

# Sulla diacronia della composizione verbo-nominale in italiano



M. Silvia Micheli (Milano) – Pavel Štichauer (Praga)

## ABSTRACT

This paper deals with, in a diachronic perspective, the Italian verb-noun compounds of the type *portalettiere* and covers the time span that goes from the Origins well into the beginnings of the Twentieth century. On the basis of a rich dataset drawn from diachronic corpora and lexicographic sources, the article presents the history of this word-formation process, emphasizing the semantic and categorial nature of these compounds. The main claim is that the current semantic variability (person, instrument, event, period of time etc.) has always been available in the linguistic history of Italian as well as, to a minor extent, the categorial ambiguity by which VN compounds also function as adjectival modifiers. It is shown that this last function is much of a diachronic innovation. This paper, which stems from the theoretical premises of Construction Morphology, reveals that the productivity of the Construction  $[V-N]_{A/N}$  is due to both the semantic and formal flexibility shown from the earliest attestations of the language, and to the productivity of some semi-specified subschemas, which have served as models for the creation of new VN compounds, especially over the past few decades.

## KEYWORDS

agent nouns, Construction Morphology, event nouns, instrument nouns, Old Italian, Present-Day Italian, productivity, verb-noun compounds

## ABSTRACT

Questo articolo si propone di esaminare, in prospettiva diacronica, i composti verbo-nominali dell'italiano lungo un arco di tempo che va dalle Origini fino alle soglie del Novecento. In base ai dati tratti da diverse fonti quali corpora e opere lessicografiche, si ripercorre la storia di questo processo di formazione di parole mettendo in rilievo la natura di tali composti a livello semantico e categoriale. Si sostiene che la variabilità semantica (persona, strumento, evento, periodo di tempo) sia sempre stata presente nella storia linguistica dell'italiano, così come, in misura minore, l'ambiguità categoriale, per cui i composti VN fungono anche da modificatori aggettivali. Tale funzione, che negli ultimi decenni ha acquisito solidità dal punto di vista quantitativo, trova infatti riscontro già nei secoli precedenti. L'articolo, che parte dai presupposti teorici del Costruzionismo, dimostra come la Costruzione  $[V-N]_{A/N}$  deve la sua produttività sia alla flessibilità semantica e formale che presenta fin dalla fase antica, sia (e soprattutto) alla fortuna di particolari sotto-schemi semi-specificati che, in particolar modo negli ultimi decenni, fungono da modello per la creazione di nuove forme.

## PAROLE CHIAVE

composti VN, italiano antico, italiano contemporaneo, Morfologia delle Costruzioni, nomi di agente, nomi di evento, nomi di strumento, produttività

## DOI

<https://doi.org/10.14712/18059635.2020.1.4>



## 1. INTRODUZIONE<sup>1</sup>

In questo articolo, ci proponiamo di indagare la diacronia della composizione verbo-nominale in italiano che serve a creare parole del tipo *guardaboschi* o *portalettere*. Tale processo rappresenta senza dubbio uno dei fenomeni più studiati nel campo della formazione delle parole dell'italiano. Difatti, i composti verbo-nominali rientrano, non soltanto in italiano ma in tutte le lingue romanze (cfr. Giurescu 1975; Gather 2001), tra le strutture significative sia per la loro natura tipologica, dato che si tratta di una interessante innovazione diacronica, sia per la loro frequenza o produttività: “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” (Bauer 2011: 543).

A prescindere dalla questione tipologica, di cui non ci occuperemo in questa sede (ma cfr., ad esempio, Ledgeway 2012: 225),<sup>2</sup> le ricerche recenti si sono concentrate su almeno due aspetti, peraltro messi in rilievo nella citazione di Bauer riportata sopra, che sono direttamente collegati allo sviluppo della linguistica dei corpora. In primo luogo, si tratta dell'aspetto quantitativo riguardante soprattutto la produttività dei composti VN che si può verificare in grandi corpora sincronici (cfr. Ricca 2005; 2010).<sup>3</sup> In secondo luogo, si tratta di analisi diacroniche condotte, anche in questo caso, su vari corpora diacronici e su dizionari storici (cfr. ad es., per lo spagnolo Moyna 2011; per il francese Rosenberg 2007; 2008; 2011; per l'italiano Micheli 2020; Štichauer 2015; 2016).

In particolare, gli studi condotti in chiave diacronica hanno consentito di mettere in luce alcune delle questioni fondamentali, quali la varietà semantica inerente ai composti VN, lo sviluppo della loro produttività, la questione della categoria di uscita di tali composti, o anche le questioni teoriche che riguardano la natura del processo stesso. Sulla scia di questi studi recenti, ci proponiamo, in questo articolo, di combinare i due aspetti quantitativi, sia quello *corpus-based*, che sarà il più importante, sia

1 Il presente lavoro è frutto della collaborazione sistematica e continuativa tra i due autori. A M. Silvia Micheli si deve la stesura delle sezioni 2, 4.1, 4.2; a Pavel Štichauer la stesura delle sezioni 1, 3, 5. La sezione 4.3 è opera di entrambi gli autori. Questo articolo rientra nell'ambito del progetto “Creativity and Adaptability as Conditions of the Success of Europe in an Interrelated World” (No. CZ.02.1.01/0.0/0.0/16\_019/0000734).

