, 1999: 378-379) e Iacobini & Thornton (1992: 44-45). Osservado in primis che tale distinzione implica riconoscere l'esistenza di due realtà omofone, una infralessicale (sotto- come prefisso) e una lessicale (sotto come preposizione), Montermini (2008) considera prefissi non prototipici tutti questi elementi.<sup>11</sup> In questo lavoro si è scelto di sostenere la posizione di Montermini (2008) e quindi di non includere nell'analisi parole come sottoscala.

Pur non essendo oggetto di analisi in questo lavoro, la composizione neoclassica merita una discussione in questa sede, in virtù dell'interazione che ha stabilito con la composizione nativa. Per composizione neoclassica si intende quel «tipo di formazione delle parole che utilizza elementi formativi tratti dalle lingue classiche per coniare termini di ambito tecnico-scientifico, usati primariamente con funzioni designative e classificatorie» (Iacobini 2004: 69). Gli elementi che prendono parte a questo tipo di composizione non possono occorrere autonomamente e possono essere considerati elementi legati: esempi di questo tipo di unità sono -logia, -forme, bio-, cardio-, etc.). <sup>12</sup> Se già a partire dal Settecento la composizione neoclassica costituisce un meccanismo a sé, alternativo alla composizione nativa e alla derivazione, è solo a partire dal Novecento

\_

<sup>&</sup>lt;sup>10</sup> Si rimanda a Amiot (1997, 2002, 2005) per casi analoghi in francese.

<sup>&</sup>lt;sup>11</sup> Per le altre motivazioni con cui si sostiene questa tesi si rimanda a Montermini (2008: 139-148).

<sup>&</sup>lt;sup>12</sup> Per identificare gli elementi che prendono parte alla composizione neoclassica sono state proposte diverse denominazioni, tra le quali *elementi formativi* (cfr. Iacobini 2010), *affissoidi* in (cfr. Migliorini 1963), *semiparole* (cfr. Scalise 1984), *confissi* (cfr. De Mauro 2009). Seguendo Iacobini (2010), in questo lavoro ci si riferirà a questo tipo di costituenti come a *elementi formativi*, in virtù della maggior neutralità di questa denominazione rispetto alla dicotomia affisso-parola e della vicinanza con i termini corrispondenti in altre lingue (cfr. *combining form* in inglese e *formant* in francese).

che un significativo numero di composti neoclassici entra a far parte della lingua comune (ad es. frigorifero, telefono, etc.), stabilendo un'interazione con il lessico e le regole di formazione di parola dell'italiano. Tale interazione ha avuto come principale conseguenza la diffusione nella lingua comune di composti, generalmente con testa a destra, costituiti da un elemento formativo<sup>13</sup> e una parola autonoma (il tipo *agriturismo*, termocoperta, biotecnologia, etc.): la combinazione che più facilmente viene integrata nella lingua comune sembra quindi essere elemento formativo + parola autonoma (idromassaggio, telecomando, etc.), in cui almeno uno dei due elementi è facilmente riconoscibile dal parlante; va inoltre notato, scorrendo la lista degli elementi formativi proposta da Iacobini & Giuliani (2010),14 che questi presentano molto spesso una struttura bisillabica. Se normalmente la parola autonoma in questo tipo di forme è un sostantivo, non è esclusa la possibilità che si tratti di un verbo (es. videoregistrare). Come messo in luce da Iacobini (2013), questo tipo di verbi composti ha origine dalla convergenza dell'esito di due processi: la retroformazione a partire da nomi o aggettivi costituiti da un elemento formativo in prima posizione e un nome deverbale (videoregistratore > videoregistrare); la conversione da composti costituiti da un elemento formativo e un nome (es. fotocopia > fotocopiare). Tuttavia, il progressivo aumento in tempi recenti di questi verbi composti induce a ritenere che si possa ormai parlare di una vera e propria regola di formazione di parola produttiva N + V > V (o, in termini costruzionisti, di una nuova costruzione, cfr. Iacobini 2014), ulteriore testimonianza dell'influenza che la composizione con un elemento formativo e un nome o un aggettivo sta avendo sulla composizione italiana.

Nel complesso, quindi, come osservato da Iacobini (2004: 78), la costante introduzione nel lessico di composti che presentano questa struttura dimostra «il crescente grado di inserimento del tipo Determinante + Determinato all'interno del sistema compositivo dell'italiano, e l'instabilità del confine che divide le parole modificate dagli elementi formativi».

٠

<sup>&</sup>lt;sup>13</sup> Gli elementi formativi che entrano in composizione con parole autonome possono essere classificati in quattro tipologie: 1) elementi formativi di origine greco-latina (es. *multi-, -termo, -logia, -cida,* etc.); 2) accorciamenti di parole autonome (es. *foto-* da *fotografia, eco-* da *ecologia,* etc.); 3) elementi ricavati per segmentazione da una parola (es. *-poli* da *tangentopoli*); 4) parole modificate attraverso la modificazione della vocale finale in *-i* o *-o* per uniformarsi allo schema formativo dei composti neoclassici (es. *insetticida, erbivoro, filmografia,* etc.). Per una trattazione più ampia si rimanda a Iacobini (2004: 72-79).

<sup>&</sup>lt;sup>14</sup> In particolare, questo lavoro si segnala perché propone una classificazione degli elementi formativi sulla base di un approccio multidimensionale.

#### 1.3 Tentativi di classificazione

Le parole composte sono state oggetto di numerosi tentativi di classificazione, basati su criteri sia grammaticali sia semantici: in questo paragrafo si focalizza l'attenzione sulla classificazione proposta da Bisetto & Scalise (2005), di cui in questo lavoro si riprenderanno alcuni aspetti fondamentali, e alla successiva versione descritta in Scalise & Bisetto (2009); la seconda parte del paragrafo è invece dedicata a una breve panoramica su altre proposte di classificazione dei composti basate su criteri di natura tipologica e semantica.

Bisetto & Scalise (2005) propongono una classificazione dei composti valida negli intenti per tutte le lingue del mondo, con l'obiettivo di mettere ordine tra i diversi schemi classificatori proposti precedentemente, alcuni basati soltanto su particolari lingue, altri su criteri ed etichette eterogenee e incoerenti (es. presenza di un elemento testa, categoria dei costituenti, relazione grammaticale, etc.). In particolare, i precedenti tentativi di classificazione (cfr. Bloomfield 1933; Marchard 1969; Spencer 1991; Fabb 1998; Olsen 2001; Bauer 2001; Booij 2005a) individuano cinque tipologie di composto: subordinativi (es. housewife), coordinati (es. poet painter), appositivi (es. woman doctor), esocentrici (es. redskin), sintetici (es. truck driver). Come messo in luce da Bisetto & Scalise (2005), tali categorie sono state individuate sulla base di criteri eterogenei e difficilmente collocabili sullo stesso piano: ad esempio, la nozione di composto esocentrico appartiene a un livello di analisi differente rispetto a quella di composto copulativo, che pure si trovano sullo stesso piano, ad esempio, nella classificazione di Haspelmath (2002), illustrata nella figura 1.

<sup>&</sup>lt;sup>15</sup> Per una trattazione sui limiti delle precedenti classificazioni si rimanda a Bisetto & Scalise (2005: 320-326). Tra i precedenti tentativi va senz'altro menzionata, per l'influenza che ha avuto sulle successive classificazioni, la classificazione dei composti proposta dai grammatici sanscriti: per una discussione sui criteri adottati e i limiti di questa classificazione si rimanda a Bauer (2017: 55-106).

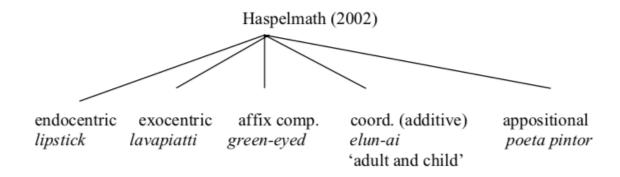


Figura 1 Classificazione dei composti proposta da Haspelmath (fonte: Scalise & Bisetto 2005)

Partendo dal presupposto che la caratteristica prototipica di un composto è quella di essere costituito dall'unione di due elementi legati da una relazione grammaticale non espressa da alcuna marca esplicita, i due autori propongono di adottare come criterio principale proprio il tipo di relazione grammaticale che intercorre tra i due costituenti, distinguendo tra:

- 1. composti subordinativi (SUB);
- 2. composti attributivi (ATTR);
- 3. composti coordinativi (COORD).

I composti subordinativi sono costituiti da due elementi legati da una relazione di dipendenza sintattica implicita: in italiano, fanno parte di questa categoria sia i composti V-N (es. *portalettere*), costituiti da un elemento verbale e un nome che ne costituisce un argomento, sia i composti N-N del tipo *capostazione*, in cui il secondo costituente è un complemento del primo.

I composti attributivi presentano un elemento che svolge la funzione di modificatore rispetto a un altro costituente: in italiano, tale modificatore può essere sia un aggettivo che qualifica il nome testa (es. *camposanto* o *pellerossa*) sia un nome che esprime una proprietà della testa (es. *pesce spada*).<sup>16</sup>

I composti coordinativi sono costituiti da due elementi legati da una relazione di coordinazione, generalmente parafrasabile attraverso la congiunzione *e* (es. un *attore*-

<sup>&</sup>lt;sup>16</sup> Spesso, ma non sempre (cfr. Radismký 2016), il nome modificatore in un composto N-N ha un'interpretazione metaforica (es. "riunione *fiume*" ma "nazione *simbolo*").

cantante è una persona che è allo stesso tempo un attore e un cantante): in italiano, possono occorrere in questo tipo di composti sia verbi (es. toccasana) sia nomi (es. studente-lavoratore) sia aggettivi (es. giallorosso).<sup>17</sup>

Il secondo criterio adottato in questa classificazione è la presenza di un elemento preminente, definito 'testa del composto': ciascuna delle tre classi di composto fin qui individuate può infatti contenere sia composti in cui è possibile individuare una testa, detti 'endocentrici', sia composti in cui non è riconoscibile una testa, detti 'esocentrici'. Il concetto di 'testa' in morfologia è mutuato dalla sintassi¹8 e indica l'elemento più importante del composto che trasmette la parte del discorso, i tratti flessivi e/o quelli semantici all'intera forma: esso può quindi essere definito da più punti di vista (morfologico, semantico, etc.). Nei composti cosiddetti esocentrici non è possibile identificare un elemento che trasmette le sue proprietà all'intero composto: fanno parte di questa categoria, ed esempio, i composti come *parapetto* o *portalettere* costituiti da un elemento verbale e un suo argomento. Né la categoria di output di queste forme, nome (o aggettivo), né i tratti flessionali e semantici sono ereditati da uno dei due costituenti: il significato di un composto esocentrico non è infatti predicibile dal significato dei singoli elementi che lo compongono, ma è posto 'fuori' dalla parola composta.

La gran parte dei composti coordinativi dell'italiano può essere considerata 'bicefala' dal punto di vista semantico: composti come *divano-letto* vengono definiti «hyponymic coordinate compound» (cfr. Arcodia, Grandi & Wälchli 2010), dal momento che il referente identificato dal composto rappresenta un iponimo delle parti, *divano* e *letto*); più rari sono gli «hyperonymic coordinate compounds», del tipo *Emilia Romagna*, in cui il referente è un iperonimo dei due costituenti.

La classificazione proposta da Bisetto & Scalise (2005) individua quindi sei tipologie di composto, esemplificate attraverso composti dell'inglese e dell'italiano nelle figure 2 e 3.

<sup>18</sup> Sulla relazione tra il concetto di testa sintattica, di testa morfologica e di testa di un composto si rimanda a Bauer (2017: 29-40).

17

<sup>&</sup>lt;sup>17</sup> Si noti inoltre che, nelle lingue romanze, i composti coordinativi possono essere ricorsivi, a differenza di tutte le altre tipologie di composto (es. *poeta pittore regista*).

SUBORD	DINATE	ATTRIB	UTIVE	COORD	INATE
ENDO	EXO <sup>9</sup>	ENDO	EXO	ENDO	EXO
apple cake	kill joy	blue cheese	white collar	actor author	mind brain
brain death	cut throat	atomic bomb	green house	priest hermit	mother child
finger print		back yard	pale face	singer bassist	north east
mail man		French kiss	long legs	dancer singer	
sun glasses		ape man	free lance	artist designer	
water pipe		ghost writer		fighter bomber	
taxi driver		key word		king emperor	
stone cutter		public opinion		merchant tailor	
arm control		sword fish			
baby care					
agora phobia					

Figura 2 Classificazione dei composti proposta da Bisetto & Scalise (2005) con esempi dall'inglese (fonte: Bisetto & Scalise 2005: 328)

Table 1. Classes of Italian compounds

	Subordinate	Attributive	Coordinate
Endocentric	capo-stazione	cassa-forte	poeta pittore
	(chief-station)	(case/box-strong)	(poet painter)
	'stationmaster'	'safe'	'poet painter'
	trasporto latte	viaggio lampo	divano-letto
	(transportation milk)	(journey lightening)	(sofa bed)
	'milk transportation'	'very fast journey'	'sofa bed'
Exocentric	porta-lettere	viso pallido	Emilia Romagna
	(carry-letters)	(face pale)	(Emilia Romagna)
	'mailman'	'facepale'	'Emilia Romagna'
	sotto-scala	piedi piatti	dormi-veglia
	(under-stairway)	(feet flat)	(sleep-wake)
	'closet under the stairway'	'cop'	'drowsiness'

Figura 3 Classificazione proposta da Bisetto & Scalise (2005) con esempi dall'italiano (fonte: Masini & Scalise 2012: 77)

Come illustrato dalla figura 3, l'italiano contemporaneo possiede composti appartenenti a tutte le sei tipologie individuate da Bisetto & Scalise (2005). In particolare, i composti esocentrici dell'italiano sono prevalentemente: 1) composti subordinativi del tipo V-N (vd. sopra, es. *guardalinee, spremiagrumi*), che nella maggior parte dei casi designano

individui o oggetti che svolgono una determinata azione espressa dal verbo in prima posizione; 2) composti attributivi del tipo N-A (es. *piedi piatti*), che generalmente designano un individuo o un animale a partire da una caratteristica fisica posseduta (es. il *pettirosso* è un uccello caratterizzato dal colore rossastro del piumaggio sul petto); 3) composti coordinativi del tipo V-V (es. *dormiveglia*), con categoria di uscita sempre nominale, che generalmente indicano un'azione, spesso svolta da un gruppo di persone. La classificazione Bisetto & Scalise (2005) ha costituito il punto di partenza per altri lavori, tra cui una serie di indagini di carattere computazionale, volte all'estrazione a alla classificazione semiautomatica dei composti N-N endocentrici (cfr. Baroni, Guevara & Pirelli 2007, 2009): partendo dalla classificazione Bisetto & Scalise (2005), gli autori ne elaborano una versione rivista, sulla base di un'analisi *corpus-based* della distribuzione e di alcune caratteristiche flessive (es. la formazione del plurale) di un campione di composti N-N dell'italiano. Lo schema rivisto da Baroni, Guevara & Pirrelli (2006) è rappresentato nella figura 4.

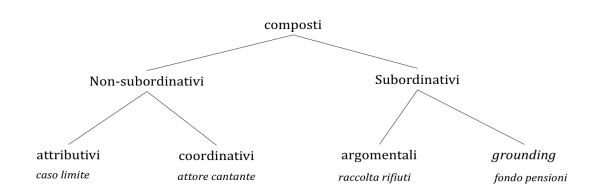


Figura 4 Classificazione dei composti proposta da Baroni, Guevara & Pirrelli (2006)

Nella classificazione sopra illustrata, i composti sono suddivisi in due macro-classi: la classe dei Non-subordinativi, che include i composti in cui non è possibile individuare una relazione sintattica di dipendenza tra i due costituenti, e la classe dei Subordinativi, in cui invece è possibile individuarla. Ciascuna delle due classi è a sua volta suddivisa in due sotto-classi: i composti non-subordinativi possono essere attributivi (es. *caso limite*) o coordinativi (es. *attore cantante*); i composti subordinativi si dividono in argomentali

(in cui la testa del composto è un deverbale e l'altro costituente un argomento del verbo, es. *raccolta rifiuti*) e *grounding* (in cui la testa non è un deverbale, es. *fondo pensioni*).

La classificazione del 2005 è stata successivamente rivista in Scalise & Bisetto (2009), in cui sono state introdotte le seguenti modifiche:

- a. la categoria dei composti attributivi è stata rinominata in ATAP (attributivi/appositivi) e suddivisa in due categorie: i composti attributivi N-A (es. *roccaforte*, *pettirosso*, etc.), costituiti da un nome e un aggettivo, e i composti appositivi N-N (es. *pesce spada*, *riunione fiume*, etc.), costituiti da due nomi;
- b. all'interno della categoria dei composti subordinativi N-N, sono state individuate due classi a seconda della relazione semantico-grammaticale che intercorre tra i due costituenti: i composti argomentali o *verbal nexus* (es. *smaltimento rifiuti*), in cui i primo nome è un deverbale e il secondo ne rappresenta un argomento, e i composti *ground* o *grounding* (es. *sala computer*).

Tali modifiche hanno portato lo schema di classificazione ad assumere la struttura illustrata, con esempi dall'italiano, nella figura 5.

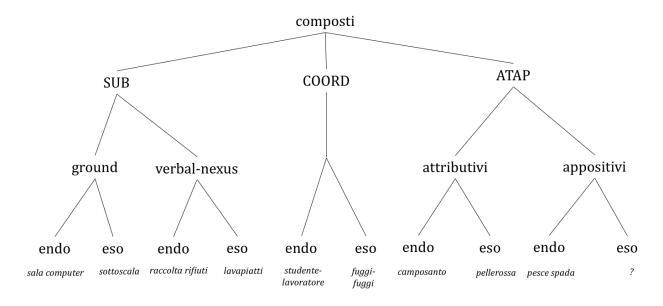


Figura 5 Classificazione Scalise & Bisetto (2009)

Come illustrato dallo schema nella figura 5, la nuova classificazione conserva la tripartizione basata sul tipo di relazione grammaticale ma inserisce due nuove tipologie all'interno della classe dei SUB (composti *ground* e composti *verbal nexus*) e rimodula la classe degli ATTR in ATAP, includendo quindi sia attributivi sia appositivi; ciascuna tipologia include sia composti esocentrici sia composti endocentrici, ad esclusione dei composti appositivi (es. *pesce spada*, *riunione fiume*, etc.) tra cui non sono attestate forme esocentriche.

L'attribuzione di alcune tipologie di composto a una delle classi identificate da Scalise & Bisetto (2009) è stata oggetto di dibattito in letteratura: è il caso dei composti del tipo ristorante-pizzeria o studente-lavoratore, considerati come coordinativi endocentrici da Scalise & Bisetto (2009), Arcodia, Grandi, Wälchli (2010) e Arcodia (in stampa); come attributivi da Radimský (2015a: 102-111). Diversi sono i criteri presi in esame per dirimere la questione: Arcodia (in stampa) propone il criterio della reversibilità per distinguerli dagli attributivi; sulla base di elementi morfologici (presenza di una testa a sinistra) e semantici (reversibilità, asimmetria) Radimský li ascrive invece alla classe degli attributivi. Un altro caso discusso, sempre legato alla categoria dei composti coordinativi, riguarda i coordinativi esocentrici N-N, che per Scalise & Bisetto (2009) sono limitati a nomi propri, come Emilia Romagna, e a binomi in cui la relazione di coordinazione è esplicitata da una congiunzione (come in sale e pepe): diversa è la posizione di Radimský (2015a: 121-125) che, attraverso uno studio quantitativo basato sul corpus itWaC (cfr. Baroni et al. 2009), individua almeno ottanta casi di composti coordinativi esocentrici costituiti anche da nomi comuni, classificabili dal punto di vista semantico in «intermediate-denoting» (come nord-est o primavera-estate) e «additive» (come carico-scarico, radio-televisione). Come mostrato dai casi presentati, la classe dei coordinativi costituisce il nodo più discusso della classificazione Scalise & Bisetto (2009).

Come verrà discusso più nel dettaglio al § 4.2, la classificazione che si adotterà in questo lavoro riprende alcuni assunti fondamentali della classificazione Bisetto & Scalise (2005) inseriti all'interno del quadro teorico della Morfologia delle Costruzioni (cfr. capitolo 2). In particolare, verranno ripresi i seguenti elementi:

- a. i tre tipi di relazione (subordinativa, coordinativa, di modificazione): in particolare, rispetto alla denominazione 'relazione attributiva' proposta in Bisetto & Scalise (2005), si è scelto di adottare quella di 'relazione di modificazione', includendo sia la relazione attributiva sia quella appositiva (cfr. Arcodia, Grandi & Montermini 2009);
- b. la distinzione tra composti esocentrici e composti endocentrici.

La relazione che intercorre tra i costituenti dei composti verrà quindi analizzata e classificata dal punto di vista grammaticale e non semantico: in questo senso, il presente lavoro non si colloca nell'ambito degli studi che hanno tentato di classificare le parole composte (in particolare, quasi esclusivamente il tipo N-N) sulla base della relazione semantica implicita tra i due costituenti (vd. oltre).

Un ulteriore tentativo di classificazione dei composti è quello proposto da Grandi (2006) e si basa sull'interazione di due parametri di natura tipologica: 1) la presenza di una struttura gerarchica tra i due costituenti del composto, che consente di distinguere tra composti di coordinazione e composti gerarchici; 2) la presenza di un'esplicita marca formale di dipendenza tra i membri del composto, che consente di distinguere tra composti in cui la relazione non è marcata, è marcata sintatticamente, è marcata morfologicamente. Sulla base di tale classificazione, Grandi (2006: 44) fornisce la seguente definizione di composto: «un composto è l'unione di due o più forme a cui i parlanti nativi attribuiscono autonomia lessicale e tra i quali vige una relazione di coordinazione o di subordinazione che è marcata mediante zero o mediante strategie puramente morfologiche». Tale definizione, valida a livello interlinguistico, sottolinea l'importanza di considerare la composizione in una prospettiva di interfaccia tra morfologia e sintassi.

Nell'ambito degli studi che hanno proposto una classificazione dei composti su base semantica possono essere individuate due posizioni: la prima è rappresentata da studi che, a partire dal tentativo di Levi (1978), hanno cercato di individuare un set chiuso di relazioni semantiche valido per tutti i composti; la seconda, sostenuta da Downing (1977), sottolinea l'impossibilità di individuare un insieme chiuso di relazioni

semantiche, in quanto queste sarebbero strettamente dipendenti dal contesto e quindi potenzialmente infinite.<sup>19</sup> La prima posizione, che raccoglie i cosiddetti *inventory approaches*, è stata quella che ha avuto più seguito ed è stata sostenuta in numerosi studi (cfr. Warren 1978; Lauer 1995; Girju et al. 2005, 2007a-b, 2009; tra gli studi di carattere computazionale, cfr. Ó Séaghdha 2007; Ó Seaghdha & Copestake 2008; Celli & Nissim 2009), ciascuno dei quali ha proposto un set di relazioni semantiche, a partire da quello proposto da Levi (1978), riportato nella tabella 1.

Relazione	Esempio
CAUSE <sub>1</sub>	flu virus
$CAUSE_2$	snow blindness
$HAVE_1$	college town
$HAVE_2$	company assets
$MAKE_1$	honey bee
$MAKE_2$	daisy chains
USE	water wheel
BE	chocolate bar
IN	mountain lodge
FOR	headache pills
FROM	bacon grease
ABOUT	adventure story

Tabella 1 Set di relazioni semantiche proposto in Levi (1978) per descrivere la semantica dei composti N-N

Accanto agli studi sopra citati, i cui set di relazioni costituiscono sostanzialmente una rivisitazione più o meno granulare di quello proposto da Levi (1978), vanno segnalati alcuni lavori basati sulle relazioni *qualia* e la teoria del Lessico Generativo elaborata da Pustejovsky (1995).<sup>20</sup> In particolare, in Johnston & Busa (1999), la semantica delle preposizioni nelle parole sintagmatiche Nome - Preposizione - Nome (es. *coltello da* 

\_

<sup>&</sup>lt;sup>19</sup> Secondo Downing (1977), infatti, «the existence of numerous novel compounds [...] guarantees the futility of any attempt to enumerate an absolute and finite class of compounding relationships». Pur ritenendo fallimentare la ricerca di un set chiuso di relazioni semantiche, Downing fornisce una lista delle dodici relazioni più utilizzate dai parlanti per la creazione di composti N-N dell'inglese: WHOLE-PART (es. duck foot), PART-WHOLE (es. pendolum clock), TIME (es. summer dust), OCCUPATION (es. coffee man), HALF-HALF (es. giraffe-cow), COMPOSITION (es. stone forniture), COMPARISON (es. pumpkin bus), PLACE (es. Easter Oregon meal), SOURCE (es. vulture shit), PRODUCT (es. honey glands), USER (es. flea wheelbarrow), PURPOSE (es. hedge hachet).

<sup>&</sup>lt;sup>20</sup> La Struttura Qualia, ispirata alla dottrina delle cause di Aristotele nell'interpretazione di Moravcsik (1975), è uno strumento teorico elaborato da Pustejovsky (1995) per rappresentare la semantica di un elemento lessicale. In particolare, si individuano quattro relazioni *qualia*, ognuna esprimente un aspetto del significato di una parola: *quale* Formale ('che cosa è x'), *quale* Costitutivo ('di che cosa è fatto x'), *quale* Agentivo ('origine di x'), *quale* Telico ('lo scopo di x').

*cucina*) delle lingue romanze viene analizzata attraverso le quattro relazioni *qualia*, con particolare attenzione alla relazione telica, che esprime lo scopo dell'evento.

In Pustejovsky *et al.* (2008) si descrive un linguaggio di *mark-up* per l'annotazione di composti N-N, basata sul Lessico Generativo. In Bouillon *et al.* (2014), si propone uno schema di annotazione per i composti N-N (es. *treno merci*) e le parole sintagmatiche Nome - Preposizione - Nome (es. *bicchiere da vino*) dell'italiano e del francese, basato su tre livelli:

- 1. relazione semantica tra i due costituenti;
- 2. tipo semantico dei costituenti;
- 3. ruolo semantico dei costituenti.<sup>21</sup>

Alcuni esempi di annotazione sono riportati in (1):

(1) a. torta\_1 gelato\_2 'ice cream cake'CONST[made \_of]1 whole / artifact2 part / substance

b. fucile\_1 da caccia\_2 'hunting rifle'TELIC[used\_for]1 instr / artifact2 event

Come mostrato dagli esempi in (1), ciascuna forma è analizzata rispetto alla relazione semantica (in (1a) si tratta della relazione made\_of del *quale* Costitutivo; in (1b) della relazione *used\_for* del *quale* Telico) e al tipo e al ruolo semantico dei singoli costituenti (ad es. in *torta gelato*, il primo costituente è un 'artefatto' che ricopre il ruolo di 'tutto', il secondo è una 'sostanza' che ricopre il ruolo di 'parte'). Da questo lavoro è stata ripresa l'idea di descrivere la semantica dei composti attraverso i tipi semantici dei costituenti: come verrà spiegato più nel dettaglio al § 3.3.2, si adotterà un tagset di tipi semantici

<sup>&</sup>lt;sup>21</sup> Lo schema di annotazione costituisce una rivisitazione di quello proposto in Pustejovky *et al.* (2008). Per i tagset e i relativi esempi si rimanda a Bouillon *et al.* (2014).

elaborato sulla base di quello proposto da Bouillon *et al.* (2014) per arricchire il formalismo costruzionista dal punto di vista semantico.

## 1.4 I composti dell'italiano

Uno degli obiettivi di questo lavoro è fornire un quadro dei composti che l'italiano ha creato nel corso della sua storia: nel capitolo 4 verranno quindi analizzate tutte le tipologie di composto attestate nel periodo preso in esame che, come verrà illustrato più nel dettaglio al § 3.1, va dalle Origini al 1947. Al fine di rendere confrontabile quanto emergerà dall'analisi con la situazione attuale, di seguito si riepilogano le tipologie di composto attestate e considerate produttive in italiano contemporaneo, identificate sulla base della categoria dei costituenti.

CATEGORIE DI INPUT	CATEGORIA DI OUTPUT	ESEMPI
NOME+NOME	NOME	cartamoneta, pescecane, posto auto
NOME+AGGETTIVO	NOME	camposanto, cassaforte
NOME+VERBO	VERBO	barcamenare
AGGETTIVO+NOME	NOME AGGETTIVO	nome: altopiano, bassorilievo, biancospino, buongusto, due pezzi, gentildonna, maleducazione, piattaforma agg: giallo crema, rosso porpora, verde bottiglia
VERBO+NOME	NOME	accendisigari, apriscatole, cantastorie, guardalinee, scolapasta
PREPOSIZIONE+NOME	NOME	doposcuola, lungotevere, oltretomba, senzatetto, sottobosco
AGGETTIVO+AGGETTIVO	AGGETTIVO	giallorosso, italoamericano, socioculturale
AVVERBIO+AGGETTIVO	AGGETTIVO	sempreverde
AVVERBIO+VERBO	VERBO	maledire, benedire

CATEGORIE DI INPUT	CATEGORIA DI OUTPUT	ESEMPI
VERBO+AVVERBIO	NOME	buttafuori, posapiano, tiratardi
VERBO+VERBO	NOME	fuggifuggi, giravolta, leccalecca, saliscendi, toccasana

Tabella 2 Tipologie di composto produttive in italiano contemporaneo (*fonte*: adattata da Iacobini 2010)

Non tutte le tipologie di composto riportate nella tabella 2 risultano essere produttive nella stessa misura: tuttavia, sebbene alcuni studi abbiano analizzato la produttività di singoli pattern in italiano contemporaneo (es. Ricca 2010 per i composti V-N), non vi sono lavori che abbiano preso in esame e misurato la produttività di tutte le tipologie in modo sistematico.

## 1.5 Rassegna dei principali studi sull'italiano e domande aperte

Alla composizione dell'italiano sono stati dedicati prevalentemente contributi di carattere sincronico, che nella maggior parte dei casi si sono focalizzati su particolari tipologie di composto, produttive in italiano contemporaneo: di seguito si riepilogano i principali lavori, mettendone in luce il contributo allo studio della composizione e le questioni aperte.

La prima trattazione dedicata alla composizione dell'italiano è l'opera di Tollemache (1945), in cui per la prima volta si focalizza l'attenzione in maniera sistematica sulle diverse tipologie di composto attestate in italiano. L'opera di Tollemache getta le basi per le trattazioni successive (cfr., tra gli altri, Dardano 1978, 2009) che, accanto alla più studiata derivazione, descrivono la composizione in quanto meccanismo attraverso cui l'italiano ha espanso il proprio lessico.

Un'esaustiva descrizione delle parole composte attestate in italiano contemporaneo è proposta da Bisetto (2004), in cui sono analizzate le proprietà formali e semantiche delle diverse tipologie di composto, classificate *in primis* sulla base della categoria di output (distinguendo quindi tra composti nominali, aggettivali, numerali, verbali e avverbiali), e in secondo luogo sul tipo di costituenti, della relazione grammaticale tra i due costituenti e dell'endo/esocentricità (vd. sopra, classificazione Bisetto & Scalise 2005).

Il più recente contributo che fornisce una panoramica sulla composizione dell'italiano contemporaneo è, infine, Masini & Scalise (2012).

I contributi fin qui citati forniscono un quadro generale su quali combinazioni di costituenti sono possibili in italiano e quali sono le tipologie di composto attualmente più produttive. La maggior parte degli studi dedicati alle parole composte dell'italiano non ha però carattere sistematico ma si sofferma su particolari tipologie. Il tipo di composti su cui più a lungo si è dibattuto è quello di *lavapiatti, spargisale*: generalmente considerata come la tipologia più produttiva, alla composizione verbonominale sono stati dedicati numerosi studi, in ottica sia sincronica sia diacronica. Ad attirare in primis l'attenzione degli studiosi è stato il primo costituente di queste forme (es. lava-, spargi-), sulla cui natura ancora oggi non si è giunti a una interpretazione condivisa (cfr. per le diverse ipotesi Meyer-Lübke 1890: 54; Prati 1931: 23; Darmesteter 1984: 112; Tollemache 1945: 28; Scalise 1992: 192; Di Sciullo & Ralli 1994: 61-75; Bisetto 1999: 530, 2005).<sup>22</sup> Accanto a studi di tipo teorico, più recentemente, questo tipo di composti è stato oggetto di analisi corpus-based: in Ricca (2005, 2010) si fornisce un quadro dettagliato delle proprietà morfologiche, semantiche e distribuzionali dei composti V-N dell'italiano contemporaneo; Štichauer (2015ab, 2016) fornisce un'analisi dei composti V-N tra Cinquecento e Ottocento, mettendone in luce la produttività e le proprietà semantiche nel corso dei secoli. Pur trattandosi di una tipologia di composti ampiamente descritta (e non solo in italiano: cfr., tra gli altri, Rosenberg 2005 per il francese), alcune questioni relative alla composizione verbonominale possono considerarsi ancora aperte: in primis, manca del tutto una descrizione dei composti V-N attestati in italiano antico. Questo tipo di analisi permetterebbe infatti di chiarire se questa tipologia è sempre stata particolarmente produttiva o se la sua fortuna può essere considerata uno sviluppo più recente (ipotesi sostenuta in Štichauer 2015a); ulteriori dati diacronici potrebbero contribuire, inoltre, alla riflessione intorno alla natura del primo costituente. Negli ultimi decenni, grande attenzione è stata riservata ai diversi tipi di composti N-N (cfr. Radismký 2015a per un quadro sistematico; Baroni, Guevara & Pirrelli 2009, 2007 per studi di tipo computazionale): il tipo subordinativo, es. treno merci, emergenza rifiuti (cfr. Terreni 2005; Lombardi Vallauri 2006); il tipo attributivo/appositivo, es. riunione

<sup>&</sup>lt;sup>22</sup> Per la discussione delle diverse ipotesi si rimanda al § 4.3.4.

*fiume, notizia bomba* (cfr. Grandi 2009; Grandi, Nissim & Tamburini 2011; Radimský 2016); il tipo coordinativo, es. *ristorante-pizzeria* (cfr. Radimský 2015b).

Come già accennato in apertura di questo paragrafo, la composizione dell'italiano è stata oggetto di studi prevalentemente in chiave sincronica: oltre ai già citati lavori di Štichauer (2015ab, 2016), un'eccezione è rappresentata dalle analisi diacroniche dedicate ai composti A-A (es. *agrodolce, giallorosso, marxista-leninista*) fornite in Grossmann & Rainer (2009) e D'Achille & Grossmann (2009, 2010), che hanno dimostrato l'utilità di adottare una prospettiva diacronica per chiarire la natura di questo tipo di composti e comprenderne meglio il comportamento in italiano contemporaneo.<sup>23</sup>

Nel complesso, quindi, la letteratura dedicata alla composizione dell'italiano è fortemente orientata all'analisi di particolari tipologie di composto produttive in epoca contemporanea, *in primis* quelle che sembrano essere emerse dal contatto con i numerosi prestiti inglesi, ossia i composti N-N. Manca quindi una trattazione sistematica che, in ottica diacronica, fornisca un quadro delle parole composte che l'italiano ha creato a partire dalla fase antica, e che metta in luce, da un lato, i rapporti con la composizione latina, dall'altro le linee di tendenza che, nel corso dei secoli, hanno determinato la situazione attuale: a questo scopo è dedicato il presente lavoro.

\_

<sup>&</sup>lt;sup>23</sup> Per una discussione dei risultati di questi studi in relazione ai dati analizzati in questo lavoro si rimanda al § 4.3.3.

#### 2 IL QUADRO TEORICO: LA MORFOLOGIA DELLE COSTRUZIONI

Questo capitolo riepiloga e discute i principali assunti della Morfologia delle Costruzioni. Dopo una breve introduzione alla Grammatica delle Costruzioni, modello teorico entro cui la Morfologia delle Costruzioni nasce e con cui condivide i principi di base, si discutono gli strumenti teorici costruzionisti, focalizzando l'attenzione sulla formazione delle parole. Nel presentare tali strumenti, se ne discutono i limiti e le potenzialità esplicative in relazione all'italiano e si fornisce una panoramica dei principali studi costruzionisti dedicati alla formazione delle parole.

#### Dalla Grammatica delle Costruzioni alla Morfologia delle Costruzioni 2.1

La Grammatica delle Costruzioni (Construction Grammar, d'ora in poi CxG) è una teoria nata dalle riflessioni di Charles Fillmore, Paul Kay e George Lakoff sugli aspetti idiomatici e irregolari del linguaggio, sviluppata in opposizione alle teorie generative che relegavano le irregolarità e gli idiom alla periferia della lingua.<sup>24</sup> Nella prospettiva generativa, infatti, l'attenzione è rivolta ai fenomeni regolari, elaborati e generati sul momento, in opposizione ai fenomeni idiomatici e irregolari, immagazzinati nella memoria del parlante. Diversamente, la teoria costruzionista si propone fin da subito come un modello unificato che analizza, con gli stessi strumenti teorici, ciò che è irregolare e idiomatico e ciò che è regolare.

Negli ultimi decenni, le riflessioni di Fillmore e colleghi hanno dato vita a una serie di approcci,<sup>25</sup> ognuno focalizzato su un particolare aspetto o livello di analisi del linguaggio, ma tutti accomunati da alcuni assunti di base, identificati da Goldberg (2013: 15-16) e discussi in Masini (2016: 39):

### 1. la nozione di Costruzione;<sup>26</sup>

<sup>&</sup>lt;sup>24</sup> L'articolo considerato il manifesto della Grammatica delle Costruzioni è Fillmore, Kay & O'Connor

<sup>&</sup>lt;sup>25</sup> In particolare, tra i modelli costruzionisti si possono menzionare la *Berkeley Construction Grammar* (cfr. Fillmore 2013), la Cognitive Construction Grammar (cfr. Goldberg 1995, 2006), la Radical Construction Grammar (cfr. Croft 2001), la Construction Morphology (cfr. questo paragrafo e sgg.) e la Sign-based Construction Grammar (cfr. Boas & Sag 2012).

<sup>&</sup>lt;sup>26</sup> Seguendo Masini (2016), in questo e nei seguenti capitoli si utilizzerà il termine *Costruzione* con l'iniziale maiuscola per riferirsi alla nozione specifica della CxG.

- 2. l'importanza della struttura superficiale;
- 3. la concezione della grammatica come *rete* di Costruzioni organizzate gerarchicamente;
- 4. il carattere language-specific delle Costruzioni;
- 5. la natura *usage-based* della teoria.

Tali principi, su cui si tornerà nei prossimi paragrafi, costituiscono i fondamenti teorici anche della Morfologia delle Costruzioni (*Construction Morphology*, d'ora in poi CxM), la teoria morfologica di stampo costruzionista elaborata da Geert Booij (cfr. Booij 2002a,b, 2005, 2010b), che indaga le Costruzioni a livello di parola.<sup>27</sup> In particolare, secondo la definizione di morfologia fornita da Booij (2014: 104): «Morphology deals with the systematic correlation between the form and the meaning of words».

Oltre ai cinque principi appena menzionati, la CxM riprende dalla CxG altri due assunti, di carattere più generale: il *continuum* lessico-grammatica e la visione non modulare della lingua. La CxM propone infatti una concezione del lessico e della grammatica come componenti poste lungo un *continuum*: a differenza di teorie di stampo dualista come i modelli *words-and-rules* (cfr. Pinker 1999), *dictionay-and-grammar* (cfr. Taylor 2012) e *rule-list fallacy* (cfr. Langacker 1987), in cui il lessico rappresenta una lista di irregolarità separata dalla grammatica, per la CxG (e, quindi, anche per la CxM), lessico e grammatica sono unificati nella nozione stessa di Costruzione.

Altro fondamento della CxG ripreso dalla CxM è la visione non modulare della lingua: la fonologia, la morfologia, la sintassi non costituiscono moduli separati e autonomi della lingua ma rappresentano le competenze dei parlanti, iscritte dentro le Costruzioni stesse (Masini 2016: 40). Entità o strutture che tradizionalmente vengono ascritte a un particolare livello di analisi (ad es. le parole al lessico, i morfemi alla morfologia, i sintagmi alla sintassi, etc.) vengono tutte rappresentate uniformemente attraverso la nozione di Costruzione. Ne discende che la CxM si occupa anche di Costruzioni lessicali più estese della singola parola, che si pongono a metà strada tra morfologia e sintassi: non è un caso, infatti, che tale modello abbia preso le mosse proprio dallo studio degli

30

<sup>&</sup>lt;sup>27</sup> Ma non solo: la CxM indaga infatti anche Costruzioni tradizionalmente considerate a metà strada tra la morfologia e la sintassi, come le espressioni multiparola (*multiword expressions*; cfr. § 2.2.3).

*idioms* e delle espressioni multiparola, come i verbi separabili con particella dell'olandese (Booij 2002a,b).

Nella prospettiva costruzionista, quindi, la morfologia non rappresenta una 'parte' della grammatica completamente separata dalla sintassi e dalla fonologia, dal momento che, come si osserverà al § 2.1, le informazioni morfologiche sono iscritte nelle Costruzioni insieme a quelle sintattiche e a quelle fonetiche. La differenza tra morfologia e sintassi è data piuttosto dal *tipo* di Costruzioni che esse tipicamente creano, dal punto di vista semantico e formale: la morfologia permette di creare Costruzioni (prevalentemente) a livello di parola, la sintassi a livello di frase. Trattandosi di meccanismi che operano sullo stesso sistema, è possibile che in alcuni casi esse entrino in competizione. <sup>28</sup> Sebbene non sia possibile delinearne i confini in maniera netta, nel quadro della CxM, la morfologia rappresenta un livello di analisi linguistica autonomo, dal momento che, come osserva Booij (2010b: 3), «there are specifically morphological generalizations or rules that cannot be reduced to either syntax or phonology».

Oltre che con la CxG, la CxM mostra evidenti legami con altri modelli teorici, quali la Morfologia lessicale (cfr. Aronoff 1976; Scalise 1984), il modello *Parallel Architecture* (cfr. Jackendoff 1997, 2002, 2013; Jackendoff & Audring 2018; Booij & Audring 2016), il modello *Word-and-paradigm* (cfr. Blevins 2006, 2018; Ackerman & Malouf 2018) e gli approcci *usage-based / exemplar-based* (cfr. Bybee 1985, 2010; Tomasello 2003). I rapporti con tali modelli verranno messi in luce nei seguenti paragrafi, parallelamente all'esposizione delle principali nozioni su cui tale teoria si basa.

### 2.2 La Costruzione: definizione, struttura, rappresentazione

La teoria morfologica elaborata da Booij si caratterizza essenzialmente come *sign-based*, *word-based*, *usage-based* (cfr. Masini & Audring, in stampa).

È una teoria *sign-based* in quanto basata sulla nozione di Costruzione (o *schema*), ossia un'associazione convenzionalizzata di una FORMA e una FUNZIONE; la definizione di Costruzione più frequentemente adottata dagli studi costruzionisti è quella di Goldberg (1995: 4):

<sup>&</sup>lt;sup>28</sup> La competizione tra Costruzioni a livello di parola e *phrasal lexemes*, accomunate dalla funzione di denominazione, è infatti oggetto di studio in ambito costruzionista (cfr. Hüning & Schlücker 2015; Masini, in stampa; Booij 2010b: 183).

C is a CONSTRUCTION iff<sub>def</sub> C is a form-meaning pair  $\langle F_i, S_i \rangle$  such that some aspect of  $F_i$ , or some aspect of  $S_i$ , is not strictly predictable from C's component parts or from other previously established constructions.

Una Costruzione è quindi un'entità linguistica dotata di almeno una proprietà formale o funzionale non strettamente predicibile dai suoi costituenti interni o da altre Costruzioni precedentemente individuate. Tale nozione è strettamente legata al concetto di *convenzionalizzazione*, applicabile a qualsiasi livello della Costruzione (formale, semantico, pragmatico, etc.).

La CxM è una teoria word-based, perché considera le parole (e non i morfemi) come le Costruzioni morfologiche minime: questa posizione allinea la CxM alle teorie morfologiche Word-and-paradigm (Matthews 1972; Anderson 1992; Blevins, Ackerman & Malouf, in stampa; Stump 2018), di contro alle teorie morpheme-based (cfr. Lieber 1992 e la Distributed Morphology, cfr. Halle & Marantz 1993; Siddiqi 2018). Per la CxM, il morfema rappresenta un'unità di analisi derivata secondariamente, dotata di rilevanza soltanto dal punto di vista astratto, non costruzionale (nel senso di Blevins 2006); ne consegue che per la CxM gli affissi non sono memorizzati autonomamente nel lessico mentale né hanno un significato indipendente dalla Costruzione entro cui occorrono (Booij 2010b: 15). Nel continuum lessico-sintassi delineato da Goldberg (2006: 5) per rendere conto dei diversi gradi di complessità delle Costruzioni, il morfema viene inizialmente indicato come Costruzione meno complessa: tuttavia, in versioni più recenti (cfr.Goldberg 2013: 17; tabella 3), l'autrice si allinea alla posizione della CxM.

Construction	Examples
Word	Iran, another, banana
Word (partially filled)	pre-N, V-ing
Idiom (filled)	Going great guns, give the Devil his due
Idiom (partially filled)	Jog <someone's> memory, <someone's> for the asking</someone's></someone's>
Idiom (minimally filled) The Xer the Yer	The more you think about it, the less you understand
Ditransitive construction: Subj V Obj1 Obj2 (unfilled)	He gave her a fish taco; He baked her a muffin
Passive: Subj aux VPpp (PPby) (unfilled)	The armadillo was hit by a car

Tabella 3. Costruzioni di diversa complessità (Fonte: adattata da Goldberg 2013: 17)

Infine, la CxM è una teoria *usage-based*, in quanto rivolge grande attenzione alla frequenza con cui le Costruzioni occorrono nell'uso e al ruolo che la dimensione quantitativa ricopre nel *processing* del linguaggio (cfr., tra gli altri, Pulvermüller, Cappelle & Shtyrov 2013). Essa mostra quindi evidenti legami con gli approcci *usage*- o *exemplar-based* (Bybee 1985, 2010; Tomasello 2003), in quanto sostiene che la competenza linguistica si forma attraverso un processo induttivo e *item-based*, a partire dagli esemplari che si incontrano nell'uso, sulla base dei quali i parlanti formulano delle generalizzazioni (ossia gli schemi costruzionali).

La CxM sostiene inoltre una visione *ridondante* del Lessico Mentale (in linea con il cosiddetto *full-entry model*), in cui sarebbero memorizzate anche le parole ottenute attraverso procedimenti produttivi, trasparenti e regolari, purché dotate di alti valori di frequenza, di contro al cosiddetto *impoverished-entry model* (cfr. Jackendoff 1975: 642) secondo cui nel Lessico Mentale sarebbero memorizzate soltanto le irregolarità e gli *idiom.*<sup>29</sup>

La struttura di una Costruzione può essere rappresentata come l'unione di più livelli (figura 6): in particolare, sul piano della FORMA si distinguono i livelli dell'informazione fonologica (PHON), morfologica (MORPH) e sintattica (SYN); il piano della FUNZIONE include invece i livelli dell'informazione semantica (SEM), pragmatica (PRAG) e discorsiva (DISC).

<sup>&</sup>lt;sup>29</sup> Tale approccio è stato adottato anche in ambito psicolinguistico, ad esempio nel *Dual Route Model* (cfr. Pinker & Ullman 2002).

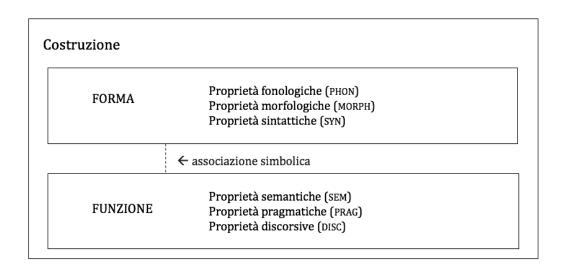


Figura 6. La struttura interna di una Costruzione *Fonte*: adattata da Masini (2016: 48)

Come la CxG, anche la CxM può quindi essere definita una teoria sia funzionale sia formale: entrambi gli aspetti sono tenuti in considerazione e inscritti nella struttura della Costruzione. Forma e funzione delle Costruzioni vengono indagate sulla base di dati provenienti da diverse fonti (cfr. Gries 2013): dati estratti da corpora sincronici (cfr. Stefanowitsch & Gries 2003) o diacronici (cfr. Bergs & Diewald 2008; Fried 2009; Hilpert 2013, 2015; Hilpert & Gries 2008), dati provenienti dalla psicolinguistica (behavioral data; cfr. ad es. in Bencini & Goldberg 2000 o Schlücker & Plag 2011). Tale metodologia, caratterizzata dall'integrazione di diversi tipi di dati, è stata definita converging evidence (cfr. Gries, Hampe & Schönefeld 2005) e rende la CxG e la CxM teorie in linea con l'elegante integrazione (graceful integration) auspicata da Jackendoff (2011: 590):30 «So to the extent that a theory of language permits a graceful integration with a plausible picture of the structure and function of the rest of the mind/brain, it places fewer demands on the genome, and therefore it is a better theory».

In particolare, a livello formale, vengono rappresentate le informazioni fonologiche, morfologiche e sintattiche: ad esempio, a livello di rappresentazione morfologica, una Costruzione codifica proprietà strutturali quali la categoria, la classe flessiva, il genere,

\_

<sup>&</sup>lt;sup>30</sup> Jackendoff (2011: 590), inoltre, mette in luce che «the criterion of graceful integration conflicts in many respects with the criterion of perfection». Il criterio della *perfezione* è legato al modello minimalista e della biolinguistica e implica una visione del tutto diversa della facoltà umana del linguaggio (ad es. sostiene l'assenza di ridondanza nella lingua).

etc. Va infatti subito notato che la CxM riconosce l'utilità delle categorie lessicali nella rappresentazione delle Costruzioni morfologiche, a differenza degli approcci più radicali della CxG, come la *Radical Construction Grammar* (Croft 2001), secondo cui le categorie lessicali e grammaticali sarebbero inadatte a rappresentare la diversità linguistica e quindi non considerate primitivi linguistici.<sup>31</sup> D'altra parte, come osserva Masini (2016: 30), «talvolta sembra essere proprio l'informazione categoriale a determinare l'esatta semantica di un processo di formazione di parola». Nella rappresentazione dei processi di formazione di parola proposta dalla CxM, si fa quindi generalmente riferimento all'informazione categoriale sia delle unità di input sia di quelle di output.<sup>32</sup> Tale posizione non è però accettata da tutti gli autori: Plag (2004), ad esempio, propone di non considerare l'informazione categoriale delle unità di input, ma limitarsi a quella delle unità di output, sostenendo che questa possa portare a formulare errate previsioni su pattern fonologici e semantici nel campo della morfologia derivazionale.

Assumendo la struttura della Costruzione proposta dalla CxG (figura 1), la CxM descrive le proprietà delle Costruzioni morfologiche su più livelli. Prendendo ad esempio la parola composta dell'italiano *portafoglio*, la sua struttura interna può essere resa con la seguente rappresentazione (figura 7):

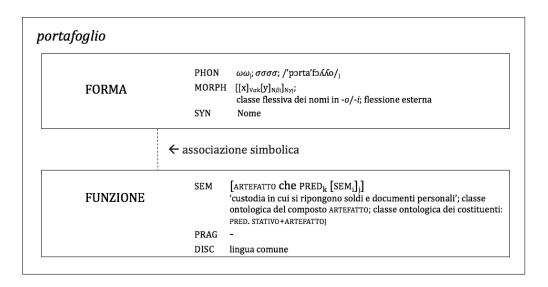


Figura 7. Struttura interna di portafoglio

<sup>&</sup>lt;sup>31</sup> L'unico primitivo linguistico riconosciuto nella RCxG è la Costruzione: «Constructions, not categories or relations, are the basic, primitive units of syntactic representation» (Croft 2001: 46).

<sup>&</sup>lt;sup>32</sup> L'utilità di considerare anche l'informazione relativa alle categorie di input è invece sostenuta da Booij (2015: 189).

Come mostrato dalla figura 7, il livello PHON include la rappresentazione fonetica della parola composta (/'pɔrta'fɔʎʎo/<sub>i</sub>) e il suo status di parola fonologica ( $\omega\omega_i$ );<sup>33</sup> un'altra proprietà analizzabile su questo livello è la struttura sillabica della parola: in questo caso, portafoglio è costituita da quattro sillabe ( $\sigma\sigma\sigma\sigma$ ).

Il livello MORPH rende conto delle categorie di input del composto (Verbo e Nome), e di quella di output (Nome); esso fornisce informazioni rispetto alla classe flessiva (nomi in -o/-i, secondo la classificazione proposta in Thornton 2001) e al tipo di plurale (flessione esterna; cfr. Micheli 2016a,b sugli altri tipi di flessione delle parole composte italiane). Tutto ciò che concerne il significato della parola è codificato nel livello SEM: il significato letterale della Costruzione (ed eventuali significati metaforici), il tipo semantico dell'intero composto e quello dei costituenti.<sup>34</sup> Nel caso di *portafoglio*, si tratta di un composto esocentrico che designa un nome di strumento (tipo ontologico ARTEFATTO): la semantica della Costruzione può essere rappresentata come [ARTEFATTO che PREDk SEMi]i, in cui PREDk è l'elemento verbale (in questo caso un verbo stativo, 'contenere') e SEMi il sostantivo che funge da secondo costituente, il cui tipo semantico è ARTEFATTO.

Le eventuali informazioni di carattere pragmatico relative alla Costruzione sono codificate nel livello PRAG: a differenza del composto portafoglio che non mostra particolari proprietà pragmatiche, Costruzioni morfologiche come i valutativi possono veicolare proprietà morfo-pragmatiche, come osservato già da Dressler & Merlini Barbaresi (1994: 55), in quanto includono «a pragmatic variable which cannot be suppressed in the description of [their] meaning».

Infine, il livello DISC rende conto delle proprietà legate alla dimensione testuale e discorsiva (ad es. registro, genere testuale, dominio, etc.): per quanto riguarda portafoglio, il composto occorre nella lingua comune senza restrizioni relative al registro o al genere testuale. Diverso è il caso da un particolare gruppo di composti italiani N-N

<sup>&</sup>lt;sup>33</sup> Si noti che nel caso dei composti italiani, a differenza ad esempio di quelli inglesi (cfr. Bauer 2017: 7-9), ciascun costituente conserva il proprio accento e quindi lo status di parola fonologica (Bisetto 2004: 34): un composto è quindi costituito da due parole fonologiche.

<sup>&</sup>lt;sup>34</sup> Come verrà discusso più ampiamente al § 3.2.2, l'analisi semantica delle Costruzioni verrà arricchita attraverso l'uso dei tipi semantici, sul modello di quanto proposto da Bouillon et al. (2012). Il tagset di tipi semantici utilizzati in questo lavoro verrà descritto al § 3.3.2.

(il tipo *emergenza rifiuti, ritiro merci*, etc.) analizzato in Baroni, Guevara & Zamparelli (2009) e tipico di una particolare tipologia testuale definita dagli autori "titolese".<sup>35</sup>

La struttura delle Costruzioni proposta dalla CxG mostra alcuni significativi punti di contatto con il modello *Tripartite Architecture of the Grammar* o *Parallel Architecture*, elaborato da Ray Jackendoff (1997, 2002, 2013).<sup>36</sup> Il modello proposto da Jackendoff prevede che ogni unità della lingua (chiamata *lexical item* e paragonabile alla Costruzione della CxG) presenti una struttura a tre livelli: fonologico (PHON), morfosintattico (MORPH/SYN) e semantico (SEM). Ciascun livello è parallelo e dotato di propri princìpi: insieme, essi forniscono simultaneamente informazioni sulle proprietà dei *lexical item* e permettono di rendere conto di fenomeni a metà strada tra i diversi livelli di analisi, ad esempio tra morfologia e sintassi. Pur differenziandosi in parte nella rappresentazione della struttura dell'unità linguistica (nella CxM essa è bipartita; nella *Parallel Architecture*, tripartita),<sup>37</sup> i due modelli sono accomunati dalla nozione di *continuum lessico-grammatica*.

In maniera più sintetica, una Costruzione morfologica può anche essere rappresentata come esemplificato in (2):<sup>38</sup>

(2) a. 
$$\langle [[x]_{V\alpha k}[y]_{N\beta i}]_{N\gamma j} \longleftrightarrow [SEM_k \text{ in } R_{SUB} \text{ con } SEM_i]_j \rangle$$
b.  $\langle [[porta]_{Vk}[y]_{N\beta i}]_{N\gamma j} \longleftrightarrow [SEM_k \text{ in } R_{SUB} \text{ con } SEM_i]_j \rangle$ 

$$| \beta = \text{cat. Nome, tipo semantico } ARTEFATTO$$

$$| \gamma = \text{cat. Nome}|Aggettivo, genere M, tipo semantico } ARTEFATTO$$
c.  $\langle [[porta]_{Vk}[foglio]_{Ni}]_{N\gamma j} \longleftrightarrow [SEM_k \text{ in } R_{SUB} \text{ con } SEM_i]_j \rangle$ 

-

<sup>&</sup>lt;sup>35</sup> Un caso simile è rappresentato dalle Costruzioni prive di soggetto come *Contains alcohol*, analizzate da Ruppenhofer & Michaelis (2010) come esempi di *labelese*.

<sup>&</sup>lt;sup>36</sup> In particolare, in Booij & Audring (2016) si discutono i legami tra CxM e *Architettura Parallela*, con particolare riferimento alla formazione delle parole.

<sup>&</sup>lt;sup>37</sup> Altre differenze tra i due modelli sono discusse in Masini (2016: 37-38).

 $<sup>^{38}</sup>$  Seguendo Masini & Audring (in stampa), in questo lavoro si utilizzeranno le seguenti norme nella notazione delle Costruzioni morfologiche: 1. lettere maiuscole per indicare le categorie lessicali: in particolare, A=Aggettivo; V=Verbo, AVV= Avverbio, N=Nome, P=Preposizione; 2. lettere minuscole in pedice per gli indici: ad es. < i >, < j >, < k >; 3. lettere dell'alfabeto greco in pedice per i *constraints* delle variabili: ad es. <  $\alpha$  >, <  $\beta$  >, <  $\gamma$  >, etc.; 4. lettere non in pedice minuscole (ad es. < x, y, z > per materiale fonologico non specificato, maiuscole (ad es. < X, Y, Z >) per categorie lessicali non specificate; 5. corsivo per il materiale fonologico specificato (ad es. *porta-, -tore*, etc.); 6. maiuscolo per gli operatori semantici (ad es. la negazione NON), il significato della Costruzione o il tipo semantico (ad es. STRUMENTO, etc.).

Gli schemi costruzionali esemplificati in (2) sono rappresentati secondo il formalismo più generalmente adottato negli studi di CxM e che si utilizzerà anche in questo lavoro. Ciascuno schema è racchiuso da parentesi uncinate (<) ed è costituito da due parti: a sinistra vengono rappresentate le proprietà formali della Costruzione (livelli PHON/MORPH/SYN), a destra quelle funzionali (livelli SEM/PRAG/DISC); le due frecce ( $\leftarrow$ ) indicano l'associazione convenzionale tra forma e significato. Le lettere greche  $<\alpha>$  e  $<\beta>$ , in pedice, indicano le restrizioni (constraints) delle variabili [x] e [y]: in questo caso, ad esempio,  $\beta$  indica le restrizioni che vincolano la scelta del secondo costituente, ossia la categoria Nome e il tipo ontologico ARTEFATTO. Le lettere <k>, <i>e <j>, in pedice, svolgono la funzione di marker di co-indicizzazione.

Ciò che distingue gli schemi in (2) è il grado di astrattezza/specificità: (2a) riproduce lo schema generale dei composti V-N dell'italiano che formano nomi di strumento (es. *portafoglio, lavastoviglie, portachiavi,* etc.); (2b) rappresenta uno schema semi-specificato, costituito da uno slot chiuso, la forma porta del verbo *portare* nel significato di 'contenere' e una variabile [y], soggetta a restrizioni formali e semantiche (ossia l'appartenenza alla classe dei Nomi e al tipo semantico ARTEFATTO); in (2c) entrambi gli slot sono specificati dai costituenti del composto *portafoglio*, che può essere definito una *istanziazione* dello schema in (2b).

Attraverso schemi come quelli riportati in (2), la CxM rende conto sia degli item concreti sia dei processi di formazione delle parole, tradizionalmente rappresentati dalle Regole di Formazione di Parola (RFP; cfr. Aronoff 1976). Le nozioni di schema costruzionale e RFP non sono tuttavia completamente sovrapponibili: entrambe esprimono la capacità generativa della lingua (Booij 2013: 257), ma solo le RFP hanno carattere procedurale e implicano di default produttività. Diversamente, la natura degli schemi costruzionali è primariamente dichiarativa: essi rappresentano generalizzazioni statiche (e non necessariamente produttive) formulate dai parlanti sulla base di un certo numero di item specifici incontrati nell'uso. Inoltre, mentre tradizionalmente le RFP sono ascritte al dominio della grammatica e le singole forme a quello del lessico, per la CxM gli schemi, i sotto-schemi e gli item specifici sono tutti strutturati secondo gli stessi principi e fanno parte di un'unica *rete*. Come si osserverà nel seguente paragrafo, tutte le Costruzioni di una lingua sono organizzate in Gerarchie di eredità, in cui schemi più astratti dominano sotto-schemi via via più specifici: tale rete di Costruzioni costituisce il Costruzionario.

### 2.3 Il Costruzionario

Il Costruzionario (*Constructicon*; cfr. Goldberg 2003: 219) raccoglie tutte le Costruzioni di una lingua, organizzate in reti (*constructional networks*): esso, infatti, non rappresenta una lista di forme (come il lessico nella concezione tradizionale bloomfieldiana) bensì una rete strutturata gerarchicamente, in grado di cogliere e rappresentare generalizzazioni (e sotto-generalizzazioni) a vari livelli di astrattezza.

Fanno quindi parte del Costruzionario sia le Costruzioni pienamente specificate (come portafoglio, stazione, mangiabile, etc.), sia gli schemi semi-specificati (ad es. i composti con porta- come primo costituente), sia gli schemi astratti (ad es. la Costruzione dei composti V-N). In linea con la nozione di *continuum* lessico-sintassi (cfr. Goldberg 2006: 5), il Costruzionario include infatti sia ciò che tradizionalmente viene chiamato "grammatica" sia le unità che tradizionalmente costituiscono il lessico di una lingua. Va notato che nella CxM, le nozioni di parola e unità lessicale non coincidono. Il parametro che definisce la 'parola morfologica' (word) è la coesione: Booij (2009: 97) osserva infatti che «If we take the notion word seriously, we might say that its defining property is coehsiveness or non-interruptability. In other words, the 'no manipulation' part of the principle of Lexical Integrity is the proper interpretation of word coehsiveness». La nozione di 'unità lessicale' (lexical unit) si sovrappone invece a quella di Costruzione. I concetti di parola e unità lessicale, quindi, possono non coincidere, dal momento che tutte le parole sono Costruzioni (pienamente specificate), ma non tutte le Costruzioni possono essere definite parole, ossia possono darsi Costruzioni che non soddisfano il parametro della coesione.

L'organizzazione interna del Costruzionario si basa su Relazioni di eredità (*Inheritance Links*); in particolare, Goldberg (1995: 75-81) ne individua quattro:

- 1. la Relazione (di eredità) di istanziazione (*Instance Link*);
- 2. la Relazione (di eredità) polisemica (Polysemy Link);
- 3. la Relazione (di eredità) metaforica (*Metaphorical Extension Link*);
- 4. la Relazione (di eredità) di sottoparte (Subpart Link).

Di seguito ci si concentra sul primo tipo di Relazione (cfr. Masini 2016: 54-58 per una più approfondita discussione degli altri tre tipi), che costituisce il tipo più sfruttato all'interno del Costruzionario, poiché rende conto del rapporto tra uno schema e le sue istanziazioni, ossia le sue (via via più concrete) realizzazioni: la Relazione di istanziazione può infatti dar vita a Gerarchie di eredità articolate in più livelli, come mostrato dalla figura 8.

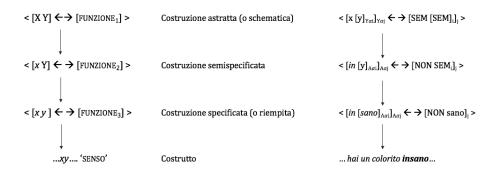


Figura 8. Costruzione della prefissazione in italiano: gerarchia di eredità<sup>39</sup> (*Fonte*: adattata da Masini 2016: 59)

A sinistra della figura 8 sono rappresentati i livelli di una Gerarchia di eredità, dalla Costruzione massimamente astratta al Costrutto, ossia la concreata realizzazione in contesto: il livello più astratto è instanziato da Costruzioni via via più specifiche che ne ereditano le proprietà. A destra, è invece esemplificato un caso particolare, ossia la Costruzione astratta che rende conto del processo di prefissazione in italiano (dove x = prefisso): essa è instanziata da una Costruzione semi-specificata (tra le tante disponibili in italiano) costituita dal prefisso negativo *in*- e una base aggettivale, a sua volta istanziata da una Costruzione specificata costituita da *in*- e l'aggettivo *sano*.

-

 $<sup>^{39}</sup>$  Seguendo Masini (2016: 59), le lettere maiuscole (X e Y) vengono utilizzate per rappresentare le variabili, le minuscole (x e y) per il materiale lessicale; i numeri associati alle funzioni indicano che il significato della Costruzione è via via più specifico; i puntini di sospensione segnalano che il Costrutto occorre nell'uso reale, all'interno di un enunciato.

<sup>&</sup>lt;sup>40</sup> Si noti che, in alcuni casi, singole parole morfologicamente complesse possono presentare proprietà non predicibili da quelle dello schema che esse istanziano: come osservano Masini & Audring (in stampa), infatti, «The properties of higher-level constructions may be *overridden* whenever a more specific (possibly contradictory) property is encountered in a lower-level construction. This produces constructions with idiosyncrasies». Cfr. Booij (2010: 27) per alcuni esempi dall'olandese.

La creazione di un Costrutto avviene riempiendo le variabili della Costruzione con parole (anch'esse considerabili Costruzioni) che soddisfano determinate restrizioni formali o semantiche (ad es., nel caso esemplificato nella figura 8, la variabile Y deve essere un aggettivo). Il meccanismo che la CxM postula per rendere conto dell'unione di più Costruzioni è chiamato *Unificazione*: così, ad esempio, l'aggettivo *sano* si unisce alla Costruzione semi-specificata dei prefissati con *in*- e dà luogo al Costrutto *insano*.

Il meccanismo di Unificazione può essere usato in due casi: per riempire uno slot aperto con materiale lessicale; per unire due schemi astratti. Il primo caso, appena osservato nell'esempio di *insano*, si ha anche nel momento in cui i parlanti creano nuove parole sul momento, occasionalismi, che non fanno parte stabilmente del lessico della lingua. Il secondo caso è fondamentale per rendere conto di casi problematici in cui processi morfologici si applicano a parole possibili ma non esistenti, come gli aggettivi inglesi del tipo *un-V-able*, es. *unbeatable* 'imbattibile', analizzati in Booij (2010a: 42) o come, in italiano, un gruppo di verbi parasintetici (Masini & Iacobini, in stampa) del tipo *de-V-izzare* (come *derattizzare*). Questo tipo di verbi è apparentemente problematico, perché presuppone una fase di derivazione in cui il prefisso *de-* si unisce a un verbo possibile ma non esistente (come \**rattizzare*): la CxM rende conto di queste forme postulando l'unificazione simultanea di due schemi, quello della prefissazione con *de-* e quello della suffissazione con *-izzare*, come illustrato nella figura 9.

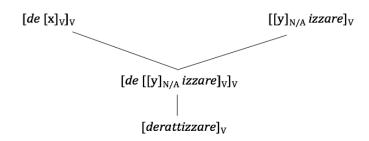


Figura 9. Esempio di Unificazione tra due schemi che produce un nuovo schema

L'Unificazione rappresentata nella figura 9 porta alla formazione di un nuovo schema, la cui produttività è indipendente da quella dei due schemi unificati.

La produttività costituisce un aspetto a cui la CxM, in quanto teoria *usage-based*, rivolge particolare attenzione. Le Costruzioni di cui si compone il Costruzionario sono infatti

dotate di diversi livelli di astrattezza e produttività: due proprietà che non sempre vanno di pari passo. L'esistenza di schemi semi-specificati, infatti, non implica necessariamente che tale schema sia produttivo: dal momento che gli schemi costituiscono generalizzazioni statiche su un set di Costrutti, essi possono essere non produttivi, ossia non più utilizzati dai parlanti per la creazione di nuove forme. D'altra parte, dal punto di vista diacronico, uno schema che ha smesso di essere produttivo può essere riattivato e dar luogo a nuove parole dopo una fase di 'inattività'.

La produttività di una Costruzione morfologica può quindi dipendere da più fattori, tra i quali l'astrattezza rispetto alle Gerarchie, le restrizioni che operano sulle variabili, la disponibilità delle basi, esigenze comunicative, la presenza di eventuali schemi in competizione. Ulteriori fattori da considerare sono il grado di radicamento (entrenchment) della Costruzione nella lingua e la sua frequenza (type frequency). Per rendere conto dei diversi gradi di produttività degli schemi, Barðdal ha proposto una scala di produttività (cline of productivity; cfr. Barðdal 2008: 172) basata sul grado di radicamento della Costruzione nella lingua.

Infine, va notato che in CxM la produttività non implica necessariamente la composizionalità: come si osserverà meglio al § 2.5), nei cosiddetti "idiom costruzionali" (constructional idiom; cfr. Jackendoff 1990; Booij 2002a), elementi che originariamente occorrono come costituenti di composto possono sviluppare un significato più astratto e non predicibile, e dar luogo a una Costruzione semi-specificata (o sotto-schema) produttiva, formata da un elemento idiomatico e lessicalmente specificato e una variabile. Un esempio è il nome olandese reuze (da reus 'gigante' con un linking element –e) quando costituisce il modificatore di un composto N-A (Masini & Audring, in stampa):

- (3) a. reuze+A: reuze-aardig 'molto gentile', reuze-leuk 'molto carino', etc.
  - b.  $\langle [[reuze]_{Nk} [y]_{Ai}]_{Aj} \leftarrow \rightarrow [MOLTO SEM_i]_j$

Nella Costruzione in (3), *reuze* esprime un valore di intensificazione, più astratto rispetto all'originario significato di 'gigante'. Per predire il significato di forme come *reuze-aardig* o *reuze-leuk* occorre quindi conoscere sia il significato dell'intera Costruzione sia quello delle singole parti: per la CxM, infatti, ogni Costruzione presenta infatti specifiche proprietà olistiche.

Un altro esempio di non composizionalità che può essere interpretato come una proprietà olistica di una Costruzione è l'esocentricità (cfr. Booij 2010: 36-39; sul concetto di esocentricità cfr. Bauer 2017: 64-70 e Scalise & Guevara 2006): oltre al già osservato caso dei composti V-N, un altro esempio in ambito italiano sono i composti V-V del tipo *fuggi-fuggi*, il cui schema è riportato in (4):<sup>41</sup>

- (4) a.  $\langle [[x_i]_{Vk} [x_i]_{Nj} \leftrightarrow [AZIONE \text{ di } [SEM]_k \text{ svolta da GRUPPO DI PERSONE}]_j \rangle$  (con x = radice verbale, omofona della forma di imperativo)
  - b. *fuggi-fuggi, corri-corri,* etc.

Il significato di 'azione svolta da un gruppo di persone', così come alcune proprietà morfologiche (ad es. la categoria di output), delle forme in (4b) sono proprietà olistiche inscritte nella Costruzione e non predicibili dai singoli costituenti.

Un altro strumento di cui si serve la CxM per rappresentare le relazioni tra schemi costruzionali sono le Relazioni Paradigmatiche,<sup>42</sup> che legano due set di parole<sup>43</sup> (e i relativi schemi) con lo stesso grado di complessità e che derivano dalla stessa base. Attraverso questo tipo di Relazioni, la CxM rende conto di fenomeni come il troncamento (cfr. Booij 2010b: 241 sulla formazione degli ipocoristici in olandese) o la sostituzione di affissi. In particolare, un caso di sostituzione di affisso in italiano è esemplificato dalle forme in (5):

(5) socialismo ~ socialista illuminismo ~ illuminista alpinismo ~ alpinista

Le coppie di parole derivate in (5) sono chiaramente legate semanticamente: il suffisso –ismo è usato per creare sostantivi che indicano un movimento politico-culturale o una

\_

<sup>&</sup>lt;sup>41</sup> Per un'analisi approfondita di questo tipo di forme si veda Masini & Thornton (2007); Thornton (2009, 2008).

<sup>&</sup>lt;sup>42</sup> Nell'ambito degli studi di CxM, il termine *relazione paradigmatica* assume il preciso significato discusso in questo paragrafo, discostandosi dal tradizionale significato saussuriano di relazione *in absentia*.

<sup>&</sup>lt;sup>43</sup> Un set di parole che condivide lo stesso affisso costituisce una "famiglia morfologica" (*morphological family* o *word family*), la cui ampiezza incide sul riconoscimento delle parole morfologicamente complesse in task di decisione lessicale (cfr. Schreuder & Baayen 1997 e De Jong, Schreuder & Baayen 2000 sul cosiddetto *family size effect*).

pratica, da cui si formano i derivati in −ista che denotano le persone che aderiscono a (o in qualche modo sono coinvolti con) tali movimenti/pratiche. La formazione dei derivati in −ista non avviene attraverso l'aggiunta del suffisso in maniera concatenativa (quindi non si hanno forme come \*illuminismista) ma tramite la sostituzione del suffisso. Per risolvere questa discrasia tra forma e significato, il modello della Morfologia lessicale (cfr. Aronoff 1976) ha proposto una regola di troncamento, per cui [illumin-ismo-ista] > [illumin-ismo-ista] > [illumin-ista]. Nel quadro della CxM, l'anomalia di tali forme è risolta introducendo la nozione di Schema di Secondo Ordine (Second Order Schema; cfr. Masini & Booij 2015), ossia l'unione di due o più schemi tramite una Relazione Paradigmatica (indicata dal simbolo ≈), come riportato in (6):

(6) 
$$\langle [x-ismo]_{N_i} \leftrightarrow [SEM_i] \rangle \approx \langle [x-ista]_{N_i} \leftrightarrow [PERSONA coinvolta in [SEM_i]]_i \rangle$$

Lo schema in (6), adattato da Booij (2013: 264), ha il vantaggio di non implicare una direzione nella derivazione: la formazione di una nuova forma si può quindi avere in entrambe le direzioni; ciò che è rilevante per la CxM è che le parole in *-ista* siano semanticamente legate a quelle in *-ismo* e non alla loro base (es. *socialista* è infatti legato a *socialismo* e non a *sociale*).

Come si può notare dagli esempi analizzati in questo paragrafo, gli strumenti elaborati dalla CxM per la rendere conto delle Costruzioni morfologiche presenti nel Costruzionario sono stati applicati prevalentemente alla formazione delle parole: nel seguente paragrafo si propone una panoramica strutturata di come tale modello rappresenta i processi con cui si formano i diversi tipi di parole, focalizzando l'attenzione in particolare sull'italiano.

## 2.4 La formazione delle parole

La formazione delle parole è il dominio a cui la CxM ha dedicato maggiore attenzione.<sup>44</sup> Nei precedenti paragrafi si è osservato che, attraverso la nozione di Costruzione, la CxM rappresenta sia le singole forme sia i processi di formazione delle parole: il Costruzionario è infatti costituito sia dai singoli Costrutti sia dagli schemi (e dai sotto-

\_

<sup>&</sup>lt;sup>44</sup> In particolare, per una panoramica generale sui processi di formazione delle parole nel quadro della CxM si veda Booij (2015).

schemi), i quali danno luogo alle Gerarchie di eredità; queste possono essere *aperte* o *chiuse*, a seconda che si tratti di un processo produttivo o no. Di seguito si fornisce una panoramica degli strumenti che la CxM fornisce per rappresentare i diversi processi (produttivi e non produttivi) di formazione di parola.

#### 2.4.1 Derivazione

Un esempio di Costruzione *aperta* che in italiano permette di formare nuove parole derivate è quella degli aggettivi deverbali in *-bile* (ad es. *leggibile*, *adattabile*, *bevibile*, etc.): la Costruzione è attualmente produttiva, come testimoniano i neologismi creati sulla base di questo schema (ad es. *taggabile*, *esodabile*, *hackerabile*). La Gerarchia di eredità di tale Costruzione è rappresentata nella figura 10.

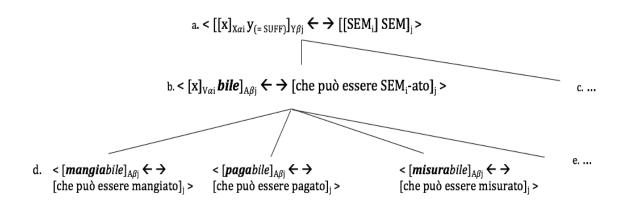


Figura 10. Gerarchia costruzionale per la suffissazione in *-bile* in italiano *Fonte*: adattato da Masini (2016: 59)

La Costruzione più astratta, rappresentata in (a), codifica il processo generale della suffissazione in italiano: x e y sono due variabili, dove y è un suffisso ed è quindi rappresentato fuori dalle prime parentesi quadre.<sup>45</sup> Dal momento che in italiano la suffissazione (contrariamente alla prefissazione) può cambiare la categoria della base,<sup>46</sup>

<sup>46</sup> In italiano, la suffissazione può anche non prevedere il cambiamento di categoria della base: si pensi ad esempio ai nomi in *–ista* come *barista* < *bar* o in *–aio* come *fioraio* < *fiore* e a tutti i derivati valutativi (cfr. Booij 2010: 29).

 $<sup>^{45}</sup>$  Si noti che, trattandosi di un affisso e non di una parola, y non ha indici: il suo contributo semantico è indicato da SEM, anch'esso non indicizzato.

nello schema in (a) vengono utilizzati due indici diversi: X rappresenta la categoria della base, Y quella dell'intera Costruzione. Tale schema domina Costruzioni meno astratte in cui i singoli suffissi sono specificati (ad es. la Costruzione con – zione o quelle con i suffissi accrescitivi – one o diminutivi – ino, – etto, etc.): il caso esemplificato in (b) è quello degli aggettivi derivati da una base verbale attraverso il suffisso – bile. In (d) sono rappresentati alcuni esempi di derivati in – bile: i puntini in (e) indicano che la Costruzione è tuttora produttiva e può instanziare nuove Costruzioni.

Un esempio di Costruzione derivata *chiusa* dell'italiano è costituito dai nomi d'azione in –*ore*, come *bruciore*, *dolore*, *amore*: si tratta di un insieme limitato a pochi sostantivi deverbali che in genere denotano stati fisici o psichici. Lo schema di questo tipo di forme può essere rappresentato come segue:

(7) 
$$< [[x]_{V\alpha i} ore]_{N\beta j} \leftarrow \rightarrow [STATO \ relativo \ a \ SEM_i]_j >$$

$$|\alpha = \text{classe flessiva del verbo } (-are, -ere, -ire)$$

$$|\beta = \text{genere maschile, classe flessiva } -e/i^{47}, \text{ tipo semantico STATO}$$

Lo schema in (7) afferma che la Costruzione è costituita da un'unità di input verbale (V) dotata di proprietà/constraints  $\alpha$ ; il suo output è un nome (N) con proprietà/constraints  $\beta$ . Dal punto di vista semantico, la Costruzione crea nomi indicanti uno STATO legato all'evento espresso dal verbo (indicato con SEM<sub>i</sub>). A differenza della Costruzione con il suffisso –bile, questo schema è instanziato soltanto da un numero limitato di sostantivi (cfr. Gaeta 2004: 350), tra i quali dolore, bruciore, amore, sentore, etc.: in questo caso, quindi, lo schema si limita a esprimere una generalizzazione sulla base di un set di parole esistenti nella lingua, ma non costituisce un modello per la formazione di nuove forme.

### 2.4.2 Composizione

L'analisi della composizione italiana con gli strumenti della CxM è l'obiettivo di questo lavoro, con cui si intende colmare una lacuna nella letteratura costruzionista sulla composizione nelle lingue romanze, e in particolare sull'italiano. Gli studi dedicati alla formazione delle parole composte nel quadro della CxM si sono infatti finora concentrati

<sup>&</sup>lt;sup>47</sup> Cfr. Thornton (2001).

prevalentemente sulle lingue germaniche.<sup>48</sup> Di seguito si propone una panoramica degli strumenti elaborati finora dalla CxM per la rappresentazione delle parole composte, rimandando al § 3.2 per una trattazione più approfondita.

Al pari della derivazione e degli altri processi morfologici, anche la composizione può essere espressa in termini di Costruzioni e schemi;<sup>49</sup> ad esempio, lo schema generale dei composti con testa a sinistra dell'italiano contemporaneo può essere rappresentato come segue:

(8) 
$$< [[x]_{X\alpha k}[y]_{Y\beta i}]_{X\alpha j} \leftarrow \rightarrow [SEM_k \text{ in } R \text{ con } SEM_i]_j >$$

Nello schema in (8) le proprietà espresse da X e  $\alpha$  sono condivise dal primo costituente, ossia l'elemento testa, e dall'intera Costruzione; i due costituenti sono legati da una relazione generica espressa dalla variabile R. Tale schema è instanziato da una serie di Costruzioni più specificate (o *sotto-schemi*): in italiano esso dà luogo alla Gerarchia di eredità illustrata nella figura 11.

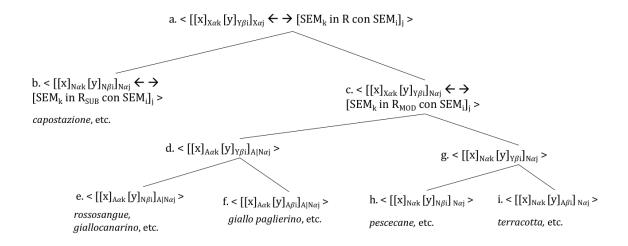


Figura 11. Gerarchia di eredità dei composti con testa a sinistra dell'italiano contemporaneo

-

<sup>&</sup>lt;sup>48</sup> Cfr. Hüning & Booij (2014); Booij (2010c, 2005a); su alcune tipologie di composto del francese si veda Van Goethem (2009); Van Goethem & Hüning (2014) e Van Goethem & Amiot (2011).

 $<sup>^{49}</sup>$  D'ora in poi, si farà riferimento alle Costruzioni morfologiche che istanziano parole composte come a  ${\rm Costruzioni_{COMP}}.$ 

Lo schema in (a), che verrà analizzato più nel dettaglio al § 4.3.1, è istanziato da due Costruzioni<sub>COMP</sub> a seconda della relazione R che intercorre tra i due costituenti: in particolare, in (b) è rappresentato lo schema dei composti subordinativi (R<sub>SUB</sub>) e in (c) quello dei composti con relazione di modificazione (R<sub>MOD</sub>).<sup>50</sup> Quest'ultimo è a sua volta sotto-specificato da ulteriori sotto-schemi a seconda della categoria lessicale dei costituenti, che in italiano possono essere A-A (es. *giallo paglierino*), A-N (es. *rosso sangue*), N-N (es. *pesce spada*) e N-A (es. *roccaforte*). Come si osserverà meglio al § 4.3.1, questi sotto-schemi sono dotati di diversi livelli di produttività: si noti per ora che le Costruzioni<sub>COMP</sub> rappresentate in (e) e (f) istanziano esclusivamente termini di colore e sono meno produttive delle altre.

Attraverso la Gerarchia di eredità rappresentata nella figura 11 si è osservato come possono essere rappresentati in termini di schemi e sotto-schemi i composti dell'italiano contemporaneo che presentano l'elemento testa a sinistra. La CxM permette inoltre di rappresentare anche i composti in cui generalmente si individuano due teste,<sup>51</sup> come i composti coordinativi del tipo *ristorante-pizzeria*, *nave-traghetto*, *divano-letto*, etc.<sup>52</sup> Dal punto di vista semantico, queste forme vengono tradizionalmente considerate come 'bicefale', dal momento che identificano entità che mostrano proprietà sia del primo sia del secondo costituente (ad es. un *divano-letto* è un artefatto che può fungere sia da divano sia da letto). Dal punto di vista formale, però, le proprietà morfologiche della Costruzione<sub>COMP</sub> vengono ereditate dall'elemento di sinistra: in *nave-traghetto*, ad esempio, il genere femminile del composto è ereditato dal primo costituente (quindi *la nave-traghetto* e non \**il nave-traghetto*); inoltre, la flessione plurale coinvolge soltanto il sostantivo *nave* (quindi *le navi-traghetto* e non \**le navi-traghetti*).<sup>53</sup> Questo tipo di composti viene rappresentato in CxM attraverso lo schema in (9) (adattato da Fábregas & Masini 2015: 93).

\_

<sup>&</sup>lt;sup>50</sup> Come verrà illustrato più nel dettaglio nel prossimo capitolo, la relazione di modificazione (R<sub>MOD</sub>) include sia la relazione attributiva (es. *camposanto*) sia quella appositiva (es. *pesce palla*), cfr. Arcodia, Grandi & Montermini (2009).

<sup>&</sup>lt;sup>51</sup> La nozione di testa costituisce un concetto ancora ampiamente dibattuto in morfologia (cfr. Bauer 2017: 29-40); sulla posizione della CxM si vedano Fábregas & Masini (2015: 91-94) e Arcodia (2012).

<sup>&</sup>lt;sup>52</sup> Cfr. Arcodia, Grandi & Wälchli (2010); Arcodia (2018); Grandi (2011) sui composti coordinativi in prospettiva tipologica.

<sup>&</sup>lt;sup>53</sup> Va però notato che alcuni composti coordinativi designanti referenti umani formano il plurale con una doppia flessione (ad es. *studenti-lavoratori*, *attori-cantanti*, etc. cfr. Radimsky 2015a).

# (9) $< [[x]_{N\alpha k}[y]_{N\beta i}]_{N\alpha j} \leftarrow \rightarrow [ENTITA che è sia SEM_k sia SEM_i]_i >$

La Costruzione<sub>COMP</sub> in (9) rende conto dei composti come *nave-traghetto* o *ristorante- pizzeria* senza ricorrere a una nozione strutturale di 'testa': attraverso la coindicizzazione (segnalata negli esempi con le lettere i e j in pedice) è rappresentata la
corrispondenza tra le proprietà formali del primo costituente e dell'intera
Costruzione<sub>COMP</sub> (N $\alpha$ ); le proprietà semantiche della Costruzione sono espresse nella
parte destra dello schema.

## 2.4.3 Tra morfologia e sintassi: le espressioni multiparola

Come già osservato più volte nel corso di questo capitolo, la CxM fornisce gli strumenti per rendere conto di fenomeni tradizionalmente considerati a metà strada tra i diversi moduli della lingua (es. morfologia/sintassi, composizione/derivazione). Un filone di ricerca a cui la CxM ha dedicato grande attenzione fin da subito è infatti quello delle *espressioni multiparola*, etichetta che include diversi tipi di fenomeni posti lungo il *continuum* lessico-sintassi (ad es. espressioni idiomatiche, proverbi, collocazioni, parole sintagmatiche, etc.). Tra i lavori di ambito costruzionista dedicati alle espressioni multiparola vi sono gli studi dedicati ai verbi con particella dell'olandese (Booij 2002ab) e dell'italiano (Iacobini & Masini 2007; Masini 2012) e ai nomi sintagmatici (cfr. Booij 2009 sui *phrasal nouns* del greco e dell'olandese; Masini 2012 su quelli dell'italiano; Masini & Benigni 2012, Benigni & Masini 2010 su quelli del russo).<sup>54</sup>

In italiano i nomi sintagmatici sono un fenomeno tutt'altro che raro: in (10a) si riporta lo schema generale di quelli che presentano la struttura Nome Preposizione Nome (N-P-N; ad es. carta di credito, battello a vapore, carta da zucchero, etc.). Tale schema domina un cospicuo numero di sotto-schemi con proprietà più specificate (tipo di preposizione, semantica dei costituenti, etc.), tra cui quello rappresentato in (10b), in cui la preposizione di è specificata e il secondo nome è un nome di evento: in questo caso la relazione che lega i due sostantivi è di natura telica.<sup>55</sup>

<sup>-</sup>

<sup>&</sup>lt;sup>54</sup> In ambito costruzionista ci si riferisce a questo tipo di Costruzioni con il termine *lexical phrases* (cfr. Booij 2009b), *phrasal lexemes* o *parole sintagmatiche* (cfr. Masini 2009; 2012; Simone 1997 per il primo studio dedicato all'italiano in cui si propone la nozione di 'verbo sintagmatico'). Grossomodo la nozione di 'parola sintagmatica' coincide con quella di *polirematica* (cfr. Voghera 2004)

<sup>&</sup>lt;sup>55</sup> La relazione telica descrive eventi che tendono verso un fine. La regolarità tra relazione telica e nome eventivo è stata osservata per la prima volta da Johnston e Busa (1999) e più recentemente da Bouillon *et* 

(10) a. 
$$\langle [x]_{Ni} [[y]_P [z]_{Nj}]_{P'}]_{Nk} \leftarrow \rightarrow [SEM_i \text{ in } R \text{ con } SEM_j]_k \rangle$$
  
b.  $\langle [x]_{Ni} [[di]_P [z]_{N\_EVENTOj}]_{P'}]_{Nk} \leftarrow \rightarrow [SEM_i \text{ con } scopo SEM_j]_k \rangle^{56}$ 

Se nel caso dei nomi sintagmatici N-P-N è più agevole tracciare un confine con i composti, data la presenza di una marca relazionale esplicita (la preposizione), meno netto appare il confine tra nomi sintagmatici con struttura N-A, come tessera sanitaria o anima gemella e forme tradizionalmente considerate parole composte (come cassaforte, camposanto).<sup>57</sup> In ottica costruzionista, i due tipi Costruzioni si collocano lungo il continuum lessicosintassi a diversi gradi di complessità, come illustrato nella figura 12, in cui si propone una rappresentazione dei diversi livelli.



Figura 12. Parole composte e parole sintagmatiche nel continuum lessico-discorso

Nonostante siano adiacenti nel continuum lessico-discorso, parole composte e parole sintagmatiche costituiscono due meccanismi distinti in CxM,<sup>58</sup> entrambi sfruttati per la denominazione di referenti stabili (contrariamente ai sintagmi liberi, la cui funzione è primariamente descrittiva) e spesso in competizione, come mostrato da Masini (in stampa) attraverso il caso delle parole composte e sintagmatiche aventi come testa il sostantivo capo (ad es. capostazione vs capo del governo). La competizione tra Costruzioni strutturalmente diverse è d'altra parte in linea con una concezione ridondante del Costruzionario, in cui sono memorizzate parole semplici, parole composte e parole sintagmatiche, insieme ai relativi schemi e sotto-schemi.

distinguere le due categorie.

al. (2012): in particolare, gli autori propongono un'analisi semantica delle preposizioni che occorrono nei nominali complessi N-P-N delle lingue romanze nel quadro del Lessico Generativo (cfr. Pustejovsky 1995).

<sup>&</sup>lt;sup>56</sup> Fonte: Masini (2016: 91). <sup>57</sup> Tale difficoltà a tracciare un confine tra composti e nomi sintagmatici N-A vale per l'italiano ma, ad esempio, non per russo e olandese, lingue nelle quali l'assenza di accordo tra i costituenti nelle parole

composte permette di distinguerli dai nomi sintagmatici (cfr. Masini & Benigni 2012; Booij 2009b). <sup>58</sup> Tale posizione è sostenuta anche in Masini & Scalise (2012: 75), a cui si rimanda per i criteri adottati per

#### 2.5 Il mutamento costruzionale e la formazione di affissoidi

Un filone di ricerca costruzionista che negli ultimi decenni ha prodotto numerosi studi è quello che indaga come si formano nuove Costruzioni nel corso del tempo e come queste evolvono nella storia delle diverse lingue. Gran parte del presente lavoro si inserisce in questo filone, in quanto si propone di analizzare la composizione italiana assumendo una prospettiva diacronica. Di seguito si traccia una panoramica degli strumenti con cui la CxM rende conto del mutamento linguistico a livello di Costruzione morfologica, focalizzando l'attenzione in particolare sulle parole composte.

La teoria del mutamento con cui il modello costruzionista ha fin da subito dialogato è la teoria della grammaticalizzazione (cfr. Lehmann 1985; Hopper & Traugott 2003): già Hopper & Traugott (1993), e in seguito Traugott (2003), avevano messo in luce l'importanza della nozione di costruzione nella spiegazione del mutamento linguistico.<sup>59</sup> Tale dialogo si è intensificato negli ultimi decenni (cfr. Giacalone Ramat 2015) e ha dato vita a una teoria del mutamento linguistico su base costruzionista (cfr. Traugott & Trousdale 2010, 2013), in cui gli assunti e gli strumenti della CxG vengono applicati ai fenomeni di mutamento e variazione.

In primis, viene ripresa dalla CxG la nozione di *continuum* lessico-grammatica che, dal punto di vista del mutamento porta a considerare *lessicalizzazione*<sup>60</sup> e *grammaticalizzazione* (cfr. Traugott 2003: 645) come due processi speculari, entrambi inclusi nella nozione di *Costruzionalizzazione*, che Traugott & Trousdale (2013: 22) definiscono così:

Constructionalization is the creation of form $_{new}$ -meaning $_{new}$  (combination of) signs. It forms new type nodes, which have new syntax or morphology and

<sup>&</sup>lt;sup>59</sup> Come osserva Trousdale (2014: 559), in questi studi si utilizza il termine costruzione (*construction*) per intendere il contesto morfosintattico e semantico-pragmatico («the surrounding context in which a particular expression comes to have a more grammatical function»), non ancora quindi nella definizione specifica elaborata dalla CxG.

<sup>&</sup>lt;sup>60</sup> Il termine 'lessicalizzazione' è usato in letteratura per intendere processi diversi (cfr. Ježek 2016: 5-7 per la discussione delle diverse interpretazioni): in ottica costruzionista, per *lessicalizzazione* si intende «[...] the change whereby in certain linguistic contexts speakers use a syntactic construction or word formation as a new contentful form with formal and semantic properties that are not completely derivable or predictable from the constituents of the construction or the word formation pattern. Over time there may be further loss of internal constituency and the item may become more lexical» (Brinton & Traugott 2005: 96).

new coded meaning, in the linguistic network of a population of speakers. It is accompanied by changes in degree of schematicity, productivity, and compositionality. The constructionalization schema always results from a succession of micro-steps and is therefore gradual. [...] Minimally constructionalization involves neoanalysis of morphosyntactic form and semantic/pragmatic meaning; discourse and phonological changes may also be implicated at various stages. Formal changes alone, and meaning changes alone cannot constitute constructionalization.

La Costruzionalizzazione è quindi il processo che porta alla formazione di nuove Costruzioni, attraverso piccoli cambiamenti posti sia sul piano morfosintattico sia su quello semantico-pragmatico. Traugott & Trousdale (2013: 26) chiariscono infatti che un cambiamento solo sul piano formale o solo sul piano funzionale non può essere definito Costruzionalizzazione ma Cambiamento costruzionale (constructional change), ossia «a change affecting one internal dimension of a construction. It does not involve the creation of a new node». I Cambiamenti costruzionali possono precedere o seguire la Costruzionalizzazione: in particolare, esempi di pre-constructionalization constructional changes (d'ora in poi PreCxzn CCs) sono i casi di mismatch tra forma e funzione o di rianalisi semantica; i post-constructionalization constructional changes (d'ora in poi PostCxzn CCs) possono riguardare la riduzione fonetica e morfologica o l'espansione a più contesti collocazionali.

La CxG fornisce strumenti utili a descrivere e spiegare i processi di Costruzionalizzazione e Cambiamento costruzionale: in particolare, la struttura su più livelli (PHON, MORPH, etc.) della Costruzione (figura 6) permette di rendere conto dei piccoli e graduali cambiamenti di cui si compone il mutamento e che agiscono sulle singole dimensioni interne. La struttura gerarchica delle Costruzioni (figura 8) permette di individuare l'inizio del mutamento (a partire dai PreCxzn) e seguirne le successive fasi.

Traugott (2007: 525) ha proposto una rappresentazione delle Gerarchie di eredità su quattro livelli, funzionali a monitorare il susseguirsi dei piccoli cambiamenti che generano il mutamento (figura 13).