2 Nella prospettiva della linguistica generativa, la variazione parametrica che riguarda l'ordine delle parole, o più precisamente *head parameter*, viene assunta anche al livello della morfologia (sia derivazionale, sia flessionale). Il principale cambiamento nell'evoluzione delle lingue romanze dal latino consiste nell'emergenza delle proiezioni funzionali e nell'inversione della posizione della testa, per cui si passa dalle strutture *head-final* a quelle *head-initial*. Tuttavia, come sostiene Ledgeway (2012: 225), “...the relevant change in directionality of the parameter did not occur at the same time in all areas of the grammar, but, rather, displays a staggered development proceeding at different rates in different morphosyntactic domains.”

3 “The main motivation for investigating this interesting word-formation procedure once again is given by the recent availability of data coming from very large textual corpora, which may be quite relevant from a theoretical perspective as well” (Ricca 2010: 237).

quello basato su alcune fonti lessicografiche, e di prendere in esame — lungo l'asse diacronico — i composti VN dell'italiano nell'arco del tempo che va dalle Origini alle soglie del Novecento.

L'articolo è organizzato come segue. In 2., descriveremo brevemente le premesse teoriche: in particolare, dopo aver definito la posizione dei composti VN nella classificazione proposta da Bisetto & Scalise (2005), introdurremo il quadro teorico della Morfologia delle Costruzioni. In 3., risumeremo le caratteristiche principali della composizione verbo-nominale in una prospettiva rigorosamente sincronica. Nella sezione 4., che rappresenta il nucleo del presente lavoro, ci occuperemo della diacronia del processo partendo, innanzitutto, da una presentazione del corpus e di altre fonti (lessicografiche) e dalla periodizzazione. In base al nostro campione, ci soffermeremo su tutte le questioni menzionate sopra. In 5., cercheremo di trarre alcune conclusioni generali.

## 2. PREMESSE TEORICHE

I composti VN dell'italiano e delle altre lingue romanze sono stati oggetto di numerosi studi (soprattutto sincronici) che, oltre a discuterne le proprietà formali e semantiche, ne hanno proposto una classificazione. Come si osserverà meglio di seguito, questo tipo di composti è generalmente considerato esocentrico, ossia privo di un elemento *testa*:<sup>4</sup> l'unica eccezione è rappresentata dalla proposta di Bisetto (1999) che invece li considera endocentrici. In particolare, secondo questa ipotesi, essi sarebbero costituiti da un nome d'agente o di strumento che manca del suffisso *-tore/trice*, per effetto di un processo di suffissazione zero o cancellazione (Bisetto 1999: 518): il composto *portalettere* andrebbe quindi interpretato come 'porta(tore) di lettere'. Tale ipotesi è tuttavia rimasta isolata rispetto alla gran parte degli studi che si è soffermata su questo tipo di composti, classificandoli come esocentrici.

Nella classificazione proposta da Bisetto & Scalise (2005), che adotteremo anche in questa sede (cfr. Figura 1), essi rappresentano un esempio di composti *subordinativi esocentrici*.

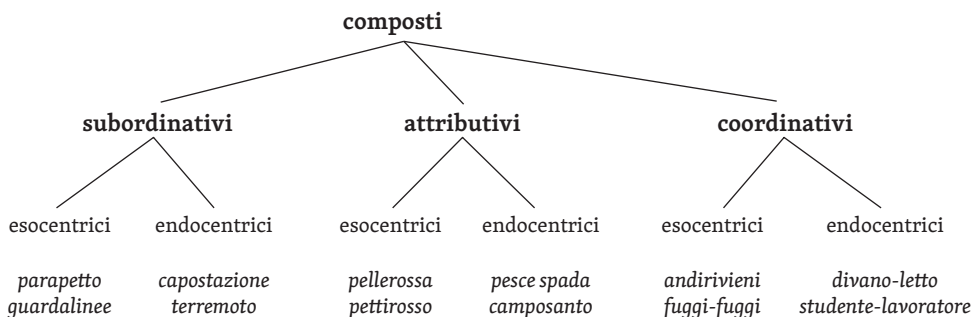


FIGURA 1. La classificazione dei composti Bisetto & Scalise (2005) con esempi dall'italiano

<sup>4</sup> Si vedano Bauer (2017: 29–40) per il concetto di testa di composto; Scalise & Guevara (2006; 2008) sull'esocentricità.



La classificazione proposta dai due autori, valida negli intenti per tutte le lingue del mondo, si basa sull'applicazione di due criteri gerarchici: il tipo di relazione grammaticale che lega i due costituenti; l'endo/esocentricità. Il primo criterio permette di distinguere tra composti subordinativi (SUB), attributivi (ATTR) e coordinativi (COORD). I composti del tipo *portalettere* fanno parte del primo tipo, poiché costituiti da un elemento verbale e un nome che ne costituisce un argomento. Il secondo criterio adottato è la presenza di un elemento testa:<sup>5</sup> ciascuna delle tre classi di composto fin qui individuate può quindi contenere sia composti endocentrici sia composti esocentrici. Questo aspetto costituisce un elemento fortemente innovativo rispetto ai precedenti tentativi di classificazione dei composti (si vedano Bloomfield 1933; Marchand 1969; Spencer 1991; Fabb 1998; Olsen 2001; Bauer 2001; Booij 2005), in cui l'esocentricità veniva considerata una classe a sé e posta sullo stesso piano di quella di "composto attributivo" o "composto subordinativo". Inoltre, diversamente dalle classificazioni tradizionali (si vedano, tra gli altri, Haspelmath & Sims 2010: 137-142; Olsen 2001), in cui i composti esocentrici erano quasi esclusivamente associati alla classe dei composti possessivi *bahvirihī* della grammatica sanscrita, in Bisetto & Scalise (2005) e Scalise & Guevara (2006; 2008), l'esocentricità è una proprietà riferita a una vasta gamma di composti, tra i quali i composti VN delle lingue romanze, che va a costituire una classe quindi tutt'altro che omogenea, dal punto di vista sia formale sia semantico.

La classificazione del 2005 è stata successivamente rivista in Scalise & Bisetto (2009): la principale innovazione introdotta riguarda la classe dei subordinativi, ripartita in due sottoclassi, i composti *ground* (il tipo *sala computer*) e *verbal-nexus* (il tipo *smaltimento rifiuti*), ciascuna delle quali presenta in italiano sia composti endocentrici (es. *raccolta rifiuti*) sia composti esocentrici (es. *portalettere*).

I composti VN romanzi sono stati oggetto di analisi anche nel quadro teorico della Morfologia delle Costruzioni (CxM; Booij 2010: 36-38; Masini 2016: 82), teoria che riprende e applica al piano morfologico i principali assunti della Grammatica delle Costruzioni (Goldberg 1995; 2013), in *primis* la nozione di *Costruzione* come associazione convenzionalizzata tra una forma e una funzione. In particolare, i composti VN vengono generalmente citati quali esempio per discutere l'esocentricità dei composti, e rappresentati come illustrato in (1) secondo il formalismo costruzionista.<sup>6</sup>

5 In Scalise & Guevara (2006: 584), si fornisce la seguente definizione di composto esocentrico: "Un composto *endocentrico* ha almeno una testa formale e almeno una testa semantica. Se un composto ha soltanto una testa formale e soltanto una testa semantica, queste devono coincidere nello stesso costituente. Tutti gli altri casi saranno considerati composti *esocentrici*."

6 Seguendo Masini & Audring (2019), in questo contributo utilizzeremo 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. < α >, < β >, < γ >, 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.).

- (1) a.  $\langle [[x]_{\text{Vak}} [y]_{\text{Npi-Nyj}}] \leftrightarrow [\text{STRUMENTO per } [\text{SEM}_k \text{ in R con SEM}_i]]_j \rangle$   
 b. *lavapiatti, asciugamani, lavastoviglie, etc.*

Nello schema riportato in (1a), l'esocentricità viene rappresentata attraverso l'uso di indici diversi nei singoli costituenti ( $\langle \alpha \rangle$  e  $\langle \beta \rangle$ ) e nell'intero composto ( $\langle \gamma \rangle$ ). Questo schema rappresenta, in particolare, i composti VN dell'italiano con semantica strumentale: come si osserverà nei seguenti paragrafi, questo tipo di composti si caratterizza infatti per una notevole varietà semantica (es. agentiva, strumentale, eventiva, etc.), rappresentabile secondo il formalismo costruzionista attraverso le perifrasi poste a destra dello schema.

Questo tipo di rappresentazione permette di rendere conto delle regolarità e sotto-regolarità di natura sia semantica sia formale riscontrabili nelle parole composte dell'italiano. Inoltre, l'uso delle perifrasi e degli indici permette di rappresentare efficacemente le proprietà non composizionali, ossia olistiche, presenti in diversi tipi di Costruzioni morfologiche, come i composti esocentrici o le reduplicazioni. Per questo motivo, nel presente lavoro, adotteremo il quadro teorico costruzionista per descrivere le proprietà formali e semantiche della composizione verbo-nominale dell'italiano.

### 3. CARATTERISTICHE PRINCIPALI DEI COMPOSTI VN NELL'ITALIANO CONTEMPORANEO

Prima di prendere in esame l'aspetto diacronico, ci soffermiamo brevemente in questa sezione sulla situazione odierna dei composti VN. Come già chiarito sopra, i composti verbo-nominali rientrano, nell'italiano contemporaneo, tra le strutture subordinative esocentriche, maggiormente produttive e ben descritte nella letteratura sull'argomento (cfr., ad es., Tollemache 1945; Scalise 1994: 134–136; Bisetto 2004: 45–47; Radimský 2006: 95–101; Scalise & Bisetto 2008: 133–135; Dardano 2009: 190–198; Ricca 2010). In quel che segue, riassumeremo brevemente le caratteristiche principali della composizione verbo-nominale, prescindendo dalla diversa impostazione teorica presente nella letteratura citata.

Per quel che riguarda la loro struttura interna, i composti VN vengono di solito definiti come un processo che consiste nella combinazione di una base verbale e una base nominale che rappresenta l'argomento interno del verbo: *portalettere*, *spazzacamina*, *lavapiatti*, *contagiri*. Vengono dunque escluse le strutture in cui si combinerebbe un elemento non-argomentale, come ad esempio *\*portacasa* (*portare a casa*), *\*lavacura* (*lavare con cura*) (cfr., ad es., Scalise 1994: 134).