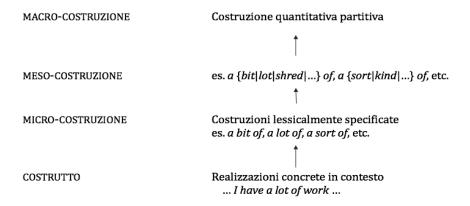


Figura 13. Gerarchia di eredità della Costruzione quantitativa partitiva (Traugott 2007)

La figura 13 riporta i quattro livelli di rappresentazione, esemplificati dal caso delle Costruzioni quantitative partitive dell'inglese, analizzate da Traugott (2007: 530-539). Come indicano le frecce, i cambiamenti linguistici hanno origine dai Costrutti, ossia le realizzazioni concrete prodotte dai parlanti in contesto: quando il cambiamento coinvolge un buon numero di occorrenze, si determina una rianalisi semantica a livello delle Micro-costruzioni e un *mismatch* tra forma e significato; gradualmente, le singole Micro-costruzioni si allineano sotto un'unica Meso-Costruzione più astratta.

Nell'ambito degli studi costruzionisti dedicati alla morfologia, la nozione di Costruzionalizzazione è stata adottata da Hilpert (2013)<sup>61</sup> per l'analisi dell'evoluzione della Costruzione con il suffisso – *ment* in inglese e dei composti inglesi Nome-Participio (il tipo *corpus-based*, *home-made*, etc.; cfr. Hilpert 2015): in entrambi gli studi, Hilpert utilizza dati di frequenza estratti da corpora diacronici per individuare la formazione di Costruzioni semi-specificate e analizzarne la produttività nel corso dei secoli.

-

<sup>&</sup>lt;sup>61</sup> In particolare, nel caso della Costruzione V-ment, Hilpert (2013: 205) parla di *contact-induced constructionalization*: «The process that is observed here is that a group of nouns were borrowed wholesale. As many of these words were morphologically transparent, that is, analyzable into a verbal stem and a suffix, speakers eventually arrived at a generalization that allowed them to coin new formations. A fitting term for this might be contact-induced constructionalization, since it is an influx of borrowed lexis that eventually triggers the emergence of a productive construction. Whether or not the *V-ment* construction is grammatical may be a matter of debate and need not concern us here».

Un caso di Costruzionalizzazione a cui la CxM ha recentemente rivolto particolare attenzione riguarda lo sviluppo di affissi e affissoidi a partire da elementi che occorrono come costituenti di composti.<sup>62</sup>

Lo sviluppo di suffissi a partire da costituenti di destra di composti è ben testimoniato in inglese: come emerge dall'analisi di Trips (2009), gli attuali suffissi –hood, –dom, –ship derivano da parole originariamente libere che hanno subito un processo di Costruzionalizzazione. Uno sviluppo intermedio è invece quello che porta alla nascita di un affissoide (affixoid),63 definito da Hüning & Booij (2014: 590-591) come «a compound constituent with an affix-like behavior which corresponds to an indipendent word with respect to its form, but not with respect to its meaning»: si tratta quindi di entità lessicali originariamente libere, che progressivamente acquisiscono proprietà formali, semantiche e distribuzionali simili a quelle degli elementi legati (prefissi o suffissi).

Un esempio di affissoide di una lingua germanica è la forma *hoofd* dell'olandese, che può occorrere con il significato di 'testa' (sia come parola autonoma sia come costituente di composto) o con quello più astratto di 'principale' (ma, in questo caso, solo come costituente di composto). L'uso in composizione di *hoofd*, definito da Booij (2017: 22) un caso di *construction-dependent polysemy*, può essere rappresentato attraverso gli strumenti della CxM come riportato nella figura 14:

\_

<sup>&</sup>lt;sup>62</sup> Hüning & Booij (2014: 600) osservano infatti che «The development of derivational affixes from compound constituents is primarily a case of constructionalization, the rise of a morphological construction, and morphological change can adequately be analyzed as constructional change at the word level». Seguendo Trousdale (2008: 156), i due autori sostengono che per riferirsi a tale processo, la nozione di 'Costruzionalizzazione' è da preferirsi a quella di 'grammaticalizzazione'.

<sup>63</sup> Il termine *affissoide* non è estraneo alla tradizione italiana di studi sulla formazione delle parole, in riferimento a elementi morfologici non autonomi, impiegati per formare composti del tipo neoclassico (in combinazione tra loro o legati a un elemento autonomo), sulla cui classificazione e delimitazione non c'è ancora accordo in letteratura (cfr. Iacobini 2010). È infatti usato da Bruno Migliorini (a cui si deve l'introduzione del termine *prefissoide*; cfr. Migliorini 1963); Dardano (1978) e Antonelli (1996). In questo lavoro, si utilizzeranno i termini *affissoide/prefissoide/suffissoide* nel senso specifico della CxM, così come definito in Hüning & Booij (2014: 590-591).

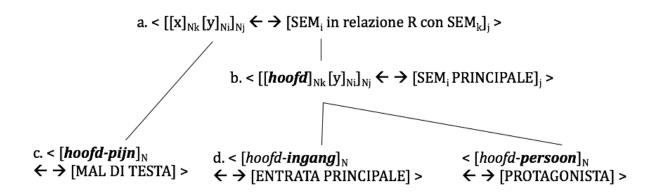


Figura 14. Gerarchia dei composti con hoofd- in olandese Fonte: adattato da Masini (2016: 84)

Nella Gerarchia riportata nella figura 14, il livello più alto è rappresentato dalla Costruzione generica dei composti N-N dell'olandese (a), che instanzia produttivamente nuovi composti, tra i quali *hoofd-pijn* 'mal di testa' (c) in cui *hoofd-* presenta il significato originario di 'testa'. Le forme in cui hoofd- presenta il significato più astratto di 'principale' (d) non dipendono direttamente da (a) ma da un livello intermedio (b), ossia una Costruzione semi-specificata, in cui il primo elemento hoofd- è fisso e il secondo è variabile, considerata da Booij (2005) un caso di constructional idiom,64 ossia uno schema posto a un livello intermedio tra uno schema astratto e i singoli Costrutti. Attraverso tale livello si rende conto della produttività di un significativo numero di forme in cui *hoofd*- presenta proprietà confrontabili a quelle di un vero e proprio prefisso (ad es. posizione fissa a sinistra, semantica più astratta). In tedesco, il processo di Costruzionalizzazione si è spinto oltre: l'elemento *Haupt*-, sostituito dalla parola *Kopf* per l'espressione del significato 'testa', è ormai utilizzato soltanto come prefisso con il valore di 'principale' (come in Hauptbahn 'linea (ferroviaria) principale') e non ha più una controparte libera.

Come si può osservare dal caso di hoofd-, l'elemento che occorre come costituente di composto acquisisce gradualmente nuove proprietà formali, semantiche e distribuzionali; su queste tre dimensioni si basano i quattro parametri proposti da Van Goethem (2010: 169-175) per individuare l'emergere di un affissoide: 65

<sup>64</sup> In particolare, Booij (2005: 129), riprendendo Jackendoff (2002), definisce constructional idiom «a construction in which one of the two positions is occupied by a specific lexical item».

<sup>65</sup> Tali parametri, qui brevemente presentati, verranno discussi più ampiamente al § 5.2 e applicati a due Costruzioni dell'italiano (cfr. capitolo 5).

- 1. *specializzazione semantica* o *risemantizzazione* (sviluppo di un nuovo significato generalmente più astratto, spesso valutativo)
- 2. *decategorializzazione* (perdita della flessione)
- 3. *paradigmaticizzazione* (inserimento nel paradigma di veri prefissi/suffissi con cui entra in competizione)
- 4. *produttività* (aumento della *type frequency* e degli hapax)

Tali parametri, basati sugli studi dedicati al processo di grammaticalizzazione (cfr., tra gli altri, Hopper 1991; Hopper & Traugott 1993; Lehmann 1995; Ten Hacken 2000; Marchello-Nizia 2006), sono stati applicati da Van Goethem (2008; 2010) ad alcuni aggettivi dell'olandese (oud 'vecchio' e nieuw 'nuovo') e del francese (nouveau 'nuovo') che, quando occorrono all'interno di parole composte, possono essere considerati parte di un constructional idiom, come nel caso di hoofd- (cfr. schema in (b) nella figura 14). In particolare, l'aggettivo francese *nouveau*, quando si combina a un participio passato (come in nouveau-né 'neonato' o nouveau(-)venu 'neoarrivato'), dà luogo alla Costruzione semi-specificata [[nouveau] [ $x]_{PART.PASS}]_{N|A}$ . Dal punto di vista semantico, se nouveau come forma libera presenta sia il significato di 'nuovo, che è recentemente accaduto' sia quello di 'secondo' o 'altro' (cfr. TFLi, s.v. nouveau), quando si trova all'interno della Costruzione, esprime invece sempre e solo il primo valore: nouveau-né è infatti usato per indicare un bambino appena nato, mai un secondogenito. Secondo Van Goethem (2010: 170), la forma ha subito un processo di specializzazione semantica, sviluppando un significato più grammaticale,66 paragonabile a quello veicolato dal prefisso néo-, con cui è entrato in competizione, come dimostra la presenza di coppie del tipo nouveauné/néonatal o nouveau-converti/néophyte, indizi dell'integrazione di nouveau nel paradigma di tale prefisso (paradigmaticizzazione). Dal punto di vista formale, inoltre, nouveau occorre in un discreto numero di casi nella forma invariabile nouveau senza flettersi rispetto al genere e al numero del participio con cui si combina, mostrando

<sup>-</sup>

<sup>&</sup>lt;sup>66</sup> La Costruzione *nouveau*+participio passato sembra infatti porsi come alternativa alla costruzione *venir de* + infinito (si confrontino *nouveau-né* con *qui vient d'être né*), attraverso cui in francese si esprime il valore (grammaticale) di passato recente (Van Goethem 2010: 174).

quindi segni di decategorializzazione. <sup>67</sup> Infine, l'alto numero di neologismi composti con nouveau- attestati nel web testimonia la produttività della Costruzione  $[[nouveau] [x]_{PART.PASS}]_{N|A}$  e ne rafforza l'interpretazione nei termini di constructional idiom in cui l'elemento specificato presenta delle proprietà assimilabili a quelle di un prefisso.

L'emergere di un nuovo affissoide e di una nuova Costruzione può non costituire l'ultimo stadio del mutamento ma dar luogo a un ulteriore sviluppo: come messo in luce da Van Goethem & De Smet (2014) e Norde & Van Goethem (2014), infatti, un affissoide nato a partire da un costituente di composto può a sua volta sviluppare proprietà di aggettivo o avverbio libero per effetto di un processo di degrammaticalizzazione definito debonding, ossia «a composite change whereby a bound morpheme in a specific linguistic context becomes a free morpheme» (Norde 2009: 186). È il caso del sostantivo tedesco *Riesen* 'gigante' che, come l'olandese *reuze* (cfr. (2); Van Goethem & Hiligsmann 2014), ha acquisito lo status di affissoide ed è stato poi rianalizzato come un aggettivo/avverbio libero dal significato di 'molto' (Norde & Van Goethem 2014), sviluppando proprietà semantiche e distribuzionali differenti dalla controparte legata (cfr. processo definito *divergence* da Hopper 1991: 24).

Lo sviluppo di affissi e affissoidi a partire da elementi che occorrono come costituenti di composti rafforza l'idea che composizione e derivazione siano meccanismi posti lungo un *continuum*: gli elementi che occorrono all'interno delle Costruzioni<sub>COMP</sub> possono infatti non conservare invariato il proprio status di costituenti e sviluppare proprietà di affissi, determinando il mutamento composto > derivato, osservabile in diacronia. Nonostante molti studi dedicati a questo fenomeno si focalizzino quasi esclusivamente sull'elemento che si grammaticalizza, va sottolineato che, in ottica costruzionista, tale processo non coinvolge soltanto il singolo costituente di composto ma l'intera Costruzione: come chiariscono Hüning & Booij (2014: 587), infatti, «It is not the isolated lexical items that become affixes; rather, it is complex words, compounds that get new interpretations and meanings».

L'analisi del mutamento a cui vanno incontro le Costruzioni<sub>COMP</sub> permette di mettere in luce che la composizione, pur essendo dotata di caratteristiche proprie che la

<sup>&</sup>lt;sup>67</sup> Come si discuterà più approfonditamente al § 5.2, la perdita della flessione è un indizio di decategorializzazione a cui sono soggetti gli elementi originariamente liberi che acquisiscono lo status di affissi.

identificano e la distinguono, si relaziona con gli altri meccanismi di formazione delle parole, non soltanto in sincronia ma anche in diacronia. Adottando tale duplice punto di vista, questo lavoro intende indagare, in particolare, la relazione tra composizione e derivazione in italiano, attraverso l'analisi dello sviluppo di prefissi e prefissoidi a cui sono dedicati gli studi di caso del capitolo 5.

#### 3. METODOLOGIA DI INDAGINE

In questo capitolo si discute la metodologia di indagine adottata in questo lavoro. Nella prima parte (§ 3.1) si descrivono le fasi di creazione del corpus diacronico CoDIt\_com e di estrazione del dataset di composti *Compos\_ita*: di entrambe le procedure si mettono in luce i limiti e i problemi metodologici; di seguito (§ 3.2), si presentano le altre risorse utilizzate nel corso del lavoro. La seconda parte del capitolo (§ 3.3) è dedicata alla discussione dei parametri formali, semantici e quantitativi adottati nell'analisi con cui verranno analizzati i dati nel capitolo 4. La struttura dell'analisi è descritta nell'ultimo paragrafo (§ 3.4).

## 3.1 CoDIt\_com: Corpus Diacronico dell'Italiano per la composizione

Il dataset di parole composte su cui si basa questo lavoro è stato estratto da un corpus diacronico costruito *ad hoc* e costituito da testi in formato .txt non annotati. Di seguito si fornisce una descrizione delle fasi di creazione del corpus; architettura e bilanciamento della risorsa vengono poi confrontati con altri corpora diacronici disponibili per l'italiano e altre lingue, di cui si mettono in luce limiti e potenzialità.

## 3.1.1 Finalità del corpus

Il CoDIt\_com (*Corpus Diacronico dell'Italiano per la composizione*) è un corpus creato per lo studio dell'italiano in prospettiva diacronica e, in particolare, della formazione delle parole composte. In quanto risorsa finalizzata all'analisi diacronica di fenomeni morfologici dell'italiano, esso si affianca al corpus MiDIA (*Morfologia dell'Italiano in DIAcronia*, cfr. Grossmann & D'Achille 2017),68 costituito da circa 7 milioni di token, lemmatizzato e annotato per parti del discorso attraverso TreeTagger (cfr. Schmid 1995); MiDIA può inoltre essere interrogato per affissi, permettendo così di compiere ricerche quantitative e qualitative sulle parole derivate.

<sup>. .</sup> 

<sup>&</sup>lt;sup>68</sup> Il corpus è liberamente consultabile all'indirizzo <u>www.corpusmidia.unito.it</u> (ultima visualizzazione in data 13/03/18).

Le ragioni per cui si è scelto di non utilizzare tale corpus, ma di ricorrere alla creazione di una nuova risorsa, sono essenzialmente due: le dimensioni; l'assenza di annotazione per le parole composte. Nonostante MiDIA sia stato pensato per lo studio della formazione delle parole in diacronia, infatti, la sua architettura risponde di fatto all'obiettivo di analizzare esclusivamente il fenomeno della derivazione, la cui alta produttività fin dal periodo delle Origini può essere osservata anche in un corpus di dimensioni limitate: lo studio dei composti, sia in sincronia sia in diacronia, necessita invece di dimensioni notevolmente più consistenti per permettere di raccogliere dati quantitativi e qualitativi relativi a entità lessicali mediamente rare. Infine, il fatto che MiDIA non preveda l'annotazione delle parole composte costituisce un ulteriore limite all'utilizzo di questa risorsa per lo scopo della presente ricerca.

### 3.1.2 Architettura del corpus

La struttura del CoDIt\_com prende a modello quella di MiDIA, di cui è stata adottata la scansione cronologica e, con alcune modifiche, la ripartizione in tipologie testuali.

La suddivisione in cinque periodi storici è scandita da eventi decisivi per la storia dell'italiano; in particolare, seguendo MiDIA, sono state individuate le seguenti date, considerate punti di svolta per le vicende storiche della lingua:

- 1. Periodo 1. Dalle prime testimonianze del volgare al 1375, anno di morte di Boccaccio e inizio dell'attività cancelleresca dell'umanista Coluccio Salutati: i limiti di questo periodo coincidono con quelli adottati nel corpus OVI<sup>69</sup> e comprendono la fase del cosiddetto *italiano antico* o *italiano delle Origini* (cfr. § 3.4);
- 2. Periodo 2. Dal 1375 al 1532: questo periodo, che si colloca tra Umanesimo e Rinascimento, comprende lo sviluppo del fiorentino "argenteo" e la successiva definizione di un canone "aureo" a partire dalla pubblicazione nel 1525 delle *Prose della volgar lingua* di Pietro Bembo. La data conclusiva di questo periodo coincide con la pubblicazione della terza edizione dell'*Orlando furioso*, adattata da Ariosto al canone fiorentino proposto da Bembo, di cui testimonia la solida affermazione;

\_

 $<sup>^{69}</sup>$  Il corpus OVI (*Opera Vocabolario Italiano*) è liberamente accessibile in rete: <a href="http://gattoweb.ovi.cnr.it/">http://gattoweb.ovi.cnr.it/</a> (ultima visualizzazione in data 14/09/2018).

- 3. Periodo 3. Dal 1533 al 1691: il periodo abbraccia tardo Rinascimento, Manierismo e Barocco e termina con la pubblicazione della terza edizione del *Vocabolario degli Accademici della Crusca* e la fondazione dell'Arcadia;
- 4. Periodo 4. Dal 1692 al 1840: il periodo comprende l'età dell'Arcadia, l'Illuminismo e il Romanticismo e termina con l'ultima edizione dei *Promessi Sposi*, modello indiscusso dell'italiano postunitario;
- 5. Periodo 5. Dal 1841 al 1947: l'ultimo periodo abbraccia il Risorgimento, la fase postunitaria e le due guerre mondiali fino alla promulgazione della Costituzione.

I testi contenuti nel CoDIt\_com appartengono a sei tipologie testuali, di cui tre di ambito letterario (a, b, d) e tre di ambito non letterario (c, e, f):

- a. poesia (letteratura in versi destinata alla lettura)
- b. prosa letteraria (racconti, novelle, romanzi, prosa d'arte)
- c. personali (testi non destinati alla pubblicazione: lettere, autobiografie, ricordi)
- d. teatro (tragedie, commedie, libretti d'opera, rappresentazioni sacre)
- e. espostivi (cronache, saggi, trattati, manuali a carattere storico-filosofico)
- f. scientifici (saggi, manuali, trattati relativi alle scienze dure)

### 3.1.3 Raccolta dei testi

I testi sono stati scaricati in formato .txt da due banche dati liberamente accessibili in rete: *Biblioteca Italiana* (BibIt)<sup>70</sup> e *Liber Liber*.<sup>71</sup> Nella scelta dei testi si è adottata, come punto di partenza, la lista disponibile sul sito di MiDIA, rispetto alla quale sono state aggiunte ulteriori opere quando ritenuto opportuno; in alcuni casi non è stato possibile includere un testo perché non reperibile in formato elettronico.

Una significativa differenza a livello metodologico tra i due corpora riguarda il campionamento. MiDIA è costituito da 25 testi per periodo storico, di ciascuno dei quali è stata selezionata una porzione di circa 8.000 token: la struttura generale del corpus, che può essere considerato un esempio di *sample corpus*, appare molto equilibrata. Per il

<sup>&</sup>lt;sup>70</sup> La banca dati è accessibile al seguente indirizzo: <a href="http://www.bibliotecaitaliana.it/">http://www.bibliotecaitaliana.it/</a> (ultima visualizzazione in data 13/03/18).

<sup>&</sup>lt;sup>71</sup>La banca dati è accessibile al seguente indirizzo: <a href="http://www.liberliber.it/">http://www.liberliber.it/</a> (ultima visualizzazione in data 13/03/18).

CoDIt\_com, invece, si è scelto di includere ciascun testo nelle sue dimensioni originali, salvaguardandone l'integrità: questa scelta ha inevitabilmente reso più difficile raggiungere un equilibrio tra le diverse sezioni del corpus ma, d'altra parte, ha permesso di evitare i limiti metodologici che il campionamento comporta (ad es. la rottura dell'unità testuale, l'alterazione della struttura complessiva dell'evento comunicativo, la perdita di dati di natura lessicale, l'impossibilità di calcolare la dispersione etc.). Per questo aspetto si è seguito il modello di altri corpora diacronici come il *Corpus del Español*, il *Corpus do Portuguese* (Davies 2002, 2009) o l'*Helsinki Corpus* dell'inglese (Kytö & Rissanen 1992): risorse molto diverse per dimensioni e architettura ma accomunate dal fatto di essere tutte esempi di *text corpora*.

## 3.1.4 Composizione del corpus

Il CoDIt\_com è composto da oltre 29 milioni di parole, distribuite in cinque sotto-corpora, corrispondenti ai cinque periodi individuati sulla base delle vicende storiche della lingua italiana: ciascun sotto-corpus è costituito da circa 6 milioni di token, eccetto il sotto-corpus corrispondente al primo periodo che, a causa delle difficoltà di reperimento dei testi del periodo delle Origini, costituisce un caso particolare.<sup>72</sup>

Di seguito si illustra l'architettura del CoDIt\_com dal punto di vista quantitativo (tabella 4).

	1	2	3	4	5	totale	%
sezione	(Origini- 1375)	(1376- 1532)	(1533- 1691)	(1692- 1840)	(1841- 1947)		
poesia	1.080.244	1.316.538	1.270.128	1.116.549	1.156.588	5.940.047	20
prosa lett.	1.393.801	1.453.174	1.543.485	1.489.027	1.558.758	7.438.245	25
personali	28.023	1.126.377	1.037.187	1.227.149	1.300.913	4.719.649	16
teatro	64.121	410.363	461.492	433.462	442.850	1.812.288	6
espositivi	1.681.422	1.227.320	1.339.114	1.105.755	1.450.335	6.803.946	23
scientifici	-	599.572	695.184	674.361	634.447	2.603.564	9
totale	4.247.611	6.133.344	6.346.590	6.046.303	6.543.891	29.317.739	100

Tabella 4 Dimensioni dei sotto-corpora del CoDIt\_com (token)

\_

<sup>&</sup>lt;sup>72</sup> Nel corso dell'analisi diacronica, i dati provenienti dal sotto-corpus 1, di dimensioni più limitate rispetto agli altri, verranno normalizzati. Come verrà spiegato più nel dettaglio ai § 3.2.2 e 3.4, per l'analisi della composizione in italiano antico, i dati del CoDIt\_com verranno integrati con quelli estratti dal Lemmario del TLIO (*Tesoro della Lingua Italiana delle Origini*).

I sottocorpora 2-5 sono costituiti, in proporzioni diverse, da testi appartenenti alle sei sezioni: testi poetici, espositivi, personali e prosa letteraria superano il milione di token in ciascun periodo; i testi teatrali contano circa 400.000 token; i testi scientifici circa 600.000 token. Va notato che il sotto-corpus 1 rispetta le proporzioni degli altri sotto-corpora solo nel caso della poesia, della prosa letteraria e dei testi espositivi; i testi personali e di letteratura teatrale hanno dimensioni molto limitate, mentre la letteratura scientifica è del tutto assente.

sezione	n. testi	token	media token	% totale
poesia	208	5.940.027	28.557	20
prosa letteraria	90	7.431.300	82.570	25
personali	57	4.719.649	82.800	16
teatro	116	1.812.288	15.623	6
espositivi	89	6.803.946	76.448	23
scientifici	43	2.612.564	60.757	9
totale	603	29.317.739	48.619	100

Tabella 5. Dimensioni delle sezioni dei generi testuali del CoDIt\_com

Come riportato nella tabella 5, i testi inseriti nel corpus sono complessivamente 603; la lunghezza media di un testo è di 48.619 token: i testi mediamente più lunghi si trovano nelle sezioni di prosa letteraria, testi personali ed espositivi; i testi più brevi appartengono alla letteratura teatrale. Le sezioni di dimensioni più consistenti sono quelle della prosa letteraria (26%), dei testi espositivi (23%) e dei testi poetici (20%); seguono la sezione dei testi personali (16%), quella dei testi scientifici (9%) e della letteratura teatrale (6%). Nella fase di raccolta dei testi si è cercato di raggiungere un equilibrio tra fonti letterarie e fonti appartenenti ad altre tipologie: i testi letterari (prosa+poesia+teatro) coprono infatti il 52% del totale del corpus, come rappresentato nel grafico riportato nella figura 16.

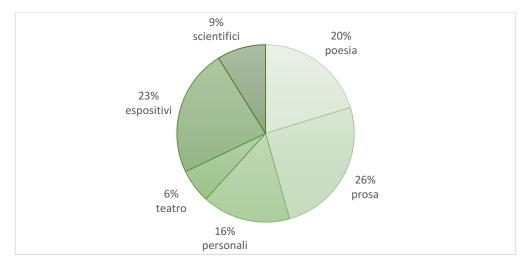


Figura 15. CoDIt\_com: proporzioni delle sezioni

## 3.1.5 Confronto con altri corpora diacronici

Come già messo in evidenza, le principali differenze tra il CoDIt\_com e MiDIA riguardano le dimensioni (dei testi e dell'intero corpus) e l'annotazione. In relazione al fine per cui è stato costruito, il CoDIt\_com ha infatti dimensioni sensibilmente più grandi e non è stato annotato (cfr. § 3.1.6).

La struttura in sezioni e sotto-corpora di MiDIA (figura 17) ha costituito il modello per quella del CoDIt\_com, eccetto per quanto riguarda la sezione di testi giuridici: a causa delle difficoltà incontrate nel reperimento di questo tipo di testi in formato elettronico, si è scelto di non dotare il CoDIt\_com di una sezione autonoma di testi giuridici ma di includere i pochi esemplari nella sezione dei testi espositivi.

Periodi	<b>Espositivi</b>	Giuridici	Personali	Poesia	Prosa	Scientifici	Teatro	<b>Totale periodo</b>
I	112385	438295	8447	207484	214503	237634	19709	1238457
II	279101	206130	251838	264792	218144	231749	194674	1646428
III	230175	261990	120181	242498	220798	268389	256270	1600301
IV	231988	199499	223654	192320	252278	206969	192704	1499412
V	263565	250541	253362	242858	224358	215559	217685	1667928
								Totale corpus
Tot. Genere	1117214	1356455	857482	1149952	1130081	1160300	881042	7652526

Figura 16. Architettura e dimensioni del corpus MIDIA (token)

Tra gli altri corpora diacronici di lingue europee, quello che per dimensioni si avvicina più al CoDIt\_com è il *Corpus do Português* (40 milioni di token), la cui architettura, illustrata nella figura 18, presenta però alcune differenze relative alla suddivisione in sotto-corpora (coincidenti con le date di inizio e fine di ciascun secolo), alle dimensioni di ciascun sotto-corpus (molto variabili e sbilanciate a favore degli ultimi due secoli) e alla composizione dell'ultimo sotto-corpus che, unico tra gli altri, è stato bilanciato per tipologie testuali (bilanciamento assente negli altri sotto-corpora).

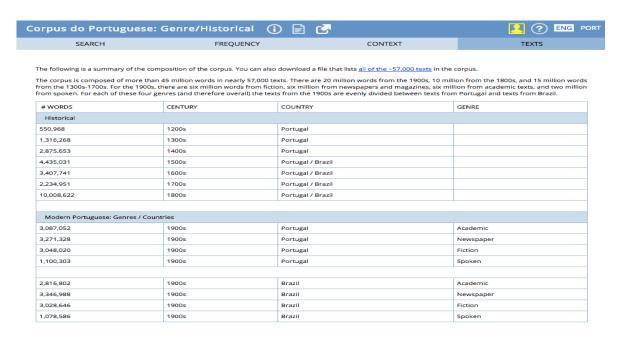


Figura 17. Architettura del Corpus do Portuguese

Di dimensioni notevolmente maggiori è il *Corpus del Español* (100 milioni di token; cfr. Davies 2002, 2009), costruito secondo un'architettura molto simile al precedente, ma potendo contare su una maggior disponibilità di testi, anche se non omogenea per tutti i secoli. In particolare, infatti, come illustrato dalla tabella riepilogativa riportata nella figura 19, il sotto-corpus del 1300 ha dimensioni molto limitate rispetto a tutti gli altri, che pure hanno un numero di token molto variabile.

CENTURY	WORDS	# TEXTS	SOURCES
Historical			
1200s	7,079,164	71	<ul> <li>[HSMS] Electronic Texts and Concordances of the Madison Corpus of Early Spanish Manuscripts and Printings. Preparado por John ONeill. (Madison y New York, 1999).</li> </ul>
1300s	2,667,810	50	ADMYTE (Archivo Digital de Manuscritos y Textos Españoles). Vol 0 y 2.     Biblioteca Virtual
1400s	8,747,963	160	Gonzalo de Berceo: Obras Completas
1500s	17,774,762	323	Biblioteca Virtual [1500s-1700s]
1600s	13,355,483	498	COMEDIA (Univ. de Arizona) [1600s]     Proyecto Filosofía en español [1700s]
1700s	10,324,328	176	
1800s	20,822,142	392 novels	Biblioteca Virtual
Modern Spa	nish: Genres		
900s-News	5,144,631	6810 articles	ABC Cultural 1991-95
			Noticias - Argentina - La Prensa Noticias - Argentina - El Cronista Noticias - Bolivia - ERBOL Noticias - Perú - Caretas Noticias - Colombia - Semana Noticias - Colombia - Semana Noticias - Cuba/EEUU - CubaNet
1900s-Lit	5,144,073	850 novels+ short stories	Biblioteca Virtual  [BYU] Humanities Research Center, Brigham Young University  [ARG] Corpus lingüístico de referencia de la lengua española en Argentina  [CHILE] Corpus lingüístico de referencia de la lengua española en Chile  Proyecto Sherezade  Badosa.com  Ficticiosa.com
1900s-Oral	5,113,249	2040+ interviews+ transcripts	Habla Culta  [ESP-ORAL] Corpus oral de referencia de la lengua española contemporánea, including CORLEC, Corpus Oral de Referencia de la Lengua Española Cortes de Castilla-La Mancha Congreso de España ABC Cultural 1991-95 Chiapas (La República) - Entrevistas Ernesto Zedillo - Entrevistas / discursos Fidel Castro - Entrevistas / discursos Partido PAN - Entrevistas Partido PRI - Entrevistas
1900s-Acad	5,138,077	2931 articles	Encarta en español
TOTAL	101.311.682	13.926	

Figura 18. Architettura del Corpus del Espanol

## 3.1.6 Estrazione dei composti

L'estrazione di parole composte da un corpus pone non pochi problemi metodologici, legati principalmente alla variabilità ortografica che caratterizza questo tipo di forme in italiano. Tra i corpora italiani (sia sincronici sia diacronici) annotati per *Part-Of-Speech* (POS), nessuno presenta un tag specifico per i composti: l'estrazione delle forme unite graficamente non può quindi che avvenire manualmente, a partire dalla lista di frequenza; l'estrazione dei composti i cui costituenti sono separati dallo spazio deve essere condotta attraverso lo spoglio manuale dei bigrammi (o *ngrams*), estratti sulla base delle parti del discorso. In entrambi i casi è possibile che la presenza di errori nel POS-tagging produca falsi positivi e/o falsi negativi, limitando l'efficacia di questa metodologia. Nel caso di corpora non annotati, l'estrazione dei composti (univerbati e non) è affidata interamente allo spoglio manuale della lista di frequenza dei lemmi e di quella degli *ngram*. Nel presente lavoro si è scelto di non annotare il CoDIt\_com, sulla

base dell'assunto che l'annotazione per POS risulterebbe molto imprecisa su testi appartenenti a periodi storici diversi, redatti con convenzioni grafiche differenti.

L'estrazione dei composti univerbati (es. *capolinea*, *portafoglio*) è stata condotta attraverso tre fasi. *In primis*, il corpus è stato caricato in formato .txt sul software AntConc (Anthony 2013), attraverso il quale è stata generata la wordlist: dal momento che il corpus non è né annotato né lemmatizzato, essa non include i lemmi ma semplicemente tutti i token di cui si compone il corpus.

Nella seconda fase, in vista dello spoglio manuale, si è cercato di ridurre le dimensioni della wordlist, eliminando le forme che, con una certa sicurezza, possono dirsi non composte, adottando come criterio dirimente la lunghezza della parola.<sup>73</sup> Il numero minimo di lettere da cui è costituita una parola composta è stato calcolato sulla base di un campione di 2.425 composti precedentemente estratti attraverso uno spoglio manuale di un dizionario, il *Devoto Oli 2014* (cfr. Devoto Oli 2014), e della lista di frequenza di un corpus di italiano contemporaneo, il *Corpus del Nuovo Vocabolario di Base* (cfr. Chiari & De Mauro 2012). Nel grafico (figura 20) si riporta la distribuzione quantitativa dei composti rispetto alla lunghezza.

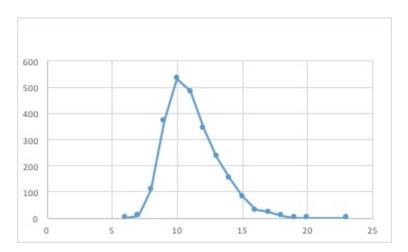


Figura 19. Lunghezza dei composti del campione estratto dal *Devoto Oli* e dal *Corpus del Nuovo Vocabolario di Base*. Sull'asse delle ascisse è riportato il numero delle lettere; sull'asse delle ordinate il numero dei lemmi.

<sup>73</sup> Dal momento che una parola composta è per definizione costituita da almeno due parole, si può legittimamente ipotizzare che questa abbia una lunghezza media superiore a quella di una parola semplice e che tale criterio possa essere utilizzato per escludere da una lista di frequenza un buon numero di parole sicuramente non composte perché di lunghezza troppo limitata.

Come si nota dal grafico riportato nella figura 20, il composto più breve (malora: mal(a) + ora) è costituito da sei lettere, il più lungo da 23 (anarcoinsurrezionalista: anar(chi)co + insurrezionalista); la maggior parte dei composti del campione è costituita da 10-13 lettere. Sulla base di questi dati si è scelto di eliminare dalla wordlist del CoDIt\_com tutte le parole costituite da meno di cinque lettere, ottenendo così un elenco di dimensioni più contenute di cui fare lo spoglio manuale.

Nella terza fase si è proceduto con lo spoglio manuale della wordlist ridotta, da cui sono state estratte tutte le parole composte appartenenti a tutte le tipologie attestate nella composizione italiana (cfr. § 3.1.7 per una descrizione del dataset); si è scelto inoltre di estrarre anche forme come *pomodoro* o *capodanno*, nate dalla cristallizzazione di espressioni sintagmatiche e contenenti una preposizione o un articolo.<sup>74</sup> I toponimi e gli antroponimi creati per composizione (soprattutto verbonominale) sono stati etichettati rispettivamente come "TOPON" e "NP" e separati, perché esclusi dall'analisi.

L'estrazione ha portato a raccogliere un campione di circa 876 composti univerbati (esclusi toponimi e antroponimi), ciascuno dei quali è stato cercato nel corpus anche nella grafia separata e con il trattino. Questa procedura ha permesso di estrarre le parole composte che occorrono almeno una volta graficamente unite, ma non quelle che occorrono sempre e solo con i due costituenti separati dallo spazio o dal trattino. Al fine di includere il maggior numero di forme possibile, si è scelto di espandere il dataset attraverso ricerche mirate dei pattern più ricorrenti in cui un costituente è specificato e l'altro è variabile (ad es. i composti con *capo*-). Una volta individuati i costituenti più produttivi in termini di type (tabella 6), si è proceduto cercandoli nella lista dei bigrammi del CoDIt\_com.

Primo costituente	Numero type (> 5)	Secondo costituente	Numero type (>5)
саро-	49	-cuore	9
malo-	32	-mano	8
male-	30	-uomo	7
batti-	20	-vendolo	6
porta-	19	-forte	6
guarda-	17		
bene-	16		
mezzo-	15		

-

<sup>&</sup>lt;sup>74</sup> In questa fase si è scelto di adottare criteri non troppo stringenti ed estrarre il maggior numero di parole composte, rimandando a una fase successiva una eventuale scrematura delle forme raccolte.

Primo costituente	Numero type (> 5)	Secondo costituente	Numero type (>5)
para-	13		
alto-	11		
salta-	10		
mangia-	10		
verde-	9		
caca-	8		
pesce-	8		
cava-	8		
piscia-	7		
taglia-	7		
passa-	7		
lecca-	6		
spacca-	6		
acqua-	6		
buono-	6		
lavare-	6		

Tabella 6. Costituenti che occorrono in più di cinque composti estratti dal CoDIt\_com

Nella tabella 6 sono riportati i primi e i secondi costituenti che occorrono in più di cinque parole composte; come si può notare dalla distribuzione tra la prima e la terza colonna, l'italiano sembra preferire il pattern con la prima posizione specificata e la seconda variabile: i primi costituenti che formano più di cinque parole composte, infatti, sono ben 24, di contro a cinque secondi costituenti. A partire dai costituenti individuati si è quindi proceduto con ricerche mirate nella lista dei bigrammi del CoDIt\_com: il campione di forme così ottenuto costituisce il dataset definitivo, la cui composizione è descritta più nel dettaglio nel paragrafo che segue.

## 3.1.7 Il dataset Compos\_ita

Il campione di parole composte estratte dal CoDIt\_com che verrà analizzato in questo lavoro è stato denominato *Compos\_ita* (*Composti dell'italiano*): esso è costituito complessivamente da 2.179 type e 37.512 token.

La tabella 7 ne fornisce un quadro generale, senza considerare la dimensione diacronica, di cui si darà conto nel corso dell'analisi al capitolo 4; le forme sono distribuite secondo le categorie di input, di cui viene riportata la frequenza dei type e dei token.

Categorie di input <sup>75</sup>	Туре	Token	Esempio
[A-N]	714	22.002	malasorte
[AVV-PART.PASS]	388	4.159	maltrattato
[V-N]	324	1.884	guardaroba
[AVV-V]	220	2.212	benvolere
[N-N]	78	1.139	capofamiglia
[AVV-A]	88	2.000	malsano
[N-A]	77	1.539	cassaforte
[A-A]	30	276	sordomuto
[AVV-PART.PRES]	32	178	maldicente
[V-V]	11	153	battisoffia
[AVV-AVV]	11	658	malvolentieri
[AVV-N]	3	443	malessere
[N-PART.PRES]	10	226	armilucente
[A-PART.PASS]	5	561	terzogenito
[V-AVV]	4	7	buttafuori
[N-V]	2	83	manomettere
Totale	1.997	37.520	

Tabella 7 Dataset *Compos\_ita*: distribuzione dei composti rispetto alla frequenza dei type e dei token

Il quadro generale fornito dalla tabella 7 permette di identificare la Costruzione<sub>COMP</sub> [A-N] come la più produttiva, in termini sia di token sia di type: il primato di questa tipologia è in parte legato alla produttività della Costruzione<sub>COMP</sub> semi-specificata mal(o/a)+x (le cui instanziazioni costituiscono il 50% dei composti A-N), che sarà analizzato al capitolo 5. Nella composizione, l'italiano sembra quindi preferire notevolmente l'ordine marcato Aggettivo – Nome (es. mezzaluna), rispetto a quello non marcato Nome – Aggettivo (es. roccaforte), il cui status morfologico non è sempre di facile definizione, come messo in luce al § 2.4.3.

In termini di frequenza dei type, è significativa la produttività delle Costruzioni<sub>COMP</sub> con un elemento avverbiale come primo costituente ([AVV-PART.PASS], [AVV-V] e [AVV-A]), dovuta prevalentemente alla fortuna della Costruzione<sub>COMP</sub> *mal(e)*-x, la cui evoluzione verrà messa in luce al capitolo 5. Non stupisce la notevole produttività (in termini sia di token sia di type) della Costruzione<sub>COMP</sub> [V-N], tradizionalmente considerato il tipo di composizione più sfruttato da tutte le lingue romanze, ad eccezione del rumeno (cfr. Grossmann 2012). I composti con ordine inverso N-V, tradizionalmente presentato dai composti latini, sono invece piuttosto rari.

Rispetto al quadro presentato in questo paragrafo, che fornisce una rappresentazione 'appiattita' dei composti che l'italiano ha creato nel corso della sua storia, al capitolo 4 si

<sup>&</sup>lt;sup>75</sup> Le Costruzioni<sub>COMP</sub> sono ordinate rispetto alla frequenza dei type.

fornirà un'analisi in prospettiva diacronica, finalizzata a mettere in luce l'evoluzione qualitativa e quantitativa dei diversi tipi di Costruzione<sub>COMP</sub>.

### 3.2 Altre fonti

#### 3.2.1 Il lessico derivazionale Word Formation Latin

La prima parte del capitolo 4 fornisce un quadro della composizione latina, funzionale a mettere in luce che cosa rimane delle parole composte latine nel lessico dell'italiano. Questa panoramica si basa su dati estratti dal lessico derivazionale *Word Formation Latin* (d'ora in poi WFL; cfr. Litta, Passarotti & Culy 2016), creato per lo studio della formazione delle parole in latino. La risorsa raccoglie 43.432 parole latine morfologicamente complesse (derivati, composti, conversioni), estratte da tre dizionari (Georges & Georges 1913-18, Glare 1982 e Gradenwitz 1904).

Per ciascun lemma, WFL fornisce informazioni su basi, affissi e Regole di Formazione di Parola (RFP)<sup>76</sup> e genera un grafico derivazionale che ne riproduce la struttura, come esemplificato dal caso di *damnicupidus* (lett. 'desideroso di sofferenza'), costituito dal sostantivo *damnum* e l'aggettivo *cupidus*, legati dalla regola N+A=A (figura 21).

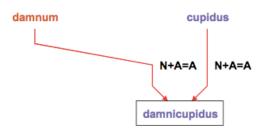


Figura 20. Grafico derivazionale fornito da WFL per il composto damnicupidus

La risorsa può essere interrogata per RFP, per costituente e per parti del discorso (POS): per estrarre il dataset di composti (tabella 8) è stata utilizzata la ricerca per RFP che ha

\_

<sup>&</sup>lt;sup>76</sup> La risorsa si basa programmaticamente sul modello *Item-and-Arrangement* (Hockett 1954): gli affissi vengono quindi considerati entrate lessicali al pari delle basi, a cui si uniscono per effetto di una RFP. Derivazione, composizione e conversione sono le tre principali regole codificate in WFL, ciascuna delle quali a sua volta si sotto-categorizza in regole più specifiche (ad es. per la composizione N+A=A, N+N=N, etc.).

permesso di raccogliere 1.798 composti latini, creati attraverso le 63 regole, di cui la tabella 8 riporta quelle che danno luogo ad almeno 2 composti.

	Regola di formazione di parola	Туре
1	N+V=A	444
2	N+V=N	250
3	A+N=A	139
4	A+V=A	137
5	N+N=N	130
6	N+N=A	119
7	V+V=V	71
8	A+V=V	55
9	N+V=V	52
10	A+N=N	52
11	V+V=A	32
12	A+V=N	29
13	I+I=I	27
14	V+N=A	22
15	PR+PR=PR	21
16	I+A=A	19
17	A+A=A	19
18	I+N=N	17
19	N+A=N	16
20	I+V=V	14
21	PR+V=PR	12
22	PR+I=I	11
23	N+A=A	11
24	I+V=A	10
25	I+V=N	9
26	V+N=N	8
27	PR+I=PR	8
28	I+PR=PR	8
29	A+I=I	7
30	I+N=A	6
31	A+A=N	6
32	V+PR=PR	4
33	PR+N=I	4
34	V+V=N	4
35	A+N=I	3
36	N+V=V	3
37	I+N=I	3
38	A+PR=PR	3
39	N+PR+N=N	2
40	PR+A=PR	2
41	I+V=I	2
42	PR+N=A	2

Tabella 8. Regole di formazione di parola dei composti latini estratti da WFL (type > 1)

I dati estratti da WFL, sulla base dei quali al § 4.1 si traccia un bilancio della composizione latina e la sua influenza su quella dell'italiano, sono di natura sia quantitativa sia qualitativa. Dal punto di vista quantitativo, infatti, WFL fornisce informazioni sulla produttività delle RFP o dei singoli costituenti: la tabella 8 illustra le RFP che danno luogo ad almeno 2 parole composte; come si osserverà più nel dettaglio al § 4.1 il pattern più produttivo è N-V, attraverso il quale il latino crea nomi e aggettivi. Sul piano qualitativo, WFL permette di delineare un quadro della struttura e della semantica di un vasto repertorio di parole composte,<sup>77</sup> che può essere confrontato con quello della composizione italiana.

# 3.2.2 Corpora di italiano antico e contemporaneo

Come già illustrato al § 3.1, il CoDIt\_com è un corpus diacronico che rappresenta l'italiano dalle Origini alla prima metà del Novecento, sulla base del quale al capitolo 4 si fornirà un'analisi della composizione italiana.

Gli studi di caso su cui si è focalizzata l'attenzione al capitolo 5 si basano, oltre che sul CoDIt\_com, anche su dati estratti da corpora di italiano contemporaneo, che rappresentano la fase più recente. In particolare, per l'analisi delle Costruzioni $_{COMP}$  peggiorative mal(e)-x, mal(o/a)-x sono stati estratti dati quantitativi e qualitativi dal web corpus itWaC (Baroni et~al.~2009), scelto in virtù delle sue dimensioni e della disponibilità in rete. RI corpus è stato interrogato attraverso l'interfaccia SketchEngine (Kilgarriff et~al.~2014) seguendo una procedura le cui fasi sono descritte al § 5.2.

L'analisi del periodo identificato come *italiano antico* (Origini-1375; cfr. § 3.4) è basata, oltre che su dati estratti dal primo sotto-corpus del CoDIt\_com, anche su un campione di forme estratte manualmente dal Lemmario del TLIO (*Tesoro Lingua italiana delle* 

 $<sup>^{77}</sup>$  Gli studi teorici finora dedicati alla composizione latina sono infatti prevalentemente di carattere qualitativo (per un quadro generale cfr. Brucale 2012) e si basano su un ristretto numero di esempi estratti dai dizionari.

<sup>&</sup>lt;sup>78</sup> L'utilizzo di un corpus di riferimento bilanciato avrebbe senz'altro costituito la scelta più corretta dal punto di vista metodologico: tuttavia, i due corpora di riferimento disponibili per l'italiano, il PErugia Corpus (d'ora in poi PEC; cfr. Spina 2014) e il CORIS (cfr. Favretti *et al.* 2002), presentavano dei limiti dovuti, nel primo caso, alle dimensioni (il PEC è costituito da circa 30 milioni di token), nel secondo, alle restrizioni sui risultati visualizzabili liberamente (il CORIS permette di consultare soltanto i primi 1000 risultati e non fornisce la wordlist dell'intero corpus). Per queste ragioni, si è scelto di utilizzare itWaC (1.5 miliardi di token), un corpus di testi estratti automaticamente dal web, liberamente consultabile in rete.

Origini)<sup>79</sup>, costruito sulla base di una versione ridotta del corpus OVI (*Opera Vocabolario Italiano*).<sup>80</sup> La scelta di utilizzare dati estratti da un'altra risorsa per questa fase della lingua risponde a due ordini di ragioni. *In primis*, nell'ambito degli studi dedicati alla composizione, l'italiano delle Origini costituisce la fase a cui è stata finora rivolta minor attenzione ed è quindi meritevole di un'indagine approfondita, che metta in luce le proprietà mostrate dalla composizione all'inizio della storia dell'italiano da confrontare con quelle delle fasi successive. D'altra parte, la disponibilità di risorse di dimensioni consistenti come l'OVI (su cui è basato il TLIO), permette di integrare il quadro delineato sulla base dei dati estratti dal sotto-corpus 1 del *CoDIt\_com*, le cui dimensioni rispondono a esigenze di bilanciamento e disponibilità di documenti liberamente accessibili.

Il campione di parole composte che sarà analizzato al § 4.2 è stato raccolto attraverso lo spoglio manuale del Lemmario, da cui sono state estratte tutte le parole composte. Tale procedura, i cui limiti sono legati ai criteri adottati per la compilazione del Lemmario in riferimento a questo tipo di parole,<sup>81</sup> ha permesso di raccogliere un campione di composti non certamente esaustivo ma in grado di fornire informazioni sufficienti per delineare un profilo della composizione in italiano antico. La composizione del campione, denominato *Compos\_ita\_TLIO*, è rappresentata nella tabella 9.82

Tipologia di composto	Туре	Esempio
A-N	109	biancospino
N-N	102	capocenso
V-N	84	cavadenti
AVV-PART.PASS	47	bencreato
N-A	46	acquaforte
AVV-PART.PRES	17	malvolente
AVV-A	17	sempreterno

\_

 $<sup>^{79}</sup>$  Il lemmario è liberamente consultabile all'indirizzo  $\frac{\text{http://tlio.ovi.cnr.it/TLIOlemm/index.php}}{\text{risorsa sono descritti in Mosti}}$  (2015).

<sup>&</sup>lt;sup>80</sup> Il corpus OVI, liberamente consultabile all'indirizzo <a href="http://gattoweb.ovi.cnr.it/">http://gattoweb.ovi.cnr.it/</a> (ultima visualizzazione in data 20/03/18) attraverso il software GATTO, rappresenta la maggior base di dati a oggi disponibile per lo studio dell'italiano anteriore al 1375 ed è costituito da 2.318 testi (23.173.538 token). La versione finora lemmatizzata che costituisce la base del Lemmario si compone invece di 1.998 testi (21.868.415 token).

<sup>&</sup>lt;sup>81</sup> In particolare, osserva Mosti (2015: 416) che «costituiscono entrata le parole composte fissate dalla lessicografia moderna in forma univerbata. [...] Diversamente, non costituiscono entrata le sequenze polirematiche presenti nella documentazione in grafia separata, che si schedano in entrata sotto il primo elemento nell'ordine delle priorità fissato dalle norme di redazione».

<sup>&</sup>lt;sup>82</sup> Per indicare le tipologie di composto si utilizzano le seguenti abbreviazioni: A= aggettivo, AVV= avverbio, N= nome, P= preposizione, PART.PASS= participio presente, PART.PRES= participio presente, V= verbo.

Tipologia di composto	Туре	Esempio
AVV-N	15	benevischio
AVV-V	15	malmettere
A-A	12	gallogreco
N-PART.PASS	11	occhibagliato
A-PART.PASS	9	primonato
V-V	9	andirivieni
N-PART.PRES	7	armipotente
N-V	5	caporicciare
AVV-AVV	3	maladagio
A-AVV	2	ognimai
Totale	510	

Tabella 9. Parole composte estratte dal Lemmario del TLIO classificate in base alle categorie di input e ordinate in base al numero di type registrati

Come appare evidente confrontando i dati della tabella 9 con quelli riportati nella tabella 7, le dimensioni del dataset *Compos\_ita\_TLIO* sono notevolmente maggiori rispetto a quelle del campione estratto dal sotto-corpus 1 del CoDIt\_com; i dati estratti dal TLIO verranno analizzati dal punto di vista qualitativo, secondo i parametri di analisi delineati nel prossimo paragrafo, integrando quelli estratti dal CoDIt\_com per i secoli successivi.<sup>83</sup>

### 3.3 Parametri di analisi

In questo paragrafo si descrivono i parametri sulla base dei quali verrà condotta l'analisi dei composti estratti dal CoDIt\_com nel prossimo capitolo: ciascun parametro viene discusso e quindi posto in relazione all'obiettivo complessivo di questo lavoro, ossia fornire un quadro della composizione italiana in prospettiva costruzionista.

Coerentemente con la definizione di Costruzione<sub>COMP</sub> e con la rappresentazione della sua struttura interna (discussa al § 2.2), l'analisi delle forme estratte dal corpus si muoverà su due livelli: il primo formale, in cui si prenderanno in esame le proprietà morfosintattiche e fonologiche di ciascun composto; il secondo funzionale, in cui si osserveranno aspetti semantici e pragmatico-discorsivi. Da intendersi come complementare o meglio trasversale allo studio di questi aspetti, è poi l'analisi della dimensione quantitativa: i dati di frequenza estratti dal CoDIt\_com permetteranno infatti

-

<sup>&</sup>lt;sup>83</sup> Si precisa che il primo sotto-corpus del CoDIt\_com costituisce un "sotto-insieme" dell'OVI, in quanto è costituito da testi che sono presenti anche in questa risorsa, di dimensioni notevolmente più grandi. Nell'analisi diacronica (cfr. § 4.3 e sgg) per ciascuna Costruzione<sub>COMP</sub> si fornirà un quadro della frequenza dei type e dei token delle diverse fasi individuate dai singoli sotto-corpora del CoDIt\_com: per la fase antica, ci si baserà su dati estratti dal primo sotto-corpus (Origini-1375), in modo da poterli confrontare con quelli delle fasi successive.

di integrare la descrizione formale e semantica delle Costruzioni<sub>COMP</sub> con informazioni relative al grado di produttività.

# 3.3.1 Parametri formali

Sul piano formale, ogni Costruzione<sub>COMP</sub> contiene informazioni di tipo fonologico, morfologico e sintattico, ciascuna delle quali può essere analizzata a diversi livelli di astrazione o specificità. In particolare, le proprietà che si è scelto di monitorare nell'analisi sono le seguenti:

- a. categorie di input e output: in linea con la posizione della CxM (cfr. § 2.2), si tengono in considerazione le categorie lessicali dei singoli costituenti e dell'intero composto. Come già osservato da Masini (2016: 30), infatti, «talvolta sembra essere proprio l'informazione categoriale a determinare l'esatta semantica di un processo di formazione di parola». Per indicare la categoria di ciascun elemento si utilizzeranno le seguenti abbreviazioni: A= Aggettivo; AVV= avverbio; P= preposizione; N= nome; PART.PASS= participio passato; PART.PRES= participio presente; V= verbo); nel caso in cui un elemento possa ricoprire più funzioni, le diverse categorie saranno indicate separate da una barra (ad es. A|N nel caso dei composti che possono essere utilizzati come nomi o aggettivi);84
- b. *classe flessiva*: le informazioni relative alla classe flessiva permettono di descrivere più dettagliatamente che tipo di parole sono implicate nel processo di composizione e se esistono delle restrizioni all'interno delle Costruzioni<sub>COMP</sub>. In particolare, nel caso dei composti V-N, è interessante osservare da un punto di vista diacronico la produttività delle coniugazioni nel processo di composizione;
- c. *relazione grammaticale tra i costituenti*: nel formalismo della CxM, la relazione che lega i due costituenti di un composto è generalmente espressa dalla lettera R

\_

<sup>&</sup>lt;sup>84</sup> La classificazione delle unità lessicali in parti del discorso pone non pochi problemi, dal momento che in molti casi uno stesso elemento può ricoprire più funzioni (ad es. un participio presente può fungere da verbo, sostantivo o aggettivo, etc.): la classificazione dei composti rispetto alle categorie di input (ad es. composti [V-N], [N-N], [AVV-PART.PRES], etc.) proposta in questo lavoro è quindi frutto di una semplificazione.

(cfr. § 2.4.2), che può essere specificata in base a parametri semantici o grammaticali.<sup>85</sup> In questo lavoro si è scelto di concentrarsi soltanto sul piano grammaticale, seguendo Bisetto & Scalise (2005), e distinguere tre tipi di relazione: R<sub>MOD</sub> per la relazione di modificazione (ad es. in *camposanto*, *rosa cipria*, *mala-sanità*, etc.); <sup>86</sup> R<sub>SUB</sub> per la relazione subordinativa (ad es. in *capostazione*, *portafoglio*, etc.); R<sub>COORD</sub> per la relazione coordinativa (ad es. in *divano-letto*, *studente-lavoratore*, etc.).

Come già osservato al § 2.2, secondo il formalismo della CxM, le proprietà di natura morfosintattica codificate nelle Costruzioni<sub>COMP</sub> sono indicate attraverso le lettere greche ( $<\alpha>$ ,  $<\beta>$ , etc.) poste in pedice, utilizzate come segnaposto dei *constraints* delle variabili; le singole proprietà vengono poi elencate di seguito alla Costruzione<sub>COMP</sub>, come esemplificato in (11):

- (11) a.  $\langle [[x]_{V\alpha k}[y]_{N\beta i}]_{N\gamma j} \rangle$ 
  - b. grattacielo
  - c.  $\alpha$  = cat. Verbo, lemma *grattare*, coniug. -*are*  $\beta$  = cat. Nome, lemma *cielo*, classe flessiva -o/-i  $\gamma$  = cat. Nome, lemma *grattacielo*, classe flessiva -o/-i

In (11a) è rappresentata la parte formale della Costruzione<sub>COMP</sub> instanziata dal composto *grattacielo*: in (11c) sono elencate le proprietà morfosintattiche dei singoli costituenti e del composto, distinte dalle lettere greche.

Oltre alle proprietà morfosintattiche, il livello forma della Costruzione<sub>COMP</sub> codifica altre informazioni di natura formale che si è scelto di considerare:

d. *grafia*: in italiano, soprattutto nelle fasi più antiche, la grafia non costituisce un criterio valido per individuare le parole composte, dal momento che queste

Ó Seaghdha & Copestake 2008).

<sup>&</sup>lt;sup>85</sup> Come già messo in luce al § 1.3, allo studio delle relazioni semantiche che legano i due costituenti di una parola composta sono stati dedicati numerosi lavori in ambito sia teorico sia computazionale (per un quadro dal punto di vista teorico cfr. Levi 1978; Downing 1977; per le applicazioni in ambito computazionale cfr. Lauer 1995; Girju *et al.* 2005, 2007a,b, 2009; Celli & Nissim 2009; Ó Séaghdha 2007;

 $<sup>^{86}</sup>$  Si noti che la relazione di modificazione ( $R_{\text{\tiny{MOD}}}$ ) include sia la relazione attributiva sia quella appositiva (cfr. Arcodia, Grandi & Montermini 2009).

possono occorre in diverse grafie; tuttavia, la grafia può essere considerata un indizio della coesione tra i due costituenti ed essere rilevante nell'analisi della formazione di una nuova Costruzione<sub>COMP</sub> in prospettiva diacronica. Tre sono i tipi di grafia in cui un composto italiano può occorrere:

- con i due costituenti UNITI graficamente
- con i due costituenti SEPARATI dallo spazio
- con i due costituenti separati dal trattino;87
- e. *erosione fonetica*: l'unione di due parole che porta alla formazione di un composto in italiano può determinare la perdita di materiale fonetico-fonologico del primo costituente. L'erosione può intaccare l'ultima sillaba (come in *granduca* < *gran(de)* + *duca*) o il morfema flessivo (come in *maltempo* < *mal(o)* + *tempo*) dell'elemento di sinistra. In alcuni casi, la riduzione del corpo fonico del primo elemento rappresenta un caso di apocope sillabica o vocalica, che si verifica secondo le stesse restrizioni che operano a livello sintattico (come in gentildonna); in altri, si tratta piuttosto di elisione della vocale finale atona del primo costituente davanti alla vocale iniziale del secondo costituente (come in malaria). L'erosione viene generalmente considerato tra i possibili indizi di grammaticalizzazione (o Costruzionalizzazione; cfr. § 2.5) che porta alla creazione di affissi/affissoidi a partire da costituenti di composti (Munske 2002; Hüning & Booij 2014; Van Goethem 2010). La riduzione del corpo fonico di uno dei due elementi, e in particolare la perdita del morfema flessivo, può infatti favorire la sua decategorializzazione, che, se accompagnata da una specializzazione semantica e un aumento di produttività, è indice della sua trasformazione in morfema legato. Monitorare questo parametro in diacronia permette dunque di ottenere informazioni sui mutamenti che coinvolgono la struttura di una Costruzione COMP;
- f. vocale di raccordo: come si osserverà più approfonditamente al § 4.1, la presenza di una vocale di raccordo (-i- o -o-) alla fine del primo elemento di un composto è un tratto caratteristico della composizione latina (es. lud-i-magister), che in italiano si conserva come elemento residuale ma non del tutto improduttivo. Se infatti

<sup>&</sup>lt;sup>87</sup> Nel corso dell'analisi, il parametro della grafia sarà esaminato distinguendo tra composti "separati" e composti "uniti".

composti come *capinera* o *pettirosso* sono di formazione piuttosto antica, i casi di *altipiano* e *bassipiano* (nati XVII e XIX secolo) testimoniano la presenza (non giustificata dalla flessione) di una -*i*- tra i due costituenti anche in composti più recenti. La presenza di questa vocale è documentata in diversi tipi di composti (principalmente [A-N] e [N-N], etc.), mentre quella di -*o*- è limitata ai composti aggettivali [A-A] del tipo *rossoneri* o *(sindrome) ansioso-depressiva*, la cui origine e il cui sviluppo diacronico sono stati descritti da D'Achille & Grossmann (2009) e Grossmann & Rainer (2009). Nel quadro della CxM, essi possono essere rappresentati come *linking element* tra i costituenti di una Costruzione<sub>COMP</sub>, come mostrato in (12):

(12)  $\langle [[x-o]_{A\alpha k} [y]_{A\beta i}]_{A\gamma j} \leftarrow \rightarrow [SEM_k \text{ in } R_{COORD} \text{ con } SEM_i]_j \rangle$ rossonero, italofrancese, ansioso-depressivo, etc.

La Costruzione<sub>COMP</sub> in (12) instanzia i composti [A-A] dell'italiano i cui costituenti sono legati da una relazione di coordinazione ( $R_{COORD}$ ) e formalmente uniti da una vocale di raccordo -o.

#### 3.3.2 Parametri semantici

Sul piano funzionale, ogni Costruzione<sub>COMP</sub> può contenere tratti semantici, pragmatici e discorsivi. Per quanto riguarda la dimensione semantica, su cui si concentrerà principalmente l'attenzione nel corso dell'analisi, si è già osservato (cfr. § 2.2) che la definizione di Costruzione<sub>COMP</sub> è strettamente legata al principio di composizionalità: una Costruzione<sub>COMP</sub> può dirsi tale solo se il suo significato non è del tutto predicibile da quello delle sue parti, ossia se presenta un certo grado di idiomaticità. L'analisi semantica terrà in considerazione il significato dei singoli costituenti ma si concentrerà su quello veicolato dall'intera costruzione.

Nel formalismo utilizzato dalla CxM, il significato dell'intera  $Costruzione_{COMP}$  è espresso nella parte destra della rappresentazione, di norma attraverso una perifrasi, come in (13): $^{88}$ 

-

<sup>88</sup> Adattato da Masini (2016: 82).

 $< [[x]_{V\alpha k}[y]_{N\beta i}]_{N\gamma i} \leftarrow \rightarrow [STRUMENTO per [SEM_k in R_{SUB} con SEM_i]]_i >$ (13)portafoglio, mangiacassette, lavastoviglie, etc.

La perifrasi utilizzata per rendere la semantica della Costruzione<sub>COMP</sub> dei composti [V-N] indicanti nomi di strumento è "strumento per [SEM<sub>k</sub> in R<sub>SUB</sub> con SEM<sub>i</sub>]", dove SEM<sub>k</sub> e SEM<sub>i</sub> rappresentano genericamente i significati dei due costituenti e R<sub>SUB</sub> la relazione di subordinazione che li lega.

Nel caso in cui uno dei due costituenti sia un verbo, esso può essere indicato nelle perifrasi in modo più specifico con PRED ('predicato'), come esemplificato in (14):

(14) 
$$< [[x]_{V\alpha k}[y]_{N\beta i}]_{N\gamma j} \leftarrow \rightarrow [STRUMENTO per [PRED_k in R_{SUB} con SEM_i]]_j >$$

Come messo in luce da Masini & Audring (in stampa), la rappresentazione semantica delle Costruzionicomp morfologiche costituisce un aspetto meritevole di riflessione, dal momento che non si è ancora arrivati a elaborare un formalismo efficace, condiviso e sufficientemente oggettivo per rappresentare il significato delle Costruzionicomp.

In questo lavoro, si propone di arricchire la descrizione della semantica delle Costruzioni<sub>COMP</sub> attraverso i tipi semantici, sul modello di quanto proposto in Bouillon *et* al. (2012) per l'analisi semantica di un campione di polirematiche dell'italiano con struttura N-P-N (es. ferro da stiro, mulino a vento) e di composti N-N (es. torta gelato, pesce palla). In Bouillon et al. (2012), tali forme sono state annotate secondo un set di tipi semantici (tabella 10) frutto della rivisitazione di un precedente set proposto da Pustejovsky et al. (2008).89

Tag		
entity	information	
abstract_entity	sensation	
human	location	
animate	time period	

<sup>89</sup> In particolare, il set proposto da Pustejovsky et al. (2008: 8) si basa sui tipi più alti della Brandeis Shallow Ontology (BSO) ed è costituita dai seguenti tipi: abstract entity, human, animate, organization, physical object, artifact, event, proposition, information, sensation, location, time period.

Tag
organization
natural
artifact
substance
event
state
proposition

Tabella 10. Tagset semantico utilizzato in Bouillon et al. 2012

I tipi semantici illustrati nella tabella 10 costituiscono il punto di partenza per l'elaborazione del set utilizzato in questo lavoro, esemplificato nella tabella 11.

Tipo ontologico	Tag	Spiegazione	Esempio
informazione	INFO_A	qualsiasi contenuto informativo (testi, canzoni, romanzi, lettere, etc.)	canta <u>storie</u>
evento	EVENT	eventi	buon <u>governo</u>
organizzazione	ORG	istituzioni, organizzazioni (università, scuola, partito, etc.)	capo <u>partito</u>
proprietà	PROPRIETÀ	proprietà (forma, colore, etc.)	petti <u>rosso</u>
sentimento	SENT	sensazioni e sentimenti provati dall'uomo (paura, amore, etc.)	mal <u>umore</u>
stato	STATO	stato	dormi <u>veglia</u>
tempo	TEMPO	periodo di tempo	mezza <u>notte</u>
parte	PARTE		
animale	ANIMALE	animali	<u>pesce</u> spada
artefatto	ARTEFATTO	entità costruite dall'uomo	porta <u>chiavi</u>
persona	PERSONA	persona	<u>capo</u> stazione
gruppo di persone	GR_PERS	gruppo di più persone (classe, squadra,	capo <u>famiglia</u>
parte del corpo	PARTE_CORPO	parte del corpo umano o animale (denti, zampa, capelli, naso, etc.)	crepa <u>cuor</u> e, copri <u>capo</u>
luogo	LUOGO	luoghi all'aperto o al chiuso (ufficio, spiaggia, casa, etc.)	capo <u>stazione,</u> <u>campo</u> santo
oggetto naturale	OGG_NAT	elementi del sistema naturale (flora, montagne, fiumi, frutta e verdura, etc.)	gira <u>sole</u>
sostanza	SOSTANZA	sostanze naturali o artificiali (petrolio, olio, etc.)	<u>acqua</u> vite
cibo	CIBO	alimenti	<u>cacio</u> cavallo, <u>pan</u> forte

Tabella 11. Tagset semantico utilizzato nell'analisi dei composti

Rispetto al set proposto in Bouillon *et al.* (2012), si è scelto di aggiungere alcuni tipi semantici necessari ad annotare un più eterogeneo campione di forme (in particolare, i tipi PROPRIETÀ, GRUPPO DI PERSONE, PARTE DEL CORPO, CIBO)<sup>90</sup> e di non utilizzare il tipo PROPOSITION (che confluisce nel più generico INFORMAZIONE).

Con l'obiettivo di affinare l'analisi semantica delle parole composte del dataset *Compos\_ita* e di migliorarne la rappresentazione in termini formali, si propone di indicare le restrizioni semantiche che limitano la scelta delle variabili e il significato dell'intera Costruzione<sub>COMP</sub> attraverso i tag dei tipi semantici, posti in pedice come esemplificato in (15):

(15)  $< [[porta]_{Vk}[y]_{Ni}]_{Nj} \leftarrow \rightarrow [SEM_k \text{ in } R_{SUB} \text{ con } SEM_{ARTEFATTO\_i}]_{ARTEFATTO\_j} >$ portacenere, portachiavi, portapranzo, etc.

La Costruzione<sub>COMP</sub> in (15) rappresenta i composti [V-N] che formano nomi di strumento (il tag semantico dell'intera Costruzione<sub>COMP</sub> è infatti ARTEFATTO) in cui il primo costituente è il verbo *portare*, che, come già osservato da Ricca (2010: 247), all'interno di questo tipo di composti assume il significato di 'contenere'. Essa costituisce quindi un caso di schema parzialmente specificato in cui il primo slot fisso è occupato dalla stringa *porta-*, tema del verbo *portare*, e il secondo slot è variabile ma soggetto ad alcune restrizioni morfologiche e semantiche: deve essere un sostantivo e appartenere al tipo semantico ARTEFATTO.

### 3.3.3 Parametri quantitativi

Nello studio della formazione delle parole, il concetto di *produttività* ricopre da sempre un ruolo centrale, testimoniato dall'ampia riflessione teorica (cfr. Iacobini 2015) che ha cercato di darne conto. Seguendo van Marle (1985), in questo lavoro si assume la seguente definizione di produttività, adottata anche da Hilpert (2014: 97) in numerosi lavori all'interno del quadro costruzionista:

-

<sup>&</sup>lt;sup>90</sup> Va notato che il tipo CIBO è stato aggiunto nel corso dell'analisi, la quale ha messo in luce come l'italiano crei attraverso la composizione un numero rilevante di nomi che designano alimenti.

The productivity of a schematic morphological construction describes the degree of cognitive ease with which speakers can produce or process new complex words on the basis of that construction.

Secondo questa definizione, la produttività morfologica è intesa come un *gradiente*: una Costruzione<sub>COMP</sub> può avere diversi gradi di produttività o può non essere più produttiva. Come già osservato al § 2.3, infatti, le Costruzioni<sub>COMP</sub> rappresentano generalizzazioni statiche su un set di Costrutti che non necessariamente vengono ancora sfruttate dai parlanti per creare nuove forme. La vitalità delle Costruzioni<sub>COMP</sub> verrà analizzata in questo lavoro attraverso i dati quantitativi estratti dal CoDIt\_com: la struttura del corpus in sezioni di dimensioni bilanciate (ad esclusione della prima, che include il periodo dalle Origini al 1375), permette di calcolare più agevolmente la produttività di una Costruzione<sub>COMP</sub> in diacronia (cfr. gli studi di caso al capitolo 5), evitando di ricorrere a normalizzazioni necessarie nel caso di sotto-corpora non comparabili perché di dimensioni molto diverse.<sup>91</sup>

L'uso di dati estratti da corpora in ambito costruzionista è stato recentemente oggetto di numerosi contributi di natura metodologica (cfr. Hilpert & Gries 2016; Hilpert 2013a), che ne hanno messo in luce le potenzialità. In Hilpert (2013b: 113-154, 2015), la riflessione teorica è applicata a specifici *case studies* di morfologia inglese di cui si misura la produttività in diacronia: nello specifico, i derivati con il suffisso in –*ment* e i composti con participio del tipo *hand-carved*.

Al di fuori del quadro costruzionista, tra gli studi sulla produttività morfologica in diacronia vanno segnalati i lavori di Štichauer (2009a,b, 2015a,b, 2016) dedicati ai suffissi –mento, –zione e –gione e alla composizione verbonominale dell'italiano tra Cinquecento e Ottocento, basati su dati estratti dal corpus LIZ 4.0 (Letteratura Italiana Zanichelli; cfr. Stoppelli & Picchi 2001) e di natura lessicografica. In particolare, assumendo la distinzione di Corbin (1987: 177) tra disponibilità (disponibilité) e redditività (rentabilité), Štichauer mette in luce che il pattern V-N, disponibile fin dal Cinquecento come meccanismo di formazione di parola in italiano, solo in epoche più recenti ha mostrato un aumento di redditività, legato alla diffusione della semantica strumentale nella quale si sono specializzate tali forme.

<sup>&</sup>lt;sup>91</sup> Per un quadro su come calcolare la produttività in corpora diacronici di dimensioni molto diverse si rimanda a Hilpert & Gries (2008, 2016).

Sul modello degli studi di Hilpert, l'analisi quantitativa delle Costruzioni<sub>COMP</sub> oggetto di analisi in questo lavoro prenderà in esame le seguenti misure:<sup>92</sup>

- a. frequenza dei token;
- b. produttività realizzata (realized productivity; frequenza dei type);
- c. produttività potenziale (potential productivity; cfr. Baayen 2005: 244).

Tali misure verranno impiegate per valutare diversi aspetti della produttività delle Costruzioni<sub>COMP</sub>: in particolare, la frequenza dei token fornirà informazioni relative a quanto una determinata Costruzione<sub>COMP</sub> sia radicata nell'uso (*entrenchment*); la frequenza dei type permetterà di valutare in che misura essa è (o è stata) sfruttata dai parlanti per la creazione di nuove forme. Un quadro più approfondito verrà infine fornito dalla produttività potenziale, data dal rapporto tra il numero degli hapax di una forma e quello di tutti i token.

#### 3.4 Struttura dell'analisi

Nel prossimo capitolo verranno analizzati i dati estratti dalle risorse descritte in questo capitolo (*CoDIt\_com*, WFL, Lemmario del TLIO, itWaC), secondo i parametri delineati al § 3.3.

La prima parte dell'analisi si focalizza sull'*italiano antico* (o *italiano delle Origini*; cfr. Renzi & Salvi 2010 per la periodizzazione dell'italiano), termine con cui tradizionalmente si individua la fase compresa tra il 1211, data del primo documento che testimonia l'uso scritto del volgare, e il 1375, data di morte di Boccaccio. In questa parte di analisi si traccia un profilo delle origini della composizione italiana, utile come termine di confronto con le fasi successive, sulla base dei dati quantitativi del primo sotto-corpus del *CoDIt\_com*, integrati con dati di natura qualitativa estratti dal Lemmario del TLIO (cfr. § 3.2.2).

La restante parte dell'analisi si basa esclusivamente sui dati estratti dai sotto-corpora del *CoDIt\_com* e ha l'obiettivo di delineare l'evoluzione delle Costruzioni<sub>COMP</sub> nel corso della storia dell'italiano, fino alla fase contemporanea. L'analisi delle forme si articola su due livelli: nel primo, vengono analizzate le Costruzioni<sub>COMP</sub> astratte, in cui sono

\_

<sup>&</sup>lt;sup>92</sup> In particolare, le prime due misure verranno analizzate per ciascuna Costruzione<sub>COMP</sub>; la produttività potenziale verrà calcolata per casi particolari (ad es. le Costruzioni<sub>COMP</sub> peggiorative, cfr. capitolo 5).

specificate soltanto le categorie dei costituenti (ad es. i composti [N-N], [V-N], [A-A], etc.), e i relativi sotto-schemi; nel secondo, si focalizza l'attenzione sulle Costruzioni<sub>COMP</sub> parzialmente specificate lessicalmente (i *constructional idiom*, cfr. § 2.5; ad es. le Costruzioni<sub>COMP</sub> con *capo*- come primo costituente). In entrambi i livelli di analisi, le Costruzioni<sub>COMP</sub> vengono descritte dal punto di vista formale, semantico e quantitativo, secondo i parametri di analisi discussi al § 3.3.

L'analisi diacronica è seguita da due studi di caso (cfr. §. 5), in cui si focalizza l'attenzione in maniera più approfondita su alcune Costruzioni<sub>COMP</sub> semi-specificate, che offrono l'occasione per indagare i rapporti tra composizione e derivazione nel corso della storia dell'italiano e nella fase contemporanea. Queste Costruzioni<sub>COMP</sub>, poste a metà strada tra composizione e derivazione, vengono analizzate, seguendo i parametri proposti da Van Goethem (2010) (cfr. § 2.5 e 5.2), come casi di *constructional idiom* in cui il primo costituente ha gradualmente assunto le proprietà di prefissoide o prefisso (cfr. Booij & Hündring 2014).

#### 4. ANALISI

Questo capitolo è dedicato all'analisi delle parole composte dell'italiano in ottica costruzionista. L'analisi si articola in tre parti. Nella prima (§ 4.1), si traccia un profilo della composizione latina, di cui si mettono in luce gli elementi di continuità e discontinuità rispetto alla composizione italiana. La seconda parte (§ 4.2) si focalizza sull'italiano antico, di cui si fornisce un quadro sincronico che delinea, da un punto di vista qualitativo, le proprietà morfologiche e semantiche delle Costruzionicomp astratte e semi-specificate. Nella terza parte (§ 4.3) si analizza l'evoluzione delle Costruzionicomp, da un punto di vista quantitativo e qualitativo, mettendone in luce le linee di sviluppo dalla fase antica alla prima metà del XX secolo.

## 4.1 Dalla composizione latina a quella italiana

Questo paragrafo traccia un breve profilo della composizione latina, funzionale a individuare gli elementi di continuità e discontinuità rispetto alla composizione italiana, che verranno discussi al § 4.2.6.

Tutte le lingue romanze presentano, fin dalle fasi più antiche, un meccanismo di formazione delle parole basato prevalentemente sulla derivazione (cfr. Dardano 2009; Tekavčić 1980b); in particolare, in italiano, l'espansione del lessico avviene principalmente attraverso la suffissazione, che rappresenta il meccanismo più sfruttato fin delle Origini. Il legame con la derivazione latina è molto chiaro: l'italiano eredita infatti dal latino gran parte dei suffissi e dei prefissi, a cui nel corso dei secoli se ne aggiungono altri presi in prestito da altre lingue (*in primis*, il francese, come ad es. – *aggio* da - *age*). Il rapporto tra la composizione latina e quella delle lingue romanze sembra invece essere caratterizzato da una forte discontinuità: secondo Dardano (2009), essa è dovuta in larga parte ai mutamenti fonetici sopravvenuti e alla conseguente perdita di trasparenza, che avrebbe di fatto reso opache (e quindi non più analizzabili) molte delle parole composte latine nella fase di transizione tra latino volgare e lingue romanze.

A tale discontinuità non è stata finora dedicata grande attenzione nell'ambito degli studi dedicati alle parole composte del latino (cfr. Oniga 1992; Fruyt 2002 e Brucale 2012 per un recente quadro sullo stato dell'arte), che nella maggior parte dei casi sono di natura teorica e mirano a fornire una descrizione esclusivamente qualitativa delle diverse

tipologie di composto. I dati forniti dal lessico derivazionale *Word Formation Latin* (WFL, cfr. § 3.2.1) permettono di fornire un quadro più approfondito, da un punto di vista sia qualitativo sia quantitativo, necessario in questo lavoro per mettere in luce come, e in che misura, la composizione latina ha esercitato un'influenza su quella dell'italiano. La principale differenza tra i due tipi di composizione, su cui vari studi si sono già soffermati (cfr. Fruyt 2002; Brucale 2012) ma che è utile rievocare in questa sede, riguarda la natura dei costituenti: generalmente, radici o temi (o altri tipi di elementi legati) formano i composti in latino; parole autonome quelli dell'italiano. La composizione latina viene infatti tradizionalmente classificata come tematica (*stembased*; cfr. Brucale 2012: 96): i tre esempi in (16) esemplificano le proprietà strutturali che ricorrono più frequentemente nei composti registrati in WFL.

- (16) a. pisciceps
  pisc-i-ceps
  piscis+LE+capio
  'pescatore'
  - b. cornupeta
    corn-u-pet-a
    cornu+LE+peto
    'che attacca con le corna'
  - c. plenilunium

    plen-i-lun-ium

    plenus+LE+luna+SUFF

    'plenilunio'

In tutti i casi le unità di input sono costituite da forme legate, unite da un elemento di raccordo (*linking element*, LE), generalmente rappresentato dal fono -i- (16a), che può alternarsi con -u- nel caso in cui il secondo costituente inizi con un fono bilabiale o labiodentale (16b) (cfr. Brucale 2012: 97). Come si può osservare da (16a), i costituenti possono occorrere all'interno di un composto in forme diverse da quella della corrispondente forma libera: in *pisciceps* 'pescatore' il secondo costituente è rappresentato da una forma del lemma *capio* 'prendere' che, quando occorre all'interno di un composto, può presentarsi come –*capio* (come in *pignoriscapio* 'atto di prendere in

pegno') o come *-ceps* (come in *vesticeps* 'che è entrato nella pubertà'). Il numero delle possibili forme con cui un lemma può occorrere può essere significativo, come nel caso di *-facio* (vd. oltre, commento alla tabella 14).

I composti latini possono inoltre presentare il suffisso –ium che ne determina la categoria lessicale di sostantivo: casi come quelli di *plenilunium* (1c) possono essere considerati esempi di *parasintesi* (cfr. Brucale 2012: 99; Melloni & Bisetto 2010 sui rapporti tra parasintesi e composizione), dal momento che i processi di composizione e derivazione avvengono simultaneamente e né la base \*plenilun- né la forma \*lunium risultano attestati come forme autonome.

La tabella 12 mostra come si distribuiscono le forme registrate in WFL rispetto all'elemento di raccordo e alla presenza del suffisso –*ium*.

	linking element			suffis	sazione	
	-i-	-u-	- <i>e</i> -	altre vocali o consonanti	-ium	nessun suffisso
Composti (type)	1.285	29	150	333	92	1.705
%	71,5%	1,6%	8,3%	18,5%	5%	95%

Tabella 12. Distribuzione dei composti di WFL (type) rispetto al *linking element* e alla presenza del suffisso -*ium* 

I dati riportati dalla tabella 12 confermano che il latino fa largo uso della vocale di raccordo –*i*– come marca di composizione (71% dei casi). D'altra parte, è significativa la presenza di composti in cui l'elemento di raccordo è la vocale –*e*– (es. *dentefaber* 'dentato', *benesuadus* 'buon consigliere', etc.), non segnalata dai principali studi teorici (cfr. Brucale 2012; Fruyt 2002). Il primo costituente di queste forme può essere un verbo (46%; in particolare *facio*, es. *stupefacio* 'stupire' da *stupeo* 'essere stupito' + *facio*), un aggettivo (22,6%, es. *purefacio* 'rendere puro' da *purus* 'puro' + facio), un sostantivo (20%, es. *dentefaber* 'dentato' da *dens* 'dente' + facio) o un avverbio (18,6%, es. *malevolens* 'malevolo' da *male* 'male' + *volo* 'volere'). La presenza della vocale –*u*– è invece piuttosto limitata e condizionata dal contesto fonetico, ossia la presenza di un fono bilabiale o labiodentale all'inizio del secondo costituente. In circa il 18% dei composti di WFL, il primo costituente termina con una vocale diversa da –*i*–, –*u*– o –*e*–

(ad esempio –o– in *mulomedicus* 'veterinario' o *sacrosantus* 'sacrosanto') o in consonante (ad es. *manceps* 'appaltatore').

Infine, l'uso della parasintesi, per quanto non particolarmente frequente (solo il 5%), costituisce un aspetto importante della composizione latina perché è indizio dello stretto rapporto che composizione e derivazione intrattengono in latino (cfr. Brucale 2012: 99). La presenza del suffisso –ium e delle vocali di raccordo, insieme alla natura legata dei costituenti, contribuiscono a rendere la struttura dei composti latini più coesa rispetto a quelli dell'italiano, in cui tipicamente entrambi i costituenti conservano un certo grado di autonomia sia morfologica (entrambi conservano il proprio morfema flessivo, es capostazione) sia fonologica (entrambi conservano il proprio accento, es. ['kapostat'tsjone]). Come si osserverà meglio al § 4.2, la composizione italiana presenta fin dalle Origini caratteristiche formali diverse, e comuni a tutte le altre lingue romanze, da quelle illustrate dalla tabella 12; la presenza della vocale –i– come linking element tra i due costituenti sopravviverà come elemento residuale soltanto in un numero limitato di forme (es. pettirosso o capinera).

I dati estratti da WFL permettono inoltre di mettere in luce informazioni più specifiche sugli elementi che prendono parte ai processi di formazione dei composti latini. Com'è noto (cfr. Oniga 1992), diversamente dall'italiano e dalle altre lingue romanze, i composti latini presentano generalmente l'ordine modificatore-testa (ad es. *ludimagister* 'maestro di scuola', *ludus* 'scuola' + *magister* 'maestro'). La tabella 13, che illustra più nel dettaglio quali categorie lessicali presentano le unità di input (ossia i costituenti) e di output dei composti in WFL,<sup>93</sup> mette in luce alcuni aspetti utili a comprendere meglio la struttura dei composti latini.

Categorie lessicali	Costituente 1	Costituente 2	Output
A	519	101	1108
I	137	55	67
N	1105	547	513
PR	64	32	53
V	177	1266	273

Tabella 13. Categorie lessicali dei costituenti e dei composti in WFL

\_

<sup>&</sup>lt;sup>93</sup> In questa tabella si segue la classificazione adottata da WFL (cfr. Litta, Passarotti & Culy 2016): A=aggettivo; I= forma invariabile (avverbio, congiunzione); N= nome; PR= pronome; V= verbo.

Per quanto riguarda le categorie delle unità di input, va osservato che i sostantivi e gli aggettivi tendono a occupare la prima posizione, generalmente sede dell'elemento modificatore, mentre i verbi la seconda, che solitamente ospita l'elemento testa. Rispetto alla categoria delle unità di output, i dati estratti da WFL mettono in luce che il latino utilizza la composizione per creare prevalentemente aggettivi, e in secondo luogo nomi e verbi. È inoltre interessante notare il fatto (non messo in luce dai precedenti studi teorici) che la composizione latina permette di creare un discreto numero di forme invariabili, come verumtamen (verum+tamen) 'cionondimeno' o aliquando (alis+quando) 'un tempo', e pronomi, come alteruter (alter+utrum) 'ciascuno dei due' o uterlibet (uter+libet) 'uno qualsiasi dei due'.

Alcune parole si segnalano per la frequenza con cui occorrono, nella forma di elementi legati, all'interno dei composti: nella tabella 14 sono illustrati gli elementi che occorrono più di 15 volte come primo e come secondo costituente.

Costituente 1	Туре	Costituente 2	Туре
multus-	66	-facio	284
aequus-	27	-fero	190
septem-	26	-gero	76
omnis-	24	-pes	64
quinque-	23	-loquor	57
sesqui-	22	-gigno	46
alte/altus-	19	-fluo	28
centum-	17	-colo	28
aurum-	16	-potens	28
bene-	16	-caedo	24
male-	16	-annus	21
		-sono	20
		-vagus	20
		-dico	20
		-capio	18
		-vir	18
		-cano	18

Costituente 1	Туре	Costituente 2	Туре
		-color	17
		-plico	17

Tabella 14. Costituenti più produttivi in WFL (frequenza dei type > 15)

I costituenti che occorrono più frequentemente in prima posizione sono prevalentemente aggettivi qualificativi (come *aequus* 'uguale', *altus* 'alto'), indefiniti (come *multus* 'molto', *omnis* 'ogni') o numerali (*septem* 'sette', *quinque* 'cinque'). Tra questi, *multus* 'molto' (ad es. *multiformis* 'multiforme'), *aequus*- 'uguale' (ad es. *aequimanus* 'ambidestro') e *omnis*- 'ogni' (ad es. *omnigenus* 'di ogni sorta') sopravvivono in italiano come elementi formativi tuttora produttivi nei composti neoclassici, sia di ambito tecnico-specialistico (ad es. *multifattoriale*, *equipotenziale*, *onnipervasivo*) sia usati nella lingua comune (ad es. *multiculturale*, *equivicinanza*, *onnicomprensivo*). <sup>94</sup> Tra gli avverbi, soltanto *bene* 'bene' e *male* 'male' mostrano una discreta produttività: come si nota dai dati riportati nella tabella 15, in alcuni casi i due avverbi si legano alle stesse parole, dando vita a coppie di vocaboli dal significato opposto, ad es. *malevolens* 'maldisposto' vs *benevolens* 'bendisposto'.

Composti con bene-	Composti con <i>male</i> -
	-
benedico 'benedire'	maledico 'maledire'
benefacio 'agire rettamente'	malefacio 'nuocere'
benemorius 'che ha buone qualità morali'	
beneolentia 'buon odore'	
beneplaceo 'aggradare'	
beneplacitum 'beneplacito'	
benesuadus 'buon consigliere'	malesuadus 'cattivo consigliere'
benevolens 'bendisposto'	malevolens 'maldisposto'
benevolus 'benevolo'	malevolus 'malevolo'
	maletractatio 'maltrattamento'
	malecastus 'impudico'
	malefaber 'perfido'

Tabella 15. Composti con bene- e male- come primi costituenti in WFL

<sup>&</sup>lt;sup>94</sup> Tutti gli esempi sono tratti dal dizionario online Il Nuovo De Mauro, liberamente consultabile all'indirizzo: <a href="https://dizionario.internazionale.it/">https://dizionario.internazionale.it/</a> (ultima visualizzazione in data 20/04/2018): in particolare, i lemmi *multifattoriale*, *equipotenziale*, *onnipervasivo* sono classificati dalla marca TS 'tecnicospecialistici'; *multiculturale*, *equivicinanza*, *onnicomprensivo* dalla marca CO 'lingua comune'.

Le forme elencate nella tabella 15 non presentano le vocali di raccordo -i- o -u-: i due avverbi, bene e male, si legano direttamente al secondo costituente, l'elemento testa, a cui apportano un contributo semantico in alcuni casi di natura valutativa, rispettivamente positiva/migliorativa il primo (come in beneolentia 'buon odore'), negativa/peggiorativa il secondo (come in malecastus 'impudico'). Come si osserverà più nel dettaglio al § 4.2.2, l'impiego di questi avverbi come costituenti di composti con testa a destra (es. malfare, maladoperato, malvolentieri, benavventuroso, benessere, etc.) frequente nella composizione italiana delle Origini e costituisce un elemento di continuità tra i due tipi di composizione.

Infine, l'unico sostantivo che occorre nella posizione di modificatore in più di 15 composti è *aurum* 'oro': esso può combinarsi con elementi verbali (come in *aurifer* 'che porta o produce oro') o sostantivali (come in *aurifodina* 'miniera d'oro).

Gli elementi che occupano più frequentemente la seconda posizione, in cui di norma si trova l'elemento testa, sono prevalentemente di natura verbale: in particolare, *facio* 'fare' e *fero* 'portare' costituiscono i verbi più produttivi. Diversamente da *fero*, che occorre nei composti soltanto nelle forme – *fer* (come in *cornifer* 'cornuto, che porta le corna') e molto raramente – *lator* (es. *legislator*<sub>N</sub> 'legislatore'), il verbo *facio* può occorrere come – *fex* (es. *artifex*<sub>N</sub> 'artigiano'), – *ficium* (es. *mellificium*<sub>N</sub> 'produzione di miele'), – *facio* (es. *calefacio*<sub>V</sub> 'riscaldare'), – *fico* (es. *fructifico*<sub>V</sub> 'fruttificare'), – *ficus* (es. *mellificus*<sub>A</sub> 'che produce miele'). Negli ultimi tre casi, – *facio*, – *fico* e – *ficus* svolgono una funzione causativa, comune anche a *facio* quando usato in collocazioni come *sacra facere* 'rendere sacro' (cfr. Fruyt 2002: 270).

In italiano, alcuni di questi costituenti sopravvivono come elementi formativi nei composti del tipo neoclassico, sia di ambito tecnico-specialistico sia della lingua comune. Il formativo *-fero*, ad esempio, è utilizzato nella creazione di numerosi composti agentivo-strumentali con il significato di 'che porta, che produce' (es. *flammifer*<sub>A</sub> 'che produce fiame, infiammato'): esso si combina a un sostantivo che occupa la prima posizione e che è generalmente interpretabile come l'argomento dell'elemento verbale

<sup>96</sup> Tutte le forme sono registrate nel Lemmario del TLIO (<a href="http://tlio.ovi.cnr.it/TLIOlemm/index.php">http://tlio.ovi.cnr.it/TLIOlemm/index.php</a> ultima visualizzazione in data 22/04/2018).

 $<sup>^{95}</sup>$  Per un'analisi semantica più approfondita dei composti latini con  $\it male\mbox{-}$  si rimanda al § 5.3.

<sup>&</sup>lt;sup>97</sup> Allo sviluppo diacronico delle Costruzioni valutative con *mal(e)*- è dedicato lo studio di caso del capitolo 5, a cui si rimanda per un'analisi più approfondita.

(es. *sonnifero* 'che provoca, induce il sonno'; cfr. Iacobini 2004: 89). Nonostante un discreto numero di queste forme siano usate stabilmente nella lingua comune (ad es. *frigorifero*, *calorifero*, *fiammifero*), tale elemento formativo è scarsamente usato per la creazione di nuove forme.<sup>98</sup>

Tra gli elementi connessi al verbo *facio*, alcuni hanno assunto lo status di affissi in italiano. In particolare, il formativo *-fic-* (cfr. Brucale & Mocciaro 2015) è usato produttivamente nella lingua comune come suffisso verbalizzante (come nel caso di *plastificare* da *plastica*) o aggettivale (come in *immaginifico*). Il formativo *-ficium*, che nei composti latini esprimeva il significato di 'arte di produrre x' (es. *panificium*<sub>N</sub> 'atto di preparare il pane, panificazione'), si è conservato in italiano come *-ficio*, cosiderato un vero e proprio suffisso da Tekavčić (1980a: 164), un «secondo elemento di origine latina» da Serianni (1989: 667-668), un «suffissoide» nella terminologia di Migliorini (1960: 718). Nella fase attuale, esso è usato in forme come *mobilificio*, *panificio*, *colorificio* con il significato di 'luogo dove si lavora/produce x' (cfr. Lo Duca 2004a: 238-239): dal punto di vista semantico, si è quindi verificato uno slittamento da un significato eventivo (astratto) a uno locativo (concreto). Sono di più recente creazione le forme come *diplomificio*, *divertimentificio*, *abortificio*, in cui vi è un'interpretazione ironica: nel caso di *diplomificio*, ad esempio, si allude a un certo tipo di scuole private considerate come un luogo dove si producono diplomi.

Il sostantivo che occupa più frequentemente la seconda posizione è *pes* 'piede': esso costituisce l'elemento testa di aggettivi e può occorrere come *–pes* (es. *cornipes*<sub>A</sub> 'dal piede di corno') e *–pedus* (es. *aequipedus*<sub>A</sub> 'dai lati equivalenti'). Anche in questo caso, il costituente si conserva in italiano come formativo neoclassico: *–pede* ricorre, con il significato di 'relativo ai piedi, che ha piedi', in forme generalmente di ambito tecnicospecialistico, come *latipede* 'che ha i piedi larghi', e in alcuni termini diffusi anche nella lingua comune, come *velocipede* 'tipo di bicicletta'.<sup>99</sup>

Un ulteriore aspetto che i dati estratti da WFL permettono di mettere in luce riguarda la presenza di participi presenti in composti aggettivali del tipo *arcitenens* 'che tiene l'arco, arciere' o *altitonans* 'altitonante', costituiti da un sostantivo o un aggettivo seguiti da un

<sup>&</sup>lt;sup>98</sup> Il repertorio dei neologismi di Treccani (liberamente consultabile all'indirizzo: <a href="http://www.treccani.it/magazine/lingua\_italiana/neologismi/searchNeologismi.jsp">http://www.treccani.it/magazine/lingua\_italiana/neologismi/searchNeologismi.jsp</a>) registra la sola forma sondaggifero (ultima visualizzazione 22/04/2018).

<sup>&</sup>lt;sup>99</sup> De Mauro (2004), s.v. -pede (https://dizionario.internazionale.it/, ultima visualizzazione in data 22/04/2018).

verbo al participio presente. Dal punto di vista semantico, queste forme si riferiscono generalmente a referenti umani o a divinità (es. *breviloquens* 'che parla in modo conciso', *tridentipotens* 'signore del tridente, appellativo di Nettuno'), con alcune eccezioni (ad es. *velivolans* 'che avanza con le vele' indica un tipo di imbarcazione).

Come già illustrato nella tabella 14 (vd. terza colonna), la forma di participio più frequente è –potens 'potente' da possum 'potere, essere in grado di', che occorre in 28 composti (es. omnipotens 'onnipotente', armipotens 'potente nelle armi, valoroso', etc.). Oltre a queste forme, WFL registra altri 47 composti costituiti da un elemento participiale in seconda posizione, come illustrato nella tabella 16.