Ci sono tuttavia composti in cui l'interpretazione argomentale presenta qualche difficoltà. A parte i casi ben noti, come *battiscopa*, *marciapiedi*, *batticuore*, *crepacuore*, ecc., in cui la relazione tra la base verbale e il nominale è diversa ma che sono generalmente lessicalizzati, ci sono, come sostiene Magni (2010), anche alcuni composti sincronicamente attivi come quelli in cui si trova la base verbale in *para-*, come in *parafango*, e che corrispondono ai loro omologhi *protège-* e *guarda-* rispettivamente del francese e dello spagnolo/portoghese. In tali composti, come sostiene Magni, la



lettura semantica della relazione tra la base verbale e quella nominale si discosta dal semplice rapporto predicato-oggetto diretto portando piuttosto all'interpretazione 'proteggere qualcosa da qualcosa/qualcuno': "Here, the overall meaning 'protect from N' implies in fact an underlying phrase with a prepositional phrase and a corresponding oblique object role for the noun constituent" (Magni 2010: 12).<sup>7</sup>

Tornando alle formazioni del tutto regolari, c'è da notare un fenomeno di cui ci occuperemo anche nella parte diacronica, e cioè che non tutti i verbi possibili si prestano alla composizione verbo-nominale, anzi il numero delle basi verbali è piuttosto basso. Seguendo Ricca (2010: 243-247), si può dire che le ragioni di tale situazione quantitativa dipendono da varie restrizioni che limitano la produttività del processo. Innanzitutto c'è, naturalmente, la già menzionata restrizione sintattica che impedisce la creazione dei composti VN a partire dai verbi non-transitivi. Tuttavia, anche nell'ambito dei verbi transitivi e agentivi la loro effettiva presenza come basi nei composti VN è alquanto ristretta. Pertanto, come sostiene Ricca (2010: 244), entrano sicuramente in gioco anche restrizioni di tipo fonologico, in particolare quella per cui vengono preferite le basi bisillabiche. Inoltre, esiste anche un vincolo marginale (nel senso che riguarda soltanto la sottoclasse dei verbi in *-ire*) che ostacola l'uso delle basi verbali uscenti in *-isci-* (con l'eccezione di *pulisci-*, *ripulisci-* e *distribuischi-*).<sup>8</sup>

Infine, gli studi condotti su basi lessicografiche (cfr. Gather 2001; Radimský 2006) e anche quelli condotti su corpora (Ricca 2010) hanno dimostrato che tra le basi verbali realmente presenti nella composizione verbo-nominale non tutte portano alla creazione di numerosi composti. Infatti, per le basi verbali principali, che Ricca (2010: 247) chiama *leading bases*, quali ad es. *porta-*, *salva-*, *mangia-*, *acchiappa-*, *ammazza-*, *copri-*, *taglia-*, *para-*, *spacca-* e *trita-*, il numero dei lemmi (ovvero *types*) cui danno origine esibisce una frequenza rapidamente decrescente: il maggior numero dei composti VN lo si osserva per la base *porta-*. Seguono poi *salva-*, *mangia-*, *acchiappa-*, ecc. (cfr. per dati precisi Ricca 2010: 247). Vale la pena introdurre, a questo punto, una distinzione su cui torneremo nella parte diacronica. Štichauer (2016: 112) considera il numero totale dei composti VN come *compound type frequency* e lo contrappone alla frequenza dei tipi cui dà origine una sola base verbale, come ad esempio *porta-*, che di fatto rappresenta la base verbale più produttiva. Quest'ultimo tipo di frequenza viene chiamata (*verbal*) *base type frequency*, ossia il numero totale dei composti costituiti da una base verbale.

Venendo infine all'aspetto semantico, i composti VN, com'è noto, esibiscono vari significati tra cui spiccano quelli di *agente* e *strumento*. Si riscontrano anche significati *eventivi* o *locativi*; e infine, si osserva che i composti VN assumono anche funzioni modificatrici che corrispondono al ruolo di un aggettivo sia qualificativo sia relazionale. Seguendo Ricca (2010: 238), si possono offrire i seguenti esempi per l'intero spettro semantico (che svilupperemo nella parte diacronica):

7 Tuttavia, se interpretiamo "parare" come "bloccare, schermare" (significato attestato in diacronia), la lettura Verbo+Oggetto diretto non pone problemi.

8 Dai dati tratti dal corpus *La Repubblica* emerge che tale restrizione non è assoluta. Infatti, Ricca (2010: 246) offre qualche esempio degno di attenzione, aggiungendo al contempo che il fatto che queste basi vengano evitate dipende anche dalla restrizione fonologica della bisillabicità.

- 1) Agente (N): *portalettere, guardaboschi*
- 2) Strumento (N): *apriscatole, portasci*
- 3) Evento (N): *baciamano, battibecco*
- 4) Luogo (N): *spartiacque, puntaspilli*
- 5) Qualificativo (A): *(panorama) mozzafiato, (discesa) spaccagambe*
- 6) Relazionale (A): *(cannone) sparaneve, (cane) annusadroga*

Ricca (2010: 250–251) offre, all'interno del suo campione, anche un'analisi della distribuzione dei vari significati, distribuzione basata sulla classe degli *hapax legomena* (e quindi focalizzata sul gruppo in cui si possono potenzialmente trovare formazioni neologiche). Innanzitutto, distinguendo i composti VN nominali da quelli aggettivali, si nota la loro distribuzione proporzionata (48,1% vs. 51,9%). All'interno dei composti nominali, anche i significati agentivo e strumentale coprono più o meno lo stesso numero di formazioni (344 agentivi rispetto a 274 strumentali). I composti eventivi sono invece in netta minoranza (29) e quelli locativi pressoché inesistenti: "(...) the hapax count confirms the well-known high productivity of both Agent and Instrument nouns (...). On the other hand, Location nouns as spartiacque 'water-divide', very rare in general, are practically not represented among hapaxes and low-frequency items. Event nouns, however, are a limiting case (...)" (Ricca 2010: 250).<sup>9</sup> Nell'ambito delle formazioni aggettivali (701 *types* ossia 51,9% degli *hapax*), dai dati di Ricca (2010: 250, tabella 6) emerge che i composti VN con funzione di aggettivo qualificativo rappresentano circa il 10%; seguono poi quelli relazionali (40%) ed altri la cui funzione corrisponde ad una modificazione in senso piuttosto lato: "(...) most VN compound modifiers which cannot be considered as relational adjectives (...) could hardly be labelled as qualifying adjectives either. Rather, they express a looser kind of modification, with transient/eventive character, sometimes bordering on the function of a (reduced) relative clause" (Ricca 2010: 251).