Participio	Туре	Esempi	
-boans (boo 'risuonare, eccheggiare')	1	altiboans 'altisonante'	
-candicans (candico 'essere biancastro')	1	margariticandidans	
-ciens ( <i>cieo</i> 'muovere')	1	omniciens 'che muove tutto'	
-colens ( <i>colo</i> 'coltivare')	1	silvicolens	
-comans (como 'avere una lunga chioma')	5	auricomans 'dai capelli d'oro'	
-creans ( <i>creo</i> 'creare')	1	omnicreans 'che genera tutto'	
-crepitans (crepito 'crepitare')	1	aericrepitans	
-distans (disto 'distare')	1	aequidistans 'equidistante'	
-ficens (facio 'fare')	1	munificens 'munificente'	
-ferens ( <i>fero</i> 'portare')	1	frugiferens 'fertile di messi'	
-fluens (fluo 'scorrere')	1	mellifluens	
-loquens (loquor 'parlare')	4	suaviloquens 'che parla soavemente'	
-medens ( <i>medeor</i> 'sanare')	1	omnimedens	
-micans (mico 'scintillare')	1	stellimicans 'che rifulge di stelle'	
-olens (oleo 'odorare')	1	graveolens 'che ha un cattivo odore'	
-opinans (opinor 'ritenere')	1	necopinans 'che non se lo aspetta'	
-pendens (pendo 'pendere')	1	lanipendens	
-pens ( <i>pendo</i> 'pendere')	1	libripens 'colui che sorregge la bilancia durante le vendite'	
-pollens (polleo 'avere potere')	3	vinipollens 'potente grazie al vino'	
-potens (possum 'potre')	28	mundipotens 'che domina il mondo'	
-putens ( <i>puteo</i> 'puzzare')	1	nariputens	
-sonans (sono 'suonare')	1	raucisonans	
-sulcans (sulco 'solcare')	1	crispisulcans 'che procede tortuosamente'	
-tenens (teneo 'tenere')	6	signitenens 'cosparso di stelle'	
-tonans (tono 'tuonare')	1	altitonans 'altitonante'	
-tuens (tueor 'vedere')	2	aedituens 'che sorveglia il tempio'	
-videns (video 'vedere')	1	omnividens 'onnivedente'	
-vigens (vigeo 'avere vigore')	1	gaudivigens	
-volens (volo 'volere')	2	benevolens 'benevolente'	

Participio	Туре	Esempi
-volans (volo 'volare')	2	altivolans 'che vola in alto'
Totale	75	

Tabella 16. Participi che occupano la seconda posizione nei composti estratti da WFL

Ad eccezione di *–potens, –comans, –loquens, –pollens* e *–tenens,* che occorrono in almeno tre forme, gli altri participi sono utilizzati in maniera occasionale all'interno di un composto. I costituenti che si trovano in prima posizione sono generalmente bisillabici (fanno eccezione soltanto cinque forme: margariticanticans, cortinipotens 'signore del treppiede, appellativo di Apollo', sagittipotens 'signore delle saette, il Sagittario', sapientipotens 'onniscente', tridentipotens 'signore del tridente, appellativo di Nettuno') e terminano con la vocale di raccordo -i- (tranne in quattro casi: graveolens 'maleodorante', necopinans 'che non se lo aspetta, incauto', benevolens 'benvolente' e malevolens 'malevolo'). Si tratta prevalentemente di sostantivi (45 type: 60%, es. stella 'stella' in *stellimicans* 'che rifulge di stelle'), aggettivi (26 type: 34.6%, es. aequus 'uguale' in aequidistans 'equidistante') e alcuni avverbi (4 type: 5.4%, es. blande 'lunsinghevolmente' in blandiloquens 'che parla lusighevolmente'). La presenza di composti costituiti da un elemento iniziale (di natura sostantivale, aggettivale o avverbiale) e un participio presente, generalmente utilizzati per indicare referenti umani, costituisce un elemento di continuità tra latino e italiano: come si osserverà nel dettaglio più avanti (§ 4.2.2 e 4.3.2), questo tipo di forme è attestato dalle fasi più antiche a quella attuale senza soluzione di continuità. 100

Delineati alcuni aspetti quantitativi e qualitativi utili a tracciare un profilo della composizione latina sulla base dei dati estratti da WFL, nel seguente paragrafo si analizzano le parole composte dell'italiano antico in ottica costruzionista; gli elementi di continuità e discontinuità tra composizione latina e italiana verranno discussi al § 4.2.6.

\_

 $<sup>^{100}</sup>$  La disponibilità di alcuni participi presenti a entrare in composizione con morfemi lessicali per designare referenti umani è testimoniata da forme, di attestazione sia antica sia recente, come altoparlante (1927), benestante (1498), chiaroveggente (1709), nullafacente (1868), tossicodipendente (1978), viandante (XIII secolo) (cfr. Lo Duca 2004b: 359). Nella fase attuale, la Costruzione  $[x_N\text{-}dipendente]_{A|N}$  (es. cibodipendente, rischiodipendente, cellularedipendente, etc.) mostra una significativa produttività: il repertorio di neologismi Treccani (online) ne registra 19 forme (http://www.treccani.it/magazine/lingua italiana/neologismi/searchNeologismi.jsp ultima visualizzazione in data 23/04/2018).

#### 4.2 Italiano antico

L'analisi della composizione in italiano antico che si propone in questo paragrafo si basa sul dataset Compos\_ita\_TLIO, estratto secondo la metodologia descritta al § 3.2.2 dal Lemmario del Tesoro della Lingua Italiana delle Origini. L'analisi è articolata su due livelli: nel primo (§ 4.2.1-4) si descrivono le proprietà formali e semantiche delle Costruzioni<sub>COMP</sub> astratte, in cui sono specificate le relazioni grammaticali tra i costituenti e la categoria lessicale delle unità di input e di quella di output; nel secondo (§ 4.2.5), si analizzano alcuni casi di Costruzionicom semi-specificate.

Come già messo in luce al § 1.3, seguendo Arcodia & Basciano (2018), le Costruzioni<sub>COMP</sub> astratte dell'italiano vengono classificate in 'famiglie di Costruzioni' (families of *constructions*)<sup>101</sup> costituite da schemi e sotto-schemi. Tale classificazione riprende alcuni criteri adottati in Bisetto & Scalise (2005), in particolare: 1) il tipo di relazione grammaticale tra i due costituenti (subordinativa, coordinativa, di modificazione); 2) la presenza di un elemento testa, che porta alla distinzione tra composti endocentrici e composti esocentrici. Un ulteriore parametro adottato in questa analisi è la posizione della testa: si distinguerà quindi tra composti (endocentrici) con testa a sinistra e con testa a destra.

Sulla base dei criteri qui delineati e dei precedenti studi dedicati alla composizione in ambito costruzionista (cfr. Booij 2010; Masini & Audring, in stampa) si individuano quindi quattro famiglie di Costruzioni  $_{\text{COMP}}$ :

- 1. Costruzione<sub>COMP</sub> dei composti endocentrici con testa a sinistra;
- 2. Costruzione<sub>COMP</sub> dei composti endocentrici con testa a destra;
- 3. Costruzione<sub>COMP</sub> dei composti coordinativi endocentrici;
- 4. Costruzione<sub>COMP</sub> dei composti esocentrici.

Come verrà mostrato nei seguenti paragrafi, la suddivisione delle Costruzionicome astratte in queste quattro classi permette di rendere conto delle regolarità e sotto-regolarità dei composti attestati in italiano antico. Ogni classe di Costruzionicomp presenta specifiche caratteristiche formali e semantiche e si compone di schemi e sotto-schemi ordinati gerarchicamente a diversi livelli, dal livello più astratto (es. Costruzione COMP) dei composti

<sup>&</sup>lt;sup>101</sup> Con il termine 'famiglia di Costruzioni', utilizzato da Arcodia & Basciano (2018) e ripreso in questo lavoro, si intende 'classi' di Costruzioni.

con testa a sinistra) a quelli via via più specifici (es. Costruzione<sub>COMP</sub> dei composti subordinativi con testa a sinistra, Costruzione<sub>COMP</sub> dei composti [N-N] subordinativi con testa a sinistra, etc.), fino alle Costruzioni parzialmente specificate lessicalmente (es. Costruzione<sub>COMP</sub> [capo-x], es. capofamiglia, caporione, etc.).<sup>102</sup> Nei seguenti paragrafi, ciascuna classe di Costruzioni<sub>COMP</sub> verrà analizzata nei diversi livelli di astrazione.

# 4.2.1 La Costruzione<sub>COMP</sub> dei composti con testa a sinistra

I composti con testa a sinistra dell'italiano antico sono costituiti da due elementi, uno dei quali (quello di sinistra) determina le proprietà morfologiche e semantiche dell'intero composto; i due costituenti possono essere legati da una relazione di subordinazione ( $R_{SUB}$ ) o di modificazione ( $R_{MOD}$ ). Nel primo caso, i due costituenti sono sempre nomi (es. *capofamiglia*); nel secondo, possono essere un nome e un aggettivo (es. *camposanto*) o due nomi di cui il secondo in funzione di modificatore (es. *pescespada*).

Nella figura 22 è rappresentata la Gerarchia di eredità di questa famiglia di Costruzioni $_{COMP}$  con i suoi sotto-schemi. In particolare, in (a) è illustrata la Costruzione $_{COMP}$  più astratta dei composti con testa a sinistra: la presenza/posizione della testa è rappresentata dalla co-indicizzazione (attraverso le lettere  $\alpha$  e X, che indicano rispettivamente le proprietà morfologiche e la categoria lessicale) del primo costituente e dell'intera Costruzione $_{COMP}$ .

Tale Costruzione<sub>COMP</sub> istanzia due schemi a seconda della relazione che intercorre tra i due costituenti: in (b) è rappresentata la Costruzione<sub>COMP</sub> dei composti subordinativi, in cui sia la relazione ( $R_{SUB}$ ) sia la categoria lessicale dei costituenti (in questo caso, entrambi nomi) sono specificate; in (c) è rappresentato lo schema dei composti i cui costituenti sono legati da una relazione di modificazione ( $R_{MOD}$ ). A questo livello della Gerarchia, è specificata soltanto la categoria lessicale della testa (e dell'intero composto: in questo caso, nome), dal momento che il modificatore può essere sia un aggettivo sia un nome. Da tale schema dipendono infatti due ulteriori sotto-schemi (rappresentati in (d) e in (e)), in cui anche le categorie lessicali dei modificatori sono specificate.

 $<sup>^{102}</sup>$  Nei seguenti paragrafi, per riferirsi alle singole Costruzioni $_{\text{COMP}}$  verrà utilizzata la formula  $[X-Y]_{Z_{-}R_{\nu}}$ , in cui  $X, Y \in Z$  indicano le categorie di input e di output e R indica il tipo di relazione grammaticale (es.  $[N-N]_{N_{-}SUB}$  'composti costituiti da due nomi, legati da una relazione di subordinazione, e con categoria di output nome').

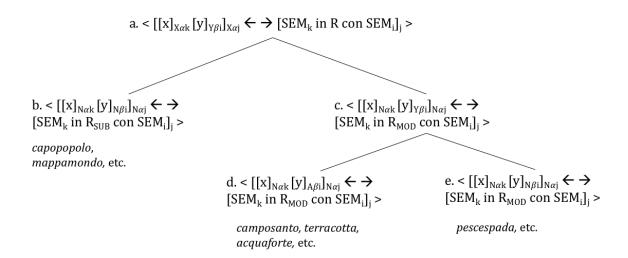


Figura 21. Gerarchia di eredità della Costruzione composti con testa a sinistra dell'italiano antico

Dal punto di vista quantitativo, i composti del dataset *Compos\_ita\_TLIO* si distribuiscono in maniera disomogenea rispetto ai vari sotto-schemi (tabella 17).

[N-N] <sub>N_SUB</sub>	$[N-A]_{N\_{MOD}}$	$[N-N]_{N\_{MOD}}$	Totale
26	26	7	59

Tabella 17. Composti con testa a sinistra dell'italiano antico: distribuzione rispetto ai sotto-schemi (type)

I dati riportati nella tabella 17 mostrano che i due sotto-schemi più istanziati sono [N-N] $_{N\_SUB}$  (es. *capofamiglia, piedistallo*) $^{103}$  e [N-A] $_{N\_MOD}$  (es. *roccaforte, terraferma*).

Lo schema [N-N]<sub>N\_SUB</sub> è istanziato da composti di origine sia latina sia italiana. In particolare, 8 composti hanno origine dall'univerbazione di espressioni tardolatine di natura sintagmatica: si tratta di *capelvenere* 'pianta medicinale' dal lat. tardo *capillum veneris*; *capopurgio* 'medicamento per purgare il capo' dal lat. medievale *caput purgium*; *caporione* 'a Roma, capo di rione' dal lat. medievale *caput regionum*; *figliuomini* 'essere mostruosi discesi da Caino' dal lat. cristiano *filiae hominum*; *finimondo* 'fine del mondo terreno che prelude al Giudizio Universale' dal lat. ecclesiastico *finem mundi*; *gommaedera* 'gommoresina prodotta dal fusto dell'edera' dal lat. medievale *gummi* 

<sup>&</sup>lt;sup>103</sup> In virtù di ragioni formali (la presenza di una preposizione) e semantiche (la relazione tra i due costituenti è parzialmente espressa dalla preposizione), le forme come *capodanno* o *malincuore*, nate dall'univerbazione di due sostantivi legati da una relazione di subordinazione e una preposizione (es. *capo d'anno, mal in cuore*) non sono considerate istanziazioni di questa Costruzione<sub>COMP</sub>.

hedere; mappamondo 'rappresentazione cartografica dell'inetro territorio terrestre' dal lat. medievale mappa mundi; palmacristo 'pianta erbacea avente i tubercoli simili al palmo di una mano' dal lat. medievale palma christi; porcospino 'nome comune dell'istrice' dal latino parlato \*porcospinum.<sup>104</sup> In questi casi, l'espressione latina entra in italiano assumendo la regolare struttura dei composti N-N che prevede la giustapposizione di due sostantivi, senza marche esplicite di composizione (elementi di raccordo, etc.).

Dei restanti 18 composti, che possono essere considerati creazioni autonome dell'italiano, 10 presentano il sostantivo *capo* in prima posizione;<sup>105</sup> gli altri 6 sono *caciocavallo*, *calendagosto*, *calendimaggio*, *pannolino*, *piedistallo*, *solleone*, esemplificati nella tabella 18.<sup>106</sup>

Composti [N-N] <sub>N_SUB</sub>	Esempi	
caciocavallo 'formaggio a pasta dura ed elastica'	«[] fa' di avere li duecento fiorini: per di qui otto dì io farò una investita di carne salata e di <i>cacio cavallo</i> , ch'io credo guadagnare sì che io ti farò buona parte» (Sacchetti, <i>Trecentonovelle</i> )	
calendimaggio 'primo giorno di maggio'	«avvenne che la sera di <i>calendimaggio</i> MCCC [], faccendosi in su la plaza di santa Trinita un gran ballo di donne []» (Boccaccio, <i>Esposizioni</i> )	
calendagosto 'primo giorno di agosto'	«Allora santo Alessandro ordinò che si facesse questa festa di Pietro in Vincoli in calendagosto, []» (Leggenda Aurea)	
madreselva 'lo stesso che caprifoglio'	«Recipe sugo di <i>madreselva</i> con acqua rosata e usalo nello occhio, è ispertissimo» (Piero Ubertino da Brescia, <i>Ricettario</i> )	
pannolino 'panno fatto con tela di lino'	«Che li corpi dei morti de la città s'invollano in <i>panno lino</i> o vero stamegna, <sup>107</sup> et portinsi in cassa» (Gangalandi, <i>Costituto del comune di Siena volgarizzato</i> )	
piedistallo 'piede che serve d'appoggio'	«[] è solo la infabile e vana loro speransa, la quale di nome Ventura si chiama e di fatto in non fermo e non stabile <i>piedistallo</i> e incollonata» (Teperto, <i>Lettera in</i> <i>prosa</i> )	

\_

<sup>&</sup>lt;sup>104</sup> TLIO, s.v. capelvenere, capopurgio, caporione, capricorno, fedecommissario, figliuomini, finimondo, gommaedera, mappamondo, palmacristo, porcospino.

<sup>&</sup>lt;sup>105</sup> All'analisi dei composti con *capo*- è dedicato il § 4.2.5.2.

<sup>&</sup>lt;sup>106</sup> Tutti gli esempi sono tratti dal corpus OVI.

<sup>&</sup>lt;sup>107</sup> Per *stamegna* si intende un 'tipo di tessuto di tela piuttosto rado, usato specialmente per rivestire i defunti' (TLIO, s.v. *stamigna*).

Composti [N-N] <sub>N_SUB</sub>	Esempi
solleone 'il sole dell'estate'	«[] e talora cominciava colle sue mani acciocché ciascuno lavorasse al caldo del <i>solleone</i> , sì gli prese una febbre continua, onde cadde forte malato» (Giovanni Villani, <i>Cronica</i> )

Tabella 18. Composti [N-N]<sub>N\_SUB</sub> con testa a sinistra di origine non latina dell'italiano antico

I composti esemplificati nella tabella 18 sono costituiti da due sostantivi legati da una relazione di dipendenza sintattica, diversa a seconda del singolo composto e di non sempre facile individuazione. Tra le forme di etimo incerto vi è *caciocavallo*, in cui il modificatore *cavallo* richiamerebbe l'uso di appendere questo tipo di formaggio *a cavallo* di una trave per farlo essiccare. In tutti gli altri casi, tra i due costituenti sembra esserci una relazione di specificazione, di norma espressa attraverso la preposizione *di*, come ad es. in *pannolino* 'panno di lino', *calendagosto* 'calende di agosto' o solleone 'sole del Leone (segno dello Zodiaco)'. Dal punto di vista semantico, non si individuano regolarità o restrizioni nella creazione di questo tipo di composti, che sono usati per indicare sia artefatti (es. *piedistallo*), sia periodi di tempo (es. *calendimaggio*), sia oggetti naturali (es. *solleone*).

Lo schema in (d) della figura 22 è istanziato da 24 composti [N-A]<sub>N\_MOD</sub> con testa a sinistra,<sup>109</sup> come *camposanto* o *roccaforte*: stabilire quali di queste forme hanno origine dalla cristallizzazione e univerbazione di espressioni sintagmatiche e quali nascono da un vero e proprio processo di composizione è nella maggior parte dei casi piuttosto complesso. A differenza di quanto osservato nel caso dei composti [N-N] subordinativi, queste forme possono essere considerate creazioni autonome dell'italiano, ad eccezione di due casi, *fiengreco* (dal lat. *faenum graecum* 'pianta erbacea') e *rosmarino* (dal lat. *ros marinus* 'rugiada di mare'). Dal punto di vista semantico, i composti si distribuiscono rispetto al tipo semantico come illustrato nella tabella 19.

<sup>&</sup>lt;sup>108</sup> Dizionario Treccani Online, s.v. caciocavallo (ultima visualizzazione in data 03/07/18).

<sup>109</sup> Come già osservato (cfr. § 2.4.3), ascrivere in modo univoco tali forme all'interno della categoria di composto non è sempre facile, dal momento che i costituenti di questi composti presentano l'ordine sintattico non marcato Nome-Aggettivo. Le forme che si analizzano in questa sede sono state estratte dal Lemmario del TLIO, all'interno del quale sono state classificate parole composte e lemmatizzate come tali. Per i criteri seguiti dai redattori del TLIO nella lemmatizzazione si rimanda a Mosti (2015).

Tipo semantico	Composti [N-A] <sub>N_MOD</sub> endocentrici	
ARTEFATTO PERSONA	ferrovecchio	
LUOGO	camposanto, roccaforte, terraferma, terrasanta	
OGG_NATURALE	codacavallina, fiengreco, gommarabica, rosamarina, rosmarino, spigaceltica, spinacervina	
CIBO	ficosecco, melacotogna, melocotogno, pomogranato, peracotogna	
SOSTANZA	acquaforte, orofino, pecegreca, terracotta	
STATO	fiammasalsa	
EVENTO	ventipiovolo	

Tabella 19. Composti  $[N-A]_{N\_MOD}$  endocentrici: distribuzione rispetto al tipo semantico dell'elemento testa

La distribuzione delle forme rispetto al tipo semantico non è affatto omogenea. In particolare, i tipi semantici più rappresentati sono OGG\_NAT e CIBO, che individuano rispettivamente gli elementi del sistema naturale e gli alimenti, come ad esempio piante (es. rosamarina o rosmarino) o frutti (es. melacotogna, pomogranato), LUOGO (camposanto, roccaforte, terraferma, terrasanta) e SOSTANZA (es. pecegreca, terracotta). Vi è poi la forma fiammasalsa, 'prurito che viene in pelle, cagionato da umor salso'<sup>110</sup> (tipo semantico STATO) e il composto ventipiovolo 'vento che fa piovere', costituito dal nome vento (tipo semantico EVENTO) e piovolo, aggettivo derivato da piova (cfr. anche piovale 'che produce/porta la pioggia'), come esemplificato in (17).

(17) «[...] e noi Greci, menati da' venti per gli nimichevoli mari, sostenemmo le saette, la notte, e' *ventipiovoli*, l'ira del cielo e quella del mare [...]» (Simintendi, *Metamorfosi d'Ovidio volgarizzate*)

Infine il composto *ferrovecchio* è caratterizzato da una polisemia agentivo (tipo semantico PERSONA)/strumentale (tipo semantico ARTEFATTO), come illustrato dagli esempi in (18).

-

 $<sup>^{110}</sup>$  Vocabolario degli Accademici della Crusca (quarta edizione, consultata online al link  $\underline{\text{http://www.lessicografia.it/}}$  il 2/06/18), s.v.  $\emph{fiammasalsa}.$ 

(18) a. «Èvi chi vende taglieri e scodelle , / chi vende liscio ed èvi 'l calzaiuolo / che vende calze e cappelline belle, / e 'l fabbro e 'l *ferrovecchio* e 'l chiavaiuolo [...]» (Pucci, *Rime*)

b. «E che neuna persona, di qualunque condictione sia, ardisca o presumma andare gridando a fregi vechi, ariento, oro, perle né pietre overo gemme o *ferrovechio* a vender per la città di Lucha, borghi o sobborghi o per le sei milia [...]» (*Statuti lucchesi*)

In (18a), *ferrovecchio* indica chi lavora o commercia ferro usato; in (18b), esso indica un insieme di oggetti metallici usati. La concentrazione di questo tipo di composti essenzialmente in tre tipi semantici può essere rappresentata in termini costruzionisti attraverso tre sotto-schemi della Costruzione<sub>COMP</sub> [N-A]<sub>N\_MOD</sub> in cui è specificato il tipo semantico dell'elemento testa (e dell'intero composto), come illustrato nella figura 23.

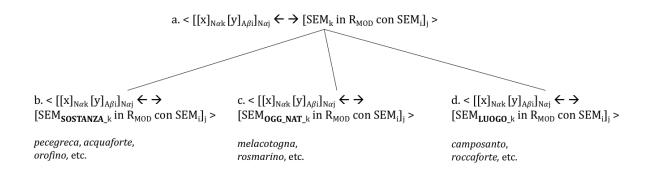


Figura 22. Sotto-schemi della Costruzione dei composti [N-A]<sub>N\_MOD</sub> con testa a sinistra

Nella Gerarchia di eredità rappresentata nella figura 23, la Costruzione  $_{COMP}$  dei composti  $[N-A]_{N\_MOD}$  con testa a sinistra domina tre sotto-schemi, in ciascuno dei quali vi è una restrizione sul tipo semantico dell'elemento testa (rappresentata dai nomi dei tipi in pedice, es.  $SEM_{LUOGO}$  per i composti come *camposanto* o *roccaforte*). I composti che non rientrano in queste tre categorie semantiche costituiscono istanziazioni dello schema più astratto rappresentato in (a).

Per quanto riguarda gli aggettivi che svolgono la funzione di modificatori di questi composti, essi si distribuiscono equamente tra qualificativi (*cotto, degno, fermo, fino, forte, magno, salso* 'salato', *santo, secco, vecchio*) e relazionali (*cavallino, greco, arabico,* 

cotogno, granato, cervino, piovolo). Dall'analisi dei dati estratti dal dataset  $Compos\_ita\_TLIO$  emerge quindi come la  $Costruzione_{COMP}$  dei composti  $[N-A]_{N\_MOD}$  sia sfruttata soprattutto per la classificazione di elementi del sistema naturale (piante, frutti, etc.) o di sostanze.

Infine, lo schema in (e) della figura 22 è istanziato da composti costituiti da due nomi giustapposti, di cui il primo in funzione di testa e il secondo in funzione di modificatore (es. *pescespada*). Il numero di queste forme all'interno del dataset *Compos\_ita\_TLIO* è piuttosto limitato: si tratta, in particolare, di 3 forme (*arcamensa*, *pesceporco*, *pescespada*) esemplificate nei seguenti contesti:

(19) a. «Per ciascuno soppedanio, cassa grande, archapredula, arcamensa, bocte et tina [...]» (Statuti della Colletta del comune di Orvieto)

b. «De' *pesci porci*, che cavano la terra sotto l'acqua; porci sono una maniera di pesci che cavano tutto giorno la terra sotto l'acqua in fondo di mare per chiedere la loro vivanda, sì come fanno i nostri porcelli» (Anonimo, *Tesoro di Brunetto Latini volgarizzato*)

c. «E di valor portate maggior pregio / che non fa il buon rubin fra l'altre pietre, / e di franchezza più che 'l *pesce spada* infra li pesci [...]» (Anonimo, *Mare Amoroso*)

Nell'esempio (19a), *arcamensa* indica una 'madia di grandi dimensioni contenente tutto ciò che serve per l'alimentazione': <sup>111</sup> il composto è costituito dal sostantivo *arca*, che in italiano antico indicava un 'cassone destinato specialmente alla conservazione di derrate agricole o di panni', <sup>112</sup> e dal modificatore *mensa*, nel significato di 'tavola imbandita'. <sup>113</sup> In (19b) e (19c) sono esemplificati due composti indicanti razze di pesci, classificati attraverso la giustapposizione di due sostantivi in funzione di modificatori: nel primo caso, *spada* richiama alla forma dell'animale; nel secondo *porco* allude all'abitudine, propria anche dei maiali, di questi pesci a scavare la terra sott'acqua.

Vi sono poi quattro forme in cui l'interpretazione della relazione tra i due costituenti non è del tutto chiara ma che si è scelto di considerare tra i composti [N-N]<sub>N MOD</sub>: si tratta dei

<sup>&</sup>lt;sup>111</sup> TLIO, s.v. arcamensa.

<sup>&</sup>lt;sup>112</sup> TLIO, s.v. *arca*.

<sup>113</sup> TLIO, s.v. mensa.

composti *gattaconiglio, melarancia, mellomela, pomarancio*. Nel primo caso si tratta di una creazione estemporanea che occorre soltanto in una novella di Franco Sacchetti, la cui unica attestazione è riportata in (20):

(20) «Essendo adunque il piovano, messer Dolcibene, e altri, fra l'altre vivande recandosi la crosta della *gattaconiglio*, ella fu buona che messer Dolcibene ne mangiò più che niuno» (Sacchetti, *Trecentonovelle*)

Come si può dedurre dall'esempio riportato in (20), la *gattaconiglio* è un 'gatto cucinato e fatto passare per coniglio':<sup>114</sup> il sostantivo *coniglio* sembrerebbe quindi modificare la testa *gatto*. Le forme *melarancia* 'altro nome dell'arancia' (mela+arancia), *mellomela* (melone?+mela, etimo incerto) e *pomarancio* 'frutto dell'albero dell'arancio' (pomo+arancio) sono nomi che designano specie di frutti:<sup>115</sup> in tutti i casi il primo costituente designa la macro-specie e il secondo ne specifica il tipo particolare. In virtù del fatto che il secondo costituente sembra classificare il primo, al pari di *pesce spada*, si è scelto di includere queste forme tra i composti [N-N]<sub>N\_MOD</sub>; tuttavia, la classificazione di queste forme rimane incerta.

L'analisi dei composti con testa a sinistra attestati nel dataset *Compos\_ita\_TLIO* ha permesso di mettere in luce almeno due aspetti che caratterizzano questo tipo di composizione in italiano antico. Il primo riguarda i composti che istanziano la Costruzione<sub>COMP</sub> [N-N]<sub>N\_SUB</sub> rappresentata in (b) nella figura 22: come è stato osservato, la maggior parte di queste forme ha origine dalla univerbazione di espressioni latine di origine sintagmatica (es. *mappamondo* da *mappa mundi, caporione* da *caput regionum*), che costituiscono il modello per la creazione di composti [N-N]<sub>N\_SUB</sub> nativi (es. *piedistallo, pannolino, capopopolo*). Questo tipo di Costruzioni<sub>COMP</sub> sembra quindi ancora legato al latino, che fornisce materiale lessicale sotto forma di prestiti o calchi, e poco produttivo autonomamente. Un'eccezione è rappresentata dai numerosi composti, di origine prevalentemente nativa, che presentano il nome *capo* come primo costituente, la cui descrizione sarà approfondita al § 4.2.5.2.

Un secondo elemento emerso dall'analisi riguarda i composti con relazione di modificazione ( $R_{MOD}$ ): il ristretto numero di composti  $[N-N]_{N\_MOD}$  sembra indicare che la

<sup>&</sup>lt;sup>114</sup> TLIO, s.v. *gattaconiglio*.

<sup>&</sup>lt;sup>115</sup> TLIO, s.v. melarancia, mellomela, pomarancio.

giustapposizione di due sostantivi, di cui il secondo in funzione di modificatore, costituisce un meccanismo poco produttivo, ma comunque già disponibile, in italiano antico. Infine, l'analisi dei tipi semantici ha messo in luce che la combinazione di un nome e un aggettivo (composti  $[N-A]_{N\_MOD}$ ), seppur maggiormente sfruttata, è utilizzata prevalentemente per la creazione di termini indicanti specie di piante, frutti, sostanze.

## 4.2.2 La Costruzione<sub>COMP</sub> dei composti con testa a destra

I composti con testa a destra dell'italiano antico sono costituiti da due elementi, uno dei quali (quello di destra) determina le proprietà morfologiche e semantiche dell'intero composto; i due costituenti possono essere legati da una relazione di subordinazione ( $R_{\text{SUB}}$ ) o di modificazione ( $R_{\text{MOD}}$ ).

Come mostrato dalla figura 24, la Gerarchia di eredità di questa famiglia di Costruzioni<sub>COMP</sub> è decisamente più articolata rispetto a quella dei composti con testa a sinistra.

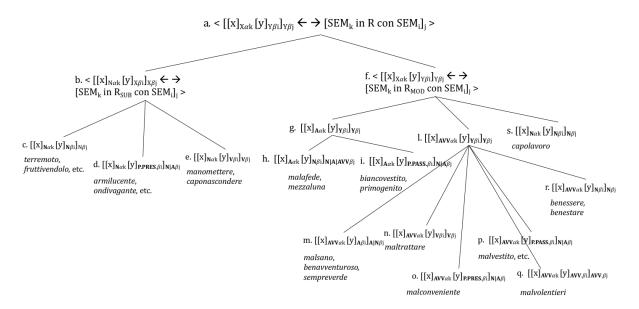


Figura 23. Gerarchia di eredità della Costruzione composti con testa a destra dell'italiano antico

In particolare, in (a) è illustrata la Costruzione<sub>COMP</sub> più astratta dei composti con testa a destra: la presenza/posizione della testa è rappresentata dalla co-indicizzazione

(attraverso le lettere  $\beta$  e X, che indicano rispettivamente le proprietà morfologiche e la categoria lessicale) del secondo costituente e dell'intera Costruzione<sub>COMP</sub>.

Tale Costruzione<sub>COMP</sub> istanzia due schemi a seconda della relazione che intercorre tra i due costituenti. In (b) è rappresentata la Costruzione<sub>COMP</sub> dei composti subordinativi, in cui sono specificate sia la relazione tra i due costituenti ( $R_{SUB}$ ) sia la categoria lessicale del primo costituente (che, nelle forse attestate, è sempre un nome). A sua volta, tale schema è istanziato da quattro sotto-schemi, a seconda della categoria lessicale dell'elemento testa: in (c) la testa è un sostantivo (es. *terremoto*, *linoseme*); in (d) è un participio presente (es. *luogotenente*, *armipotente*); in (e) è un verbo (es. *manomettere*, *capolevare*).

In (f) è invece rappresentato lo schema dei composti i cui costituenti sono legati da una relazione di modificazione (R<sub>MOD</sub>): in questo caso, soltanto la relazione tra costituenti è specificata, dal momento che entrambi i costituenti possono avere più di una categoria lessicale. Tale schema è infatti a sua volta istanziato da due sotto-schemi. In particolare, lo schema in (g) è costituito da un aggettivo più una base non specificata, che può essere un nome (es. *maladefe, mezzaluna*), come rappresentato in (h), o un participio passato (es. *biancovestito, secondogenito*) come rappresentato in (i). Lo schema in (l) è costituito da un avverbio più una base variabile che può presentare diverse categorie lessicali: aggettivo (es. *sempreverde, benavventuroso*; cfr. (m)), verbo (es. *malvolere, chiarefare*; cfr. (n)), participio presente (es. *maldicente, altitonante*; cfr. (o)), participio passato (es. *benavviato, malavveduto*; cfr. (p)), avverbio (es. *oggimai* 'ormai', *malvolentieri*; cfr. (q)) e nome (es. *benevischio, benestare*; cfr. (r)).

I composti subordinativi che istanziano la Costruzione<sub>COMP</sub> rappresentata in (b) si distribuiscono in maniera piuttosto disomogenea nei quattro sottoschemi, come illustrato nella tabella 20.

[N-N] <sub>N_SUB</sub>	$[N-PART.PRES]_{N A\_SUB}$	$[N-PART.PASS]_{N A\_SUB}$	$[N-V]_{V\_SUB}$	Tot
38	4	3	5	50

Tabella 20 I composti subordinativi con testa a destra: distribuzione dei type rispetto ai sotto-schemi individuati dalla categoria lessicale del secondo costituente

Il sotto-schema istanziato da composti subordinativi  $[N-N]_{N\_SUB}$  come *pescivendolo* o *terremoto* appare decisamente più produttivo rispetto agli altri: questo insieme di forme

è costituito da 7 composti che presentano il sostantivo *capo*- come primo costituente non testa (es. capopurgio, capomorbo)<sup>116</sup> e un discreto numero di composti (o espressioni multiparola) ereditati dal latino. In particolare, si tratta di animavversione da animadversionem; arcobalestro da arcuballista; capopurgio da caputpurgio; caprifico da caprificus; caprifoglio da caprifolium; corpobalsamo da carpobalsamum; fedecommissario da fideicommissarius; manubalestro da manuballista; nottetempo da noctistempore; panicuocolo da panicoculus; pescivendolo da piscivendolus; senatoconsulto da senatus consultum; terremoto da terrae motum. L'origine latina di tali forme è rivelata, oltre che dall'ordine dei costituenti tipico della composizione latina, anche dalla presenza della -icome vocale finale del primo costituente (es. caprifoglio, pescivendolo). Sul modello di pescivendolo e panicuocolo si formano cavallivendolo, lanovendolo, mercivendola, pannotendolo, terratremolo. Dal punto di vista semantico, i composti con -vendolo indicano sempre un individuo che vende una determinata merce; anche la forma pannotendolo si riferisce a un agente (tipo semantico PERSONA), in particolare un 'tenditore di lana'; l'unica eccezione è rappresentata da terratremolo 'terremoto' che, come esemplificato in (21), indica un EVENTO.

(21) «Et ne lo tempo suo quasi tutte le provincie de lo munno se lli revellaro et poi foro molti *terratremoli* et crescimento de fulmina et molte tempestate e pestilentie et molte locuste [...]» (Anonimo, *Storie de Troia e de Roma*)

In tutti i casi il primo costituente costituisce l'oggetto diretto del verbo da cui è derivato, attraverso il suffisso –olo, il secondo costituente. Queste forme sono infatti costituite da un sostantivo e un nome deverbale derivato attraverso il suffisso –olo: dal momento che né le forme \*vendolo, \*tendolo e \*tremolo né \*pescivende o \*pannotende risultano attestate come forme autonome, la formazione di questo tipo di composti può essere interpretata in chiave costruzionista come un caso di unificazione simultanea di due schemi (cfr. § 2.3, vd. il caso di derattizzare): quello di composizione (a) e quello di suffissazione in –olo (b), come illustrato nella figura 25.

<sup>&</sup>lt;sup>116</sup> Per un'analisi più approfondita dei composti con *capo*- si rimanda al § 4.2.5.2.

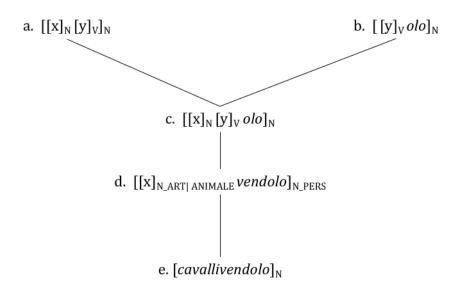


Figura 24 Unificazione di schemi che permette la formazione dei composti del tipo pescivendolo

L'Unificazione rappresentata nella figura 25 porta alla formazione di un nuovo schema (c), la cui produttività è indipendente da quella dei due schemi unificati: tale schema è istanziato da composti come *lanovendolo*, *pannotendolo* o *terratremolo*, costituiti da un nome e un derivato in –olo. All'interno di questo gruppo di forme è possibile individuare un sottoinsieme di composti che indicano un referente umano e in cui il secondo costituente è –vendolo: tali forme costituiscono un sotto-schema (rappresentato in (d)), formato da un primo elemento variabile ma soggetto a restrizioni semantiche (il tipo semantico può infatti essere ARTEFATTO o ANIMALE, ossia prodotti che possono essere venduti) e un secondo costituente fisso, ossia –vendolo. I composti di questo tipo presentano come tipo semantico sempre PERSONA, in quanto indicano referenti umani. I restanti composti [N-N]<sub>N\_SUB</sub> con testa a destra creati autonomamente dall'italiano sono riportati ed esemplificati nella tabella 21. 117

Composti [N-N] <sub>N_SUB</sub> con testa a destra	Esempio			
artemagico 'esperto di magia, augure'	«[] alcuno, nominato Simone mago			
	perch'era artemagico, della grazia di cristo gli			
	chiese, cioè ch'egli il facesse d'i suo' apostoli,			
	promettendogli che della sua ricchezza assai			

-

<sup>&</sup>lt;sup>117</sup> Quando non diveramente indicato, le definizioni riportate nella tabella 10 e nelle seguenti sono tratte dal TLIO, ad eccezione della definizione di *salgemma*, che è tratta dalla prima edizione del Vocabolario della Crusca (consultabile in rete al seguente link: <a href="http://www.lessicografia.it">http://www.lessicografia.it</a> ultima visualizzazione 18/06/18). Gli esempi riportati sono stati estratti dall'OVI.

Composti [N-N] <sub>N_SUB</sub> con testa a destra	Esempio
	gli darebbe» (Jacopo Alighieri, <i>Chiose all'Inferno</i> )
becchipuzzola 'upupa'	«[] Progne diventò rondine e Filomena usignuolo, e Iti diventò pettorosso e Tereo becchipuzzola» (Boccaccio, Teseida delle nozze d'Emilia. Chiose)
terrafine 'confine'	«[] onde per tutte queste cose fu mandato a terrafine a Leon sopra Rodano, là onde era stato natìo, acciò che vi morisse in vergogna de la sua gente» (Anonimo, Leggenda aurea)
castrocampo 'campo fortificato con un muro di cinta'	«Castrocampo, intendi per campo di gente d'arme, afforzato di fosso e di steccato, o dell'uno, o dell'altro» (Romuleo volg.)
fioripanno pezzo di stoffa usato per pulire le stoviglie'	«[] quasi paniculus, idest <i>fluripannu</i> , quo terguntur scutelle» (Senisio, <i>Declarus</i> )
gattafodero 'fodero confezionato con pelle di gatto'	«Egli avea una foggia alta presso a una spanna, con uno <i>gattafodero</i> che parea una pelle d'orsa, tanto era morbido, e avea uno collaretto a un suo guarnaccione []» (sacchetti, <i>Trecentonovelle</i> )
linoseme 'tipo di graniglia costituita dai semi del lino'	«Diedi a Stefano Giovanni, per lui a Nuto Donati, che cti reghoe ad Onbrone, che nde conperasti <i>lino seme</i> , dì soprascricto» (Anonimo, <i>Libro di entrate e uscite di Mino tesoriere</i> )
manroverso 'colpo dato con la parte conversa della mano'	«Guarda che tratta è stata un <i>manroverso</i> » (Sacchetti, <i>Rime</i> )
salgemma 'spezie di sale minerale, lucido, e traparente'	«prende [] di pepe bianco cinque granella, e di salgemma due denari, e pestalo bene insieme sottilemente []» (Anonimo, Trattato del governo delle malattie e guarigioni de' falconi)

Tabella 21 Composti  $[N-N]_{N\_SUB}$  di origine nativa attestati in italiano antico

I composti esemplificati nella tabella 21, e i relativi costituenti, presentano tipi semantici piuttosto variegati; in particolare, i tipi semantici in output attestati sono i seguenti: PERSONA (es. *artemagico*), ARTEFATTO (es. *gattafodero*, *fioripanno*), OGG\_NATURALE (es. *linoseme*, *salgemma*), ANIMALE (es. *becchipuzzola*), EVENTO (es. *manroverso*) e LUOGO (es. *castrocampo*, *terrafine*). Per queste forme non è quindi possibile individuare sotto-

schemi che presentino restrizioni sul tipo semantico delle unità di input né in quelle di output. Il ristretto numero di questo tipo di composti induce a ritenere che, come per i composti [N-N]<sub>N\_SUB</sub> con testa a destra, si tratti di un meccanismo poco produttivo.

Il sotto-schema  $[N-PART.PRES]_{N|A}$  istanzia 5 composti subordinativi costituiti da un nome e un participio presente, riportati ed esemplificati nella tabella 22.

Composti [N-PART.PRES] <sub>N A</sub>	Esempi
armipotente 'valoroso in guerra, potente in armi'	«O armipotente combattrice, che se' sopra le battaglie, vergine Minerva, rompi con la tua mano la lancia di questo ladrone di Troia» (Guido da Pisa, Fatti di Enea)
luogotenente 'comandante'	«[] e 'l detto Pipino fu fatto patrice di Roma, cioè <i>luogotenente</i> d'imperio e padre della repubblica de' Romani» (Villani, <i>Cronica</i> )
panchiedente 'mendicante'	«Già tanto non sarebbe pro' né saggio / Ched i' non ne facesse panchiedente» (Dante Alighieri, Fiore)
verodicente 'che dice la verità'	«Seneca dice, che non falla a' grandi se non i <i>verdicenti</i> , che delli maldicenti, e de' menzogneri hanno elli grande mercato intorno di loro» (Bencivenni Zucchero, <i>Esposizione del Paternostro</i> )
verosimigliante 'verosimile'	«Lo verosimigliante argomento è di quelle cose che o sono usate di venir spesso, o delle cose che l'uomo pensa che siano []» (Anonimo, Tesoro di Brunetto Latini volgarizzato da Bono Giamboni)

Tabella 22 Composti [N-PART.PRES]<sub>N|A</sub> attestati in italiano antico

Come illustrato dagli esempi, il nome può costituire l'oggetto diretto del participio presente: è il caso di *luogotenente*, *panchiedente*, *verodicente*; nel caso di *armipotente* si tratta di un composto di origine latina (*armipotens* 'potente nelle armi'). Come già osservato al § 4.1, la combinazione di un nome e di un participio presente è largamente attestata in latino classico e funge da modello per la creazione di composti italiani nativi con testa a destra. Dal punto di vista semantico, come in latino, i composti attestati nel dataset *Compos\_ita\_TLIO* indicano sempre referenti umani; lo schema che istanzia tali composti può quindi essere arricchito dell'informazione relativa al tipo semantico come illustrato in (22):

(22) 
$$< [[x]_{N\alpha k}[y]_{PART.PRES.\beta i}]_{N|A\beta j} \leftarrow \rightarrow [[SEM_k \text{ in } R_{SUB} \text{ con } SEM_i]_{PERSONA\_j} >$$

La Costruzione<sub>COMP</sub> [N-PART.PASS]<sub>N|A</sub> non sembra essere particolarmente produttiva, dal momento che istanzia soltanto 3 composti, costituiti da un nome e un participio passato: si tratta di *fedecommesso*, *occhibagliato*, *plebeuscito*, esemplificati in (23):

(23) a. «E quisto etiandio s'entenda se cotale occidente fosse sostituito a l'uciso directamente overo per fedecommesso [...]» (Statuto del Comune del popolo di Perugia)

b. «Quando pensati al domandar segondo, / se vu' siti zentili o *plebesciti*, / non siti vu' smariti / e vergognosi a far cotal bataglia?» (Antonio da Ferrara, *Rime*)

c. «[...] una luce smisurata circunfulse me e' compagni, ch'erano con meco, sicchè tutti cademmo a terra *occhibagliati* per quello isplendore» (Cavalca, *Atti degli apostoli volgarizzati*)

Nel primo caso (23a), *fedecommesso* ha origine dal composto latino *fideicommissum* 'affidato alla lealtà', attestato in italiano antico con un significato diverso: 'disposizione testamentaria che vincola l'erede diretto a conservare i beni ricevuti e a trasmetterli alla sua morte ad altro erede designato dal testatore'. In (23b), il composto *plebeuscito*, costituito dal sostantivo *plebe* e il participio passato *uscito*, indica 'chi per nascita appartiene al popolo'; in (23c), *occhibagliato* si riferisce a una persona accecata da una luce eccessiva (lett. 'abbagliato nei occhi'). La forma deriva dal verbo *occhibagliare* 'restare abbagliato', istanziazione insieme ad altre 4 forme (vd. tabella 23) della Costruzione<sub>COMP</sub> [N-V]<sub>V</sub>.

Composti [N-V] <sub>V</sub>	Esempi
capolevare 'cadere rovinosamente'	«[] per la qual cosa <i>capolevando</i> questa tavola con lui insieme se n'andò quindi giuso» (Boccaccio, <i>Decameron</i> )

<sup>118</sup> TLIO, s.v. fedecommesso.

\_\_

Composti [N-V] <sub>V</sub>	Esempi
caporicciare 'provare raccapriccio'	«[] et accapricciare è levare li capelli ritti, come avviene per paura; cioè caporicciare; e però si dice: Io ebbi un grande raccapriccio; cioè uno arricciamento de' capelli del capo, che significa la paura» (Francesco di Bartolo dal Buti, Commento all'Inferno)»
crocesegnare 'investire della missione della crociata'	«Questo santo Bernardo, dell'Ordine di Chiaravalle, crocesegnò Currado III imperatore, il quale creò dugeto navi di pellegrini, trapassò il mare, anni Domini 1145» (Anonimo, Ottimo commento della Commedia)
manomettere 'metter mano, offendere, guastare'	«Là ov'elli cadde, era la pressa molto grande de' chavalieri, che molto lo calpestaro e manumisero» (Binduccio dello Scelto, <i>La</i> Storia di Troia)
occhibagliare 'restare abbagliato'	«Subitamente lo circonfuse una luce dal cielo, per la quale <i>occhibaglioe</i> e cadde in terra» (Cavalca, <i>Atti degli Apostoli volgarizzati</i> )

Tabella 23 Composti che istanziano la Costruzione [N-V]<sub>V</sub> attestati in italiano antico

Come illustrato nella tabella 23, la creazione di verbi attraverso l'incorporazione di un nome a un elemento verbale (largamente attestata in latino, cfr. § 4.1) trova riscontro anche in italiano antico, seppur in pochi casi. Tra le forme attestate nel dataset *Compos\_ita\_TLIO*, soltanto una ha origine latina: si tratta di *manomettere*, dal latino *manumittere* 'affrancare uno schiavo' (lett. 'mandar(lo) libero dalla potestà (*manus*) del padrone'). In italiano antico, il composto è stato interpretato come 'metter mano' e, nella maggior parte dei contesti, assume un'interpretazione peggiorativa di 'offendere, guastare', come illustrato nell'esempio riportato nella tabella 23, in cui *manomettere* vale 'danneggiare, offendere'. Sul modello di tale forma, l'italiano ha creato autonomamente altri composti in cui il nome è un partecipante del verbo in seconda posizione (es. *caporicciare*). 121

<sup>&</sup>lt;sup>119</sup> Sul fenomeno dell'incorporazione si rimanda all'ampia trattazione fornita da Mithun (1984, 1986). La creazione di verbi composti N-V non è estranea alla formazione delle parole delle lingue romanze, soprattutto nelle fasi più antiche: si veda Klingebiel (1989) per un quadro su questo tipo di composti in latino e in quattro lingue romanze (francese, occitano, catalano, spagnolo).

<sup>&</sup>lt;sup>120</sup> Dizionario Treccani online, s.v. manomettere (ultima visualizzazione in data 18/06/18).

<sup>&</sup>lt;sup>121</sup> Per l'analisi dei due composti con *capo*- si rimanda al § 4.2.5.2.

Sulla base dei dati analizzati finora, si è osservato che, tra i sotto-schemi della Costruzione<sub>COMP</sub> dei composti subordinativi con testa a destra, il più produttivo è costituito da due sostantivi giustapposti (ossia [N-N]<sub>N\_SUB</sub>): tale schema è istanziato da un significativo numero di composti di origine latina (es. *terremoto*), adattati fonomorfologicamente all'italiano, e da un più limitato numero di composti di origine nativa. Si è inoltre osservato che, all'interno di questo insieme di forme, è possibile individuare un ulteriore sotto-schema costituito da un nome e un derivato deverbale attraverso il suffisso –olo (es. *pescivendolo*, *pannotendolo*), rappresentabile in ottica costruzionista come il risultato dell'Unificazione di due schemi: tale sotto-schema ha origine a partire da due forme latine, *pescivendolo* e *panicuocolo*, sul modello delle quali l'italiano ha creato autonomamente altre forme, la maggior parte delle quali presenta – *vendolo* come secondo costituente e alcune restrizioni relative al tipo semantico del primo costituente e dell'intera forma. Gli altri sotto-schemi ([N-PART.PRES]<sub>N|A</sub>, [N-PART.PASS]<sub>N|A</sub>, [N-V]<sub>V</sub>), attestati già in latino classico, sono istanziati da un numero di forme piuttosto limitato ma sufficiente a testimoniarne la vitalità in italiano antico.

La Costruzione<sub>COMP</sub> dei composti con testa a destra i cui costituenti sono legati da una relazione di modificazione, rappresentata in (f) nella figura 24, è istanziata da due sottoschemi in cui il primo slot è sempre riempito da un aggettivo o un avverbio e il secondo (ossia l'elemento testa) è variabile.

Nel primo caso si hanno le Costruzioni<sub>COMP</sub> [A-N]<sub>N|A|AVV</sub> e [A-PART.PASS]<sub>N|A</sub>. La Costruzione<sub>COMP</sub> [A-N]<sub>N|AVV</sub> permette la creazione di parole composte come *biancospino*, *maltempo*, *tuttavolta*, *vanagloria*; il dataset *Compos\_ita\_TLIO* contiene 119 composti che istanziano questa Costruzione. Come illustrato nella figura 26, in cui è rappresentata la Gerarchia di eredità, la Costruzione<sub>COMP</sub> [A-N]<sub>N|AVV</sub> (in (b)) costituisce un'instanziazione della più astratta Costruzione<sub>COMP</sub> dei composti italiani con testa a destra (in (a)) ed è a sua volta instanziata da due sotto-schemi a seconda della categoria di output del composto. In particolare, (c) rappresenta lo schema dei composti [A-N]<sub>N|A</sub> (es. *buonafede*, *biancospino*, etc.); in (g) è rappresentato lo schema dei composti [A-N]<sub>AVV</sub> (es. *tuttora*, *tuttodì*, *ognora*, etc.). Da entrambi gli schemi dipendono alcuni sotto-schemi parzialmente specificati lessicalmente (rappresentati in (d), (e), (f), (h)), costituiti da uno slot specificato e uno variabile.

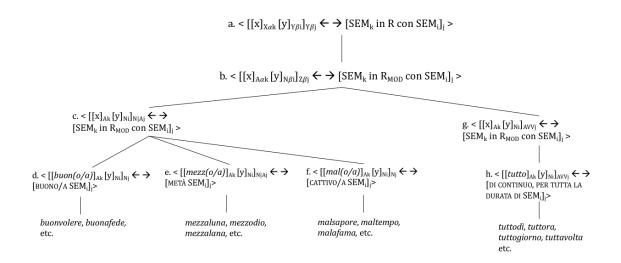


Figura 25. Gerarchia di eredità delle Costruzioni<sub>COMP</sub> [A-N]<sub>A|N|AVV</sub>

Nello schema dei composti nominali [A-N]<sub>N|A</sub>, rappresentato in (c), la prima posizione è occupata da un numero ristretto di aggettivi: in particolare, come illustrato nel grafico della figura 27, si tratta degli aggettivi malo (46 type), buono (17 type), mezzo (11 type), cento (4 type) e gentil(e) (3 type).

Nel caso dei composti [A-N]<sub>N|A</sub> che presentano *cento*- e *gentil(e)*- come primo costituente, non si tratta di creazioni autonome dell'italiano ma prevalentemente di forme di origine esogena. In particolare, i composti con *cento*- hanno origine latina: *centofoglie* da *centifolius* 'tipo di pianta che prende nome dall'elevanto numero di foglie', *centogemino* da *centumgeminus* 'dotato di cento braccia', *centopiedi* da *centipes* 'insetto simile a un verme, con molte zampe; appellativo di Giove'.<sup>122</sup> L'unica forma di cui non è attestato un corrispettivo latino è *centopezze* 'trippa ricavata dall'omaso dei ruminanti, centopelli', in cui l'aggettivo *cento* si combina con il sostantivo *pezza*, attestato in italiano antico con il significato di 'quantità di pesce o carne tagliata'.<sup>123</sup> I composti con l'aggettivo *gentil(e)*-sono invece prestiti dal francese: *gentiluomo* da *gentilhomme*, *gentildama* da *gentildame*. Sul modello di queste forme, l'italiano ha creato autonomamente il composto *gentildonna* e un insieme di composti costituiti da un aggettivo e il sostantivo *uomo* (o,

114

<sup>122</sup> TLIO, s.v. centofoglie, centogemino, centopiedi.

<sup>123</sup> TLIO, s.v. centopezze, pezza.

in un caso, donna), quali giovanuomo, nobiluomo, poveruomo, piccioluomo, valentuomo, nobildonna.

Le forme che presentano gli aggettivi qualificativi gradabili *buono-, malo-* e *mezzo-* rappresentano il 70% dei composti che istanziano la Costruzione<sub>COMP</sub> [A-N]<sub>N|A</sub> e possono essere considerate istanziazioni di Costruzioni semi-specificate [*buon(o/a)-x*], [*mal(o/a)-x*] e [*mezz(o/a)-x*] (rappresentate in (d), (e) e (f) nella figura 27).<sup>124</sup>

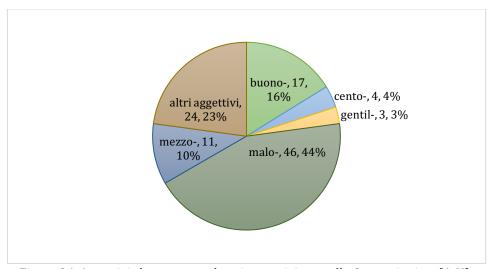


Figura 26. Aggettivi che occupano la prima posizione nelle Costruzioni<sub>comp</sub> [A-N]<sub>N</sub>

I composti che presentano un aggettivo diverso da *buono*, *cento*, *gentile*, *malo*, *mezzo* rappresentano il 23% del totale dei composti [A-N]<sub>N|A</sub>: essi costituiscono istanziazioni della Costruzione<sub>COMP</sub> rappresentata in (c) nella Gerarchia di eredità (cfr. figura 26), in cui sono specificate soltanto le categorie di input e di output. Analizzando più nel dettaglio questo insieme di forme si può notare che più della metà (16 type, 67%) è di origine esogena: in particolare, due composti sono prestiti dal francese (*biancomangiare* da *blanc mangier* e *granmangiere* da *grant mangier* 'pietanza cotta nel latte') e 10 hanno origine latina (*altisono* da *altisonus*, *biancospino* da *alba spina*, *cinquefoglie* da *quinquefolium*, *duramadre* da *dura mater* 'membrana esterna e dura del cervello', *longanime* da *longanimis*, *millefoglie* da *millefolium*, *piamadre* da *pia mater* 'membrana

\_

L'analisi diacronica della Costruzione<sub>COMP</sub> [mal(o/a)-x] costituisce uno studio di caso di questo lavoro, per il quale si rimanda al § 5.4 e seguenti. Alle Costruzioni<sub>COMP</sub> [buon(o/a)-x] e [mezz(o/a)-x] in italiano antico è dedicato il § 4.2.5.1.

interna che avvolge il cervello', *primiscrineo* da *primiscrinius* 'capo di un segretario della cancelleria imperiale romana', *primotempo* da *primo ver* 'primavera', *settiforme* da *septiformis, vanagloria* da *vana gloria*). La restante parte di questo insieme di forme è costituito dai 6 composti con *-uomo/-donna* in seconda posizione (es. *nobiluomo, valentuomo*, etc., creati, come si è già osservato, sul modello di *gentiluomo* e *gentildonna*) e dalle forme *granduca* e *bellegote* (epiteto dal significato non accertato).<sup>125</sup>

Oltre che sostantivi, la Costruzione<sub>COMP</sub> [A-N] può creare anche avverbi, il cui schema è rappresentato in (g) (figura 26): si tratta esclusivamente di avverbi di tempo (tabella 24), costituiti da un aggettivo indefinito (in particolare, *nessuno*, *ogni*, *tale* e *tutto*) e un sostantivo indicante un periodo di tempo (es. *fiata*, *giorno*, *ora*).

Composti [A-N	$I]_{AVV}$	Esempi
nessunora	(1) 'mai'	(1) «Sì me prendisti, quando resguardai / vostre belleze, angellica figura, / che <i>nesunora</i> me ne posso partire».
ognanno	(2) 'tutti gli anni, di anno in anno'	(2) «Un cavalero per devotione <i>ognanno</i> vegnìa a visitare lo corpo de la Magdalena»
ognidì	(3) 'tutti i giorni, quotidianamente'	(3) «E Antonio molto rallegrandosi [] ammoniva li suoi monaci di sempre crescere in fervore e <i>ognindì</i> migliorare, come se <i>ognendì</i> dovessero morire».
ognora	(4) 'sempre, di continuo'	(4) «Tutte le cose volontariamente cercan lor corso e, allegrate in quello, fan lor virtute <i>ognora</i> più possente».
talora	(5) 'alle volte'	(5) «[] e io rimasi in tanta tristizia, che alcuna lagrima <i>talora</i> bagnava la mia faccia».
talvolta	(6) 'alle volte'	(6) «[] per carnale amore l'uno diviene nimico dell'altro, seguendone molti mali e <i>talvolta</i> morte».
tuttafiata	(7) 'ancora, tuttora'	(7) «Contra me non si conviene muovere di ciò questione, ma contra Catellina che ll'àe fatto e fa tuttafiata».
	(8) 'molto spesso, continuamente'	(8) «[] ella tuttanotte sì lo tenea in braccio e basciavalo <i>tutta fiata</i> , ned altro diletto ella non credea che ffosse []».

<sup>125</sup> TLIO, s.v. bellegote.

\_

Composti [A-N]	vv	Esempi
	(9) 'in ogni caso, comunque'	(9) «Et tu acciò ti sforza quanto puoi, e <i>tutta fiata</i> con modo e con ragione».
	(10) 'però, tuttavia'	(10) «Parmi questo uomo da privarlo di vita; <i>tutta fiata</i> iscrivi quello che a te ne pare».
tuttavia	(11) 'continuamente, sempre'	(11) «[] Chi lauda ed ama <i>tuttavia</i> / madonna santa Maria, / fermamente sicuro sia / che buon luogo arà a trovare».
	(12) 'tuttavia, però' (in combinazione con <i>ma</i> )	(12) « <i>Ma tuttavia</i> voglio dire che sua sentenzia non è niente convenevole al nostro Comune».
tuttavolta	(13) 'continuamente'	(13) «Bella figliuola, e io mi penai <i>tuttavolta</i> di bene fare e di digiunare per fuggire lo fuoco d'inferno e per guadagnare questo santo luogo».
	(14) 'tuttavia, però'	(14) «[] e poi dee scrivere la sua affezione, cioè quello che desidera che venga a colui che riceve la lettera, sì come salute o altro che sia avenante, <i>tuttavolta</i> guardando che questa affezione sia di quella guisa e di quelle parole che ssi convengono al mandante et al ricevente».
tuttodì	(15) 'sempre'	(15) «Per quanti modi e casi potemo <i>tuttodì</i> morire!».
tuttogiorno	(16) 'sempre'	(16) «[] in tal maniera che quando alcuno perde sua mogliere, o alcuna perde suo marito, elli non si congiungono mai con altrò, né con altra, tutto il tempo della vita sua, anzi vanno <i>tuttogiorno</i> soli per la foresta».
tuttora	(17) 'di continuo'	(17) «Noi potemo mettere i nostri adversarii in dispetto degli uditori, cioè farli tenere a vile et a neente, se noi diremo che sono uomini nescii sanza arte e sanza senno, de neuno uopo e da neuna cosa; o che sono neghettosi, che <i>tuttora</i> si stanno e dormono e non si muovono se non come per sonno; o diremo che sono lenti e tardi a tutte cose».

Composti [A-N]	VV	Esempi
tuttotempo	(18) 'sempre'	(18) «L'om pote in sé aver tal disianza ch'affanna tutto tempo e non v'aviene, e foria me' s'avesse temperanza al primo che giungesse ne le pene».

Tabella 24. Composti [A-N] con funzione avverbiale attestati nel dataset Compos\_ita\_TLIO

Le forme riportate ed esemplificate nella tabella 24 costituiscono istanziazioni del sottoschema [A-N]<sub>AVV</sub>: esso presenta delle restrizioni a livello semantico che riguardano entrambi i costituenti: il primo slot deve essere riempito da un aggettivo indefinito (tra *ogni, tale, tutto*), il secondo da un sostantivo indicante un periodo di tempo (tipo semantico TEMPO, es. *ora, fiata,* etc.). Tali composti hanno origine da espressioni sintattiche progressivamente cristallizzatesi e unitesi graficamente, sul modello delle quali si sono create altre forme con questa struttura.

Dal punto di vista semantico, tali avverbi esprimono primariamente un valore temporale, legato alla semantica del sostantivo, con sfumature diverse a seconda del tipo di aggettivo: i composti con ogn(i)- rimandano a un'azione che avviene quotidianamente, a cadenza regolare; i composti con tal(e)- indicano un'azione che avviene saltuariamente; i composti con tutto- indicano che l'azione è svolta continuativamente, senza interruzione. In alcuni casi (esempi (10), (12), (14) nella tabella 24), le forme tuttavia, tuttavolta e tuttafiata esprimono, accanto al significato temporale, un valore concessivo ('nondimeno, però'), soprattutto quando usate in combinazione con la congiunzione avversativa ma (come in (12)). Come mostrato in Mauri & Giacalone Ramat (2008), nel corso del tempo, tale uso si imporrà fino a diventare l'unico possibile (si pensi al valore di tuttavia in italiano contemporaneo); d'altra parte, le forme tuttavolta e tuttafiata usciranno progressivamente d'uso (cfr. § 4.3.2). 126 Lo sviluppo di connettivi interfrasali con valore avversativo a partire da avverbi di tempo che esprimono continuità è stato analizzato da Mauri & Giacalone Ramat (2008, 2009, 2012) come un caso di grammaticalizzazione (nel senso di Lehmann 1995; Hopper & Traugott 2003; Traugott 1989): in particolare, nel caso di tuttavia, si tratta di una costruzione avverbiale circostanziale che gradualmente acquisisce lo status di costruzione connettiva controaspettativa.

-

 $<sup>^{126}</sup>$  Entrambe le forme sono registrate nel dizionario Il Nuovo De Mauro (2014) ma contrassegnate dalla marca d'uso OB per i lemmi obsoleti.

L'analisi della Costruzione<sub>COMP</sub> [A-N] permette di mettere in luce alcuni aspetti interessanti relativi a questo tipo di composti. *In primis*, si è osservato che nella maggior parte delle forme (il 70%) la prima posizione è occupata da un ristretto numero di aggettivi qualificativi gradabili (*buono*, *malo*, *mezzo*) o indefiniti (*ogni*, *tale*, *tutto*). Nel primo caso, la Costruzione<sub>COMP</sub> [A-N] permette di formare sostantivi, che talvolta possono occorrere come aggettivi (sotto-schema [A-N]<sub>A|N</sub>); nel secondo, si tratta di avverbi che esprimono un valore temporale (sotto-schema [A-N]<sub>AVV</sub>].

La restante parte dei composti (circa il 30%) è di origine esogena (in particolare, si tratta di prestiti dal latino e dal francese) o fa parte di un insieme di composti costituiti da un aggettivo indicante qualità o stati riferibili a referenti umani e i sostantivi uomo/donna (es. nobiluomo, poveruomo, etc.), creati sul modello dei prestiti francesi gentilhomme e gentildame. L'apporto del francese e, soprattutto, del latino in questa tipologia di composti appare quindi significativo. D'altra parte, gli schemi semi-specificati buon(o/a)-, mal(o/a)- e mezz(o/a)-, i quali (come si vedrà ai § 4.2.5.1 e 5.4) non hanno origine latina, testimoniano la produttività di questa Costruzione<sub>COMP</sub> in italiano antico.

La Costruzione<sub>COMP</sub> [A-PART-PASS]<sub>N|A</sub> istanzia 7 nomi (o aggettivi) composti: si tratta di *biancovestito* 'chi indossa abiti di colore bianco', *malacquisto* 'ciò che viene ottenuto illecitamente o con l'inganno; *maltolto*, furto', *primogenito* 'nato per primo tra i figli di uno stesso genitore', *primegenite* 'insieme dei diritti di cui gode il primogenito', *primonato* 'lo stesso che primogenito', *salvocondotto*, *secondogenito* 'nato per secondo'. Tra queste forme, due sono di origine latina: *primogenito* e *primegenite*, rispettivamente da *primogenitus* e *primogenita*; sul modello di *primogenito*, si formano sia *secondogenito* (e le altre forme utilizzate per riferirsi a figli nati da uno stesso genitore rispetto all'ordine di nascita, es. *terzogenito*, *quartogenito*, etc.) sia *primonato*, con cui condivide lo stesso significato.

Il sotto-schema [AVV-X]<sub>X</sub> istanzia 105 composti costituiti da un avverbio e un elemento testa variabile, che può essere un aggettivo (es. *malsano, sempreverde*), un verbo (es. *maltrattare*), un participio presente (es. *malconveniente*), un participio passato (es. *malvestito*), un avverbio (es. *malvolentieri*) o un nome (es. *benevischio, benessere*). Sulla

base della categoria lessicale del secondo costituente è possibile individuare quindi 6 sotto-schemi, la cui produttività è illustrata nella tabella 25.

[AVV-	[AVV-	[AVV-	[AVV-	[AVV-	[AVV-	Tot
A] <sub>A N</sub>	V] <sub>V</sub>	P.PRES] <sub>A N</sub>	PART.PASS] <sub>A N</sub>	AVV] <sub>AVV</sub>	N] <sub>N</sub>	
14	14	16	48	4	8	105

Tabella 25 Sotto-schemi della Costruzione [AVV-X]<sub>X</sub>: distribuzione dei composti (type) rispetto alla categoria lessicale del secondo costituente

Il sotto-schema [AVV-A]<sub>A|N</sub> è istanziato da 16 composti che presentano come primo costituente un avverbio tra *bene* (*benavventuroso*, *benvoglioso*), *male* (*malacconcio*, *malaccorto*, *malcontento*, *malcornuto*, *malfermo*, *malfido*, *malsaccente*, *malsano*, *malventurato*) e *sempre* (*sempreterno*, *sempreverde*, *sempreviva*).<sup>127</sup> L'uso di *sempre* all'interno di un composto è già attestato in latino classico: delle tre forme attestate nel dataset *Compos\_ita\_TLIO*, due hanno origine latina (*sempreterno* 'la cui durata non ha fine' da *sempiternus*; *sempreviva* 'pianta erbacea perenne del genere Semprevivo' da *semperviva*), mentre *sempreverde* può essere considerato un composto nativo dell'italiano.

Come già osservato al § 4.1, anche gli avverbi male e bene occorrono come primi costituenti di composti già in latino classico, seppur raramente in combinazione con un aggettivo (es. malecastus 'impudico'). Questo tipo di composti acquista particolare produttività in italiano antico, in cui ben(e)- e, soprattutto mal(e)-, si combinano a numerose categorie lessicali e, come si osserverà al § 5.3 e sgg. possono essere considerati una Costruzione<sub>COMP</sub> parzialmente specificata lessicalmente. Nella figura 28 è illustrata la percentuale dei composti con mal(e)- nei diversi sotto-schemi.

 $<sup>^{127}</sup>$  Di seguito non verranno analizzati i composti che presentano l'avverbio *male* come primo costituente, ai quali sono dedicati il  $\S$  5.3 e sgg.

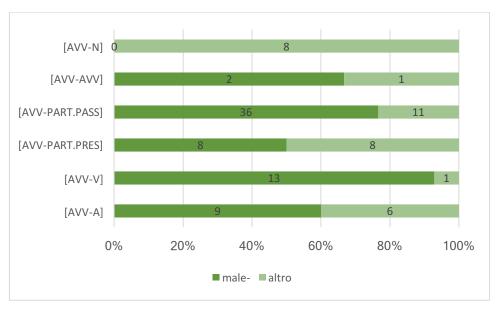


Figura 27 Distribuzione dei composti con mal(e)- nei sotto-schemi della Costruzione [AVV-X]X

Come illustrato dal grafico riportato nella figura 28, i sotto-schemi [AVV-V]<sub>V</sub>, [AVV-PART-PASS]<sub>A|N</sub> e [AVV-A]<sub>N|A</sub> sono istanziati da un significativo numero di composti con mal(e)- in prima posizione. Il sotto-schema [AVV-PART.PRES]<sub>N|A</sub> è istanziato da composti costituiti, oltre che da mal(e)-, dagli avverbi ben(e)- (in particolare, benesonante, benevolente, benparlante, benpiacente, benvivente) e alto- (in particolare, altitonante e altipossente). Entrambi i composti con alto- sono di origine latina (rispettivamente da altitonans e altipotens): anche l'uso dell'avverbio alto come primo elemento di composto è infatti già attestato in latino classico. Il sotto-schema [AVV-AVV]<sub>AVV</sub> è istanziato da due composti con mal(e)- (in particolare, malvolentieri 'controvoglia' e malordinemente 'senza modo, smisuratamente') e da oggimai 'ormai'. L'unico sotto-schema che non presenta mai mal(e)- come primo costituente è [AVV-N]<sub>N</sub>, le cui istanziazioni sono riportate ed esemplificate nella tabella 26.

Composti [AVV-N] <sub>N</sub>	Esempi
retroguardia 'parte posteriore di un esercito'	«Ordenò VIII schiere, a ciascuna evve V uomini, e lo rimanente misa a la <i>retriguardia</i> dietro» (Anonimo, <i>Fatti di Cesare</i> )
benentrata 'per ironia, cattiva accoglienza'	«Qui pone fine al sesto canto, e entra nel IIIJ circulo, là dove per <i>benentrata</i> truova Pluto ministro di questo luogo, e re d'Inferno,

Composti [AVV-N] <sub>N</sub>	Esempi
	secondo li poeti []» (Anonimo, <i>Ottimo</i> commento alla Commedia)
benefattore 'chi fa del bene agli altri in modo disinteressato'	«Lo cane è uno animale che conosce molto li soi <i>benefactori</i> , et è loro molto fedele []» (Anonimo, <i>Bestiario toscano</i> )
beneplacimento 'approvazione, consenso'	«[] noi mostriamo per la pruova delle vostre virtudi voi essere accetti a Dio, aprovati al suo beneplacimento» (Villani, Cronica)
beneplacito 'volontà libera e incondizionata'	«E così fatti, dentro da me lei poi fero tale, che lo mio <i>beneplacito</i> fue contento a disposarsi a quella imagine» (Dante, <i>Convivio</i> )
benessere 'stato di prosperità'	«[] ma al loro <i>benessere</i> , cioè a ddire colore e bellezza, e lle quali se ffatte non fossero già però l'elezzione non sarebbe né ppiù forte né meno» (Anonimo, <i>Libro del difenditore della</i> pace volgarizzato)
benevischio 'buonvischio, altea'	«[] e saldare li malori, tòlli e fa unguento di queste chose: malva, bramchorsina, e benevischio e cipola salvaticha []» (Ruberto di Guido Bernardi, <i>Ricette</i> )
oggidì 'oggigiorno'	«E così oggidì fanno molti, e così Iddio gli giudica, perché amano più i figliuoli che Iddio» (Cavalca, <i>Specchio dei peccati</i> )

Tabella 26 Composti che istanziano la Costruzione (AVV-N) ni ni italiano antico

Tra i composti esemplificati nella tabella 26, due hanno origine latina: si tratta di *beneplacito* da *beneplacitum* e *benefattore* da *benefactor*; <sup>128</sup> le altre forme possono essere considerate composti nativi dell'italiano. Come illustrato dagli esempi, l'avverbio *bene*-, combinandosi con un nome, assume un valore paragonabile a quello dell'aggettivo *buono* ed esprime una valutazione positiva rispetto alla semantica del secondo costituente. Come si è finora osservato, la vitalità dei sotto-schemi che presentano un avverbio come primo costituente è strettamente legata alla diffusione dei composti con *mal(e)*-: a

1

 $<sup>^{128}</sup>$  Il primo costituente del composto *benefattore* potrebbe essere anche interpretato come un sostantivo (lett. 'che fa il bene' invece che 'che agisce bene, in modo virtuoso'): in questo caso andrebbe considerato come un'istanziazione della Costruzione<sub>COMP</sub> dei composti subordinativi con testa a destra [N-N]<sub>N</sub>. Va tuttavia ricordato che, diversamente dall'italiano, in latino la forma *bene* ha solo valore di avverbio: per questo motivo, anche *benefattore* è stato incluso tra le istanziazioni della Costruzione<sub>COMP</sub> [AVV-N]<sub>N</sub>.

questo tipo di forme è dedicata un'analisi più approfondita al § 5.3, che ne metterà in luce aspetti formali e semantici, quantitativi e qualitativi, in prospettiva diacronica.

L'analisi della Costruzione<sub>COMP</sub> dei composti con testa a destra proposta in questo paragrafo ha permesso di mettere in luce alcuni aspetti rilevanti della composizione dell'italiano antico. *In primis*, si è osservato come la Gerarchia di eredità di questa famiglia di Costruzioni<sub>COMP</sub> (figura 24) sia molto più articolata in schemi e sotto-schemi rispetto a quella dei composti con testa a sinistra. La struttura di questo tipo di composti, Modificatore-Testa, richiama senz'altro il modello latino, la cui influenza deve aver giocato un ruolo nella diffusione di questo pattern: si è infatti notato che, nella maggior parte dei casi, la creazione di composti nativi avviene sul modello di una o più forme latine ereditate (es. i composti con *-vendolo* da *pescivendolo*, i composti con *mal(e)-* e *ben(e)-*). L'anteposizione del modificatore può inoltre essere considerato un fattore in grado di garantire maggior riconoscibilità e coesione al composto rispetto alle forme con testa a sinistra che presentano l'ordine non marcato Testa-Modificatore e sono più difficilmente distinguibili dai sintagmi.

## 4.2.3 La Costruzione<sub>COMP</sub> dei composti coordinativi endocentrici

I composti coordinativi endocentrici sono costituiti da due elementi legati da una relazione di coordinazione, ossia parafrasabile attraverso la congiunzione *e* (es. *agrodolce* 'che è agro e dolce allo stesso tempo').<sup>129</sup> Come già osservato al § 1.3 e 2.5, queste forme vengono considerate 'bicefale' dal punto di vista semantico, in quanto individuano entità che presentano proprietà sia del primo sia del secondo costituente. Dal punto di vista formale, la categoria lessicale dei costituenti coincide con quelle dell'intero composto, mentre le proprietà morfologiche del composto (come il genere) sembrano venir ereditate dal secondo costituente, come illustrato dagli esempi in (24).

(24) «*Mazzafrusto* è quello che ha asta lunga quattro braccia, e legatovi una fonda gittan la pietra a due mani a modo di manganella» (Bono Giamboni, *Vegezio*)

-

<sup>&</sup>lt;sup>129</sup> Per un quadro sui composti coordinativi (anche) dell'italiano contemporaneo e in prospettiva tipologica si rimanda a Arcodia, Grandi & Waelchi (2010); Grandi (2011).

b. «[...] girò a morire, lasso, / com fece il *lepretasso*, / c'à si grande tormento / ca di vita ha spavento» (Anonimi, *Rime della scuola siciliana*)

c. «[...] lo color dell'aire, allora che l'aire è senza nebbie, e la tiepida *ventipiova* non còncita l'acque pluviale» (Anonimo, *Arte d'Amare di Ovidio volgarizzata*)

Negli esempi in (24), i composti *mazzafrusto* 'arma da lancio, costituita da una parte fissa cui era assicurata una fionda' (*mazza+frusto* 'vergo o corda dotata di manico'), *lepretasso* 'animale favoloso'<sup>130</sup> e *ventipiova* 'vento misto a pioggia' (*vento+piova* 'pioggia')<sup>131</sup> presentano lo stesso genere del secondo costituente. Questo aspetto è particolarmente interessante perché differenzia i composti coordinativi endocentrici della fase antica da quelli dell'italiano contemporaneo, che generalmente ereditano il genere dal primo costituente. Dal momento che, come osservato da Masini & Scalise (2012: 81-82), questa caratteristica sarebbe legata al fatto che la posizione canonica della testa nei composti dell'italiano contemporaneo è quella di sinistra, i dati relativi all'italiano antico, seppur molto limitati, dimostrerebbero che nella fase antica la posizione canonica della testa fosse quella di destra, sul modello latino.

In ottica costruzionista, queste forme possono essere considerate istanziazioni della Costruzione<sub>COMP</sub> rappresentata in (a) nella figura 29.

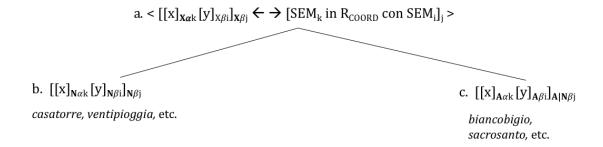


Figura 28 Gerarchia di eredità della Costruzione<sub>COMP</sub> dei composti coordinativi endocentrici dell'italiano antico

\_

<sup>&</sup>lt;sup>130</sup> Nel caso di *lepretasso* va però considerato che, in italiano antico, il sostantivo *lepre* occorre sia al femminile sia al maschile (TLIO, s.v. *lepre*).

 $<sup>^{131}</sup>$  Si noti che in questo caso, e in quello di *ventipioggia*, entrambi attestati al genere femminile, il primo costituente presenta la vocale finale -i, sul modello del *linking element* latino.

La Gerarchia è dominata dallo schema astratto dei composti coordinativi endocentrici (rappresentato in (a)), in cui la categoria lessicale delle unità di input e dell'unità di output (indicata dalla lettera 'X' posta in pedice) coincidono e le proprietà morfologiche del secondo costituente sono condivise dall'intera Costruzione (tale corrispondenza è rappresentata attraverso la co-indicizzazione attraverso la lettera ' $\beta$ ' posta in pedice). Tale schema è istanziato da due sotto-schemi in cui le categorie lessicali sono specificate: in particolare, nel sotto-schema in (b) la coordinazione è tra due sostantivi (es. *casatorre*, *ventipioggia*); nel sotto-schema in (c), i composti sono formati da due aggettivi (es. *gallogreco*, *sacrosanto*).

Il dataset *Compos\_ita\_TLIO* attesta 16 composti di questo tipo, di cui 10 istanziazioni del sotto-schema (b) e 6 istanziazioni del sotto-schema (c).

Il sotto-schema in (b) è istanziato da 10 composti costituiti da due nomi indicanti entità che presentano proprietà sia del primo sia del secondo nome. Dal punto di vista semantico, tali forme (e i relativi costituenti) si distribuiscono tra tre tipi semantici (ANIMALE, ARTEFATTO, EVENTO), come illustrato dalla tabella 27.

Tipo semantico	Composti coordinativi endocentrici [N-N] <sub>N</sub>
ARTEFATTO	casatorre 'edificio a uso abitativo, costruito accentuando lo sviluppo verticale';
	cassapanca 'mobile a forma di cassa usato sia per riporre oggetti che per sedersi';
	mazzafionda 'lo stesso che mazzafrusto';
	mazzafrusto 'arma da lancio, costituita da una parte fissa cui era assicurata una fionda';
	orcipoggia 'nel linguaggio marin., speciale paranco di riserva che faceva eventualmente servizio di orza o di poggia';
ANIMALE	leonepardo 'lo stesso che leopardo'; lepretasso 'animale favoloso'; vermocane 'parassita letale degli animali' <sup>132</sup>

 $<sup>^{132}</sup>$  La forma vermocane è qui collocata tra i composti coordinativi endocentrici, nonostante non sia ben chiaro quale sia il rapporto tra i due costituenti; un'ipotesi alternativa sarebbe considerarla un composto  $[N-N]_{N\_MOD}$  con testa a sinistra.

Tipo semantico	Composti coordinativi endocentrici [N-N] <sub>N</sub>
EVENTO	acquivento 'pioggia con vento';
	ventipioggia (o ventipiova) 'vento misto a pioggia';

Tabella 27 Composti coordinativi endocentrici [N-N]<sub>N</sub> dell'italiano antico: distribuzione rispetto al tipo semantico

Nei composti riportati nella tabella 27 la coordinazione tra i due costituenti è di tipo accidentale (cfr. Grandi 2011: 883-884): le due entità sono cioè legate da un legame transitorio, accidentale, non di tipo naturale. I composti risultanti dalla giustapposizione dei due elementi indicano una terza entità che condivide i tratti dei due coordinandi, come nel caso del composto *casatorre* che indica un artefatto che funge da abitazione e la forma di una torre. Come appare chiaro dalla distribuzione dei composti rispetto al tipo semantico, questa Costruzione comp è sfruttata dall'italiano principalmente per la creazione di nomi che indicano artefatti che svolgono due funzioni, eventi atmosferici, animali (fantastici). Va notata l'assenza di composti coordinativi che indicano luoghi adibiti a due scopi o persone che svolgono due mestieri, entrambi attestati e produttivi in italiano contemporaneo (es. *ristorante-pizzeria*, *attore-cantante*, etc.).

Il sotto-schema (c) è istanziato da 6 composti costituiti da due aggettivi, riportati ed esemplificati nella tabella 28.

Composti coordinativi endocentrici [A-A] <sub>A N</sub>	Esempi
agrodolce 'che ha allo stesso tempo sapore agro e dolce'	«[] giungivi succhio di melangole e zucchero, che sarà <i>acrodolze</i> » (Anonimo, <i>Libro della cocina</i> )
biancobigio 'di colore bianco tendente al grigio, perlato'	«Dicono anore ch'egli [scil.] è di colore biancobigio» (Zanobi da Strada, Volgarizzamento dei Moralia di Gregorio Magno)
celtibero 'abitante della regione centrale della penisola iberica, detta Celtiberia'	«Reggevi Macedonia con Galazia, / illirici, celtiberi e Dalmazia, / li numantini e li parti

\_

 $<sup>^{133}</sup>$  Un esempio di coordinazione naturale in italiano è rappresentato dai binomi coordinati del tipo *marito* e moglie o bianco e nero (cfr. Masini 2008).

Composti coordinativi endocentrici [A-A] <sub>A N</sub>	Esempi
	veloci, / e variate voci / d'altri reami e paesi diversi» (Boccaccio, <i>Rime</i> )
gallogreco 'lo stesso che Galati'	«Nell'oste e nelle moltitudini de' <i>Gallogreci</i> da Gneo Manilio consolo nel monte Olimpo in parte vinti e in parte presi []» (Valerio Massimo, <i>prima red</i> .)
libifenicio 'che appartiene alla popolazione di origine fenicia stanziata nei pressi di Cartagine'	«[] un terzo nuovo da Annibale in luogo d'Ippocrate mandato, di generazione libifenicio Ipponiate, il quale i popolani chiamavano Mutine []» (Anonimo, Deca terza di Tito Livio volgarizzata)
sacrosanto 'che è sacro e santo'	«E niente sono valuti li prieghi del padre per lo figliuolo, né del fratello per lo fratello, i quali sono tribuni della plebe, la quale dignità è sacrosanta, e fu fatta a difensione di libertà» (Filippo da Santa Croce, Deca prima di Tito Livio volgarizzata)

Tabella 28 Composti coordinativi endocentrici [A-A]<sub>A|N</sub> dell'italiano antico

Delle forme riportate nella tabella 28, soltanto *biancobigio* può essere considerato un composto nativo dell'italiano, dal momento che in tutti gli altri casi si tratta di forme di origine latina. In particolare, *sacrosanto* deriva *sacrosanctus*, *libifenicio* da *libyphoenices*, *gallogreco* da *gallograeci*, *celtibero* da *celtiber*; infine, *agrodolce* è probabilmente formata sul modello del composto latino *dulcacidus* 'dolce-amaro'. Dal punto di vista semantico, queste forme indicano prevalentemente la provenienza geografica di un individuo.

L'analisi dei dati estratti dal dataset *Compos\_ita\_TLIO* mette in luce che, seppur scarsamente produttiva, la Costruzione<sub>COMP</sub> dei composti coordinativi è attestata in italiano antico e permette la formazione di composti che designano entità o proprietà che partecipano di tratti semantici di entrambi i costituenti. Dal punto di vista morfologico, va notato che, diversamente da quanto osservato in italiano contemporaneo, nei composti [N-N]<sub>N</sub> coordinativi endocentrici della fase antica, le proprietà flessionali (come il genere) sono ereditate dal secondo costituente, possibile indizio che la posizione canonica della testa nei composti dell'italiano antico fosse quella di destra.

### 4.2.4 La Costruzione<sub>COMP</sub> dei composti esocentrici

Come già discusso al § 2.3, i composti esocentrici rappresentano un caso di Costruzioni<sub>COMP</sub> dotate di proprietà olistiche, non predicibili dalle singole parti e quindi non composizionali: le proprietà morfologiche e semantiche dell'intera Costruzione<sub>COMP</sub> non sono quindi ereditate da uno dei due costituenti o da entrambi, come osservato nelle forme analizzate nei precedenti paragrafi, ma sono proprie della Costruzione<sub>COMP</sub> stessa. Il dataset *Compos\_ita\_TLIO* attesta 93 composti che istanziano tale Costruzione<sub>COMP</sub>, appartenenti a tre tipologie ([V-N], [N-A], [V-V]) e i cui costituenti possono essere legati da una relazione subordinativa, coordinativa o di modificazione. In ottica costruzionista, i composti esocentrici dell'italiano antico possono quindi essere rappresentati in una Gerarchia di eredità articolata in schemi e sotto-schemi, come illustrato dalla figura 30.

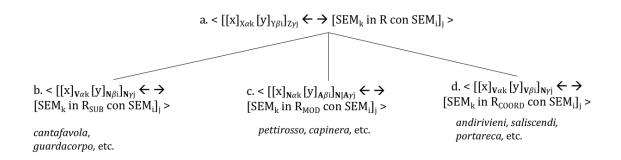


Figura 29 Gerarchia di eredità della Costruzione COMP dei composti esocentrici dell'italiano antico

Nella figura 30, la Costruzione  $_{COMP}$  astratta dei composti esocentrici è riprodotta in (a): l'esocentricità è rappresentata attraverso l'uso di indici diversi per le proprietà dei due costituenti  $(X\alpha \ e \ Y\beta)$  e l'intera Costruzione  $_{COMP}$   $(Z\gamma)$ . Tale Costruzione  $_{COMP}$  istanzia alcuni sotto-schemi, individuati dalle diverse relazioni che intercorrono tra i costituenti e dalle categorie lessicali dei costituenti. In (b) è rappresentato il sotto-schema dei composti esocentrici i cui costituenti sono legati da una relazione di subordinazione: tale schema è istanziato da composti  $[V-N]_N$  (es. guardacorpo, cantafavola). I composti esocentrici che presentano una relazione di modificazione (cfr. (c)) sono costituiti da un nome e un aggettivo (es. pettirosso, capinera). Infine, in (d) è rappresentato lo schema dei composti coordinativi esocentrici, costituiti sempre da due verbi (es. portareca, andirivieni).

La distribuzione nei diversi sotto-schemi dei composti esocentrici attestati nel dataset *Compos\_ita\_TLIO* è rappresentata nella tabella 29.

[V-N] <sub>N</sub>	[N-A] <sub>N A</sub>	[V-V] <sub>N</sub>	tot
78	10	6	94

Tabella 29 La Costruzione<sub>COMP</sub> dei composti esocentrici dell'italiano antico: distribuzione dei type rispetto ai sotto-schemi

I composti esocentrici non si distribuiscono equamente tra i vari sotto-schemi: in particolare, il sotto-schema più produttivo è quello dei composti  $[V-N]_N$ , tradizionalmente considerati il pattern di composizione romanzo per eccellenza. Dei 78 composti  $[V-N]_N$  attestati in italiano antico, 135 13 entrano in italiano sotto forma di prestiti o calchi dal francese antico o dal latino (tabella 30).

francese antico	latino
bagnalcore (cfr. se baigneur); cantafavola (cfr. cantefable) castigabriccone (cfr. chastiermusart) covriceffo (cfr. couvre-chef) guardacorpo (cfr. garde-corps) guardacuore (cfr. garde-nappe) guardaroba (cfr. garde-robe) tagliaborse (cfr. coupe-bourse)	girasole (cfr. heliotropium) mordigallina (cfr. morsus gallinae) placamarito (cfr. viriplaca) volgicuore (cfr. verticordia)

Tabella 30 Composti [V-N]N di origine esogena dell'italiano antico

<sup>134</sup> Come osserva Bauer (2011: 543), «The most important innovation in Romance is the dramatic increase of adjectival and nominal compounds based on verb + complement noun. Attested in all Romance languages in all periods and representing an open class, they are the richest type». I numerosi contributi dedicati ai composti V-N dell'italiano sono prevalentemente di natura sincronica (cfr. Bisetto 1999, 2004: 45-47; Dardano 2009: 190-98; Ricca 2005, 2010, 2015); un'eccezione è rappresentata dall'analisi diacronica (dal Cinquecento all'Ottocento) fornita da Štichauer (2015, 2016), i cui risultati verranno confrontati con quelli emersi nel corso di questa indagine (vd. oltre e § 4.3.4). Sui composti V-N delle altre lingue romanze cfr. Nielsen Whitehead (2012) per un quadro generale; Moyna (2011) per lo spagnolo; Rosenberg (2007, 2008, 2011) per il francese.

<sup>&</sup>lt;sup>135</sup> I composti [V-N]<sub>N</sub> attestati in italiano antico sono generalmente costituiti da un elemento verbale come primo costituente, seguito da un sostantivo; in alcuni rari casi, tra verbo e sostantivo compare anche la preposizione *in* (come in *guardancanna* e *sperandìo*) o l'articolo determinativo *il* (come in *bagnalcore*, *battiloro*, *giralsole*, *tornalsol*). Si tratta di casi piuttosto rari che non verranno inclusi nell'analisi, dal momento che presentano caratteristiche formali diverse rispetto a quelle della Costruzione<sub>COMP</sub> [V-N]<sub>N</sub>.

Come messo in luce dai dati riportati nella tabella 30, un discreto numero di composti [V-N]<sub>N</sub> entra in italiano in prestito dal francese antico: in particolare, i composti con *guarda*- come primo costituente hanno tutti origine esogena. Diversamente da quanto osservato per altre tipologie di composto, il modello latino influisce limitatamente nella formazione di composti [V-N]<sub>N</sub>: tra le forme riportate nella seconda colonna della tabella 30, soltanto *verticordia* presenta una struttura V-N corrispondente a quella del composto italiano; negli altri casi si tratta di calchi imperfetti (ad es. in *placamarito* i costituenti presentano l'ordine inverso rispetto alla base latina; in *morsus gallinae* > *mordigallina* si ha il passaggio da una forma N-N a una V-N).<sup>136</sup>

Analizzando le forme attestate nel dataset *Compos\_ita\_TLIO*, è possibile mettere in luce alcune caratteristiche formali e semantiche di questa tipologia di composti.

Per quanto riguarda l'elemento verbale che funge da primo costituente, la tabella 31 illustra la distribuzione quantitativa dei composti  $[V-N]_N$  rispetto alle tre classi flessive dei verbi dell'italiano.<sup>137</sup>

-are	-ere	- <i>ire</i>
60 type	16 type	2 type
affettapane, baciapolvere, beccalite, beccamorto, brucamaglia, cacapensieri, cacciapensieri, cacciapensieri, cacciapensieri, cassapiedi, castigabriccone, castigapazzi, cavadenti, crepacuore, curaorecchie, drizzacrine, fattibello, foraterra, girasole, grattacacio, guardabanco, guardacorpo, guardacuore, guardamorto, guardanappa, guardapasso, guardaroba, guardaspensa, guastamondo,	batticuore, battifolle, battigote, battilana, battipetto, battisacco, battisfancello, chiudiborsa, fottivento, mesciroba, mettifuoco, mordecascio, mordigallina, pascibietola, pascipeco, volgicuore	coprifuoco, covriceffo

13

 $<sup>^{136}</sup>$  Come già osservato al § 4.1, in latino è largamente sfruttato il pattern  $[N-V]_N$  (es. *domiporta* 'chiocciola', lett. 'colei che porta la propria casa' da *domus* 'casa' + *porto* 'portare'). La produttività del pattern romanzo V-N, con ordine inverso a quello latino, è stata messa in relazione al generale mutamento sintattico nell'ordine dei costituenti da OV a VO (cfr. Nielsen Whitehead 2012: 159; Ledgeway 2012: 225 per un'analisi nella prospettiva della linguistica generativa). Tale prospettiva supporterebbe l'ipotesi, proposta da Gaeta (2008), secondo cui, in relazione all'ordine dei costituenti, la morfologia non sarebbe autonoma dalla sintassi.

<sup>&</sup>lt;sup>137</sup> Dressler & Thornton (1991: 5) propongono invece per l'italiano contemporaneo una classificazione dei verbi in due macroclassi flessive: una produttiva (verbi in -*are*) e una non produttiva (verbi in -*ere* e -*ire*). Trattandosi di dati di italiano antico, qui si mantiene la tradizionale distinzione tra le tre coniugazioni.

-are	-ere	-ire
60 type	16 type	2 type

guazzabuglio, lavaceci, lavascodelle, leccamuffo, leccascodelle, mirasole, pappaculo, parabordone, paracuore, paramosche, parapetto, parasole, pelacane, piantamalanno, picchiapetto, placamarito, portapesi, raffacane, rizzacrino, salvadanaio, sbiecaciglio, scannadio, scialacquavino, squarciasacco, tagliaborse, tornafine, tornasole, vuotacessi

Tabella 31 I composti [V-N]

N dell'italiano antico: distribuzione dei type rispetto alla classe flessiva del verbo che funge da primo costituente

Come illustrato dalla tabella 31, i composti attestati nel dataset *Compos\_ita\_TLIO* si formano prevalentemente con verbi appartenenti alla classe flessiva in *-are*: in particolare, tra gli elementi verbali più produttivi si segnalano *guarda-*, che forma 8 composti, e *para-*, primo costituente di 5 forme. Un discreto numero di composti è costituito da verbi di seconda coniugazione, tra i quali *battere* è il più produttivo: esso ricorre infatti in *batticuore*, *battifolle*, *battigote*, *battilana*, *battipetto*, *battisacco*, *battisfancello*. L'unico verbo in *-ire* che entra in composizione con un sostantivo è *coprire*, in *covriceffo* e *coprifuoco*. Nella selezione delle basi verbali coinvolte nella creazione di questi composti sembrano operare due tipi di restrizioni: una di natura sintattica, per la quale la base verbale deve essere un verbo transitivo e il secondo costituente un argomento interno del verbo;<sup>138</sup> la seconda di natura fonologica, per la quale, ad eccezione di tre casi (*affetta-*, *castiga-* e *scialacqua-*), le basi verbali sono sempre bisillabiche.<sup>139</sup>

L'elemento verbale è stato oggetto di numerosi studi che ne hanno indagato la natura; tre sono le ipotesi formulate dagli studiosi, secondo le quali esso andrebbe interpretato

<sup>138</sup> Come si vedrà meglio oltre, in alcuni casi l'interpretazione argomentale presenta alcune difficoltà.

<sup>&</sup>lt;sup>139</sup> Tale aspetto è stato osservato anche da Ricca (2010: 245) in italiano contemporaneo: «Therefore, it can be safely said thet the preference for two-syllabe verb bases is specific of VN compounding and opposes it to deverbal derivation».

come: 1) un imperativo di seconda persona singolare; 2) un presente indicativo di terza persona singolare; 3) un puro tema verbale. L'ipotesi (1) è stata sostenuta da Darmesteter (1984: 112), Meyer-Lübke (1890: 54) e Prati (1931: 23); quest'ultimo in particolare ha concentrato la sua analisi su antiche forme cognominali, osservando come queste in origine indicassero un individuo che era solito compiere una determinata azione e che poi finiva con l'essere identificato con l'azione stessa, espressa all'imperativo: questo meccanismo sarebbe poi diventato un modello per la creazione di altre forme. L'ipotesi imperativale, accantonata per lungo tempo, è stata recentemente riproposta da alcuni studiosi, che hanno parlato di «imperativo non attualizzato» (Di Sciullo & Ralli 1994: 61-75). L'ipotesi (2) ha trovato l'appoggio di Tollemache (1945: 28), secondo cui nell'interpretare qualsiasi composto non si può mai prescindere dal suo significato, anche quando la forma presenta delle incertezze: se si considera il significato di una forma come *lavapiatti*, cioè 'strumento che lava i piatti', appare chiaro che l'elemento verbale è un presente indicativo di terza persona singolare. Un elemento che indebolisce tale ipotesi è la presenza di una -i, invece che una -e, nei verbi di seconda coniugazione, come ci si aspetterebbe da un indicativo presente (per cui spremiagrumi e non \**spremeagrumi*); secondo Tollemache la presenza della –*i* non rappresenterebbe un problema, perché si sarebbe formata, come per le forme capifuoco o pettirosso, in analogia alla -i linking element del latino.