Come vedremo in seguito, tale presenza significativa dell'uso aggettivale dei composti VN rappresenta una vera e propria novità considerata la loro scarsa attestazione nei periodi qui presi in esame.

#### 4. ANALISI DIACRONICA DEI COMPOSTI VN

##### 4.1 METODOLOGIA DI INDAGINE

Come già accennato, questo lavoro si basa su dati estratti sia da corpora sia da risorse lessicografiche. In particolare, i corpora utilizzati sono due: il corpus diacronico *CoDI\_t\_com* (*Corpus Diacronico dell'Italiano per la composizione*); il corpus LIZ 4.0 (*Lettera-*

9 In seguito, Ricca (2010: 251) allude anche all'aspetto diacronico degli *event nouns*: "The fact that there are undoubtedly very few new formations with Event semantics (2.1% of the hapax total) with respect to both Agentive and Instrument VN compounds is not enough to state that the former have undergone a substantial loss of productivity, because they are a minority among established words as well (4.8% of all the non-hapax types), and it is likely that this semantic output has always been marginal in the history of Italian."



tura Italiana Zanichelli). Nel primo caso si tratta di un corpus diacronico costruito *ad hoc* per lo studio della composizione italiana in diacronia (cfr. Micheli 2020: 49–58), costituito da testi non annotati<sup>10</sup> scritti nel periodo compreso tra le Origini e il 1947;<sup>11</sup> la risorsa si articola in cinque sottocorpora delimitati da date significative per la storia della lingua italiana.<sup>12</sup> La struttura del corpus è illustrata nella Tabella 1.

	1	2	3	4	5
	Origini–1375	1376–1532	1533–1691	1692–1840	1841–1947
token	4.247.611	6.133.344	6.346.590	6.046.303	6.543.891

Totale token: 29.317.739

Totale testi: 603

**TABELLA 1.** Struttura del corpus CoDIt\_com

Il corpus è stato interrogato attraverso il software AntConc (Anthony 2013); l'estrazione dei composti VN dai sottocorpora 2-3-4-5, presi in considerazione nel corso dell'analisi diacronica (vedi oltre per la scansione temporale adottata), è avvenuta attraverso lo spoglio manuale di una versione ridotta della *wordlist* dell'intero corpus.<sup>13</sup> Il campione di forme ottenuto è costituito da 311 composti VN.

Nel secondo caso, il corpus LIZ è stato utilizzato soltanto per il periodo post-cinquecentesco e soltanto per integrare i dati già presenti nel primo corpus (in particolare là dove contiene alcuni testi interessanti in cui si possono trovare composti verbo-nominali ancora non riscontrati). LIZ 4.0, infatti, raccoglie testi italiani dalle Origini all'inizio del Novecento, ma si tratta di un corpus eterogeneo, non bilanciato sia dal punto di vista quantitativo che qualitativo. Quantitativamente, il corpus contiene, anche per i periodi più recenti, un numero di testi non comparabile (ad esempio, per il Cinquecento, la LIZ 4.0 offre più di 200 testi la cui dimensione, in termini di *tokens*, è di circa 10 milioni di occorrenze, mentre per il Seicento, la dimensione si riduce a soli 49 testi con circa 3 milioni di occorrenze). Dal punto di vista qualitativo, la LIZ 4.0 deve essere considerata con maggiore cautela dal momento che contiene anche testi che rientrano in varietà locali diverse dal *toscano* e dall'italiano letterario post-cinquecentesco (e, ovviamente, tali testi non sono stati presi in esame).

10 Si è scelto di utilizzare un corpus di testi non annotati, dati i numerosi problemi che sarebbero emersi nell'annotazione per parti del discorso di testi appartenenti a fasi storiche precedenti a quella contemporanea.

11 Il corpus non è ancora disponibile in rete; l'annotazione e la pubblicazione del corpus avverranno in una fase successiva.

12 Il bilanciamento del CoDIt\_com riprende quello del corpus MIDIA, al quale si rimanda per un approfondimento sui criteri adottati per la periodizzazione. Il corpus MIDIA e la relativa documentazione sono disponibili in rete al seguente indirizzo: <http://www.corpusmidia.unito.it/> (ultimo accesso: 10.06.2019).

13 In particolare, la *wordlist* totale è stata ridotta escludendo le parole più corte di 6 caratteri, lunghezza minima di una parola composta (es. *malora* < *mal(a)+ora*), stimata sulla base di un campione di parole composte dell'italiano contemporaneo.





Inoltre, si è fatto ricorso anche a fonti lessicografiche che, pur rappresentando risorse problematiche, possono offrire esempi isolati, degni di particolare attenzione riguardo alle questioni teoriche che ci siamo posti sopra. Pertanto, ci rifaremo anche alla testimonianza delle quattro edizioni<sup>14</sup> del *Vocabolario degli Accademici della Crusca*, che vanno dal 1612 al 1738, e ad altre opere lessicografiche coeve (quali, ad esempio, il celebre dizionario bilingue italiano-inglese di John Florio)<sup>15</sup>, per arrivare fino al *Tommaseo*, il maggior vocabolario ottocentesco.