Infine, l'ipotesi del puro tema verbale è stata sostenuta da Pagliaro (1930), secondo cui «la possibilità della composizione [...] è condizionata indubbiamente dalla nozione che si ha del tema come esponente di un significato fondamentale indipendente da qualsiasi rapporto»: il primo elemento sarebbe quindi un puro tema verbale, perché così è percepito nella coscienza linguistica del parlante.

Rispetto alle tre ipotesi, i dati relativi ai composti [V-N] dell'italiano antico permettono di contribuire alla discussione; in particolare, sembra rilevante osservare che le basi verbali di seconda coniugazione (es. *batti*-) possono presentare una –*e* come vocale finale: è il caso di *battilana*, *battiloro*, *battigote*, *mettifuoco*, esemplificati in (25).

(25) a. «Conciosiacosachè a la cità di Pisa, da pogo tempo in qua, piò facienti oro et argento filato, li quali si chiamino *bacteloro*, avegnano et siano, et in de la dicta cità di Pisa la dicta arte adoperino et faciano cotidianamente» (*Breve dei consoli della corte dell'Ordine de' Mercatanti dell'anno MCCCXXI*)

b. «[...] per due chirlande e uno battechote» (Libro dare e avere di Nesone)

c. «L'arte dei testore e de la *bactelana* uno camorlengo e tre rectore e per agionte doie rectore» (*Statuto del Comune e del Popolo di Perugia del 1342 in volgare*)

d. «Come li *mettefuoco* sieno puniti» (*Costituto del comune di Siena volgarizzato*)

Gli esempi in (25) mostrano come la base verbale dei verbi in -ere possa occorrere anche nella forma con la vocale finale -e: questo dato sembrerebbe supportare l'ipotesi (2) secondo cui il primo elemento di un composto [V-N] andrebbe interpretato come un presente indicativo di terza persona singolare che esprime l'azione svolta abitualmente da una persona o un oggetto. La chiusura della -e e la diffusione (e sistematizzazione) della -i come vocale finale del primo costituente possono essere stati causati da due fattori: la generale tendenza alla chiusura di e protonica (cfr. Serianni 1998: 54-55);  $^{140}$  l'analogia con gli altri composti che presentano una -i sul modello del *linking element* del latino.

Quanto ai secondi costituenti, si può notare che gli unici sostantivi che ricorrono più di due volte sono *cuore* (5), *sole* (4) e *petto* (3); anche per i nomi è ravvisabile una tendenza a preferire forme bisillabiche, pur trattandosi di una restrizione meno stringente rispetto a quella che opera sui verbi.

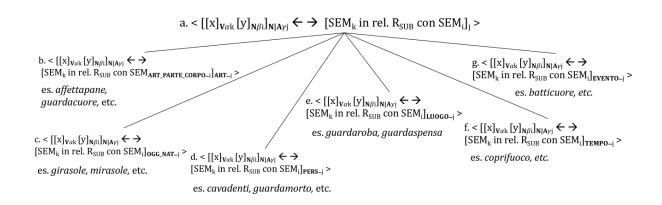
Infine, va osservato che la categoria di output di questo tipo di composti è sempre quella di nome: questo aspetto appare rilevante se confrontato con la situazione mostrata in italiano contemporaneo in cui la Costruzione<sub>comp</sub> [V-N] istanzia sia nomi sia aggettivi (cfr. Ricca 2005, 2010; Štichauer 2015: 141-2 individua nel Settecento un primo esempio, probabilmente un'eccezione, di uso aggettivale di un composto V-N in *vascello da guerra guardacoste*).

Dal punto di vista semantico, i composti  $[V-N]_N$  si distribuiscono tra sei tipi semantici di output: ARTEFATTO (es. *paramosche*, etc.), PERSONA (es. *battipetto* 'chi dimostra una devozione religiosa affettata', *cavadenti* 'chi per mestiere estrae i denti'), LUOGO (es. *guardaroba* 'locale adibito alla conservazione di beni e provviste, *guardaspensa* 'locale o

133

<sup>&</sup>lt;sup>140</sup> In particolare, nel caso dei composti [V-N], si tratta di una vocale intertonica, dal momento che è posta tra i due accenti del primo e del secondo costituente. La posizione intertonica è spesso interessata da un indebolimento che può portare alla chiusura o alla perdita di una vocale interna alla parola.

mobile di arredamento adibito alla conservazione di provviste') e EVENTO (es. *batticuore* 'il battito accelerato del cuore causato da forte emozione', *cantafavola* 'discorso lungo e pieno di sciocchezze'), TEMPO (es. *coprifuoco* 'ora a partire dalla quale era vietato ai cittadini uscire di casa') e OGG\_NATURALE (es. *girasole* 'fiore che segue il corso del sole', *mirasole* 'lo stesso che cicoria'). Tali regolarità semantiche possono essere rappresentate in chiave costruzionista come sotto-schemi in cui è specificato il tipo semantico di output del composto e, in alcuni casi, quello del secondo costituente (figura 31).



 $Figura~30~Costruzione_{\text{\tiny COMP}}~dei~composti~V-N~dell'italiano~antico:~sotto-schemi~rispetto~al~tipo~semantico~di~output$ 

Come illustrato nella figura 31, le regolarità semantiche dei composti [V-N]<sub>N</sub> possono essere rappresentate come sotto-schemi in cui il tipo semantico di output è specificato; le restrizioni semantiche possono riguardare anche il secondo costituente, come nei composti con semantica ARTEFATTO (rappresentati in (b)): in questo caso, infatti, il secondo costituente presenta sempre il tipo semantico ARTEFATTO (es. *affettapane* 'coltello per affettare il pane') o PARTE\_CORPO (es. *guardacuore* 'soprabito foderato di foggia ampia').

I sotto-schemi individuati dal tipo semantico di output non hanno tutti la stessa produttività, come illustrato nella tabella 32.

ARTEFATTO	PERSONA	LUOGO	ТЕМРО	OGG_NAT	EVENTO
23	35	4	1	6	8

Tabella 32 Distribuzione dei composti [V-N] (type) dell'italiano antico rispetto al tipo semantico di output

I dati riportati nella tabella 32 mettono in luce che la composizione verbonominale in italiano antico viene utilizzata prevalentemente per creare nomi indicanti persone o artefatti; sono pochi, ma comunque significativi, i casi in cui la semantica del composto designa un evento o un oggetto naturale; la creazione di nomi che si riferiscono a luoghi o periodi di tempo è invece piuttosto rara. I dati del dataset *Compos\_ita\_TLIO* consolidano quindi l'ipotesi proposta da Štichauer (2015: 145), secondo cui «la polisemia dei composti VN è effettivamente un fenomeno ben presente anche in diacronia».

Tra i composti agentivi si può individuare un nutrito gruppo di appellativi con connotazione negativa: si tratta prevalentemente di nomi che identificano individui che eccedono nel proprio comportamento, tra i quali baciapolvere 'chi dimostra troppa devozione', battipetto 'chi dimostra una devozione religiosa affettata', lavaceci 'persona dappoco', leccascodelle 'chi mangia ingordamente fin gli ultimi avanzi', pascibietola 'inetto, stupido', scialacquavino 'chi abusa del vino'. Nella maggior parte dei casi, tali composti costituiscono creazioni estemporanee per designare personaggi connotati negativamente all'interno di un'opera letteraria: come si osserverà meglio al § 4.3.4, infatti, la maggior parte di queste forme non sopravviverà nel corso della storia dell'italiano. Più rari sono gli appellativi che indicano l'azione svolta abitualmente da un individuo per mestiere, come cavadenti 'chi per mestiere estrae i denti' o portapesi. 141 Minoritaria, ma comunque significativa, è la presenza di composti che indicano un evento o un luogo: nel primo caso si tratta delle forme batticuore, battilana 'arte della battitura della lana', cantafavola, crepacuore, guazzabuglio 'confusione, chiasso', leccamuffo 'percossa sul viso'; i composti dal significato locativo sono guardaroba, guardaspensa 'dispensa', salvadanaio e tornafine 'lo stesso che confine'.

In alcuni casi, i verbi che occorrono in questo tipo di forme presentano un significato che non coincide con quello con cui più frequentemente è utilizzato come forma libera: è il

<sup>&</sup>lt;sup>141</sup> Nel caso di *portapesi*, è interessante osservare che la base verbale *porta*- presenta nella fase antica il significato di 'trasportare': in particolare, nel corpus OVI *portapesi* è usato come sinonimo di 'facchino', come illustrato dal seguente esempio: «[...] diedi a' *portapesi* che recaro detti leni» (*Testi orvietani*). Come si osserverà al § 4.3.4, a questo significato si aggiungerà anche quello di 'contenere', molto produttivo nella fase contemporanea, come notato anche da Ricca (2010).

caso di *guarda*-<sup>142</sup> e *para*-,<sup>143</sup> che occorrono sempre nel significato di, rispettivamente, 'proteggere, conservare' e 'proteggere'.

Per quanto riguarda la relazione tra l'elemento verbale e il nome, generalmente il secondo costituente rappresenta l'oggetto diretto del verbo (es. *affettapane, guardaroba*); tuttavia, sono attestati anche casi in cui il nome rappresenta il soggetto, come in *batticuore* o *crepacuore*. Va inoltre notato che nei composti con *para*- sono possibili due interpretazioni, come esemplificato in (26):<sup>144</sup>

(12) a. «Di mezzo agosto hanno bisogno del *parasole*» (Anonimo, *Libro delle segrete cose delle donne*)

b. «Poi vi feciono fare i fiorentini in Arezzo un altro piccolo castello sopra la porta del piano che va a Laterina, [...] tra 'l muro grande e 'l *parapetto* per i cavalieri, e su per le mura per i pedoni [...]» (Villani, *Cronica*)

In (26a), la forma *parasole* indica un piccolo ombrello atto a riparare dai raggi solari: il secondo costituente rappresenta quindi il complemento di causa efficiente (lett. '(strumento che) ripara dal sole'); in *parapetto* (26b), invece, il secondo elemento costituisce l'oggetto diretto del verbo (lett. '(strumento che) protegge il petto'). L'intepretazione del composto appare legata alla semantica del secondo costituente: ad esempio, nel caso si tratti di una parte del corpo, il verbo va interpretato come nell'esempio in (26b). In ottica costruzionista, i composti con *para*- possono essere rappresentati come sotto-schemi, individuati dalla diversa relazione tra costituenti, della Costruzione<sub>COMP</sub> dei composti [V-N]<sub>N</sub> con tipo semantico di output ARTEFATTO.

L'analisi dei composti  $[V-N]_N$  attestati nel dataset  $Compos\_ita\_TLIO$  permette di mettere in luce almeno due aspetti rilevanti:  $in\ primis$ , si è osservato che questa  $Costruzione_{COMP}$  è disponibile fin dalla fase antica e particolarmente produttiva nella creazione di forme

<sup>143</sup> Il verbo *parare* è attestato nel TLIO soltanto con il significato di 'preparare' o 'imparare'. Il significato di 'proteggere' è espresso dalla forma *riparare*.

 $<sup>^{142}</sup>$  La particolare semantica di *guarda*- può essere legata all'origine esogena di tali composti, che entrano in italiano come prestiti dal francese.

<sup>&</sup>lt;sup>144</sup> Questa particolarità dei composti con *para*- in italiano contemporaneo e in altre lingue romanze (cfr. *protège*- e *guarda*- rispettivamente in francese e spagnolo/portoghese) è stata messa in luce da Magni 2010. I dati del dataset *Compos\_ita\_*TLIO confermano che la duplice interpretazione era possibile già in italiano antico (vs. campione di composti V-N analizzati in Štichauer 2015: 128, in cui sono attestati soltanto casi di *para*- con valore transitivo).

occasionali che denotano referenti umani. Inoltre, l'analisi dei tipi semantici di output ha evidenziato che la semantica agentiva non è l'unica attestata, ma si accompagna ad altri significati, in particolare a quello strumentale (tipo semantico ARTEFATTO). I dati analizzati in questo paragrafo consolidano quindi le ipotesi avanzate da Štichauer (2015: 145-146) secondo cui la composizione verbonominale è sempre stata disponibile, fin dalle fasi più antiche, e caratterizzata dalla compresenza diacronica di vari output semantici. Infine, si è messo in luce che la Costruzione (V-N) istanzia, nella fase antica, esclusivamente sostantivi: la funzione aggettivale, ben attestata in italiano contemporaneo, rappresenta quindi un'evoluzione dei secoli successivi.

Il sotto-schema  $[N-A]_{N|A}$  istanzia composti esocentrici costituiti da un nome e un aggettivo, legati da una relazione di modificazione: il dataset  $Compos\_ita\_TLIO$  ne attesta 10 forme, esemplificate nella tabella 33.

Composti esocentrici [N-A] <sub>N A</sub>	Esempi
bocchiduro 'cavallo duro di bocca'	«Per cammin si concia soma / e lo bocchiduro si doma / per continuar li modi /ch'ànno a dissolver li nodi» (Francesco da Barberino, Documenti d'amore)
capinera 'uccello dei Silvidi'	«La <i>capinera</i> canti "ci ci ricì" / e 'l grillo salti e dica spesso "cri", / e mugghi forte, se ci fosse, il bo'» (Sacchetti, <i>Rime</i> )
capocanino 'dalla testa di cane'	«[] soprappognallo a Priapo o a ciascuno iddio <i>capocanino</i> o a Febro []» (Anonimo, <i>Della città di Dio di sant'Agostino volgarizzato</i> )
capochino 'triste, afflitto'	«Homo capochino com mal viso» (Anonimo, Glossario provenzale-italiano della Laurenziana)
capofitto 'a capo all'ingiù'	«O morte in croce, che lieta prendesti / per seguir il tuo sommo patriarca, / a capofitto per maggior martiro!» (Sacchetti, Rime)
caporiccio 'raccapriccio'	«[] e' le feci una mostra di moneta; / quella mi disse; "Avresti <i>caporiccio</i> ?"» (Meo dei Tolomei, <i>Rime</i> )

<sup>&</sup>lt;sup>145</sup> Si utilizza il termine disponibilità nel senso di *disponibilité* (cfr. Corbin 1987).

Composti esocentrici [N-A] <sub>N A</sub>	Esempi
caposalvo 'corrisponde alla formula latina ad salvum caput'	«Anchor vetammo che nessuno chierego per sé o per altro in palese o in secreto preste ad usura on a caposalvo, né facia alcun'altra cosa o conceda per la quale se possa recevere alcuna cosa oltra la sorte o guadagno usuraio provenire» (Costituzioni egidiane)
palafitta 'struttura di pali di lego conficcati in terra'	«[] a' luoghi dove i fiumi per loro rabbia fanno rotture, un poco più suso si facciano palafitte forti, secondo l'impeto del fiume» (Piero de' Crescenzi volg.)
pettirosso 'piccolo passeriforme della famiglia dei Muscicapidae'	«Di che quando Teseo s'accorse, correndo loro dietro per ucciderle, Progne diventò rondine e Filomena usignuolo e Iti diventò pettorosso e Tereo becchipuzzola» (Boccaccio, Chiose Teseida)
codimozzo 'che ha la coda mozzata'	«Onde il podestà giudicò che 'l buon uomo se ne menasse l'asino suo, e se non volesse, begnai se lo tenesse tanto elli che rimettesse la coda, e poi glielo rendesse. Begnai rimase libero, e 'l villano se nel menò a casa sua così codimozzo per lo migliore» (Sacchetti, Trecentonovelle)

Tabella 33 Composti esocentrici [N-A]<sub>N|A</sub> dell'italiano antico

La maggior parte di questi composti<sup>146</sup> può essere ascritta alla categoria dei composti 'possessivi' che identificano un'entità esterna al composto esprimendone una proprietà caratteristica posseduta e possono essere parafrasati come 'qualcosa che ha XY come caratteristica principale'. Questa Costruzione<sub>COMP</sub> è particolarmente produttiva nelle lingue germaniche e permette di creare soprattutto sostantivi che denotano persone, animali o piante attraverso una caratteristica da essi posseduta (cfr. Saracco 2017). Anche i composti attestati in italiano antico mostrano tali proprietà semantiche: in particolare, *bocchiduro*, *codimozzo*, *capinera*, *pettirosso*, designano animali mediante una caratteristica fisica; *capochino* o *capocanino* si riferiscono invece a referenti umani. Questi composti possono essere considerati istanziazioni di un sotto-schema della Costruzione<sub>COMP</sub> dei composti [N-A]<sub>N|A</sub> esocentrici che presenta restrizioni sul tipo semantico di output, come illustrato nella figura 32.

.

<sup>&</sup>lt;sup>146</sup> Fa eccezione la forma caposalvo che ha origine dall'espressione latina ad salvum caput (vd. § 4.2.5.2).

$$a. < [[x]_{\mathbf{N}\alpha\mathbf{k}}[y]_{\mathbf{A}\beta\mathbf{i}}]_{\mathbf{N}|\mathbf{A}\gamma\mathbf{j}} \longleftrightarrow [\operatorname{SEM}_{\mathbf{k}} \text{ in rel. } \mathbf{R}_{\text{MOD}} \text{ con SEM}_{\mathbf{i}}]_{\mathbf{j}} > \\ \\ b. < [[x]_{\mathbf{N}\alpha\mathbf{k}}[y]_{\mathbf{A}\beta\mathbf{i}}]_{\mathbf{N}|\mathbf{A}\gamma\mathbf{j}} \longleftrightarrow [\operatorname{SEM}_{\mathbf{k}} \text{ in rel. } \mathbf{R}_{\text{MOD}} \text{ con SEM}_{\mathbf{i}}]_{\mathbf{ANIMALE}|\mathbf{PERS}-\mathbf{j}} > \\ \\ \text{es. } \textit{capinera, capochino, etc.}$$

Figura 31 Sotto-schema della Costruzione  $_{\text{COMP}}$  dei composti [N-A]  $_{\text{N|A}}$  esocentrici

Fanno eccezione il composto *palafitta*, che deriva dall'univerbazione dell'espressione latina *pala fixa*, e *caporiccio* che indica uno stato mediante l'effetto che esso produce (ossia il rizzarsi dei capelli sul capo per effetto della paura): nella Gerarchia di eredità rappresentata nella figura 32, queste forme si collocano a livello della Costruzione<sub>COMP</sub> in (a) dal momento che non rispettano le restrizioni semantiche dello schema.

Infine, i composti esocentrici dell'italiano antico possono essere costituiti da due verbi, legati da una relazione di coordinazione: essi istanziano la Costruzione<sub>COMP</sub> rappresentata in (d) nella figura 30. In particolare, il dataset *Compos\_ita\_TLIO* ne attesta 6 forme, riportate ed esemplificate nella tabella 34.

Composti esocentrici [V-V] <sub>N</sub>	Esempi
andirivieni 'movimento d'entrata e d'uscita attraverso uno spazio circoscritto, percorso tortuoso'	«et allora Dedalo cavò in una grotta d'uno monte, e facevi molte pareti con molti usci, che tutti aprivano in dentro, e molti <i>andirivieni</i> e pose nell'entrata molto imagini che faceano grande paura a ch' v'entrava» (Francesco da Buti, <i>Commento all'Inferno</i> )
battisoffia 'paura che fa venire la palpitazione al cuore'	«[] io dovessi essere rimaso nelle valli tra le cannnucci; ché m'hai dato sì fatta <i>battisoffia</i> che io non sarà mai lieto, e forse me ne morrò» (Sacchetti, <i>Trecentonovelle</i> )
giravolta 'rapido movimento rotatorio, repentina inversione di marcia di un esercito'	«[] elli, sì come savi, quando li nemici andaro per piglialli, egli diero una <i>giravolta</i> tuttavia fuggendo, e vennero verso le navi e quini

	montaro tutti incontamente» (Anonimo, <i>Milione di Marco Polo toscano</i> )
portareca 'gente vile' (fig.)	«Quante minacce ha' fatte con novelle / per farti udire spesso al <i>portareca</i> / acciò che l'altru' sangue torni in pelle!» (Sacchetti, <i>Rime</i> )
saliscendi 'una delle serrature dell'uscio'	«[] e giugnendo all'uscio, mettendo la chiave nel serrame, e volgendola per aprirlo, il frate, sentendo il <i>saliscendo</i> , subito si leva []» (Sacchetti, <i>Trecentonovelle</i> )
sanamunda 'pianta erbacea della famiglia Timeleacee, adoperata per proprietà purganti e revulsive'	«Prindi erba <i>sanamunda</i> et erba venti et agrimonia et frapparia et pinpinella []» (Anonimo, <i>Thesaurus pauperum in volgare</i> )

Tabella 34 Composti [V-V]<sub>N</sub> esocentrici dell'italiano antico

I composti  $[V-V]_N$  attestati in italiano antico sono sempre costituiti da due verbi diversi legati da una relazione di coordinazione. Dal punto di vista formale, si tratta di due basi verbali che compaiono in forma omofona all'imperativo singolare; si osserva una netta preferenza per i verbi bisillabici: fa eccezione soltanto il composto *andirivieni* in cui il secondo costituente presenta il prefisso ri-.

Dal punto di vista semantico, questa Costruzione<sub>COMP</sub> istanzia nomi indicanti eventi (*giravolta, andirivieni*), stati (*battisoffia*), persone (*portareca*), artefatti (*saliscendi*) ed elementi naturali (*sanamunda*).<sup>147</sup> Non sono quindi attestati composti costituiti dalla reduplicazione dello stesso verbo confrontabili con quelli del tipo *fuggi-fuggi* (cfr. Thornton 2010) che indicano un'azione svolta da un gruppo di persone.

## 4.2.5 Le Costruzioni semi-specificate

Nei seguenti paragrafi si analizzano alcune Costruzioni $_{COMP}$  parzialmente specificate lessicalmente, costituite quindi da uno slot chiuso e uno variabile: in particolare, il § 4.2.5.1 è dedicato a due sotto-schemi della Costruzione $_{COMP}$  dei composti [A-N] $_{A|N}$  in cui l'elemento specificato è un aggettivo; il § 4.2.5.2 prende in esame i composti che presentano *capo*- come primo costituente, di cui vengono analizzate le proprietà semantiche nelle diverse Costruzioni $_{COMP}$  che questo insieme di forme istanzia.

<sup>147</sup> Si noti però che questa forma è un prestito dal latino thymelaea sanamunda (cfr. TLIO, s.v. sanamunda).

# 4.2.5.1 Le Costruzioni [buon(o/a)-x]<sub>NIA</sub> e [mezz(o/a)-x]<sub>NIA</sub>

La Costruzione $_{COMP}$  [buon(o/a)-x] $_N$  permette di creare sostantivi ed è costituita dall'aggettivo buono più una base variabile. Dal punto di vista formale, l'aggettivo buono può perdere la vocale finale quando si unisce a un sostantivo (come in buonfatto o buonora).

Semanticamente, la Costruzione veicola una valutazione positiva da parte del parlante rispetto al significato del sostantivo, come si può osservare negli esempi riportati in (27):

(27) a. «[...] e Dio è colui che adopera in noi il *buonvolere*» (Cavalca, *Esp. simbolo*)

b. «Certo, donzella, quelli che primamente mi disseno novelle de le vostre bellesse non mi mentino di nulla, ansi ne funo alquanto scarsi di dirmi la verità; ché, se Dio mi dia *buona aventura*, in voi à troppo maggiore bellesse che ll'omo non mi dicea [...]» (*Palamedés* pis.)

In (27a), il composto *buonvolere* indica una 'buona disposizione d'animo', una 'propensione ad agire positivamente'; in (27b), *buona aventura* fa riferimento a un 'destino favorevole, positivo'. 

148 La forma *bonavventura* può occorrere anche all'interno di due locuzioni avverbiali: *per bonavventura* 'per fortuna' e *in bonavventura* 'formula di augurio'. 

Possono occorrere in locuzioni avverbiali anche le forme *buonafede* e *buonora*: in particolare, *buonafede* 'dimostrazione di fedeltà e servizio leale' è utilizzato in formule giuridiche con il significato di 'convinzione di operare per il bene comune da parte di chi svolge una funzione pubblica' nelle le locuzioni *a buona fede, in buona fede, per buona fede*; il composto *buonora*, utilizzato per indicare l'inizio della giornata, occorre esclusivamente nelle locuzioni *a buonora*, come esemplificato in (28).

(28) a. «[...] e andavano nella piazza *a buon'ora* per avere buono luogo presso alle bigoncie del grano» (Dom. Benzi, *Specchio umano*, 1347)

b. «Fatto il mugnaio abate, la mattina *di buon'ora* si mise in cammino [...]» (Sacchetti, *Trecentonovelle*)

-

<sup>&</sup>lt;sup>148</sup> TLIO, s.v. buonvolere e bonavventura.

Le basi che occorrono nei composti [buon(o/a)-x] sono prevalentemente nomi astratti: in particolare, si tratta di nomi indicanti stati d'animo (tipo semantico SENTIMENTO, es. buonavoglia), stati (STATO, es. bonavventura), unità di tempo (tipo TEMPO, es. buonora) ed eventi (tipo EVENTO, es. buonfatto). Fanno eccezione soltanto i composti buonuomo (tipo PERSONA) e buonvischio 'altea' (tipo OGG\_NATURALE).

La Costruzione<sub>COMP</sub> [mezz(o/a)-x] è costituita dall'aggettivo mezzo e una base variabile; essa permette di creare sostantivi che, in rari casi (due, in particolare, tra le forme del dataset  $Compos_ita_TLIO$ ), possono occorrere anche in funzione di aggettivo, come esemplificato in (29):

(29) a. «Per ciascuna soma di panni *mezzalane*, veronese, altronese, fiorentino [...]» (Statuti della colletta del comune di Orvieto)

b. «[...] per lo lino, che è dentro da riverso in questi panni di *mezzalana*, s'intende l'astuzia, e la malizia delli ipocriti [...]» (Domenico Cavalca, *Esposizione del simbolo degli Apostoli*)

c. «Mandai Lippo a vendere in Pisa panni XXVIIJ in due some, i quali fuoro sei tintillana mezokolore [...]» (Conto delle mercanzie di Pisa tenuto da Stefano Soderini)

In (29a), *mezzalane* occorre come modificatore del sostantivo *panno* per indicare un tipo di stoffa costituita per una metà da lana, per l'altra da cotone; come esemplificato in (29b), esso può occorrere, con lo stesso significato, anche come sostantivo. Il composto *mezzocolore* è attestato invece soltanto come modificatore: nel caso esemplificato in (29c), esso modifica il sostantivo *tintillano* (indicante un 'panno fino di lana, tinta prima di essere lavorata').

Semanticamente, la base è costituita da un sostantivo che può indicare un artefatto (tipo ARTEFATTO, es. *mezzocannone* 'salterio di dimensioni più piccole del cannone'), un'unità di tempo (tipo TEMPO, es. *mezzogiorno*), un essere animato (tipo ANIMALE, es. *mezzofiera* 'essere mezzo uomo mezzo animale'), una proprietà (tipo PROPRIETÀ, es. *mezzocolore* 'colore di mezzo tra due colori').

In combinazione con tali basi, l'aggettivo *mezzo* può veicolare il significato di 'metà' (come nel caso di *mezzaluna* (30a)) o indicare che un'entità è a metà strada tra due classi

(come in *mezzofiera* (30b)) o di minori dimensioni rispetto allo standard (come in *mezzocannone* (30c)).

(30) a. «[...] e ripiegato col tronco corpo, saltò nell'acque, e l'ultima coda ee divisa come sono i corni della *mezza luna*» (Simintendi, *Metamorfosi di Ovidio volgarizzate*)

b. «[...] non si possono saziare di riguardare i terribili occhi e 'l volto e 'l petto velluto di setole e li spenti fuochi nella strozza del *mezzofiera*» (Lancia, *Eneide volgarizzata*)

c. «[...] llei, e cierti altri ch'erano ivi ad albergo la fecion ballare al suono d'uno *mezzo cannone*» (Francesco da Barberino, *Reggimento e costumi di donna*)

Entrambe le Costruzioni<sub>COMP</sub> non hanno riscontro in latino, ma possono essere considerate creazioni autonome dell'italiano, la cui evoluzione verrà analizzata al § 4.3.2.

## 4.2.5.2 La Costruzione<sub>COMP</sub> [capo-x]<sub>N</sub>

Come osservato nei paragrafi dedicati alle Costruzioni<sub>COMP</sub> astratte, un nutrito gruppo di composti attestati in italiano antico presenta il nome *capo* come primo costituente: esso deriva dal latino *caput*, che già in epoca classica presenta un notevole numero di significati, elencati nella prima colonna della tabella 35; la quasi totalità dei significati trova riscontro in italiano antico, come mostrato dagli esempi riportati nella terza colonna.<sup>149</sup>

Significati di <i>caput</i>	Significati di <i>capo</i>	Esempio in italiano antico
1. 'testa, come parte estrema del corpo o sede dell'intelletto'	1a. 'parte superiore del corpo' 1b. 'complesso delle facoltà umane intellettive'	1a. « [] vino citrino dicise caldissimo e fumuso, / dolore in capo genera, troppo è furiuso» (Regimen Sanitatis) 1b. «Così sen va, e quivi m'abbandona / lo dolce padre, e io rimagno in forse, / che sì e no nel capo mi tenciona» (Dante, Commedia, <i>Inferno</i> )
2. 'vita di una persona' (anche come pegno per varie offese')	2. 'condanna per decapitazione'	2. «Asdrubale con pochi fuggio in Lilibeo, ed essendo assente da quelli di Cartagine fue nel <i>capo</i> dannato» (Bono Giamboni, <i>Orosio</i> )

\_

<sup>&</sup>lt;sup>149</sup> Le definizioni latine (traduzioni mie) sono tratte dall'*Oxford Latin Dictionary* (Glare 1982); le definizioni e gli esempi in italiano antico sono tratti dal TLIO.

Significati di <i>caput</i>	Significati di <i>capo</i>	Esempio in italiano antico
3. 'persona individuo'		
4. 'parte estrema di qualcosa, somma, vertice'	4. 'uno degli estremi entro cui si estende in lunghezza o in larghezza uno spazio'	4. «[] e in capo della canna segnare con brocco e poi tagliare o panno []» (Statuto dell'arte di Calimala del 1334)
5. 'parte iniziale di un processo o di un periodo di tempo'	5. 'momento iniziale rispetto a una durata temporale'	5. «A capo dell'anno, ov'è la festa del suo idolo, lo padre col figliuolo menanto questo montone dinanzi a lo suo idolo, e fannogli grande riverenza con tutti li figliuoli» (Anonimo, Milione di Marco Polo, versione toscana)
6. 'estremità di una strada, origine di un fiume'	6. 'estremità, inizio'	6. «E poscia in prima che andasse contra Ispartaco, il quale era coll'oste a capo del fiume di Silaro, i Galli e i Germani vinse []» (B. Giamboni, <i>Orosio</i> )
7. 'origine, punto di partenza di cose astratte o concrete'	7. 'ciò che sta all'origine di qualcosa'	7. «[] queste cose contate / son di superbia nate, / di cui il savio dice /ched è <i>capo</i> e radice / del male e del peccato» (Brunetto Latini, <i>Tesoretto</i> )
8. 'iniziatore, leader, guida'	8. 'chi occupa una posizione di comando o di responsabilità o gode di autorevolezza rispetto a un gruppo o a una comunità'	8. «E dice che stando per due mesi in sua corte gli era fatto multo grande onore da tutta gente, ma da Pugliesi e da Ciciliani più, i quali n'aveano fatto loro capo di lui» (Anonimo, Leggenda di messer Gianni da Procida)
9. 'città capitale'	9. 'città più importante di una determinata regione; capitale di uno stato'	«E dice ch'elli fia salute di questa umile e bassa Italia, nella quale è Roma, che fu capo del mondo, e sedia dello Imperio» (Ottimo, Commento della Commedia, Inferno)
10. 'cosa essenziale,	10. 'elemento principale'	10. «Qui presso ha una virtù che s'apella Fede Cristiana, la quale è capo e

Significati di <i>caput</i>	Significati di <i>capo</i>	Esempio in italiano antico
argomento fondamentale'		fondamento di tutte l'altre Virtù a coloro che vogliono intendere al servizio di Dio» (B. Giamboni, <i>Il Libro de' Vizii e delle Virtudi</i> )
11. 'la prima parte di uno scritto o un discorso'	11. 'parte iniziale di uno scritto o di un discorso'	11. «Questa ragione di questa faccia è iscritta in capo del libro» ( <i>Registro di entrata e uscita di Santa Maria di Cafaggio</i> )
12. 'la parte iniziale di una parola o una frase'	12. 'prima lettera di una frase o di un verso'	12«Or giongi questo Dentro con lo capo de li ttre versi dicti e nominati di sora, e concorderai il testo como dee stare» (Maramauro, Expositione sopra l'Inferno di Dante Alighieri)
13. 'sezione, paragrafo all'interno di un testo'	13. 'ciascuna delle divisioni in cui è ripartito uno scritto o un discorso'	13. «[] i quali danari ne ritrovamo a ricolliare per lo libro de' conti per lo posciaio dì d'ottobre 308 partitamente in vintessei capita di dette []»
	14. '(riferito a un insieme di uomini, oggetti, merci, bestiame) la singola unità appartenente a quell'insieme'	14. «[] e all'isola della Capraia scesono in terra, e levarono preda di mille <i>capi</i> di bestie []» (Matteo Villani, <i>Cronica</i> )

Tabella 35. *Caput* in latino e *capo* in italiano antico: definizioni e relativi esempi

Come messo in luce dagli esempi riportati nella tabella 35, sia *caput* sia *capo* possono indicare:

- a. una parte del corpo (la testa, come in (1a-b));
- b. la parte estrema o l'origine di un elemento concreto (un artefatto, come in (4); un fiume o una strada, come in (6); un testo, come in (11); una parola o un verso, come in (12)) o astratto (come in (7) e (5));
- c. un elemento fondamentale, essenziale (come in (10));
- d. un elemento preminente rispetto agli altri (ad es. una città, come in (9), o un individuo, come in (8)).

Non trova riscontro in italiano antico l'uso di *caput* esemplificato in (3), in cui il nome si accompagna a un epiteto descrittivo (come in *caput liberum* 'uomo libero'), a un nome proprio, a una carica o a un dimostrativo per indicare un individuo. Si noti però l'uso, attestato in italiano antico, di *capo* per indicare una singola unità appartenente a un insieme di uomini, oggetti, merci o bestiame, come esemplificato in (14).

In latino classico, l'uso di *caput* come costituente di composto<sup>150</sup> è registrato nei dizionari soltanto nella forma (plurale) *capitecensi* (lett. 'censiti per la persona'), attestata anche in italiano antico come *capocenso*, con il significato sia di 'chi nell'antica Roma era contato per la sola persona, nullatenente' (significato comune anche alla forma latina) sia di 'imposizione fiscale per testa, testatico'.<sup>151</sup> Nel composto latino, la forma *caput* svolge la funzione di modificatore (si trova infatti al caso ablativo) dell'elemento testa, il participio passato *census* 'censito', e ha un significato confrontabile con quello esemplificato in (14).

Diversamente dal latino classico, in latino medievale caput occorre in un certo numero di combinazioni Nome Nome, registrate dal dizionario Du Cange  $(1982)^{152}$  e riportate nella tabella  $36.^{153}$ 

Caput + x	
caput anni	ʻinizio dell'anno'
caputaqua	'fonte d'acqua'
caput burgi	'capo del borgo'
caput camerae	'lo stesso che prafectus cubiculi'
caput castri	'capo dell'accampamento'
caput contubernii	'lo stesso che <i>decano</i> '
caput domus	'lo stesso che <i>caput mansi</i> '
caput ecclesiae	'abate, piore, o qualsiasi altra figura preposta alla guida di una chiesa o un monastero'
caput finis	'confine'
caput fundorum	
caput Germaniae	'detto del Reno, confine della Germania'
caput guerrae	'capo militare, lo stesso che dux militum'
caput mansi	'capo famiglia'
caput monasterii	'lo stesso che <i>abate</i> '
caputpurgium	'medicamento per il capo'

<sup>&</sup>lt;sup>150</sup> I principali dizionari del latino non registrano alcun lemma composto in cui *caput* rappresenta il primo costituente ad eccezione di *capitecensi* (cfr. Glare 1982; Georges & Georges 1972; Forcellini 1940).

<sup>&</sup>lt;sup>151</sup> TLIO, s.v. *capocenso* (cfr. Forcellini 1940, s.v. *capitecensi*).

<sup>&</sup>lt;sup>152</sup> Il dizionario è liberamente consultabile online al link: <a href="http://ducange.enc.sorbonne.fr/">http://ducange.enc.sorbonne.fr/</a> (ultima visualizzazione in data 05/06/18).

 $<sup>^{153}</sup>$  Per ciascun lemma viene proposta una glossa in italiano sulla base degli esempi e delle definizioni riportati dal dizionario.

Caput + x	
caput scholae	'chi presiede una scuola militare'
caput senatus	ʻchi presiede il Senato'

Tabella 36. Combinazioni *caput* + x registrate dal dizionario di latino medievale Du Cange (1982)

Nella maggior parte delle forme estratte dal dizionario Du Cange (1982), e riportate nella tabella 36, *caput* designa un individuo che svolge la funzione di guida o di autorità all'interno di un luogo o di un'istituzione: <sup>154</sup> così il *caput ecclesiae* è colui che amministra e guida una chiesa o un monastero, il *caput scholae* è colui che presiede una scuola militare, il *caput castri* è l'autorità che amministra e controlla un accampamento. In alcuni casi, come *caput senatus*, *caput guerrae* e *caput camerae*, la forma *caput* è preferita a termini utilizzati in epoca classica per designare specificatamente il comandate/l'autorità di un gruppo, come *dux* o *praefectus*.

Non presentano invece una semantica agentiva le forme *caput anni*, *caputaqua*, *caput finis*, *caput Germaniae* e *caputpurgium*: nei primi due casi, *caput* rinvia al significato di 'inizio', 'origine', designando la prima parte di un periodo di tempo e la fonte di un corso d'acqua; in *caput finis* e *caput Germanie*, esso indica l'estremità di un territorio; infine, in *caputpurgium*, *caput* occorre nel significato concreto di 'testa' e, diversamente dagli altri casi, in questo caso costituisce il modificatore del composto.

La semantica mostrata da *caput* all'interno delle forme medievali si articola quindi attorno a tre significati: 1) 'testa'; 2) 'inizio o origine di qualcosa'; 3) 'elemento preminente rispetto ad altri'. Quest'ultimo, in particolare, costituisce il significato più frequente all'interno del (seppur limitato) campione di forme di origine medievale.

Il quadro fin qui delineato trova riscontro in italiano antico, come mostrano i composti con *capo* estratti dal dataset *Compos\_ita\_TLIO*; in queste forme *capo* può svolgere diverse funzioni: in particolare, esso può fungere da testa di un composto subordinativo o attributivo (endocentrico o esocentrico) o può costituire l'elemento modificatore. La distribuzione delle forme con *capo-* rispetto al tipo di Costruzione<sub>COMP</sub> è illustrata dalla tabella 37.

-

<sup>&</sup>lt;sup>154</sup> Si tratta prevalentemente di istituzioni politiche o militari della Roma repubblicana (es. il Senato nel caso di *caput senatus*) o di luoghi/istituzioni di natura religiosa (es. il monastero in *caput monasterii*).

Tipologia di composto	Forme con <i>capo</i> -
(a) $[N-N]_N$ subordinativi con testa a sinistra	capodieci 'capo di dieci soldati', capocroce 'crocevia', capofamiglia 'capo della famiglia', capoguardie 'capo delle guardie', capomastro 'capo delle maestranze', capomese il primo giorno del mese', capopopolo 'capo di parte popolare', caporione 'a Roma, capo del rione', caposcala 'sommità di una scala', capoverso 'inizio di capitolo, periodo o verso'
(b) $[N-N]_N$ subordinativi con testa a destra	capocenso 'imposizione fiscale per testa', capogirlo 'vertigine, lo stesso che capogiro', capomaglio 'maglia di acciaio che poteva coprire anche il capo', capomorbo 'malattia infettiva', capopurgio 'medicamento per far purgare il capo', caposoldo 'primo nella divisione del soldo', capovolta 'rivolgimento'
(c) [N-V] <sub>V</sub> subordinativi con testa a destra	capolevare 'cadere rovinosamente', caporicciare 'provare raccapriccio'
(d) [N-A] <sub>N A</sub> attributivi (esocentrici)	capinera 'piccolo uccello di tinte bruno- cenerognole', capocanino 'dalla testa di cane', capochino 'triste, afflitto', capofitto 'a capo all'ingiù', caporiccio 'raccapriccio', caposalvo 'clausola giuridica che prevede la restituzione materiale degli animali consegnati'
(e) $[N-N]_N$ subordinativi esocentrici	capoletto 'drappo da appendere in capo al letto'
(e) [N-N] <sub>N</sub> coordinativi	capopiede 'sottosopra'

Tabella 37. Composti con capo-: distribuzione rispetto alla Costruzione COMP

I composti come *capofamiglia* o *capomese* costituiscono istanziazioni della Costruzione $_{\text{COMP}}$  dei composti  $[N-N]_N$  subordinativi con testa a sinistra (figura 33): in particolare, essi rappresentano uno schema parzialmente specificato lessicalmente  $[capo-x_N]_N$ . Sulla base delle proprietà semantiche di *capo* e del secondo costituente possono essere individuati due ulteriori sotto-schemi, rappresentati nella figura 33 ai punti (c) e (d).

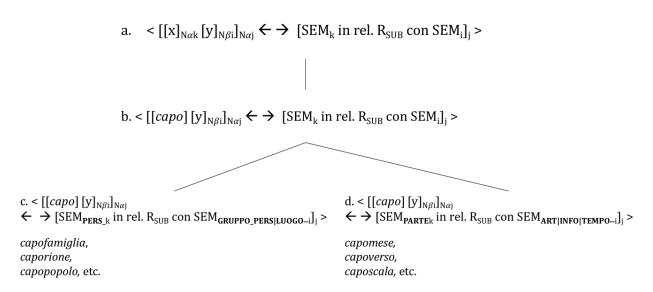


Figura 32. Sotto-schemi con *capo*- della Costruzione<sub>comp</sub> dei composti N-N subordinativi con testa a sinistra

In particolare, il sotto-schema in (c) istanzia i composti in cui *capo* indica un referente umano che detiene l'autorità o il comando rispetto a un gruppo di persone (es. *capofamiglia*, *capopolo*) o a un luogo (*caporione* 'a Roma, capo di rione'): tali restrizioni semantiche sono espresse mediante l'indicazione del tipo semantico di *capo* (PERSONA) e del secondo costituente (GRUPPO\_PERS O LUOGO).

Il sotto-schema (d) istanzia i composti in cui capo indica la parte iniziale/estremità di un periodo di tempo (tipo semantico TEMPO; es. *capomese* 'primo giorno del mese') o di un elemento concreto (tipo semantico ARTEFATTO, es. *caposcala* 'sommità d'una scala'; tipo semantico INFO, es. *capoverso* 'inizio di capitolo, periodo o verso'). Anche in questo caso le restrizioni semantiche sono indicate dai tipi semantici posti in pedice.

I composti come *capocenso* o *caporicciare* costituiscono invece istanziazioni della  $Costruzione_{COMP}$  dei composti  $[N-N]_N$  e  $[N-V]_v$  subordinativi con testa a destra. La figura 34 ne rappresenta la Gerarchia di eredità.

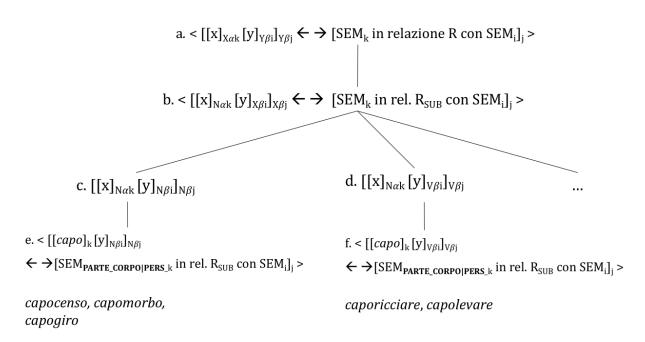


Figura 33 Sotto-schemi con  $\it capo$ - della Costruzione $_{\tiny COMP}$  dei composti  $[N-N]_N$  e  $[N-V]_V$  subordinativi con testa a destra

I composti come *capocenso* e *capomorbo* costituiscono istanziazioni di un sotto-schema parzialmente specificato lessicalmente della Costruzione<sub>COMP</sub> rappresentata in (c), che istanzia i composti [N-N]<sub>N</sub> subordinativi con testa a destra: come indicato dai tipi semantici posti in pedice nel sotto-schema (e), in queste forme *capo* ha sempre il significato di 'testa' (tipo semantico PARTE\_CORPO, es. *capogiro*, lett. 'giro del capo'); solo nel caso di *capocenso* (lett. 'censito per la persona'), esso si riferisce a un individuo (tipo semantico PERSONA).

I composti come *caporicciare* e *capolevare* rappresentano istanziazioni dello schema in (d) dei composti  $[N-V]_V$  subordinativi con testa a destra; anche in questo caso *capo* indica la 'testa', come esemplificato in (31).

- (31) a. «Andreuccio dentro sicuramente passato, gli venne per ventura posto il piè sopra una tavola, la quale dalla contraposta parte sconfitta dal travicello sopra il quale era, per la qual cosa *capolevando* questa tavola con lui insieme se n'andò quindi giuso» (Boccaccio, *Decameron*)
  - b. «[...] et accapricciare è levare li capelli ritti, come avviene per paura; cioè *caporicciare*; e però si dice: Io ebbi un grande raccapriccio; cioè uno arricciamento

de' capelli del capo, che significa la paura» (Francesco di Bartolo dal Buti, Commento all'Inferno)

Nell'esempio (31a), *capolevare* allude al 'volgersi sottosopra, capovolgersi' di Andreuccio da Perugia nella celebre novella di Boccaccio; in (31b) si spiega letteralmente il significato di *caporicciare* come la sensazione scatenata dal sentimento della paura che porta i capelli a rizzarsi sul capo.

Oltre che come costituente di composti subordinativi, *capo* può combinarsi anche con aggettivi dando luogo a composti attributivi esocentrici del tipo di *capinera* o *capocanino* ('composti possessivi', cfr. § 4.2.4): si tratta di epiteti che designano individui o animali sulla base di qualità possedute.

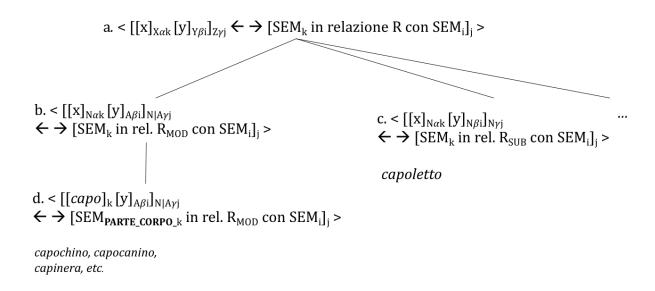


Figura 34 Sotto-schema con *capo*- della Costruzione<sub>COMP</sub> dei composti N-A esocentrici

Nella Gerarchia di eredità illustrata nella figura 35, i composti come *capinera* costituiscono un sotto-schema della Costruzione<sub>COMP</sub> dei composti [N-A]<sub>N|A</sub> esocentrici (b); questa, a sua volta, rappresenta uno dei sotto-schemi della più astratta Costruzione<sub>COMP</sub> dei composti esocentrici (a), tra i quali ad esempio vi è lo schema che istanzia i composti [V-N]<sub>N</sub>. Anche in questo tipo di forme, *capo* ha sempre il significato di 'testa': così la *capinera* è un uccello 'con una calotta nera sulla fronte e la cervice' e *capochino* è usato per designare un 'uomo afflitto, triste'.

Tra i composti esocentrici con capo vi è inoltre capoletto, unico caso attestato nel dataset di composto esocentrico [N-N]<sub>N</sub> subordinativo (vd. (c)); come esemplificato in (32), esso indica un 'drappo da appendere in capo al letto':155

(32) «Quindi essendo in più luoghi per fatti letti e tutti dal discreto siniscalco di sarge francesche e di *capoletti* intorniati e chiusi, con licenzia del re, a cui piacque, si poté andare a dormire» (Boccaccio, *Decameron*)

Infine, il dateset Compos ita TLIO attesta la forma capopiede 'sottosopra, al contrario'. 156

(33) «[...] e poi presa la bandiera fu strascinata per lo campo, come stimava messer Farinata Uberti, e poi in Firenze il simile, ed appiccata capopiedi come era usanza [...]» (Marchionne, Cronaca fiorentina)

Come esemplificato in (33), il composto (unico di tutte le forme con capo) è usato in funzione avverbiale: i costituenti capo e piede sembrano essere legati dalla stessa relazione di *sottosopra*.

L'analisi proposta in questo paragrafo ha messo in luce che l'uso in composizione del sostantivo capo può esser fatto risalire all'epoca medievale, quando iniziano a diffondersi combinazioni *caput* + x in cui il primo costituente può esprimere il valore di 'testa', 'inizio/estremità' o 'elemento/individuo preminente'. Tali valori si riscontrano anche in italiano antico, nell'uso di capo all'interno di parole composte di varie tipologie (es. [N-N] subordinativi, [N-A] attributivi esocentrici, etc.). Analizzando questo tipo di forme in ottica costruzionista, si è osservato che i composti con capo possono essere rappresentati come sotto-schemi di Costruzioni più astratte (es. la Costruzione dei composti [N-A]<sub>NIA</sub> esocentrici) e che tali sotto-schemi possono presentare delle restrizioni semantiche che limitano sia i diversi valori di capo sia la semantica del secondo costituente: ad esempio, si è osservato che nella Costruzione COMP dei composti [N-N]<sub>N</sub> subordinativi con testa a sinistra, quando *capo* assume il valore di 'autorità/guida' (es. capofamiglia), il secondo costituente indica sempre un gruppo di persone o un luogo; quando presenta quello di 'inizio/estremità' (es. capomese), il secondo costituente può essere un periodo di tempo, un artefatto o un testo.

<sup>&</sup>lt;sup>155</sup> TLIO, s.v. capoletto.

<sup>&</sup>lt;sup>156</sup> TLIO, s.v. *capopiede*.

L'uso del termine per indicare primariamente la parte superiore del corpo all'interno di parole composte trova riscontro in altre lingue, quali il francese (es. *chef lieu* 'capoluogo'), l'inglese (es. *chief engineer* 'ingegnere capo'), il tedesco (es. *Hauptbahn* 'linea (ferroviaria) princiale'), l'olandese (es. *hoofd-pijn* 'mal di testa'), etc. Come già osservato al § 2.5, in ambito costruzionista sono state studiate le Costruzioni con *Haupt*-in tedesco e *hoofd*- in olandese (cfr. Booij 2017: 22; Masini 2016: 84). In particolare, in olandese la forma *hoofd* può occorrere con il significato di 'testa' (sia come parola autonoma sia come costituente di composto) o con quello più astratto di 'principale' (ma, in questo caso, solo come costituente di composto). L'uso in composizione di *hoofd*, definito da Booij (2017: 22) un caso di *construction-dependent polysemy*, può essere rappresentato attraverso gli strumenti della CxM come riportato nella figura 36.

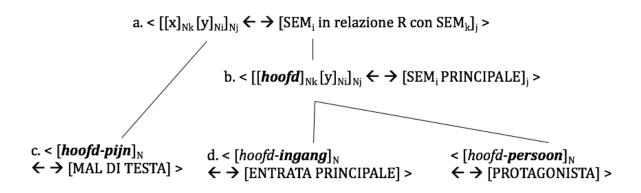


Figura 35. Gerarchia dei composti con *hoofd*- in olandese Fonte: adattato da Masini (2016: 84)

Nella figura 36 è rappresentata la Gerarchia dei composti con *hoofd*- in olandese: come già osservato al § 2.5, il livello più alto è rappresentato dalla Costruzione generica dei composti N-N dell'olandese (a), che instanzia produttivamente nuovi composti, tra i quali *hoofd-pijn* 'mal di testa' (c) in cui *hoofd*- presenta il significato originario di 'testa'. Le forme in cui *hoofd*- presenta il significato più astratto di 'principale' (d) non dipendono direttamente da (a) ma da un livello intermedio (b), ossia una Costruzione semi-specificata, in cui il primo elemento *hoofd*- è fisso e il secondo è variabile, considerata da Booij (2005) un caso di sotto-schema semi-specificato. In tedesco, l'elemento *Haupt*-, sostituito dalla parola *Kopf* per l'espressione del significato 'testa', è

ormai utilizzato soltanto come prefisso con il valore di 'principale' (come in *Hauptbahn* 'linea (ferroviaria) principale') e non ha più una controparte libera.

L'emergere di sotto-schemi semi-specificati con *hoofd*- e *Haupt*- può essere considerato un caso di Costruzionalizzazione (ossia la formazione di una nuova Costruzione, cfr. Traugott & Trousdale 2013: 22), dal momento che tali elementi hanno sviluppato un significato più astratto di 'principale'. Confrontando il caso dell'olandese e quello dei composti dell'italiano antico, emergono alcuni aspetti utili a delineare le caratteristiche di questo tipo di forme in italiano. In primis, va notato che nei composti italiani il costituente *capo*- può presentare fin dalle origini tre valori (ossia quello di 'testa', quello di 'inizio o origine di qualcosa' e quello di 'elemento preminente rispetto ad altri'): l'analisi dei tipi semantici dei costituenti ha poi messo in luce che il valore presentato da capo- dipende dalla semantica della base. Come in olandese, quindi, anche in italiano capo- può talvolta assumere significati astratti come quello di 'inizio': contrariamente ai casi attestati nelle lingue germaniche (in cui hoofd- e Hapt- presentano il significato astratto di 'principale' soltanto all'interno di un composto), tuttavia, capo esprime tale valore anche quando usato come forma libera, al di fuori di una parola composta. I casi in cui capo- presenta il valore di 'inizio' sono comunque ancora piuttosto limitati, e si affiancano agli altri valori. L'analisi di queste forme da un punto di vista diacronico (cfr. § 4.3) permetterà di valutare se i diversi valori di capo individuati per i composti dell'italiano antico si stabilizzeranno dando vita a constructional idiom come nelle lingue germaniche.

## 4.2.6 Quid manet? Un bilancio sulla composizione dal latino all'italiano antico

I precedenti paragrafi hanno fornito un quadro della composizione dell'italiano in chiave costruzionista. Di seguito si riepilogano gli aspetti emersi durante l'analisi, utili per delineare un profilo delle parole composte della fase antica e a mettere in luce il rapporto con la composizione latina, di cui sono state descritte le principali caratteristiche al § 4.1. La tabella 38 illustra come si distribuiscono i composti del dataset *Compos\_ita\_TLIO* nelle quattro famiglie di Costruzioni individuate al § 4.2.

Composti co sinistra	n testa a	Composti con destra	testa a	Composti coordinat		Composti esoco	entrici
[N-N] <sub>N_SUB</sub>	26	[N-N] <sub>N_SUB</sub>	38	[N-N] <sub>N</sub>	10	[V-N] <sub>N_SUB</sub>	77
$[N-A]_{N_{-MOD}}$	26	$[N-P.PRES]_{N A}$	4	$[A-A]_{N A}$	6	$[V-V]_{N\_{COORD}}$	6
$[N-N]_{N\_{MOD}}$	7	[N-P.PASS] <sub>N A</sub>	3	Totale	16	$[N-A]_{N\_{MOD}}$ $[N-N]_{N A\_{MOD}}$	10 2
Totale	59	$[N-V]_V$ $[A-N]_{N A AVV}$ $[A-P.PASS]_{N A}$ $[AVV-X]_X$ $Totale$	5 119 7 105 281			Totale	95

Tabella 38 Dataset Compos\_ita\_TLIO: distribuzione dei type rispetto al tipo di Costruzione COMP

I dati quantitativi riportati nella tabella 38 mettono in luce la diversa produttività, in termini di type, delle diverse Costruzionicome dell'italiano antico. Un primo aspetto rilevante riguarda la notevole presenza di composti con testa a destra: si è infatti osservato che questa famiglia di Costruzioni<sub>COMP</sub>, che presenta l'ordine dei costituenti tipico della composizione latina, si articola in una serie di sotto-schemi, alcuni dei quali molto produttivi. È il caso delle Costruzioni che presentano un aggettivo o un avverbio anteposti a una base variabile (es. buonora, mezzaluna, malafede, benpensante, etc.): questi schemi, a loro volta, istanziano schemi semi-specificati, ad es. quello dei composti [mal(e)-x] o [ben(e)-x] in cui il primo costituente generalmente esprime un valore valutativo rispetto alla semantica della base. La fortuna di questa famiglia di Costruzionicomp può essere attribuita alla persistenza del modello latino: nel corso dell'analisi si è infatti osservato che quasi tutti i sotto-schemi sono istanziati da un certo numero di parole composte di origine latina, adattate fonomorfologicamente all'italiano, che fungono da modello per la creazione di forme native (es. i composti del tipo pescivendolo, i composti con mal(e)-, etc.). Inoltre, va osservato che l'ordine marcato dei costituenti, proprio di questa famiglia di Costruzionicome, può essere considerato un fattore in grado di migliorare la coesione del composto, non sempre garantita nel caso dei composti con testa a sinistra (soprattutto nel caso del tipo [N-A]<sub>N</sub> che è spesso difficilmente distinguibile da un sintagma). La Costruzione dei composti con testa a sinistra sembra infatti poco produttiva in italiano antico: gran parte dei composti di questo tipo ha infatti origine dall'univerbazione di espressioni sintagmatiche latine e non costituisce il risultato di un meccanismo di formazione di parola produttivo.

Il secondo 'motore' della composizione dell'italiano antico è costituito dalla  $Costruzione_{COMP}$  dei composti  $[V-N]_N$ : i dati analizzati in questo capitolo confermano che la composizione verbonominale è disponibile e produttiva fin dalla fase più antica della lingua; in questo caso, però, la diffusione di questo tipo di forme non è imputabile al modello latino, ma può essere considerata un'innovazione delle lingue romanze. Va inoltre osservato che la fortuna di questa  $Costruzione_{COMP}$  può essere dovuta anche alla particolare coesione formale e semantica e alla regolarità di questo tipo di forme, che si specializza dal punto di vista semantico nell'espressione di referenti umani o artefatti.

Le altre  $Costruzioni_{COMP}$  attestate nel dataset costituiscono meccanismi decisamente meno produttivi, alcuni strettamente legati al modello latino (es. i composti coordinativi  $[A-A]_{A|N}$  del tipo *gallogreco*, *libifenicio*), altri considerabili innovazioni dell'italiano (es. i composti  $[V-V]_N$  del tipo *andirivieni*, etc.).

Il quadro emerso dall'analisi del dataset *Compos\_ita\_TLIO* restituisce quindi un profilo della composizione dell'italiano antico come essenzialmente ancora legata al modello latino: l'influenza della composizione latina, da cui comunque l'italiano si discosta dal punto di vista formale per quanto riguarda la natura dei costituenti,<sup>157</sup> è ancora pienamente riconoscibile nella fase antica e in grado di orientare la formazione di nuove parole composte native verso la Costruzione<sub>COMP</sub> dei composti con testa a destra. È essenzialmente grazie a questa Costruzione<sub>COMP</sub> e a quella dei composti [V-N]<sub>N</sub> che l'italiano inizia ad ampliare il proprio lessico attraverso (oltre che la derivazione) la formazione di parole composte.

Nel prossimo paragrafo, si analizzeranno in prospettiva diacronica le linee di sviluppo delle diverse Costruzioni<sub>COMP</sub> dalla fase antica al XX secolo.

## 4.3 Analisi diacronica

Nei seguenti paragrafi, verrà analizzata l'evoluzione diacronica di ciascuna famiglia di Costruzioni<sub>COMP</sub>, di cui saranno messe in luce le proprietà formali e semantiche, oltre che

<sup>&</sup>lt;sup>157</sup> Si è però osservato che alcuni aspetti formali della composizione latina classica sopravvivono come elementi residuali anche nella composizione italiana nativa: in particolare, la vocale finale –*i*– del primo costituente, che occorre in composti come *pettirosso*, *capinera*, sul modello del *linking element* latino.

la frequenza type/token, sulla base dei composti estratti dal corpus CoDIt\_com e raccolti nel dataset *Compos\_ita*.

## 4.3.1 La Costruzione<sub>COMP</sub> dei composti con testa a sinistra

Nella fase antica, la Costruzione<sub>COMP</sub> dei composti con testa a sinistra è istanziata dai seguenti sotto-schemi:

- 1.  $[N-N]_{N\_SUB}$  (es. *capofamiglia*)
- 2.  $[N-A]_{N\_MOD}$  (es. terracotta, camposanto)
- 3.  $[N-N]_{N\_MOD}$  (es. *pescespada*)

L'analisi fornita al § 4.2.1 ha messo in luce come le istanziazioni di questa famiglia di Costruzioni<sub>COMP</sub> si distribuiscano essenzialmente tra il sotto-schema (1) e il (2); rari sono i composti  $[N-N]_{N\_MOD}$  attestati nel dataset *Compos\_ita\_TLIO*.

Tutti i sotto-schemi attestati nella fase antica sono istanziati senza soluzione di continuità da nuovi composti attestati nel CoDIt\_com; inoltre, i dati estratti dal corpus permettono di individuare due ulteriori sotto-schemi:

- 4.  $[A-N]_{A|N}$  (es. rosso sangue, nero carbone)
- 5.  $[A-A]_{A|N}$  (es. *verdebruno*)

Sulla base dei dati estratti dal CoDIt\_com è possibile ordinare tali sotto-schemi in una Gerarchia di eredità, rappresentata nella figura 37.

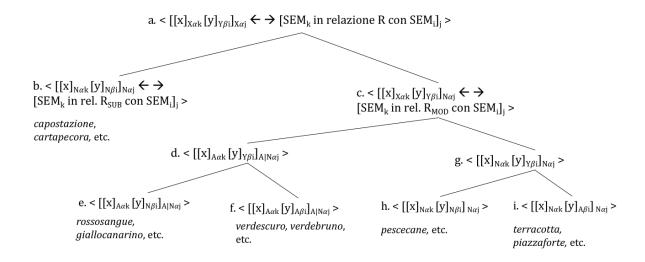


Figura 36 Gerarchia di eredità della  $Costruzione_{COMP}$  dei composti con testa a sinistra attestati nel  $CoDIt\_com$ 

Come illustrato dalla figura 37, la Gerarchia di eredità che rende conto dei diversi gradi di astrattezza della Costruzione $_{\text{COMP}}$  dei composti con testa a sinistra attestati nel CoDIt\_com è più articolata rispetto a quella proposta per l'italiano antico (cfr. § 4.2.1). In particolare, lo schema dei composti che presentano una relazione di modificazione (rappresentato in (c)) si arricchisce del sotto-schema con testa aggettivale e modificatore variabile (cfr. schema in (d)), che a sua volta si specifica nei due sotto-schemi [A-N]<sub>A|N</sub> e [A-A]<sub>A|N</sub>.

La tabella 39 illustra come si distribuiscono i type e i token di ciascun sotto-schema attestato nel CoDIt\_com.

	$[N-N]_{N\_SUB}$	[N-N] <sub>N_MOD</sub>	[N-A] <sub>N_MOD</sub>	$[A-N]_{A A\_{MOD}}$	[A-A] <sub>A N_MOD</sub>	tot
type	42	17	48	5	8	120
token	533	216	1.243	68	52	2.112

Tabella 39 Costruzione<sub>COMP</sub> dei composti con testa a sinistra: distribuzione dei type e dei token attestati nel CoDIt\_com rispetto ai sotto-schemi

Come si osserva dai dati riportati nella tabella 39, se osservati all'interno di tutto il corpus (non considerando la scansione cronologica), i composti con testa a sinistra si distribuiscono in maniera più omogenea tra i tre sotto-schemi rispetto all'italiano antico.

In particolare, va notato che il sotto-schema  $[N-N]_{N\_MOD}$ , la cui produttività era minima nella fase antica, istanzia circa il 17% dei composti attestati nel corpus. Analogamente a quanto osservato per l'italiano antico, si può notare che i due sotto-schemi  $[N-N]_{N\_SUB}$  e  $[N-A]_{N\_MOD}$  presentano quasi lo stesso numero di type; tuttavia, il secondo mostra una maggior produttività in termini di token, ossia istanzia composti più frequenti all'interno del CoDIt\_com.

Di seguito si descrivono in ottica diacronica le proprietà formali e semantiche e la frequenza di ciascun sotto-schema.

Il sotto-schema  $[N-N]_{N\_SUB}$  istanzia 42 composti i cui costituenti sono legati da una relazione di subordinazione. I cinque composti più frequenti all'interno del CoDIt\_com sono riportati ed esemplificati nella tabella  $40.^{158}$ 

	1	2	3	4	5	tot	Esempio
pannolino	12	104	20	2	31	169	«Mandoti le sacha il primo dì che cci sarà, da cciò ti rimanderò di quelle che ttu m'ai mandate di chostà e altre chose che io ti voglio mandare, chome pannilini e schanpoli di panolano e altre chose» (Datini, Lettere alla moglie Margherita)
acquavite	-	2	1	7	39	49	«Paragonerei la frode all' <i>acquavite</i> : pare che sul momento dia forza, ma poi vi lascia pi spossato di prima» (D'Azeglio, <i>Racconti</i> )
cartapecora	-	5	8	7	9	29	«[] ma la pelle del viso era gialla come la <i>cartapecora</i> dei libri di vostro nonno e si piegava tutta in mille rughe» (Serao, <i>Leggende napoletane</i> )
mappamondo	6	3	7	6	6	28	«[] lo quale tramontare chiaro appare nel <i>mappamondo</i> considerando le elevazioni de' segni e li suoi occasi, come si può vedere nella spera materiale»

-

 $<sup>^{158}</sup>$  In questa e nelle seguenti tabelle di questo capitolo, si sottolineano i composti già attestati nella fase antica, nel dataset  ${\it Compos\_ita\_TLIO}.$ 

	1	2	3	4	5	tot	Esempio
							(Della Lana, Commento alla Commedia)
<u>caporione</u>	1	1	4	10	8	24	«[] perché era alla difesa una parte della gioventù romana sotto i loro <i>caporioni</i> e bandiere del popolo» (Guicciardini, <i>Storia</i> d'Italia)

Tabella 40 Composti [N-N]<sub>N\_SUB</sub> con testa a sinistra più frequenti nel CoDIt\_com (token)

Come si osserva dalla tabella 40, il composto  $[N-N]_{N\_SUB}$  più frequente è *pannolino* (*panno+lino*): il termine indica un 'piccolo panno in lino, salvietta' ed è particolarmente frequente nel secondo sotto-corpus; è infatti solo negli ultimi decenni che il composto ha assunto il significato attuale di 'striscia di stoffa più o meno imbottita, o, oggi più comunemente, di ovatta, cellulosa o altro materiale assorbente, spesso sagomata a forma di mutandina, usata per assorbire e contenere urina e feci di lattanti e neonati' (cfr. De Mauro 2004, s.v. *pannolino*).

Alcuni tra i composti più frequenti hanno origine latina: è il caso di *acquavite* (da *aqua vitae*), *mappamondo* (da *mappa mundi*), il composto *cartapecora* ha origine dall'espressione antica 'carta di pecora', attestata nei primi due sotto-corpora e di seguito esemplificata:

(34) a. «E sentendosi cuocere e alquanto movendosi, parve nel muoversi che tutta la cotta pelle le s'aprisse e ischiantasse, come veggiamo avvenire d'una *carta di pecora* abrusciata se altri la tira» (Boccaccio, *Decameron*)

b. «[...] e farne una somma in su uno libro, il quale sia tutto di *carta di pecora*; e in su quello libro sempre si possa vedere tutte le spese che si sono fatte anno per anno» (Filarete, *Trattato di architettura*)

Gli esempi riportati in (34) mostrano come nelle fasi più antiche della lingua la pergamena fosse indicata dal nome sintagmatico *carta di pecora*, che in seguito si è ridotta ai due costituenti senza la preposizione *di*.

Nel caso di *caporione* si tratta di un composto costituito da *capo* più un nome subordinato, in cui *capo* indica un referente umano che detiene l'autorità o il comando

rispetto a un gruppo di persone (es. *capofamiglia*, *capopolo*) o a un luogo (*caporione* 'a Roma, capo di rione'): come osservato al § 4.2.5.2, questo tipo di forme costituisce un sotto-schema della Costruzione<sub>COMP</sub> dei composti [N-N]<sub>N\_SUB</sub> con testa a sinistra, che presenta due restrizioni semantiche relative al tipo semantico di *capo* (PERSONA) e a quello del secondo costituente (GR\_PERS o LUOGO).

Dal punto di vista quantitativo, la frequenza dei type e dei token della Costruzione<sub>COMP</sub> è illustrata dai grafici nelle figure 38 e 39.

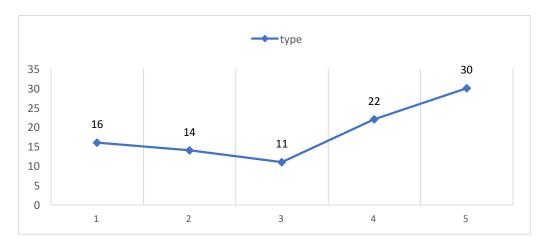


Figura 37 Composti  $[N-N]_{N\_SUB}$ : distribuzione dei type nei sotto-corpora del CoDIt\_com

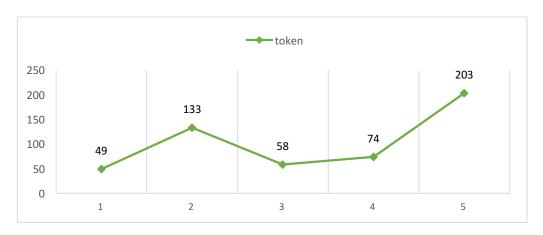


Figura 38 Composti  $[N-N]_{N\_SUB}$ : distribuzione dei token nei sotto-corpora del CoDIt\_com

Il grafico relativo alla frequenza dei type mette in luce un progressivo incremento di questo tipo di composti nel corso dei secoli, nonostante si registri una lieve diminuzione in corrispondenza del terzo sotto-corpus. L'evoluzione della frequenza dei token è invece

meno regolare, con due picchi di crescita: uno durante il periodo individuato dal secondo sotto-corpus, l'altro nella fase più recente; anche per la frequenza dei token si registra infatti un deciso aumento nell'ultimo sotto-corpus.

Dal punto di vista semantico, i composti si distribuiscono come illustrato nella tabella 41 rispetto al tipo semantico di output.

ANIMALE	ARTEFATTO	EVENTO	LUOGO	OGG_NAT	CIBO	PERSONA
porcospino	pannolino cartapecora mappamondo pannolano capoletto capofuoco mazzapicchio saccapane	finimondo pallacorda	capotavola pianoterra caciovacca capopagina capoverso	acquarosa <u>capelvenere</u>	caciocavallo	caporione capostazione capostazione caposervizio caposervizio capobanda capoufficio capocomune caposquadra capodivisione caposcuola capofila caposetta capocaccia capofamiglia capolista capovilla capogestione capoparte caposcena capotamburo
1	8	2	5	3		20

Tabella 41 Composti  $[N-N]_{N\_SUB}$  attestati nel CoDIt\_com: distribuzione rispetto al tipo semantico di output

La Costruzione<sub>COMP</sub> [N-N]<sub>N\_SUB</sub> istanzia un ristretto numero di nomi che indicano luoghi, oggetti naturali, cibi ed eventi; prevalentemente, essa istanzia nomi che designano artefatti e referenti umani. In particolare, nel primo caso si tratta di forme di antica attestazione (*capoletto*, *mazzapicchio*, <sup>159</sup> *pannolino*, *saccapane*) e di un composto di origine latina (*mappamondo*). Nel secondo caso, si tratta esclusivamente di forme che presentano *capo* come primo costituente e che, come già osservato al § 4.2.5.2, costituscono un sotto-schema semi-specificato della Costruzione<sub>COMP</sub> dei composti [N-N]<sub>N\_SUB</sub> con testa a sinistra, produttivo fin dalla fase antica (cfr. ad es. *caporione*,

\_

<sup>&</sup>lt;sup>159</sup> Il composto *mazzapicchio* è costituito dal sostantivo *mazza* e dalla forma deverbale *picchio* (da *picchiare*) ed è usato con il significato di 'mazza per picchiare'.

capofamiglia attestati nel dataset *Compos\_ita\_TLIO*). Tale sotto-schema istanzia nomi che designano referenti umani (tipo semantico PERSONA) che occupano una posizione di autorità rispetto a un gruppo di persone o un luogo.

Analizzando la frequenza dei type e dei token delle forme in cui *capo* indica un referente umano, è possibile indagare l'evoluzione in termini di produttività di questo sottoschema (figure 40 e 41).

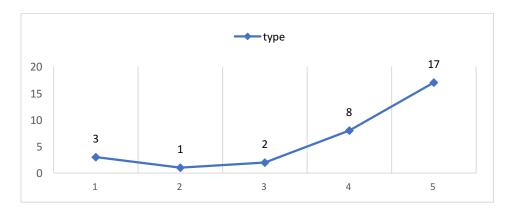


Figura 39 Sotto-schema [capo<sub>PERSONA</sub>-x]<sub>PERSONA</sub>: distribuzione dei type nei sotto-corpora del CoDIt\_com



Figura 40 Sotto-schema [capopersona-x]persona: distribuzione dei token nei sotto-corpora del CoDIt\_com

Come illustrato dai grafici nelle figure 40 e 41, lo schema [capopersona-x]persona è scarsamente produttivo nel periodo compreso tra le Origini e il terzo sotto-corps: nell'analisi dei composti dell'italiano antico fornita al § 4.2.5.2 si era infatti già osservato che, tra i composti con capo, quelli in cui esso designa un referente umano erano un numero piuttosto limitato a confronto delle forme in cui il primo costituente presenta il

significato di 'testa'. È quindi solo a partire dal periodo 1533-1691 che questo schema diventa realmente produttivo, in termini sia di token sia di type, fino alla fase contemporanea in cui è stabilmente affermato. I composti più frequenti all'interno del CoDIt\_com sono riportati nella tabella 42.

	1	2	3	4	5	totale
caporione	1	1	4	10	8	24
capostazione	-	-	-	-	22	22
capopopolo	1	-	-	3	14	18
caposervizio	-	-	-	-	9	9
capobanda	-	-	-	4	4	8

Tabella 42 Composti [capo<sub>PERS</sub>-x]<sub>PERS</sub> più frequenti nel CoDIt\_com

Come mostrato dai dati riportati nella tabella 42, ad eccezione di *caporione* e *capopolo*, attestati fin dalla fase antica, si tratta di composti di recente formazione. Questo tipo di forme è stato oggetto di studio in ottica sincronica da parte di Masini (in stampa) che ha analizzato la competizione tra Costruzioni sintagmatiche con *capo* (in particolare, i nomi sintagmatici Nome+Preposizione+Nome, es. *capo dello Stato*) e Costruzioni di natura morfologica (ossia i composti Nome Nome del tipo *capofamiglia*). Tale studio ha messo in luce come parole morfologicamente complesse ed espressioni multiparola siano meccanismi effettivamente in competizione, in grado di determinare un effetto di blocco bidirezionale: la presenza di nome sintagmatico ben stabilizzato nel lessico (o, in ottica costruzionista, nel Costruzionario) come *capo dello stato* (o *capo di stato*) può di fatto bloccare la creazione di un corrispettivo composto, ossia \**capostato*, e viceversa. <sup>160</sup> La prospettiva diacronica offerta dal CoDIt\_com permette di approfondire tale analisi, confrontando il numero delle occorrenze dei composti attestati nel corpus con quello delle possibili combinazioni '*capo di/del* x' corrispondenti, come illustrato dalla tabella 43.

Composti [capo <sub>PERS</sub> -x] <sub>PERS</sub>	occ.	Combinazioni <i>capo di/del</i> x	occ.
caporione	24	capo del rione	0
capostazione	22	capo della stazione	0

1

<sup>&</sup>lt;sup>160</sup> I risultati di questo studio supportano inoltre l'ipotesi per la quale sia le parole sia le espressioni multiparola vengono entrambe immagazzinate nel lessico mentale, in virtù della condivisa funzione denominatrice.

Composti [capo <sub>PERS</sub> -x] <sub>PERS</sub>	occ.	Combinazioni <i>capo di/del</i> x	occ.
capopopolo	18	capo del popolo	1
caposervizio	9	capo del servizio	0
capobanda	8	capo della banda	0
capoufficio	8	capo dell'ufficio	0
capocomune	7	capo del comune	0
caposquadra	7	capo della squadra	0
capodivisione	5	capo della divisione	0
caposcuola	4	capo della scuola	1
capofila	3	capo della fila	0
caposetta	3	capo della setta	11
capocaccia	2	capo della caccia	0
capofamiglia	2	capo della famiglia	8
capolista	2	capo della lista	0
capovilla	2	capo della villa	0
capogestione	1	capo della gestione	0
capoparte	1	capo della parte	0
caposcena	1	capo della scena	0
capotamburo	1	capo del tamburo	0

Tabella 43 Competizione tra composti e combinazioni *capo di/del x* (dati CoDIt\_com)

Come illustrano i dati riportati nella tabella 43, eccetto rari casi, nel CoDIt\_com non risultano attestati nomi sintagmatici corrispondenti a composti con *capo*-. Vanno però segnalate due eccezioni, *caposetta* (3) vs *capo della setta* (11); *capofamiglia* (2) vs *capo della famiglia* (8), le occorrenze dei due (potenziali) nomi sintagmatici si distribuiscono tra i sotto-corpora del CoDIt\_com come riportato di seguito nella tabella 44.

	1	2	3	4	5	Esempi
capo della setta	-	1	-	-	10	«Dunque il supremo capo della setta nel reame di Napoli, non poteva essere che un grande unitario, non poteva essere che uno []» (Settembrini, Ricordanze)
capo della famiglia	-	1	-	3	4	«La sua prima visita fu pel <i>capo della famiglia</i> ; ma, giunto dinanzi al portone del palazzo, vide con stupore che era chiuso, col solo sportello aperto» (De Roberto, <i>I Vicerè</i> )

Tabella 44 Combinazioni *capo di/del x* in competizione con composti con *capo*-: distribuzione delle occorrenze nei sotto-corpora del CoDIt\_com

Entrambi le combinazioni 'capo di/del x' sono attestate prevalentemente nell'ultimo sotto-corpus, indizio del fatto che non si tratta di un caso in cui un nome sintagmatico N-P-N si è poi ridotto a N-N (vd. sopra, il caso di carta di pecora > cartapecora) ma di un'effettiva competizione tra le due forme, a vantaggio di quella sintagmatica, nel periodo individuato dal sotto-corpus più recente (1841-1947). Sulla base dei dati estratti dal CoDit\_com, si può notare come, nel caso di caposetta, il meccanismo perdente nella competizione sia la composizione; sembra simile la situazione di capofamiglia, anche se va considerato che attualmente in italiano contemporaneo la forma decisamente più diffusa è il composto capofamiglia. Nel complesso, i dati estratti dal CoDit\_com supportano quindi la tesi sostenuta da Masini (in stampa), secondo cui Costruzioni sintagmatiche e Costruzioni morfologiche possono trovarsi in competizione e bloccare la formazione del proprio corrispettivo, pur mettendo in luce rari casi in cui è attestata la compresenza di entrambe (seppur a vantaggio di un tipo, es. capo della famiglia).

Tra i composti in cui *capo*- non indica un referente umano, il CoDIt\_com attesta le forme *capoletto*, *capoverso*, *capotavola*, *capopagina*, *capofuoco*, in cui il primo costituente indica una 'estremità': ad eccezione di *capofuoco*,<sup>162</sup> si tratta di composti già attestati in italiano antico, che si tramandano nel corso dei secoli, senza però costituire il modello per la creazione di nuove forme. Il sotto-schema semi-specificato individuato nella fase antica, costituito da *capo* nel senso di 'estremità' e un nome, sembra quindi aver esaurito la produttività nel periodo successivo alla fase antica, dal momento che non risultano attestate nel CoDIt\_com sue nuove istanziazioni.

Nel complesso, se si escludono i composti con *capo*- e quelli già attestati nella fase antica, i dati estratti dal CoDIt\_com permettono di osservare che la Costruzione<sub>COMP</sub> dei composti con testa a sinistra istanzia un numero limitato di forme nel corso dei secoli: è infatti lo schema semi-specificato con *capo*- a trainare la produttività dell'intera Costruzione<sub>COMP</sub>.

\_

<sup>&</sup>lt;sup>161</sup> Nel corpus di italiano contemporaneo itTenTen, il composto *capofamiglia* occorre 6.209 volte; la combinazione *capo della famiglia* registra 474 occorrenze.

<sup>&</sup>lt;sup>162</sup> Il composto, attestato (anche nella forma *capifuoco*) a partire dal terzo sotto-corpus, ha lo stesso significato di *alare*, ossia indica uno strumento di ferro usato per muovere o sorreggere la legna nel caminetto. In questo caso il significato di *capo* non è perfettamente chiaro, ma probabilmente è legato a quello di 'estremità'.

Lo schema [N-N]<sub>N\_MOD</sub> istanzia 17 composti costituiti da due sostantivi legati da una relazione di modificazione. Come già osservato al § 4.2.1, questa Costruzione<sub>COMP</sub> è scarsamente produttiva nella fase antica, ma comunque già presente: in particolare, il dataset *Compos\_ita\_TLIO* attesta le forme *pescespada*, *pesceporco* e *arcamensa*, delle quali soltanto la prima è attestata nei secoli successivi rappresentati dai sotto-corpora del CoDIt\_com. I cinque composti che occorrono più frequentemente nel CoDIt\_com sono riportati nella tabella 45.

Composti [N-N] <sub>N_MOD</sub>	1	2	3	4	5	tot	Esempio
pescecane	-	1	-	-	52	53	«Raggiata e rombo, occhiata e pescecane, / la triglia, il ragno e 'l corvallo e 'l salmone, / lo scurpin colle punte aspre e villane []» (Pulci, Morgante)
							«Il palombo (Mustelus) è un pesce della famiglia degli squali ossia de' pescicani, e perciò in alcuni paesi il palombo si chiama pescecane» (Artusi, La scienza in cucina e l'arte di mangiar bene)
melarancia	-	35	1	-	41	40	«E se quelle <i>melarancie</i> si gastano, mandamene anche un paniere e daròne a Nicholò, e ttu anche ne dà chostì a chi tti pare» (Datini, <i>Lettere alla moglie Margherita</i> )
ceralacca	-	-	3	26	3	32	«La ceralacca adunque ed il vetro pigliano collo strofinio elettricità, ma l'una è all'altra opposta []» (Scina, Della fisica sperimentale)
cavolfiore	-	-	1	-	29	30	«Ma noi ci metteremo pure i broccoli, e i <i>cavolifiori</i> Buoni, eh?» (Verga, <i>I Malavoglia</i> )
<u>pescespada</u>	-	-	13	-	8	21	«Al pesce spada / O terror d'ogni rete e d'ogni nassa, / pesce guerrier, che la lucente spada / vibri per l'acque, e de l'algosa strada / cerchi la parte

Composti [N-N] <sub>N_MOD</sub>	1	2	3	4	5	tot	Esempio
							più riposta e bassa» (Marino, <i>Rime marittime</i> )

Tabella 45 Composti [N-N]<sub>N MOD</sub> più frequenti nel CoDIt\_com

Il composto  $[N-N]_{N\_MOD}$  più frequente è *pescecane*, attestato a partire dal secondo sottocorpus (ma come hapax) e in maniera consistente nell'ultimo periodo: in particolare, le occorrenze di questo composto si concentrano per la maggior parte in un'opera, *Pinocchio* di Collodi. Il sostantivo *pesce* occorre come primo costituente di questo tipo di composti anche in *pescespada*, che come già ricordato è uno degli esempi più antichi di composto  $[N-N]_{N\_MOD}$ , *pescecavallo* (1), *pesceluna* (2) e *pescesega* (1). Nel caso di *ceralacca*, il composto indica un tipo di cera usata per sigillare: il modificatore *lacca* ha origine araba (ar. lakk) ed è attestato già in italiano antico con il significato di 'gommoresina granulosa prodotta in oriente'.  $^{163}$ 

Il composto *melarancia* indica il frutto dell'arancio: il sostantivo *arancia* modifica la testa del composto classificando il tipo di frutto.<sup>164</sup> Anche nel composto *cavolfiore*, il secondo costituente classifica il tipo di cavolo, in questo caso quello che ha una infiorescenza carnosa compatta;<sup>165</sup> un caso simile è quello di *cavolrapa* (hapax).

Dal punto di vista quantitativo, la Costruzione $_{COMP}$  [N-N] $_{N\_MOD}$  segue l'evoluzione delineata nei grafici riportati nelle figure 42 e 43.

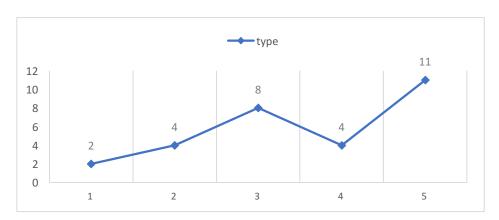


Figura 41 Composti [N-N]<sub>N MOD</sub>: distribuzione dei type nei sotto-corpora del CoDIt\_com

-

<sup>163</sup> TLIO, s.v. lacca1.

<sup>&</sup>lt;sup>164</sup> Non si tratta quindi di incrocio tra due specie di frutti come in *mandarancio*, in cui i costituenti sono legati da una relazione di coordinazione.

<sup>165</sup> De Mauro (2004), s.v. cavolfiore.



Figura 42 Composti [N-N]<sub>N\_MOD</sub>: distribuzione dei token nei sotto-corpora del CoDIt\_com

Come illustrato dai grafici, la produttività della Costruzione $_{\text{COMP}}[N-N]_{N\_\text{MOD}}$ , in termini sia di type sia di token, si fa significativa solo a partire dal terzo sotto-corpus e, in particolare, nella fase più recente. La frequenza dei type registra un picco in corrispondenza del terzo sotto-corpus, a cui non corrisponde un aumento in termini di frequenza dei token, indizio del fatto che si tratta di composti che occorrono un numero limitato di volte nei testi.

Come illustrato nella tabella 46, composti  $[N-N]_{N\_MOD}$  si specializzano dal punto di vista semantico.

ANIMALE	ARTEFATTO	OGG_NAT	CIBO	PERSONA	SOSTANZA
pescecane pescecavallo pescesega pescespada pesceluna formicaleone grillotalpa	cartastraccio	erbaspada	melarancia cavolfiore cavolrapa melarancio casciofiore	capostipite	ceralacca gommalacca
8	1	6		1	2

Tabella 46 Composti  $[N-N]_{N\_MOD}$  attestati nel CoDIt\_com: distribuzione rispetto al tipo semantico di output

La Costruzione<sub>COMP</sub>  $[N-N]_{N\_MOD}$  è sfruttata soprattutto per la creazione di nomi che indicano specie di animali (in particolare, pesci) e cibi (in particolare, frutta e verdura):

come già osservato per i composti più frequenti (vd. sopra), generalmente il secondo costituente classifica il tipo animale o pianta/frutta espresso dal costituente testa.

Nel periodo di tempo rappresentato dal CoDIt\_com non risultano quindi attestati composti  $[N-N]_{N\_MOD}$  del tipo di *parola chiave, riunione fiume, città simbolo* (cfr. Baroni *et al.* 2009; Grandi *et al.* 2011; Radimský 2015) che costituiscono istanziazioni di schemi semi-specificati con una testa varabile e un modificatore fisso, come [x-chiave], [x-fiume], [x-simbolo], etc: la diffusione di questi schemi va quindi collocata in un periodo successivo al 1947.

Il sotto-schema  $[N-A]_{N\_MOD}$  istanzia 48 composti, distribuiti nei diversi periodi rappresentati dal CoDIt\_com. I cinque composti più frequenti sono riportati nella tabella 47, in cui si riporta il numero di token per ciascun sotto-corpus.

	1	2	3	4	5	tot	
argentovivo	3	8	122	9	15	157	
terraferma	16	22	59	9	6	112	
palcoscenico	-	-	-	8	92	100	
pangrattato	-	-	-	4	93	97	
camposanto	5	1	19	2	72	99	

Tabella 47 Composti [N-A]<sub>N\_MOD</sub> più frequenti nel CoDIt\_com (token)

Si tratta di composti (alcuni di carattere specialistico, altri diffusi nella lingua comune) che si distribuiscono in maniera piuttosto disomogenea tra i vari sotto-corpora: argentovivo, termine specialistico per indicare il mercurio, si concentra nel terzo sotto-corpus (in particolare in due opere: i Saggi di Magalotti e le Lezioni accademiche di Torricelli); il composto terraferma è piuttosto frequente nei primi tre sotto-corpora (in particolare nell'opera storica Della Repubblica de' Viniziani di Giannotti). Di più recente formazione sono palcoscenico e pangrattato, quest'ultimo concentrato soprattutto ne La scienza in cucina e l'arte di mangiar bene di Pellegrino Artusi. Più omogenea la distribuzione di camposanto, seppur con due addensamenti in corrispondenza del terzo e, soprattutto, del quinto sotto-corpus.

Per quanto riguarda la frequenza dei type e dei token dei composti che istanziano questa Costruzione<sub>COMP</sub>, si può ossevare un generale aumento nel corso dei secoli (figure 44 e 45).<sup>166</sup>

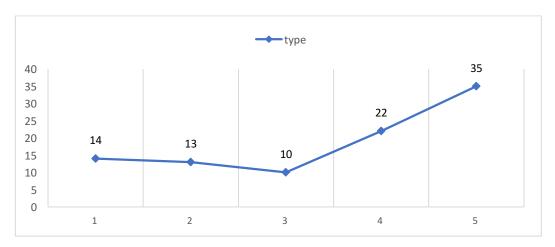


Figura 43 Composti [N-A]<sub>N\_MOD</sub>: distribuzione dei type nei sotto-corpora del CoDIt\_com

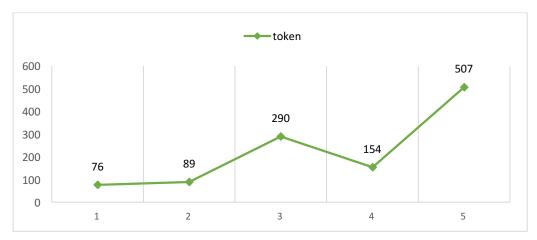


Figura 44 Composti [N-A] $_{N\_MOD}$ : distribuzione dei token nei sotto-corpora del CoDIt\_com

Come illustrato dai grafici riportati nelle figure 44 e 45, la Costruzione $_{\text{COMP}}$  [N-A] $_{\text{N\_MOD}}$  mostra un graduale aumento in termini sia di type sia di token: in particolare, tranne che per un lieve decremento in corrispondenza del terzo sotto-corpus, il numero dei type sale costantemente dalla fase antica a quella contemporanea; il numero dei token oscilla maggiormente ma registra una rapida crescita nella fase più recente. Il picco di crescita

<sup>&</sup>lt;sup>166</sup> In queste tabelle, come in tutte quelle discusse nell'analisi diacronica, i dati relativi alla frequenza dei type/token sono stati normalizzati, per ovviare alle (seppur lievi) differenze nelle dimensioni dei sottocorpora.

che si registra in corrispondenza del terzo sotto-corpus è dovuto al significativo numero di occorrenze del composto *argentovivo* (vd. sopra) che, essendo concentrato in un solo testo, sbilancia l'intera curva.

Dal punto di vista semantico, i composti  $[N-A]_{N\_MOD}$  coprono una vasta gamma di significati, come illustrato nella tabella 48.

ANIMALE	ARTEFATTO	EVENTO	LUOGO	OGG_NAT	CIBO	PERSONA	SOSTANZA
cinciallegra pescepersico picciongrosso	arcoscenico cappacorta cappalunga cappamagna cartapesta cartavelina cassaforte ferravecchio	cantofermo cantopieno girotondo	camposanto caposaldo casamatta palcoscenico pianterreno piazzaforte roccaforte terraferma	codacavallina erbabianca <u>fiengreco</u> legnosanto	melacotta peracotta carnesecca pancotto panforte pangiallo pangrattato pastafrolla pastasciutta	guardiamarina padreterno capobugio	acquabianca acquaforte argentovivo gommarabica gommelastica terracotta vinsanto pecegreca
3	8	4	8	4	9	2	8

Tabella 48 Composti  $[N-A]_{N\_MOD}$  attestati nel CoDIt\_com: distribuzione rispetto al tipo semantico di output

Come emerge dai dati riportati nella tabella 48, questa Costruzione<sub>COMP</sub> istanzia soprattutto nomi che cibi, sostanze e artefatti; più sporadica la formazione di nomi che designano eventi, luoghi, persone o animali. Tale distribuzione ricalca grossomodo quella osservata al § 4.2.1 per la fase antica: già in italiano antico, infatti, questo tipo di composti era sfruttato soprattutto per classificare oggetti naturali e alimenti.

Nel complesso, la Costruzione  $_{\text{COMP}}$  [N-A] $_{\text{N\_MOD}}$ , se osservata in diacronia, appare stabilmente sfruttata dall'italiano per creare soprattutto nomi che classificano oggetti naturali o sostanze e artefatti. Dal punto di vista quantitativo, si è notato che il numero di istanziazioni di tale schema aumenta progressivamente, nonostante in alcuni casi la frequenza delle singole forme non sia equamente distribuita all'interno del corpus ma concentrata in particolari opere.

Infine, i dati estratti dal CoDIt\_com permettono di analizzare due Costruzioni<sub>COMP</sub> con testa a sinistra non individuate nella fase antica, lo schema [A-N]<sub>A|N</sub> (es. *rossosangue*) e

lo schema [A-A] $_{\rm A|N}$  (es. rossobruno): $^{167}$  entrambi gli schemi istanziano esclusivamente nomi che designano colori. $^{168}$ 

In particolare, il primo è istanziato da cinque composti, riportati ed esemplificati nella tabella 49.

Composti [A-N] <sub>A N</sub>	1	2	3	4	5	tot	Esempi
biancoavorio	-	-	-	-	1	1	«[] Egli trascinò e stese su le mattonelle di maiolica il bel tappeto amarantino, variato d'azzurro cupo e di <i>bianco avorio</i> , morbido e intenso come l'antico velluto di Lucca» (D'Annunzio, <i>Forse che sì forse che no</i> )
giallocanarino	-	-	-	-	1	1	«Il rosso pareva molto bello a Domenico; mentre Anna, come le aveva anche detto qualche conoscente, avrebbe voluto scegliere o un celeste o un giallo canarino» (Tozzi, Con gli occhi chiusi)
rossociliegia	-	-	-	1	-	1	«I risultamenti di queste sperienze danno che i raggi calorifici provenienti dal ferro riscaldato al rosso-bruno ed al <i>rosso-ciliegia</i> sono intercettati completamente da uno strato d'acqua []» (Scina, <i>Della fisica</i> particolare)
rossosangue	-	-	-	-	-	2	« L'Ospedale, alto su le mura, rosso sangue, lo vedeva doventare del colore della terra bruciata; il turchino del cielo, bigio» (Tozzi, Con gli occhi chiusi)
verdeterra	-	47	1	-	-	48	«Verde è un color naturale di terra, il quale si chiama <i>verdeterra</i> . Questo colore ha più proprietà: prima, ch'egli è grassissimo colore, e buono a lavorare in visi, in vestiri, in

\_

<sup>&</sup>lt;sup>167</sup> In questa analisi i termini di colore sono considerati aggettivi (piuttosto che nomi) seguendo D'Achille & Grossmann (2013). Sullo status categoriale dei termini di colore in italiano si rimanda a Marello & Onesti (2016).

<sup>&</sup>lt;sup>168</sup> Lo schema [A-A]<sub>A|N</sub> (es. *rossobruno*) istanza forme che si pongono a metà strada tra la sintassi e la morfologia. Seguendo la letteratura dedicata ai termini di colore (cfr. D'Achille & Grossmann 2016, 2017; Grossmann & D'Achille 2016), che tratta queste forme insieme agli altri composti aggettivali che esprimono tonalità di colori, in questa analisi esse vengono considerate istanziazioni di una Costruzione<sub>COMP</sub>, in virtù di alcune proprietà (es. non interrompibilità) che le assimilano a entità di tipo morfologico.