Infine, per l'analisi della composizione verbo-nominale nella fase antica, è stato consultato il lemmario del TLIO (*Tesoro della Lingua Italiana delle Origini*),<sup>16</sup> da cui sono stati estratti manualmente 79 composti VN.

L'analisi proposta nei seguenti paragrafi si articola in due parti principali che corrispondono a una scansione cronologica ben motivata. Nella prima parte (§4.2), infatti, ci occuperemo del periodo che si estende dalle Origini all'inizio del Cinquecento e, nella seconda parte (§4.3), prenderemo in esame il periodo post-cinquecentesco. Tale divisione copre dunque, da una parte, i primi due sottocorpora del corpus diacronico CoDIt\_com (dalle Origini al 1375, dal 1375 al 1532), e dall'altra i tre sottocorpora che corrispondono al periodo dal 1533 al 1947.<sup>17</sup> In termini teorici, la motivazione per tale scansione risiede nella natura dell'evoluzione dell'italiano letterario che, sulla scia di ciò che sostiene Tesi (2001; 2005), cercheremo di riassumere.

L'italiano letterario, attraverso i periodi che prenderemo in considerazione, può essere definito in base ad alcuni parametri introdotti da Tesi (2001: 5-9): in particolare, la natura della comunità linguistica e la norma linguistica che viene utilizzata

- 
- 14 Com'è noto, la quinta edizione della *Crusca*, in preparazione dal 1863, fu interrotta alla lettera O nel 1923 e la preparazione di un nuovo vocabolario nazionale fu affidata all'Accademia d'Italia (cfr. Marazzini 2009: 383-385). Per la consultazione del *Vocabolario*, ci siamo avvalsi delle edizioni liberamente accessibili sul sito ufficiale creato e gestito dall'*Accademia della Crusca* (<http://www.lessicografia.it/>).
- 15 Il dizionario di John Florio, *Queen Anna's New World of Words, or a Dictionarie of the Italian and English Tongues*, pubblicato nella sua seconda edizione nel 1611 — e quindi quasi contemporaneamente alla prima edizione della *Crusca* — rappresenta uno dei dizionari bilingui più importanti soprattutto per la sua ricchezza (cfr. Marazzini 2009: 107-108; Pizzoli 2004: 32-35) che si rispecchia anche nel numero dei composti VN registrati che supera nettamente quello della prima *Crusca*. Il dizionario di Florio è consultabile in rete al seguente indirizzo: <http://www.pbm.com/~lindahl/florio/> (ultimo accesso: 29.07.2019). Inoltre, esiste anche l'edizione critica della prima edizione del 1598, *A Worlde of Wordes. A critical edition with an introduction by Herman W. Haller*, Lorenzo da Ponte Italian Library series. Toronto: University of Toronto Press, 2013.
- 16 Il TLIO è un dizionario dell'italiano antico basato sul corpus OVI (*Opera vocabolario Italiano*), disponibile in rete: <http://gattoweb.ovi.cnr.it/>. Il lemmario del TLIO è liberamente accessibile al seguente indirizzo: <http://tlio.ovi.cnr.it/TLIOlemm/index.php> (ultimo accesso: 10.06.2019).
- 17 Si precisa che l'analisi del periodo Origini-1375 è basata sui dati estratti dal Lemmario del TLIO; il periodo 1376-1525 è indagato sulla base dei dati estratti dal secondo sotto-corpus del CoDIt\_com. Infine, per l'analisi diacronica fornita al paragrafo 4.3 si sono utilizzati i sottocorpora 3-4-5 del CoDIt\_com.



come un criterio di periodizzazione “interna”. Come è ben noto, l’italiano odierno, che esiste ormai sia nella variante scritta che in quella parlata, si basa, per certi aspetti (ma non per tutti) su una lingua di uso esclusivamente scritto. Da un punto di vista della comunità linguistica si può proporre, come prima definizione, di considerare l’italiano antico come basato su un *codice scritto* (cfr. Tesi 2001: 6). Tale codice scritto crea una norma linguistica che può essere sia *esplicita*, basata su una codificazione in forma, ad esempio, di grammatiche, sia *implicita* che “estrae” tale codificazione da testi che fungono da modello. L’italiano, dal Duecento al Cinquecento, è stato, di fatto, una lingua con una norma implicita basata su tali testi-modello. La distinzione tra norma implicita e norma esplicita consente di proporre una periodizzazione più ampia che, come afferma Tesi (2001: 8), “*si basa sull’alternanza di norme linguistiche (o standard linguistici) che ridisegnano il comportamento linguistico della collettività.*” In questa prospettiva, l’italiano letterario, dal punto di vista diacronico, conosce due periodi principali: il primo, caratterizzato dalla norma implicita basata su un canone di testi letterari, che va dal Duecento ai primi del Cinquecento (per la precisione fino al 1525); il secondo, basato sulla norma esplicita (basata già non soltanto su testi modello ma anche su opere esplicitamente normative), che va dal secondo Cinquecento fino ad oggi (cfr. Tesi 2001: 8).