Composti [A-N] <sub>A N</sub>	1	2	3	4	5	tot	Esempi
							casamenti, in fresco, in secco, in muro, in tavola, e dove vuoi» (Cennini, <i>Trattato dell'arte</i> )

Tabella 49 Costruzione<sub>COMP</sub> [A-A]<sub>A|N</sub>: composti attestati nel CoDIt\_co

Come illustrato dalla tabella 49, si tratta generalmente di composti attestati a partire dal sotto-corpus più recente, ad esclusione di *verdeterra*, forma che ricorre quasi esclusivamente all'interno dell'opera di Cennino Cennini, il *Trattato dell'arte*, in cui si descrivono accuratamente molteplici tonalità di colore. Tutti i composti sono costituiti da un nome di colore e un secondo costituente che designa un elemento che prototipicamente ne possiede una determinata tonalità: così, ad esempio, il canarino identifica una particolare tonalità di giallo, definita *giallocanarino*. Tale forma può quindi essere parafrasata come 'giallo come il canarino'. In termini costruzionisti, questo insieme di composti può essere considerato uno schema della famiglia di Costruzionicomp con testa a sinistra, in cui sono identificabili restrizioni semantiche che operano sia sul primo costituente sia sull'interno composto: entrambi presentano infatti il tipo semantico proprietà.

Il sotto-schema  $[A-A]_{A|N}$  (es. rossobruno) è istanziato da otto composti, riportati ed esemplificati nella tabella 50.

Composti [A-A] <sub>A N</sub>	1	2	3	4	5	tot	Esempi
verdebruno	-	3	3	3	3	12	«Il cloroscisto è di clorite mescolata ad altre sostanze, sfoglioso, tenero, verdebruno» (Regnani, Elementi di fisica universale)
verdescuro	-	6	1	2	3	12	«Mentre le due vecchie portavano il pittoresco vestiario del paese: rete nera o <i>verde-scuro</i> in capo, come i briganti nell'Ernani []» (D'Azeglio, <i>Racconti</i> )
rossobruno	-	-	-	1	9	10	«Ho sempre invidiato quel suo colore rossobruno, per i miei capelli. Ah, se tu potessi mai cavare una nota che gli somigli!» (D'Annunzio, Forse che sì forse che no)

Composti [A-A] <sub>A N</sub>	1	2	3	4	5	tot	Esempi
verdechiaro	-	2	-	1	2	5	«Il dolce fiume di Lucchesia divideva col suo nastro <i>verdechiaro</i> i boschi di San Rossore dai boschi di Migliarino, il dominio regio dal dominio ducale» (D'Annunzio, <i>Forse che sì forse che no</i> )
rossochiaro	-	-	-	1	4	5	«[] nella maggior parte dei luoghi il fondo Ë formato da diverse gradazioni di <i>rosso chiaro</i> , o di aranciato o di giallo» (Schiapparelli, <i>La vita sul pianeta Marte</i> )
rossoscuro	-	-	1	-	3	4	«Riscaldato al <i>rosso scuro</i> evapora in tanta abbondanza, da divenir prima aeriforme che fuso []» (Regnani, <i>Elementi di fisica universale</i> )
giallobruno	-	-	-	2	1	3	«Nell'Adriatico se ne rinvengono individui della lunghezza di più di un pollice, la cui tinta è bianco giallognola con istrisce ondeggianti e angolari di colore giallo bruno ovvero sia giallo rossiccio []» (Brocchi, Conchiologia fossile subappennina)
bianco latteo	-	-	-	1	-	1	«In questo gli anfratti sono indivisi ed elegantemente scannellati per lungo, ed ha un colore di miele con una fascia intorno alla sutura superiore di un bianco latteo, ornata di una serie di macchiette rosse» (Brocchi, Conchiologia fossile subappennina)

Tabella 50:  $Costruzione_{COMP}$  [A-A]<sub>A|N</sub>: composti attestati nel CoDIt\_com

Le istanziazioni della  $Costruzione_{COMP}$  [A-A]<sub>A|N</sub> designano colori, identificati attraverso aggettivi qualificativi che ne indicano la tonalità, ad eccezione di *biancolatteo*, costituito da un aggettivo di relazione e interpretabile come 'bianco come il latte'.

Diversamente dai composti aggettivali [A-N]<sub>A|N</sub> (es. *rossosangue*), questo tipo di forme è di più antica attestazione, nonostante la maggior parte delle occorrenze si collochi nel periodo più recente rappresentato dal corpus. Va inoltre notato che gran parte delle occorrenze è attestata in opere di carattere tecnico-specialistico e designa tonalità di colore di specie animali o sostanze.

I dati estratti dal CoDIt\_com relativi alle Costruzioni<sub>COMP</sub> che istanziano nomi di colori permetto di confermare quando osservato da D'Achille & Grossmann (2013: 535), secondo cui, nelle precedenti fasi evolutive, l'italiano «adottava prevalentemente strategie di carattere sintattico, accostando ad aggettivi di colore di base altri aggettivi indicanti la luminosità». La Costruzione<sub>COMP</sub> costituita da un aggettivo di colore di base e un nome indicante un elemento che prototipicamente possiede una determinata sfumatura (es. *giallocanarino*) sembra essere una Costruzione<sub>COMP</sub> più recente, destinata ad avere fortuna in italiano contemporaneo.

Nel complesso, l'analisi dei dati estratti dal CoDIt\_com ha permesso di mettere in luce alcuni aspetti relativi all'evoluzione diacronica della Costruzione dei composti con testa a sinistra. In primis, si è notato che, rispetto al quadro delineato per la fase antica, nei secoli successivi tale famiglia di Costruzionicomp si arricchisce di due nuovi sottoschemi (ossia [A-N]<sub>AIN MOD</sub> e [A-A]<sub>AIN MOD</sub>), semanticamente specializzati nell'espressione di termini di colore. I sotto-schemi attestati in italiano antico ([N-N]<sub>N\_SUB</sub>, [N-N]<sub>N\_MOD</sub> e [N-A]<sub>N MOD</sub>) continuano, nel corso dei secoli rappresentati dai sotto-corpora del CoDIt\_com, a essere istanziati da nuove forme senza soluzione di continuità: in tutti e tre i casi, e in particolare in quello dei composti [N-N]<sub>N\_MOD</sub>, si è osservato un aumento di frequenza type/token nella fase più recente (1841-1947). Rispetto allo schema ([N-N<sub>N SUB</sub>, va notato che la sua produttività è trainata soprattutto dalla vitalità del sottoschema semi-specificato [capo<sub>PERS</sub>-x]<sub>PERS</sub>, costituito da capo- nel significato di 'autorità' e un sostantivo indicante generalmente un gruppo di persone o un luogo su cui tale autorità è esercitata. Tale schema, seppur attestato sporadicamente nella fase antica, si fa più produttivo a partire dal periodo individuato dal terzo sotto-corpus (1533-1691). Ad eccezione di questo caso, la Costruzione<sub>COMP</sub> [N-N]<sub>N\_SUB</sub> appare piuttosto conservativa rispetto alla fase antica: una situazione simile è quella della Costruzione<sub>COMP</sub> [N-A]<sub>N MOD</sub>, stabilmente produttiva nel corso dei secoli, ma orientata prevalentemente (come nel periodo delle Origini) alla creazione di forme che designano specie di piante o sostanze.

## 4.3.2 La Costruzione<sub>COMP</sub> dei composti con testa a destra

Come emerso dall'analisi proposta al § 4.2.2, in italiano antico, la Costruzione<sub>COMP</sub> dei composti con testa a destra si caratterizza per una notevole complessità rispetto a quella dei composti con testa a sinistra. In particolare, dall'analisi dei composti attestati nel dataset *Compos\_ita\_TLIO* è emerso che, nella fase antica, essa si articola nei seguenti sotto-schemi:

- 1.  $[N-N]_{N\_SUB}$  (es. terremoto, lanovendolo)
- 2. [N-PART.PRES]<sub>N|A\_SUB</sub> (es. *panchiedente, armipotente*)
- 3. [N-PART.PASS]<sub>N|A\_SUB</sub> (es. *manrovescio*, *occhibagliato*)
- 4. [N-V]<sub>V\_SUB</sub> (es. *manomettere*, *caporicciare*)
- 5. [A-N]<sub>N|A|AVV\_MOD</sub> (es. *malafede*, *tuttavolta*)
- 6. [A-PART.PASS]<sub>NIA MOD</sub> (es. *biancovestito*, *primonato*)
- 7. [AVV-A]<sub>A|N\_MOD</sub> (es. *malsano*, *benavventuroso*)
- 8.  $[AVV-V]_{V\_MOD}$  (es. *maltrattare*, *chiarefare*)
- 9. [AVV-PART.PRES]<sub>N|A\_MOD</sub> (es. *malconveniente*)
- 10. [AVV-PART.PASS]<sub>N|A\_MOD</sub> (es. *malvestito*)
- 11. [AVV-N]<sub>N\_MOD</sub> (es. benessere)
- 12. [AVV-AVV]<sub>AVV MOD</sub> (es. *malvolentieri*)
- 13. [N-N]<sub>N MOD</sub> (es. *capoluogo*, *capolavoro*)

Tutti i sotto-schemi individuati nella fase antica sono attestati nel CoDIt\_com in almeno una istanziazione (di antica o più recente formazione), ad eccezione dello schema [N-PART.PASS] $_{N|A\_SUB}$  (es. *occhibagliato*, *plebeuscito*) che sembra aver esaurito la propria produttività. I dati estratti dal CoDIt\_com permettono inoltre di identificare un ulteriore sotto-schema non attestato in italiano antico, ossia [N-N] $_{N\_MOD}$  (es. *capoluogo*, *capolavoro*).

I type e i token dei composti attestati nel CoDIt\_com si distribuiscono come illustrato nella tabella 51.

Costruzioni <sub>COMP</sub> con testa a destra	type	token
[N-N] <sub>N_SUB</sub>	19	390
[N-PART.PRES] <sub>N A_SUB</sub>	10	226
[N-PART.PASS] <sub>N A_SUB</sub>	-	-
$[N-V]_{V\_SUB}$	2	83
[A-N] <sub>N A AVV_MOD</sub>	714	22.002
[A-PART.PASS] <sub>N A_MOD</sub>	5	561
[AVV-A] <sub>A N_MOD</sub>	88	2.000
$[AVV-V]_{V\_MOD}$	220	2.212
$[AVV-PART.PRES]_{N A\_MOD}$	32	178
$[AVV-PART.PASS]_{N A\_MOD}$	388	4.159
[AVV-N] <sub>N_MOD</sub>	3	443
[AVV-AVV] <sub>AVV_MOD</sub>	11	658
$[N-N]_{N\_MOD}$	8	376
totale	1500	33.288

Tabella 51 Costruzione<sub>COMP</sub> dei composti con testa a destra: distribuzione dei type e dei token attestati nel CoDIt\_com rispetto ai sotto-schemi

Rispetto alla fase antica, gran parte delle  $Costruzioni_{COMP}$  con testa a destra risulta attestata nel  $CoDIt\_com$ , nonostante alcune siano istanziate da un numero limitato di forme, prevalentemente di antica formazione. È il caso delle  $Costruzioni_{COMP}$  [N-PART.PRES]<sub>N|A\\_SUB</sub>, [A-PART.PASS]<sub>N|A\\_MOD</sub> e [N-V]<sub>V\\_SUB</sub>. Nel primo caso, lo schema è istanziato dalle forme riportate ed esemplificate nella tabella 52.

Composti [N-PART.PRES] <sub>N A_SUB</sub>	1	2	3	4	5	tot	Esempi
armilucente	-	-	-	-	1	1	«Ospite all'onde sacre, e pieno gli occhi / Del greco sole, armilucente Atena, / Già non vedrò, come bramai gran tempo / Nel sogno mio, le tue beate rive» (Prati, <i>Poesie scelte</i> )
<u>armipotente</u>	1	-	-	3	2	6	«A udir quell'omaccino armipotente / Isbigottisce la povera gente, / E dice: Veramente / Cotestui studia per le invenzioni / Di verseggiar le bombarde e i cannoni» (Carducci, Juvenilia)
bombisonante	-	1	-	-	-	1	«Alcuni sono nel mezo vacui, benché nel principio alquanto compareno; chi è affectato et bombisonante; chi exile, arido et ieiuno, dale rime, como da

Composti [N-PART.PRES] <sub>N A_SUB</sub>	1	2	3	4	5	tot	Esempi
							concitato torrente, se lassa portare []» (Equicola, <i>Libro de natura de amore</i> )
bronzosonante	-	-	-	-	1	1	«Il giovinetto udÏ strepere trombe, / gemere conche, ed ascoltò soavi, / tra l'immensa manìa bronzosonante, / squillare i doppi flauti di loto» (Pascoli, Poemi conviviali)
<u>luogotenente</u>	17	100	56	6	28	207	«[] e da altra parte, con cautela italiana, consentì che don Alfonso suo primogenito si conducesse col duca di Milano con cento cinquanta uomini d'arme, con titolo di luogotenente delle sue genti» (Guicciardini, Storia d'Italia)
manoprendente	1	-	-	-	-	1	«Uomo lietissimo, piacentiere, reca a' nimici, grave e soperchiante agli amici, a niuno fedele, consigliere malvagio e leggiero, signore mercenaio, lingua vana, manoprendente, la gola grande, magagnato di cose turpissime, che non son degne di nominare» (Brunetto Latini, Fiore di filosofi)
nullatenente	-	-	-	-	4	4	«Oggi gli oziosi e i vagabondi sono di due specie: ricchi e nullatenenti. I primi vivono di rendita e vanno a Montecarlo; i secondi vivono di audacia e vanno in prigione []» (Guerrini, Brani di vita)
ondisonante	-	-	-	2	-	2	«[] cioè la metafora non fosse più semplicemente delle armi ruggenti, ma cambiate in fiere o assomigliate alle fiere e così ruggenti, per una enallage pindarica) fa forza alla lingua nelle voci (come le composte alla greca: <i>ondisonante</i> ec. che la nostra lingua non ama) nelle forme trasportate dal greco e

Composti [N-PART.PRES] <sub>N A_SUB</sub>	1	2	3	4	5	tot	Esempi
							latino infelicemente» (Leopardi, <i>Zibaldone</i> )
ondivagante	-	-	-	2	-	2	«Un consiglio tessean, perch'io cadessi / Novamente ne' guai. Come lontano / Da terra fu l' <i>ondivagante</i> legno, / Il negro m'apparì giorno servile» (Pindemonte, <i>Odissea</i> )
pennipotente	-	-	-	-	1	1	«[] ccòmodi al tibìcine, ogni specie / pennipotente altivolante cadde / per la forza degli archi miei cidonii / e de' miei dardi gnossi» (D'Annunzio, Laudi)

Tabella 52 Composti [N-PART.PRES] $_{N|A\_SUB}$  attestati nel CoDIt\_com: distribuzione dei type e dei token rispetto ai sotto-corpora

Come illustrato dagli esempi riportati nella tabella 52, si tratta di composti che occorrono prevalentemente all'interno di opere poetiche, creati sul modello di forme latine come armipotens. In particolare, armilucente, armipotente e luogotenente sono composti di origine latina, già attestati nella fase antica. Dai dati estratti dal CoDIt\_com, la Costruzione<sub>COMP</sub> [N-PART.PRES]<sub>N|A\_SUB</sub>, particolarmente produttiva in latino classico (cfr. § 4.1) e ben attestata in italiano antico, sembrerebbe quindi sfruttata solo per estemporanee creazioni poetiche sul modello latino, ad eccezione di alcuni rari casi (come nullatenente) che occorrono anche in testi non poetici. Tuttavia, prendendo in considerazione dati relativi a un periodo successivo rispetto a quello rappresentato dal CoDIt\_com (quindi dopo il 1947), si può osservare la diffusione, nella lingua comune, di un certo numero di questo tipo di composti. In particolare, negli ultimi decenni, si è diffuso un numero considerevole di forme che presentano il participio dipendente come secondo costituente e che possono essere considerate un sotto-schema semi-specificato [x-dipendente] della Costruzione<sub>COMP</sub> [N-PART.PRES]<sub>NIA SUB</sub>. Nel De Mauro (2004) sono registrate le seguenti forme: 169 alcoldipendente (1963), tossicodipendente (1978), teledipendente (1981), farmacodipendente (1983), videodipendente (1981). Tali forme sono costituite dal participio presente dipendente in funzione di testa e un nome o un

<sup>169</sup> Tra parentesi si riporta la data di prima attestazione registrata nel De Mauro (2004).

elemento formativo (è il caso di *video-, tossico-* e *tele-*) come modificatore. Alla formazione e alla fortuna di questo sotto-schema avrà probabilmente contribuito la diffusione del termine *tossicodipendente.*<sup>170</sup> L'attuale produttività di questo sotto-schema è confermata dalla presenza di composti con *-dipendente* nel repertorio di neologismi Treccani, che registra le seguenti forme: *aerodipendente, blog-dipendente, cellularedipendente, cibodipendente, email-dipendente, fagianodipendente, franco-dipendente, internet-dipendente, messaggiodipendente, pornodipendente, psicanalista dipendente, rischiodipendente, silviodipendente, spesadipendente, storiadipendente, sudokudipendente, telefoninodipendente. Semanticamente, tale sotto-schema istanzia nomi o aggettivi che designano referenti umani: tale restrizione semantica può essere espressa in termini costruzionisti mediante il tipo semantico PERSONA riportato in pedice ([x-dipendente]<sub>N|A\_PERSONA</sub>.* 

Nel complesso, le istanziazioni della Costruzione<sub>COMP</sub> [N-PART.PRES]<sub>N|A\_SUB</sub> rimangono residuali nel corso dei secoli successivi alla fase antica, e strettamente limitati alla lingua poetica; solo in tempi più recenti, la fortuna di alcuni composti che presentano – *dipendente* come secondo costituente ha determinato la formazione di uno schema semi-specificato con il primo slot variabile e il secondo fisso. È interessante notare che, diversamente da quanto osservato finora (ad eccezione del sotto-schema [x-vendolo], cfr. § 4.2.2), si tratta di un sotto-schema semi-specificato in cui lo slot fisso è costituito dal secondo costituente, l'elemento testa, e non dal modificatore in prima posizione (come nel caso dei composti con mal(e)-, ben(e)-, capo-x, etc).

La Costruzione<sub>COMP</sub> [A-PART.PASS]<sub>N|A\_MOD</sub> è istanziata dai composti *biancovestito* (3), *primogenito* (253), *secondogenito* (47), *terzogenito* (9) e *salvacondotto* (249), già attestati in italiano antico: dal momento che non si registrano composti di più recente formazione, questo schema sembra aver esaurito la sua produttività.

Simile è la sorte della Costruzione $_{\text{COMP}}$  [N-V] $_{\text{V\_SUB}}$ , istanziata dai composti *caponascondere* e *manomettere*, esemplificati nella tabella 53.

<sup>&</sup>lt;sup>170</sup> La forma risulta essere decisamente la più frequente di quelle registrate dal De Mauro (2004) nel corpus di italiano contemporaneo itTenTen (Jakubíček et al. 2013), di cui si riportano le frequenze: alcoldipendente (80), farmacodipendente (88), teledipendente (900), tossicodipendente (21.227), videodipendente (71). Oltre al modello di tossicodipendente, va considerata la diffusione di composti dell'inglese che presentano il participio –addicted come secondo elemento, spesso tradotti con un composto italiano x-dipendente.

Composti [N-V] <sub>V_SUB</sub>	1	2	3	4	5	tot	Esempi
caponascondere	-	-	-	-	1	1	(a) «[] e tutto andava a finire nello stampare foglietti di sotterfugio, nello scrivere col carbone una minaccia sul muro, nel fare a caponascondi colle spie, e ciò era detto missione» (Giusti, Cronaca dei fatti di Toscana)
<u>manomettere</u>	1	13	5	25	38	82	(b) «vedendole ora vittoriose star contente a magri accordi, ora vinte esser calpestate e manomesse con incredibile crudeltà, potrà darsi ragione di fatti che altrimenti riescirebbero strani ed inesplicabili» (D'Azeglio, Racconti)

Tabella 53 Costruzione<sub>COMP</sub> [N-V]<sub>V\_SUB</sub>: composti attestati nel CoDit\_com

Le due forme riportate nella tabella 53 costituiscono le uniche istanziazioni attestate nel CoDIt\_com della Costruzione $_{COMP}$  [N-V] $_{V\_SUB}$ . Nel primo caso si tratta di un composto attestato nel sotto-corpus più recente, indicante un gioco probabilmente simile al nascondino; nel secondo, *manomettere* è una forma di origine latina, già ben attestata nella fase antica nel senso di 'offendere, guastare': come illustrato nell'esempio riportato in (b), in cui *manomettere* vale 'danneggiare, violare', questo significato si conserva anche nei secoli successivi.

Più produttiva è la Costruzione<sub>COMP</sub> [N-N]<sub>N\_SUB</sub>, che istanzia 19 composti costituiti da due nomi legati da una relazione di subordinazione, riportati ed esemplificati nella tabella 54.

Composti [N-N] <sub>N_SUB</sub>	1	2	3	4	5	tot	Esempi
ferrovia	-	-	-	2	148	150	(a) «Ora la campagna romana comincia ad aprirsi alle <i>ferrovie</i> . Al tempo della mia gioventù non c'erano di questi lussi []» (D'Azeglio, <i>Racconti</i> )
<u>piedistallo</u>	1	2	7	9	44	63	(b) «I suoi zoccoli o piedistalli che gli vogliam chiamare, dove posano le colonne, sono quadri di proporzione, con l'avere da piè la sua fascia soda e così un'altra di sopra, che lo ricinga in cambio di cornice» (Vasari, Vite)
<u>nottetempo</u>	11	1	4	6	15	37	(c) «Si trova abbondante nell'Adriatico ove viene pescato di <i>nottetempo</i> con grandi reti a sacco, fortemente piombate alla bocca, le quali raschiando il fondo del mare sollevano il pesce insieme colla sabbia e col fango in cui giace» (Artusi, La scienza in cucina e l'arte di mangiar bene)
arcobaleno	-	3	6	2	16	27	(d) «L'aria era torbida come quando apparisce l'arcobaleno, e più, sicchè senz'altra pioggia bagnava; il Sole si scorgeva chiaramente []» (Sassetti, Lettere)
capogiro/capogirlo	1	1	7	2	15	26	(e) «[] ed ora non può peranco muover passo ove non sia sorretto da qualcuno, e se vuole tenersi ritto non istà sicuro senza un saldo sostegno che lo salvi dalle conseguenze di un capogiro» (Belli, Lettere)
							(f) «E che altro è il volere e il non volere, che col transito del sì e del no gli mette il cervello in compressa, che il

Composti [N-N] <sub>N_SUB</sub>	1	2	3	4	5	tot	Esempi
							litigio che tuttavia hanno insieme i <i>capogirli</i> del no e del sì?» (Aretino, <i>Lettere sull'arte</i> )
<u>caprifoglio</u>	1	-	-	-	14	15	«Mentre questi scompaiono nelle bocche, saranno sparsi nella sala, con spruzzatori, dei soavissimi profumi di rosa, gelsomino, caprifoglio e gaggia, la cui dolcezza nostalgica e decadente sarà brutalmente rifiutata dai combattenti che si porranno con fulmineità la maschera contro i gas asfissianti» (Marinetti, Cucina futurista)
<u>caprifico</u>	-	2	-	2	4	12	« Allora si lasciò vincere dalla tentazione di riposare due minuti su quelle quattro pietre lisce messe in fila all'ombra del <i>caprifico</i> che c'è accanto alla cappelletta []» (Verga, <i>Malavoglia</i> )
<u>manrovescio</u>	1	3	1	2	5	12	«Io vidi intorno a questi un cerchio fare, / e seguir cose che non sono umane, / chè si sentì una spada fischiare d'un certo manrovescio tondo e giusto / ch'a venti il capo levò dallo imbusto» (Pulci, Morgante)
erbivendolo	-	-	-	-	8	8	«Eran vari giorni che l'erbivendola, ogni volta che passava accanto a Derossi, lo guardava, lo guardava con una espressione di grande affetto» (De Amicis, Cuore)
resoconto	-	-	-	-	8	8	«Ciò radica sempre più nel concetto delle masse che la magistratura non sia che il rifugio dei legulei che non trovano da difendere un ladro di galline, e questa dolorosa convinzione si legge troppo chiaramente

Composti [N-N] <sub>N_SUB</sub>	1	2	3	4	5	tot	Esempi
							tra le righe dei resoconti parlamentari, come si sente a troppo chiare note nei discorsi d'ogni giorno, fino tra i contadini» (Guerrini, Brani di vita)
<u>pescivendolo</u>	1	-	1	2	4	8	«Quando Agnese si trovò al punto dove la strada che conduceva al suo tugurio si divideva da quella che dovevan fare i pescivendoli per giungere a casa loro, cioè quando ebbe passato il ponte dell'Adda []» (Manzoni, Fermo e Lucia)
animavversione	1	-	-	2	3	6	«[] sarebbe stato lo stesso che rinunziare al privilegio di asilo, screditare il convento presso tutto il popolo, attirarsi l'animavversione di tutti i capuccini dell'universo per aver lasciato ledere il diritto di tutti, tirarsi contra tutte le autorità ecclesiastiche, le quali allora si consideravano come tutrici di questo diritto» (Manzoni, Fermo e Lucia)
boccaporto	-	-	-	1	3	4	«Tumultuando dalle boccaporte / Il capitano e i marinari tutti / Irruppero sul ponte; e molti a prua / Corsero a moí di gregge agglomerandosi []» (Graf, Poesie)
borgomastro	-	-	-	1	2	3	«Pochi anni fa il borgomastro di Gratz fu vittima di un monomaniaco religioso che alcuni anni prima aveva minacciato un'altra esistenza» (Lombroso, L'uomo delinquente)

Composti [N-N] <sub>N_SUB</sub>	1	2	3	4	5	tot	Esempi
fruttivendolo	-	-	-	-	3	3	«Le fruttivendole e i maniscalchi lungo la via guardavano, senza capire, l'inseguimento di quei due uomini affannati e gocciolanti di sudore sotto il solleone» (D'Annunzio, Novelle della Pescara)
robivendolo	-	-	-	-	2	2	«Infatti un ebreo robivendolo, tutte le volte che non aveva da pagare il conto alla trattoria, le portava a far vedere ogni specie di oggetti vecchi e glieli lasciava sul banco []» (Tozzi, Con gli occhi chiusi)
lattivendolo	-	-	-	-	1	1	«Le carriuole dei <i>lattivendoli</i> stridevano sulle loro ruote mal ferme, i vetturini facevano scoppiettare le loro fruste []» (Tarchetti, <i>Fosca</i> )
straccivendolo	-	-	-	1	-	1	«Dico il vero, Ciro mio che volentieri ti avrei udito, perché m'immagino che lo scartarello, come tu dici, non fosse poi roba da gittarsi allo straccivendolo» (Belli, Lettere)

Tabella 54 Composti  $[N-N]_{N\_SUB}$  attestati nel CoDIt\_com: distribuzione delle occorrenze rispetto ai sottocorpora

Il composto [N-N]<sub>N\_SUB</sub> più frequente all'interno del CoDIt\_com è *ferrovia*, calco dal tedesco *Eisenbahn*, attestato a partire dal quarto sotto-corpus. Un significativo numero di forme è di origine latina e risulta già attestato nella fase antica: si tratta di *animavversione, capogiro/capogirlo* (o *capogirolo*), *caprifoglio*, *caprifico*, *manrovescio*, *nottetempo*, *pescivendolo*, *piedistallo*. Notevole è il numero di composti costituiti da *vendolo* come secondo costituente, i quali (come già osservato al § 4.2.2) rappresentano istanziazioni di un sotto-schema semi-specificato [xartefatto|Animale-vendolo]<sub>Persona</sub> in cui il primo elemento è variabile, ma soggetto a restrizioni semantiche (il tipo semantico

può infatti essere artefatto o animale, ossia prodotti che possono essere venduti), e il secondo è fisso, ossia *-vendolo*. I composti di questo tipo presentano come tipo semantico sempre pers, in quanto indicano referenti umani. Tale schema risulta essere ancora produttivo nei due sotto-corpora più recenti, nei quali occorrono cinque istanziazioni (*erbivendolo*, *fruttivendolo*, *lattivendolo*, *robivendolo*, *straccivendolo*) designanti individui che vendono una determinata merce.

Il tipo semantico di output più rappresentato tra i composti  $[N-N]_{N\_SUB}$  è PERSONA, come illustrato dalla tabella 55.

ARTEFATTO	EVENTO	LUOGO	OGG_NAT	PERSONA	SENTIMENTO	ТЕМРО
<u>piedistallo</u>	<u>capogiro</u> resoconto <u>manrovescio</u>	boccaporto ferrovia <sup>171</sup>	arcobaleno <u>caprifoglio</u> <u>caprifico</u>	borgomastro erbivendolo fruttivendolo lattivendolo pescivendolo robivendolo straccivendolo	<u>animavverisone</u>	<u>nottetempo</u>
1	4	2	3	7	1	1

Tabella 55 Composti [N-N]<sub>N\_SUB</sub> attestati nel CoDIt\_com: distribuzione rispetto al tipo semantico di output

Come si osserva dalla distribuzione dei composti riportata nella tabella 55, la  $Costruzione_{COMP}$  [N-N]<sub>N\_SUB</sub> (e in particolare il sotto-schema [ $x_{ARTEFATTO\_ANIMALE}$ -vendolo]<sub>PERSONA</sub>) istanzia prevalentemente nomi che designano referenti umani; occasionale è la creazione di nomi che designano eventi, luoghi, oggetti naturali.

Nel complesso, le istanziazioni della Costruzione $_{COMP}$  [N-N] $_{N\_SUB}$  attestate nel CoDIt\_com si concentrano all'interno del sotto-schema[ $x_{ARTEFATTO|ANIMALE}$ -vendolo] $_{PERS}$ : analogamente a quanto osservato per la Costruzione $_{COMP}$  dei composti con testa a sinistra, la produttività della Costruzione è trainata da quella di un sotto-schema: le altre istanziazioni sono generalmente creazioni occasionali o composti di origine latina o antica.

<sup>&</sup>lt;sup>171</sup> Il composto *ferrovia* può in realtà essere descritto sia dal tag LUOGO sia dal tag ARTEFATTO: esso rappresenta, dal punto di vista semantico, un *tipo complesso* (cfr. Jezek 2016: 59), ovvero presenta una polisemia inerente (cfr. *inherent polysemy*, Pustejovsky 2008) in cui i diversi significati di una parola si attivano simultaneamente.

La Costruzione<sub>COMP</sub> [A-N]<sub>N|A|AVV\_MOD</sub> è istanziata da 714 composti costituiti da un aggettivo e un nome; come già osservato per la fase antica al § 4.2.2, gli aggettivi che occorrono in prima posizione nei composti sono un numero limitato: in particolare, gli aggettivi che occorrono in un maggior numero di forme sono *buono* (2%), *malo* (90%), *mezzo* (2%). La Costruzione<sub>COMP</sub> [A-N]<sub>N|A|AVV\_MOD</sub> è quindi istanziata prevalentemente da composti costituiti da mal(o/a) come primo costituente fisso e un nome variabile (es. *malasorte, maltempo, malpasso*, etc.): come già osservato al § 4.2.2, questo tipo di forme rappresenta un sotto-schema semi-specificato di notevole produttività, a cui si è scelto di riservare particolare attenzione e un'analisi specifica nel quadro della teoria sulla formazione di affissoidi a partire da costituenti di composti (cfr. § 5.4 e sgg). Di seguito si analizzano quindi i composti che istanziano la Costruzione<sub>COMP</sub> [A-N]<sub>N|A|AVV\_MOD</sub> e che non fanno parte di tale sotto-schema.

Oltre che dai composti con mal(a/o), la Costuzione<sub>COMP</sub> [A-N]<sub>N|A|AVV\_MOD</sub> è istanziata da altri 63 composti: in particolare, 54 nomi/aggettivi e 9 avverbi. Come già riscontrato nella fase antica (cfr. § 4.2.2), questa Costruzione<sub>COMP</sub> istanzia infatti due sotto-schemi distinti in base alla categoria di output nominale/aggettivale ([A-N]<sub>N|A</sub>, es. *sempreverde*) o avverbiale ([A-N]<sub>AVV\_MOD</sub>, es. *tuttafiata*).

Tra le istanziazioni del sotto-schema  $[A-N]_{N|A}$ , è possibile isolare due gruppi di forme che presentano gli aggettivi buon(o)- e mezz(o)-, riportati nella tabella 56.

Composti con buon(a/o)-	Composti con mezz(a/o)-
buonamano (7)	mezzacanna (6)
<u>buonanotte</u> (115)	mezzacoda (1)
buoncuore (112)	<u>mezzaluna</u> (30)
<u>buongiorno</u> (79)	mezzamattina (10)
buongoverno (65)	<u>mezzanotte</u> (439)
buongustaio (5)	mezzapicca (1)
<u>buonora</u> (267)	mezzobusto (3)
buonsenso (153)	<u>mezzocannone</u> (8)
<u>buonuomo</u> (403)	mezzocuore (1)
<u>buonafede</u> (191)	<u>mezzodì (</u> 66)
<u>buonumore</u> (83)	<u>mezzogiorno (</u> 546)
	mezzosoprano (2)

Tabella 56 Composti con buon(a/o)- e mezz(a/o)- attestati nel CoDIt\_com

Come si è già osservato al § 4.2.5.1, i composti con buon(a/o)- e mezz(o/a)- non hanno riscontro in latino ma possono essere considerati una creazione autonoma dell'italiano,

attestata fin dalla fase antica. Nel caso dei composti con buon(a/o)-, si tratta prevalentemente di forme già attestate in italiano antico; dal punto di vista semantico, buon(o/a)- presenta il valore migliorativo proprio dell'aggettivo come forma libera. Nei composti con mezz(o/a)-, il primo elemento veicola il significato di 'metà', riferito a un periodo di tempo (mezzogiorno, mezzodì, mezzamattina, mezzanotte), a un artefatto (mezzobusto, mezzacanna, mezzacoda) o a un referente umano (mezzosoprano); nel caso di mezzocannone 'salterio di dimensioni più piccole del cannone', mezz(o/a)- allude a dimensioni minori rispetto a uno standard.

Gli altri composti  $[A-N]_{N|A}$  si distribuiscono come illustrato nella tabella 57 in base all'aggettivo che occorre in prima posizione.

Aggettivi in prima posizione	Composti [A-N] <sub>N A</sub>
albo	albaspina (4)
alto	altogatto (1), altopiano (30), altosuono (2)
basso	bassofondo (17), bassorilievo (114)
bello	belladonna (4), belvedere (6)
bianco	<u>biancospino</u> (22)
falso	falsacappa (2), falsamoneta (5), falsariga
,	(10)
franco	francobollo (26)
galante	galantuomo (555)
gentile	gentiluomo (1804), gentildonna (297)
grande	grancassa (13), <u>granduca</u> (284),
	granduchessa (46), granguardia (1),
	grantempo (821), granturco (54)
madre	madreperla (35), <u>madreselva</u> (4)
nobile	<u>nobiluomo</u> (7)
onesto	onestuomo (15)
piatto	piattaforma (27)
povero	poveruomo (190)
sommo	sommoscapo (3)
valente	<u>valentuomo</u> (94)
vano	<u>vanagloria</u> (258)

Tabella 57 Composti  $[A-N]_{N|A}$  attestati nel CoDIT\_com (eccetto composti con *buono-, malo-, mezzo-*): distribuzione rispetto all'aggettivo in prima posizione

I composti riportati nella tabella 57 sono costituiti da aggettivi prevalentemente gradabili; gran parte dei composti di più antica attestazione è costituita dai nomi *uomo/donna* come secondo costituente: si tratta di un insieme di forme già presente in italiano antico, creato probabilmente sul modello di forme francesi (cfr. *gentilhomme*,

*gentildame*, § 4.2.2). Inoltre, va notato che i composti con *alto-, basso-* e *falso-* costituiscono un'innovazione dei secoli successivi alla fase antica.<sup>172</sup>

Il sotto-schema [A-N]<sub>AVV\_MOD</sub> è istanziato da otto composti, costituiti da un aggettivo indefinito (*ogni*, *tale*, *tutto*) e un sostantivo che individua un periodo di tempo (*fiata*, *giorno*, *ora*, *tempo*); l'unica eccezione è l'avverbio di modo (o conclusivo) *ognimodo*. Non tutti gli avverbi di tempo identificati nella fase antica risultano attestati nei periodi successivi, come illustrato dalla tabella 58, dove sono riportati ed esemplificati i composti [A-N]<sub>AVV\_MOD</sub> che occorrono nei sotto-corpora del CoDIt\_com.

Composti [A-N] <sub>AVV_MOD</sub>	1	2	3	4	5	tot	Esempi
ognimodo	-	47	57	107	183	394	(a) «e pur quando questo non se potesse, vorrei che V. S. facesse trovare un bon cavallante con un bon cavallo, et accordarsi seco in tanto al mese, e mandarmelo ognimodo, perch'io ne ho estremo bisogno, e piacesse a Dio ch'io havesse Martino» (Castiglione, Lettere)  (b) « A ogni modo, quelle storie parlavano alla credulità, non al bon gusto, che non era ancora nato» (Manzoni, Del romanzo storico)
ognora	118	84	156	86	38	482	(c) «Nervosa per temperamento, pallida ognora nel volto, di complessione gracile, dotata di soverchia e però funesta sensibilità, io non prometteva di giungere alle proporzioni di un organismo fortemente costituito» (Caracciolo, Misteri del chiostro napoletano)
talora	155	141	502	222	545	1568	(d) «S'alzava ella <i>talora</i> con impeto a mezzo il

-

<sup>&</sup>lt;sup>172</sup> Composti con *alto* in funzione di avverbio sono invece già attestati sia in latino sia in italiano antico.

Composti [A-N] <sub>AVV_MOD</sub>	1	2	3	4	5	tot	Esempi
							discorso, come se temesse in quel momento di esser tenuta, e passeggiava pel parlatorio; talvolta dava in risa smoderate, talvolta levando gli occhi, senza che se ne intendesse una cagione, prorompeva in sospiri; talvolta dopo una lunga e manifesta distrazione, si risentiva, ed approvava con negligenza ragionamenti che la sua mente non aveva avvertiti» (Manzoni, Promessi sposi)
talvolta	170	169	358	926	578	2201	vd. sopra
tuttafiata	10	6	-	1	-	17	(e) «Et se tutta fiata ve ne capitasse alcuno ne le mane fatine subito fare executione senza expectarni altra nostra commissione» (Boiardo, Lettere)
tuttavia	326	448	1102	818	764	3458	(f) «Tutto questo, sebbene apparisca forse principalmente nella letteratura, tuttavia si può applicare a molti altri rami del sapere, o di altri pregi umani» (Leopardi, Zibaldone)
tuttavolta	18	158	111	97	42	426	(g) «Dico, dunque, ch'io non biasimo l'aggiunta che vi fate de l'avverbio per forza; tuttavolta non mi par necessaria: perciochè il tirare è uno de' quattro moti violenti; che sono, il portare, il sospingere, il girare a torno, e'l tirare; ciascun de' quali si fa per forza» (Tasso, Lettere)

Composti [A-N] <sub>AVV_MOD</sub>	1	2	3	4	5	tot	Esempi
tuttogiorno	22	22	15	105	8	172	(h) «[] e l'idea di giusto e d'ingiusto che pure è deposta come un germe nel cuore di tutti gli uomini, svolta nel suo, fin dal principio insieme con le passioni del terrore e della cupidigia servile, accomodata per abito ai principi che tuttogiorno sentiva predicare, ed alle azioni che vedeva compiersi e alle quali ella partecipava, era divenuta una applicazione mostruosa di tutte queste idee e di tutte quelle passioni» (Manzoni, Fermo e Lucia)
tuttora	61	23	3	121	122	330	(i) «M'è facile di ricordarmi dei propositi che feci allora, prima di tutto perché ne feci d'identici in epoca più recente, eppoi perché li annotai su un foglio di carta che conservo tuttora» (Svevo, Coscienza di Zeno)
tuttotempo	12	1	1	-	-	14	(l) «La qual marito non vuole già et servasi vergine tutto tempo, ma bene d'essere amata et vagheggiata si contenta» (Bembo, Asolani)

Tabella 58 Composti [A-N]<sub>AVV\_MOD</sub> attestati nel CoDIt\_com

Dei composti attestati in italiano antico, il CoDIt\_com non registra le forme *nessunora*, *ognanno*, *ognidì*, *tuttogiorno*, che sembrano quindi uscire dall'uso molto presto; una sorte simile spetta a *tuttotempo*, che occorre fino al periodo individuato dal terzo sottocorpus. Non è invece attestata nel dataset *Compo\_ita\_TLIO* la forma *ognimodo*: il composto è infatti attestato per la prima volta nel secondo sotto-corpus, come testimoniato nell'esempio in (a), con valore di 'in qualunque maniera possibile'; nei

secoli successivi esso occorre, con lo stesso significato o con valore conclusivo, quasi esclusivamente all'interno della locuzione avverbiale *a ogni modo* (cfr. es. in (b)).

Nell'analisi della fase antica, si è osservato che *tuttafiata*, *tuttavia* e *tuttavolta* possono esprimere il valore concessivo di 'nondimeno', soprattutto se precedute da *ma*: come mostrano gli esempi in (e), (f) e (g), le tre forme sono attestate prevalentemente con questo valore nel CoDIt\_com. I dati estratti dal corpus permettono inoltre di osservare che l'uso di *tuttafiata* e *tuttavolta* diminuisce (nel primo caso, fino a scomparire) nel corso dei secoli, mentre *tuttavia* si impone come costruzione connettiva controaspettativa (cfr. Mauri & Giacalone Ramat 2008, 2009, 2012). I composti *ognora*, *talora*, *talvolta* continuano a essere utilizzati per indicare che un'azione è svolta 'sempre, di continuo', nel caso di *ognora* (cfr. esempio in (c)), o 'alle volte', nel caso di *talora* e *talvolta* (cfr. esempio in (d)); *tuttora*, attestato nella fase antica con il significato di 'di continuo', occorre nelle fasi più recenti con il valore di 'ancora, ancora adesso', come esemplificato in (i).

Nel complesso, la Costruzione<sub>COMP</sub> [A-N]<sub>AVV\_MOD</sub> continua grossomodo a essere istanziata dalle stesse forme attestate in italiano antico: eccetto alcuni composti che escono d'uso molto presto, gran parte degli avverbi attestati nella fase antica è ancora in uso nel sottocorpus più recente, sebbene in alcuni casi mostri di aver subito mutamenti dal punto di vista semantico-funzionale.

Nelle Costruzioni<sub>COMP</sub> attestate nel CoDIt\_com con primo elemento avverbiale il primo costituente è quasi sempre l'avverbio mal(e)- (circa il 90%): al sotto-schema [mal(e)-x] è dedicato il § 5.3 e sgg, a cui si rimanda per un'analisi approfondita nel quadro della teoria sulla formazione di affissoidi a partire da costituenti di composti. Eccetto male, gli unici avverbi attestati in questo tipo di composti con testa a destra sono alto e bene, le cui istanziazioni sono di seguito descritte ed esemplificate.

L'avverbio *alto* si combina a un participio presente o passato, come illustrano gli esempi riportati nella tabella 59.

Composti [alto- PART.PRES PART.PASS] <sub>NIA</sub>	1	2	3	4	5	tot	Esempi		
altitonante	-	3	2	9	-	14	«L' <i>altitonante</i> marito,	di	Giunon

Composti	1	2	3	4	5	tot	Esempi
[ <i>alto-</i> PART.PRES PART.PASS] <sub>NIA</sub>							
							Voglia che il dì del mio ritorno spunti, / Com'io nel dolce ancor nido nativo / Sempre, qual dea, t'onorerò: chè fosti / La mia salvezza tu, fanciulla illustre» (Pindemonte, Odissea)
altochiomante	-	-	-	-	1	1	«E giunti presso, videro sul mare, in una punta, l'antro, alto, coperto / di molto lauro, e v'era intorno il chiuso / di rozzi blocchi, e lunghi pini e quercie / altochiomanti. E il vecchio Eroe parlava []» (Pascoli, Poemi conviviali)
altocinto	-	-	-	1	7	8	«[] e la Marchesa dal profilo greco, / altocinta, l'un piede ignudo in mano, / si riposava all'ombra d'uno speco / arcade, sotto un bel cielo pagano» (Gozzano, I colloqui)
altolocato	-	-	-	-	3	3	«A Siracusa (egli continua) la costituzione si mutò per una querela amorosa che spinse all'insurrezione due giovani altolocati e i loro seguaci []» (Lombroso, L'uomo delinquente)
altorecato	-	-	-	-	1	1	«Il frutto <i>altorecato</i> assai mi tenta: / balzo sul plinto, il dono della Terra / tolgo alli acuti simboli di Guerra, / avvincendomi all'erma sonnolenta» (Gozzano, <i>Poesie sparse</i> )
altoriverso	-	-	-	-	1	1	«Bel fiore del carbone e dell'acciaio / Miss Ketty fuma e zufola giuliva / altoriversa nella sedia a sdraio» (Gozzano, Poesie sparse)
altosonante	-	-	2	-	4	6	«A che dunque tante parole altisonanti sull'arca santa

Composti	1	2	3	4	5	tot	Esempi
[alto-							
PART.PRES PART.PASS] <sub>N A</sub>							
							delle nostre istituzioni?» (Guerrini, <i>Brani di vita</i> )
altovolante	-	-	-	2	2	4	«Un'aquila comparve altovolante, / Che avea colomba trepida tra l'ugne» (Pindemonte, Odissea)

Tabella 59 Composti [alto- PART.PRES|PART.PASS]<sub>N|A</sub> attestati nel CoDIt\_com

Come illustrato dagli esempi riportati nella tabella 59, i composti con l'avverbio *alto-*, già attestati in latino classico, occorrono quasi esclusivamente in opere poetiche e in funzione di aggettivo riferito a referenti umani o divinità. Non è attestata nel CoDIt\_com la forma *altoparlante* (creata sul modello dell'inglese *loud-speaker* e attestata a partire dal 1927) che invece designa un artefatto. Questo sotto-schema continua quindi a essere produttivo nel corso dei secoli, fino alla fase più recente, ma istanzia composti che riproducono il modello latino e occorrono quasi esclusivamente in opere poetiche.

L'avverbio *ben(e)* può combinarsi con un notevole numero di basi, come illustrato nella tabella 60.

[bene-A] <sub>A N</sub>	[bene- AVV] <sub>AVV</sub>	[bene-N] <sub>N</sub>	[bene- PART.PRES] <sub>N A</sub>	[bene- PART.PASS] <sub>N A</sub>	[bene-V] <sub>V</sub>
benavventuroso (12) bensicuro (66) bencerto (118) bendegno (137) bennoto (143)	benvolentieri (48) benaltrimenti (32) bentosto (112) benpresto (190)	benessere (185) benestare (24) benfare (234)	benpensante (5) benparlante (17)	benaffetto (11) benamato (12) bendisposto (123) benaccetto (1) benarrivato (1) beninteso (99) benservito (9) bentornato (25) benvenuto (145) benvoluto (34) bendetto (60)	bendire (136) benvenire (68) benvolere (36) benguardare (121) benconsiderare (92)

[bene-A] <sub>A N</sub>	[bene- AVV] <sub>AVV</sub>	[bene-N] <sub>N</sub>	[bene- PART.PRES] <sub>N A</sub>	[bene- PART.PASS] <sub>N A</sub>	[bene-V] <sub>V</sub>
				bennato (141) benfornito (78) benpagato (33) bentornato (25) benvestito (81)	

Tabella 60 Composti con ben(e)- attestati nel CoDIt\_com

I dati estratti dal CoDIt\_com permettono di osservare che i composti con ben(e)costituiscono un sotto-schema semi-specificato che a sua volta istanzia ulteriori sottoschemi a seconda della base a cui ben(e)- si combina, che può essere un aggettivo
(gradabile), un avverbio, un nome, un participio presente o passato o un verbo. Il sottoschema più produttivo in termini di type è quello in cui la base è rappresentata da un
participio passato.

Dal punto di vista semantico, *ben(e)*- esprime una valutazione migliorativa nei confronti della base. In alcuni casi, esso può veicolare un significato intensivo, in particolare quando si combina a un aggettivo o a un avverbio, come si può notare dagli esempi riportati in (35).

- (35) «[...] quello sgabello son *ben sicura* d'averlo bruciato: e pure quando colei arriva, si trova sempre a quel posto, ed ella vi si adagia, e non vuol partire» (Manzoni, *Fermo e Lucia*)
  - b. «Capitandomi l'occasione *ben volentieri* adempirò quel che Ella m'ingiunge colla Signora Masucci, e col Marchese Melchiori, che ora passano la più parte dell'anno in campagna» (Leopardi, *Epistolario*)

Negli esempi in (35), *ben(e)*- intensifica il significato della base a cui si riferisce: così *ben sicura* vale 'molto sicura, sicurissima' e *ben volentieri* 'molto volentieri'.

Dal punto di vista quantitativo, i dati estratti dal CoDIt\_com permettono di osservare come evolve la produttività della Costruzione $_{\text{COMP}}$  [ben(e)-x] nel corso dei secoli.

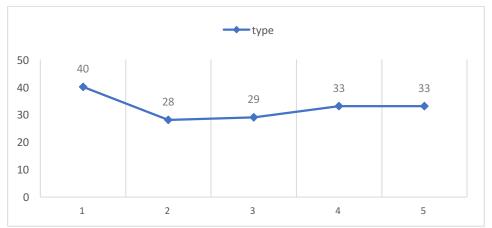


Figura 45 Costruzione<sub>COMP</sub> [ben(e)-x]: distribuzione dei type nei sotto-corpora del CoDIt\_com



Figura 46 Costruzione<sub>COMP</sub> [ben(e)-x]: distribuzione dei type nei sotto-corpora del CoDIt\_com

Come mostrato dai grafici riportati nelle figure 46 e 47, la Costruzione $_{\text{COMP}}$  [ben(e)-x] è particolarmente produttiva nelle fasi più antiche; nei secoli più vicini all'epoca contemporanea, il numero di type e soprattutto di token diminuisce lievemente. Questo tipo di Costruzione $_{\text{COMP}}$ , già attestata in latino classico (cfr. § 4.2.1), si conserva quindi stabilmente nel corso della storia dell'italiano: dal punto di vista semantico, ben(e)-esprime una valutazione migliorativa o intensiva (a seconda della categoria della base) nei confronti dell'elemento a cui si riferisce. Dal punto di vista formale e semantico, tale

Costruzione<sub>COMP</sub> è senz'altro legata alla Costruzione<sub>COMP</sub> [mal(e)-x], di cui rappresenta la controparte migliorativa/positiva.<sup>173</sup>

Infine, come già segnalato in apertura di questo paragrafo, i dati estratti dal CoDIt\_com permettono di identificare un sotto-schema non attestato nella fase antica, costituito da due sostantivi legati da una relazione di modificazione. In particolare, si tratta delle forme *capoluogo* e *capolavoro*: il primo è un calco dal francese *chef-lieu*, attestato a partire dal quarto sotto-corpus; il secondo, attestato fin dal secondo sotto-corpus, può essere considerato una creazione autonoma dell'italiano. In entrambi i casi, il sostantivo *capo* svolge la funzione di modificatore ed esprime il significato di 'principale'. Il CoDIt\_com attesta inoltre altri sei composti che presentano *capo* in prima posizione e in funzione di modificatore, riportati ed esemplificati nella tabella 61.<sup>174</sup>

Composti [capo-x] <sub>N_MOD</sub>	1	2	3	4	5	tot	Esempi
capobirro	-	-	-	-	1	1	«Non fu così d'un certo Paolini, capobirro famoso, il quale se uscì vivo di mano al popolo di Firenze, e' può attaccarne il voto []» (Giusti, Cronaca dei fatti di Toscana)
capocarceriere	-	-	-	1	-	1	«Mezz'ora dappoi, ecco stridere le chiavi; la porta s'apre: il <i>capocarceriere</i> mi portava una brocca d'acqua» (Pellico, <i>Le mie prigioni</i> )
capocomico	-	-	-	-	275	277	«Per far dispetto al capocomico, prima che cominciasse la prova generale, la piccola Giustina, col gomito appoggiato sul ginocchio d'una gamba accavalciata sull'altra []»

\_

<sup>&</sup>lt;sup>173</sup> Per ragioni di tempo non si approfondisce il confronto tra le due Costruzioni<sub>comp</sub>, per il quale si rimanda a lavori futuri.

<sup>&</sup>lt;sup>174</sup> In italiano contemporaneo, questo tipo di composti sembra essere particolarmente produttivo; il Devoto Oli ne attesta 16 forme: *capo-operaio*, *capobandito*, *capobrigante*, *capocannoniere*, *capocarceriere*, *capocomico*, *capocommessa*, *capocronista*, *capocuoco*, *capogabbiere*, *capoguardia*, *capomacchinista*, *caporedattore*, *caposcafista*, *capotecnico*, *capovaccaio*.

Composti [capo-x] <sub>N_MOD</sub>	1	2	3	4	5	tot	Esempi
capomaestro	1	13	5	-	-	19	«L'apportatore di questa sarà Bernardo Basso, capomaestro dell'Opera di San Lorenzo, al quale io vi prego pagiate la pigione m'avete a dare []» (Buonarroti, Lettere)
capomastro	1	2	2	1	16	22	«Un capomastro campagnuolo, ch'io so, condotto a ristaurare un ponte già fabbricato da' matematici, e poscia per venti anni, con evidentissimi calcoli e con mezza l'entrata delle gabelle annue del comune []» (Foscolo, Viaggio sentimentale)
caposoldo	-	2	-	-	-	2	«[] vostra excellentia ha fatto gratia libera d'una condennatione che gli havea data detto capitano, e la discordia è che'l Capitano voria exigere il caposoldo, cioè dui bolognini per libra de detta condennatione» (Ariosto, Lettere)

Tabella 61 Composti [capo-x]<sub>N MOD</sub> attestati nel CoDIt\_com

I composti riportati nella tabella 61, costituiti da *capo* e un sostantivo, generalmente riferito a un referente umano, possono essere interpretati in diversi modi. Una prima ipotesi li classifica come composti subordinativi con testa a sinistra: secondo questa intepretazione, *capo* rappresenterebbe l'elemento testa, da cui dipenderebbe il secondo elemento (il *capocarceriere* è il 'capo dei carcerieri'). Questi composti possono inoltre essere interpretati come coordinativi, in cui *capo* e l'altro elemento si trovano sullo stesso piano e designano un individuo che è sia capo sia carceriere. Una terza ipotesi vuole invece che *capo* sia il modificatore di un elemento testa posto a destra del composto (il *capocarceriere* è il 'carceriere che tra tutti è il capo'). Tali forme possono essere considerate istanziazioni di uno schema semi-specificato costituito da *capo* e una base ristretta, dal punto di vista semantico, a nomi che designano referenti umani: [*capo-*

x<sub>PERSONA</sub>]<sub>PERSONA</sub>. Un elemento che può contribuire alla classificazione di questi composti è la formazione del plurale (cfr. Micheli 2016a,b, 2018): infatti, a differenza dei composti con *capo*- chiaramente subordinativi (il tipo *capostazione*), che formano il plurale regolarmente modificando soltanto il primo costituente (es. *capistazione*, *capiscuola*, etc.), questi composti occorrono sia nella forma con doppia flessione (es. *capiredattori*), sia nella forma con flessione esterna, sul secondo costituente (es. *caporedattori*). <sup>175</sup> Questo dato supporterebbe quindi sia l'ipotesi secondo cui si tratterebbe di composti coordinativi, sia quella secondo cui si tratterebbe di composti con testa a destra e relazione di modificazione.

I dati estratti dal CoDit\_com permettono di tracciare un quadro delle Costruzionicomp con testa a destra che integra quanto osservato per la fase antica. In primo luogo, l'analisi ha rivelato che quasi tutte le Costruzionicomp con testa a destra individuate nella fase antica risultano attestate nel CoDIt\_com, con l'eccezione dello schema [N-PART.PASS]<sub>NIA SUB</sub> (es. occhibagliato, plebeuscito); tuttavia, si è osservato che alcuni sotto-schemi (in particolare, i sotto-schemi [N-PART.PRES]<sub>N|A\_SUB</sub>, [A-PART.PASS]<sub>N|A\_MOD</sub> e [N-V]<sub>V\_SUB</sub>) sono istanziati quasi esclusivamente da composti di antica attestazione o di origine latina. Tali schemi sembrano quindi aver smesso di istanziare nuove forme nel corso dei secoli rappresentati dal CoDIt com. L'analisi dello schema [N-PART.PRES]<sub>NIA</sub> (es. luogotenente) dimostra però come una Costruzione come che apparentemente sta esaurendo la propria produttività, limitandosi a istanziare forme sul modello latino e utilizzate soltanto nel linguaggio poetico (ondivagante, armilucente), possa tornare a essere produttiva in una fase successiva, grazie soprattutto alla diffusione di un sottoschema semi-specificato (in questo caso, lo schema [x-dipendente] PERS, tossicodipendente, cibodipendente, cellularedipendente, etc.). L'analisi dei composti con testa a destra attestati nel CoDIt\_com ha infatti messo in luce come, in non pochi casi, la vitalità di una Costruzione comp sia strettamente legata a quella di particolari sotto-schemi semi-specificati: si pensi al caso degli schemi [mal(a/o)-x] e [mal(e)-x] (a cui sarà dedicato il capitolo 5), le cui istanziazioni costituiscono la quasi totalità dei composti [A- $N_{N|A} e [AVV-X]_X$ .

\_

<sup>&</sup>lt;sup>175</sup> In particolare, nel corpus itWaC, la forma *caporedattori* occorre 64 volte; *capiredattori* registra 72 occorrenze (cfr. Micheli 2018: 54)

Infine, un ulteriore elemento messo in luce durante l'analisi è l'emergere di una  $Costruzione_{COMP}$  [N-N]<sub>N\_MOD</sub>, istanziata dalle forme *capoluogo*, *capolavoro* e non attestata in italiano antico.

La Gerarchia di eredità delle Costruzioni<sub>COMP</sub> con testa a destra, elaborata sulla base dei dati estratti dal CoDIt\_com, è rappresentata nella figura 48.

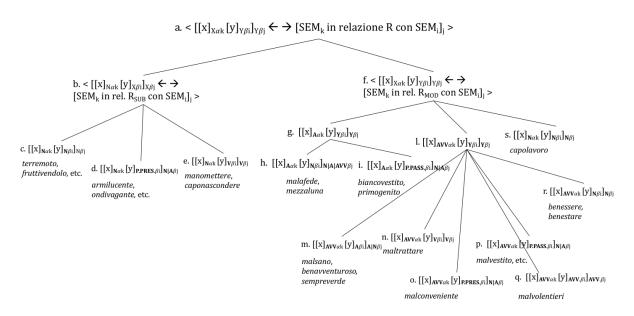


Figura 47 Gerarchia di eredità delle Costruzioni<sub>COMP</sub> con testa a destra, elaborata sulla base dei dati estratti dal CoDIt\_com

Rispetto alla Gerarchia dei composti con testa a destra dell'italiano antico, quella elaborata sulla base del CoDIt\_com presenta soltanto tre sotto-schemi (rappresentati ai punti (c), (d) (e)) in cui i due costituenti sono legati da una relazione di subordinazione. Il ramo delle Costruzioni<sub>COMP</sub> che presentano una relazione di modificazione si arricchisce invece dello schema in (s), costituito da due nomi (es. *capoluogo*).

#### 4.3.3 La Costruzione<sub>COMP</sub> dei composti coordinativi endocentrici

Nel § 4.2.3 si è osservato che la Costruzione<sub>COMP</sub> dei composti coordinativi endocentrici dell'italiano antico è istanziata da due sotto-schemi:

#### 1. $[N-N]_N$ (es. cassapanca)

## 2. [A-A]<sub>A|N</sub> (es. *libifenicio*, *sacrosanto*)

I composti  $[A-A]_{A|N}$  (sottoschema in (2)), in italiano antico, sono prevalentemente aggettivi che indicano la provenienza geografica di un individuo, in gran parte di origine latina (es. *celtibero*, *gallogreco*).

Nel CoDIt\_com sono attestati 35 composti che istanziano la Costruzione $_{\text{COMP}}$  dei coordinativi endocentrici; in particolare, si tratta di 10 composti [N-N] $_{\text{N}}$ , 25 composti [A-A] $_{\text{A|N}}$ : rispetto alla fase antica, quindi, tale Costruzione $_{\text{COMP}}$  istanzia gli stessi sotto-schemi. I composti [N-N] $_{\text{N}}$  sono riportati ed esemplificati nella tabella 62.

Composti [N-N] <sub>N</sub>	1	2	3	4	5	tot	Esempio
caffelatte	-	-	-	-	8	8	(a) «A vostra norma, questa crosta di gusto delicatissimo, deve avere il colore del <i>caffè latte</i> . Al caffè, volendo, si può sostituire la cioccolata infusa» (Artusi, <i>La scienza in cucina e l'arte di mangiar bene</i> )
cassabanca/ cassabanco	-	3	-	5	-	8	(b) «Un gran <i>cassabanco</i> , che la sera si trasforma in letto per chi dorme in sala, coperto da un panno verde a frangie usato e tempestato di frittelle d'una lucerna» (D'Azeglio, <i>Racconti</i> )
cassapanca	1	4	-	-	5	10	(c) «Con quelle povere mani sbiancate e raspose, che sapevano ogni fatica, levava dalla vecchia cassapanca d'abete, lunga e stretta che pareva una bara []» (Pirandello, Scialle nero)
lattemiele	-	-	-	-	1	1	(d) «"Vivanda-ritratto dell'amico bruno": gote ben modellate di pasta frolla — baffi e capelli di cioccolato — grandi cornee di <i>lattemiele</i> — pupille di liquerizia» (Marinetti, <i>Cucina futurista</i> )
mazzafrusto	1	14	4	1	-	20	(e) «[] costui soleva cavalcando in battaglia portare un <i>mazzafrusto</i> di ferro, col quale virilmente gl'inimici abbatteva» (D'Azeglio, <i>Racconti</i> )
pianoforte (o pianforte)	-	-	-	25	127	152	(f) «[] se tu mi assicurerai di amarmi ancora dove sei, e di ricordarti spesso di me che ti scrivo al suono del pianoforte di Carlo, che non mi lascia

Composti [N-N] <sub>N</sub>	1	2	3	4	5	tot	Esempio
							distinguere i spropositi che faccio, per terminare i quali ti lascio» (Leopardi, <i>Epistolario</i> )
							(g) «Il <i>pianforte</i> e ogni altro mobile stanno in questa nuova stanza assai ben situati []» (Belli, <i>Lettere</i> )
piantanimale	-	-	1	-	-	1	(h) «L'ostriche, le spugne, e gli altri simili animali, che <i>piantanimali</i> nelle scuole son chiamati» (Redi, <i>Esperienze intorno alla generazione degli insetti</i> )
servopadrone	-	-	-	1	-	1	(i) «In questo frattempo il già mio Piemonte, celtizzato anch'egli, scimmiando ogni cosa dei suoi servipadroni, cambiò l'Accademia sua delle Scienze, già detta Reale, in un Istituto Nazionale a norma di quel di Parigi, dove avean luogo, e le belle lettere, e gli artisti» (Alfieri, Vita)
tragicommedia	-	-	2	11	2	15	(l) «Volle fare una tragicommedia, mescolanza di elementi tragici e comici in un ordito largo e ricco, dove fossero innestate più azioni» (De Sanctis, Storia della letteratura)

Tabella 62 Composti [N-N]<sub>N</sub> coordinativi endocentrici attestati nel CoDIt\_com

I composti [N-N]<sub>N</sub> esemplificati nella tabella 62 designano artefatti che svolgono una duplice funzione (come *cassabanca/cassabanco, cassapanca, mazzafrusto,*<sup>176</sup> *pianoforte*)<sup>177</sup> o cibi costituiti dall'insieme di due ingredienti (*caffelatte, lattemiele*); tra i nomi che indicano referenti animati vi sono le forme *piantanimali* e *servopadrone*. Nel primo caso, si tratta di un calco del composto latino di origine greca *zoophitum*, come suggerito dell'esempio riportato dalla Terza edizione del Vocabolario degli Accademici della Crusca:

\_

<sup>&</sup>lt;sup>176</sup> Il *mazzafrusto* (*mazza+frusta*) è un'arma costituita da una mazza, all'estremità della quale si trova una catena con appesa una sfera munita di numerosi puntoni: il composto rientra quindi tra i nomi che designano entità che partecipano delle proprietà espresse dai due costituenti.

<sup>&</sup>lt;sup>177</sup> Il pianoforte, nato negli ultimi anni del XVII secolo, era denominato originariamente *gravicembalo col piano e col forte* o *fortepiano*: in questo caso, i due costituenti alludono alla capacità di tale strumento di suonare sia il piano sia il forte.

(36) «I quali i Greci componendo, ec. due nomi insieme, gli chiamaron zoofita, cioè animali, e piante, come se noi dicessimo Animalipiante, o veramente *Piantanimali*, come sono le spugne marine, e alcuni altri, ec.» (Varchi, *Lezzioni*, 1590)

L'esempio in (36) spiega come *piantanimali* sia usato per indicare essere viventi che partecipano sia della natura di piante sia di quella di animali: il termine è probabilmente un calco creato in modo estemporaneo da Benedetto Varchi per classificare animali come le spugne marine, etc.

Nel caso di *servopadrone*, si tratta di un hapax che indica il duplice status di servi e padroni ricoperto dai francesi dominatori del Piemonte: questo tipo di composto coordinativo, indicante una persona che ricopre un duplice status e rappresentato da quest'unico caso nel CoDIt\_com, acquisirà maggior produttività dopo la seconda metà del Novecento, come testimonia la diffusione di forme come *studente-lavoratore*, *attorecantante*, etc. (cfr. Arcodia, Grandi & Wälchli 2010).

Infine, il composto *tragicommedia*, attestato nel CoDIt\_com a partire dal terzo sotto-corpus, ha origine dalla forma latina *tragicomoedia* e indica un componimento teatrale che presenta elementi sia del genere comico sia di quello tragico.

La Costruzione<sub>COMP</sub> [A-A]<sub>A|N</sub> istanzia 22 composti costituiti da due aggettivi legati da un rapporto di coordinazione. All'interno di questo gruppo di forme è possibile isolare un sotto-gruppo costituito da 12 composti che presentano come tipo semantico di output PROPRIETÀ e, in particolare, indicano una tonalità di colore. Tali forme sono riportate ed esemplificate nella tabella 63.

Composti [A-A] <sub>A N</sub>	1	2	3	4	5	tot	Esempi
biancoazzurro	-	-	-	-	2	2	(a) «La cupola dai dodici occhi tondi / il bianco-azzurro fregio dei festoni / i fiori i frutti gli òvoli i dentelli []» (D'Annunzio, Laudi)
biancogiallo	-	-	-	-	1	1	(b) «È sempre quella, la battaglia. È sempre quella, il Papa contro il Re, il <i>biancogiallo</i> contro il tricolore, la tirannide contro la libertà []» (Guerrini, <i>Brani di vita</i> )
celesteazzurro	-	-	-	2	-	2	(c) «[] in più luoghi il Manto che di <i>celeste</i> azzurro la ricopriva, in dimostranza forse del Patrocinio del Cielo, sotto cui continuamente è vissuta» (Cascina, <i>Alfea</i> )

Composti [A-A] <sub>A N</sub>	1	2	3	4	5	tot	Esempi
gialloarancione	-	-	-	-	1	1	(d) «Gli ovoli, come sapete, sono di color <i>giallo-arancione</i> ; i più giovani sono chiusi in forma d'uovo, i più maturi sono aperti e spianati» (Artusi, <i>La scienza in cucina e l'arte di mangiar bene</i> )
giallorosso	-	-	-	-	1	1	(e) «Così formerete dei crostini che potranno contornare un piatto d'arrosto i quali avendo per base il pane fritto coperto dal verde degli spinaci, col bianco e il <i>giallo-rosso</i> delle uova figureranno di tre colori; ma sono più belli che buoni» (Artusi, <i>La scienza in cucina e l'arte di mangiar bene</i> )
gialloverde	-	-	-	-	1	1	(f) «O rabido ventare di scirocco che l'arsiccio terreno <i>gialloverde</i> bruci; e su nel cielo pieno di smorte luci trapassa qualche biocco di nuvola» (Montale, <i>Ossi di seppia</i> )
grigioroseo	-	-	-	-	1	1	(g) «Esterina, i vent'anni ti minacciano, grigiorosea nube che a poco a poco in sé ti chiude» (Montale, Ossi di seppia)
neroazzurro	-	-	-	-	2	2	(h) «[] le labbra arcuate erano fatte per quel sorriso lungo, profondo e cosciente che poche donne conoscono; le trecce folte, brune, s'incupivano in un nero azzurro» (Serao, Leggende napoletane)
neroblu	-	-	-	-	2	2	(i) «La salvatichetta si stringeva contro il braccio di Vana, lanciando di tratto in tratto un'occhiata torva all'ospite, di sotto il lustro nero blu della sua capellatura» (D'Annunzio, Forse che sì forse che no)
rossogiallo	-	-	-	-	1	1	(l) «Saprete che la pesca cotogna è quella rosso-giallo o semplicemente giallastra, con la polpa attaccata al nocciolo» (Artusi, La scienza in cucina e l'arte di mangiar bene)
verdeazzurro	-	8	-	2	6	16	(m) «Verde azzurro. Questo è un minerale di cobalto, il quale deve il suo colore verde al rame, al ferro e allo zinco con cui è combinato. Ed è un colore di assai durata. Verde è un colore, el quale è mezzo naturale; e questo si fa artifizialmente, chè si fa d'azzurro della Magna; e questo si chiama verde azzurro» (Cennini, Libro dell'arte)

Composti [A-A] <sub>A N</sub>	1	2	3	4	5	tot	Esempi
verdegiallo	-	2	4	-	4	10	(n) «Nuove arme ritrovò, nuovo cavallo, / E sopravveste nera e scudo nero / Portò, fregiato a color <i>verdegiallo</i> » (Ariosto, <i>Orlando furioso</i> )  (o) «Tutto giallo, del <i>verde-giallo</i> che hanno i bambini malati, poveretto!» (Verga, <i>Novelle rusticane</i> )

Tabella 63 Composti coordinativi [A-A]<sub>A|N</sub> che indicano tonalità di colore attestati nel CoDIt\_com

Con l'eccezione di *verdeazzurro* e *verdegiallo*, i composti esemplificati nella tabella 63 sono attestati soltanto nel sotto-corpus più recente (1841-1947). Il composto più antico è infatti *verdeazzurro*, la cui prima attestazione è nel trattato sull'arte, scritto per la prima volta in volgare intorno all'inizio del Quattrocento, di Cennino Cennini; *verdegiallo* è invece attestato per la prima volta circa un secolo dopo, nella terza edizione dell'*Orlando fuorioso* (1532). In entrambi i casi si tratta di tonalità di colore intermedie tra verde e azzurro/giallo; dello stesso tipo è il composto *celesteazzurro*, con cui è descritto il colore del cielo. I composti *giallo-arancione*, *giallo-rosso* e *rosso-giallo* occorrono soltanto nel ricettario di Pellegrino Artusi e sono funzionali a descrivere il colore delle uova e delle pesche: anche in questo caso si tratta di tonalità a metà strada tra i colori espressi dai due costituenti. Similmente, le forme *neroazzurro* e *neroblu* individuano tonalità di capelli neri che presentano riflessi bluastri. Diverso è invece il caso di *biancoazzurro* e *biancogiallo*, entrambi attestati nella prima metà del Novecento, utilizzati per descrivere, rispettivamente, un fregio e una bandiera che presentano entrambi i colori.

Sulla falsariga di quanto proposto da Grossmann & D'Achille (2013: 527), i composti coordinativi  $[A-A]_{A|N}$  attestati nel CoDIt\_com possono essere classificati in due tipi sulla base di proprietà semantiche: il primo con significato intersettivo (parafrasabile in 'tra  $A_1$  e  $A_2$ ', come *giallorosso* nell'esempio (e)), il secondo additivo (parafrasabile in ' $A_1$  e  $A_2$ ', come *biancogiallo*, nell'esempio (b)). I dati estratti dal CoDIt\_com mettono in luce come il tipo additivo possa essere considerato più recente rispetto a quello intersettivo, attestato stabilmente nel corso della storia dell'italiano fin dalle Origini (cfr. al § 4.2.3 il composto *biancobigio* 'bianco tendente al grigio', attestato nel dataset *Compos\_ita\_TLIO*).

Oltre alle forme che individuano tonalità di colori, nel CoDIt\_com occorrono altri 10 composti che istanziano la Costruzione<sub>COMP</sub> [A-A]<sub>AIN</sub>, riportati nella tabella 64.

Composti [A-A] <sub>A N</sub>	1	2	3	4	5	totale
agrodolce	-	-	-	-	4	4
analogicoorganico	-	-	-	-	1	1
angloamericano	-	-	-	1	-	1
celtibero	1	3	6	-	-	10
chiaroscuro	-	4	9	11	12	36
duroliquido	-	-	-	-	1	1
eroicomico	-	-	-	-	2	2
sacrosanto	5	6	39	29	39	118
sordomuto	-	-	-	1	8	9
tragicomico	-	-	1	-	1	2

Tabella 64 Composti [A-A]<sub>A|N\_COORD</sub> (esclusi gli aggettivi di colore) attestati nel CoDIt\_com

In due casi si tratta di composti coordinativi che esprimono la provenienza geografica: celtibero, già attestato in italiano antico e di origine latina; angloamericano attestato nel sotto-corpus più recente; gli altri composti indicano la compresenza di due proprietà riferita a persone (es. sordomuto) o entità concrete/astratte (es. agrodolce, analogicoorganico). Dal punto di vista formale, questo tipo di forme è frequentemente caratterizzato da un primo costituente bisillabico, spesso terminante in -o (o apocopato), e un secondo elemento di lunghezza variabile. Come già osservato in Grossmann & Rainer (2009), la presenza di una -o come vocale finale del primo costituente è caratteristica di una parte dei composti aggettivali [A-A]<sub>A</sub>: in particolare, analizzando questo tipo di composti in diacronia, i due autori identificano un primo tipo di composti in -o, vocale mutuata dai composti latini del tipo comico-tragicus e divenuta marca di composizione (il tipo 'discipline fisico-matematiche'), e un secondo tipo 'asindetico', più recente, in cui il primo costituente presenta altre vocali e i due aggettivi sono semplicemente giustapposti senza ulteriori marche di composizione (il tipo 'corrente marxista-leninista'). Il tipo asindetico sembra essersi formato più recentemente dalla rianalisi del tipo in -o, nel tentativo di adattare tale tipo alla struttura dei composti italiani attraverso la coordinazione asindetica. I dati estratti dal CoDIt\_com permettono di confermare l'ampia diffusione del tipo in -o, ma non offre ulteriori spunti rispetto agli studi precedenti (cfr. anche D'Achille & Grossmann 2009, 2010).

Semanticamente, si tratta di aggettivi riferiti a referenti umani, come anche nel caso di *sordomuto*, calco dal francese *sourd-muet*.<sup>178</sup> Non sono attestati nel corpus ma appartengono a questa Costruzione<sub>COMP</sub> (e, in particolare, al tipo *angloamericano*) anche composti di più recente formazione, come *italoamericano*,<sup>179</sup> che indicano generalmente la doppia nazionalità e hanno quindi un significato additivo. La decurtazione del primo costituente è presente anche in *eroicomico* e *tragicomico* (derivato da *tragicommedia*): entrambi gli aggettivi si riferiscono a eventi (tipo semantico EVENTO) o componimenti (tipo semantico INFO) le cui proprietà sono date dalla somma di quelle espresse dai due costituenti.

La figura 49 rappresenta la Gerarchia di eredità della Costruzione<sub>COMP</sub> dei composti coordinativi endocentrici, sulla base dei dati estratti dal CoDIt\_com.

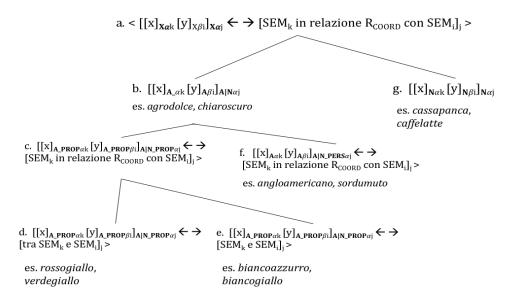


Figura 48 Gerarchia di eredità della Costruzione<sub>COMP</sub> dei composti coordinativi endocentrici (dati estratti dal CoDIt\_com)

Rispetto a quanto osservato per la fase antica, i composti che istanziano la  $Costruzione_{COMP}$  dei composti coordinativi endocentrici presentano alcuni elementi di continuità e alcune differenze. Per quanto riguarda lo schema  $[N-N]_{N\_COORD}$ 

\_

<sup>&</sup>lt;sup>178</sup> Cfr. De Mauro (2004), s.v. sordomuto.

<sup>&</sup>lt;sup>179</sup> Attestato a partire dal 1950, cfr. De Mauro (2004), s.v. *italoamericano*.

(rappresentato al punto (g)), esso istanzia generalmente nomi che designano entità animate o (soprattutto) inanimate che presentano proprietà di due entità. Nel periodo di tempo rappresentato dal CoDIt\_com (Origini-1947), tale Costruzione<sub>COMP</sub> istanzia prevalentemente composti con tipo semantico di output ARTEFATTO e solo sporadicamente PERSONA; essa, inoltre, non sembra istanziare composti con tipo semantico di output LUOGO, indicanti luoghi adibiti a due funzioni (come *ristorante-pizzeria*, *bar-pasticceria*, etc.).

Per quanto riguarda lo schema [A-A]<sub>A|N\_COORD</sub> (rappresentato al punto (b)), esso istanzia un considerevole numero di composti che indicano tonalità intermedie tra due colori (semantica intersettiva) o, ma solo nella fase più recente, la somma di due colori (semantica additiva). Questo tipo di forme, scarsamente attestato nella fase antica, si afferma nel corso dei secoli. In ottica costruizionista, le restrizioni semantiche che operano su queste forme possono essere rappresentate attraverso i tipi semantici di input e output (tutti PROPRIETÀ) e le perifrasi semantiche poste a destra degli schemi, come illustrato ai punti (c) (d) e (e) nella figura 49.

Si è inoltre osservato che questa Costruzione<sub>COMP</sub> continua a istanziare aggettivi riferiti prevalentemente a referenti umani (tipo semantico PERSONA, es. *sordomuto*, vd. sottoschema rappresentato in (f)); in particolare, sono attestati composti che indicano la provenienza geografica: gran parte delle forme di origine latina attestate nella fase antica si perde (es. *libifenicio*, *gallogreco*), ma se ne formano altre (es. *angloamericano*) che conservano la struttura con primo costituente bisillabico. Tale restrizione relativa alla struttura del costituente di sinistra è presente in tutti i composti attestati nel CoDIt\_com, ad eccezione di *analogicoorganico*, attestato soltanto una volta nel sotto-corpus 5. La Costruzione<sub>COMP</sub> [A-A]<sub>A|N</sub> è quindi generalmente sfruttata per creare aggettivi che designano la somma di due proprietà compresenti, riferite a persone (come nel caso di *sordomuto*) o entità astratte/concrete (come *eroicomico*, *agrodolce*).

## 4.3.4 La Costruzione<sub>COMP</sub> dei composti esocentrici

Come osservato al § 4.2.4, la Costruzione $_{\text{COMP}}$  dei composti esocentrici dell'italiano antico è istanziata da tre sotto-schemi:

- 1.  $[V-N]_{N\_SUB}$  (es. parasole, crepacuore)
- 2. [N-A]<sub>N|A\_MOD</sub> (es. *pettirosso*, *capochino*)
- 3.  $[V-V]_{N\_{COORD}}$  (es. portareca, andirivieni)

Oltre a questi sotto-schemi, che continuano a istanziare composti senza soluzione di continuità nel corso della storia dell'italiano, i dati estratti dal CoDIt\_com permettono di identificare un ulteriore sottoschema:

# 4. [V-AVV]<sub>N\_MOD</sub> (es. buttafuori, mettimale)

Di seguito si delinea un quadro delle proprietà formali e semantiche e la frequenza type/token di ciascun sotto-schema in ottica diacronica.

Il CoDIt\_com attesta 324 composti che istanziano la Costruzione [V-N], ben attestata fin dalla fase antica e che si mantiene produttiva nel corso dei secoli: la frequenza dei type e dei token dei composti nei diversi sotto-corpora del CoDIt\_com è illustrata nelle figure 50 e 51.

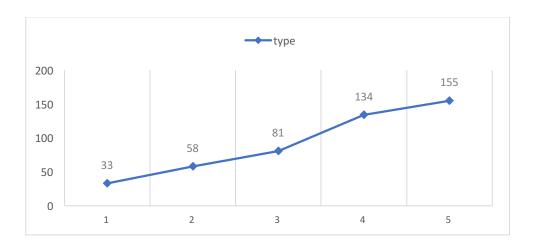


Figura 49. Frequenza dei type della Costruzione COMP [V-N] nei sotto-corpora del CoDIt\_com

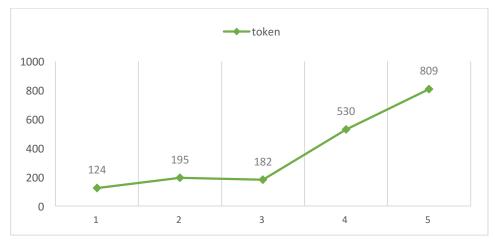


Figura 50 Frequenza dei token della Costruzione<sub>COMP</sub> [V-N] nei sotto-corpora del CoDIt\_com

La produttività della Costruzione<sub>COMP</sub> [V-N] in termini sia di type sia di token aumenta gradualmente nel corso del tempo: in particolare, è a partire dal periodo individuato dal terzo sotto-corpus che il numero di composti [V-N] aumenta sensibilmente.

I cinque composti [V-N] più frequenti sono riportati ed esemplificati nella tabella 65.

Composti [V-N]	1	2	3	4	5	tot	Esempi
<u>parapetto</u>	2	53	6	13	55	129	(a) «Scalone a stucchi del tempo, al quale per compimento ci sarebbe voluto, secondo lo stile, un parapetto a colonnette di marmo o almen di stucco» (D'Azeglio, Racconti)
passaporto	-	-	-	88	37	125	(b) «Quando poi si potesse aver subito un <i>passaporto</i> per l'interno, senza bisogno dei detti mezzi, gradirei di averlo, giacchè con esso non sarà difficile, io credo, di ottener l'altro ai confini» (Leopardi, <i>Epistolario</i> )
passatempo	-	3	25	47	37	112	(c) «Queste cose le scrivo così minutamente, non solo perchè togliendomi la lontananza di poter ragionar con lei, mi giova di scriverli famigliarmente molte di quelle cose de le quali le

Composti [V-N]	1	2	3	4	5	tot	Esempi
							parlerei per <i>passatempo</i> » (Tasso, <i>Lettere</i> )
guardaroba	1	8	25	12	24	70	(d) «Non seppe o non volle dirmi nulla: ma, cammin facendo, scoprii dal famiglio che mio marito ha fatto mettere in sesto un appartamento del castello; e tratto ogni vecchio arnese dalla <i>guardaroba</i> , ne ha addobbate le camere facendo egli stesso da tappezziere ed apparatore» (Nota, <i>La fiera</i> )
							(e) «Abbiamo di là un intero guardaroba, tutto di costumi del tempo, eseguiti a perfezione, su modelli antichi» (Pirandello, Enrico IV)
marciapiede	-	-	-	-	53	53	(f) «Nei primi giorni, quando cominciò ad uscire in carrozza, accanto alla madrigna, la gente si fermava sui <i>marciapiedi</i> , l'aspettava al varco, dinanzi al portone del palazzo, per figgerle gli occhi addosso, a bocca aperta []» (De Roberto, <i>I Vicerè</i> )

Tabella 65 Composti [V-N] più frequenti nel CoDIt\_com

Tra i composti [V-N] più frequenti vi sono sia forme di antica attestazione (come *parapetto*, *passatempo* e *guardaroba*) sia forme create nella fase più recente (come *passaporto* e *marciapiede*); in due casi, *guardaroba* e *marciapiede*, si tratta di prestiti dal francese (rispettivamente, *garde-robe* e *marchepied*), entrati in italiano in diversi momenti. Come mostrano gli esempi in (d) e (f), il composto *guardaroba* occorre fino alla seconda metà dell'Ottocento al genere femminile, probabilmente per influenza del francese: risalgono ai primi del Novecento le prime attestazioni d'uso al maschile, genere con cui è comunemente usato in italiano contemporaneo.

Nei composti attestati nel CoDIt\_com occorrono 136 basi verbali diverse: le più produttive sono *batti-* (18 type), *porta-* (18 type), *guarda-* (17 type), *para-* (12 type), *mangia-* (10 type) e *salta-* (10 type). Si tratta di basi verbali con struttura prevalentemente bisillabica (fanno eccezione soltanto 17 forme su 136):<sup>180</sup> tale restrizione fonologica, già riscontrata in italiano antico, sembra quindi operare senza soluzione di continuità nel corso dei secoli. I verbi coinvolti appartengono prevalentemente alla classe flessiva in *-are*: in particolare, la distribuzione dei type rispetto alla classe flessiva è riportata nella tabella 66.

-are	-ere	-ire	tot
117	16	3	136

Tabella 66 Distribuzione dei composti [V-N] nel CoDIt\_com rispetto alla classe flessiva dell'elemento verbale

Tra i verbi in –ere che occorrono all'interno dei composti [V-N], i più produttivi sono battere (18 type, es. batticuore), rompere (5 type, es. rompicapo) e torcere (4 type, es. torciorecchia), pascere (3 type, es. pascivento) e perdere (2 type, es. perdigiorno); negli altri casi si tratta di basi verbali che occorrono in un solo composto. Gli unici verbi in – ire attestati in questo tipo di forme sono coprire (3 type, es. copricapo), aprire (1 type, es. apricessi) e fuggire (1 type, es. fuggifatiche).

Prendendo in esame il periodo più recente testimoniato dal CoDIt\_com (1841-1947), si può notare che le basi verbali più frequenti sono porta- (9 type: portabaldacchini, portacenere, portafiori, portamantelli, portapanni, portaritratti, portasigarette, portaspilli, portorologio), spacca- (3 type: spaccamondi, spaccamontagne, spaccapietre) e mangia- (3 type: mangialiberali, mangiamaccheroni, mangiamoccoli); i restanti 42 composti si distribuiscono tra ben 35 basi diverse: questo dato induce a ritenere che la produttività mostrata da questo tipo di composti riguardi l'intera Costruzione astratta [V-N] e non soltanto particolari sotto-schemi semi-specificati (come ad esempio [porta-x]). In particolare, il sotto-schema costituito da porta- e un sostantivo variabile istanzia nomi che indicano un individuo (tipo semantico PERSONA) che trasporta un oggetto o uno

<sup>&</sup>lt;sup>180</sup> In particolare, si tratta delle seguenti basi verbali: fa-, attacca-, affitta-, arruffa-, scrutina-, asciuga-, strascina-, acchiappa-, cavalca-, imbratta-, azzecca-, schiccara-, acconcia-, appiccia-, spaventa-, stuzzica-, scavezza-.

strumento (tipo semantico ARTEFATTO) che contiene un oggetto (tipo semantico ARTEFATTO): i due significati della base verbale *porta*- sono esemplificati in (37).

(37) a. «Per la prima ipotesi, oltre all'essere caso raro, non vi crederanno alla cieca, tanto più adesso che abbiamo quasi i baffi; giacché ci vogliono fisionomie speciali, e perfino speciali barbe per *portabaldacchini*» (Raiberti, *L'arte di convitare*)

b. «[...] sul piano di bardiglio, una boccia d'acqua col bicchiere capovolto, un portacenere, un portorologio» (Pirandello, Maschere nude)

Il composto *portabaldacchini*, esemplificato in (37a), costituisce l'unico esempio in cui *porta*- veicola il significato di 'trasportare': il composto indica una persona che trasporta un baldacchino; negli altri casi, esemplificati da *portacenere* e *portorologio*, si tratta di oggetti che contengono qualcosa. <sup>181</sup> Come osservato al § 4.2.4, questo secondo significato di *porta*- non è attestato nella fase antica, in cui il sotto-schema [*porta*-x] istanzia soltanto la forma *portapesi*, attestata nel significato di 'facchino' (quindi di 'persona che trasporta dei pesi'). La prima attestazione di *porta*- 'che contiene' nel CoDIt\_com è costituita dal composto *portafoglio*, attestato a partire dal quarto sotto-corpus: la forma è un calco dal francese *porte-feuille* ed è utilizzata per indicare una custodia che contiene fogli e documenti, come esemplificato dagli esempi in (38).

(38) a. «Incontrare l'esercito francese, che si gettava rotto e disordinato sul Reno dopo la battaglia di Lipsia, vestito da auditore con un *portafoglio* sotto braccio, era cosa da non piacere a tutti» (D'Azeglio, *Racconti*)

b. «Mi sovviene che in una sua lettera che non ho qui sott'occhio, ma che aveva messa nel mio *portafoglio* da viaggio per risponderle a voce a Firenze, Ella mi fa qualche ricerca rispetto ai regolamenti di questa Censura in riguardo ai romanzi» (Leopardi, *Epistolario*)

I due esempi di *portafoglio* esemplificati in (38) mostrano come il composto sia utilizzato per indicare un oggetto (tipo semantico ARTEFATTO) che contiene fogli o documenti,

214

<sup>&</sup>lt;sup>181</sup> Si noti che tra i significati del verbo *portare* come forma libera non è registrato quello di 'contenere', che occorre soltanto quando *porta-* costituisce il primo elemento di un composto [V-N] (cfr. De Mauro 2004, s.v. *portare*).

simile quindi a una 'cartellina' piuttosto a che all'accessorio che nell'italiano del XXI secolo chiameremmo *portafoglio*. Dal suo ingresso nel lessico dell'italiano, il composto ha avuto notevole fortuna in termini di frequenza (esso occorre 13 volte nel quarto sottocorpus e 27 nel quinto) e può aver certamente influito nella creazione e diffusione del significato '(strumento) che contiene x' del sotto-schema [*porta*-x], come mostrato in (39):

(39)  $\langle [[porta]_{Vk}[x]_{N\beta i}]_{N\gamma j} \leftarrow \rangle$  [ARTEFATTO che CONTIENE [SEM]<sub>i</sub>]<sub>j</sub>  $\rangle$ 

Come già osservato più volte precedentemente, un prestito/calco può contribuire alla creazione e alla fortuna di un particolare sotto-schema che costituisce un'innovazione rispetto dal punto di vista della forma o, come in questo caso, della semantica.

Per quanto riguarda la categoria di output, i dati estratti dal CoDIt\_com permettono di collocare nel periodo individuato dal secondo sotto-corpus (1376-1532) le prime attestazioni di composti [V-N] con valore attributivo, come mostrato dai seguenti esempi:

(40) a. «Questo Vescovo *lavaceci*, vogliendo ammaestrare nel vizio della gola, riprendea li Fiorentini dicendo» (Sacchetti, *Trecentonovelle*)

b. «E a dì 4 d'agosto 1511, affogò 3 uomini *vuotacessi*, in un certo pozzo nero [...]» (Landucci, *Diario fiorentino*)

c. «Ell'ha quegli occhi tanto *rubacuori*, che la trafiggere' con egli un muro» (Lorenzo de' Medici, *Nencia da Barberino*)

Gli esempi riportati in (40) mostrano le prime attestazioni di composti [V-N] in funzione di modificatori di nomi: in particolare, in (40a) e (40b), i due composti *lavaceci* e *vuotacessi* si riferiscono a referenti umani; in (40c), il composto *rubacuori* è riferito agli occhi di una donna. Inoltre, va notato che i primi due composti occorrono all'interno del corpus anche come sostantivi, mentre il terzo esclusivamente come modificatore.

L'uso in funzione attributiva dei composti [V-N] rimane comunque raro nel corso di tutta la storia della lingua italiana: dei 136 composti attestati nel CoDIt\_com, soltanto 12 occorrono anche come modificatori, nella maggior parte dei casi riferiti a persone. In

particolare, si tratta delle forme: castraporci, lavaceci, mangiapolenta, perdigiorni, portalettere, rompicollo, rubacuori, scannapagnotte, scansafatiche, storcileggi, tagliapane, vuotacessi. Di questi composti, alcuni occorrono esclusivamente in funzione di modificatore: si tratta di 'medico castraporci', 'occhi rubacuori', 'dottore storcileggi', 'coltelli tagliapane'.

Sebbene la diffusione di composti [V-N] con funzione di aggettivo assuma rilievo soltanto nella fase contemporanea, come osservato da Ricca (2005, 2010), i dati estratti dal CoDIt\_com inducono a ritenere che quest'uso fosse possibile anche nei secoli precedenti e che quindi la Costruzione $_{\text{COMP}}$  [V-N] $_{\text{N}|A}$ , a partire dalla fine del Trecento, istanzi forme che possono essere nomi o aggettivi, a seconda del contesto.

Dal punto di vista semantico, è interessante osservare la distribuzione dei composti rispetto ai tipi semantici di output, sia in prospettiva diacronica sia prendendo in considerazione l'intero dataset di composti [V-N]. In particolare, nella tabella 67 è illustrata la distribuzione dei composti [V-N] attestati in tutto il CoDIt\_com rispetto ai tipi semantici di output.

ANIMALE	ARTEFATTO	EVENTO	INFO	LUOGO	OGG_NAT	PARTE_CORPO	PERSONA	ТЕМРО	tot
5	81	48	6	6	8	2	166	2	324

Tabella 67 Distribuzione dei composti [V-N] (type) attestati nel CoDIt\_com rispetto al tipo semantico di output

Prendendo in considerazione l'intero CoDIt\_com (appiattendo quindi la dimensione diacronica), si può osservare che i composti [V-N] attestati si distibuiscono principalmente tra i tipi semantici di output PERSONA, ARTEFATTO e EVENTO. Piuttosto rara è la creazione di nomi che designano animali (ANIMALE), contenuti informativi (INFO), luoghi (LUOGO), parti del corpo (PARTE\_CORPO) e periodi di tempo (TEMPO): i composti che presentano questi tipi semantici di output sono riportati nella tabella 68.

ANIMALE	INFO	LUOGO	PARTE_CORPO	ТЕМРО
beccafico	cantafavola	battifolle	pappagorgia	coprifuoco
cangiacolore	cantafola	guardacamera 'anticamera'	pascipeco 'organo sessuale	fattifeste 'periodo di
forasiepi	contafavola	ammonu	maschile'	festa'

ANIMALE	INFO	LUOGO	PARTE_CORPO	ТЕМРО
fottivento	lasciapassare	guardacucina 'tinello'		
picchiamartelli	segnacaso			
	stuccalettori	guardaroba		
		marciapiede		
		salvaroba		
		'guardaroba'		

Tabella 68 Composti [V-N] attestati nel CoDIt\_com che presentano tipi semantici di output poco frequenti

Come mostrano i dati riportati nella tabella 68, la composizione verbonominale, seppur sfruttata prevalentemente per creare nomi che designano persone, artefatti ed eventi, permette di creare anche nomi con altri significati.

Dal punto di vista diacronico, è interessante focalizzare l'attenzione sui tre tipi più attestati e analizzarne la produttività nel corso dei secoli: la distribuzione dei composti [V-N] (type) che presentano i tipi semantici ARTEFATTO, PERSONA E EVENTO nei sottocorpora del CoDIt com è illustrata nella figura 52.

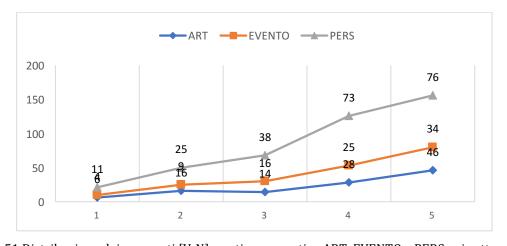


Figura 51 Distribuzione dei composti [V-N] con tipo semantico ART, EVENTO e PERS nei sotto-corpora del CoDIt\_com

Come illustrato dal grafico riportato nella figura 52, il numero di composti che presentano i tre tipi semantici più frequenti aumenta gradualmente, dalla fase antica a quella contemporanea. In particolare, un sensibile incremento è registrato dai [V-N] indicanti un referente umano dal periodo successivo al terzo sotto-corpus; gli altri due tipi semantici aumentano in termini di frequenza dei type in modo più graduale.

Tali risultati sembrerebbero non confermare l'analisi proposta da Štichauer (2015), secondo cui il generale aumento di composti [V-N] sarebbe legato soprattutto alla fortuna della semantica strumentale, testimoniata dall'ingresso nei dizionari dell'epoca di un numero considerevole di forme verbonominali indicanti strumenti. In realtà, l'apparente discordanza tra le due analisi può essere spiegata pensando al fatto che lo studio di Štichauer (2015) è basato su dati estratti prevalentemente da opere lessicografiche che, si può ragionevolmente ipotizzare, registrano con maggior probabilità forme che designano strumenti piuttosto che appellativi occasionali riferiti a persone; va inoltre considerato che nel lavoro di Štichauer (2015) vengono analizzate soltanto le basi verbali più frequenti.

L'analisi sembra quindi mettere in luce una crescente produttività dei tre tipi semantici di output, in linea con l'aumento di produttività registrato da tutta la Costruzione<sub>COMP</sub> e illustrato nella figura 52; va inoltre osservato che, accanto alla semantica agentiva e strumentale, la composizione verbonominale è costantemente sfruttata fino alla fase contemporanea anche per la creazione di nomi che indicano eventi (si pensi ad esempio a composto di formazione piuttosto recente come *firmacopie* con cui si indica l'incontro in cui un autore autografa le copie del suo libro ai lettori). Focalizzando l'attenzione proprio sulla fase più recente rappresentata dal CoDIt\_com, corrispondente al sottocorpus 5 (1841-1947), si può avere un'idea più chiara della produttività dei sotto-schemi individuati dai tipi semantici di output, analizzando come vi si distribuiscono gli hapax.<sup>182</sup>

ARTEFATTO	EVENTO	OGG_NAT	PERSONA	tot
20	8	2	25	55

Tabella 69 Distribuzione degli hapax rispetto ai tipi semantici di output nel sotto-corpus più recente (1841-1947)

I dati riportati nella tabella 69 mettono in luce che, dei tipi semantici attestati nei composti [V-N] nel corso della storia dell'italiano, i sotto-schemi produttivi nella fase recente sono prevalentemente quelli che presentano i tipi semantici ARTEFATTO e PERSONA

<sup>&</sup>lt;sup>182</sup> Il numero degli hapax è fondamentale per valutare la produttività di un sotto-schema, in quanto riflette in che misura esso è sfruttato dai parlanti per la creazione di forme occasionali, estemporanee (cfr. Ricca 2010: 250-51 sui composti VN dell'italiano contemporaneo).

e, in misura minore, EVENTO (battibaleno, giracapo, rompicapo, scaccianoia, serramanico, strippapelle, struggibuco, torcipiede). 183

In conclusione all'analisi diacronica della Costruzione COMP [V-N], si possono mettere in luce alcuni aspetti, che completano il quadro delineato per la fase antica, grazie ai dati estratti dal CoDIt com. *In primis*, si è osservato che questa Costruzione<sub>COMP</sub> è ampiamente sfruttata senza soluzione di continuità dalla fase antica e non costituisce quindi un'innovazione dell'età contemporanea; il graduale aumento del numero di type e di token testimonia una crescente produttività di questa Costruzione comp che permette di formare nomi che indicano prevalentemente persone, artefatti e eventi, ma anche oggetti naturali, luoghi, periodi di tempo: la varietà semantica è infatti un elemento che caratterizza costantemente questa Costruzione<sub>COMP</sub>, come già osservato da Štichauer (2015: 145). Infine, l'analisi dei composti [V-N] ha messo in luce che, a differenza di quanto si è osservato per altre Costruzioni<sub>COMP</sub> (ad es. i composti [N-N]<sub>N SUB</sub> con testa a sinistra o i composti [N-PART.PRES]<sub>NIA</sub> con testa a destra), in questo caso è l'intera Costruzione<sub>COMP</sub> a essere stabilmente produttiva nel corso dei secoli e non un singolo sotto-schema.

La Costruzione<sub>COMP</sub> [V-AVV]<sub>N MOD</sub>, non attestata in italiano antico, è istanziata da quattro forme, buttafuori (2 occ.), cacasotto(1 occ.), mettimale (3 occ.) e sputagrave (1 occ.), attestate a partire dal 4 sotto-corpus: si tratta quindi di una delle Costruzionicome di più recente attestazione. La classificazione di questo tipo di forme non è del tutto chiara: se infatti il secondo costituente è un avverbio che modifica il verbo, aspetto che porterebbe a classificarli tra gli esocentrici con R<sub>MOD</sub>, va senz'altro considerato il rapporto che li lega ai  $[V-N]_{N\_SUB}$ . Inoltre, nel caso di *buttafuori*, *cacasotto* e *mettimale* va notato che si tratta di verbi sintagmatici (buttare fuori, cacare sotto, mettere male).

La Costruzione<sub>COMP</sub> [N-A]<sub>N|A</sub>, produttiva fin dalla fase antica, istanzia composti esocentrici del tipo pettirosso, capinera; il CoDIt\_com attesta 28 composti esocentrici costituiti da un nome e un aggettivo: la distribuzione dei type e dei token nei sotto-corpora del CoDIt\_com è illustrata dai grafici nelle figure 53 e 54.

<sup>183</sup> Il tipo OGG\_NAT è rappresentato soltanto dalle due forme forasacchi e pisciacani e può quindi difficilmente essere considerato produttivo.

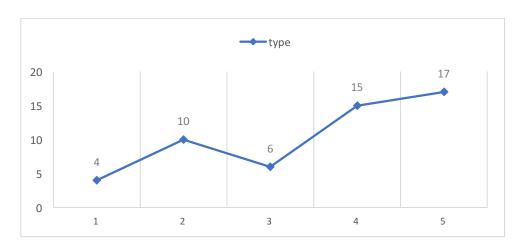


Figura 52 Costruzione  $_{\text{COMP}}$  dei composti [N-A] esocentrici: distribuzione dei type nei sotto-corpora del CoDIt\_com



Figura 53  $Costruzione_{COMP}$  dei composti [N-A] esocentrici: distribuzione dei token nei sotto-corpora del CoDIt\_com

Come mostrano i grafici nelle figure 53 e 54, la Costruzione<sub>COMP</sub> dei composti [N-A]<sub>N|A\_MOD</sub> esocentrici è stabilmente produttiva nel corso della storia dell'italiano, seppur con un andamento non costante: in particolare, il grafico rivela un picco in negativo nella frequenza dei type in corrispondenza del terzo sotto-corpus (1533-1691), a seguito del quale si registra un graduale aumento di produttività fino alla fase contemporanea; la frequenza dei token subisce un notevole aumento a partire dal quarto sotto-corpus. Tale aumento è dovuto all'aumento di frequenza nell'ultimo sotto-corpus di particolari composti, quali *capofitto* (48), *capinera* (23), *pellerossa* (14), *pettirosso* (17).  $^{184}$ 

 $<sup>^{\</sup>rm 184}$  I numeri tra parentesi si riferiscono alle occorrenze dei composti nel sotto-corpus 5.

Dal punto di vista semantico, i 28 composti [N-A]<sub>NIA MOD</sub> si distribuiscono tra sei tipi semantici di output: ANIMALE, PERSONA, ARTEFATTO EVENTO, LUOGO (tabella 70).

ANIMALE	ARTEFATTO	EVENTO	PERSONA
capirosso pettirosso capinera codilungo codirosso capinero capogrosso codimozzo	capomorto terrapieno	capochino caporiccio capofitto manomorta	bracciobianco chiomazzurro facciamara pellerossa ficamoscia capameno capomozzo cappanera paneperso occhiazzurro caposcarico cuorcontento cazzomatto

Tabella 70 Composti [N-A]<sub>N|A MOD</sub> esocentrici attestati nel CoDIt\_com: distribuzione rispetto al tipo semantico di output

Come già osservato per la fase antica (cfr. § 4.2.4), questa Costruzione<sub>COMP</sub> istanzia prevalentemente nomi che indicano referenti umani e animali, identificati sulla base di una caratteristica fisica (es. codirosso, pellerossa, cuorcontento, etc.). Queste forme rappresentano istanziazioni di un sotto-schema che presenta restrizioni semantiche sia rispetto all'output (ANIMALE PERSONA) sia rispetto al primo costituente, che indica sempre una parte del corpo (PARTE\_CORPO). 185 Meno produttiva è la formazione di nomi che designano eventi (es. capochino); residuale la presenza di nomi di artefatti (capomorto 'residuo, scoria inutile', dal latino caput mortuum, terrapieno 'opera costituita da una massa di terra accumulata secondo un progetto prestabilito, in modo da sostenersi da sola). 186 Focalizzando l'attenzione sul sotto-corpus più recente, è interessante notare che gli hapax sono soltanto tre: capogrosso (ANIMALE), caposcarico (PERSONA), cuorcontento (PERSONA): questo dato induce a ritenere che non si tratti di una Costruzione COMP particolarmente produttiva in italiano contemporaneo. Tra le ragioni della scarsa produttività di questo schema va senz'altro menzionato che si tratta di forme non sempre facilmente identificabili come composti e distinguibili da espressioni

<sup>185</sup> Si noti che tali restrizioni erano già presenti in italiano antico.

<sup>&</sup>lt;sup>186</sup> Dizionario Treccani Online, s.v. capomorto e terrapieno (ultima visualizzazione in data 23/07/18).

multiparola o sintagmi, dal momento l'ordine dei costituenti ricalca quello di norma attestato nella sintassi.

Il dataset *Compos\_ita* attesta 11 composti esocentrici che istanziano la Costruzione<sub>COMP</sub> [V-V]<sub>N\_COORD</sub> (vd. tabella 71). Si è precedentemente osservato che, in italiano antico (cfr. 4.2.4), tale Costruzione<sub>COMP</sub> [V-V]<sub>N</sub> istanziava composti costituiti da due basi verbali, sempre diverse e preferibilmente bisillabiche (es. *giravolta*, *battisoffia*); semanticamente, è stato notato che, a differenza dei composti [V-V]<sub>N</sub> attestati in italiano contemporaneo (il tipo *fuggifuggi*, cfr. Thornton 2010), queste forme non indicano un'azione svolta da un gruppo di persone. Secondo Thornton (2010: 527), l'origine di questi composti, caratteristici dell'italiano,<sup>187</sup> è delocutiva: «I sostantivi si sono formati per lessicalizzazione di citazioni di discorso diretto, usate in origine nella funzione che è stata denominata da Spitzer (1918; 1951-52) "imperativo descrittivo". Si è avuta cioè sostantivizzazione di enunciati comunemente pronunciati o pronunciabili nelle circostanze cui si riferisce il nome poi lessicalizzato». A conferma di quanto sostenuto da Thornton (2010), va notato che nel CoDIt\_com sono attestati discorsi diretti costituiti da due imperativi che occorrono anche all'interno di questi composti, come esemplificato in (41).

(41) a. «La calca grande andava de continuo costoro sequendo con gridi, cifolare e urlare; e in ogne luoco gridandosi: "Para! Piglia!", e cui loro sassi traendo [...]» (Masuccio Salernitano, Novellino)

b. «E vi fu alcun che gridò *serra*, *serra*, / Per disfar l'arte dell'indovinare; / Ma la Sibilla fece scongiurare / Lucifero nel centro della terra» (Burchiello, *Rime*)

c. «Pare come ne' drammi: *tira*, *tira*: c'è ancora due atti prima che vadano a letto, o prima che muoiano» (Tommaseo, *Fede e bellezza*)

Gli esempi in (25) testimoniano la giustapposizione di due imperativi in un discorso diretto che possono essere considerati antecedenti dei corrispettivi composti

<sup>&</sup>lt;sup>187</sup> Come segnala Thornton (*idem*), negli studi di riferimento relativi alle altre lingue romanze, questo tipo di forme, in particolare i composti costituiti dalla reduplicazione di una stessa base verbale, non è registrato (cfr. Guilbert 1971; Rainer 1993: 263).

lessicalizzati. Di seguito si riportano ed esemplificano i composti  $[V-V]_N$  attestati nel CoDIt\_com.

Composti [V-V] <sub>N</sub>	1	2	3	4	5	tot	Esempio
battisoffia 'paura, tremore'	-	1	-	-	-	1	(a) «Se tu me l'avessi detto, non che io ci fosse albergato, ma io sarei camminato più oltre parecchie miglia, se io dovessi essere rimaso nelle valli tra le cannucci; ché m'hai dato sì fatta battisoffia che io non sarò mai lieto, e forse me ne morrò» (Sacchetti, Trecentonovelle)
<u>andirivieni</u>	1	1	3	14	27	46	(b) «[] perché al tempo dell'imbandigione del pranzo non ci sia un <i>andirivieni</i> continuo deì cuochi, che trasportano le vivande aventi sottoposte le brage []» (Caro, <i>Lettere di Seneca</i> )
							(c) «I tedeschi si strisciano sempre intorno e appiedi alla verità; di rado l'afferrano con mano robusta: la seguono indefessamente per tutti gli andirivieni di questo laberinto della natura, []» (Leopardi, Zibaldone)
cantimplora 'vaso'	-	-	2	-	-	2	(d) «T'ama, ti brama, lacrima e sospira / Bacco, ebro e stracco accanto <i>cantimplora</i> , / E ti agogna, e qual spogna a sé ti tira» (Leporeo, <i>Leporeambi</i> )
							(e) «Prendi, o garzon, del gran grotton la chiave, / Tranne del vino che il pretino beve, / Metti la cantimplora or ora in neve / Con piselli novelli e fresche fave» (Leporeo, Leporeambi)

Composti [V-V] <sub>N</sub>	1	2	3	4	5	tot	Esempio
giravolta	2	-	9	5	8	24	(f) «Questi tre, quando videro tutti gli usci chiusi e la strada deserta, uscirono in fretta, come se si fossero avvisti d'aver fatto tardi, e dicendo di voler andar subito a casa; diedero una giravolta per il paese, per venire in chiaro se tutti eran ritirati» (Manzoni, Promessi sposi, 1840)
							(g) «Quel pensare per le strade mi dava la <i>giravolta</i> ; spesso più ripensavo e più mi si guastava il pensiero o la frase []» (De Sanctis, <i>La giovinezza</i> )
							(h) «Egli, invece, continuò ad andarle incontro; ma ella fece una giravolta, rasentandolo senza farsi toccare» (Tozzi, Con gli occhi chiusi)
parapiglia	-	1	2	5	12	20	(i) «V'era stato un <i>parapiglia</i> , avea sentito i pugnali scalcinar le mura intorno []» (D'Azeglio, <i>Racconti</i> )
serraserra	-	1	1	9	6	17	(l) «Alle grida s'aggiunsero gli urtoni; di maniera che, in poco tempo, parte con le gambe proprie, parte con le gomita altrui, ottenne ciò che più gli premeva in quel momento, d'esser fuori di quel serra serra» (Manzoni, Promessi sposi, 1840)
							(m) «M'ero addormentato con la cara imagine sulle labbra, e la mattina, nel serra serra del vestirmi in fretta sotto gli occhi grifagni dell'abataccio, non potei che nasconderla sotto alle lenzuola» (Guerrini, Brani di vita)
<u>saliscendi</u>	-	-	1	6	9	16	(n) «Accostatevi all'uscio, che tirerò la fune del <i>saliscendi</i> ; intendete? » (Guarini, <i>La idropica</i> )
							(o) «Non ci mettere il <i>saliscendi</i> all'uscio, perché aspetto Rocco Spatu

Composti [V-V] <sub>N</sub>	1	2	3	4	5	tot	Esempio
							che deve venire a parlarmi» (Verga, <i>I Malavoglia</i> )
							(p) «Non potresti immaginarti mai di trovare a pochi passi dall'Appennino un <i>saliscendi</i> di poggi più o meno alti, sparsi di paesi a uno a due e tre miglia di distanza l'uno dall'altro» (Giusti, <i>Epistolario</i> )
dormiveglia	-	-	-	-	15	15	(q) «[] ma se poi stampava fuori, o non era molestato, o la molestia si limitava a sequestrargli i libri stampati, e non era difficile eludere le dogane e la dormiveglia della polizia» (Giusti, Cronaca dei fatti di Toscana)
							(r) «Nel dormiveglia ricordo che il mio testo asserisce che con questo sistema si può arrivar a ricordare la prima infanzia, quella in fasce» (Svevo, La coscienza di Zeno)
grattapugia 'pennello per levigare il metallo'	-	1	-	-	-	1	(s) «Piglia le grattapugie o un pannuccio / ruvido, e netta bene ogni cantuccio []»(de' Medici, Rime in forma di ballata)
tiratira	-	-	1	3	6	10	(t) «Naturalmente a non voler fare una tragedia d'un atto, bisogna che per altri quattro sia un continuo fare a <i>tira tira</i> fra Enea e Didone []» (D'Azeglio, <i>Racconti</i> )
							(u) «Davvero, appena letta la sua lettera, mi s'è messa di qua la Superbia e di là la Modestia, e prendendomi una per un braccio e una per un altro, hanno fatto un gran pezzo a tira tira» (Giusti, Epistolario)
portareca 'messo'	1	-	-	-	-	1	(v) «Quante minacce ha' fatte con novelle, / per farti udire spesso al portareca, / acciò che l'altru' sangue torni in pelle!» (Sacchetti, Rime)

Tabella 71 Composti [V-V]<sub>N</sub> attestati nel CoDIt\_com

Come emerge osservando i dati riportati nella tabella 71, i composti  $[V-V]_N$  più antichi, che si sono conservati fino alla fase contemporanea senza soluzione di continuità, sono

andirivieni e giravolta. Dal punto di vista semantico, tali forme hanno subito alcuni mutamenti. Nel caso di andirivieni, il composto è attestato nella fase antica con il significato di 'movimento d'entrata e d'uscita attraverso uno spazio circoscritto; percorso tortuoso'<sup>188</sup> (cfr. 4.2.4). Nei secoli successivi, esso occorre anche nel significato di 'viavai di persone' come nell'esempio (b) in cui si fa riferimento a un viavai di cuochi in una cucina, pur conservando parallelamente l'antico significato, come testimoniato dall'esempio in (c). In italiano contemporaneo, il composto è descritto dai dizionari esclusivamente come sinonimo di 'viavai'. 189 Anche nel caso di giravolta, si notano cambiamenti sul piano semantico: in italiano antico, il composto è usato per indicare un 'rapido moto rotatorio' o una 'inversione di marcia di un esercito schierato';<sup>190</sup> il primo significato si conserva fino all'età contemporanea, come testimoniato dall'esempio (h). Gli esempi (f) e (g) mostrano però ulteriori usi di *giravolta* per indicare, rispettivamente, un giro per le strade di un paese e un sentimento di capogiro suscitato dal camminare per strade tortuose. Il composto saliscendi può indicare un artefatto per chiudere porte o finestre (significato attestato dalla fase antica a quella contemporanea; cfr. es. (n)-(o)) o un una serie continua di salite e discese. Forma piuttosto antica è anche parapiglia: esso indica fin dalle prime attestazioni un'agitazione di persone che si scontrano (cfr. es. (i)).

Le forme *battisoffia*, *cantimplora*, <sup>191</sup> *grattapugia* e *portareca* sono attestate nel corso dei secoli ma non nel periodo più recente corrispondente al sotto-corpus 5 del CoDIt\_com (1841-1947): tali composti sono costituiti da due verbi diversi e indicano referenti umani (*portareca*), artefatti (*grattapugia*, *cantimplora*) e stati (*battisoffia*). Nessuno di questi esprime quindi un'azione svolta da un gruppo di persone: tale significato è espresso, oltre che da *andirivieni* e *parapiglia*, da *serraserra* (cfr. es. (l)), attestato anche con il significato di 'rapide azioni svolte di fretta' (cfr. es. (m)).

Infine, il composto *dormiveglia*, attestato soltanto nel sotto-corpus più recente, è utilizzato per indicare uno stato tra il sonno e la veglia; va inoltre notata l'oscillazione

\_

<sup>&</sup>lt;sup>188</sup> TLIO, s.v. andirivieni.

<sup>&</sup>lt;sup>189</sup> Dizionario Treccani Online e De Mauro (2004), s.v. andirivieni.

<sup>&</sup>lt;sup>190</sup> TLIO, s.v. *giravolta*.

<sup>&</sup>lt;sup>191</sup> Si noti che tale composto è utilizzato anche come verbo, come nell'es. (d) (unico caso attestato nel CoDIt\_com), con il significato di 'cantare e implorare': in questo caso, quindi non si tratta di un composto esocentrico ma di un coordinativo endocentrico.

nell'assegnazione del genere tra maschile (r), impostosi nell'uso attuale, e femminile (q). 192

L'analisi dei composti  $[V-V]_N$  attestati nel CoDIt\_com mette in luce che la Costruzione<sub>COMP</sub> istanzia un numero limitato di forme, alcune delle quali si perdono nel corso dei secoli. Dal punto di vista formale, a partire da sotto-corpus 4 (periodo 1692-1840) iniziano a diffondersi composti costituiti da due basi verbali uguali,<sup>193</sup> come *serraserra* o *tiratira*. Va notato che queste forme, date dalla reduplicazione di una base verbale, possono essere classificate come 'coordinative' soltanto dal punto di vista grammaticale: il loro significato, infatti, non coincide con la parafrasi 'X *e* X' (es. *serra* e *serra*), ma veicola intensificazione (o plurazionalità) rispetto all'azione espressa dal verbo.

#### 4.4 Discussione

Analogamente a quanto proposto per la fase antica, di seguito si fornisce un riepilogo dell'analisi diacronica. Nella tabella 72 si illustrano le famiglie di Costruzioni<sub>COMP</sub> attestate nel CoDIt\_com, per ciascuna delle quali viene riportata la distribuzione dei type e dei token nei diversi sotto-schemi.

Costruzioni <sub>COMP</sub> con testa a sinistra			Costruzioni <sub>comp</sub> con testa a destra			
	type	token		type	token	
$[N-N]_{N\_SUB}$	42	533	$[N-N]_{N\_SUB}$	19	390	
$[N-N]_{N\_MOD}$	17	216	[N-PART.PRES] <sub>N A_SUB</sub>	10	226	
$[N-A]_{N\_MOD}$	48	1243	[N-PART.PASS] <sub>N A_SUB</sub>	-	-	
$[A-N]_{A N\_{MOD}}$	5	68	$[N-V]_V$	2	83	
$[A-A]_{A N\_MOD}$	8	52	$[A-N]_{N A AVV\_MOD}$	714	22.002	
			$[A-PART.PASS]_{N A\_MOD}$	5	561	
			$[AVV-X]_{X\_MOD}$	770	9648	
			$[N-N]_{N\_{MOD}}$	8	376	
tot	120	2.112		814	33.286	

Costruzioni <sub>COMP</sub> coordinative endocentriche			Costruzioni <sub>COMP</sub> esocen	triche	
	type	token		type	token
[N-N] <sub>N_COORD</sub>	9	216	$[V-N]_{N A\_SUB}$	324	1884

1

<sup>&</sup>lt;sup>192</sup> L'assegnazione del genere femminile è probabilmente dovuta al fatto che il secondo costituente *veglia* può esser stato analizzato come sostantivo piuttosto che come verbo.

<sup>&</sup>lt;sup>193</sup> Va però segnalato che Thornton (2010: 527) colloca la diffusione di questo tipo di composti a partire dal XVI secolo (cfr. *bolli bolli*, 1565; *corri corri*, av. 1587).

Costruzioni <sub>COMP</sub> coordinative endocentriche			Costruzioni <sub>comp</sub> esocen	ntriche	
[A-A] <sub>N A_COORD</sub>	22	224	[N-A] <sub>N A_MOD</sub> [V-V] <sub>N_COORD</sub> [V-AVV] <sub>N</sub>	28 11 4	289 153 7
tot	31	440	tot	367	2.333

Tabella 72 Costruzioni<sub>COMP</sub> attestate nel CoDIt\_com: riepilogo generale

Un primo aspetto che si può notare osservando i dati riportati nella tabella 72 è che le categorie lessicali non si distribuiscono in maniera omogenea tra primo e secondo costituente nelle diverse famiglie di Costruzioni<sub>COMP</sub>: in particolare, nei composti con testa a sinistra, la prima posizione è occupata sempre da un nome o da un aggettivo;<sup>194</sup> nei composti con testa a destra, la seconda posizione può essere un nome, un aggettivo, un verbo, un participio. La struttura con testa a destra permette quindi una maggior varietà in termini di categorie lessicali che possono costituire la testa del composto.

Provando a tracciare un quadro generale sulla base dell'analisi diacronica, si può *in primis* osservare la sostanziale continuità tra fase antica e fasi successive: gran parte delle Costruzioni<sub>COMP</sub> istanziate in italiano antico trova riscontro nel CoDIt\_com, seppur, in alcuni casi, con diversi livelli di produttività. In particolare, nel caso delle Costruzioni<sub>COMP</sub> con testa a sinistra, si è osservato un aumento di produttività dello schema [N-N]<sub>N\_MOD</sub> (es. *pescecane*), scarsamente attestato in italiano antico, ma istanziato da un considerevole numero di forme nei secoli successivi e destinato ad avere grande diffusione nella fase recente (cfr. composti come *riunione fiume*, *notizia bomba*, *parola chiave*, non attestati nel CoDIt\_com).

Tra le Costruzioni $_{COMP}$  con testa a destra, alcuni sotto-schemi (in particolare, i sotto-schemi [N-PART.PRES] $_{N|A\_SUB}$ , [A-PART.PASS] $_{N|A\_MOD}$  e [N-V] $_{V\_SUB}$ ) sono istanziati quasi esclusivamente da composti di antica attestazione o di origine latina e possono quindi essere considerati scarsamente produttivi nei secoli presi in esame; tuttavia, l'analisi della Costruzione $_{COMP}$  [N-PART.PRES] $_{N|A\_SUB}$  (es. *nullatenente*) ha mostrato come uno schema possa tornare a essere produttivo dopo un periodo di 'inattività', grazie alla fortuna di un sotto-schema semi-specificato: in questo caso, quello dei composti [x-

-

<sup>&</sup>lt;sup>194</sup> Va inoltre notato che i composti con testa a sinistra sono costituiti esclusivamente da nomi e aggettivi; maggior varietà in termini di categorie lessicali è invece caratteristica dei composti con testa a destra.

dipendente]<sub>PERS</sub>. L'unico schema che invece non risulta essere istanziato da nessuna forma è [N-PART.PASS]<sub>N|A\_SUB</sub> (es. occhibagliato, plebeuscito).

Si è inoltre osservato che la notevole produttività degli schemi [A-N]<sub>N</sub> e [AVV-X]<sub>X</sub> dipende strettamente da quella di alcuni sotto-schemi (*in primis*, le Costruzioni<sub>COMP</sub> [mal(e)-x] e [mal(o/a)-x]); questa situazione è riscontrabile anche nella famiglia delle Costruzioni<sub>COMP</sub> con testa a sinistra, in cui lo schema [N-N]<sub>N\_SUB</sub> deve la sua produttività *in primis* al sotto-schema semi-specificato [ $capo_{PERSONA}$ -x]<sub>PERSONA</sub>. D'altra parte, l'analisi dello schema [V-N]<sub>N|A</sub> ha permesso di osservare un caso di Costruzione<sub>COMP</sub> produttiva nel suo insieme e non per merito di un sotto-schema semi-specificato. In particolare, i composti [V-N]<sub>N|A</sub> costituiscono la Costruzione<sub>COMP</sub> astratta più produttiva e stabile nel corso della storia dell'italiano, ampiamente sfruttata dalla fase antica a quella contemporanea per la creazione di nomi (o aggettivi) che prevalentemente designano individui, artefatti ed eventi.

Si pongono in netta continuità con l'italiano antico anche le altre Costruzioni $_{COMP}$  esocentriche ([N-A] $_{N|A\_SUB}$  e [V-V] $_{N\_COORD}$ ) e quelle coordinative endocentriche ([N-N] $_{N\_COORD}$  e [A-A] $_{N|A\_COORD}$ ), che nella fase successiva a quella antica continuano a istanziare un numero limitato ma costante di composti, nella maggior parte dei casi conservando la semantica presentata dalle forme più antiche (es. lo schema [N-A] $_{N|A\_SUB}$  continua a istanziare prevalentemente nomi di specie di piante o animali e alimenti, lo schema [A-A] $_{N|A\_COORD}$  aggettivi/nomi che indicano la provenienza geografica, ma anche la compresenza di due proprietà).

Infine, l'analisi diacronica ha messo in luce l'emergere di quattro nuovi schemi: in particolare, si tratta degli schemi  $[A-N]_{A|N\_MOD}$  (es. *giallo canarino*),  $[A-A]_{A|N\_MOD}$  (es. *rossobruno*), appartenenti alla famiglia delle Costruzioni<sub>COMP</sub> con testa a sinistra e semanticamente ristretti a termini di colore; dello schema con testa a destra  $[N-N]_{N\_MOD}$ , istanziato esclusivamente da forme che presentano *capo* come primo costituente in funzione di modificatore (es. *capoluogo*); dello schema esocentrico  $[N-AVV]_N$  (es. *buttafuori*).

Dall'analisi diacronica emerge il quadro di una composizione che si pone in forte continuità con la fase antica: le famiglie di Costruzioni<sub>COMP</sub> attestate in italiano antico si conservano grossomodo nel corso dei secoli ma con diversi livelli di produttività. In particolare, le Costruzioni<sub>COMP</sub> in cui si concentra la maggior produttività sono quella dei

composti [V-N]<sub>N|A</sub> e alcuni sotto-schemi semi-specificati ([ $capo_{PERSONA}$ -x]<sub>PERSONA</sub>, [mal(e)-x], [mal(o/a)-x], [ben(e)-x], etc). I dati estratti dal CoDIt\_com confermano quindi che la composizione dell'italiano è costituita essenzialmente da composti esocentrici [V-N]<sub>N|A</sub> e composti con testa a destra. Sulle ragioni che motivano la preferenza dell'italiano per questi tipi di composti si proporrà qualche riflessione nel capitolo conclusivo, alla luce dell'analisi diacronica delle Costruzioni<sub>COMP</sub> peggiorative [mal(e)-x] e [mal(o/a)-x], a cui è dedicato il prossimo capitolo, che integrerà il quadro sulla composizione dell'italiano.

# 5. LE COSTRUZIONI VALUTATIVE [MAL(E)-X] E [MAL(O/A)-X]

In questo capitolo vengono analizzati due casi di Costruzioni<sub>COMP</sub> semi-specificate di cui si è osservata la rilevanza nel corso dell'analisi (§ 4.2.2 e 4.3.2): in particolare, si tratta delle Costruzioni<sub>COMP</sub> [mal(e)-x] e [mal(o/a)-x]. Adottando i parametri discussi al § 5.2, entrambe le Costruzioni<sub>COMP</sub> vengono analizzate dalla fase antica a quella contemporanea (§ 5.3 e sgg), mettendone in luce i mutamenti formali e semantici subiti nel corso dei secoli.

#### 5.1 Introduzione

Le CostruzionicomP con mal(e)- e mal(o/a)- permettono di analizzare la formazione di affissoidi/affissi a partire da costituenti di composto in italiano: come già osservato al § 2.2.5, tale mutamento, ampiamente descritto nelle lingue germaniche, è stato finora scarsamente trattato nelle lingue romanze. Inoltre, i due casi discussi in questo capitolo consentono di colmare una lacuna nell'ambito degli studi dedicati alla morfologia valutativa dell'italiano, finora concentratisi principalmente sulla valutazione quantitativa, piuttosto che su quella qualitativa. La ragione di questa disparità va probabilmente ricercata nella preminenza della valutazione descrittiva su quella valutativa: Grandi (2002: 85) osserva infatti che le strategie morfologiche attuate per esprimere i valori BIG e SMALL possono essere reimpiegate, attraverso estensioni semantiche, per esprimere anche i valori GOOD e BAD, ma non viceversa. L'univocità delle estensioni semantiche, possibili solo nella direzione marche descrittive > marche valutative (es. SMALL > BAD: it. attricetta, dottorino, etc.), sarebbe giustificata dal fatto che «con l'esclusione dei pochi suffissi peggiorativi e dei prefissi accrescitivomigliorativi, la polarità GOOD/BAD non dispone di specifiche marche linguistiche davvero produttive» (Grandi 2002: 86). Inoltre, è stato osservato che la strategia morfologica più ampiamente sfruttata per esprimere i valori BIG/SMALL è la suffissazione, che di conseguenza è la principale candidata a esprimere anche i valori qualitativi BAD/GOOD.

In italiano, il valore BAD è espresso mediante suffissi specifici di origine latina (es. –astro, -accio, -ucolo, etc.) o suffissi originariamente accrescitivi (ad es. -one in mangione): secondo Grandi (2002), infatti, per esprimere la valutazione qualitativa, l'italiano

utilizzerebbe quasi esclusivamente la suffissazione, limitando la prefissazione al solo valore GOOD (ad es. con *super-, extra-, arci-,* etc.). In particolare, come già messo in luce in Grandi & Montermini (2003: 281), «la classe dei prefissi valutativi è probabilmente quella più instabile nella storia dell'italiano», nonostante si tratti di una strategia disponibile e sfruttata dalla lingua, soprattutto nelle precedenti fasi evolutive. Il caso delle Costruzioni $_{\text{COMP}}$  [mal(e)-x] e [mal(o/a)-x] permette quindi di arricchire il quadro sull'uso della prefissazione nella valutazione, analizzando la nascita di un prefisso peggiorativo a partire da un costituente di composto.

Come già discusso al § 2.6, la formazione di affissoidi costituisce l'esito di un processo di grammaticalizzazione/costruzionalizzazione che interessa unità lessicali originariamente libere, che assumono progressivamente lo status di elementi legati. Nei seguenti paragrafi si analizzerà come tale processo, nel corso della storia dell'italiano, abbia interessato in gradi diversi l'avverbio male e l'aggettivo malo, quando posti all'interno di una Costruzione di natura morfologica. In particolare, si osserverà come mal(e)-, da modificatore di composti con testa a destra, abbia acquisito proprietà formali e semantiche tipiche di un prefisso, mentre mal(o/a)-, progressivamente uscito dall'uso come forma libera, si sia conservato all'interno della Costruzionecomp semi-specificata [mal(o/a)-x] e dei suoi sotto-schemi.

## 5.2 Parametri di analisi e metodologia di indagine

Al fine di valutare se, e in che misura, *male* e *malo* abbiano subito un processo di grammaticalizzazione/costruzionalizzazione nel corso della loro storia, si è scelto di adottare i quattro parametri proposti da Van Goethem (2010) per lo studio della formazione di affissoidi in olandese e francese. Tali parametri si basano su numerosi studi dedicati ai processi di grammaticalizzazione (cfr., tra gli altri, Hopper 1991; Hopper & Traugott 1993; Lehmann 1995; Ten Hacken 2000; Marchello-Nizia 2006) e sono di natura semantica, fono-morfologica e quantitativa.

Il primo parametro è la *specializzazione semantica* (o *risemantizzazione*): lo sviluppo di un nuovo, o più ristretto, significato è infatti indizio della grammaticalizzazione di un'entità lessicale autonoma in elemento legato, il cui significato può essere più generico o rappresentare un'estensione metaforica di quello presentato dalla forma libera. In particolare, gli studi dedicati agli affissoidi nelle lingue germaniche (cfr., tra gli altri, Booij

& Hüning 2014; Hüning & Booij 2014) hanno messo in luce che, in molti casi, essi costituiscono il risultato di un processo di *subjectification*, ossia «the development of a grammatically identifiable expression of speaker belief or attitude to what is said» (Traugott 1995: 32). Un caso ben noto è quello del tedesco *reuze-* 'gigante' che, quando compare come primo costituente di un composto, assume la funzione di intensificatore (es. *reuzeklein*, lett. 'gigante-piccolo', 'molto piccolo').

Il secondo parametro è la *produttività*: un aumento di produttività, infatti, caratterizza tipicamente un processo di grammaticalizzazione. Per valutare la produttività morfologica sono state proposte diverse misure (Bauer 2001; Baayen & Lieber 1991; Baayen 1992, 1993; Baayen & Renouf 1996; Baayen 2009), discusse da Hilpert (2013: 132) in relazione allo studio del mutamento in chiave costruzionista da un punto di vista diacronico. Seguendo Hilpert, in questa analisi si terranno in considerazione due misure: la *produttività realizzata* (ossia la frequenza dei type) e la *produttività potenziale* (Baayen 2005: 244), data dal rapporto tra il numero degli hapax e quello dei token di una Costruzione.

Il terzo parametro è la *decategorializzazione* (Hopper & Traugott 1993: 103-13), ossia la perdita di proprietà morfologiche o sintattiche da parte di un elemento che subisce un processo di grammaticalizzazione; un esempio è l'apocope del morfema flessivo in sostantivi e aggettivi che stanno assumendo lo status di prefissi (ad es. *nouveau*- in francese, cfr. Van Goethem 2010).

L'ultimo parametro è la *paradigmaticizzazione* (Lehmann 1995: 135): gli affissoidi si trovano infatti in competizione, o in distribuzione complementare, con reali affissi con cui condividono la semantica (ad es. il prefissoide olandese *oud-* si alterna al prefisso *ex-*, con cui condivide la semantica di 'ex, precedente', cfr. Van Goethem 2010: 165).

Tali parametri verranno applicati ai composti con mal(e)- e mal(o/a)- in prospettiva diacronica: dopo una breve descrizione dei composti con male- in latino, ci si concentrerà sull'italiano antico e sulle successive fasi evolutive, fino all'italiano contemporaneo.

I dati su cui si basa l'analisi delle Costruzioni<sub>COMP</sub> con *mal(e)*- e *mal(o/a)*- provengono da fonti diverse. Per il latino, sono stati consultati il lessico derivazionale *Word Formation Latin* (cfr. § 3.2.1; Litta, Passarotti & Culy 2016) e il corpus del *Packard Humanities* 

Institute (PHI), interrogato attraverso il software *Diogenes*;<sup>195</sup> per l'italiano dalle Origini al 1947, è stato utilizzato il corpus CoDIt\_com (cfr. § 3.1); per l'italiano contemporaneo, i dati sono stati estratti dal corpus itWaC (Baroni *et al.* 2009) e, in alcuni casi, dall'archivio del quotidiano *La Stampa*.<sup>196</sup> La procedura di estrazione delle forme dai corpora ha seguito le seguenti fasi.<sup>197</sup> *In primis*, sono state estratte tutte le forme inizianti con la stringa *mal*-; l'elenco così ottenuto è stato ripulito manualmente, raccogliendo soltanto le forme in cui *mal*- coincide con *malo/mala* o *male* e costituisce il primo costituente di un composto. Successivamente, sono stati estratti tutti i bigrammi delle forme *male* e *mal(o-a-i)*, al fine di raccogliere un dataset che comprendesse i composti nelle due varianti grafiche in cui possono comparire nei testi, ossia nella forma unita e con i due costituenti separati dallo spazio. Il dataset totale su cui si basa l'analisi oggetto dei seguenti paragrafi è costituito da 67.375 token e 935 type della Costruzione<sub>COMP</sub> [*mal(e)*-x] e 111.978 token e 742 type della Costruzione [*mal(o/a)*-x].

## 5.3 La Costruzione<sub>COMP</sub> [mal(e)-x]

La Costruzione<sub>COMP</sub> [mal(e)-x] è costituita dalla stringa mal(e)- (dall'avverbio male), più una base variabile che può essere un verbo (es. maltrattare), un participio presente (es. malconveniente) o passato (es. malnato), un aggettivo (es. malsano) o un avverbio (es. malvolentieri). In italiano contemporaneo, tali parole (quando lemmatizzate) sono di norma considerate dai dizionari come composti e possono occorrere con i due costituenti separati dallo spazio o uniti graficamente.

Nei seguenti paragrafi se ne propone un'analisi diacronica, dal latino all'italiano contemporaneo, secondo i parametri discussi al § 5.2, da un punto di vista sia qualitativo che quantitativo.

<sup>&</sup>lt;sup>195</sup> Il software è scaricabile al seguente indirizzo: <a href="https://community.dur.ac.uk/p.j.heslin/Software/Diogenes/">https://community.dur.ac.uk/p.j.heslin/Software/Diogenes/</a>.

<sup>&</sup>lt;sup>196</sup> L'archivio, consultabile liberamente al sito <u>www.archiviolastampa.it</u>, è costituito da oltre 12 milioni di articoli dal 1867 al 2005.

<sup>&</sup>lt;sup>197</sup> La stessa procedura è stata seguita per il corpus usato per l'italiano contemporaneo, itWaC, e per ciascun sotto-corpus del CoDIt\_com.

### 5.3.1 Dal latino all'italiano

L'avverbio italiano *male* deriva dal latino *male* 'male, sfavorevolmente, malvagiamente, eccessivamente' (cfr. Glare 1982). Già nel latino classico, *male* si trova attestato come primo costituente di forme tradizionalmente considerate come parole composte dalle fonti lessicografiche e dagli studi teorici dedicati alla composizione latina (cfr. ad es. Brucale 2012). Tali composti, quantitativamente non molto numerosi, <sup>198</sup> sono costituiti dall'avverbio *male* più un verbo (42a), un participio (42b), un aggettivo (42c) o un nome (42d).

```
(42) a. malefacio
'nuocere' (lett. 'che agisce in vista del male di qualcuno')
b. malevolens
'malintenzionato, maligno' (lett. 'che vuole il male di qualcuno')
c. malefidus
'infido' (lett. 'non degno di fiducia')
d. malefaber
'subdolo, ingannevole' (lett. 'artefice del male')
```

Quando occorre all'interno di un composto, *male* presenta un significato più ristretto rispetto a quando è utilizzato come forma libera: in particolare, in (42a), (42b) e (42d), esso modifica il secondo costituente nel senso di 'malvagiamente', mentre in combinazione con un aggettivo (42c) sembra avere funzione negativa e veicolare un significato molto vicino a 'poco, non'. Tale interpretazione sembra essere confermata da due glosse, riportate in (43), nelle quali il grammatico Mauro Servio Onorato osserva che *male* può significare a volte 'poco', a volte 'per nulla'. 199

\_

<sup>&</sup>lt;sup>198</sup> In particolare, la risorsa *Word Formation Latin* contiene 12 forme, classificate come parole composte e costituite dall'avverbio *male* più un secondo costituente: *maledicax, maledico, malefaber, malefacio,* maleficio, *maleficus, malefidus, malesuadium, malesuadus, maletractatio, malevolens, malevolus*.

<sup>&</sup>lt;sup>199</sup> Gli esempi in (2) sono stati estratti dal corpus di lingua latina, *Thesaurus Linguae Latinae*, interrogato attraverso il software Diogenes.

(43) *«male sana* non plene sana, amore vitiata; 'male' enim plerumque 'non', plerumque 'minus' significat, sicut 've', ut 'vecors' et 'vesanus'»<sup>200</sup>

male sana significa 'non completamente sana', 'turbata dall'amore'; infatti, male a volte significa 'non' a volte 'poco', così come ve in vecors 'folle' e vesanus 'insano'.

b. *«male pinguis* non pinguis, id est sterilis: nam 'male' apud Vergilium saepe pro 'non' significat, ut "male sana sororem"» <sup>201</sup>

*male pinguis* vuol dire 'non fertile', ossia sterile: infatti, *male* in Virgilio spesso significa 'non', come in *male sana sororem*.

Sebbene si tratti ancora di un numero limitato di casi, *male* sembra aver subito già in latino uno slittamento semantico: quando si combina a verbi e participi, esso seleziona un significato più specifico rispetto a quelli che veicola come forma libera; con gli aggettivi ha il valore più astratto di 'poco' o 'non'.

### 5.3.2 Italiano antico

In italiano antico, *mal(e)* può combinarsi con verbi (44a), participi presenti (44b) e passati (44c), aggettivi (44d) e avverbi (44e), come mostrano i seguenti esempi:<sup>202</sup>

(44) a. «[...] io no(n) vidi ancho filiuoli che *maltractasseno* li padri et le madre, deli quali Dio non facesse vendecta» (Anonimo, *Trattati di Albertano da Brescia volgarizzati*)

b. «e guarda ben la *mal tolta* moneta ch'esser ti fece contra Carlo [...]» (Dante, *Commedia*)

c. «A cui Guglielmo, udendo il suo *mal conveniente* parlare, rispose: [...]» (Boccaccio, *Decameron*, I, 8)

d. «Onde i buoni cittadini popolani erano *malcontenti*, e biasimavano l'ufficio de' Priori [...]» (Dino Compagni, *Cronica*, 1310-12, cap. 5)

<sup>&</sup>lt;sup>200</sup> Maurus Servius Honoratus, Vergilii Aeneidos Libros (Servii Grammatici Qui Feruntur in Vergilii Carmina Commentarii. Vols. 1–2, ed. G. Thilo, 1878–1884).

<sup>&</sup>lt;sup>201</sup> Maurus Servius Honoratus, *Vergilii Georgicon Libros* (Servii Grammatici Qui Feruntur in Vergilii Bucolica et Georgica Commentarii. Vol. 3:1, ed. G. Thilo, 1887).

<sup>&</sup>lt;sup>202</sup> Tutti gli esempi sono tratti dal primo sottocorpus del CoDIt\_com.

e. «Or ecco dunque, come Dio volentieri perdona e *malvolentieri* punisce». (Cavalca, *Esp. simbolo*, 1342)

Quando si combina con verbi o participi passati, come in (44ab), mal(e) esprime una valutazione peggiorativa da parte del parlante rispetto a un evento o uno stato, svolgendo quindi una funzione dispregiativa. Due forme fanno eccezione: 1) maledire, dal latino maledicere, che presenta un significato lessicalizzato e non composizionale (45a); 2) malmenare, prestito dal provenzale malmenar, in cui mal- mostra piuttosto una funzione intensiva (45b).

(45) a. «Questo Papa *maledisse* e scomunicò Arrigo terzo imperatore in uno concilio di vescovi». (*Cronica fiorentina* XIII)

b. «Con grande volontà si mise a ferire tra la troiana gente; quanti ne truova, tutti li *malmena*». (Armannino giudice da Bologna, *Fiorita*, 1325)

A questo stadio, sono tuttavia ancora possibili usi legati e non legati di *mal(e)*- in combinazione con un verbo, che convivono dal punto di vista sincronico: ad esempio, mentre in (45a) *mal(e)*- presenta delle proprietà tipiche delle forme legate (ad es. grafia univerbata, perdita della vocale finale), in altri casi, come in (46), la distinzione tra *mal(e)*- come costituente di una Costruzione morfologica e *mal(e)* come avverbio libero non è sempre chiara, dal momento che in italiano antico (diversamente da quanto accade nella fase contemporanea) la posizione postverbale non era l'unica disponibile per l'avverbio *male*.

(46) «Dice Lucano che quine aveva barbari che *male sapevan* regiare in battaglia». (Anonimo, *Fatti di Cesare*, XIV cent.)

Quando invece mal(e)- si combina con participi presenti, aggettivi e avverbi, esso si comporta in modo simile a un prefisso negativo e veicola il significato di 'non' o 'poco' a seconda della semantica della base e del contesto, come si può notare dagli esempi (44c), in cui malconveniente si riferisce a un comportamento inappropriato, sconveniente, o (44d), in cui malcontenti sta per 'scontenti'; infine, anche in malvolentieri, mal(e)- assolve la funzione di negare la semantica dell'avverbio volentieri. Tra i participi presenti, due forme fanno eccezione perché conservano il significato mostrato dai composti latini: sia

*maldicente* (47a) sia *malparlante* (47b), infatti, si riferiscono a una persona che dice o parla malvagiamente.

- (47) a. «Maggiore uccisione commette la lingua del *maldicente*, che veruno coltello tagliente». (Matteo Corsini, 1373)
  - b. «Or priego qui ciascun fedele amante / che sia in questo mia difesa e scudo / contro a ogni invidioso e *mal parlante* [...]» (Boccaccio, *Ninfale fiesolano*)

Riassumendo, i dati estratti dal sottocorpus di italiano antico del CoDIt\_com mostrano che *mal(e)*- è usato per esprimere una valutazione peggiorativa rispetto a un verbo o a un participio passato, mentre ha valore negativo con aggettivi, avverbi e (anche se con alcune eccezioni) participi presenti.

Nel quadro della CxM, i composti con *mal(e)*- possono essere interpretati come istanziazioni di una Costruzione morfologica che sviluppa una gerarchia di eredità, come mostrato nella figura 55 (vd. oltre).

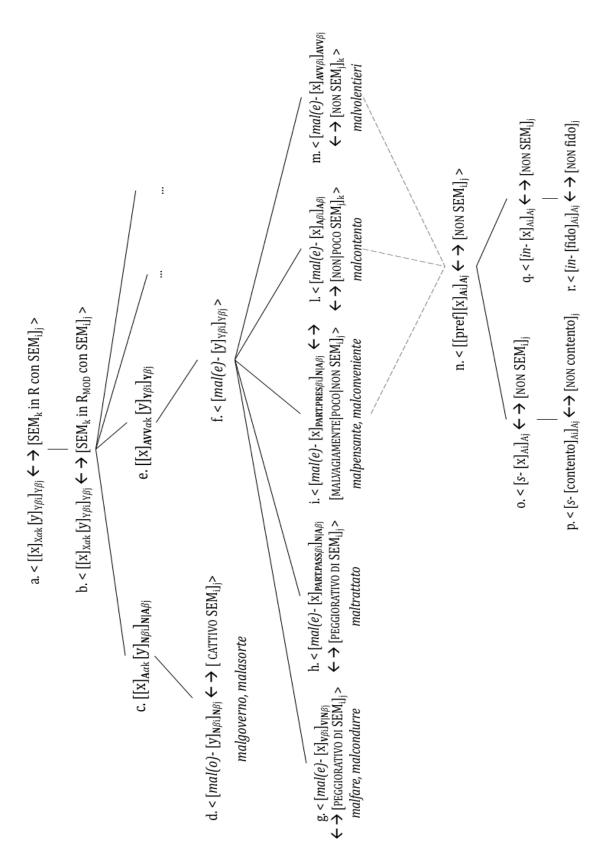


Figura 54 Gerarchia di eredità della Costruzione [mal(e)-x] in italiano antico

La Costruzione<sub>COMP</sub> semi-specificata in (f) è uno schema produttivo in cui una posizione è riempita da una specifica unità lessicale (in questo caso, mal(e)) e l'altra è variabile. Tale Costruzione<sub>COMP</sub> eredita le sue proprietà dalla Costruzione<sub>COMP</sub> dei composti italiani con testa a destra e primo elemento avverbiale (e); questa, fa parte delle istanziazioni della famiglia di Costruzioni<sub>COMP</sub> con testa a destra, al pari dello schema costituito da un aggettivo e un nome (d). Come illustrato nella figura 55, la Costruzione<sub>COMP</sub> [A-N] istanzia il sotto-schema semi-specificato [mal(o/a)-x] di cui si parlerà al § 5.4.

Lo schema semi-specificato con mal(e)- è istanziato da cinque sotto-schemi produttivi (rappresentati ai punti (g-m)), che si differenziano rispetto alla categoria della base (es. [mal(e)-VERBO], [mal(e)-AGG], [mal(e)-AVV], etc.). Ad eccezione dei primi due sotto-schemi ([mal(e)-VERBO] e [mal(e)-PART.PASS]), che esprimono una valutazione peggiorativa verso un evento o uno stato veicolato dalla base, tutti gli altri sono legati da una relazione paradigmatica alla Costruzione in (n) dei derivati attraverso un prefisso negativo.

Dal punto di vista quantitativo, i dati estratti dal primo sotto-corpus del CoDIt\_com permettono di analizzare la produttività della Costruzione<sub>COMP</sub> [mal(e)-x] e dei suoi sotto-schemi in italiano antico. Nella tabella 73 sono riportati i dati relativi alla frequenza dei type e dei token e agli hapax, per ciascun sotto-schema.

mal(e)-x	mal(e)+	mal(e)+	mal(e)+	mal(e)+	mal(e)+	Totale
	AGG	PART.PRES	PART.PRES	VERBO	AVVERBIO	
Type	24	8	99	143	5	279
Token	234	34	603	625	45	1541
Hapaxes	13	4	47	66	4	134
$\overline{P}$	0,06	0,10	0,08	0,10	0,10	0,10

Tabella 73. Frequenza type/token, hapax e produttività potenziale di ciascun sotto-schema della Costruzione<sub>COMP</sub> [mal(e)-x]

*In primis*, va osservato che i sotto-schemi più produttivi, dal punto di vista della *produttività realizzata*, sono [mal(e)-VERBO] e [mal(e)-PART.PASS]: essi mostrano inoltre un alto numero di hapax, indizio del fatto che già in questa fase erano utilizzati dai parlanti per creare nuove parole.

Nel complesso, la Costruzione [mal(e)-x] ha un valore di produttività potenziale (P) pari a 0,10 e può quindi dirsi produttiva in italiano antico, seppur con gradi diversi a

seconda del sotto-schema: le Costruzioni $_{COMP}$  [mal(e)-AGG] e [mal(e)-PART.PRES], ad esempio, risultano meno sfruttate delle altre.

Rispetto al parametro della *decategorializzazione*, nel caso di *mal(e)*- non è possibile parlare di perdita della flessione, dal momento che si tratta di una forma invariabile; tuttavia, va osservato che in circa la metà delle occorrenze (il 54% dei token), esso occorre nella forma apocopata senza vocale finale e, in un numero considerevole di casi (circa il 41% dei token), graficamente unito al secondo costituente. La distribuzione delle occorrenze dei cinque sotto-schemi in rapporto alla variante grafica è mostrata nella tabella 74.

	mal(e)-	mal(e)-	mal(e)-	mal(e)-	mal(e)-	Totale
	AGG	PART.PRES	PART.PASS	VERBO	AVVERBIO	
Separati	59	16	301	499	31	906
Uniti	175	18	302	126	14	635
Totale (token)	234	34	603	625	45	1541

Tabella 74 Distribuzione dei cinque sotto-schemi rispetto alla variante grafica in italiano antico

Come illustrato dalla tabella 74, mal(e)- generalmente si trova staccato dal secondo costituente, tranne quando occorre in combinazione con un aggettivo: in questo caso, esso mostra una netta tendenza a unirsi graficamente alla base. La perdita della vocale finale insieme all'unione grafica alla base contribuiscono alla *decategorializzazione* di mal(e)- e lo rendono più simile, dal punto di vista formale, a un prefisso.

Si può considerare un indizio di *paradigmaticizzazione* il fatto che in molti casi esso possa essere sostituito dal prefisso negativo *s*- (o, sebbene più raramente, da *in*- e *dis*-) senza apportare cambiamenti dal punto di vista semantico: si pensi a coppie come *malcontento* – *scontento* (48), *malsano* – *insano*, o *malagiato* – *disagiato*.

- (48) a. [...] come addiviene di molti religiosi, e di molte religiose, che sono *malcontenti* di loro stato (Cavalca, *Specchio de' peccati*, 1333)
  - b. Il prode uomo non sarà già mai *scontento* per cosa che gli avvenga [...] (Anonimo, *Bibbia volgare*, XIV cent.)

Lo spoglio dei dati ha quindi rivelato che già in italiano antico mal(e)- presenta delle proprietà formali, semantiche e distribuzionali che lo avvicinano a un prefissoide, ossia «a hybrid element combining characteristics of lexemes and affixes» (Hüning & Booij 2014): come mostrato nell'analisi, esso soddisfa i quattro parametri proposti da Van Goethem (2010) per stabilire la grammaticalizzazione di un'unità lessicale originariamente libera, nonostante a questo stadio non sia sempre facile distinguere tra gli usi legati o liberi di mal(e)-. Seppur non ancora sistematicamente, mal(e)- tende quindi a essere utilizzato come prefissoide per esprimere una valutazione peggiorativa o valore negativo; come emergerà nei seguenti paragrafi, questa tendenza si rafforzerà nel corso del tempo.

#### 5.3.3 Analisi diacronica

In questo paragrafo si esaminerà l'evoluzione diacronica della Costruzione [mal(e)-x] e dei suoi sotto-schemi, analizzandone i mutamenti in termini di produttività, forma e significato, prendendo in esame i dati estratti da ciascuno dei cinque sotto-corpora del CoDIt\_com. Per quanto riguarda la produttività, un primo aspetto interessante riguarda la frequenza, in termini di token e type, dell'intero dataset della Costruzione [mal(e)-x], i cui valori sono riportati nei due grafici nelle figure 56 e 57. $^{203}$ 

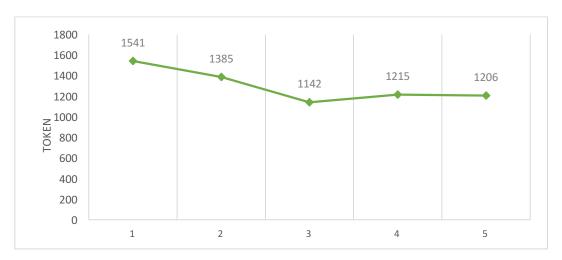


Figura 55. Frequenza dei token dell'intero dataset della Costruzione<sub>COMP</sub> [mal(e)-x]

-

 $<sup>^{203}</sup>$  I due grafici riportano le frequenze normalizzate di token e type sull'asse delle y e i numeri dei sotto-corpora del CoDIt\_com sull'asse delle x.

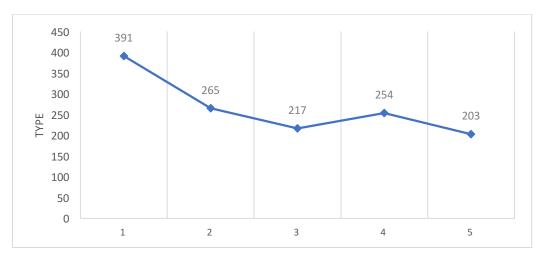


Figura 56 Frequenza dei type dell'intero dataset della Costruzione<sub>COMP</sub> [mal(e)-x]

I due grafici mettono in luce che la Costruzione<sub>COMP</sub> [mal(e)-x] si è mantenuta produttiva in termini sia di token sia di type di nel corso del tempo, come ci si aspetterebbe da un *constructional idiom*. In particolare, entrambe le curve mostrano un iniziale decremento di produttività fino al terzo sotto-corpus (corrispondente al periodo 1533-1691), a cui segue una ripresa, stabile per quanto riguarda la frequenza dei token, meno duratura nel caso di quella dei type.

È quindi interessante paragonare tali valori, relativi all'intero dataset, con quelli dei singoli sotto-schemi, le cui frequenze type/token sono illustrate nelle figure 58 e 59.

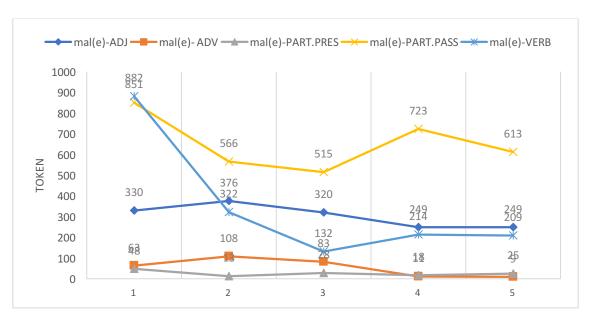


Figura 57. Frequenza dei token di ciascun sotto-schema della Costruzione<sub>COMP</sub> [mal(e)-x]

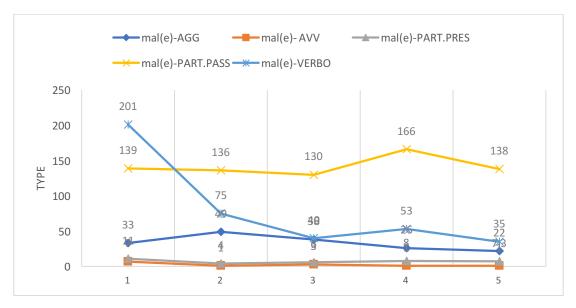


Figura 58 Frequenza dei type di ciascun sotto-schema della Costruzione<sub>COMP</sub> [mal(e)-x]

Come si osserva dai due grafici, ogni sotto-schema mostra uno specifico comportamento. I due sotto-schemi inizialmente meno produttivi, [mal(e)-AVV] e [mal(e)-PART.PRES], mantengono bassi valori di frequenza type/token in tutti i sotto-corpora. D'altra parte, le Costruzioni $_{\text{COMP}}$  più produttive in italiano antico, [mal(e)-VERBO] e [mal(e)-PART.PASS], vanno incontro a due destini differenti: mentre il primo subisce un drastico calo di frequenza type/token intorno al periodo rappresentato dal terzo sotto-corpus, seguito da una lieve ripresa, il secondo mostra un andamento meno instabile, seppur caratterizzato da un punto di svolta in corrispondenza del terzo sotto-corpus. Infine, la produttività del sotto-schema [mal(e)-AGG] si mantiene stabile nel corso del tempo. Anche per quanto riguarda la produttività potenziale (P), illustrata dai grafici nelle figure 60 e 61, si osserva un'evoluzione diversa per ciascun sotto-schema.

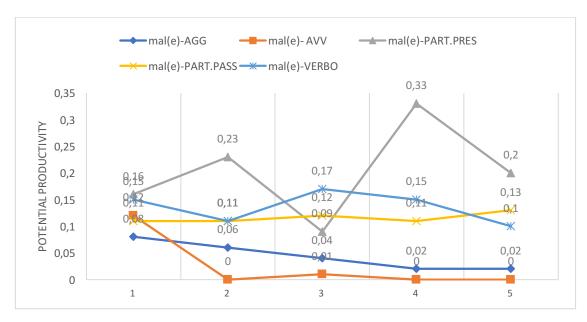


Figura 59. Produttività potenziale di ciascun sotto-schema e dell'intero dataset della Costruzione [mal(e)-x]

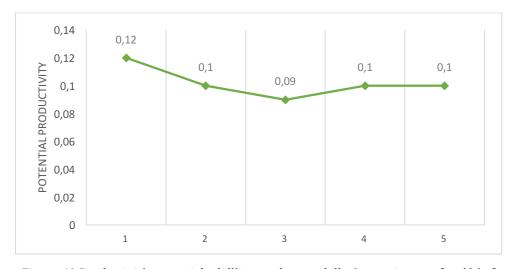


Figura 60 Produttività potenziale dell'intero dataset della Costruzione<sub>COMP</sub> [mal(e)-x]

Il grafico relativo ai singoli sotto-schemi (figura 60) permette di mettere in luce la loro evoluzione in termini di frequenza type/token nel corso della storia dell'italiano. Il primo dato osservabile è che il sotto-schema [mal(e)-AVV] cessa di essere produttivo già intorno al XV secolo: a partire da quel periodo, l'unica forma che lo istanzia è malvolentieri. Focalizzando l'attenzione sui sotto-corpora più recenti, l'unico sotto-schema che mostra un aumento di produttività è [mal(e)-PART.PASS], laddove gli altri ([mal(e)-AGG] e [mal(e)-VERBO]) presentano valori rispettivamente stabili e in lieve declino: tale schema sembra quindi esser stato stabilmente utilizzato dai parlanti per la

creazione di nuove forme. Infine, va osservato che la curva relativa alla produttività del sotto-schema [*mal(e)*-PART.PRES] presenta un andamento fortemente instabile, caratterizzato da frequenti picchi, probabilmente dovuto al ridotto numero di occorrenze su cui è stato calcolato il valore di *P*.

Il grafico riportato nella figura 61 fornisce invece una panoramica sulla produttività potenziale dell'intero dataset di forme con mal(e)-: anche in questo caso, il terzo sottocorpus rappresenta un punto chiave nell'evoluzione della costruzione, la cui produttività aumenta a partire dal periodo 1533-1691 e si mantiene costante nei sotto-corpora più recenti. Va notato che tale periodo che può essere considerato cruciale nella storia della lingua italiana in quanto immediatamente successivo alla fortunata codificazione normativa proposta da Pietro Bembo nel trattato *Prose della volgar lingua* (1525).

Nel complesso, i dati quantitativi estratti dal CoDIt\_com permettono di osservare che la costruzione [mal(e)-x], seppur con notevoli differenze a seconda del singolo sottoschema, ha conservato una discreta produttività nel corso dei secoli, fungendo stabilmente da modello per la creazione di nuove forme.

Dal punto di vista formale, è interessante analizzare quali mutamenti abbiano coinvolto la Costruzione in relazione alla grafia e alla vocale finale. La distribuzione delle occorrenze dell'intero dataset rispetto al primo parametro è illustrata nella figura 62.

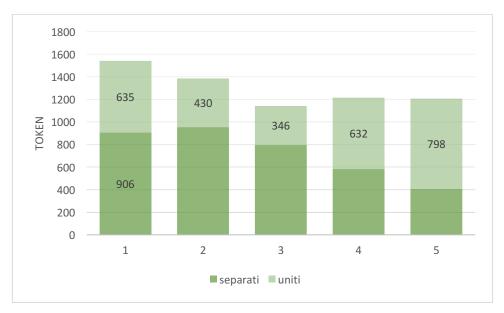


Figura 61. Distribuzione delle occorrenze dell'intero dataset rispetto alla grafia

Il grafico nella figura 62 mostra che fino al periodo rappresentato dal terzo sotto-corpus (1533-1691), mal(e)- occorre più frequentemente graficamente separato dal secondo costituente;<sup>204</sup> nei secoli successivi, invece, esso mostra una graduale tendenza a essere unito alla base: la coesione della Costruzione sembra quindi rafforzarsi gradualmente nel corso dei secoli. Per un quadro più dettagliato, si può osservare nel dettaglio il comportamento dei singoli sotto-schemi negli ultimi due sotto-corpora, più affidabili per quanto riguarda le convenzioni grafiche (figure 63 e 64).

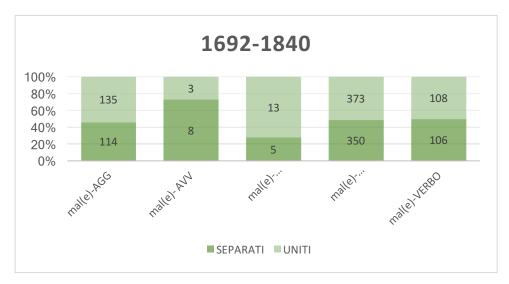


Figura 62. Distribuzione delle occorrenze dei singoli sotto-schemi in relazione alla grafia nel sotto-corpus 4 del CoDIt\_com

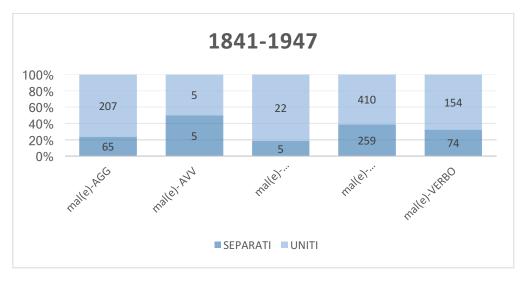


Figura 63 Distribuzione delle occorrenze dei singoli sotto-schemi in relazione alla grafia nel sotto-corpus 5 del CoDIt\_com

2

<sup>&</sup>lt;sup>204</sup> È tuttavia opportuno precisare che, specialmente per i primi due sotto-corpora, il criterio ortografico non è del tutto affidabile, dal momento che si tratta di periodi in cui le convenzioni grafiche erano ancora molto instabili.

Entrambi i grafici mettono in luce un generale incremento delle occorrenze con i due costituenti uniti graficamente per ciascun sotto-schema, ad eccezione dello schema [mal(e)-AVV] che conserva stabilmente entrambe le varianti grafiche;<sup>205</sup> i sotto-schemi più coesi sembrano essere [mal(e)-AGG] e [mal(e)-PART.PRES].

Rispetto al secondo parametro formale, la presenza/assenza della vocale finale, il grafico riportato nella figura 65 mostra chiaramente che in tutti i sotto-corpora *mal(e)*- occorre più frequentemente nella forma apocopata.

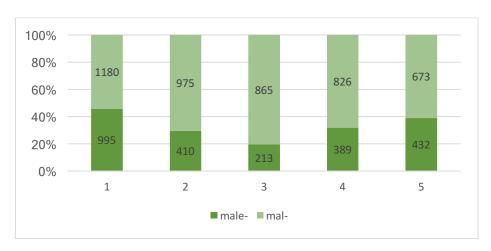


Figura 64. Distribuzione delle occorrenze della costruzione [mal(e)-x] rispetto alla presenza/assenza della vocale finale

Tuttavia, osservando le distribuzioni relative agli ultimi tre sotto-corpora, si può notare un graduale incremento delle forme non apocopate, contrario rispetto a quanto ci si aspetterebbe da un elemento che subisce un processo di decategorializzazione. Per comprendere le ragioni di questa apparente anomalia, è opportuno focalizzare l'attenzione sul comportamento dei singoli sotto-schemi negli ultimi tre sotto-corpora (figure 66, 67 e 68).

-

<sup>&</sup>lt;sup>205</sup> Tale comportamento può essere dovuto al fatto che questo sotto-schema cessa di essere produttivo già a partire dal XV secolo.



Figura 65. Distribuzione delle occorrenze di ciascun sotto-schema rispetto alla presenza/assenza della vocale finale nel sotto-corpus 3 del CoDIt\_com

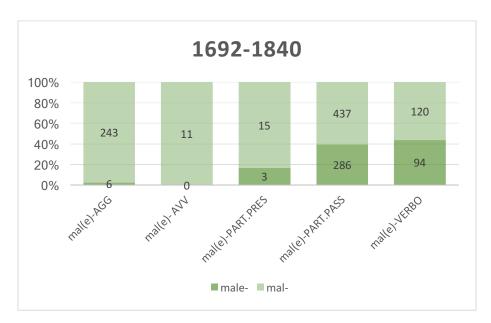


Figura 66 Distribuzione delle occorrenze di ciascun sotto-schema rispetto alla presenza/assenza della vocale finale nel sotto-corpus 4 del CoDIt\_com

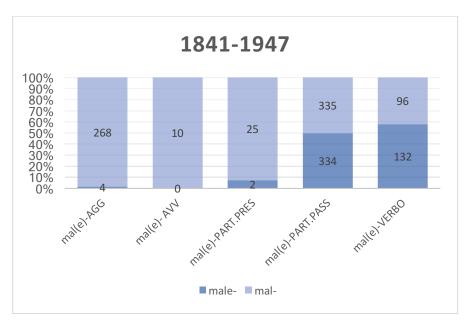


Figura 67 Distribuzione delle occorrenze di ciascun sotto-schema rispetto alla presenza/assenza della vocale finale nel sotto-corpus 5 del CoDIt\_com

Esaminando i dati riportati nei tre grafici relativi ai sotto-corpora 3-4-5, si può notare che soltanto in due sotto-schemi, ossia [mal(e)-PART.PASS] e [mal(e)-VERBO], si riscontra un incremento in termini di frequenza dei token delle forme in cui *mal(e)*- non è apocopato. Tale comportamento può essere spiegato dall'altissima frequenza di due type in particolare, maledetto e maledire, in cui mal(e)- occorre sistematicamente con la vocale finale. Tali forme, ormai opache poiché molto frequenti e di formazione molto antica (da latino *maledicere*), si sono lessicalizzate e cristallizzate in questa forma. Questa ipotesi è confermata dal nuovo scenario mostrato nella figura 69, nelle quali è riportata la distribuzione delle occorrenze della Costruzione nei sotto-corpora del CoDItcom, senza prendere in considerazione i valori di frequenza delle forme maledetto e *maledire*: il grafico mostra infatti una graduale diminuzione delle forme non apocopate, alla tendenza all'unione insieme grafica, costituisce un indizio decategorializzazione, che tipicamente coinvolge un elemento che sta acquisendo lo status di prefisso.

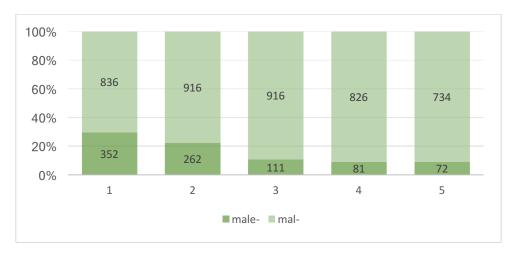


Figura 68 Distribuzione delle occorrenze tra le varianti mal-/male- senza i token delle forme maledire e maledetto

Dal punto di vista semantico, ad eccezione del sotto-schema [mal(e)-AVV] che smette di essere produttivo, negli altri sotto-schemi mal(e)- conserva nel corso del tempo gli stessi significati attestati in italiano antico, come esemplificato dai seguenti contesti:

(48) a. «Vedete dunque in quali angustie conducono i *mal fondati* principii, o, per dir meglio, le mal tirate consequenze da principii buoni» (G. Galilei, *Dialogo sopra i massimi sistemi*, 1624)

b.  $\ll$ [...] e ben l'errore scorgo or del vulgo che *mal scerne* il vero» (Della Casa, *Rime*, 1558)

(49) a. «io dispiaccio altrui perché piacciono i miei *mal fortunati* componimenti» (Tasso, *Lettere*, 1579)

b. «Rimasi spiacevolmente colpita dallo squallore, dal sudiciume, dal misero aspetto di quelle vittime di *malcauti* amori» (Caracciolo, *Misteri del chiostro napoletano*, 1864)

(50) a. «contro quel fier nemico di pietade fu *mal possente* a far ripari o schermi» (Marino, *Adone*, 1623)

b. «[...] cognito sotto il nome del Passatore, capo di una truppa di *malviventi*, che da molto tempo era il flagello della provincia di Bologna» (Chigi, *Il tempo del papa-re*, 1830-55)

Come mostrano gli esempi in (48), quando si lega a verbi o participi passati, mal(e)esprime una valutazione peggiorativa rispetto a un evento o uno stato; d'altra parte, in
combinazione con un aggettivo, esso presenta un significato più astratto di negazione:
negli esempi in (49), mal(e)- concorre infatti con i prefissi negativi, con cui condivide
proprietà semantiche e distribuzionali (malcauto - incauto, malfortunato - sfortunato).
Infine, quando mal(e)- si lega a un participio passato può assumere sia valore negativo
(come in malpossente, 50a) sia dispregiativo (come in malvivente, 50b).

Concludendo questa sezione dedicata all'analisi diacronica della costruzione [mal(e)-x], dai dati finora analizzati emerge un quadro piuttosto chiaro: a partire dall'italiano antico, mal(e)ha acquisito progressivamente, per effetto di un processo di grammaticalizzazione, un comportamento paragonabile a quello di un prefisso valutativo o negativo, a seconda della base a cui si lega. Questo processo di prefissizzazione porterà mal(e)- ad assumere lo status di vero e proprio prefisso produttivo in italiano contemporaneo, come emergerà dall'analisi fornita nel seguente paragrafo.

### 5.3.4 Italiano contemporaneo

In questo paragrafo si analizzerà da un punto di vista sincronico il comportamento della costruzione [mal(e)-x] e dei suoi sotto-schemi, sulla base di dati estratti dal web corpus itWaC (Baroni et~al.~2009).  $^{206}$ 

	mal(e)-	mal(e)-	mal(e)-	mal(e)-	mal(e)-	tot
_	AGG	PART.PRES	PART.PASS	VERBO	AVVERBIO	
Type	71	29	301	63	1	465
Token	13.313	5.545	36.838	3.635	928	60.259
Нарах	21	13	62	13	0	109
P	0,002	0,002	0,002	0,004	0	0,002

Tabella 75 Frequenza type/token e produttività potenziale dei sotto-schemi della Costruzione [mal(e)-x] in italiano contemporaneo

 $<sup>^{206}</sup>$  È opportuno sottolineare che l'analisi sincronica e quella diacronica si basano su due risorse sensibilmente differenti sia per contenuto sia per dimensioni: i dati analizzati ai § 5.3.2-3 e 5.3.4 sono di fatto solo parzialmente confrontabili, ma contribuiscono a delineare un quadro generale della Costruzione $_{\text{COMP}}$  [mal(e)-x] nel corso della storia dell'italiano.

Nella tabella 75 sono riportati i valori relativi alla frequenza type/token e alla produttività potenziale (P) dei singoli sotto-schemi della costruzione. In termini di produttività realizzata, ossia frequenza dei type, il sotto-schema più produttivo è [mal(e)-PART.PASS], mentre rispetto alla produttività potenziale il sotto-schema [mal(e)-VERBO] mostra i valori più alti. Tuttavia, va notato che il valore P del sotto-schema [mal(e)-PART.PASS] risulta particolarmente basso perché molte delle sue istanziazioni sono particolarmente frequenti: un alto numero di token influenza infatti, abbassandolo, il valore della produttività potenziale. In particolare, le istanziazioni più frequenti del sotto-schema [mal(e)-PART.PASS] sono malcapitato (4094 occ.), maltrattato (3814 occ.) e maleducato (2727 occ.). Tale sotto-schema è ampiamente sfruttato per creare aggettivi che esprimono una valutazione peggiorativa rispetto a un evento o uno stato subito o esperito da qualcuno o qualcosa, come esemplificato dai seguenti contesti:

(51) a. «Genova è solitamente conosciuta come un luogo grigio, piovoso, inquinato e *mal frequentato* come tutte le città portuali»

b. «Da turista non voglio più visitare posti dove ci sono cani e gatti *malnutriti*, malati e maltrattati in giro»

Se in *malfrequentato* e *maltrattato*, *mal(e)*- funge da peggiorativo, in *malnutrito* esso indica che l'azione di nutrire è stata compiuta 'al di sotto di una soglia di accettabilità': *malnutrito* non significa infatti 'nutrito malvagiamente' o 'non affatto nutrito', ma 'nutrito poco, non sufficientemente'. In questo caso, *mal(e)*- svolge una funzione simile a quella di *sotto*- in *sottopagato* (accanto al corrispettivo *malpagato*). Tale valore di 'al di sotto di una determinata soglia'<sup>207</sup> può slittare verso la negazione quando *mal(e)*- si combina a un aggettivo, come mostrato nell'esempio (52), in cui *maladatto* si riferisce a un abbigliamento inadatto, inappropriato:

(52) «Secondo me è anche molto bella, almeno la sera, quando non è vestita di camicie maschili *maladatte* e giacche e sottane un bel po' troppo corte...»

-

<sup>&</sup>lt;sup>207</sup> In questo caso di tratta quindi di attenuazione.

Usi eufemistici di *mal(e)*- sono attestati anche quando si trova in combinazione con un verbo, come nel caso di *malsopportare* nell'esempio (53), che può essere interpretato a seconda del contesto sia come 'non sopporta' (come probabilmente in questo caso), sia come 'sopporta poco':

(53) «[...] a questo fine hanno presentato un disegno di legge ispirato ad una filosofia xenofoba che *malsopporta* la presenza di immigrati»

La scelta tra le due interpretazioni sembra essere strettamente dipendente non solo dal significato della base lessicale ma anche dal contesto: come già osservato da Dressler & Merlini Barbaresi (1994: 55), infatti, i valutativi includono «a pragmatic variable which cannot be suppressed in the description of [their] meaning».

Infine, come già osservato nel corso dell'analisi diacronica, il sotto-schema [mal(e)-AVV] si conferma non più produttivo per la creazione di nuove forme: l'unica istanziazione attestata nel corpus di italiano contemporaneo è infatti malvolentieri (928 occorrenze).

Dal punto di vista formale, possiamo osservare la distribuzione delle occorrenze rispetto alla grafia e alla presenza/assenza della vocale finale rispettivamente nella tabella 76 e 77.

	mal(e)- AGG	mal(e)- PART.PRES	mal(e)- PART.PASS	mal(e)- VERBO	mal(e)- AVV	tot
separati	239	144	12.622	1320	179	14.504
uniti	13.074	5401	24.216	2315	749	45.755

Tabella 76 Distribuzione dei token rispetto alla grafia dei sotto-schemi della Costruzione [mal(e)-x]

	mal(e)- AGG	mal(e)- PART.PRES	mal(e)- PART.PASS	mal(e)- VERBO	mal(e)- AVVERBIO	tot
male-	32	1.479	2.649	97	0	4.257
mal-	13.281	4066	34.189	3.538	928	56.002

Tabella 77 Distribuzione dei token rispetto alla presenza/assenza della vocale finale nei sottoschemi della Costruzione [mal(e)-x]

Come ci si aspetterebbe da un'entità lessicale originariamente libera che acquisisce gradualmente lo status di elemento legato, mal(e)- tende a occorrere unito graficamente alla base (nel 76% dei casi) e nella forma apocopata mal- (92% dei token).

In conclusione, si può affermare che in italiano contemporaneo mal(e)- presenta da un punto di vista formale, semantico e distribuzionale le proprietà di un elemento legato e può quindi essere considerato un vero e proprio prefisso con funzione peggiorativa o negativa.

#### 5.3.5 Discussione

In questa sezione si riassumono i principali risultati dell'analisi diacronica della Costruzione [mal(e)-x], che verranno discussi più ampiamente nel quadro della CxM al § 5.5.

In primis, si è osservato che, già in latino, l'avverbio *male* può occorrere come primo costituente di un numero limitato di composti: se dal punto di vista formale esso non presenta modificazioni, dal punto di vista semantico, invece, mostra di aver subito uno slittamento di significato. Quando fa parte di un composto, *mal(e)*- seleziona infatti uno specifico significato tra quelli disponibili come forma libera: in particolare, quando si combina a verbi e participi, modifica la base lessicale con il valore di 'malvagiamente'; quando si lega ad aggettivi, invece, assume un valore più astratto di 'non' o 'poco'.

Analizzando gli usi di mal(e) in italiano antico si è osservato che già a partire da questo stadio, esso tende a legarsi a una base lessicale con cui forma una Costruzione di tipo morfologico. Quando si combina con verbi e participi passati, mal(e)- esprime una valutazione peggiorativa rispetto a un evento o uno stato; con aggettivi e avverbi, invece, ha un significato vicino a quello di prefisso negativo. Nel caso dei participi presenti, inoltre, si è osservato che entrambe le interpretazioni sono possibili. Applicando i parametri proposti da Van Goethem (2010) per individuare la grammaticalizzazione degli affissoidi, si è mostrato che essi vengono soddisfatti da mal(e)-, che può quindi essere considerato un elemento che ha subito un processo di grammaticalizzazione. Sulla base di dati estratti dai sotto-corpora del corpus diacronico CoDIt\_com, si è poi visto che mal(e)- ha acquisito progressivamente le proprietà formali, semantiche e distribuzionali di un prefisso: tale processo di prefissizzazione risulta ormai concluso in italiano contemporaneo, come hanno mostrato i dati estratti da itWaC.

Tali risultati inducono a ritenere che *mal(e)*- debba essere classificato come un prefisso dell'italiano: in particolare, esso può fungere da prefisso dispregiativo o negativo a seconda della base lessicale a cui si lega e del contesto d'uso. Oltre alla suffissazione (cfr. suffissi *-accio -one/ino*, *-astro*, etc.), la funzione valutativa peggiorativa può quindi essere espressa in italiano anche attraverso la prefissazione, diversamente da quanto sostenuto da precedenti studi sulla valutazione (Grandi 2002; Grandi & Montermini 2003).

# 5.4 La Costruzione [mal(o/a)-x]

La Costruzione [mal(o/a)-x] è costituita dall'aggettivo mal(o/a), <sup>208</sup> dal latino malus 'cattivo, nocivo', più una base nominale ed è utilizzata per esprimere un'attitudine peggiorativa del parlante rispetto a qualcuno o qualcosa. Come si vedrà nei seguenti paragrafi, fin dalle fasi più antiche, essa mostra legami, da un punto di vista sia semantico sia formale, con la costruzione [mal(e)-x]. L'uso dell'aggettivo malo come forma libera, possibile anche se già molto limitato in italiano antico, scomparirà nel corso dei secoli; nella fase contemporanea, malo ricorre soltanto all'interno di Costruzioni morfologiche con funzione valutativa, di cui verrà analizzata l'evoluzione diacronica nei seguenti paragrafi.

### 5.4.1 Italiano antico

L'aggettivo italiano *malo* deriva dall'aggettivo latino *malus* 'cattivo, nefasto, di scarsa qualità, nocivo'. Diversamente da *male*, *malus* non occorre all'interno di parole composte latine: le risorse lessicografiche (ad es. Glare 1982) segnalano però alcune combinazioni di *malus*, attestate nelle fonti scritte, con determinati sostantivi (ad es. *mala res* 'guaio, sfortuna', *mala aetas* 'tempi passati', *mala fides* 'malafede', *mala conscientia* 'coscienza sporca', etc.) in cui l'aggettivo si trova sempre preposto al nome, coerentemente con l'ordine dei costituenti del sintagma nominale del latino (Aggettivo - Nome). Alcune di queste combinazioni risultano attestate anche in italiano antico, come si osserva nei seguenti esempi tratti da testi di tipo religioso:

-

 $<sup>^{208}</sup>$  Nel corso dei paragrafi, se non specificato diversamente, si farà riferimento genericamente a mal(o/a), intendendo anche la forma femminile mala, le forme plurali mali/male e quella apocopata mal-.

(54) a. «Onde dice San Gregorio, che la *mala coscienza* impedisce la lingua, che non può favellare con ardire» (Cavalca, *Specchio di croce*)

b. «Ma, che peggio è, irato dio contra la nostra *mala Fede*,<sup>209</sup> e contra li nostri peccati, permette, che gli uomini, che li adorano, ricevano quello c'essi dimandano alcuna volta, perché più si confermino in questo errore» (Cavalca, *Esposizione del simbolo degli Apostoli*)

In (54), *mala coscienza* e *mala fede* ricalcano le corrispondenti espressioni latine, con cui condividono la semantica.

Nel sotto-corpus di italiano antico del CoDIt\_com, l'aggettivo malo occorre quasi esclusivamente prima di un sostantivo a cui attribuisce una valutazione peggiorativa, come esemplificato in (55); in particolare, il campione di combinazioni  $[mal(o/a)-x_N]$  estratto dal sotto-corpus consta di 274 type (di cui 146 hapax) e 995 token.

(55) a. «S'alquno metteràe le foglie de rramerino i' la botte, sì guarda e conserva il vino da ogne acierbitate e da ognie *malsapore*» (Anonimo, *Virtù del ramerino*, 1310)

b. «Che neuna meretrice o vero femina di *malafama* stia presso la casa de li signori Nove» (*Costituto del comune di Siena volgarizzato*, 1309)

Nelle forme malsapore e malafama, considerate tradizionalmente parole composte appartenenti alla Costruzione<sub>COMP</sub> [A-N], mal(o/a)- ricopre una funzione valutativa paragonabile a quella del suffisso -accio in italiano contemporaneo.<sup>210</sup> Dal punto di vista formale, si può notare che mal(o/a), seppur in un numero abbastanza limitato di casi, può occorrere unito graficamente alla base (68 occ.; 7% dei token) e nella forma apocopata mal- (265 occ.; 27% dei token).

<sup>&</sup>lt;sup>209</sup> Si noti che, nel caso di *mala fede*, la forma non presenta il significato lessicalizzato di 'consapevolezza di frodare e mentire' (De Mauro 2014) con cui è usata in italiano contemporaneo, ma ha ancora significato composizionale.

<sup>&</sup>lt;sup>210</sup> Come osservato da Mutz (2015: 151), in italiano antico il suffisso –accio svolgeva sia la funzione peggiorativa sia quella accrescitiva; in italiano contemporaneo si è specializzato nell'espressione del valore peggiorativo. Rispetto ai due casi riportati in (55), è interessante osservare che in un corpus di italiano contemporaneo (ItTenTen, Jakubíček et al. 2013) siano attestati sia famaccia (1 occ.) sia saporaccio (183 occ.).

Nel quadro teorico della CxM, i composti  $[mal(o/a)-x_N]$  possono essere considerati istanziazioni di una Costruzione $_{COMP}$  costituita da mal(o/a) e da una base nominale variabile: come la Costruzione [mal(e)-x], si tratta di una Costruzione semi-specificata che eredita le sue proprietà da quella più astratta dei composti italiani con testa a destra (cfr. § 4.2.2 e 4.3.2).

Le due Costruzioni $_{COMP}$  [mal(o/a)- $x_N$ ] e [mal(e)-x] presentano già in italiano antico alcune somiglianze da un punto di vista funzionale e formale: in primis, entrambe rappresentano strategie morfologiche attraverso cui l'italiano antico esprime la valutazione peggiorativa; inoltre, sono costituite da elementi molto simili dal punto di vista formale, ossia mal(e) e mal(o/a), che a seguito della caduta della vocale finale possono venire a coincidere nella stringa mal- e confondersi. Non sono infatti rari i casi in cui, in combinazione con un sostantivo, al posto di malo e accanto a mal, compaia male (o viceversa), come esemplificato in (56):

(56) a. «[...] e la voglia avea pure grande di ritornare al primo stato, lasciòssi vincere, e consentì al *male consiglio* del suo castaldo» (Jacopo Passavanti, *Specchio della vera penitenza*, 1355)

b. «[...] dal quale senator erano posti in croce li principi li quali per malo consiglio guidavano loro bataglie» (Anonimo, *Il volgarizzamento B del secondo libro di Valerio Massimo*, 1326)

c. «Ove è la fede della real casa di Francia, caduta per *mal consiglio*, non temendo vergogna?» (Dino Compagni, *Cronica delle cose occorrenti ne' tempi suoi*, 1310)

Questi casi sembrano indicare un legame tra le due Costruzioni<sub>COMP</sub> che, come si vedrà nei seguenti paragrafi, si intensificherà nel corso dei secoli, dando vita a un network costruzionale.

I dati analizzati finora permettono di osservare che, fin dall'italiano antico, l'aggettivo *malo* occorre quasi esclusivamente come costituente di composto, all'interno di uno schema semi-specificato, la cui base è un sostantivo variabile a cui è attribuita una valutazione dispregiativa. Tuttavia, seppur limitati (16 occ.), risultano attestati anche

casi in cui mal(o/a) si trova posposto al sostantivo (57a) o in coordinazione con un altro aggettivo (57b):

(57) a. «Poi fummo dentro al soglio della porta che è l'amor *malo*, ch'hanno avuto le anime nella prima vita» (Della Lana, *Commento alla Commedia*)

b. «[...] e conserva l'anima e guardala dal timore *malo* e reo» (Giamboni, *Giardino di consolazione*)

Gli esempi in (57) testimoniano che, in italiano antico, mal(o/a) può essere utilizzato anche come forma libera (es. in posizione non marcata e in coordinazione con altri aggettivi), nonostante tenda già significativamente a occorrere soltanto all'interno della Costruzione<sub>COMP</sub> [mal(o/a)- $x_N$ ].

#### 5.4.2 Analisi diacronica

In questo paragrafo si analizzerà, da un punto di vista diacronico, la Costruzione $_{\text{COMP}}$  [mal(o/a)- $x_N$ ], prendendo in esame i dati estratti dai diversi sotto-corpora del CoDIt\_com.

Un aspetto su cui occorre concentrarsi prima di analizzare l'evoluzione di tale Costruzione<sub>COMP</sub>, riguarda il destino dell'aggettivo *malo* come forma libera: come mostrano i dati riportati nella tabella 78, l'uso di *malo* posposto al sostantivo o in coordinazione con un altro aggettivo diminuisce progressivamente, fino a scomparire quasi del tutto nel sotto-corpus più recente del CoDIt\_com.

	1	2	3	4	5	tot	
occorrenze di <i>malo</i> come forma libera	16	12	8	5	2	43	

Tabella 78 Occorrenze di malo come forma libera nei sotto-corpora del CoDIt\_com

Diversamente da quanto osservato nel caso della Costruzione $_{\text{COMP}}$  [mal(e)-x], in cui l'avverbio male continua a essere utilizzato anche come forma libera, l'aggettivo malo esce quindi progressivamente dall'uso nel corso della storia dell'italiano, conservandosi solo all'interno della Costruzione [mal(o/a)-x<sub>N</sub>], in cui ricopre lo status di affissoide

valutativo. Per valutarne il grado di grammaticalizzazione, di seguito se ne propone un'analisi secondo i parametri discussi al § 5.2.

Sul piano quantitativo, è interessante osservare come sia evoluta la produttività della Costruzione, analizzandone la frequenza dei type e dei token (figura 70).

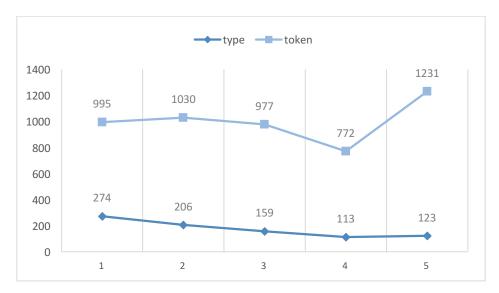


Figura 69. Frequenza dei type e dei token della costruzione  $[mal(o/a)-x_N]$  nei sotto-corpora del CoDIt\_com

Come illustrato dalla figura 70, la frequenza dei type segue uno sviluppo lineare, diminuendo progressivamente nel corso dei secoli, fino al periodo più recente in cui si registra una lieve ripresa. Diversamente, il numero dei token mostra un andamento meno regolare, con un picco negativo in corrispondenza del quarto sotto-corpus e uno positivo nell'ultimo periodo. La progressiva perdita di produttività della Costruzione $_{\text{COMP}}$  è confermata anche dai dati sulla *produttività potenziale* (P). Come riportato in figura 71, i valori di P diminuiscono nel tempo, senza tuttavia arrestarsi del tutto.

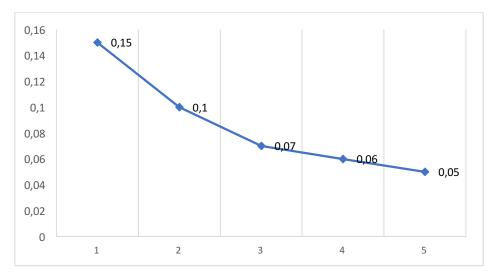


Figura 70. Produttività potenziale (P) della costruzione [mal(o/a)- $x_N$ ] nei sotto-corpora del CoDIt\_com

Nel complesso, quindi, i dati estratti dai sotto-corpora del CoDIt\_com mettono in luce che lo schema  $[mal(o/a)-x_N]$ , che in italiano antico mostra una significativa produttività, nel corso dei secoli viene utilizzato sempre di meno per la formazione di nuove parole composte, tuttavia senza uscire completamente dall'uso.

Dal punto di vista formale, è interessante monitorare la grafia con cui tali composti occorrono nei testi (figura 72) e la presenza/assenza della vocale finale in mal(o/a) (figura 73), che, come già messo in luce, costituiscono degli indizi di decategorializzazione.

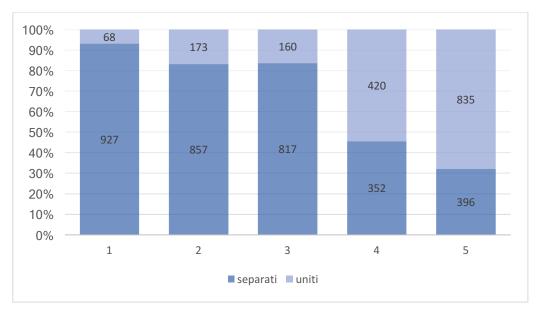


Figura 71. Distribuzione dei token della costruzione  $[mal(o/a)-x_N]$  rispetto alla grafia nei sotto-corpora del CoDIt\_com

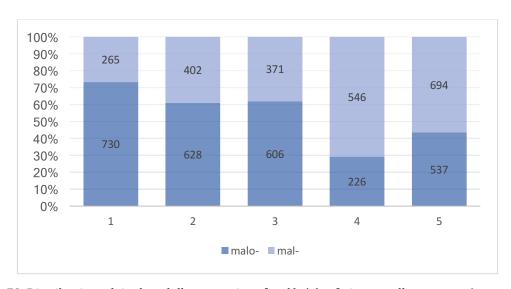


Figura 72. Distribuzione dei token della costruzione  $[mal(o/a)-x_N]$  rispetto alla presenza/assenza della vocale finale di mal(o/a)- nei sotto-corpora del CoDIt\_com

I grafici riportati nelle figure 72 e 73 mettono in luce un progressivo aumento dei casi in cui mal(o/a)- si unisce graficamente alla base e perde la vocale finale, come osservato anche per mal(e)-. Rispetto a quest'ultimo parametro, tuttavia, nel sotto-corpus più recente del CoDIt\_com si registra un lieve aumento dei casi in cui mal(o/a)- non subisce l'apocope. In queste forme, la vocale finale -a/-o sembra conservare il suo valore flessivo in relazione sia al genere (non si danno mai casi in cui il primo e il secondo costituente non concordino) sia al numero (nella quasi totalità dei casi, al plurale si ha una doppia flessione). Tale aspetto sembrerebbe indicare che, nel corso dei secoli, mal(o/a)- non abbia del tutto perso le proprietà di aggettivo (ad es. quella di concordare in genere e numero con il nome a cui si lega) e quindi non possa essere considerato un vero e proprio affissoide, ma ancora un costituente di composto o, in ottica costruzionista, un esponente di una Costruzione morfologica.