#### 4.2 DALLE ORIGINI ALL’INIZIO DEL CINQUECENTO

Il lemmario del TLIO include 79 composti costituiti da un elemento verbale e un nome.<sup>18</sup> L’origine di questo insieme di forme è, nella maggior parte dei casi, nativa: sono infatti solo 12 i composti VN che entrano in italiano come prestiti o calchi dal francese antico o dal latino, come illustrato dalla Tabella 2.

Francese antico	Latino
<i>cantafavola</i> (cfr. <i>cantefable</i> )	<i>girasole</i> (cfr. <i>heliotropium</i> )
<i>castigabriccone</i> (cfr. <i>chastiermusart</i> )	<i>mordigallina</i> (cfr. <i>morsus gallinae</i> )
<i>covriceffo</i> (cfr. <i>couvre-chef</i> )	<i>placamarito</i> (cfr. <i>viriplaca</i> )
<i>guardacorpo</i> (cfr. <i>garde-corps</i> )	<i>volgicuore</i> (cfr. <i>verticordia</i> )
<i>guardacuore</i> (cfr. <i>garde-corps</i> )	
<i>guardanappa</i> (cfr. <i>garde-nappe</i> )	
<i>guardaroba</i> (cfr. <i>garde-robe</i> )	
<i>tagliaborse</i> (cfr. <i>coupe-bourse</i> )	

**TABELLA 2.** Composti VN attestati in italiano antico di origine esogena

<sup>18</sup> Sono invece 6 i composti che presentano la preposizione *in* o l’articolo determinativo *il* tra l’elemento verbale e il nome, ossia *bagnalcore*, *battiloro*, *giralsole*, *guardancanna*, *saltamin-dosso*, *sperandio*. Se semanticamente queste forme non presentano sostanziali differenze con i composti VN senza preposizione/articolo, dal punto di vista formale esse appaiono come univerbazioni di espressioni sintattiche cristallizzate, piuttosto che il risultato di un procedimento puramente morfologico. Per questo motivo, si è scelto di separarle dal campione di composti VN analizzato.



I dati riportati nella Tabella 2 mettono in luce una notevole influenza del francese antico, in particolare per quanto riguarda i composti con *guarda-* ‘proteggi, conserva’ come primo costituente, il quale, come vedremo, diventerà una base verbale particolarmente produttiva. D'altra parte, il legame con il latino sembra piuttosto limitato: tra le forme riportate nella tabella, soltanto *verticordia* presenta una struttura VN corrispondente a quella italiana; negli altri casi si tratta di calchi imperfetti (ad es. *placamarito* da *viriplaca*).<sup>19</sup>

Dal punto di vista morfologico si può osservare nella Tabella 3 la distribuzione dell'intero campione di composti VN rispetto alla coniugazione dell'elemento verbale.<sup>20</sup>

-are	-ere	-ire
61	16	2

**TABELLA 3.** Distribuzione dei composti VN rispetto alla coniugazione dell'elemento verbale

Come si osserva dai dati riportati nella Tabella 3, i composti VN dell'italiano antico si formano prevalentemente da verbi appartenenti alla classe flessiva in *-are*: in particolare, *guarda-* e *para-*, che occorrono rispettivamente in 8 e 5 composti, sono le basi verbali più frequenti (in termini di *verbal base type frequency*, che abbiamo introdotto sopra). La produttività dei verbi di seconda coniugazione è data prevalentemente dalla base *batti-* che occorre in sette forme (*batticuore*, *battifolle*, *battigote*, *battilana*, *battipetto*, *battisacco*, *battisfancello*). Infine, l'unico verbo della terza coniugazione attestato è *coprire*, che occorre in *covriceffo* e *coprifuoco*.

Per quanto riguarda la categoria di output dei composti attestati nel lemmario del TLIO, va notato che questi vengono utilizzati sempre come sostantivi e non sembrano svolgere la funzione di modificatori/aggettivi largamente attestata in italiano contemporaneo, come già menzionato sopra.

Dal punto di vista semantico, i composti VN dell'italiano antico mostrano una notevole varietà semantica (che oltrepassa le etichette semantiche proposte da Ricca, 2010 e introdotte sopra): come mostrato dai seguenti esempi, possono infatti indicare animali (es. *fottivento* ‘lo stesso che gheppio’ in 2a), strumenti (es. *affettapane* ‘coltello per affettare il pane’ in 2b), persone (es. *cavadenti* ‘chi per mestiere estrae i denti’ in 2c), eventi (es. *crepacuore* ‘grande afflizione dell'animo’ in 2d), luoghi (es. *tornafine* ‘confino, confine’ in 2e), piante (es. *girasole* in 2f), periodi di tempo (es. *coprifuoco* ‘ora a partire dalla quale era vietato ai cittadini di uscire di casa’ in 2g).<sup>21</sup>

19 Questa situazione è in linea con quanto brevemente descritto all'inizio dell'articolo. Infatti, il latino non possedeva la struttura VN esibendo invece il *pattern* NV, ad es. *lucifer*, *vexillifer* ecc. (cfr. Brucale 2012; Štichauer 2016: 110–111).

20 In questo contributo non ci soffermeremo sulla questione relativa alla natura del primo costituente dei composti VN, a cui sono stati dedicati numerosi contributi (si veda Bisetto, 1999 per un quadro delle diverse posizioni). Si rimanda a Micheli (2020: 118–120) per una proposta di analisi, sulla base dei dati estratti dal lemmario del TLIO.