Dal punto di vista semantico, si può notare che un discreto numero di forme di antica attestazione (come maltempo, malessere, malgoverno o malasorte) va incontro a un processo di lessicalizzazione che determina una più o meno significativa opacizzazione del significato globale del composto e del valore peggiorativo di mal(o/a)-. In altri casi, l'opacizzazione del significato originario ha determinato lo sviluppo di un nuovo significato che ha soppiantato il precedente (il caso di malafede) o gli si è affiancato (il

caso di *malavita*). In particolare, a partire dal periodo compreso tra XVI e XVII secolo, *malafede* è usato con un significato diverso da quello attestato in italiano antico (cfr. (54b)), di 'fede contaminata dal peccato', come testimonia il seguente esempio tratto dal terzo sottocorpus del CoDIt\_com:

(58) «Per la qual cosa io mi cominciai a dolere, e loro mi squadernavano addosso il contratto; di modo che per via della loro *malafede* io non mi potevo aiutare» (Cellini, *Vita*, 1558-62)

In (58), *malafede* mostra infatti già il significato non trasparente di 'consapevolezza di frodare o mentire', che si è imposto su quello originario e si è conservato fino alla fase contemporanea.

Nel caso di *malavita*, la forma è attestata fin dall'italiano antico con il significato di 'vita dedita al vizio, immorale', come in (59a), in cui *mala*- veicola lo stesso valore peggiorativo di forme come *malasorte*. La forma continua a essere utilizzata con tale valore, senza soluzione di continuità ma con una frequenza gradualmente più bassa e in combinazioni parzialmente fisse come *donna di malavita* (cfr. (59b) tratto da itWaC), fino alla fase contemporanea.

(59) a. «Quando quello che è stato lungo tempo in peccato vuole lasciare suo peccato e sua *mala vita*, e vuole ritornare al consiglio di Santa Chiesa, molto deve essere lieto, quando egli sa che per lo male ch'egli ha fatto egli non puote salire alla gloria del paradiso» (Anonimo, *Bibbia volgare*, XIV-XV sec)

b. «Così praticava il De Rossi per i carcerati adulti: ma altrettanto faceva per i giovanetti racchiusi nelle carceri correzionali, e per le donne di *malavita*, per le quali, a procurarne l'emendazione e a bene ancora del pubblico, presentò le più calde suppliche alla sa. me.»

Le occorrenze di *malavita* con il significato esemplificato in (59) tendono gradualmente a ridursi parallelamente al diffondersi di un nuovo significato, attestato a partire dalla prima metà del Novecento, di 'insieme di persone dedite ad attività illecite'; in (60) se ne

riporta una delle prime attestazioni, tratta da un numero del quotidiano La Stampa del  $1902:^{211}$ 

(60) «Dice di Musolino, picciotto, che, durante la latitanza, visse con soccorsi della malavita, e spiega la sua convinzione che la *malavita* effettivamente esista in quella regione».

Nell'esempio in (60), *malavita* viene utilizzato in senso collettivo per riferirsi alla criminalità organizzata, significato con cui occorre più frequentemente nella fase odierna (anche nella variante decurtata gergale *mala*).

Accanto alle forme lessicalizzate, l'analisi diacronica rivela un certo numero di composti creati nel corso dei secoli che testimoniano come il pattern con mal(o/a)- in funzione valutativa non smetta mai del tutto di essere produttivo: la presenza di hapax di questo tipo negli ultimi tre sotto-corpora (tabella 79) sembrerebbe infatti indicare che esso continua a essere sfruttato dai parlanti per la formazione di nuove parole.

	sottocorpus	sottocorpus	sottocorpus
	3	4	5
hapax mal(o/a)-x	71	50	57

Tabella 79 Hapax *mal(o/a)*-x negli ultimi tre sotto-corpora del CoDIt\_com

In tali forme, mal(o/a)- conserva il valore peggiorativo dell'aggettivo malo (ad es. in malastruttura, malaparentela, malaunione: forme attestate a partire dal XIX secolo): esso non sembra quindi subire il processo di risemantizzazione che di norma coinvolge i costituenti di composti che acquisiscono lo status di affissoidi.

Riassumendo quanto analizzato finora da un punto di vista sia qualitativo sia quantitativo, si è osservato come la produttività della Costruzione [mal(o/a)-x] vada diminuendo nel corso dei secoli senza tuttavia esaurirsi. Accanto alle forme di più antica attestazione, che vanno incontro a un processo di opacizzazione semantica più o meno significativo a seconda dei casi, sono attestati anche hapax che testimoniano l'utilizzo di questo schema per la formazione di nuove parole in cui mal(o/a)-, ormai uscito dall'uso come forma libera, veicola un significato peggiorativo rispetto a una base. Circa lo status

<sup>&</sup>lt;sup>211</sup> In particolare, si tratta del numero 153 pubblicato il 4/06/1902.

di mal(o/a)-, analizzandone in diacronia le proprietà formali e semantiche, si è messo in luce che questo non può essere considerato un affissoide: dal punto di vista formale, infatti, mal(o/a)- presenta ancora in molti casi una vocale finale con valore flessivo; da quello semantico, esso non subisce un processo di risemantizzazione ma conserva il valore peggiorativo dell'aggettivo malo. Dai dati finora analizzati sembra quindi che, almeno fino alla prima metà del Novecento, la Costruzione $_{\text{COMP}}$  [mal(o/a)-x] sia ancora più vicina alla composizione che alla derivazione e che quindi mal(o/a)- conservi il suo status di costituente di composto. Come si osserverà nel seguente paragrafo, è in epoca più recente che tale schema acquisirà produttività.

### 5.4.3 Italiano contemporaneo

In questo paragrafo si analizzerà la Costruzione<sub>COMP</sub> [mal(o/a)- $x_N$ ] nella fase contemporanea, sulla base di dati estratti dal corpus itWaC (Baroni et~al.~2009), con l'obiettivo di metterne in luce i mutamenti formali e semantici più recenti. Diversamente dalla Costruzione<sub>COMP</sub> [mal(e)-x], infatti, nella quale mal(e)- mostra fin dall'italiano antico proprietà formali e semantiche, via via più definite, paragonabili a quelle di un prefisso negativo o peggiorativo, la Costruzione<sub>COMP</sub> con mal(o/a)- sviluppa solo negli ultimi decenni aspetti semantici e formali innovativi.

Nel paragrafo precedente si è osservato che la produttività dello schema semi-specificato  $[mal(o/a)-x_N]$  tende a diminuire nel corso dei secoli, senza tuttavia interrompersi del tutto. Semanticamente, ad eccezione dei composti che si lessicalizzano significativamente (dal caso di malafede a quello di malgrado o malapena), tali forme possono essere considerate istanziazioni di una Costruzionecomp valutativa, in quanto esprimono un'attitudine peggiorativa del parlante nei confronti di qualcosa o qualcuno. L'analisi delle proprietà formali e semantiche induce a considerare mal(o/a)- ancora come un costituente di composto, diversamente da quanto osservato per mal(e)-, che ha invece sviluppato proprietà di prefissoide e, nella fase contemporanea, di vero e proprio prefisso. Nel caso di mal(o/a)-, la fase recente sarà caratterizzata dalla formazione, negli ultimi decenni, di un nuovo sotto-schema esito di un processo di Costruzionalizzazione (Traugott & Trousdale 2014).

La tabella 80 illustra, da un punto di vista quantitativo, il campione di dati estratto da itWaC: come si può notare, l'alto numero di type e di hapax suggerisce che la Costruzione<sub>COMP</sub> è ancora ampiamente sfruttata dai parlanti per creare nuove forme.

Туре	Token	Нарах
479	106.973	205

Tabella 80 Frequenza type/token e hapax della costruzione [mal(o/a)-x] in italiano contemporaneo (dati dal corpus itWaC)

Tuttavia, è l'analisi qualitativa a fornire le informazioni più importanti per cogliere i cambiamenti a cui è andata incontro tale Costruzione $_{\text{COMP}}$  negli ultimi decenni: lo spoglio del campione estratto da itWaC permette infatti di distinguere, sulla base di aspetti formali e semantici, tre tipi di forme che presentano mal(o/a)- come primo costituente. Il primo tipo è costituito da composti già attestati nei secoli precedenti che si sono conservati fino alla fase attuale: si tratta di forme come malasorte, maltempo o malgoverno, il cui significato (come già osservato nel precedente paragrafo) ha in parte subito un processo di lessicalizzazione che, in alcuni casi, non ha oscurato del tutto l'originario valore dispregiativo di mal(o/a)- (come in maltempo, malgoverno), in altri sì (come in malocchio o malora).

Si segnalano, inoltre, alcune forme che si sono conservate (e vengono oggi utilizzate) soltanto all'interno di locuzioni avverbiali:

- (61) a. «Concepire una riforma della scuola in termini quasi esclusivamente estensivi, significa chiudere gli occhi di fronte ad un sistema scolastico dove si impara a *malapena* a leggere e a scrivere».
  - b. «Delusi e demoralizzati, quindi, hanno accettato *di malavoglia* la decisione, e prepareranno tutto il necessario per il prossimo mese».

In (61) sono esemplificati i casi di *malapena* e *malavoglia*, composti che nella fase attuale occorrono soltanto all'interno di una locuzione avverbiale con il significato rispettivamente di 'a stento' e 'malvolentieri'; condividono il medesimo destino anche le forme *malpartito* e *malomodo*, che in itWaC occorrono soltanto nelle locuzioni *a malpartito* 'in una situazione difficile' e *in malo modo* 'scortesemente'.

Appartengono al secondo tipo le forme costituite dalla stringa mal- e da un nome d'azione, come ad esempio malnutrizione, malformazione, malfunzionamento: a differenza delle forme appartenenti al primo tipo, in cui la vocale finale di mal(o/a)compare solo nelle forme femminili, in questi casi il primo costituente occorre sempre nella variante apocopata,<sup>212</sup> a prescindere dal genere della base, in virtù del legame con il verbo corrispondente nel quale la stringa *mal*- corrisponde, almeno diacronicamente, all'avverbio male ed esprime un valore peggiorativo o negativo (cfr. §. 5.3). Tali forme, infatti, si inseriscono all'interno di serie morfologiche costituite da verbo / participio passato / sostantivo (es. maltrattare / maltrattato / maltrattamento o malnutrire / malnutrito / malnutrizione), che rimandano a un evento di cui il parlante fornisce una valutazione peggiorativa o negativa (cfr. § 5.3.4). Esse possono essere considerate istanziazioni di un sotto-schema (ossia  $[mal-x_{NOME D'AZIONE}]$ ) che i parlanti utilizzano per la creazione di nuove forme, in cui il primo elemento occorre sempre come mal- e la scelta della base è ristretta ai nomi d'azione. Tale schema può essere rappresentato nella gerarchia di eredità come legato alle Costruzionicomp [mal(e)-PART.PASS] e [mal(e)-VERBO], da cui eredita le proprietà semantiche e formali.

Il terzo tipo comprende le forme costituite dalla stringa *mala*- e da una base lessicale semanticamente ristretta a nomi che indicano realtà sociali e/o istituzioni: esse occorrono soprattutto nel linguaggio politico-giornalistico, per denunciare pratiche/gestioni scorrette o negligenti, illegali o abusive (ad es. *mala-sanità*, *mala-politica*, *mala-urbanistica*, etc.), come mostrano i seguenti esempi:

(62) «Si parla tanto di *malasanità*, di interventi chirurgici errati, di diagnosi sbagliate o formulate con ritardo, di erronea interpretazione di analisi e accertamenti, di superficialità degli operatori sanitari».

<sup>&</sup>lt;sup>212</sup> Si segnalano soltanto alcuni casi in cui il primo costituente compare nella forma non apocopata, sia come *male*- (ad es. in «Questo significa che una percentuale di morti per cancro è imputabile al *malefunzionamento* delle strutture sanitarie»), sia come *mala-/malo*- (ad es. in «A proposito, non era una Nike, le ali sono un *maloassemblamento*, un'invenzione di restauro "creativo", pare rappresentasse invece la polena di una nave o chissà che altro, ma sicuramente senza ali» o «Il direttore ha così spiegato che le preoccupazioni sorte in questi mesi sono dovute a una *mala-interpretazione* delle parole del ministro Sirchia»).

b. «Il Csm è tornato ieri l'altro a far sentire la sua voce contro il disegno di legge Cirielli sulla riduzione dei tempi di prescrizione di determinati reati prefigurando scenari apocalittici di *malagiustizia*, con mafiosi, assassini, criminali fuori di galera».

c. «Ne abbiamo parlato in quanto è da anni uno dei più sfortunati esempi di *mala edilizia* o *mala urbanizzazione* della nostra provincia».

«Che *L'isola dei famosi* - dichiara - come già il *Grande fratello*, rappresenti un tipico esempio di *malatelevisione* che non rispetta i diritti dei minori, è diseducativa e disvaloriale è fuori di dubbio».

Nei casi riportati in (62), il primo costituente *mala*- non veicola un generico significato peggiorativo ma lo specifico valore di 'illecito' o 'negligente': in particolare, con malasanità ci si riferisce a fatti di cronaca che riflettono un disservizio nel funzionamento o nella gestione di una struttura sanitaria; malagiustizia allude a casi di iniquità o inefficienza nella gestione della giustizia italiana; mala edilizia e mala urbanizzazione si riferiscono a pratiche negligenti o abusive nella gestione del settore abitativo di una città. Alcune di queste forme, registrate solo da alcuni dei principali dizionari dell'uso (cfr. De Mauro 2014; Devoto Oli 2018), sono segnalate nel repertorio di neologismi di Adamo & Della Valle (2008):213 in particolare, si tratta di malacooperazione (1993), malaeconomia (1995), malafinanza (1985), malagestio (1992), malainformazione (1991), malamministrazione (1991), malapolitica (1970). Il composto di meno recente attestazione sembrerebbe quindi essere malapolitica: tale forma risulta in realtà già attestata nel quinto sottocorpus del CoDIt\_com, ma con un significato vicino a 'malgoverno'. 214 È infatti solo nella seconda metà del Novecento che malapolitica assume il significato di 'gestione inconcludente e inefficace della politica, nella quale prevalgono episodi di costume e corruzione' (Treccani Online).<sup>215</sup> Un ruolo

2

<sup>&</sup>lt;sup>213</sup> Tra parentesi si riporta l'anno di prima attestazione registrato nella banca dati *dell'Osservatorio neologico della lingua italiana* (Adamo & Della Valle 2008).

<sup>&</sup>lt;sup>214</sup> Se ne riporta il contesto d'uso, tratto dall'opera di Cesare Balbo, *Delle speranze d'Italia* (1844): «sarebbe arra dataci dall'Austria, del suo rivolgersi dalla *mala politica* sua occidentale a quella nuova e buona orientale».

<sup>&</sup>lt;sup>215</sup> In particolare, il primo contesto d'uso registrato per *malapolitica* con questo significato è il seguente: «L'opposizione afferma tuttavia che la Giunta è stata costretta alla resa dalla denuncia della "*malapolitica*" e dei favoritismi» (*Stampa*, 1° aprile 1970, p. 10, Interno).

significativo per la creazione di questo nuovo pattern è stato però senz'altro ricoperto anche dalla forma *mala-sanità*, attestato a partire dall'inizio degli anni Novanta, come testimoniato dal seguente esempio, tratto da un articolo del quotidiano *La Stampa* del 1992:

(63) «Un altro episodio incredibile - ma vero - si aggiunge alla ormai troppo lunga serie dei momenti di «malasanità» italiana che, pur clamorosa, giunti a questo punto, sembra non impressionare nessuno più di tanto».<sup>216</sup>

In (63), la forma *malasanità*, le cui virgolette sono probabile indizio che il parlante percepisce la parola come neoformazione non ancora stabilizzatasi nel lessico, allude a un mancato arrivo di un'ambulanza per soccorrere un bambino in condizioni di salute critiche: essa presenta quindi, fin dalle prime attestazioni, il significato di 'gestione negligente di x' proprio delle forme create su questo pattern. Il ruolo di possibile *leader word*, o forma modello, di *malasanità* è confermato anche dal punto di vista quantitativo, dal momento che si tratta della forma con *mala*- più frequente all'interno di itWac (649 occ.), seguita da *malagestione* (221 occ.) e *malagiustizia* (131 occ.).

Un ulteriore aspetto interessante di questo tipo di Costruzioni<sub>COMP</sub> riguarda le restrizioni semantiche e morfologiche che operano nella selezione della base, dal momento che la scelta sembra essere limitata (o quantomeno fortemente orientata) a nomi indicanti realtà sociali, astratti e femminili: un fatto raro in morfologia derivazionale, in cui proprietà come il genere del nome di base non sono generalmente pertinenti.

Dal punto di vista formale, inoltre, va osservato che *mala*- tende a occorrere esclusivamente nella forma non apocopata con la vocale finale, il cui valore flessivo (tanto per il numero quanto per il genere) sembra però indebolirsi progressivamente, come testimoniano i seguenti casi:

(64) «Esiste una politica estera italiana, da 150 anni a questa parte, fatta di sotterfugi, cambi di bandiera, tradimenti, ruberie, furbate e *malagestioni* che hanno lasciato il segno».

-

<sup>&</sup>lt;sup>216</sup> L'articolo è tratto dal numero 56 del 1992, uscito il 26/02/1992.

b. «Ogni volta che il *malafidanzato* propone cenette in casa, non riesco neppure a pensare a quelle file davanti al banco dei salumi, dove appunto i numeri hanno ritmi tutti loro».

c. «Fortunatamente in molti non hanno impiegato troppo tempo per "aprire gli occhi" e capire che Fabio Conte era un *malasindaco*».

d. «La Lazio, è un caso a parte, ma sintesi del *malacalcio* che ha distrutto l'Italia».

Gli esempi riportati in (64) mostrano un forte indebolimento del valore flessivo della vocale del primo costituente che sembra essersi cristallizzato nella forma femminile singolare *mala*-: in particolare, in (64a) *mala*- non concorda per numero con la base *gestioni*, mentre nei casi (64b-d) *mala*- si lega a sostantivi di genere maschile. Anche in questi casi, *mala*- non veicola un generico valore dispregiativo (come in *malasorte* o *malafama*), ma un significato più specifico: in (64a) e (64d), indica che l'attività espressa dalla base (o la gestione di tale attività) è stata svolta in modo illecito o negligente; in (64b) e (64c), designa persone che non hanno ricoperto il loro ruolo in maniera efficiente o apprezzabile. Inoltre, gli esempi (64b) e (64c), seppur marginali, sembrano indicare un indebolimento delle restrizioni semantiche che operano sulla base, che in questi casi è un nome indicante un individuo.

Da un punto di vista quantitativo, è interessante osservare come si distribuisce il campione di forme estratte da itWaC rispetto ai tre tipi individuati (tabella 81).

	Type	Token	Нарах	Р
$[[mal(o/a)]-x_N]$	324	80.630	141	0,002
[mal-X <sub>NOME D'AZIONE</sub> ]	67	24.636	22	0,001
$[mala-x_N]$	88	1.707	42	0,02

Tabella 81 Frequenza type/token, hapax e produttività potenziale dei composti con mal(o/a)- estratte da itWaC

I tre pattern individuati, la cui posizione all'interno del network costruzionale verrà discussa al § 5.5, sembrano essere tutti produttivi, seppur in gradi diversi. Come mostrano i dati riportati nella tabella 81, lo schema più produttivo è quello più recente, in cui mala- presenta proprietà formali e semantiche di prefissoide. La Costruzione<sub>COMP</sub> [mal- $x_{NOME\ D'AZIONE}$ ] mostra invece valori di P piuttosto bassi, dovuti in parte all'alta frequenza in termini di token di alcune forme, come maltrattamento (6.430 token),

malformazione (4.795 token) o malnutrizione (2.785 token). Infine, il discreto numero di hapax in cui mal(o/a)- esprime l'originario valore peggiorativo testimonia che tale pattern compositivo non ha del tutto esaurito la sua produttività, come mostrano gli esempi riportati in (65).

(65) a. «La C : R : I ha messo a disposizione una unità di pronto intervento con cucina di campo per tè, panini; una unità di rianimazione per eventuali malori e una serie di tende riscaldate per accogliere donne, minori o persone in *malasalute*».

b. «Un giro d' orizzonte a tutto campo per cercare di mettere tutti d'accordo sulla necessità di colmare le lacune del «Mattarellum», bonificare la palude dello scorporo e mandare in pensione la *mala-abitudine* delle liste civetta».

Le forme riportate in  $(65)^{217}$  sono create sul modello della Costruzione<sub>COMP</sub> [mal(o/a)-x] in cui il primo costituente funge da peggiorativo rispetto alla base: in malasalute e malaabitudine, mala- non veicola il significato di 'illecita, negligente' (come in mala-sanità) ma di 'cattiva salute' e 'cattiva abitudine' (così come maltempo è sinonimo di 'cattivo tempo').

Riassumendo, l'analisi delle forme costituite dalla stringa mal(o/a)- (o mal-) e un sostantivo ha permesso di individuare tre tipi di Costruzioni<sub>COMP</sub> produttive:

- a) la Costruzione<sub>COMP</sub> [mal(o/a)- $x_N$ ], con mal(o/a)- in funzione di peggiorativo, produttiva fin dall'italiano antico (es. mala-abitudine, mala-scrittura, etc.);
- b) la Costruzione<sub>COMP</sub> [*mal*-x<sub>NOME D'AZIONE</sub>] che dà origine a sostantivi creati a partire da verbi o participi passati con il prefisso *mal* (etimologicamente legato all'avverbio *male*) con cui crea una serie morfologica verbo / participio passato / sostantivo (es. *maltrattare* / *maltrattato* / *maltrattamento*);
- c) la Costruzione $_{COMP}$  [mala- $x_N$ ], emersa solo negli ultimi decenni, in cui malapresenta proprietà semantiche (ad es. il significato più ristretto di 'illecito,

-

<sup>&</sup>lt;sup>217</sup> Entrambe le forme sono hapax nel corpus itWaC e non sono registrate dai dizionari di italiano.

negligente') e formali (ad es. perdita del valore flessivo della vocale finale) proprie di un prefissoide.

In tutti e tre i casi si tratta di schemi semi-specificati, costituiti dalle stringhe mal(o/a)-, mal- e mala- e una base variabile nei limiti delle restrizioni evidenziate nel corso dell'analisi: nel primo caso, mal(o/a)- veicola un significato peggiorativo rispetto alla semantica della base, che non è soggetta a particolari restrizioni semantiche; nel secondo caso mal-si lega soltanto a nomi d'azione e veicola un significato peggiorativo o negativo; nel secondo, *mala*- si lega a nomi designanti realtà sociali, generalmente femminili, con il significato di 'illecito', 'negligente', salvo alcuni rari casi in cui si unisce a nomi dalla semantica diversa, che sembrerebbero indicare un'espansione di classe della base. Applicando i parametri di grammaticalizzazione di Van Goethem (2010) alla costruzione [mala-x] si può concludere che mala- presenta le proprietà tipiche degli affissoidi. In particolare, dal punto di vista semantico si è osservato un processo di risemantizzazione che ha portato mala- ad assumere, da un più astratto valore peggiorativo, quello di 'negligente, illecito'. Dal punto di vista formale, i dati estratti da itWaC hanno mostrato come mala- stia attraversando un processo di decategorializzazione per cui la vocale finale -a sta progressivamente perdendo valore flessivo sia rispetto al numero sia rispetto al genere. Sul piano quantitativo, l'alto numero di neoformazioni di questo tipo conferma che si tratta di uno schema produttivo, che si inserisce all'interno del network costruzionale costituito dalle altre Costruzioni<sub>COMP</sub> con mal(o/a)-, la cui discussione nel quadro della CxM è oggetto del seguente paragrafo.

#### 5.4.4 Discussione

In questo paragrafo si riepilogano i principali risultati dell'analisi diacronica dei composti con mal(o/a)-, inquadrandoli nel framework della CxM.

In primis, è emerso che già in italiano antico l'aggettivo malo tende a occorrere quasi esclusivamente premesso a un sostantivo di cui esprime una valutazione peggiorativa, come si è osservato nei casi di malasorte, maltempo, malgoverno: tali forme, possono essere considerate istanziazioni di uno schema semi-specificato,  $[mal(o/a)-x_N]$ , costituito da un primo elemento fisso, mal(o/a), e il secondo variabile ma ristretto alla categoria dei sostantivi; tale schema costituisce a sua volta un'istanziazione della

Costruzione [AVV-X]<sub>X</sub> dei composti con testa a destra (§ 4.2.2 e 4.3.2). Come emerso dall'analisi diacronica, nel corso dei secoli, tale Costruzione mostra un progressivo calo in termini di produttività: diminuiscono, infatti, sia il numero dei type sia quello degli hapax. Parallelamente, gran parte dei composti formatisi in epoca antica va incontro a un processo di lessicalizzazione che ne oscura il significato e rinforza la coesione tra i due costituenti, che occorrono nella maggior parte dei casi uniti graficamente: tale processo di opacizzazione favorisce la diminuzione in termini di produttività della Costruzione  $[mal(o/a)-x_N]$  con mal(o/a)- in funzione di peggiorativo. Diversamente da quanto osservato per la Costruzione<sub>COMP</sub> [mal(e)-x], in questo caso, il primo costituente dello schema, potenzialmente candidato ad acquisire lo status di prefisso come *mal(e)*sulla base di proprietà formali, distribuzionali e semantiche presentate in italiano antico, non si grammaticalizza a tal punto nel corso dei secoli, ma può essere considerato fino alla fase contemporanea un esponente di costruzione con valore peggiorativo, in virtù del fatto che l'uso dell'aggettivo malo come forma libera (quindi non preposta a un sostantivo), diviene residuale a partire dal Cinquecento e scompare nei secoli successivi. Se nel caso della Costruzione [mal(e)-x] l'analisi dei dati estratti da un corpus di italiano contemporaneo ha confermato l'esito del percorso di grammaticalizzazione di mal(e)-, che da prefissoide ha progressivamente acquisito le proprietà di vero e proprio prefisso valutativo/negativo dell'italiano, in quello di  $[mal(o/a)-x_N]$  i dati estratti da itWaC hanno invece messo in luce almeno due elementi di innovazione.

In primis, si è osservata la presenza di un discreto numero di forme, assenti nei secoli precedenti, costituite dalla variante apocopata mal- e un nome d'azione, retroformato a partire da un verbo o un participio passato che istanziano la Costruzione [mal(e)-x] (es. malnutrizione da malnutrito; malinterpretazione da malinterpretato/malinterpretare). In queste forme, che si inseriscono all'interno di reti lessicali nome / participio passato / verbo, mal- esprime valore peggiorativo/negativo e occorre sempre nella forma ridotta, senza la vocale finale: in virtù del forte legame semantico e morfologico con i corrispondenti participi/verbi, lo schema [mal- $x_{NOME D'AZIONE}$ ] può essere considerato un sotto-schema delle Costruzioni [mal(e)-VERBO] e [mal(e)-PART.PASS], piuttosto che di [mal(o/a)- $x_N$ ].

Il secondo elemento di innovazione emerso dall'analisi sincronica riguarda la nascita di un nuovo schema semi-specificato costituito dalla stringa *mala*- e da un sostantivo soggetto a restrizioni di tipo semantico: in particolare, si tratta di nomi che designano realtà sociali o istituzioni, di cui si denuncia la gestione illecita o negligente, come malasanità, mala-politica, mala-urbanistica. In queste forme, mala- compare sempre nella forma femminile non apocopata e mostra una tendenza a non flettersi per genere (come in malacalcio, malatennis, etc.) e numero (come in malapolitiche). In ottica costruzionista, si tratta di un sotto-schema della Costruzione<sub>COMP</sub> [mal(o/a)-x], in cui mala- sembra stia acquisendo lo status di un prefissoide, di cui presenta le principali proprietà (cfr. Van Goethem 2010): in particolare, dal punto di vista formale, la vocale finale del primo costituente sembra stia perdendo valore flessivo (parametro decategorializzazione); dal punto di vista semantico, mala- veicola un significato diverso e più ristretto rispetto al più astratto valore peggiorativo di mal(o/a) (parametro risemantizzazione). Inoltre, in un numero ancora limitato di forme si è osservato un indebolimento delle restrizioni semantiche che vincolano la scelta del secondo costituente ai nomi di realtà sociali o istituzioni: forme come malafidanzato, malasindaco e *malasuocera* sembrano infatti indicare un possibile allargamento della semantica della Costruzione. In termini quantitativi, come emerso dai dati riportati nella tabella 81, tale schema risulta essere il più produttivo (parametro produttività). La nascita di questa Costruzione può essere considerata un esempio di Costruzionalizzazione (cfr. Traugott & Trousdale 2014: 22): a partire da un certo numero di Costrutti (in questo caso, probabilmente, le forme mala-sanità e mala-politica), infatti, si è innescato un meccanismo di rianalisi che ha portato alla formazione di un nuovo schema semispecificato con caratteristiche formali e semantiche nuove.

### 5.4 Osservazioni conclusive

In questa sezione si riepilogano e si discutono nel quadro della CxM, i principali risultati emersi dall'analisi diacronica delle Costruzioni $_{\text{COMP}}$  [mal(e)-x] e [mal(o/a)-x].

I due schemi possono essere rappresentati all'interno della Gerarchia di eredità che discende dalla Costruzione astratta dei composti italiani con testa a destra (figura 74) e si specifica nei due schemi semi-specificati con mal(o/a)- e mal(e)-, legati da similarità semantiche e formali ma che, come si è visto, presentano un'evoluzione e uno status sincronico differente.

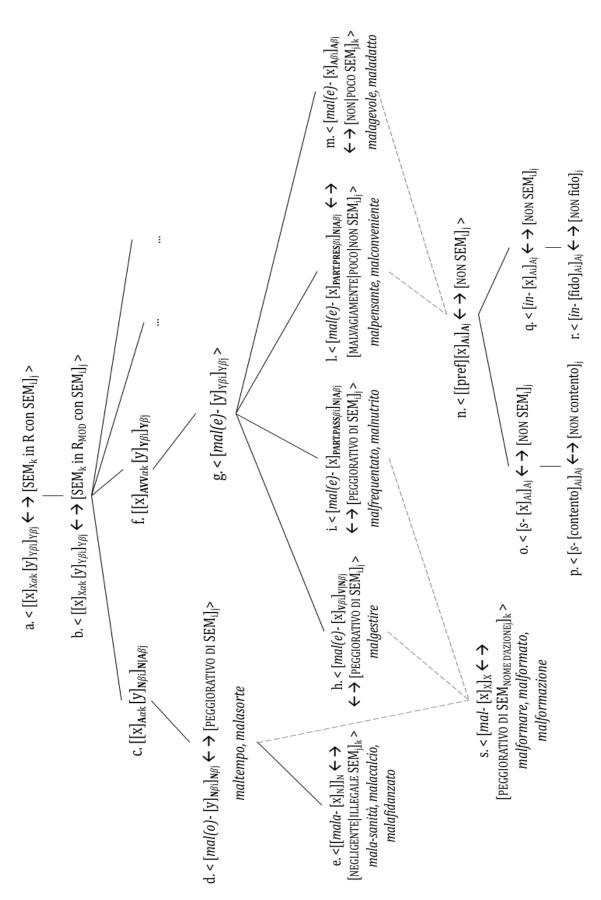


Figura 73 Gerarchia di eredità delle Costruzioni $_{\text{COMP}}$  [mal(e)-x] e [mal(o/a)-x] in italiano contemporaneo

Come raffigurato nella figura 74, le due Costruzioni<sub>COMP</sub> semi-specificate [mal(o/a)- $x_N$ ] e [mal(e)-x] danno luogo a un network di Costruzioni sotto-specificate a livelli più bassi della Gerarchia.

In particolare, la Costruzione<sub>COMP</sub> [mal(e)-x], rappresentata in (g), si sotto-specifica in quattro sotto-schemi (i-m), tutti già attestati fin dall'italiano antico, a seconda della categoria del secondo costituente. In virtù del fatto che con nel sotto-schema [mal(e)-AGG], mal(e)- entra in distribuzione complementare con alcuni prefissi negativi, come diso s-, nella figura 74 esso si lega alla Costruzione dei derivati con prefissazione negativa (n). Inoltre, i sotto-schemi [mal(e)-PART.PASS] e [mal(e)-VERBO] permettono la creazione di sostantivi costituiti dalla stringa mal- e da un nome d'azione, creati a partire dai corrispettivi participi passati e verbi con *mal(e)*- (cfr. schema in (s)). Sulla base delle proprietà semantiche, formali e distribuzionali osservate nel corso dell'analisi, si è concluso che *mal(e)*- ha ormai acquisito lo status di prefisso, con valore peggiorativo o negativo,<sup>218</sup> e che la Costruzione<sub>COMP</sub> [mal(e)-x] può essere considerata più vicina alla derivazione che alla composizione.

La Costruzione<sub>COMP</sub> [mal(o/a)-x], in cui mal(o/a)-, rappresentata in (d), conserva invece il proprio status di costituente di composto e permette ancora nella fase contemporanea la creazione di nuove parole composte in cui il primo elemento conserva l'originario valore peggiorativo. Questa si sotto-specifica in un nuovo sotto-schema, rappresentato in (e), emerso negli ultimi decenni, costituito dalla stringa mala- e da un sostantivo indicante una realtà sociale o una istituzione: in virtù delle proprietà formali e semantiche analizzate nei precedenti paragrafi, si è osservato che mala- può essere considerato un prefissoide nel senso di Hüning & Booij (2014) e Van Goethem (2010) e che la formazione di tale sotto-schema rappresenti un caso di Costruzionalizzazione (Traugott & Trousdale 2014).

I due casi analizzati permettono, infine, di discutere, nel quadro della CxM, alcuni aspetti teorici relativi alla composizione e alla morfologia valutativa.

In primo luogo, si è osservato che anche in italiano i costituenti di composto possono grammaticalizzarsi e acquisire gradualmente lo status di affissoide e/o di vero e proprio

<sup>&</sup>lt;sup>218</sup> Come osservato in Iacobini & Thornton (2016: 146-149), in italiano contemporaneo, la negazione

costituisce, insieme alla ripetizione, il significato più produttivo nella prefissazione verbale ed è molto diffuso in quella nominale e aggettivale.

affisso. I parametri proposti da Van Goethem (2010) si sono dimostrati efficaci per individuare, in prospettiva diacronica, i mutamenti riguardanti le proprietà formali, semantiche e distribuzionali di mal(e)- e mal(o)- ed evidenziarne i due diversi percorsi di grammaticalizzazione.

Un secondo aspetto meritevole di attenzione riguarda la struttura delle forme analizzate: in entrambi i casi si tratta di composti con testa a destra, il cui modificatore a sinistra ha subito un processo di grammaticalizzazione che, nel caso dell'avverbio mal(e)-, ha portato alla trasformazione in vero e proprio prefisso, mentre in quello di mal(o)- ha determinato solo negli ultimi decenni la formazione di un affissoide, mala-, esito di un processo di Costruzionalizzazione. La prima posizione sembra quindi essere favorita in italiano per lo sviluppo di affissoidi a partire da costituenti di composto, diversamente da quanto osservato nelle lingue germaniche, nelle quali entrambi i costituenti possono acquisire lo status di elemento legato. Le ragioni di tale preferenza possono essere almeno tre. *In primis*, la rilevanza, nel sistema morfologico italiano, della flessione, che ostacolerebbe la decategorializzazione di un elemento posto a destra di un composto. In secondo luogo, l'influenza di due pattern di formazione di parola: quello neoclassico, costituito da due elementi legati di origine greco-latina con testa a destra (es. fonografo, termometro, etc.) e, soprattutto, quello di formazioni costituite da un elemento formativo con funzione di determinante che si aggiunge a una base libera con funzione di determinatore (es. audiocassetta, cicloraduno, psicolinquistica, etc.), la cui fortuna nell'italiano contemporaneo degli ultimi decenni è stata osservata e discussa da Thornton & Iacobini (1992: 42-54); Iacobini & Thornton (2016: 143).<sup>219</sup> La diffusione di tali forme conferma l'integrazione del pattern neoclassico nella formazione delle parole dell'italiano contemporaneo e ne motiva l'analisi a fianco dei meccanismi di composizione e derivazione, tra i quali sembra collocarsi.

Infine, si è osservato che le due  $Costruzioni_{COMP}$  analizzate possono essere considerate dal punto di vista semantico e funzionale come valutative: mal(e)- e mal(o)- si legano infatti a un base di cui esprimono una valutazione peggiorativa. L'analisi della  $Costruzione_{COMP}$ 

\_

<sup>&</sup>lt;sup>219</sup> La determinazione a sinistra è individuata dagli autori come una tendenza morfologica in crescente diffusione negli ultimi decenni. Queste forme «non appartengono esclusivamente a linguaggi tecnicospecialistici, ma comprendono sempre di più anche termini di uso comune, la cui composizionalità è motivata per il parlante non specialista» (Thornton & Iacobini 1992: 48). Secondo gli autori, tuttavia, tale processo formativo, che testimonia un ormai consolidato legame tra elementi neoclassici e lingua comune, è più vicino alla prefissazione che alla composizione.

[mal(e)-x], in particolare, ha messo in luce almeno due aspetti rilevanti nell'ambito dello studio della morfologia valutativa dell'italiano. *In primis*, si è osservato lo sviluppo di un prefisso valutativo a partire da un costituente di composto: tale prefisso esprime esclusivamente una valutazione qualitativa (in particolare, peggiorativa). Questo aspetto assume rilevanza se messo in relazione con i precedenti studi dedicati alla valutazione: in particolare, analizzando la morfologia valutativa delle lingue del Mediterraneo, Grandi (2006: 86) osserva che «con l'esclusione dei pochi suffissi peggiorativi e dei prefissi accrescitivo-migliorativi, la polarità GOOD-BAD non dispone di specifiche marche linguistiche davvero produttive».

	Prefissi	Suffissi
SMALL	mini-, micro-	-ino, -etto, -ello, ecc.
	es. miniappartamento, minishow, microcriminalità, miniconsultazione	es. appartamentino, gattino, casetta, pastorello
BIG	maxi-, mega-	-one
	es. maxi-schermo, maxi-pranzo,	es. gattone, concorsone, barcone,
	maxi-concorso, megaconcerto	bambolona
GOOD	super-, extra-, stra-, arci-, ultra-	-issimo (?)
	es. super-lavoro, super-ricercato,	es. presidentissimo, finalissima,
	stracontento, ultrasottile, ecc.	campionissimo, occasionissima
BAD		-accio, -astro
		es. ragazzaccio, postaccio,
		poetastro

Figura 74 Classificazione dei prefissi e dei suffissi valutativi dell'italiano proposta da Grandi & Montermini (2003)

Nella classificazione dei prefissi e suffissi valutativi dell'italiano proposta in Grandi & Montermini (2003) e illustrata nella figura 75, infatti, si ritiene che il valore BAD venga espresso in italiano soltanto tramite la suffissazione (in particolare, per mezzo dei suffissi -accio, -astro, etc.). Mettendo in luce l'esistenza di un prefisso valutativo deputato esclusivamente alla valutazione peggiorativa, i risultati di questa indagine inducono a rivedere tale classificazione, ascrivendo mal(e)- alla classe dei prefissi peggiorativi dell'italiano.

# 6. CONCLUSIONI

In questo capitolo si riepilogano i risultati emersi da questo lavoro. In particolare, il § 6.1 evidenzia il contributo di questa ricerca rispetto al modello teorico adottato, la Morfologia delle Costruzioni; il 6.2 fornisce un bilancio conclusivo sulla composizione dell'italiano.

# 6.1 La composizione dell'italiano nel quadro della Morfologia delle Costruzioni

In questo lavoro, la composizione dell'italiano è stata analizzata in prospettiva diacronica attraverso gli strumenti teorici proposti dalla Morfologia delle Costruzioni (CxM): come discusso al capitolo 2, per rappresentare le parole composte che l'italiano ha creato nel corso dei secoli è stata proposta la nozione di Costruzione<sub>COMP</sub>, la cui struttura è articolata in schemi e sotto-schemi caratterizzati da diversi livelli di astrattezza e di produttività. L'adozione di questo quadro teorico ha permesso di mettere in luce alcuni aspetti, che di seguito si riepilogano, utili a comprendere come opera (e ha operato in passato) il meccanismo della composizione in italiano. Lo studio della composizione italiana ha d'altra parte consentito di arricchire il quadro teorico, finora adottato quasi esclusivamente per analizzare la formazione delle parole composte nelle lingue germaniche (cfr., tra gli altri, Booij 2010c, 2015b).

Un primo contributo è legato alla rappresentazione della semantica delle Costruzioni<sub>COMP</sub>: in particolare, in questo lavoro si è proposto di integrare il formalismo tradizionalmente adottato per la descrizione di fenomeni morfologici (cfr. Masini & Audring, in stampa) con l'indicazione del tipo semantico dei costituenti e dell'intero composto, al fine di rendere conto delle regolarità e delle restrizioni semantiche dei costituenti osservate nel corso dell'analisi. Questo strumento si è rivelato particolarmente utile per descrivere schemi piuttosto complessi dal punto di vista semantico, come la Costruzione<sub>COMP</sub> [V-N], che fin dalla fase antica è caratterizzata da una significativa varietà semantica, o il sottoschema [capo-x] in cui il primo costituente può assumere diversi significati e selezionare, in alcuni casi, il secondo elemento. Il significato delle parole composte è stato quindi rappresentato attraverso un set di tipi semantici, invece che attraverso le perifrasi

tradizionalmente adottate negli studi di CxM, al fine di identificare la presenza di regolarità nella semantica delle unità di input e/o di output.<sup>220</sup>

Dal punto di vista teorico, un primo aspetto rilevante da osservare è che la classificazione in 'famiglie di Costruzioni' proposta da Arcodia & Basciano (2018) per la descrizione delle parole composte del cinese si è dimostrata efficace anche per rendere conto della composizione dell'italiano: l'individuazione di quattro classi di Costruzioni<sub>COMP</sub> astratte, articolate in schemi e sotto-schemi, sulla base di tre criteri (relazione grammaticale tra costituenti, endo/esocentricità, posizione della testa; cfr. § 1.3 e 4.2) ha permesso di descrivere l'intero repertorio di composti estratti dalle risorse utilizzate (ossia il TLIO per la fase antica, il CoDIt\_com per il periodo compreso tra le Origini e il 1947, e itWaC per la fase contemporanea). Questo lavoro consolida quindi tale macro-classificazione, all'interno della quale sono stati individuati schemi e sotto-schemi specifici della composizione dell'italiano (es. i composti [V-N]<sub>N|A</sub>, [V-V]<sub>N</sub>, etc.).

L'analisi proposta al capitolo 4 ha evidenziato che le famiglie di Costruzioni<sub>COMP</sub> presentano un numero variabile di sotto-schemi caratterizzati da diversi livelli di produttività: ad esempio, è stato osservato che la famiglia di Costruzioni<sub>COMP</sub> con testa a destra istanzia un numero notevole di sotto-schemi, alcuni molto produttivi (in particolare, lo schema [AVV-X]x), altri meno (ad es. lo schema [N-PART.PASS]<sub>N|A</sub>). L'analisi diacronica ha inoltre messo in luce che la produttività di uno schema può variare nel corso del tempo: ad esempio, la Costruzione<sub>COMP</sub> [N-N]<sub>N\_MOD</sub> con testa a sinistra (es. *pescespada*; cfr. § 4.2.1 e 4.3.1), scarsamente produttiva nella fase antica, acquisisce rilevanza nel corso dei secoli e costituisce una Costruzione<sub>COMP</sub> particolarmente produttiva in italiano contemporaneo, come testimoniano i numerosi composti del tipo *notizia bomba, riunione fiume, città simbolo,* etc. (cfr. Grandi, Nissim & Tamburini 2011; Grandi 2009; Radimský 2015a, 2016).<sup>221</sup>

D'altra parte, l'analisi ha evidenziato che la Costruzione<sub>COMP</sub> [V-N] si mantiene altamente produttiva nel corso dei secoli, dalla fase antica a quella contemporanea: senza soluzione

<sup>220</sup> La natura della relazione semantica che lega i due elementi di un composto non è invece stata integrata nel formalismo adottato, né analizzata nel corso del lavoro: su questo aspetto è auspicabile quindi che si concentrino studi futuri

concentrino studi futuri.

 $<sup>^{221}</sup>$ È inoltre interessante osservare che in italiano contemporaneo, a differenza di quanto osservato nei composti analizzati in questo lavoro, i composti [N-N]\_{N\_{LMOD}} con testa a sinistra occorrono quasi esclusivamente nella grafia separata dallo spazio (cfr. Radimský 2015a).

di continuità, tale schema istanzia nomi (e, in una fase più recente, anche aggettivi) che designano prevalentemente persone, artefatti, eventi (cfr. § 4.2.4 e 4.3.4)

Un altro caso che testimonia come la produttività di un sotto-schema possa variare nel corso del tempo è quello della Costruzione [N-PART.PRES]\_N|A (es. armilucente, luogotenente, etc.; cfr. § 4.2.2 e 4.3.2): tale schema, produttivo già in latino e nella fase antica, nel corso dei secoli istanzia un numero limitato di forme che occorrono quasi esclusivamente in testi poetici che riprendono il modello latino, ma nella fase recente riacquista vitalità grazie alla fortuna del sotto-schema semi-specificato [x-dipendente]\_N|A. Quest'ultimo caso, insieme ad altri, mette in luce come la produttività di una Costruzione  $_{\text{COMP}}$  possa essere legata alla fortuna di un sotto-schema semi-specificato, costituito da uno slot chiuso e uno variabile; nel corso dell'analisi sono stati messi in luce gli schemi semi-specificati più rilevanti, di seguito elencati:

```
1. [capo-x] (cfr. § 4.2.5.2 e 4.3.1)
```

- 2. [x-vendolo] (cfr. § 4.2.2 e 4.3.2)
- 3. [ben(e)-x] (cfr. § 4.3.2)
- 4. [buon(o/a)-x] (cfr. § 4.2.5.1 e 4.3.2)
- 5. [mezz(o/a)-x] (cfr. § 4.2.5.1 e 4.3.2)
- 6. [mal(e)-x] (cfr. § 5.3)
- 7. [mal(o/a)-x] (cfr. § 5.4)

Ad esclusione di [x-vendolo], si tratta di sotto-schemi in cui lo slot chiuso è rappresentato dall'elemento di sinistra, che può essere un nome (nel caso di [capo-x]) ma più frequentemente un avverbio o un aggettivo. Dal punto di vista formale, va notato che questi elementi presentano sempre una struttura bisillabica, che può ridursi in seguito alla perdita della vocale finale (come nel caso di [mal(e)-x] o [ben(e)-x]).

Come messo in luce al § 4.2.5.2, il sotto-schema [capo-x] costituisce un caso interessante dal punto di vista semantico, dal momento che anche in altre lingue (ad es. olandese, tedesco, francese) il termine per indicare la testa occorre all'interno di composti e può sviluppare significati più astratti. L'analisi dei composti con capo- in italiano ha messo in luce che, nella fase antica, capo- può assumere diversi significati sia concreti ('testa',

'persona dotata di autorità')<sup>222</sup> sia astratti ('elemento preminente', 'inizio'). Esaminando tale sotto-schema dal punto di vista diacronico, si è però verificato che i composti con *capo*- che vengono formati nei secoli successivi alla fase antica indicano prevalentemente una persona dotata di autorità su un gruppo di persone (es. *capogruppo*) o su un luogo (es. *capoclasse*): il sotto-schema [*capo*<sub>PERSONA</sub>-X<sub>LUOGO</sub>|GR\_PERS]<sub>PERSONA</sub> è quindi l'unico che si conserva produttivo in diacronia fino alla fase contemporanea. In italiano, quindi, *capo*- non subisce lo stesso processo di grammaticalizzazione attestato in olandese o in tedesco, per il quale *Haupt*- e *hoofd*-, quando occorrono all'interno di composto, esprimono il significato astratto di 'principale'.

L'analisi degli schemi [mal(e)-x] e [mal(o/a)-x], a cui è stato dedicato il capitolo 5, ha consentito di mettere in luce che, seppur in maniera più limitata rispetto a quanto accade nelle lingue germaniche, anche in italiano le entità che prendono parte al meccanismo della composizione possono subire un processo di grammaticalizzazione. In particolare, applicando i parametri proposti da Van Goethem (2010) per lo studio di questo fenomeno in olandese e francese, è stato dimostrato che mal(e)- ha gradualmente acquisito lo status di prefisso, con valore peggiorativo o negativo a seconda della base a cui si lega, e può quindi essere considerato una marca morfologica per la valutazione peggiorativa. Questo studio di caso ha contribuito a ridefinire il quadro della morfologia valutativa dell'italiano: contrariamente a quanto sostenuto da studi precedenti (cfr. Grandi & Montermini 2003; Grandi 2002), esso ha dimostrato che l'italiano utilizza anche la prefissazione (accanto alla suffissazione, vd. -accio, -astro) per esprimere il polo 'BAD' della valutazione. Il caso di *mal(e)*- ha confermato inoltre un importante principio della CxM, secondo cui composizione e derivazione sono poste lungo un continuum e le unità che vi prendono parte possono trovarsi a diverse distanze tra i due poli: se analizzato dal punto di vista diacronico, questo fenomeno si configura come un mutamento che porta un costituente di composto a diventare un elemento legato, passando dal polo della composizione a quello della derivazione.

L'analisi diacronica dello schema [mal(o/a)-x] ha permesso di osservare un più articolato percorso di grammaticalizzazione, in cui da una Costruzione<sub>COMP</sub> originariamente soltanto peggiorativa (istanziata da forme come maltempo, malasorte)

<sup>&</sup>lt;sup>222</sup> Più precisamente, quando *capo* assume il significato di 'autorità' (es. *capostazione*), il contributo semantico del primo costituente è astratto (essendo l'autorità un concetto astratto), mentre l'intero composto ha semantica concreta.

è emerso un sotto-schema [mala-x] (es. malapolitica, malaurbanistica, malauniversità, etc.) in cui il primo costituente *mala*- può essere considerato un affissoide che esprime il valore (non più genericamente peggiorativo) di 'negligente, illecito' e la base è soggetta a restrizioni semantiche (ossia è generalmente ristretta a nomi astratti di entità sociali o istituzioni) e morfologiche (si tratta quasi sempre di nomi femminili). La nascita di tale sotto-schema è stata collegata alla fortuna del termine malasanità, che può essere considerata la leaderword sulla base di cui, attraverso un processo di astrazione e analogia, si è formato lo schema [mala-x] che ha istanziato numerose altre forme. L'analisi dei dati estratti da itWaC ha inoltre messo in luce l'esistenza del sotto-schema  $[mal-x_{NOME\ D'AZIONE}]$  in cui il primo costituente occorre sempre senza vocale finale mal- e la base è sempre un nome d'azione: tale schema rappresenta un punto di congiunzione tra la Costruzione<sub>COMP</sub> [mal(e)-x] e la Costruzione<sub>COMP</sub> [mal(o/a)-x], dal momento che istanzia forme che si inseriscono all'interno di serie lessicali di verbo / participio passato / nome (es. maltrattare / maltrattare / maltrattamento, etc.). Il nesso tra le due Costruzionicomp ha rafforzato l'idea che, all'interno del Costruzionario, le Costruzioni costituiscono un network dinamico basato su legami di natura semantica e formale.

Due ulteriori elementi sono meritevoli di attenzione: a differenza delle lingue germaniche, in cui la grammaticalizzazione può coinvolgere sia i costituenti posti a sinistra sia quelli posti a destra, sia il modificatore sia la testa del composto (cfr., tra gli altri, Bauer 2005 sul suffissoide –*voll* in tedesco, Booij 2008: 60 sul prefissoide *oud*-dell'olandese), in italiano la posizione privilegiata per la grammaticalizzazione sembra essere quella di sinistra e l'elemento che subisce tale processo costituisce il modificatore del composto. Inoltre,va notato che l'emergere di sotto-schemi semi-specificati sembra essere favorito nella famiglia di Costruzioni<sub>COMP</sub> con testa a destra:<sup>223</sup> dall'analisi emerge infatti che né la famiglia delle Costruzioni<sub>COMP</sub> esocentriche né quella delle Costruzioni<sub>COMP</sub> coordinative endocentriche istanziano sotto-schemi con uno slot chiuso e uno variabile. L'analisi ha infine evidenziato che la creazione di un nuovo schema può essere dovuta all'influenza di un'altra lingua: la composizione dell'italiano deve al latino e al modello di altre lingue, *in primis* il francese, la creazione di nuovi schemi più o meno fortunati. In particolare, il legame con il latino è ancora chiaramente vivo nella fase antica, come

\_

<sup>&</sup>lt;sup>223</sup> Ad esclusione del caso di [*capo-*x], che appartiene in realtà a più famiglie a seconda della relazione che lega *capo-* al secondo costituente e del suo significato.

testimoniato dai numerosi composti attestati nel dataset *Compos\_ita\_TLIO* che hanno origine dall'univerbazione di espressioni latine (come *terremoto* da *terrae motus*) o costituiscono prestiti adattati dal latino. In più casi è stato infatti osservato che, sul modello di un prestito particolarmente fortunato, ha avuto origine uno schema che ha istanziato nuove forme, come nel caso dello schema [x-vendolo] da cui *fruttivendolo*, *lanovendolo*, etc.

Nel complesso, quindi, il quadro teorico della CxM si è dimostrato uno strumento efficace per rendere conto della formazione delle parole composte dell'italiano, la loro organizzazione all'interno del Costruzionario, e per individuare i mutamenti che coinvolgono le unità di input, il cui status categoriale può non rimanere costante nel corso del tempo.

## 6.2 Un bilancio sulla composizione dell'italiano

Lo scopo di questo lavoro è stato delineare un profilo diacronico della composizione dell'italiano, mettendone in luce le principali caratteristiche e individuando le linee di tendenza che ne hanno guidato lo sviluppo dalle Origini alla fase contemporanea. Di seguito si riepilogano i principali risultati, mettendone in luce il contributo allo studio della composizione e gli interrogativi sollevati.

Dall'analisi dei composti raccolti nel dataset *Compos\_ita\_TLIO* è emerso *in primis* che l'italiano utilizza la composizione per creare nuove parole fin dalle Origini: le parole composte sono infatti ampiamente attestate fin dal XIII secolo. È inoltre rilevante osservare che gran parte delle tipologie di composto che occorrono in italiano contemporaneo risulta già attestata in italiano antico.<sup>224</sup>

L'analisi dedicata alle Origini ha messo in luce che, in quella fase, la gran parte delle forme attestate è costituita da composti [V-N] o da prestiti dal latino, che presentano prevalentemente la struttura con testa a destra. In particolare, la composizione verbonominale appare fin da subito uno dei motori della composizione, attraverso cui l'italiano crea nomi che designano persone, artefatti o eventi: come già osservato da Štichauer (2015, 2016) per il periodo Cinquecento-Ottocento, la varietà semantica è infatti una delle caratteristiche peculiari dei composti [V-N]. D'altra parte è stato

\_

<sup>&</sup>lt;sup>224</sup> Fanno eccezione quattro schemi:  $[A-N]_{N|A}$  (es. *rosso sangue*),  $[A-A]_{A|N}$  (es. *verdescuro*),  $[N-N]_{N}$  (es. *capocarceriere*) e  $[V-AVV]_{N}$  (es. *buttafuori*).

osservato che molti composti costituiscono prestiti di origine latina e presentano caratteristiche strutturali simili a quelle dei composti latini (tra cui la testa a destra e il primo costituente bisillabico e terminante in vocale, es. *altitonante, caprifoglio, patrisonno*, etc.): in non pochi casi, sul modello di prestiti dal latino si sono formate altre forme native, come nel caso dei composti con *mal(e)*- e *ben(e)*- (es. *malfattore, benevolente*), già attestati in latino e divenuti (nel primo caso) una delle tipologie più produttive in assoluto. La composizione dell'italiano antico appare quindi essenzialmente verbonominale (quindi esocentrica) o con testa a destra: in particolare, all'interno di quest'ultimo insieme di forme, va segnalato il significativo numero di composti che presentano un aggettivo o un avverbio come modificatore posto a sinistra (*buon(o)*-x, *mezzo*-x, *mal(e)*-x, *alto*-x, *ben(e)*-x). Dal punto di vista formale, è stata osservata una generale preferenza per una struttura con primi elementi bisillabici terminanti in vocale e secondi elementi con lunghezza variabile: questa tendenza è molto netta nei composti [V-N] ma è riscontrabile in quasi tutte le tipologie.

Paragonati ai composti con testa a destra, quelli con testa a sinistra non sono particolarmente numerosi nella fase antica, come osservato al § 4.2.1: la posizione canonica della testa dei composti dell'italiano antico non sembra essere quindi quella di sinistra. All'interno di questo gruppo di forme, il tipo [N-N] più attestato è quello in cui i costituenti sono legati da una relazione di subordinazione (es. *capofamiglia*); più rari sono invece i composti appositivi (es. *pescespada*), la cui produttività in italiano contemporaneo è invece piuttosto significativa (cfr. Radimský 2015a, 2016; Grandi, Nissim & Tamburini 2011).

La coordinazione è una relazione sfruttata in maniera limitata, ma comunque già attestata, in italiano antico: come illustrato ai §§ 4.2.3 e 4.2.4, il dataset  $Compos\_ita\_TLIO$  attesta i composti coordinativi endocentrici  $[A-A]_{N|A}$  (es. celtibero, gallogreco) e  $[N-N]_N$  (es. cassapanca, casatorre) e quelli esocentrici  $[V-V]_N$  (es. andirivieni, saliscendi). Mentre nel primo caso (composti  $[A-A]_{N|A}$ ) si tratta prevalentemente di forme ereditate dal latino, negli altri due si tratta quasi esclusivamente di composti creati autonomamente dall'italiano; il pattern  $[V-V]_N$ , in particolare, non ha riscontro in latino e può essere considerato un'innovazione dell'italiano.

La composizione nei secoli successivi evolve in sostanziale continuità con la fase antica: per quasi tutte le tipologie di composto è stata osservata una produttività stabile (o in

crescita) nel corso dei secoli. La composizione verbonominale si è confermata uno dei principali motori della composizione, stabilmente sfruttata per la creazione di nuove forme, alcune delle quali entrate stabilmente nel lessico (es. *portafoglio*, *guardaroba*, etc.), altre rimaste occasionalismi creati per soddisfare esigenze comunicative estemporanee. Alla fortuna di questo tipo di forme possono aver contribuito molteplici fattori: dal punto di vista formale, esse presentano una struttura per la quale è stata osservata una generale preferenza in italiano, ossia un primo costituente bisillabico terminante in vocale e un secondo elmento di lunghezza variabile; dal punto di vista semantico, i composti [V-N] mostrano senza soluzione di continuità una significativa varietà semantica. Inoltre, a partire dal XIV-XV secolo, questi composti sono usati, seppur sporadicamente, anche (o, in alcuni rari casi, esclusivamente) in funzione di aggettivo: questa tendenza, ancora debole nei secoli presi in esame, si consoliderà negli ultimi decenni, come mostrato da Ricca (2005, 2010).

Accanto ai composti  $[V-N]_N$ , sono stabilmente produttivi i composti con testa a destra costituiti da un avverbio o un aggettivo in prima posizione, in particolare i composti con mal(e)- e mal(o/a)-, le cui vicende diacroniche sono state analizzate al capitolo 5: in entrambi i casi si tratta di composti con modificatore bisillabico e terminante in vocale (che può cadere in determinati contesti) che esprime un significato valutativo. In particolare, l'analisi diacronica dei composti con mal(e)- ha dimostrato che un costituente di composto posto in prima posizione può acquisire le proprietà di un affisso e che quindi, se osservate da un punto di vista diacronico, le unità che prendono parte ai processi di formazione di parola possono non mantenere invariato il proprio status categoriale.

Tra i composti con testa a sinistra attestati nel CoDIt\_com, un nutrito gruppo è costituito da forme che presentano *capo*- (nel senso di 'autorità') come primo costituente (es. *capostazione, capodipartimento*, etc.), già attestate nella fase antica e particolarmente frequenti anche in italiano contemporaneo. Ad esclusione di questo tipo di forme, i composti [N-N]<sub>N</sub> legati da una relazione di subordinazione non si sono distinti per particolare produttività: la fortuna di questo pattern, particolarmente sfruttato in italiano contemporaneo (cfr. Lombardi Vallauri 2008; Terreni 2005; Radimský 2015a), sembra essere una caratteristica della fase più recente. Costituiscono invece un'innovazione rispetto all'italiano antico i termini di colore [A-N]<sub>N|A</sub> (es. *verde acqua*,

giallo canarino, etc.) costituiti da un aggettivo indicante un colore e un secondo costituente che designa un elemento che prototipicamente ne possiede una determinata tonalità.

I composti coordinativi (endocentrici [N-N]<sub>N</sub> e [A-A]<sub>AIN</sub> ed esocentrici [V-V]<sub>N</sub>) sono attestati in maniera limitata ma senza soluzione di continuità nel corso dei secoli. In particolare, i composti [N-N]<sub>N</sub> designano prevalentemente entità animate o (soprattutto) inanimate che presentano proprietà di due entità; sono quindi rari o del tutto assenti composti che designano individui che svolgono due funzioni (es. studente-lavoratore, attore-cantante) o luoghi adibiti a due funzioni (come ristorante-pizzeria, barpasticceria, etc.): la diffusione di questo tipo di composti, rilevante in italiano contemporaneo, va quindi collocata in un periodo successivo a quello rappresentato dal CoDIt\_com. Per quanto riguarda i composti [A-A]<sub>A|N</sub>, si è osservata la presenza di un considerevole numero di composti che indicano tonalità intermedie tra due colori (semantica intersettiva, es. gialloarancione) o, ma solo nella fase più recente, la somma di due colori (semantica additiva, es. biancogiallo):<sup>225</sup> quest'ultimo tipo, scarsamente attestato nella fase antica, si afferma nel corso dei secoli. Inoltre, dal punto di vista formale, in queste forme si è osservata una generale preferenza per una struttura con primo costituente bisillabico e terminante in vocale (molto spesso -o). I composti  $[V-V]_N$ , che nella fase antica sono costituiti da due basi verbali diverse, a partire dal periodo 1692-1840 iniziano a presentare due basi verbali uguali. Dal punto di vista semantico, essi designano generalmente un artefatto o un evento; il significato di 'azione compiuta da un gruppo di persone', attestato in un considerevole numero di composti dell'italiano contemporaneo (es. fuggi-fuggi; cfr. Thornton 2008, 2009, 2010; Masini & Thornton 2007), è minoritario nel sotto-corpus più recente.

Nel complesso, quindi, considerando la composizione dell'italiano in prospettiva diacronica nel periodo rappresentato dal CoDIt\_com, gli elementi di continuità sono più numerosi di quelli di discontinuità: gran parte di quello che è la composizione nella fase contemporanea esisteva già in epoca antica, seppur con le differenze analizzate nel corso dell'analisi.

\_

 $<sup>^{225}</sup>$  Cfr. D'Achille & Grossmann (2013) per la terminologia utilizzata per descrivere la semantica di questi composti.

Prendendo in esame le categorie delle unità di input che prendono parte alla formazione dei composti italiani e la posizione della testa, è stato osservato (cfr. tabella 72) che i composti con testa a sinistra sono costituiti esclusivamente da nomi e aggettivi; maggior varietà in termini di categorie lessicali è invece caratteristica dei composti con testa a destra. Nei composti con testa a sinistra, il primo costituente può quindi essere soltanto un nome o un aggettivo (composti [N-N]<sub>N\_SUB|MOD</sub>, es. *capostazione*, *pescespada*; [A-N]<sub>A|N</sub>, es. *rosso sangue*); nei composti con testa a destra, invece, l'elemento testa può essere un nome, un aggettivo, un verbo, un participio (es. [N-PART.PRES]<sub>N|A\_SUB</sub>, [AVV-V]<sub>V</sub>, etc.). La struttura con testa a destra permette quindi a più parti del discorso di ricoprire la funzione di testa di un composto.

Dal punto di vista formale, è stata più volte messa in luce nel corso dell'analisi una preferenza per una struttura con un primo costituente bisillabico e terminante in vocale; tra le ragioni di questa preferenza vi è senz'altro il modello esercitato dai composti del latino, la cui influenza si esercita sull'italiano principalmente attraverso due modalità. Nella fase antica, il latino fornisce un numero consistente di forme (composti o espressioni multiparola N-N o N-A, es. *finem mundi* 'finimondo', *caput regionum* 'caporione') che entrano in italiano e vengono adattate alla struttura propria dei composti dell'italiano. Queste forme arricchiscono il nucleo della composizione italiana e fungono da modello per la creazione di nuove forme nei secoli successivi; sono infatti poche le tipologie di composto che non hanno riscontro in latino e che possono dirsi quindi innovazioni dell'italiano: *in primis*, i composti esocentrici [V-V]<sub>N</sub> e in parte quelli [V-N]<sub>N</sub>, la cui rilevanza in latino è decisamente marginale rispetto a quella che questa assume in italiano e nelle altre lingue romanze. L'influenza del latino sulle prime fasi della lingua è stata valutata attraverso l'analisi del dataset *Compos\_ita\_TLIO*, mettendo in luce il legame che emerge tra i due tipi di composizione.

Non sono state invece analizzate le parole composte del latino medievale, tranne che per il caso dei composti con *capo*- (cfr. 4.2.5.2): l'analisi delle forme con *caput* registrate dal vocabolario Du Cange (1982) ha permesso di ricostruire la storia di questo tipo di composti, in particolare di quelli in cui *capo*- esprime il significato di 'autorità', assenti nel latino classico ma attestati in italiano fin dalla fase antica e con crescente produttività nel corso dei secoli. Questo caso di studio, seppur limitato, ha evidenziato l'importanza

di considerare in futuro anche dati di latino medievale nello studio della composizione italiana, in quanto appartenenti a un periodo di transizione tra le due lingue.

Non è stata inoltre oggetto di analisi diretta la composizione neoclassica, che pure ha esercitato (ed esercita) un'influenza sulle parole composte dell'italiano: è in particolare a partire dal Settecento che in italiano iniziano a diffondersi, in primo luogo nei linguaggi specialistici, composti costituiti da elementi formativi del greco e del latino, che presentano caratteristiche formali molto vicine a quelle dei composti del latino. Inizialmente limitati alla terminologia specialistica, nell'ultimo secolo molti di questi composti sono entrati nel lessico della lingua comune: come osservato da Iacobini (2004a) e Iacobini & Thornton (1992), la composizione con elementi formativi non coinvolge più soltanto elementi di origine greco-latina ma anche elementi nativi, derivati da accorciamenti o per segmentazione di una parola etc. (es. foto- da fotografia, -poli da tangentopoli). Una delle conseguenze dell'interazione tra i due tipi di composizione è stata la diffusione di composti con testa a destra, costituiti da un primo elemento legato, bisillabico e terminante in vocale, e un elemento libero in funzione di testa (es. videocassetta, eurozona, etc.): per tale struttura, ampiamente riscontrata nel campione di composti qui analizzati, l'italiano sembra quindi avere una netta preferenza.

Riassumendo, i rapporti tra composizione nativa, latina (classica e medievale) e neoclassica sono illustrati dalla figura 76.



Figura 75 Composizione italiana, composizione latina, composizione neoclassica

Nello schema, la composizione italiana e quella latina e neoclassica si trovano su due livelli diversi, in virtù delle diverse caratteristiche formali identificate al § 4.1. Tra la

composizione latina classica e l'italiano antico si colloca la composizione del latino medievale, ancora priva di studi sistematici e per questo contrassegnata da un punto di domanda. Il legame tra la composizione latina e quella italiana della fase antica è rappresentato da una freccia rossa tratteggiata che indica la trasmissione di prestiti che, una volta adattati, arricchiscono il nucleo dei composti italiani, accanto a quelli nativi. Nel corso dei secoli successivi al periodo delle Origini, la composizione italiana continua la sua storia senza significativi mutamenti, mentre quella del latino si interrompe, salvo riacquistare produttività nella forma della composizione neoclassica a partire dal Settecento. Il legame tra composizione latina e quella neoclassica è rappresentato da una linea tratteggiata, che rende conto, da un lato, della similarità formale, dall'altro della distanza temporale che le separa. Nella fase contemporanea, composizione nativa e neoclassica convivono come due meccanismi di formazione di parola di cui l'italiano si serve per espandere il proprio lessico: entrambe permettono di creare nuove parole ascrivibili alla categoria di composti, ma con caratteristiche strutturali differenti. Tuttavia, l'interazione tra i due meccanismi ha determinato la diffusione di forme che presentano un primo elemento non testa legato e un elemento libero: questo tipo di composizione è rappresentato dal rettangolo arancione posto tra quello blu della composizione nativa e quello verde della composizione neoclassica.

Il quadro delineato in questo lavoro presenta l'italiano come una lingua che fin dalle Origini sfrutta la composizione, accanto alla derivazione, per la creazione di nuove forme: il tipo di composizione più tipicamente italiano è quello verbonominale, le cui caratteristiche formali e semantiche le garantiscono produttività costante nel corso dei secoli. Sul modello di composti latini entrati in uso come prestiti, l'italiano ha creato forme appartenenti ad altre tipologie (ad es. i composti [N-N]<sub>N\_SUB</sub> con testa a destra) e ha arricchito il proprio repertorio di composti anche di nuove tipologie, sconosciute al latino, come il tipo [V-V]<sub>N</sub>. Oltre agli aspetti già messi in evidenza, l'analisi diacronica proposta in questo lavoro porta a riconsiderare un altro elemento discusso da studi precedenti, ossia la rilevanza della struttura con testa a destra nei composti dell'italiano: come dimostrato dalla significativa presenza di queste forme nel *dataset* esaminato, questo pattern, ereditato dal latino, è stabilmente produttivo nel corso dei secoli e decisamente più sfruttato rispetto a quello con testa a sinistra (almeno fino al 1947), tradizionalmente considerata la posizione canonica dei composti dell'italiano

contemporaneo (cfr. Masini & Scalise 2012). Inoltre, l'analisi ha dato modo di osservare che i composti con testa a destra presentano una maggior varietà di unità di input (nomi, aggettivi, participi, etc.). Tra le ragioni della preferenza della composizione italiana per la struttura con testa a destra, oltre alla già ricordata influenza del latino, si può senz'altro segnalare l'ordine dei costituenti Modificatore-Testa, marcato rispetto a quello della sintassi (Testa-Modificatore): esso infatti segnala che una data entità è un composto, distinguendola da un sintagma libero. La posizione della testa è quindi usata come marca di composizione, in assenza di altri elementi morfologici deputati a distinguere esplicitamente gli elementi morfologici da quelli sintattici (come ad es. i *linking element* del latino o del russo).

Infine, come si è messo in luce in quest'ultimo capitolo, questa ricerca risponde ad alcuni degli interrogativi lasciati aperti dai precedenti studi e ne solleva inevitabilmente di altri. In particolare, tre sembrano essere le direzioni verso cui orientare futuri studi dedicati a questo settore della formazione delle parole. *In primis*, come già messo in luce, l'analisi delle parole composte del latino medievale permetterebbe di colmare il gap tra composizione classica e italiano-antica, contribuendo a chiarire l'origine della composizione italiana. In secondo luogo, il quadro qui delineato andrebbe confrontato più approfonditamente con un profilo della composizione dell'italiano contemporaneo formulato sulla base di dati estratti da corpora, come è stato fatto per alcune particolari tipologie di composto (ad es. i composti [V-N], cfr. Ricca 2005, 2010), al fine di delineare un quadro completo della composizione dell'italiano. In terzo luogo, sembra rilevante approfondire, alla luce dei risultati di questo studio, la questione relativa alla posizione della testa dei composti, analizzando i neologismi che presentano tale struttura e quella Elemento formativo + Elemento libero, per convalidare l'ipotesi che questa struttura sia (e in che misura) produttiva in italiano contemporaneo. A questi tre aspetti è necessario che si rivolgano i futuri studi dedicati alla composizione dell'italiano, la quale, come dimostrato in questo lavoro, costituisce un meccanismo produttivo e dotato di caratteristiche ben definite già nella fase antica.

## Bibliografia

- Adamo, G. & Della Valle, V. (2008). *Il Vocabolario Treccani. Neologismi. Parole nuove dai giornali.* Roma: Istituto della Enciclopedia italiana.
- Amiot, D. (1997). *L'antériorité temporelle dans la préfixation en français*. Villeneuve d'Ascq: Presses Univ. Septentrion.
- Amiot, D. (2002). De l'utilité de la notion de métaphore pour décrire le sens d'un préfixe: le cas de sur. *Verbum (Presses Universitaires de Nancy)*, 24(3), 269-282.
- Amiot, D. (2005). Between Compounding and Derivation: Elements of Word-formation Corresponding to Prepositions. In Dressler, W., Kastovsky, D., Pfeiffer, O. & Rainer, F. (a cura di), *Morphology and Its Demarcations. Selected papers from the 11th Morphology Meeting*, Vienna, February 2004. Amsterdam: John Benjamins, 183-195.
- Amiot, D. & Van Goethem, K. (2012). A constructional account of French-clé 'key' and Dutch sleutel-'key' as in mot-clé/sleutelwoord 'key word'. *Morphology*, 22(3), 347-364.
- Anderson, S. R. (1992). *A-morphous morphology* (Vol. 62). Cambridge: Cambridge: University Press.
- Anthony, L. (2013). Developing AntConc for a new generation of corpus linguists. *Proceedings of the Corpus Linguistics Conference* (CL 2013), July 22-26, Lancaster University, UK, 14-16.
- Antonelli, G. (1996). Sui prefissoidi dell'italiano contemporaneo. *Studi di lessicografia italiana*, 13, 253-293.
- Arcodia, G. F. (2011). A Construction Morphology account of derivation in Mandarin Chinese. *Morphology*, 21(1), 89-130.
- Arcodia, G. F. (2012). Constructions and headedness in derivation and compounding. *Morphology*, 22(3), 365-397.
- Arcodia, G. F. (in stampa). Coordinating nominal compounds: Universals vs. areal tendencies. *Linguistics.*
- Arcodia, G. F., & Basciano, B. (2018). The Construction Morphology Analysis of Chinese Word Formation. In Booij, G. (a cura di), *The Construction of Words. Advances in construction Morphology*. Cham Springer, vol. 4, 219-253
- Arcodia, G. F., Grandi, N., & Montermini, F. (2009). Hierarchical NN compounds in a cross-linguistic perspective. *Rivista di Linguistica*, 21(1), 11-33.
- Arcodia, G. F., Grandi, N., & Wälchli, B. (2010). Coordination in compounding. In Scalise, S. & Vogel, I. (a cura di), *Cross-disciplinary issues in compounding*. Philadelphia: Benjamins, 177-198.
- Aronoff, M. (1976). Word formation in generative grammar. *Linguistic Inquiry Monographs Cambridge, Mass.*, (1), 1-134.
- Aronoff, M. (1994). Morphology by itself: Stems and inflectional classes (No. 22). MIT press.
- Baayen, R. H. (1992). Quantitative aspects of morphological productivity. In Booij, G. & van Marle, J. (a cura di), *Yearbook of Morphology (1991)*. Dordrecht: Kluwer Academic Publishers, 109-149
- Baayen, R. H. (2005). Morphological productivity. In Köhler, R., Altmann, G. & Piotrowski, R. G. (a cura di), *Quantitative Linguistics: An international handbook*. Berlin: Walter de Gruyter, 243–256.
- Barðdal, J. (2008). *Productivity: Evidence from case and argument structure in Icelandic.* Philadelphia/Amsterdam: John Benjamins Publishing.

- Baroni, M., Guevara, E., & Pirrelli, V. (2007). NN compounds in Italian: Modelling category induction and analogical extension. *Lingue e linguaggio*, *6*(2), 263-290.
- Baroni, M., Bernardini, S., Ferraresi, A., & Zanchetta, E. (2009). The WaCky wide web: a collection of very large linguistically processed web-crawled corpora. *Language resources and evaluation*, 43(3), 209-226.
- Baroni, M., Guevara, E., & Zamparelli, R. (2009). The dual nature of deverbal nominal constructions: evidence from acceptability ratings and corpus analysis. *Corpus Linguistics and Linguistic Theory*, *5*(1), 27-60.
- Baroni, M., Guevara, E., & Pirrelli, V. (2006). Sulla tipologia dei composti N+ N in italiano: principi categoriali ed evidenza distribuzionale a confronto. *Linguistica e modelli tecnologici di ricerca*, 73-95.
- Bauer, L. (1978). *The grammar of nominal compounding: With special reference to Danish, English and French* (Vol. 4). University Press of Southern Denmark.
- Bauer, L. (1998). When is a sequence of two nouns a compound in English?. *English Language & Linguistics*, *2*(1), 65-86.
- Bauer, L. (2001). *Morphological productivity* (Vol. 95). Cambridge: Cambridge University Press.
- Bauer, L. (2005). The borderline between derivation and compounding. In Dressler, W., Kastovsky, D., Pfeiffer, O. & Rainer, F. (eds), *Morphology and its Demarcations*. Amsterdam/Philadelphia: Benjamins, 97-108.
- Bauer, L. (2011). Word formation. In: Maiden, M., Smith, J. C. & Ledgeway, A. (a cura di), *The Cambridge History of the Romance Languages. Volume I. Structures.* Cambridge: Cambridge University Press, 532–563.
- Bauer, L. (2017). Compounds and compounding. Cambridge: Cambridge University Press.
- Bayeen, R. H. (1993). On frequency, transparency, and productivity. In Booij, G. & van Marle, J. (a cura di), *Yearbook of Morphology (1992)*. Dordrecht: Kluwer Academic Publishers, 181-208.
- Bayeen, R. H. (2005). Morphological productivity. In Kohler, R., Altmann, G. & Piotrowski, R. G. (a cura di), *Quantitative Linguistics: An international handbook*. Berlin: Walter de Gruyter, 243-256.
- Bayeen, R. H. (2009). Corpus linguistics in morphology: Morphological productivity. In Lüdeling, A., Kytö, M. & McEnery, T. (a cura di), *Corpus Linguistics. An International Handbook*, vol. 2. Berlin: Mouton De Gruyter, 899-919.
- Bayeen, R. H. & Lieber, R. (1991). Productivity and English derivation: A corpus-based study. *Linguistics*, 29, 801-843.
- Bayeen, R. H. & Renouf, A. (1996). Chronicling The Times: Productive lexical innovations in an English newspaper. *Language*, 72(1), 69-76.
- Bencini, G. M., & Goldberg, A. E. (2000). The contribution of argument structure constructions to sentence meaning. *Journal of Memory and Language*, 43(4), 640-651.
- Benigni, V., & Masini, F. (2010). Nomi sintagmatici in russo. Studi slavistici, 7(1), 145-172.
- Bergs, A., & Diewald, G. (a cura di). (2008). *Constructions and language change* (Vol. 194). Berlin: Walter de Gruyter.
- Bisetto, A. (1999). Note sui composti VN dell'italiano. In Benincà, P., Vanelli, L., & Mioni, A. (a cura di), Fonologia e morfologia dell'italiano e dei dialetti d'Italia: atti del 31. Congresso della Società di linguistica italiana, Padova, 25-27 settembre 1997. Roma: Bulzoni, 505-38.
- Bisetto, A. (2004). Composizione con elementi italiani. In Grossmann, M. & Rainer, R. (a cura di), *La formazione delle parole in italiano*. Tübingen: Niemeyer, 33-51.
- Bisetto, A., & Melloni, C. (2008). Parasynthetic compounding. Lingue e linguaggio, 7(2), 233-260.

- Bisetto, A., & Scalise, S. (2005). The classification of compounds. *Lingue e linguaggio*, 4(2), 319-20.
- Blevins, J. (2006). Word-based morphology. *Journal of Linguistics*, 42, 531–573.
- Blevins, J., Ackerman, F. & Robert Malouf, R. (in stampa). Word-and-Paradigm Morphology. In Audring, J. & Masini, F. (a cura di), *The Oxford Handbook of Morphological Theory*. Oxford: Oxford University Press.
- Bloomfield, L. (1933). Language. New York: Holt.
- Boas, H. C., & Sag, I. A. (2012). *Sign-based construction grammar*. CSLI Publications/Center for the Study of Language and Information.
- Bonami, O., Boyé, G., Dal, G., Giraudo, H., & Namer, F. (2018). *The lexeme in descriptive and theoretical morphology*. Language Science Press (http://langscipress.org/catalog/book/165).
- Booij, G. (2002a). Constructional idioms, morphology, and the Dutch lexicon. *Journal of Germanic linguistics*, *14*(4), 301-329.
- Booij, G. (2002b). Separable complex verbs in Dutch: A case of periphrastic word formation. *Verb-particle explorations*, 21-41.
- Booij, G. (2005a). Compounding and derivation: Evidence for construction morphology. In Dressler, W. U., Rainer, F., Kastovsky, D. & Pfeiffer, O. (a cura di), *Morphology and its demarcations*. Amsterdam: John Benjamins, 109–132.
- Booij, G. (2005b). Construction-dependent morphology. *Lingue e Linguaggio*, 4, 163–178.
- Booij, G. (2008a). Composition et morphologie des constructions. In Amiot, D. (a cura di), *La composition dans une perspective typologique*, 49-74.
- Booij, G. (2008b). Constructional idioms as products of language change: The *aan het* + INFINITIVE construction in Dutch. In Bergs, A. & Diewald, G. (a cura di), *Construction grammar and language change* (pp. 79–104). Berlin: Mouton de Gruyter.
- Booij, G. (2009a). Lexical integrity as a formal universal: a constructionist view. In Scalise, S., Magni, E. & Bisetto, A. (a cura di) *Universals of language today*. Dordrecht: Springer, 83-100.
- Booij, G. (2009b). Phrasal names: A constructionist analysis. Word Structure, 3, 219-240.
- Booij, G. (2010a). Construction morphology. *Language and Linguistics Compass*, 4, 543–555.
- Booij, G. (2010b). Construction morphology. Oxford: Oxford University Press.
- Booij, G. (2010c). Compound construction: Schemas or analogy? A construction morphology perspective. In Scalise, S. & Vogel, I. (a cura di), *Cross-disciplinary issues in compounding*, 93-108.
- Booij, G. (2013). Morphology in CxG. In Hoffmann, T. & Trousdale, G. (a cura di), *The Oxford handbook of construction grammar*. Oxford: Oxford University Press, 255-273.
- Booij, G. (2014). Morphology: the structure of words. In Allan, K. (a cura di), *The Routledge Handbook of Linguistics*. London & New York: Routledge, 104-117.
- Booij, G. (2015a). The nominalization of Dutch particle verbs: Schema unification and second order schemas. *Nederlandse Taalkunde*, *20*, 285–314.
- Booij, G. (2015b). Word formation in Construction Grammar. In Müller, P. O., Ohnheiser, I., Olsen, S. & Rainer, F. (a cura di), *Word formation. An international handbook of the languages of Europe*. Berlin: De Gruyter Mouton, 188-202.
- Booij, G. (2016). Construction morphology. In Hippisley, A. & Stump, G. T. (a cura di), *The Cambridge handbook of morphological theory*. Cambridge, U.K.: Cambridge University Press, 245-275.

- Booij, G. (2017). Inheritance and motivation in construction morphology. In Gisborne, N. & Hippisley, A. (a cura di), *Default inheritance*. Oxford: Oxford University Press.
- Booij, G. & Hüning, M. (2014). Affixoids and constructional idioms. In Boogaart, R., Colleman, T. & Rutten, G. (a cura di), *Extending the scope of Construction Grammar*. Berlin: de Gruyter, 77-1.
- Booij, G., & Audring, J. (2016). Construction morphology and the Parallel Architecture of grammar. *Cognitive Science*, 41 (S2), 277-302.
- Booij, G., & Hüning, M. (2014). Affixoids and constructional idioms. In Boogaart, R., Colleman, T. & Rutten, G. (a cura di), *Extending the scope of construction grammar*. Berlin: De Gruyter Mouton, 77-105.
- Booij, G., & Masini, F. (2015). The role of second order schemas in the construction of complex words. In *Semantics of complex words*. Springer, Cham, 47-66.
- Bosque, I. (1989). Constricciones morfológicas sobre la coordinación. *Lingüística Española Actual*, 9, 83-100.
- Bouillon, P., Ježek, E., Melloni, C., & Picton, A. (2012). Annotating qualia relations in Italian and French complex nominals. In Calzolari, N., Choukri, K., Declerck, T., Uğur Doğan, M., Maegaard, B., Mariani, J., Odijk, J. & Piperidis, S. (a cura di), *The eighth international conference on Language Resources and Evaluation (LREC)*, 1527-1532.
- Brinton, L. J., & Traugott, E. C. (2005). *Lexicalization and language change*. Cambridge University Press.
- Brucale, L. (2012). Latin compounds. *Probus*, 24, 93-117.
- Brucale, L., & Mocciaro, E. (2016). Composizione verbale in latino: il caso dei verbi in facio, fico. *LATINITATIS RATIONES: Descriptive and Historical Accounts for the Latin Language*, 279-297.
- Bybee, J. (2010). Language, usage and cognition. Cambridge: Cambridge University Press.
- Bybee, J. L. (1985). *Morphology: A study of the relation between meaning and form* (Vol. 9). John Benjamins Publishing.
- Celli, F., & Nissim, M. (2009). Automatic identification of semantic relations in Italian complex nominals. In *Proceedings of the Eighth International Conference on Computational Semantics*. Association for Computational Linguistics, 45-60.
- Chiari, I. & De Mauro, T. (2012). The new basic vocabulary of Italian: problems and methods. *Rivista di statistica applicata/Italian Journal of Applied Statistics*, 22, 21-35.
- Corbin, D. (1987). *Morphologie dérivationnelle et structuration du lexique* (Vol. 193). Berlin: Walter de Gruyter.
- Croft, W. (2001). *Radical construction grammar: Syntactic theory in typological perspective*. Oxford: Oxford University Press.
- D'Achille, P. & Grossmann, M. (2013). I composti «colorati» in italiano tra passato e presente. In Casanova Herrero, E. & Calvo Rigual, C. (a cura di), *Actas del XXVI Congreso Internacionalde. Lingüística y de Filología Románicas (Valencia, 6 11 septiembre 2010)*, III, Berlin: De Gruyter, 523-537.
- D'Achille, P. & Grossmann, M. (2009). *Stabilità e instabilità dei composti aggettivo+ aggettivo in italiano*. Roma: Bulzoni Editore, 143-171.
- D'Achille, P. & Grossmann, M. (2017). Italian colour terms in the BROWN area. Synchrony and diachrony. *Lingua e Stile*, *52*(1), 87-118.
- Dardano, M. (1978). La formazione delle parole nell'italiano di oggi (Vol. 148). Roma: Bulzoni.
- Dardano, M. (2009). Costruire parole: la morfologia derivativa dell'italiano. Bologna: Il Mulino.

- Darmesteter, A. (1877). De la création actuelle de mots nouveaux dans la langue française: et des lois qui la régissent. F. Viewieg, Librairie-Éditeur.
- Davies, M. (2002). Un corpus anotado de 100.000. 000 palabras del español histórico y moderno. *Procesamiento del lenguaje natural*, 29(1), 27.
- Davies, M. (2009). Creating useful historical corpora: A comparison of CORDE, the Corpus del español, and the Corpus do português. *Diacronía de las lenguas iberorrománicas: Nuevas aportaciones desde la lingüística de corpus*, 137-166.
- De Jong IV, N. H., Schreuder, R., & Harald Baayen, R. (2000). The morphological family size effect and morphology. *Language and cognitive processes*, *15*(4-5), 329-365.
- De Mauro, T. (2009). Grande dizionario dell'uso della lingua italiana. Torino: Utet.
- Devoto, G. & Oli, G. C. (2014). *Il Devoto Oli. Vocabolario della lingua italiana*, a cura di Luca Serianni e M. Trifone, Firenze, Le Monnier.
- Di Sciullo, A. M., & Ralli, A. (1994). Argument structure and inflection in compounds: some differences between English, Italian and Modern Greek. In *Proceedings of the workshop on compound nouns: Multilingual aspects of nominal composition. Geneva: ISSCO*, 61-76.
- Downing, P. (1977). On the creation and use of English compound nouns. *Language*, 810-842.
- Dressler, W. U., & Barbaresi, L. M. (1994). *Morphopragmatics: Diminutives and intensifiers in Italian, German, and other languages* (Vol. 76). Berlin: De Gruyter.
- Dressler, W. U., & Thornton, A. M. (1996). Italian nominal inflection. *Wiener Linguistische Gazette*, 55(57), 1-26.
- Dugas, E. (2014). The pragmatics of morphological negation: Pejorative and euphemistic uses of the prefix *non* in French. *Taikomoji kalbotyra*, 4, 1-21.
- Fabb, N. (1998). Compounding. In Spencer, A. & Zwicky, A. (a cura di), *The handbook of morphology*. Oxford: Oxford University Press, 66-83.
- Fábregas, A., & Masini, F. (2015). Prominence in morphology: the notion of head. *Lingue e linguaggio*, 14(1), 79-96.
- Favretti, R. R., Tamburini, F. & De Santis, C. (2002). CORIS/CODIS: A corpus of written Italian based on a defined and a dynamic model, In Wilson, A., Rayson, P. & McEnery, T. (a cura di), *A Rainbow of Corpora: Corpus Linguistics and the Languages of the World*, Munich: Lincom-Europa.
- Fillmore, C. J. (2013). Berkeley construction grammar. In Hoffmann, T. & Trousdale, G. (a cura di). *The Oxford handbook of construction grammar*. Oxford: Oxford University Press.
- Fillmore, C. J., Kay, P., & O'connor, M. C. (1988). Regularity and idiomaticity in grammatical constructions: The case of let alone. *Language*, 501-538.
- Fleischer, W. (1969). *Wortbildung der Deutschen Gegenwartsprache*. Leipzig: VEB Bibliographisches Institut.
- Forcellini, E. (1940). Lexicon totius latinitas. Padova: Typis Seminarii.
- Fried, M. (2009). Construction grammar as a tool for diachronic analysis. *Constructions and Frames*, 1(2), 261-290.
- Fruyt, M. (2002). Constraints and productivity in Latin nominal compounding. *Transactions of the Philological Society*, *100*(3), 259-287.
- Gaeta, L. (2004). Suffissi non produttivi. In Grossmann, M. & Rainer, F. (a cura di), *La formazione delle parole in italiano*. Tübingen: Niemeyer, 349-351.
- Gaeta, L. (2008). Constituent order in compounds and syntax: typology and diachrony. *Morphology*, 18(2), 117-141.
- Georges K. E. & H. Georges, Ausführliches Lateinisch-Deutsches Handworterbuch, Hannover: Hahn.

- Giacalone Ramat, A. (2015). Variazione sincronica e mutamento diacronico: il caso di alcuni connettori dell'italiano. In Engwall, G. & Fant, L. (a cura di), *Festival Romanistica*, Stockholm: Stockholm University Press, 13-36.
- Girju, R. (2007a). Improving the interpretation of noun phrases with cross-linguistic information. In *Proceedings of the 45th annual meeting of the association of computational linguistics*, 568-575.
- Girju, R. (2007b). Experiments with an annotation scheme for a knowledge-rich noun phrase interpretation system. In *Proceedings of the Linguistic Annotation Workshop*. Association for Computational Linguistics, 168-175.
- Girju, R., Moldovan, D., Tatu, M., & Antohe, D. (2005). On the semantics of noun compounds. *Computer speech & language*, 19(4), 479-496.
- Glare, G. W. (1982). Oxford Latin Dictionary. Oxford: Oxford University Press.
- Goldberg, A. (2003). Constructions: A new theoretical approach to language. *Trends in Cognitive Science*, 7, 219–224
- Goldberg, A. (2006). *Constructions at work. The nature of generalization in language*. Oxford: Oxford University Press.
- Goldberg, A. E. (1995). *Constructions: A construction grammar approach to argument structure*. Chicago: University of Chicago Press.
- Goldberg, A. E. (2013). Constructionist approaches. In Hoffmann, T. & Trousdale, G. (a cura di), *The Oxford handbook of construction grammar*, Oxford: Oxford University Press, 15-31.
- Gradenwitz, O. (1904). Laterali Vocum Latinarum. Leipzig: Hirzel.
- Grandi, N. (2002). *Morfologie in contatto: Le costruzioni valutative nelle lingue del Mediterraneo* (Vol. 33). Milano: Franco Angeli.
- Grandi, N. (2006). Considerazioni sulla definizione e la classificazione dei composti. *Annali dell'Università di Ferrara*. Sezione di Lettere online 1, 31-52.
- Grandi, N. (2011). La coordinazione tra morfologia e sintassi: tendenze tipologiche ed areali. In Massariello Merzagora, G & Dal Maso, S. (a cura di). *I luoghi della traduzione. Atti del XLIII Congresso internazionale di Studi della Società di Linguistica Italiana*. Roma: Bulzoni, 881-895.
- Grandi, N. (2017). Evaluatives in Morphology. In Oxford Research Encyclopedia of Linguistics.
- Grandi, N. & Körtvélyessy, L. (a cura di) (2015). *The Edinburgh handbook of evaluative morphology*. Edinburgh: Edinburgh University Press.
- Grandi, N., & Montermini, F. (2003). Valutativi suffissali e valutativi prefissali: un'unica categoria? In Grossmann, M. & Thornton, A. M. (a cura di), *La formazione delle parole, Atti del XXXVII Congresso Internazionale di Studi della Società di Linguistica Italiana*. Roma: Bulzoni, 271-287.
- Grandi, N., Nissim, M. & Tamburini, F. (2011). Noun-clad adjectives. On the adjectival status of non-head constituents of Italian attributive compounds. *Lingue e linguaggio*, 10(1), 161-70.
- Gries, S. T. (2013). Data in construction grammar. In T. Hoffmann & G. Trousdale (a cura di), *The Oxford handbook of construction grammar*. Oxford: Oxford University Press, 93-108.
- Gries, S. T., Hampe, B., & Schönefeld, D. (2005). Converging evidence: Bringing together experimental and corpus data on the association of verbs and constructions. *Cognitive linguistics*, 16 (4), 635-76.
- Grossmann, M. (2012). Romanian compounds. *Probus International Journal of Latin and Romance*, 147-173.
- Grossmann, M. & Rainer, F. (a cura di) (2004). *La formazione delle parole in italiano*. Tübingen: Niemeyer.

- Grossmann, M., & D'Achille, P. (2016). Italian colour terms in the BLUE area: synchrony and diachrony. In Grossmann, M., Dworkin, S. N., Strudsholm, E., Stala, E., Dębowiak, P., & Marello, C. (a cura di), *Colour and colour naming: crosslinguistic approaches* (Vol. 5). Centro de Linguística da Universidade de Lisboa/Universidade de Aveiro.
- Grossmann, M., & D'Achille, P. (2017). *Per la storia della formazione delle parole in italiano: un nuovo corpus in rete (MIDIA) e nuove prospettive di studio.* Firenze: Franco Cesati Editore.
- Grossmann, M., & Rainer, F. (2009). Italian adjective-adjective compounds: Between morphology and syntax. *Italian Journal of Linguistics*, *21*(1), 71-96.
- Guevara, E., & Scalise, S. (2009). Searching for universals in compounding. In Scalise, S., Magni, E. & Bisetto, A. (a cura di). *Universals of language today*, Dordrecht: Springer, 101-128
- Guiraud, P. (1954). *Les Charactères Statistiques du Vocabulaire. Essai de méthodologie.* Presses Uni- versitaires de France, Paris.
- Halle, M., & Marantz, A. (1993). Distributed morphology and the pieces of inflection. In Hale, K. & Keyser, J. (a cura di), The View from Building 20. Cambridge: MA MIT Press, 111-176.
- Hilpert, M. (2013a). Corpus-based approaches to constructional change. In T. Hoffmann & G. Tousdale (a cura di), *The Oxford handbook of construction grammar*. Oxford: Oxford University Press, 458-476.
- Hilpert, M. (2013b). *Constructional change in English: Developments in allomorphy, word formation, and syntax*. Cambridge: Cambridge University Press.
- Hilpert, M. (2015). From hand-carved to computer-based: Noun-participle compounding and the upward strengthening hypothesis. *Cognitive Linguistics*, 26(1), 1-36.
- Hilpert, M., & Gries, S. T. (2008). Assessing frequency changes in multistage diachronic corpora: Applications for historical corpus linguistics and the study of language acquisition. *Literary and Linguistic Computing*, 24(4), 385-401.
- Hilpert, M., & Gries, S. T. (2016). Quantitative approaches to diachronic corpus linguistics. *The Cambridge Handbook of English Historical Linguistics*, 36-53.
- Hockett, C. F. (1954). Two models of grammatical description. Word, 10(2-3), 210-234.
- Hopper, P. J. (1991). On some principles of grammaticization. In E.C. Traugott & B Hiene (a cura di), *Approaches to grammaticalization*, Vol. 1. Amsterdam/Philadelphia: John Benjamins, 17-35
- Hopper, P. J. & Traugott, E. (1993). *Grammaticalization*. Cambridge: Cambridge University Press. Hopper, P. J. & Traugott, E. C. (2003). *Grammaticalization*. Cambridge: Cambridge University Press.
- Hüning, M., & Booij, G. (2014). From compounding to derivation. The emergence of derivational affixes through "constructionalization". *Folia Linguistica*, 48(2), 579-604.
- Hüning, M., & Schlücker, B. (2015). Multi-word expressions. In Muller, P. O., Ohnheiser, I., Olsen,S. & Rainer, F. (a cura di), Wordformation. An international handbook of the languages of Europe. Berlin: De Gruyter, 450-467.
- Iacobini, C. (1992). *La prefissazione nell'italiano contemporaneo*. Tesi di dottorato, Università di Roma La Sapienza.
- Iacobini, C. (1999). I prefissi dell'italiano. In Benincà, P., Mioni, A. M. & Vanelli, L. (a cura di), Fonologia e morfologia dell'italiano e dei dialetti d'Italia. Atti del XXXI Congresso della Societa di Linguistica Italiana, Roma: Bulzoni, 369-99.
- Iacobini, C. (2004a). Composizione con elementi neoclassici. In Grossmann, M. & Rainer, F. (a cura di), *La formazione delle parole in italiano*. Tübingen: Max Niemeyer Verlag, 69-95.

- Iacobini, C. (2004b). Prefissazione. In Grossmann, M. & Rainer, F. (a cura di), *La formazione delle parole in italiano*. Tübingen: Max Niemeyer Verlag, 97-163.
- Iacobini, C. (2010). Elementi formativi. In Enciclopedia dell'italiano. Roma: Treccani.
- Iacobini, C. (2013). Il tipo videoregistratore: da retroformazione all'emergere di un nuovo processo compositivo fondato su di un vecchio modello romanzo. In Born, J. & Pöckl, W. (a cura di), Wenn die Ränder ins Zentrum drängen... Außenseiter in der Wortbildung (sforschung), Berlin: Frank & Timme, 189-212.
- Iacobini, C. (2014). A constructionist account of the emergence of a new lexeme-formation pattern: Italian right-headed verbal compounds. In Augendre, S., Couasnon-Torlois, G., Lebon, D., Michard, C., Boyé, G., & Montermini, F. (a cura di), *Proceedings of the Décembrettes 8th International conference on morphology (December 6-7, 2012)*, CNRS & Université Toulouse Jean Jaurès, ( = Carnets de grammaire, 22), 189-204.
- Iacobini, C. (2015). Produttività morfologica tra sistema e norma. In Bernal, E. & DeCesaris, J. ( cura di), *Los afijos: variación, rivalidad y representación*. Barcellona: Institut Universitari de Lingüística Aplicada Universitat Pompeu Fabra, 65-79.
- Iacobini, C., & Giuliani, A. (2010). A multidimensional approach to the classification of combining forms. *Italian journal of linguistics*, *22*(2), 287-316.
- Iacobini, C., & Masini, F. (2007). The emergence of verb-particle constructions in Italian: locative and actional meanings. *Morphology*, *16*(2), 155-188.
- Iacobini, C., & Thornton, A. M. (1992). Tendenze nella formazione delle parole nell'italiano del ventesimo secolo. In Moretti, B., Petrini, D. & Bianconi, S. (a cura di), *Linee di tendenza dell'italiano contemporaneo*, Atti del XXV Congresso della Società di Linguistica Italiana, Roma: Bulzoni, 25-55.
- Iacobini, C., & Thornton, A. M. (2016). Morfologia e formazione delle parole. In Lubello, S. (a cura di), *Manuale di linguistica italiana*, 13, Berlino: De Gruyter, 123-154.
- Jackendoff, R. (1975). Morphological and semantic regularities in the lexicon. *Language*, 639-671.
- Jackendoff, R. (1990). Semantic structures (Vol. 18). Cambridge: MIT press.
- Jackendoff, R. (1997). The architecture of the language faculty (No. 28). Cambridge: MIT Press.
- Jackendoff, R. (2002). *Foundations of language: Brain, Meaning, Grammar, Evolution*. Oxford: Oxford University Press.
- Jackendoff, R. (2011). What is the human language faculty?: Two views. *Language*, 87(3), 586-624.
- Jackendoff, R. (2013). Constructions in the parallel architecture. In Hoffmann, T. & Trousdale, G. (a cura di), *The Oxford Handbook of Construction Grammar*. Oxford: Oxford University Press.
- Jackendoff, R., & Audring, J. (2018). Morphology in the Parallel Architecture. In Audring, J. & Masini, F. (in stampa) (a cura di), *The Oxford Handbook of Morphological Theory*. Oxford: Oxford University Press.
- Jakubíček, M., Kilgarriff A., Kovář V.V., Rychlý, P. & Suchomel, V. (2013). The TenTen corpus family. In *7th International Corpus Linguistics Conference*, Lancaster, 125-127
- Ježek, E. (2016). The Lexicon: An introduction. Oxford: Oxford University Press.
- Johnston, M., & Busa, F. (1999). Qualia structure and the compositional interpretation of compounds. In *Breadth and depth of semantic lexicons*. Dordrecht: Springer, 167-187.
- Kilgarriff, A., Baisa, V., Bušta, J., Jakubíček, M., Kovář, V., Michelfeit, J., ... & Suchomel, V. (2014). The Sketch Engine: ten years on. *Lexicography*, 1(1), 7-36.

- Klingebiel, K. (1989). *Noun+verb compounding in Western Romance* (Vol. 113). Berkeley: University of California Press.
- Kytö, M. & Rissanen, M. (1992). A language in transition: The Helsinki Corpus of English Texts. ICAME Journal, 16, 7-27.
- Langacker, R. W. (1987). *Foundations of cognitive grammar: Theoretical prerequisites* (Vol. 1). Stanford: Stanford University Press.
- Lauer, M. (1995). Corpus statistics meet the noun compound: some empirical results, In *Proceedings of the 33rd annual meeting on Association for Computational Linguistics*, Association for Computational Linguistics, 47-54.
- Lehmann, C. (1985). The role of grammaticalization in linguistic typology. *Language invariants* and mental operations, 41-52.
- Lehmann, C. (1995). Thoughts on grammaticalization. München: Lincom Europa.
- Levi, J. N. (1978). The syntax and semantics of complex nominals. New York: Academic Press.
- Lieber, R. (1992). *Deconstructing morphology: Word formation in syntactic theory*. Chicago: University of Chicago Press.
- Litta, E., Passarotti, M., & Culy, C. (2016). Formatio formosa est. Building a Word Formation Lexicon for Latin. In *Proceedings of the Third Italian Conference on Computational Linguistics* (CLiC-it 2016), 185-189.
- Lo Duca, M. (2004a). Nomi di strumento. In Grossmann, M. & Rainer, F. (a cura di), *La formazione delle parole in italiano*. Tübingen: Niemeyer, 227-233.
- Lo Duca, M. (2004b). Nomi di agente. In Grossmann, M. & Rainer, F. (a cura di), *La formazione delle parole in italiano*. Tübingen: Niemeyer, 351-363.
- Lombardi Vallauri, E. (2008). Composti intitolativi in italiano: un'oscillazione. In Cresti, E. (a cura di), *Prospettive nello studio del lessico italiano. Atti del IX congresso della Società Internazionale di Linguistica e Filologia Italiana (Firenze, 14-17 giugno 2006*), Firenze: Firenze University Press, 3 voll., vol. 2°, 555-562.
- Magni, E. (2010). From the periphery to the core of Romance [VN] compounds. *Lingue e linguaggio*, 9(1), 3-40.
- Marchand, H. (1969). The categories and types of present-day English word-formation: A synchronic-diachronic approach. Beck.
- Marchello-Nizia, C. (2006). *Grammaticalisation et changement linguistique*. De Boeck: Duculot.
- Marello, C., & Onesti, C. (2016). Colourful microstructures: How Italian dictionaries see colour terms. In Silvestre, J. P., Cardeira, E. & Villalva, A. (a cura di), *COLOUR AND COLOUR NAMING: CROSSLINGUISTIC APPROACHES*, 91-104.
- Masini, F. (2008). Binomi coordinati in italiano. In Cresti, E. (a cura di), *Prospettive nello studio del lessico italiano. Atti SILFI 2006*. Firenze: Firenze University Press, 563-571.
- Masini, F. (2009). Phrasal lexemes, compounds and phrases: A constructionist perspective. *Word structure*, *2*(2), 254-271.
- Masini, F. (2012). Parole sintagmatiche in italiano. Roma: Caissa Italia.
- Masini, F. (2016). Grammatica delle Costruzioni. Un'introduzione. Roma: Carocci.
- Masini, F. (2017). Polirematiche 'di colore' in italiano: uno studio quantitativo. In Iannaccaro, G., D'Alessandro, R., Passino, D. & Thornton, A. M. (a cura di), *Di tutti i colori. Studi linguistici per Maria Grossmann*, Utrecht: Utrecht University, 203-215.
- Masini, F. (in stampa). Competition between morphological words and multiword expressions. In Gardani, F., Luschützky, H. C., Rainer, F. & Dressler, W. (a cura di), *Competition in Inflection and Word formation*. Berlin: Springer.

- Masini, F. & Audring, J. (in stampa). Construction Morphology. In Audring, J. & Masini, F. (a cura di), *The Oxford Handbook of Morphological Theory*. Oxford: Oxford University Press.
- Masini, F., & Benigni, V. (2012). Phrasal lexemes and shortening strategies in Russian: the case for constructions. *Morphology*, 22(3), 417-451.
- Masini, F., & Scalise, S. (2012). Italian compounds. Probus, 61-91.
- Masini, F., & Thornton, A. (2007). Italian VeV lexical constructions. In *Mediterranean Morphology Meetings*, Vol. 6, 148-189.
- Matthews, P. H. (1972). *Inflectional morphology: A theoretical study based on aspects of Latin verb conjugation* (Vol. 6). Cambridge: CUP Archive.
- Mauri, C., & Giacalone Ramat, A. (2012). The development of Italian adversative connectives: factors at play and stages of grammaticalization. *Linguistics*, *50*(2), 191-239.
- Meyer-Lübke, W. (1890). Italienische Grammatik (Vol. 2). Leipzig.
- Micheli, M. S. (2016a). Limiti e potenzialità dell'uso di dati empirici in lessicografia: il caso del plurale delle parole composte, in Jezek, E. & Chiari, I. (a cura di), *Dati empirici e risorse lessicali*, RiCognizioni, vol. 3, n. 6, 15-33.
- Micheli, M. S. (2016b). Sul plurale delle parole composte nell'italiano contemporaneo. *Studi di Lessicografia italiana*, 33, 229-256.
- Micheli, M. S. (2018). Number inflection in AN and NA Italian compounds. In Koutsoukos, N., Audring, J. & Masini, F. (a cura di), *Morphological variation: synchrony and diachrony Proceedings of the 11th Mediterranean Morphology Meeting*, 51-62.
- Migliorini, B. (1963). I prefissoidi (il tipo 'aeromobile, radiodiffusione'). In Id., *Saggi sulla lingua del Novecento*. Firenze: Sansoni, 9-60.
- Mithun, M. (1984). The evolution of noun incorporation. Language, 60(4), 847-894.
- Mithun, M. (1986). On the nature of noun incorporation. Language, 62(1), 32-37.
- Montermini, F. (2003). Prefissi vs. preposizioni nell'italiano di oggi. In Maraschio, N. & Poggi Salani, T. (a cura di), *Italia linguistica anno Mille, Italia linguistica anno Duemila. Atti del XXXIV Congresso Internazionale di studi della Società di Linguistica Italiana*, 635-647.
- Montermini, F. (2008). *Il lato sinistro della morfologia. La prefissazione in italiano e nelle lingue del mondo.* Milano: Franco Angeli.
- Montermini (2010). Units in compounding. In Scalise, S. & Vogel, I. (a cura di), *Cross-disciplinary issues in compounding*. Amsterdam/Philadelphia: John Benjamins, 77-92.
- Mosti, R. (2015). Il lemmario del *Tesoro della Lingua Italiana delle Origini. Bollettino dell'Opera del Vocabolario Italiano*, XIX-XX, 405-425.
- Moyna, M. I. (2011). *Compound words in Spanish. Theory and history*. Amsterdam / Philadelphia: John Benjamins.
- Mutz, K. (2015). Evaluative morphology in a diachronic perspective. In Grandi, N. & Körtvélyessy, L. (a cura di), *Edinburgh handbook of evaluative morphology*, 142-154.
- Nielsen Whitehead, B. (2012). *Pickpocket compounds from Latin to Romance*. Doctoral dissertation, Faculty of the Humanities, Leiden University.
- Norde, M. (2009). *Degrammaticalization*. Oxford: Oxford University Press.
- Norde, M., & Van Goethem, K. (2014). Bleaching, productivity and debonding of prefixoids: A corpus-based analysis of 'giant' in German and Swedish. *Lingvisticae Investigationes*, 37(2), 256-274.
- Olsen, S. (2001). Copulative compounds: a closer look at the interface between syntax and morphology. In *Yearbook of Morphology 2000*, Dordrecht: Springer, 279-320.

- Ó Séaghdha, D. (2007). Annotating and learning compound noun semantics. In *Proceedings of the* 45th Annual Meeting of the ACL: Student Research Workshop. Association for Computational Linguistics, 73-78.
- Ó Séaghdha, D., & Copestake, A. (2008). Semantic classification with distributional kernels. In *Proceedings of the 22nd International Conference on Computational Linguistics*-Volume 1. Association for Computational Linguistics, 649-656.
- Oniga, R. (1992). Compounding in Latin. *Rivista di linguistica*, 4(1), 97-116.
- Pagliaro, A. (1930). Sommario di linguistica arioeuropea. Roma: L'Universale.
- Pinker, S. (2015). Words and rules: The ingredients of language. Basic Books.
- Plag, I. (2004). Syntactic category information and the semantics of derivational morphological rules. *Folia Linguistica*, *38*(3-4), 193-226.
- Prati, A. (1931). Composti imperativi quali casati e soprannomi. *Revue de Linguistique Romane*, 7, 250-264.
- Pulvermüller, F., Cappelle, B., & Shtyrov, Y. (2013). Brain basis of meaning, words, constructions, and grammar. In *The Oxford handbook of construction grammar*.
- Pustejovsky, J. (1995). The Generative Lexicon. Cambridge MA: MIT Press.
- Pustejovsky, J., Rumshisky, A., Moszkowicz, J. L., & Batiukova, O. (2008). GLML: A Generative Lexicon markup language. In *Proceedings of the Generative Lexicon Workshop*, Instituto di Linguistica Computazionale (CNR), Pisa, Italy.
- Radimský, J. (2015a). *Noun + noun compounds in Italian: A corpus-based study* (Vol. 3). University of South Bohemia.
- Radimský, J. (2015b). Les composés de coordination en italien: esquisse d'une typologie. *Studia Romanica Posnaniensia*, 42(1), 97-111.
- Radimský, J. (2016). I composti N-N attributivi nel corpus ItWac.
- Ralli, A., & Dimela, E. (2007). On the borderline between prefixation and compounding: The saadverbs in the dialectal varieties of Lesbos, Kydonies (Aivali) and Moschonisia. In *Proceedings of the 3rd international conference of Modern Greek dialects and linguistic* theory. Nicosia.
- Ramat, A. G., & Mauri, C. (2008). From cause to contrast: a study in semantic change. In *Studies on Grammaticalization*, Vol. 1. Berlin: Mouton de Gruyter, 303-321.
- Renzi, L. & Salvi, G. (2010). Italiano antico. In *Enciclopedia dell'Italiano* (<a href="http://www.treccani.it/enciclopedia/italiano-antico">http://www.treccani.it/enciclopedia/italiano-antico</a> (Enciclopedia-dell'Italiano)/ ultima visualizzazione in data 22/03/2018).
- Ricca, D. (2005). Al limite tra sintassi e morfologia: i composti aggettivali VN nell'italiano contemporaneo. In Grossmann, M. & Thornton, A. M. (a cura di), *La formazione di parole. Atti del XXXVII Congresso della SLI*. Roma: Bulzoni, 465–486.
- Ricca, D. (2010). Corpus data and theoretical implications: With special reference to Italian VN compounds. In Scalise, S. & Vogel, I. (a cura di), *Cross-disciplinary Issues in Compounding*. Amsterdam / Philadelphia: John Benjamins, 237-254.
- Ricca, D. (2015). Verb-noun compounds in Romance. In Müller, P. O., Ohnheiser, I., In Olsen, S., & Rainer, F. (a cura di), *Word-formation: an international handbook of the languages of Europe* (Vol. 40). Berlin: de Gruyter.
- Rosenberg, M. (2005). Agent nouns, productivity and diachrony: an analysis of [VN/A] N/A compounds and-eur derivations in French. In *Mediterranean Morphology Meeting*, Vol. 5, 359-378.

- Ruppenhofer, J., & Michaelis, L. A. (2010). A constructional account of genre-based argument omissions. *Constructions and Frames*, *2*(2), 158-184.
- Saracco, C. (2017). I composti possessivi dell'antico frisone. Linguistica e Filologia, 37, 155-188.
- Scalise, S. (1984). Morfologia lessicale. Padova: Clesp.
- Scalise, S. (1990). Constraints on the Italian suffix-mente. *Contemporary morphology*, 87-98.
- Scalise, S. & Bisetto, A. (2009). The classification of compounds. In Lieber, R. & Steckauer, P. (a cura di), *The Handbook of Compounding*. Oxford: Oxford University Press, 34-53.
- Scalise, S., & Guevara, E. (2006). Exocentric compounding in a typological framework. *Lingue e linguaggio*, (2), 185-206.
- Schlücker, B., & Plag, I. (2011). Compound or phrase? Analogy in naming. *Lingua*, 121(9), 1539-1551.
- Schmid, H. (1995). Treetagger: a language independent part-of-speech tagger. *Institut für Maschinelle Sprachverarbeitung, Universität Stuttgart, 43,* 28.
- Schreuder, R., & Baayen, R. H. (1997). How complex simplex words can be. *Journal of memory and language*, *37*(1), 118-139.
- Serianni, L. (1998). Lezioni di grammatica storica italiana (Vol. 350). Roma: Bulzoni.
- Simone, R. (1997). Esistono verbi sintagmatici in italiano?. *Lessico e grammatica. Teorie linguistiche e applicazioni lessicografiche.* Roma: Bulzoni, 155-170.
- Spencer, A. (1991). *Morphological theory: An introduction to word structure in generative grammar* (Vol. 2). Oxford: Basil Blackwell.
- Spina, S. (2014). Il Perugia *Corpus*: una risorsa di riferimento per l'italiano. Composizione, annotazione e valutazione. In Basili, R., Lenci, A. & Magnini, B. (a cura di), *Proceedings of the First Italian Conference on Computational Linguistics CLiC-it 2014*. vol. 1. Pisa: Pisa University Press, 354-359.
- Stefanowitsch, A., & Gries, S. T. (2003). Collostructions: Investigating the interaction of words and constructions. *International journal of corpus linguistics*, 8(2), 209-243.
- Štichauer, P. (2009a). Morphological productivity in diachrony: the case of the deverbal nouns in *-mento*, *-zione* and *-gione* in Old Italian from the 13th to the 16th century. In *Selected Proceedings of the 6th Décembrettes*, 138-147.
- Štichauer, P. (2009b). *La produttività morfologica in diacronia: i suffissi-mento,-zione e-gione in italiano antico dal Duecento al Cinquecento*. Prague: Karolinum Press.
- Štichauer, P. (2015a). La formazione delle parole in diacronia: studi di morfologia derivazionale dell'italiano tra il Cinquecento e l'Ottocento. Prague: Karolinum Press.
- Štichauer, P. (2015b). The meaning of Italian VN compounds in a diachronic perspective: the development of a word-formation rule and its semantic instruction. In *Mediterranean Morphology Meetings*, Vol. 9, 138-142.
- Štichauer, P. (2016). Verb-noun compounds in Italian from the 16th century onwards: an increasing exploitation of an available word-formation pattern. *Morphology*, 26(2), 109-131.
- Stoppelli, P., & Picchi, E. (2001). LIZ 4.0. *Letteratura italiana Zanichelli*.
- Taylor, J. R. (2012). *The mental corpus: How language is represented in the mind*. Oxford: Oxford University Press.
- Tekavčić, P. (1980a). Grammatica storica dell'italiano: Morfosintassi (Vol. 2). Bologna: Il Mulino.
- Tekavčić, P. (1980b). Grammatica storica dell'italiano: Lessico (Vol. 3). Bologna: Il Mulino.
- Ten Hacken, P. (2000). Derivation and compounding. *Morphology/Morphologie*, 349-359.

- Thornton, A. M. (2001): Some reflections on gender and inflectional class assignment in Italian. In Schaner-Wolles, C., Rennison, J. R. & Neubarth, F. (a cura di): *Naturally! Linguistic studies in honour of Wolfgang Ulrich Dressler presented on the occasion of his 60th birthday.* Torino: Rosenberg & Sellier, 479-487.
- Thornton, A. M. (2008). Italian Verb-Verb reduplicative action nouns. *Lingue e linguaggio*, 7(2), 209-232.
- Thornton, A. M. (2009). Italian verb reduplication between syntax and the lexicon. *Italian Journal of Linguistics*, *21*, 235-261.
- Thornton, A. M. (2010). Il tipo fuggifuggi. In M. Iliescu, H. Siller-Runggaldier & P. Danler (a cura di), Actes du XXVe CILPR (Congrès International de Linguistique et de Philologie Romanes), Innsbruck, 3 8 septembre 2007, Berlin: De Gruyter, 527-536.
- Tollemache, F. (1945). *Le parole composte nella lingua italiana*(Vol. 2). Edizioni Rores di N. Ruffolo.
- Tomasello, M. (2003). *Constructing a language: A usage-based approach to child language acquisition*. Harvard: Harvard University Press.
- Traugott, E. C. (1989). On the rise of epistemic meanings in English: An example of subjectification in semantic change. *Language*, 31-55.
- Traugott, E. C. (1995). Subjectification in grammaticalization. *Subjectivity and subjectivisation*, 31-54.
- Traugott, E. C. (2003). Constructions in grammaticalization. *The Encyclopedic Dictionary of Applied Linguistics: A Handbook for Language Teaching*, 624-647.
- Traugott, E. C. (2007). The concepts of constructional mismatch and type-shifting from the perspective of grammaticalization. *Cognitive Linguistics*, 18(4), 523-557.
- Traugott, E. C., & Trousdale, G. (2013). *Constructionalization and constructional changes* (Vol. 6). Oxford: Oxford University Press.
- Traugott, E. C., & Trousdale, G. (2010). (a cura di) *Gradience, gradualness and grammaticalization* (Vol. 90). John Benjamins Publishing.
- Trips, C. (2009). *Lexical Semantics and Diachronic Morphology: The Development of-hood, -dom and-ship in the History of English* (Vol. 527). Berlin: Walter de Gruyter.
- Trousdale, G. (2014). On the relationship between grammaticalization and constructionalization. *Folia Linguistica*, 48(2), 557-578.
- Van Goethem, K. (2008a). Oud-leerling versus ancien élève: A Comparative Study of Adjectives Grammaticalizing into Prefixes in Dutch and French. *Morphology*, 18(1), 27-49.
- Van Goethem, K. (2008b). The interaction between word structure and grammaticalization. Evidence from word-formation with French entre-and Dutch tussen. *Word Structure*, 1(1), 65-82.
- Van Goethem, K. (2009). Choosing between A+ N compounds and lexicalized A+N phrases: The position of French in comparison to Germanic languages. *Word Structure*, *2*(2), 241-253.
- Van Goethem, K. (2010). The French construction nouveau+ past participle revisited: Arguments in favour of a prefixoid analysis of nouveau. *Folia Linguistica*, 44(1), 163-178.
- Van Goethem, K. (2011). From adjective to affix in Dutch and French: The influence of word order patterns on grammaticalization. *Studies in Language. International Journal sponsored by the Foundation "Foundations of Language"*, 35(1), 194-216.
- Van Goethem, K., & Amiot, D. (2011). Un problème vraiment clé? Degrammaticalization of French N+ N compounds. In *Eighth Mediterranean Morphology Meeting*.

- Van Goethem, K., & De Smet, H. (2014). How nouns turn into adjectives: The emergence of new adjectives in French, English and Dutch through debonding processes. *Languages in Contrast*, 14(2), 251-277.
- Van Goethem, K., & Hiligsmann, P. (2014). When two paths converge: Debonding and clipping of Dutch reuze. *Journal of Germanic linguistics*, 26(1), 31-64.
- Van Goethem, K., & Hüning, M. (2014, September). Debonding of compounds in French, English, Dutch and German. In *MGDLT6: 6th International Conference on Modern Greek dialects and linguistic theory*.
- Van Marle, J. (1985). *On the paradigmatic dimension of morphological creativity* (Vol. 18). Foris Pubns USA.
- Voghera, M. (2004). Polirematiche. In Grossmann, M. & Rainer, F. (a cura di), *La formazione delle parole in italiano*, Tübingen: Niemeyer, 6-69.
- Warren, B. (1978). Semantic patterns of noun-noun compounds. Gothenburg: Studies in English.
- Zagona, K. T. (1990). Mente adverbs, compound interpretation and the pojection principle. *Probus* 2(1), 1-30.

## APPENDICE I: DATASET COMPOS\_ITA\_TLIO

Composto	Categorie di input
Composto acquaforte	N-A
acquivento	N-N
affettapane	V-N
agrodolce	A-A
alicorno	N-N
alititonante	AVV-PART.PRES
	AVV-PART.PRES
altipossente altisono	A-N
andirivieni	V-V
animavversione	N-N
arcamensa	N-N
arcobalestro	N-N
armipotente	N-PART.PRES
arretroguardia	AVV-N
artemagico	N-N
baciapòlvere	V-N
barbagianni	N-N
batticuore	V-N
battigote	V-N
battilana	V-N
battipetto	V-N
battisacco	V-N
battisfancello	V-N
battisoffia	V-V
beccalite	V-N
beccamorto	V-N
becchipuzzola	N-N
bellegote	A-N
benavventuroso	AVV-A
bencreato	AVV-PART.PASS
bendisposto	AVV-PART.PASS
beneamato	AVV-PART.PASS
benefacente	AVV-PART.PRES
benefatto	AVV-PART.PASS
benefattore	AVV-N
benemerito	AVV-PART.PASS
benentrata	AVV-N
beneplacimento	AVV-N
beneplacito	AVV-N
benesonante	AVV-PART.PRES
benessere	AVV-N
benevischio	AVV-N
benevolente	<b>AVV-PART.PRES</b>
benfatto	AVV-PART.PASS
benformato	AVV-PART.PASS
benparlante	AVV-PART.PRES
benpiacente	<b>AVV-PART.PRES</b>
benevolente benfatto benformato benparlante	AVV-PART.PRES AVV-PART.PASS AVV-PART.PASS AVV-PART.PRES

capomaestrato	N-N
capomaglio	N-N
capomastro	N-N
capomese	N-N
capomorbo	N-N
caponaballe	N-AVV
capopiede	N-N
capopopolo	N-N
capopurgio	N-N
caporicciare	N-V
caporiccio	N-A
caporione	N-N
caposalvo	N-A
caposcala	N-N
caposoldo	N-N
capoverso	N-N
capovolta	N-N
capricorno	N-N
caprifico	N-N
caprifoglio	N-N
casatorre	N-N
cassapanca	N-N
cassapedi	V-N
castigabriccone	V-N
castigapazzi	V-N
castrocampo	N-N
cavadenti	V-N
cavallivendolo	N-N
celtiberico	A-A
centofoglie	A-N
centogemino	A-N
centopezze	A-N
centopiedi	A-N
chiarefare	AVV-V
chiudiborsa	V-N
cinquefoglie	A-N
codacavallina	N-A
codimozzo	N-A
coprifuoco	V-N
corgaudente	N-PART.PRES
corpobalsamo	N-N
covriceffo	V-N
crepacuore	V-N
crocesegnare	N-V
curaorecchie	V-N
drizzacrine	V-N
duramadre	A-N
duramale	V-AVV
fattibello	V-N
fedecommesso	N-PART.PASS
fedecommissario	N-N
fededegno	N-A
· · · · · · · · · · · · · · · · · · ·	

ferrovecchio	N-A
fiammasalsa	N-A
ficosecco	N-A
fienogreco	N-A
figliuomini	N-N
finimondo	N-N
fioripanno	N-N
foraterra	V-N
formicaleone	N-N
fottivento	V-N
gallogrecese	A-A
gallogreci	A-A
gattaconiglio	N-N
gattafodero	N-N
gentildama	A-N
gentildonna	A-N
gentiluomo	A-N
giovanuomo	A-N
girasole	V-N
giravolta	V-V
gommaedera	N-N
gommarabica	N-A
grancane	A-N
granduca	A-N
granmangiere	A-N
grattacacio	V-N
guardabanco	V-N
guardacorpo	V-N
guardacuore	V-N
guardamorto	V-N
guardanappa	V-N
guardapasso	V-N
guardaroba	V-N
guardaspensa	V-N
guastamondo	V-N
guazzabuglio	V-N
lanovendolo	N-N
lavaceci	V-N
lavascodelle	V-N
leccamuffo	V-N
leccascodelle	V-N
leonaparda	N-N
leonepardo	N-N
lepretasso	N-N
libifenicio	A-A
linoseme	N-N
longanime	A-N
luogotenente	N-PART.PRES
madreselva	N-N
malacconciato	AVV-PART.PASS
malacconcio	AVV-A
malaccorto	AVV-A

malacquistato	AVV-PART.PASS
malacquisto	A-PART.PASS
maladagio	AVV-AVV
maladoperato	AVV-PART.PASS
malafama	A-N
malafede	A-N
malaffare	A-N
malagente	A-N
malaguida	A-N
malalbergo	A-N
malalingua	A-N
malamorte	A-N
malandanza	A-N
malanimo	A-N
malapasqua	A-N
malaugurato	AVV-PART.PASS
malaugurio	A-N
malaventura	A-N
malavoglia	A-N
malavoglienza	A-N
malavveduto	AVV-PART.PASS
malavventura	A-N
malavviato	AVV-PART.PASS
malbailire	AVV-V
malbailito	AVV-PART.PASS
malcalzato	AVV-PART.PASS
malcapitare	AVV-V
malconciato	AVV-PART.PASS
malconcio	AVV-PART.PASS
malconoscente	AVV-PART.PRES
malconsiglio	A-N
malcontento	AVV-A
malcornuto	AVV-A
malcostume	A-N
malcreato	AVV-PART.PASS
malcubato	AVV-PART.PASS
malcucinato	AVV-PART.PASS
maldesinato	AVV-PART.PASS
maldicente	AVV-PART.PRES
maldicenza	A-N
maldire	AVV-V
maldisposto	AVV-PART.PASS
maleabbiato	AVV-PART.PASS
maleavviato	AVV-PART.PASS
malebolge	A-N
malebranche	A-N
maledetto	AVV-PART.PASS
maledire	AVV-V
malefatta	AVV-PART.PASS
malessere	A-N
malfare	AVV-V
malfatto	AVV-PART.PASS

10	
malfattore	A-N
malfattoria	A-N
malfattura	A-N
malfermo	AVV-A
malferuto	AVV-PART.PASS
malfido	AVV-A
malformato	AVV-PART.PASS
malfornito	AVV-PART.PASS
malgoverno	A-N
malgradito	AVV-PART.PASS
malgrado	A-N
malguadagno	A-N
malincorpo	N-N
malincuore	N-N
malinfamare	AVV-V
malmenare	AVV-V
malmerito	A-N
malmesso	<b>AVV-PART.PASS</b>
malmettere	AVV-V
malmodo	A-N
malnato	<b>AVV-PART.PASS</b>
malora	A-N
malordinemente	AVV-AVV
malpaga	A-N
malparato	AVV-PART.PASS
malparlante	AVV-PART.PRES
malparlare	AVV-V
malparliere	A-N
malpartito	A-N
malpaziente	AVV-PART.PRES
malpensante	AVV-PART.PRES
malpensare	AVV-V
malsaccente	AVV-A
malsano	AVV-A
malsapore	A-N
malsembiante	AVV-PART.PRES
malservente	AVV-PART.PRES
maltagliato	AVV-PART.PASS
maltalento	A-N
maltempo	A-N
maltenuto	AVV-PART.PASS
maltolto	AVV-PART.PASS
maltrattamento	A-N
maltrattare	AVV-V
maltratto	AVV-V AVV-PART.PASS
	AVV-PART.PASS AVV-PART.PASS
maltraverso	
maltrovato	AVV-PART.PASS
maluomo	A-N
malusanza	A-N
maluso	A-N
malvedere	AVV-V
malveduto	AVV-PART.PASS

malventurato malventurato malvenuto malvenuto malvezzo malvizio malvoglianza malvoglianza malvolente malvolente malvolore malvolontà malvoluto malvoluto malvoluto manomettere manroverso manubalestro manzzafionda mazzafionda mazzafirusto mazzascudo melagrana melagrana melocotogna me		
malvenuto AVV-PART.PASS malvezzo A-N malvizio A-N malvoglianza A-N malvolente AVV-PART.PRES malvolere AVV-V malvoluto AVV-PART.PASS manomettere N-V manroverso N-N manubalestro N-N mappamondo N-N mazzafascio N-N mazzafionda N-N mazzafionda N-N mazzafrusto N-N mazzariusto N-N mazzariusto N-N melacotogna N-A melagrana N-PART.PASS melagranata N-PART.PASS melagranata N-PART.PASS meloctogno N-A melograno N-PART.PASS mercivendola N-N mesciroba V-N mettifuoco V-N mezzaboccia A-N mezzaluna A-N mezzaluna A-N mezzaluna A-N mezzanotte A-N mezzodio A-N mezzodio A-N mezzodio A-N mezzofiera A-N	malventura	A-N
malvezzo A-N malvizio A-N malvoglianza A-N malvoglianza A-N malvolente AVV-PART.PRES malvolere AVV-V malvolontà A-N malvoluto AVV-PART.PASS manomettere N-V mannowetso N-N manubalestro N-N mappamondo N-N mazzafascio N-N mazzafionda N-N mazzafascio N-N mazzafascio N-N mazzafrusto N-N mazzafrusto N-N mazzafrusto N-N melagrana N-PART.PASS melagrana N-PART.PASS melagranata N-PART.PASS meloctogno N-A melograno N-PART.PASS mercivendola N-N mesciroba V-N mettifuoco V-N mezzaboccia A-N mezzaluna A-N mezzaluna A-N mezzaluna A-N mezzocolore A-N mezzocolore A-N mezzodio A-N millefoglie A-N millefoglie A-N millefoglie A-N millefoglio A-N	malventurato	AVV-A
malvizio A-N malvoglianza A-N malvolente AVV-PART.PRES malvolere AVV-V malvolontà A-N malvoluto AVV-PART.PASS manomettere N-V manroverso N-N manubalestro N-N mappamondo N-N mazzafascio N-N mazzafionda N-N mazzafrusto N-N mazzafrusto N-N melacotogna N-A melagrana N-PART.PASS melagranata N-PART.PASS melagranata N-PART.PASS meloreta N-N mellomela N-N melocotogno N-A melograno N-PART.PASS mercivendola N-N mesciroba V-N mettifuoco V-N mettifuoco V-N mezzaboccia A-N mezzagrana A-N mezzaluna A-N mezzaluna A-N mezzocolore A-N mezzocolore A-N mezzodi A-N mezzofiera A-N mezzodi A-N mezzofiera A-N mezzodi A-N me	malvenuto	AVV-PART.PASS
malvoglianza malvolente malvolere malvolere malvolontà A-N malvoluto AVV-PART.PRES manomettere N-V manroverso N-N manubalestro N-N mappamondo N-N mazzafascio N-N mazzafionda M-N mazzafionda M-N mazzapicchio M-V mazzascudo N-N melacotogna melagrana N-PART.PASS melagranata N-PART.PASS melagranata N-PART.PASS melograno N-A mellomela N-N mellomela N-N mercivendola N-N mesciroba V-N mettifuoco M-PART.PASS metifuoco V-N metzzagrana A-N mezzalana M-N mezzalore A-N mezzalore A-N mezzalore A-N mezzalore A-N mezzolore A-N mezzoolore A-N mezzodio A-N mezzodio A-N mezzodio A-N mezzofiera A-N mezzofiera A-N mezzodio A-N mezzofiera A-N mezzodio M-N millefoglio A-N millefoglio A-N millefoglio A-N millefoglio A-N mordecascio W-N mordecascio W-N mordigallina N-N nobiluomo A-N nobiluomo A-N nobiluomo A-N nottetempo N-N occhibagliare N-V	malvezzo	A-N
malvolente malvolere malvolere malvolontà A-N malvoluto AVV-PART.PASS manomettere M-V manroverso N-N manubalestro mappamondo M-N mazzafascio M-N mazzafionda M-N mazzafionda M-N mazzafiondo mazzapicchio M-V mazzascudo M-N melacotogna M-A melagrana M-PART.PASS melagranata M-PART.PASS melarancia M-N mellomela M-N melocotogno M-A melograno M-PART.PASS mercivendola M-N mesciroba M-N mesziroba M-N mezzaboccia M-N mezzaluna M-N mezzaluna M-N mezzaluna M-N mezzaluna M-N mezzocolore M-N mezzoolore M-N mezzoolore M-N mezzoolore M-N mezzofiera M-N mobiluomo A-N mobiluomo N-N nobiluomo N-N N-N N-N N-N N-N N-N N-N N-N N-N N-	malvizio	A-N
malvolere malvolontà A-N malvoluto AVV-PART.PASS manomettere N-V manroverso N-N manubalestro mappamondo M-N mazzafascio N-N mazzafionda M-N mazzafrusto M-N mazzapicchio mazzascudo M-N melacotogna melagrana melagrana mellomela M-N mellomela M-N mesciroba M-PART.PASS mercivendola Metifuoco Metifuoco Mezzagrana M-PART.PASS Merzagrana M-PART.PASS Mercivendola M-N mesciroba M-N mesciroba M-N mezzaboccia M-N mezzaluna M-N mezzaluna M-N mezzaluna M-N mezzaluna M-N mezzocolore M-N Mezzodio M-N	malvoglianza	A-N
malvolontà A-N malvoluto AVV-PART.PASS manomettere N-V manroverso N-N manubalestro N-N mappamondo N-N mazzafascio N-N mazzafionda N-N mazzafrusto N-N mazzapicchio N-V mazzascudo N-N melacotogna N-A melagrana N-PART.PASS melagranata N-PART.PASS melagranata N-PART.PASS melocotogno N-A melocotogno N-A melocotogno N-A melocotogno N-A melocotogno N-A melocarono N-PART.PASS mercivendola N-N mesciroba V-N mettifuoco V-N mettifuoco V-N mezzaboccia A-N mezzaluna A-N mezzaluna A-N mezzaluna A-N mezzocolore A-N mezzocolore A-N mezzodi A-N mezzodi A-N mezzofiera A-N	malvolente	<b>AVV-PART.PRES</b>
malvoluto AVV-PART.PASS manomettere N-V manroverso N-N manubalestro N-N mappamondo N-N mazzafascio N-N mazzafionda N-N mazzafionda N-N mazzafiondo N-V mazzascudo N-N mazzascudo N-N melacotogna N-A melagrana N-PART.PASS melagranata N-PART.PASS melarancia N-N mellomela N-N melocotogno N-A melograno N-PART.PASS mercivendola N-N mesciroba V-N mettifuoco V-N mezzaboccia A-N mezzalana A-N mezzaluna A-N mezzaluna A-N mezzocolore A-N mezzocolore A-N mezzodi A-N mezzodi A-N mezzofiera A-N	malvolere	AVV-V
manomettere manroverso manubalestro manubalestro mappamondo mazzafascio N-N mazzafionda N-N mazzafrusto M-N mazzapicchio M-V mazzascudo M-N melacotogna melagrana melagrana M-PART.PASS melagranata M-PART.PASS melarancia M-N mellomela M-N melocotogno M-A melograno M-PART.PASS mercivendola M-N mesciroba W-N mettifuoco W-N mettifuoco W-N mezzaboccia M-N mezzalana M-N mezzaluna M-N mezzaluna M-N mezzocolore M-A Mezzocolore Mezzodi M-N Mezzofiera M-N Mezzofi	malvolontà	A-N
manroverso N-N manubalestro N-N manubalestro N-N manubalestro N-N mappamondo N-N mazzafascio N-N mazzafonda N-N mazzafrusto N-N mazzapicchio N-V mazzascudo N-N melacotogna N-A melagrana N-PART.PASS melagranata N-PART.PASS melarancia N-N mellomela N-N melocotogno N-A melograno N-PART.PASS mercivendola N-N mesciroba V-N mettifuoco V-N mezzaboccia A-N mezzagrana A-N mezzaluna A-N mezzaluna A-N mezzocolore A-N mezzocolore A-N mezzodio A-N mezzodio A-N mezzofiera A-N modilefoglio A-N millefoglio A-N millefoglio A-N millefoglio A-N mordecascio V-N mordecascio V-N mordigallina V-N musacorna N-N nessunora A-N nobiluomo A-N nobiluomo A-N notetempo N-N notetetempo N-N notetempo N-N not	malvoluto	AVV-PART.PASS
manubalestro mappamondo mappamondo n-N mazzafascio n-N mazzafionda n-N mazzafrusto n-N mazzapicchio m-V mazzascudo n-N melacotogna melagrana melagrana melagranata m-PART.PASS melagranata mellomela n-N melocotogno n-A melograno mercivendola mesciroba mettifuoco mezzaboccia mezzalana mezzalana mezzaluna mezzaluna mezzaluna mezzalotte mezzocolore mezzodi mezzodio mezzodio mezzodio mezzodio mezzofiera mezzogiorno millefoglie millefoglio mirasole v-N mordecascio v-N mordecascio mordigallina nobiluomo nottetempo occhibagliare N-V	manomettere	N-V
mappamondo mazzafascio mazzafionda mazzafionda mazzafrusto mazzascudo m-N mazzascudo melacotogna melagrana melagrana melagranata mellomela melocotogno melograno mesciroba mettifuoco mezzagocia mezzagrana mezzalana mezzaluna mezzaluna mezzalore mezzocolore mezzodi mezzodi mezzodi mezzodi mezzodi mezzodi mezzodi mezzofiera mezzodi mezzofiera A-N mezzofiera mezzofiera A-N mezzofiera mezzofiera A-N mezzofiera A	manroverso	N-N
mazzafascio N-N mazzafionda N-N mazzafrusto N-N mazzascudo N-V mazzascudo N-N melacotogna N-A melagrana N-PART.PASS melagranata N-PART.PASS melarancia N-N mellomela N-N melocotogno N-A melograno N-PART.PASS mercivendola N-N mesciroba V-N mettifuoco V-N mezzaboccia A-N mezzagrana A-N mezzalana A-N mezzaluna A-N mezzalona A-N mezzaotte A-N mezzocolore A-N mezzodi A-N mezzofiera A-N	manubalestro	N-N
mazzafionda mazzafrusto M-N mazzapicchio M-V mazzascudo N-N melacotogna M-A melagrana M-PART.PASS melagranata M-PART.PASS melarancia M-N mellomela M-N melocotogno M-A melograno M-PART.PASS mercivendola M-N mesciroba M-N mesciroba M-N mettifuoco M-N mettifuoco M-N mezzaboccia M-N mezzalana M-N mezzalana M-N mezzaluna M-N mezzodio M-A Mezzocolore M-N mezzodio M-N mezzodio M-N mezzodio M-N mezzofiera M-N millefoglio M-N millefoglio M-N millefoglio M-N millefoglio M-N mordecascio V-N mordecascio V-N mordigallina V-N musacorna N-N nessunora N-N nobildonna nobiluomo A-N nottetempo N-N N	mappamondo	N-N
mazzafrusto mazzapicchio mazzascudo N-N melacotogna melagrana melagrana melagranata melagranata mellomela melocotogno melograno melograno mesciroba mettifuoco mezzaboccia mezzagrana mezzaluna mezzaluna mezzaluna mezzocolore mezzodò mezzodò mezzodò mezzodò mezzodò mezzodò mezzodò mezzodo mezzodo mezzodo mezzodo mezzodò nez mez ne	mazzafascio	N-N
mazzapicchio mazzascudo N-N melacotogna M-A melagrana M-PART.PASS melagranata M-PART.PASS melarancia M-N mellomela N-N melocotogno M-A melograno M-PART.PASS mercivendola N-N mesciroba M-V-N mettifuoco Mezzaboccia M-N mezzalana Mezzalana M-N mezzaluna Mezzaluna Mezzaluna Mezzaoctie A-N Mezzocolore Mezzocolore Mezzodi M-N Mezzodio A-N Mezzodio M-N Mezzofiera Mezzogiorno Millefoglie A-N millefoglie M-N millefoglie M-N mirasole M-N mordecascio M-N mordecascio M-N mordigallina M-N nessunora N-N nessunora N-N nobildonna N-N nottetempo N-N nottetempo N-N nottetempo N-N notchibagliare N-V	mazzafionda	N-N
mazzascudo melacotogna melacotogna M-A melagrana M-PART.PASS melagranata M-PART.PASS melarancia N-N mellomela M-N melocotogno M-A melograno M-PART.PASS mercivendola N-N mesciroba M-N mettifuoco M-N mezzaboccia A-N mezzagrana A-N mezzaluna Mezzaluna Mezzaluna Mezzaotte A-N mezzocolore M-N mezzodi A-N mezzofiera A-N mezzogiorno M-N millefoglie A-N millefoglie A-N mirasole V-N mordecascio V-N mordigallina V-N musacorna N-N nessunora nobildonna A-N nobiluomo A-N nottetempo N-N occhibagliare N-V	mazzafrusto	N-N
melacotogna melagrana melagranata M-PART.PASS melagranata M-PART.PASS melarancia M-N mellomela N-N melocotogno N-A melograno M-PART.PASS mercivendola N-N mesciroba M-V-N mettifuoco Mezzaboccia M-N mezzaloccia M-N mezzalana Mezzaluna Mezzaluna Mezzaluna Mezzocolore Mezzocolore Mezzodi Mezzodi Mezzodi Mezzodi Mezzodi  Mezzodi  Mezzodi  Mezzodi  Mezzodi  Mezzodi  Mezzodi  Mezzodi  Mezzodi  Mezzodi  Mezzodi  Mezzodi  Mezzodi  Mezzodi  Mezzodi  M-N Mezzodi  Mezzodi  Mezzodi  M-N Mezzodi  Mezzodi  M-N Mezzodi  M-N Mezzodi  M-N Millefoglie  A-N Millefoglie  M-N Millefoglio  M-N Mordecascio  V-N Mordigallina  V-N Mordecascio  N-N Mordigallina  N-N N N N N N N N N N N N N N N N N N	mazzapicchio	N-V
melagrana N-PART.PASS melagranata N-PART.PASS melagranata N-N mellomela N-N melocotogno N-A melograno N-PART.PASS mercivendola N-N mesciroba V-N mettifuoco V-N mezzaboccia A-N mezzagrana A-N mezzaluna A-N mezzaluna A-N mezzocolore A-N mezzodò A-N mezzodio A-N mezzodio A-N mezzofiera A-N mezzogiorno A-N millefoglie A-N millefoglio A-N mirasole V-N mordecascio V-N mordigallina V-N mordigallina N-N nessunora A-N nobildonna A-N nobildonna A-N nobildonna N-N nottetempo N-N occhibagliare N-V	mazzascudo	N-N
melagranata melarancia M-N mellomela M-N melocotogno M-A melograno M-PART.PASS mercivendola M-N mesciroba M-N mesciroba M-N mettifuoco Mezzaboccia M-N mezzaloccia M-N mezzalana M-N mezzaluna Mezzaluna Mezzaluna Mezzocolore Mezzocolore Mezzodò Mezzodò Mezzodò M-N Mezzodio Mezzodio M-N Mezzofiera Mezzogiorno Millefoglie M-N millefoglie M-N millefoglio M-N mordecascio M-N mordecascio M-N mordigallina M-N messunora N-N nessunora N-N nessunora N-N nobildonna N-N nottetempo N-N occhibagliare N-V	melacotogna	N-A
melarancia N-N mellomela N-N melocotogno N-A melograno N-PART.PASS mercivendola N-N mesciroba V-N mettifuoco V-N mezzaboccia A-N mezzagrana A-N mezzaluna A-N mezzaluna A-N mezzocolore A-N mezzocolore A-N mezzodi A-N mezzodi A-N mezzodi A-N mezzodio A-N mezzodio A-N mezzodio A-N mezzodio A-N mezzogiorno A-N millefoglie A-N millefoglie A-N millefoglio A-N millefoglio A-N mirasole V-N mordecascio V-N mordecascio V-N mordigallina V-N musacorna N-N nessunora A-N nobildonna A-N nobildonna A-N nobildonna A-N nottetempo N-N occhibagliare N-V	melagrana	
mellomela melocotogno M-A melograno M-PART.PASS mercivendola M-N mesciroba W-N mettifuoco W-N metzaboccia A-N mezzagrana A-N mezzalana A-N mezzaluna Mezzaluna Mezzocolore Mezzodio Mezzodio A-N Mezzodio A-N Mezzofiera Mezzofiera Mezzogiorno Millefoglie A-N millefoglie M-N millefoglio M-N mordecascio W-N Mordigallina W-N Mordigallina N-N N N N N N N N N N N N N N N N N N	melagranata	N-PART.PASS
melocotogno melograno N-A melograno N-PART.PASS mercivendola N-N mesciroba V-N mettifuoco Mezzaboccia A-N mezzagrana A-N mezzalana A-N mezzaluna A-N mezzocannone Mezzocolore A-N mezzodi A-N mezzodi A-N mezzodi A-N mezzodio A-N mezzodio A-N mezzofiera Mezzogiorno Millefoglie A-N millefoglie A-N mirasole V-N mordecascio V-N mordigallina V-N musacorna N-N nessunora N-N nobildonna A-N nobildonna N-N notetempo N-N occhibagliare N-V		N-N
melograno N-PART.PASS mercivendola N-N mesciroba V-N mettifuoco V-N mezzaboccia A-N mezzagrana A-N mezzalana A-N mezzaluna A-N mezzocannone A-N mezzocolore A-N mezzodì A-N mezzodì A-N mezzodio A-N mezzogiera A-N mezzogiera A-N mezzogiera A-N mezzogiera A-N mezzogiono A-N millefoglie A-N millefoglio A-N mirasole V-N mordecascio V-N mordigallina V-N musacorna N-N nessunora A-N nobildonna A-N nobiluomo A-N notetempo N-N occhibagliare N-V	mellomela	N-N
mercivendola N-N mesciroba V-N mettifuoco V-N mezzaboccia A-N mezzagrana A-N mezzalana A-N mezzaluna A-N mezzoconore A-N mezzocolore A-N mezzodì A-N mezzodio A-N mezzofiera A-N mezzogiorno A-N millefoglie A-N millefoglie A-N millefoglio A-N mirasole V-N mordecascio V-N mordigallina V-N musacorna N-N nessunora A-N nobiluomo A-N nobiluomo A-N nottetempo N-N occhibagliare N-V	melocotogno	
mesciroba mettifuoco V-N mezzaboccia A-N mezzagrana A-N mezzalana Mezzaluna Mezzanotte A-N mezzocannone Mezzocolore Mezzodi Mezzodio A-N Mezzofiera Mezzofiera Mezzogiorno Millefoglie Millefoglio Mirasole W-N Mordecascio W-N Mordecascio W-N Mordigallina W-N Mordigallina N-N N Nocchibagliare N-V	9	N-PART.PASS
mettifuoco mezzaboccia A-N mezzagrana A-N mezzalana A-N mezzaluna Mezzanotte A-N mezzocannone Mezzocolore Mezzodi A-N mezzodi A-N Mezzodio A-N Mezzofiera Mezzogiorno Millefoglie Millefoglie Millefoglie Millefoglio Millefog	mercivendola	
mezzaboccia A-N mezzagrana A-N mezzalana A-N mezzaluna A-N mezzanotte A-N mezzocannone A-N mezzocolore A-N mezzodì A-N mezzodio A-N mezzofiera A-N mezzogiorno A-N millefoglie A-N millefoglie A-N millefoglia A-N mirasole V-N mordecascio V-N mordigallina V-N mordigallina V-N nessunora A-N nobildonna A-N nobiluomo A-N notetempo N-N occhibagliare N-V	mesciroba	V-N
mezzagrana A-N mezzaluna A-N mezzaluna A-N mezzanotte A-N mezzocannone A-N mezzocolore A-N mezzodì A-N mezzodio A-N mezzofiera A-N mezzogiorno A-N millefoglie A-N millefoglie A-N millefoglia A-N millefoglia A-N mirasole V-N mordecascio V-N mordigallina V-N mordigallina V-N nessunora A-N nobildonna A-N nobildonna A-N nobildonna A-N nobiluomo N-N occhibagliare N-V		V-N
mezzalana A-N mezzaluna A-N mezzanotte A-N mezzocannone A-N mezzocolore A-N mezzodio A-N mezzofiera A-N mezzogiorno A-N millefoglie A-N millefoglie A-N mirasole V-N mordecascio V-N mordigallina V-N musacorna N-N nessunora A-N nobildonna A-N nobiluomo A-N nottetempo N-N occhibagliare N-V	mezzaboccia	A-N
mezzaluna A-N mezzanotte A-N mezzocannone A-N mezzocolore A-N mezzodio A-N mezzofiera A-N mezzogiorno A-N millefoglie A-N millefoglie A-N mirasole V-N mordecascio V-N mordigallina V-N musacorna N-N nessunora A-N nobildonna A-N nobiluomo A-N nottetempo N-N occhibagliare N-V	_	
mezzanotte A-N mezzocannone A-N mezzocolore A-N mezzodì A-N mezzodio A-N mezzofiera A-N mezzogiorno A-N millefoglie A-N millefoglie A-N mirasole V-N mordecascio V-N mordigallina V-N mordigallina A-N nobildonna A-N nobildonna A-N nobildomo A-N nottetempo N-N occhibagliare N-V		
mezzocannone mezzocolore mezzodì A-N mezzodio A-N mezzofiera A-N mezzogiorno A-N millefoglie A-N millefoglie A-N millefoglio A-N mirasole V-N mordecascio V-N mordigallina V-N musacorna nobildonna nobiluomo nottetempo occhibagliare N-V	mezzaluna	
mezzocolore mezzodì A-N mezzodio A-N mezzofiera A-N mezzogiorno A-N millefoglie A-N millefoglie A-N millefoglio A-N mirasole V-N mordecascio V-N mordigallina V-N musacorna N-N nessunora nobildonna A-N nobiluomo N-N nottetempo occhibagliare N-V	mezzanotte	
mezzodì A-N mezzodio A-N mezzofiera A-N mezzogiorno A-N millefoglie A-N millefoglie A-N mirasole V-N mordecascio V-N mordigallina V-N musacorna N-N nessunora A-N nobildonna A-N nobiluomo A-N nottetempo N-N occhibagliare N-V		
mezzodio A-N mezzofiera A-N mezzogiorno A-N millefoglie A-N millefoglio A-N mirasole V-N mordecascio V-N mordigallina V-N musacorna N-N nessunora A-N nobildonna A-N nobiluomo A-N nottetempo N-N occhibagliare N-V		
mezzofiera A-N mezzogiorno A-N millefoglie A-N millefoglio A-N mirasole V-N mordecascio V-N mordigallina V-N musacorna N-N nessunora A-N nobildonna A-N nobiluomo A-N nottetempo N-N occhibagliare N-V		
mezzogiorno A-N millefoglie A-N millefoglio A-N mirasole V-N mordecascio V-N mordigallina V-N musacorna N-N nessunora A-N nobildonna A-N nobiluomo A-N nottetempo N-N occhibagliare N-V		
millefoglie A-N millefoglio A-N mirasole V-N mordecascio V-N mordigallina V-N musacorna N-N nessunora A-N nobildonna A-N nobiluomo A-N nottetempo N-N occhibagliare N-V		
millefoglio A-N mirasole V-N mordecascio V-N mordigallina V-N musacorna N-N nessunora A-N nobildonna A-N nobiluomo A-N nottetempo N-N occhibagliare N-V	9	
mirasole V-N mordecascio V-N mordigallina V-N musacorna N-N nessunora A-N nobildonna A-N nobiluomo A-N nottetempo N-N occhibagliare N-V	_	
mordecascio V-N mordigallina V-N musacorna N-N nessunora A-N nobildonna A-N nobiluomo A-N nottetempo N-N occhibagliare N-V	_	
mordigallina V-N musacorna N-N nessunora A-N nobildonna A-N nobiluomo A-N nottetempo N-N occhibagliare N-V		
musacorna N-N nessunora A-N nobildonna A-N nobiluomo A-N nottetempo N-N occhibagliare N-V		
nessunora A-N nobildonna A-N nobiluomo A-N nottetempo N-N occhibagliare N-V	mordigallina	
nobildonna A-N nobiluomo A-N nottetempo N-N occhibagliare N-V	musacorna	
nobiluomo A-N nottetempo N-N occhibagliare N-V		
nottetempo N-N occhibagliare N-V		
occhibagliare N-V		
	=	
occhibagliato N-PART.PASS	_	
	occhibagliato	N-PART.PASS

. 1\	ATTT N
oggidì 	AVV-N
oggimai 	AVV-AVV
ogginnanzi	AVV-P
ognanno	A-N
ognicavelle	A-N
ognidì 	A-N
ognimai	A-AVV
ognissanti	A-A
ognomo	A-N
ognora	A-N
orcipoggia	N-N
orofino	N-A
palafitta	N-A
palmacristo	N-N
panchiedente	N-PART.PRES
panicuocolo	N-N
pannolino	N-N
pannotendolo	N-N
pappaculo	V-N
parabordone	V-N
paracuore	V-N
paramosche	V-N
parapetto	V-N
parasole	V-N
pascibietola	V-N
pascipeco	V-N
patrisonno	N-N
pecegreca	N-A
pelacane	V-N
peracotogna	N-A
pesceporco	N-N
pescespada	N-N
pescivendolo	N-N
pettirosso	N-A
piamadre	A-N
piantamalanno	V-N
picchiapetto	V-N
piccioluomo	A-N
piedistallo	N-N
placamarito	V-N
plebeuscito	N-PART.PASS
pomagranata	N-PART.PASS
pomarancio	N-N
pomogranato	N-A
porcospino	N-N
portapesi	V-N
portareca	V-V
poveruomo	A-N
primegenite	A-PART.PASS
primiscrineo	A-N
primogenito	A-PART.PASS
primogemeo	A-PART.PASS
r	

primotempo	A-N
raffacane	V-N
rizzacrino	V-N
roccaforte	N-A
rosamarina	N-A
rosmarino	N-A
sabatodì	N-N
sacrosanto	A-A
salgemma	N-N
saliscendi	V-V
saltabeccare	V-V
saltabellare	V-V
salvacondotto	A-PART.PASS
salvadanaio	V-N
sanamunda	V-V
sbiecaciglio	V-N
scannadio	V-N
scialacquavino	V-N
secondogenito	A-PART.PASS
secondogenito	A-N
sempremai	AVV-AVV
sempreterno	AVV-A
sempreverde	AVV-A
sempreviva	AVV-A
semprevivo	AVV-A
senatoconsulto	N-N
settiforme	A-N
solleone	N-N
spigaceltica	N-A
spinacervina	N-A
spinamagna	N-A
spinapesce	N-N
squarciasacco	V-N
squarciavela	V-N
squassacapo	V-N
stracciasacco	V-N
tagliaborse	V-N
talora	A-N
talvolta	A-N
telaragna	N-N
terracotta	N-A
terraferma	N-A
terrafine	N-N
terrasanta	N-A
terratremolo	N-N
terremoto	N-N
tornafine	V-N
tornasole	V-N
tuttafiata	A-N
tuttavia	A-N
tuttavolta	A-N
tuttodì	A-N

tuttogiorno	A-N
tuttora	A-N
tuttotempo	A-N
ultimogenito	A-N
usufrutto	N-N
valentuomo	A-N
vanagloria	A-N
vanpiedare	V-V
ventipioggia	N-N
ventipiova	N-N
ventipiovolo	N-A
veracesomigliante	A-PART.PRES
verdebruno	A-A
vermocane	N-N
verodicente	N-PART.PRES
verosimigliante	N-PART.PRES
verosimile	N-A
vivamai	V-AVV
volgicuore	V-N
vuotacessi	V-N

## APPENDICE II: DATASET COMPOS\_ITA

Composto	Categorie di input	1	2	3	4	5	Occ. totali
acchiappaporci	V-N	0	0	0	0	4	4
acconciateste	V-N	0	0	0	1	0	1
acquabianca	N-A	0	0	0	0	3	3
acquaforte	N-A	1	3	2	12	17	35
acquarosa	N-N	4	5	11	0	0	20
acquavita	N-N	0	1	2	9	7	19
acquavite	N-N	0	2	1	7	39	49
affittacamere	V-N	0	0	0	0	1	1
affittascagni	V-N	0	0	0	3	0	3
agrodolce	A-A	0	0	0	2	2	4
albaspina	A-N	0	0	0	0	4	4
altitonante	AVV-PART.PRES	0	3	2	9	0	14
altochiomante	AVV-PART.PRES	0	0	0	0	1	1
altocinto	AVV-PART.PASS	0	0	0	1	7	8
altogatto	A-N	0	0	0	0	1	1
altolocato	AVV-PART.PASS	0	0	0	0	3	3
altopiano	A-N	0	14	0	0	16	30
altorecato	AVV-PART.PASS	0	0	0	0	1	1
altoriverso	AVV-PART.PASS	0	0	0	0	1	1
altosonante	AVV-PART.PRES	0	0	2	0	4	6
altosono	A-N	2	0	0	0	0	2
altovolante	AVV-PART.PRES	0	0	0	2	2	4
analogicoorganico	A-A	0	0	0	0	1	1
andirivieni	V-V	1	1	3	14	27	46
angloamericano	A-A	0	0	0	1	1	2
animavversione	N-N	1	0	0	2	3	6
appicciafoco	V-N	0	0	0	1	0	1
apricessi	V-N	0	0	1	0	0	1
arcobaleno	N-N	0	3	6	2	16	27
arcoscenico	N-A	0	0	0	0	1	1
argentovivo	N-A	3	8	122	9	15	162
armacollo	V-N	0	1	1	8	8	18
armilucente	N-PART.PRES	0	0	0	0	1	1
arruffacervello	V-N	0	0	0	0	1	1
arruffapopolo	V-N	0	0	0	0	7	7
asciugamano	V-N	0	0	0	1	8	9
attaccabrighe	V-N	0	0	0	0	4	4

attaccaferro	V-N	0	0	0	1	0	1
attaccapanni	V-N	0	0	0	3	8	11
azzeccagarbugli	V-N	0	0	0	0	5	5
baciamano	V-N	0	0	1	9	4	14
baciapile	V-N	0	0	0	0	3	3
bancarotta	N-A	0	0	0	0	26	26
bassofondo	A-N	0	1	4	1	11	17
bassorilievo	A-N	0	24	53	10	27	114
battibaleno	V-N	0	0	0	0	1	1
battibecco	V-N	0	0	0	0	9	9
battibuglio	V-N	0	0	0	3	0	3
batticarne	V-N	0	0	0	0	2	2
batticulo	V-N	0	0	1	0	0	1
batticuore	V-N	0	0	2	16	8	26
battiferro	V-N	0	1	0	0	0	1
battifolle	V-N	57	0	0	0	1	58
battifuoco	V-N	0	0	0	1	0	1
battilana	V-N	1	0	0	0	0	1
battiloro	V-N	0	5	2	2	0	9
battimano	V-N	0	0	0	1	21	22
battipalo	V-N	0	1	0	0	0	1
battiponti	V-N	0	2	0	0	0	2
battisfancello	V-N	1	0	0	0	0	1
battisoffia	V-V	0	1	0	0	0	1
battisoffiola	V-N	0	0	0	0	2	2
battistrada	V-N	0	0	0	3	0	3
bbattimuro	V-N	0	0	0	1	0	1
beccafico	V-N	0	8	4	2	6	20
beccamorto	V-N	3	6	6	15	5	35
belladonna	A-N	0	0	0	0	4	4
belvedere	A-N	0	0	0	0	6	6
benaffetto	AVV-PART.PASS	0	0	4	6	1	11
benaltrimenti	AVV-AVV	0	0	0	12	20	32
benamato	AVV-PART.PASS	0	2	2	2	6	12
benavventuroso	AVV-A	6	0	4	2	0	12
bencerto	AVV-A	17	33	25	19	24	118
benconsiderare	AVV-V	6	23	44	13	6	92
bendegno	AVV-A	12	12	78	31	4	137
bendetto	AVV-PART.PASS	9	16	20	10	5	60
bendire	AVV-V	33	43	37	10	13	136
bendisposto	AVV-PART.PASS	37	32	30	11	13	123
beneaccetto	AVV-PART.PASS	0	0	0	0	1	1

benearrivato	AVV-PART.PASS	0	0	0	1	0	1
benessere	AVV-N	8	13	28	56	80	185
benestare	AVV-N	7	5	3	3	6	24
benfare	AVV-N	113	63	39	14	5	234
benfornito	AVV-PART.PASS	24	29	16	7	2	78
benguardare	AVV-V	51	34	12	12	12	121
beninteso	AVV-PART.PASS	2	8	20	32	37	99
bennato	AVV-PART.PASS	6	18	73	36	8	141
bennoto	AVV-A	6	17	75	37	8	143
benpagato	AVV-PART.PASS	2	17	7	4	3	33
benparlante	AVV-PART.PRES	9	3	1	3	1	17
benpensante	AVV-PART.PRES	0	0	0	0	5	5
benpresto	AVV-AVV	0	8	22	108	52	190
benservito	AVV-PART.PASS	10	25	15	10	6	9
bensicuro	AVV-A	5	11	21	10	19	66
bentornato	AVV-PART.PASS	4	7	5	6	3	25
bentornato	AVV-PART.PASS	4	7	5	6	3	25
bentosto	AVV-AVV	3	8	27	61	13	112
benvenire	AVV-V	3	23	18	8	16	68
benvenuto	AVV-PART.PASS	24	51	34	13	23	145
benvestito	AVV-PART.PASS	5	17	35	4	20	81
benvolentieri	AVV-AVV	3	4	5	29	7	48
benvolere	AVV-V	13	10	3	7	3	36
benvoluto	AVV-PART.PASS	1	11	9	11	2	34
bianco latteo	A-A	0	0	0	1	0	1
biancoavorio	A-N	0	0	0	0	1	1
biancoazzurro	A-A	0	0	0	0	2	2
biancogiallo	A-A	0	0	0	0	1	1
biancospino	A-N	0	0	0	2	21	23
biancovestito	A-PART.PASS	1	0	0	0	2	3
boccaporto	N-N	0	0	0	1	3	4
borgomastro	N-N	0	0	0	1	2	3
bracciobianco	N-A	0	0	0	1	0	1
bruciapelo	V-N	0	0	0	2	13	15
buonamano	A-N	0	2	2	3	0	7
buonanotte	A-N	3	10	15	38	49	115
buoncuore	A-N	5	11	14	44	38	112
buongiorno	A-N	0	3	11	19	46	79
buongoverno	A-N	0	9	15	26	15	65
buongustaio	A-N	0	0	0	0	5	5
buonora	A-N	14	101	78	43	31	267
buonsenso	A-N	1	0	7	46	99	153

buonumore	A-N	0	0	0	26	57	83
buonuomo	A-N	14	44	33	53	43	187
buttafuori	V-AVV	0	0	0	0	2	2
cacafaccia	V-N	0	0	0	0	1	1
cacalettere	V-N	0	0	1	0	0	1
cacanido	V-N	0	0	0	2	0	2
cacapensieri	V-N	0	1	2	0	0	2
cacasangue	V-N	0	1	7	1	0	9
cacasotto	V-AVV	0	0	0	1	0	1
cacastecchi	V-N	0	1	0	0	0	1
cacazibetto	V-N	0	0	1	1	0	2
cacciacristo	V-N	0	1	0	0	0	1
cacciacuore	V-N	0	0	1	0	0	1
cacciafusto	V-N	1	0	0	0	0	1
caccialepre	V-N	0	0	2	0	4	6
cacciapensieri	V-N	1	0	0	0	0	1
caciocavallo	N-N	1	6	0	1	3	11
caciovacca	N-N	0	0	1	0	0	1
caffelatte	N-N	0	0	0	0	6	6
calcafoglio	V-N	0	0	0	0	1	1
calcatreppolo	V-N	0	1	0	0	0	1
cambiavalute	V-N	0	0	0	4	0	4
camposanto	N-A	5	1	19	2	72	99
cangiacolore	V-N	1	0	1	0	0	2
cantafavola	V-N	0	1	3	0	1	5
cantafola	V-N	0	1	2	0	0	3
cantambanco	V-N	0	0	0	2	0	2
cantastorie	V-N	0	0	0	1	12	14
cantimplora	V-V	0	0	2	0	0	2
cantofermo	N-A	0	0	0	0	2	2
cantopieno	N-A	0	1	0	0	0	1
capameno	N-A	0	0	0	1	10	10
capelvenere	N-N	1	0	1	1	12	15
capinera	N-A	1	0	0	0	23	24
capinero	N-A	0	1	1	0	6	8
capirosso	N-A	0	1	0	0	0	1
capobanda	N-N	0	0	0	4	4	8
capobirro	N-N	0	0	0	0	1	1
capobugio	N-A	0	0	0	0	1	1
capocaccia	N-N	0	0	1	0	1	2
capocarceriere	N-N	0	0	0	1	0	1
capochino	N-A	15	22	13	3	48	101

capocomico	N-N	0	0	0	0	275	275
capocomune	N-N	0	0	0	0	7	7
capodivisione	N-N	0	0	0	0	5	5
capofamiglia	N-N	0	0	0	0	2	2
capoficco	N-A	0	0	0	0	1	1
capofila	N-N	0	0	0	1	2	3
capofitto	N-A	1	1	1	0	18	21
capofuoco	N-N	1	0	2	1	0	4
capogestione	N-N	0	0	0	0	1	1
capogirlo	N-N	1	1	5	0	0	7
capogiro	N-N	0	0	2	2	15	19
capogrosso	N-A	0	0	0	0	1	1
capolavoro	N-N	0	1	0	6	38	45
capoletto	N-N	4	2	0	0	0	6
capolista	N-N	0	0	0	0	2	2
capoluogo	N-N	0	0	0	3	8	11
capomaestro	N-N	1	13	5	0	0	19
capomastro	N-N	1	2	2	1	16	22
capomorto	N-A	0	0	0	2	0	2
capomozzo	N-A	0	2	0	2	3	7
caponascondi	N-V	0	0	0	0	1	1
capopagina	N-N	0	1	0	0	0	1
capoparte	N-N	0	0	0	0	1	1
capopopolo	N-N	1	0	0	3	14	18
caporiccio	N-A	1	0	0	0	0	1
caporione	N-N	1	1	4	10	8	24
caposaldo	N-A	0	0	0	0	12	12
caposcarico	N-A	0	0	0	0	1	1
caposcena	N-N	0	0	0	1	0	1
caposcuola	N-N	0	0	0	2	3	4
caposervizio	N-N	0	0	0	0	9	9
caposetta	N-N	0	0	0	2	1	3
caposoldo	N-N	0	2	0	0	0	2
caposquadra	N-N	0	0	0	7	0	7
capostazione	N-N	0	0	0	0	22	22
capostipite	N-N	0	0	0	1	4	5
capotamburo	N-N	0	0	0	0	1	1
capotavola	N-N	0	0	0	2	2	4
capoufficio	N-N	0	0	0	0	8	8
capoverso	N-N	0	1	0	2	1	4
capovilla	N-N	0	0	0	0	2	2
cappacorta	N-A	0	0	0	0	1	1

cappalunga	N-A	0	0	0	0	1	1
cappamagna	N-A	0	0	0	3	2	5
cappanera	N-A	0	0	0	4	3	7
caprifico	N-N	0	2	0	2	4	12
caprifoglio	N-N	0	0	0	0	14	14
carnesecca	N-A	2	6	6	2	39	55
cartapecora	N-N	0	5	8	7	9	29
cartapesta	N-A	0	0	0	0	6	6
cartastraccio	N-N	0	0	1	0	0	1
cartavelina	N-A	0	0	0	27	4	31
casamatta	N-A	0	8	0	1	8	17
cascamorto	V-N	0	0	0	4	4	8
casciofiore	N-N	0	0	0	1	0	1
cassabanca	N-N	0	3	0	0	0	3
cassabanco	N-N	0	0	0	5	0	5
cassaforte	N-A	0	0	0	0	4	4
cassapanca	N-N	0	4	0	0	5	9
castigamatti	V-N	0	0	0	1	0	1
castrapopoli	V-N	0	0	0	0	1	1
castraporci	V-N	0	0	1	0	0	1
cavaceci	V-N	0	0	0	1	0	1
cavadenti	V-N	0	1	2	2	4	9
cavafango	V-N	0	2	0	0	0	2
cavalcavia	V-N	0	1	0	0	0	1
cavalocchio	V-N	1	0	0	0	0	1
cavamacchie	V-N	0	0	0	0	1	1
cavaoro	V-N	0	0	0	0	2	2
cavastraccio	V-N	0	0	0	1	0	1
	V-N	0	0	0	0	1	1
cavatappi cavolfiore	N-N	0	0	1	0	29	30
cavolrapa	N-N	0	0	0	0	1	1
cazzomatto	N-A	0	0	0	2	0	2
celesteazzurro	A-A	0	0	0	2	0	2
celtibero	A-A	2	3	5	0	0	10
ceralacca	N-N	0	0	3	26	3	32
chiappanuvoli	V-N	0	0	0	0	2	2
chiaroscuro	A-A	0	4	9	1	13	27
chiarosonante	AVV-PART.PRES AVV-PART.PRES	0	0	0	0 2	3 2	3
chiaroveggente chiavacuore		0					4
chiomazzurro	V-N	0	0	2	0	0	2
	N-A	0	0	1 0	1	0 9	2 9
cinciallegra	N-A	U	0	U	0	9	9

circavicino	AVV-AVV	2	0	0	0	0	2
codacavallina	N-A	1	3	0	0	0	4
codilungo	N-A	0	1	0	0	0	1
codimozzo	N-A	0	2	0	0	0	2
codirosso	N-A	0	1	0	0	0	1
collotorto	N-PART.PASS	0	0	2	0	3	5
compravendita	V-N	0	0	0	0	2	2
conciabocca	V-N	0	0	0	1	0	1
conciabrocca	V-N	0	0	0	0	2	2
conciastoria	V-N	0	0	0	1	0	1
contafavola	V-N	0	0	1	0	0	1
contasecondi	V-N	0	0	0	0	3	3
copialettere	V-N	0	15	0	2	1	18
copricapo	V-N	0	0	0	1	0	1
copriceffo	V-N	1	0	0	0	0	1
coprifuoco	V-N	0	1	0	0	0	1
crepacuore	V-N	0	0	2	10	17	29
crepapancia	V-N	0	0	1	0	1	2
crepapelle	V-N	0	0	0	1	9	10
cuorcontento	N-A	0	0	0	0	1	1
domacavallo	V-N	0	0	0	1	0	1
dormiveglia	V-V	0	0	0	0	15	15
erbabianca	N-A	0	0	0	0	6	6
erbivendolo	N-N	0	0	0	0	8	8
eroicomico	A-A	0	0	0	0	2	2
facciamara	N-A	0	0	0	1	0	1
falegname	V-N	0	1	0	17	34	52
falsacappa	A-N	0	0	0	2	0	2
falsamoneta	A-N	0	0	0	0	5	5
falsamonete	V-N	0	2	0	0	0	2
falsariga	A-N	0	0	0	0	10	10
fattefeste	V-N	0	0	1	0	0	1
fattibello	V-N	2	0	0	0	0	2
ferravecchio	N-A	0	5	9	3	4	21
ferrovia	N-N	0	0	0	2	148	150
fiaccagote	V-N	0	0	0	0	1	1
ficamoscia	N-A	0	0	0	1	0	1
ficcanaso	V-N	0	0	0	5	2	7
fiengreco	N-A	0	0	0	0	1	1
finimondo	N-N	1	1	4	1	11	18
fiutaprofumi	V-N	0	0	0	1	0	1
forasacchi	V-N	0	0	0	0	1	1

forasiepi	V-N	0	0	1	0	0	1
formicaleone	N-N	0	0	0	0	1	1
fottivento	V-N	0	0	1	0	1	2
francobollo	A-N	0	0	0	0	26	26
frustamattoni	V-N	0	0	0	1	0	1
fruttivendolo	N-N	0	0	0	0	3	3
fuggifatiche	V-N	0	0	0	1	0	1
gabbaminchioni	V-N	0	0	0	0	1	1
gabbamondo	V-N	0	0	0	0	3	3
galantuomo	A-N	13	1	15	275	251	555
gattaconiglio	N-N	0	1	0	0	0	1
gentildonna	A-N	116	45	104	18	14	297
gentiluomo	A-N	308	108	1171	82	135	1804
gialloarancione	A-A	0	0	0	0	1	1
giallobruno	A-A	0	0	0	2	1	3
giallocanarino	A-N	0	0	0	0	1	1
giallorosso	A-A	0	0	0	0	1	1
gialloverde	A-A	0	0	0	0	1	1
giocoforza	N-N	0	0	0	2	4	6
giracapi	V-N	0	0	0	0	1	1
girarrosto	V-N	0	0	0	0	7	7
girasole	V-N	0	0	2	1	19	22
giravolta	V-V	2	0	9	5	8	24
girotondo	N-A	1	0	0	0	1	2
gommalacca	N-N	0	1	0	5	0	6
gommarabica	N-A	1	8	0	0	3	12
gommelastica	N-A	0	0	0	0	14	14
gonfianuvoli	V-N	0	0	0	0	2	2
grancassa	A-N	0	0	0	0	13	13
granduca	A-N	6	19	182	26	51	284
granduchessa	A-N	0	0	41	5	0	46
granguardia	A-N	0	0	0	1	0	1
grantempo	A-N	122	168	161	268	102	821
granturco	A-N	0	0	0	6	48	54
grattacapi	V-N	0	0	0	0	2	2
grattacascia	V-N	0	0	0	2	0	2
grattapanze	V-N	0	0	0	2	0	2
grattapopolo	V-N	0	0	0	1	0	1
grattapugie	V-V	0	1	0	0	0	1
grigioroseo	A-A	0	0	0	0	1	1
grillotalpa	N-N	0	0	1	0	0	1
guardaboschi	V-N	0	0	0	6	2	8

guardacaccia	V-N	0	0	0	0	3	3
guardacamera	V-N	0	7	7	0	0	14
guardacucina	V-N	0	4	0	0	0	4
guardacuore	V-N	1	0	0	0	0	1
guardadonna	V-N	0	1	0	0	0	1
guardamacchie	V-N	0	0	0	1	0	1
guardamantili	V-N	0	1	0	0	0	1
guardamarina	V-N	0	0	0	0	1	1
guardamorto	V-N	1	0	0	0	0	1
guardanappa	V-N	1	11	0	0	0	12
guardancanna	V-N	0	1	0	0	0	1
guardaporta	V-N	0	0	0	0	1	1
guardaportone	V-N	0	0	0	2	6	8
guardaroba	V-N	1	8	25	12	24	70
guardasigilli	V-N	0	0	1	0	4	5
guardiamarina	N-A	0	0	0	0	4	4
guardinfante	V-N	0	0	1	2	6	9
guastamestiere	V-N	0	0	0	5	2	7
guastamondo	V-N	1	0	0	0	0	1
guastavillano	V-N	0	0	1	0	0	1
guazzabuglio	V-N	0	3	4	17	13	38
iermattina	AVV-AVV	5	14	0	1	10	30
imbrattamestiere	V-N	0	0	1	0	0	1
lasciapassare	V-N	0	0	0	18	2	20
latinobarbaro	N-A	0	0	0	16	0	16
lattemiele	N-N	0	0	0	0	1	1
lattivendolo	N-N	0	0	0	0	1	1
lavaceci	V-N	2	2	0	0	0	4
lavacecia	V-N	1	0	0	0	0	1
lavacoglione	V-N	0	0	0	1	0	1
lavamano	V-N	0	0	2	2	6	10
lavapiatti	V-N	0	0	0	0	32	32
lavascodelle	V-N	0	0	2	0	0	2
leccaboni	V-N	0	1	0	0	0	1
leccamuffo	V-N	1	0	0	0	0	1
leccapiatti	V-N	0	1	0	0	0	1
leccapiedi	V-N	0	0	0	0	1	1
leccasanti	V-N	0	0	0	0	1	1
leccascodelle	V-N	0	0	0	1	0	1
legnosanto	N-A	0	0	0	0	1	1
lunganime	A-N	0	0	0	0	3	3
lunganimo	A-N	2	0	0	0	0	2

luogotenente	N-PART.PRES	17	99	58	11	28	213
madreperla	A-N	0	1	8	10	16	35
madreselva	A-N	0	0	0	0	4	4
madrevite	N-N	0	0	2	0	2	5
magnapagnotte	V-N	0	0	1	0	0	1
magnapane	V-N	0	0	1	0	0	1
magnatordi	V-N	0	0	0	1	0	1
mandiritto	N-A	0	0	3	4	0	7
mangiacarte	V-N	0	0	0	0	2	2
mangiafagioli	V-N	0	0	1	0	0	1
mangiaguadagno	V-N	0	1	0	0	0	1
mangiaguerra	V-N	0	0	0	0	3	5
mangialiberali	V-N	0	0	0	0	1	1
mangiamaccheroni	V-N	0	0	0	0	1	1
mangiamoccoli	V-N	0	0	0	0	1	1
mangiapane	V-N	0	0	0	0	10	10
mangiapolenta	V-N	0	0	0	0	2	2
mangiapreti	V-N	0	0	0	0	2	2
manomettere	N-V	1	13	5	25	38	82
manomorta	N-A	0	0	0	2	13	15
manrovescio	N-N	1	3	1	2	5	12
mappamondo	N-N	6	3	7	6	6	28
marciapiede	V-N	0	0	0	0	53	53
mazzafrusto	N-N	0	14	3	1	0	18
mazzapicchio	N-N	0	1	0	0	2	3
melacotta	N-A	0	0	0	5	0	5
melarancia	N-N	0	35	1	0	4	40
melarancio	N-N	2	14	2	0	1	19
melograno	N-A	0	0	0	0	41	41
menadito	V-N	0	0	0	1	3	4
mesciroba	V-N	0	8	0	1	2	11
mettimale	V-AVV	0	0	0	1	2	3
mettiscandali	V-N	0	0	0	0	1	1
mezzacanna	A-N	0	0	0	4	2	6
mezzacoda	A-N	0	0	0	0	1	1
mezzaluna	A-N	0	3	5	1	21	30
mezzamattina	A-N	0	0	0	7	3	10
mezzanotte	A-N	50	59	73	103	154	439
mezzapicca	A-N	0	0	0	0	1	1
mezzobusto	A-N	0	0	0	0	3	3
mezzocannone	A-N	1	2	5	0	0	8
mezzocotto	A-N	0	0	1	0	0	1

mezzocuore	A-N	0	0	0	1	0	1
mezzodì	A-N	0	3	2	2	1	8
mezzodie	A-N	58	0	0	0	0	58
mezzogiorno	A-N	33	46	109	162	196	546
mezzosoprano	A-N	0	0	0	1	1	2
mietigambe	V-N	0	0	1	0	0	1
mozzorecchio	V-N	0	0	1	13	1	15
nericiglia	A-N	0	0	0	0	1	1
neroazzurro	A-A	0	0	0	0	2	2
neroblu	A-A	0	0	0	0	2	2
nobiluomo	A-N	1	0	1	4	1	7
nottetempo	N-N	11	1	4	6	15	37
nullatenente	N-PART.PRES	0	0	2	1	6	9
occhiazzurro	N-A	0	0	0	2	1	3
oggigiorno	N-N	0	0	3	13	15	31
ognimodo	A-N	0	47	57	107	183	394
ognissanti	A-N	18	55	16	10	8	107
ognora	A-N	118	84	156	86	38	482
onestuomo	A-N	0	0	0	7	7	14
padreterno	N-A	8	8	43	9	21	89
palcoscenico	N-A	0	0	0	8	92	100
pallacorda	N-N	0	0	0	3	0	3
pancotto	N-A	0	0	0	8	4	12
paneperso	N-A	0	0	0	0	4	4
panforte	N-A	0	0	0	0	9	9
pangiallo	N-A	0	0	0	7	0	7
pangrattato	N-A	0	0	0	4	93	97
pannolano	N-N	2	2	0	1	1	6
pannolino	N-N	12	104	20	2	31	169
pappafichi	V-N	0	0	1	0	0	1
pappagorgia	V-N	0	0	0	1	5	6
pappalardo	V-N	0	0	0	1	0	1
pappalasagne	V-N	0	0	1	0	0	1
paracadute	V-N	0	0	0	0	1	1
paracalunnie	V-N	0	0	0	1	0	1
paracarro	V-N	0	0	0	1	1	2
paracenere	V-N	0	0	0	0	1	1
paracqua	V-N	0	0	0	0	4	4
parafulmine	V-N	0	0	0	6	4	10
parafuoco	V-N	0	0	0	1	3	4
paralume	V-N	0	0	0	0	9	9
paramano	V-N	0	0	0	3	0	3

parapetto	V-N	2	53	6	13	55	129
parapiglia	V-V	0	1	2	5	12	20
parasole	V-N	0	0	0	0	4	4
paravento	V-N	0	0	0	2	11	13
pascibietola	V-N	3	0	0	1	0	4
pascipeco	V-N	0	5	0	0	0	5
pascivento	V-N	0	1	0	0	0	1
passacarte	V-N	0	0	0	0	1	1
passamano	V-N	0	0	1	1	1	3
passaporto	V-N	0	0	0	88	37	125
passatempo	V-N	0	3	25	47	37	112
passavento	V-N	0	0	1	0	0	1
passavolante	V-PART.PRES	0	18	1	0	0	19
pastafrolla	N-A	0	0	0	1	35	36
pastasciutta	N-A	0	0	0	0	11	11
pecegreca	N-A	0	0	4	0	4	8
pellerossa	N-A	0	0	0	1	14	15
peracotta	N-A	0	0	0	3	0	3
perdigiorni	V-N	0	0	0	1	7	8
perditempo	V-N	0	0	0	3	9	12
pescecane	N-N	0	1	0	0	52	53
pescecavallo	N-N	0	0	1	0	0	1
pesceluna	N-N	0	0	0	0	2	2
pescepersico	N-A	0	0	0	0	1	1
pescesega	N-N	0	0	0	0	1	1
pescespada	N-N	0	0	13	0	8	21
pescespina	N-N	0	1	0	0	0	1
pescivendolo	N-N	0	0	1	2	4	7
pettirosso	N-A	0	1	2	7	17	26
pianoforte	N-N	0	0	0	25	127	152
pianoterra	N-N	0	0	0	0	2	2
piantanimale	N-N	0	0	1	0	0	1
piantastecchi	V-N	0	0	0	0	1	1
pianterreno	N-A	0	11	0	16	26	53
piattaforma	A-N	0	0	1	0	26	27
piazzaforte	N-A	0	0	1	2	4	7
picchiamartelli	V-N	0	0	0	1	0	1
picchiapetto	V-N	1	1	0	1	0	3
picciongrossi	N-A	0	0	25	0	0	25
piedistallo	N-N	1	2	7	9	44	63
pisciabotte	V-N	0	0	0	4	0	4
pisciacani	V-N	0	0	0	0	1	1

pisciacquasanta	V-N	0	0	0	1	0	1
pisciadottrine	V-N	0	0	1	0	0	1
piscialletto	V-N	0	0	0	0	1	1
pisciaquindi	V-N	0	0	1	0	0	1
pisciasangue	V-N	0	0	1	0	0	1
porcospino	N-N	1	1	0	1	2	5
portabaldacchini	V-N	0	0	0	0	1	1
portabandiera	V-N	0	0	0	1	5	6
portabiglietti	V-N	0	0	0	0	3	3
portacenere	V-N	0	0	0	0	1	1
portafiori	V-N	0	0	0	0	1	1
portaflagello	V-N	0	0	1	0	0	1
portafoglio	V-N	0	0	0	13	27	40
portalettere	V-N	0	0	2	19	4	25
portamantelli	V-N	0	0	0	0	1	1
portamonete	V-N	0	0	0	0	7	7
portapanni	V-N	0	0	0	0	1	1
portapollastro	V-N	0	0	0	1	0	1
portareca	V-V	1	0	0	0	0	1
portaritratti	V-N	0	0	0	0	1	1
portasigarette	V-N	0	0	0	0	1	1
portaspilli	V-N	0	0	0	0	1	1
portatasche	V-N	0	0	0	1	0	1
portavento	V-N	0	0	1	0	0	1
portavoce	V-N	0	0	0	0	9	9
posalume	V-N	0	0	0	0	1	1
poveruomo	A-N	5	14	32	61	78	190
primogenito	A-PART.PASS	44	52	41	37	79	253
pungitopo	V-N	0	0	1	1	3	5
purosangue	A-N	0	0	0	0	3	3
rendiconto	V-N	0	0	0	3	11	14
resoconto	N-N	0	0	0	0	8	8
retrobottega	AVV-N	0	0	0	0	5	5
robivendolo	N-N	0	0	0	0	2	2
rodipopolo	V-N	0	0	0	0	1	1
rompicapo	V-N	0	0	0	0	1	1
rompicoglioni	V-N	0	0	0	2	0	2
rompicollo	V-N	0	0	3	11	17	33
rompiscatole	V-N	0	0	0	0	2	2
rompitesta	V-N	0	0	1	0	0	1
rossobruno	A-A	0	0	0	1	9	10
rossochiaro	A-A	0	0	0	1	4	5

rossociliegia	A-N	0	0	0	1	0	1
rossogiallo	A-A	0	0	0	0	1	1
rossosangue	A-N	0	0	0	0	2	6
rossoscuro	A-A	0	0	1	0	3	4
rubacuori	V-N	0	1	0	0	1	2
saccapane	N-N	0	0	0	0	1	1
sacrosanto	A-A	8	6	42	30	40	126
saliscendi	V-V	0	0	1	6	9	16
saltafossi	V-N	0	0	0	1	0	1
saltamontoni	V-N	0	0	0	1	1	2
salvacondotto	A-PART.PASS	11	80	112	20	26	249
salvadanaio	V-N	0	1	1	0	9	11
salvaguardia	V-N	0	0	1	7	10	18
salvamano	V-N	0	0	6	0	0	6
salvaroba	V-N	0	0	1	0	0	1
sbucciafratte	V-N	0	0	0	1	0	1
scaccafava	V-N	0	1	0	0	0	1
scacciamosche	V-N	0	0	0	0	1	1
scaccianoia	V-N	0	0	0	0	1	1
scacciapensieri	V-N	0	0	0	1	2	3
scacciaragno	V-N	0	0	0	1	0	1
scalapertico	V-N	0	0	1	0	0	1
scaldaletto	V-N	0	0	1	1	1	3
scaldasedie	V-N	0	0	0	1	0	1
scalzacane	V-N	0	0	2	1	3	6
scampaforche	V-N	0	0	0	1	1	2
scannafossi	V-N	0	1	0	0	0	1
scannaminestre	V-N	0	0	1	0	0	1
scannapagnotte	V-N	0	0	2	1	1	4
scannapane	V-N	0	0	1	0	0	1
scansafatiche	V-N	0	0	0	0	4	4
scappafatica	V-N	0	0	0	2	0	2
scaricabarile	V-N	0	0	0	1	0	1
scavezzacollo	V-N	0	1	1	3	2	7
scendidame	V-N	0	0	0	0	1	1
schiaccianoci	V-N	0	0	0	0	1	1
schiappalegne	V-N	0	1	0	0	0	1
schiccaraforzieri	V-N	0	0	1	0	0	1
schivafatiche	V-N	0	0	0	1	0	1
sciupateste	V-N	0	0	0	0	1	1
scolabrodo	V-N	0	0	0	1	0	1
scopamestieri	V-N	0	0	0	1	0	1

scoppiacorpo	V-N	0	1	0	0	0	1
scorribanda	V-N	0	0	1	1	2	4
scroccafuso	V-N	0	0	0	1	0	1
scrutinaparole	V-N	0	0	0	1	0	1
scrutinapensieri	V-N	0	0	0	1	0	1
seccafistole	V-N	0	0	0	1	0	1
secondogenito	A-PART.PASS	3	16	1	7	20	47
segnacaso	V-N	0	0	0	10	1	11
serramanico	V-N	0	0	0	0	1	1
serraserra	V-V	0	1	1	9	6	17
servopadrone	N-N	0	0	0	1	0	1
sgrignapappola	V-N	0	0	0	1	0	1
sommoscapo	A-N	0	3	0	0	0	3
sordomuto	A-A	0	0	0	2	10	12
spaccalegna	V-N	0	0	0	1	2	3
spaccamondi	V-N	0	0	0	0	1	1
spaccamontagne	V-N	0	0	0	0	1	1
spaccamonti	V-N	0	0	0	1	1	2
spaccapietre	V-N	0	0	0	0	1	1
spaccastrommola	V-N	0	0	1	0	0	1
sparafucile	V-N	0	0	0	0	1	1
spassapensieri	V-N	0	0	0	1	0	1
spassatempo	V-N	0	0	1	1	2	4
spaventapasseri	V-N	0	0	0	1	2	3
spazzacamini	V-N	0	0	1	4	9	14
spazzacampagna	V-N	0	0	0	2	0	2
spazzacoperta	V-N	0	1	0	0	0	1
spazzaforno	V-N	0	1	1	0	0	2
sputabalsamo	V-N	0	1	0	0	0	1
sputagrave	V-AVV	0	0	0	0	1	1
sputaperle	V-N	0	0	0	1	0	1
sputasenno	V-N	0	1	0	0	0	1
sputasentenze	V-N	0	0	0	1	1	2
squarciagola	V-N	0	0	0	0	15	15
squarciasacco	V-N	1	1	0	0	0	3
squassaforche	V-N	0	0	0	0	1	1
squassapennacchi	V-N	0	0	1	0	0	1
sticcalegna	V-N	0	0	0	2	0	2
storcileggi	V-N	0	0	0	1	0	1
stracciabrache	V-N	0	0	1	0	0	1
stracciafoglio	V-N	0	6	0	0	0	6
stracciasacco	V-N	0	1	0	1	0	2

straccivendolo	N-N	0	0	0	1	0	1
strappazucca	V-N	0	1	0	0	0	1
strascinafaccende	V-N	0	0	0	0	1	1
strippapelle	V-N	0	0	0	0	1	1
strozzapreti	V-N	0	0	0	2	1	3
struggibuco	V-N	0	0	0	0	1	1
stuccalettori	V-N	0	0	1	0	0	1
stuzzicadenti	V-N	0	0	1	2	0	3
tagliaborse	V-N	1	3	0	2	1	7
tagliacantoni	V-N	0	0	2	1	0	3
tagliacarte	V-N	0	0	0	0	1	1
taglialegna	V-N	0	0	0	0	8	8
tagliapane	V-N	0	0	0	0	1	1
tagliapietre	V-N	0	2	0	1	0	3
tagliaricotte	V-N	0	0	1	0	0	1
talora	A-N	163	143	502	224	546	1578
talora	A-N	155	141	502	222	545	1568
talvolta	A-N	170	169	358	926	578	2201
terracotta	N-A	16	8	17	7	13	61
terraferma	N-A	16	22	59	9	6	120
terrapieno	N-A	0	4	5	11	9	29
terzogenito	A-PART.PASS	1	0	3	1	4	9
tirapiedi	V-N	0	0	0	0	3	3
tiratira	V-V	0	0	1	3	6	10
torcicollo	V-N	0	0	0	0	4	4
torcifeccia	V-N	0	1	0	0	0	1
torciorecchio	V-N	0	0	0	3	0	3
torcipiede	V-N	0	0	0	0	1	1
tornaconto	V-N	0	0	0	3	24	27
tornagusto	V-N	0	0	0	2	2	4
tornasole	V-N	0	0	0	1	2	3
tragicomico	A-A	0	0	1	1	1	3
tragicommedia	N-N	0	0	2	9	3	14
tritacarne	V-N	0	0	0	0	3	3
troncabraccia	V-N	0	0	1	0	0	1
tuttafiata	A-N	10	6	0	1	0	17
tuttavia	A-N	326	448	1102	818	764	3458
tuttavolta	A-N	18	158	111	97	42	426
tuttogiorno	A-N	22	22	15	105	8	172
tuttora	A-N	61	23	3	121	122	330
tuttotempo	A-N	12	1	1	0	0	14
valentuomo	A-N	2	11	41	34	6	94

vanagloria	A-N	176	24	30	9	19	258
verdazzurro	A-A	0	0	0	0	6	6
verdebiondo	A-A	0	0	0	0	1	1
verdebruno	A-A	3	3	3	0	3	12
verdechiaro	A-A	0	2	0	1	2	5
verdegiallo	A-A	0	3	3	0	6	12
verdemezzo	A-A	0	3	0	0	1	4
verdescuro	A-A	0	6	1	2	3	12
verdeterra	A-N	0	47	1	0	0	59
vermocane	N-N	0	8	0	0	0	8
vinsanto	N-A	0	0	0	2	4	6
voltafaccia	V-N	2	0	0	0	3	5
voltastomaco	V-N	0	0	0	1	0	1
vuotacessi	V-N	0	3	0	0	0	3
vuotapitali	V-N	0	0	0	0	1	1
zappaterra	V-N	0	0	2	0	2	4