21 I significati dei composti riportati qui e di seguito sono tratti dal TLIO.

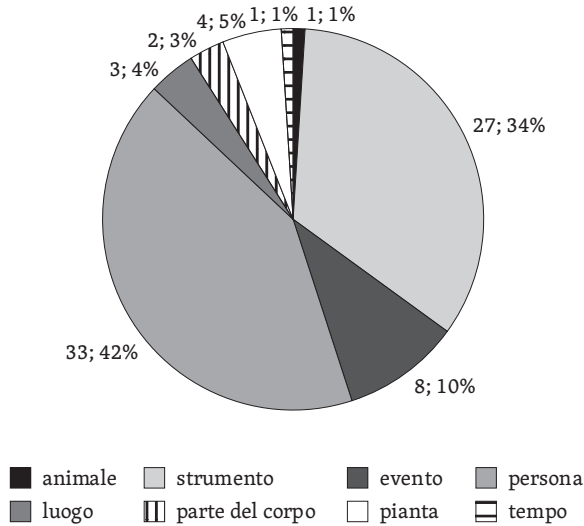


- (2) a. Piglia lo sterco del *fottivento*, e stemperalo con acqua piovana (Anonimo, *Libro della cura delle malattie*)
- b. A noia m'è chi chon l'*afettapane* / bruttura taglia, cielatto o palese, / po' l'usa a mensa da sera e di mane. (Pucci, *Le noie*)
- c. Se questa medicina non giova, fa di mestiere servirsi del *cavadenti*. (Anonimo, *Libro della cura delle malattie*)
- d. [...] perché non pate lo cuore dell'uomo di essere suddito a chi non è suo Signore per ragione, anzi è troppo gran *crepacore* di vedersi in capo per maggiore colui, che dee esser pari, o minore. (Cavalca, *Esposizione del Simbolo degli Apostoli*)
- e. Ma santo Eusepio e Beda ne le loro Croniche non dicono che fosse mandato a *tornafine*, ma pur che, cagendo elli in molte miserie, con la sua propria mano s'uccise. (Anonimo, *Leggenda aurea*)
- f. "Io mirava". Qui poetando vuole mostrare, sì come le foglie dell'erba chinate per lo notturno freddo si rilevano, e seguono quasi l'andare del Sole, sì come appare nel *girasole*; così il suo sguardo seguiva ogni movimento della detta femina. (Anonimo, *Ottimo commento alla Commedia, Purgatorio*)
- g. Perdio, aprimi, ché tosto sonerà *coprifuoco*; e s'io sarò trovato qui, io sarò preso e messo in prigione e domattina battuto. (Anonimo, *Libro dei Sette Savi*)

Gli esempi riportati mostrano come la composizione verbo-nominale sia ampiamente sfruttata già in italiano antico per la creazione di parole dalla semantica molto varia. Un quadro più chiaro della diffusione di questo tipo di composti può essere ottenuto osservando, dal punto di vista quantitativo, come questi si distribuiscono tra i diversi tipi semantici.

Come illustrato dal grafico nella Figura 2, i composti VN dell'italiano antico sono prevalentemente utilizzati per indicare persone (42%), strumenti (34%) ed eventi (10%); gli altri tipi semantici risultano nel complesso poco sfruttati.

Analizzando più nel dettaglio i raggruppamenti semantici individuati (cfr. Tabella 4), si possono notare ulteriori regolarità: in particolare, tra i composti VN che designano referenti umani è possibile identificare un nutrito gruppo di appellativi con connotazione negativa, tra i quali ad esempio *baciapolvere* 'chi dimostra un affettato eccesso di devozione', *leccascodelle* 'chi mangia ingordamente fin gli ultimi avanzi', *scialacquavino* 'chi beve vino in maniera eccessiva, ubriacone'. Nella maggior parte dei casi, si tratta di creazioni estemporanee per designare individui che eccedono nel loro comportamento, destinate a non radicarsi nel lessico. Sarà questo il motivo per cui la registrazione di tali composti nei dizionari (sia nelle prime edizioni della *Crusca*, all'inizio del Seicento, che in altre opere lessicografiche) risulterà particolarmente debole. Più rari sono invece i composti VN agentivi che esprimono l'azione svolta abitualmente da un individuo per mestiere: è il caso di *cavadenti* 'dentista' e *portapesi* 'facchino'.

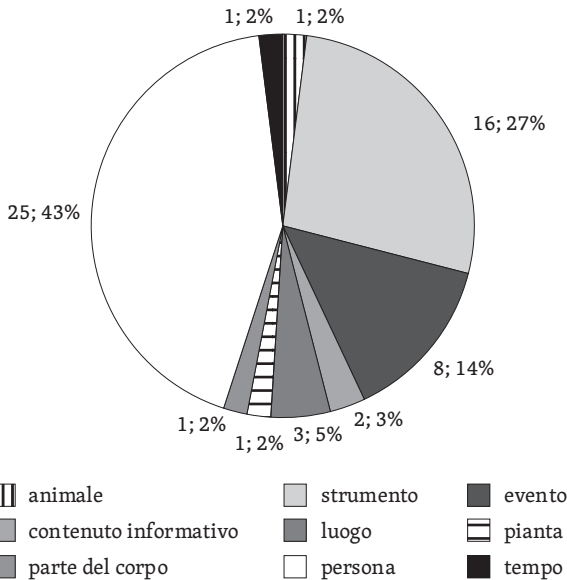


**FIGURA 2.** Distribuzione dei composti VN dell'italiano antico rispetto al tipo semantico di output

TIPO SEMANTICO	COMPOSTI VN dell'italiano antico
ANIMALE	<i>fottivento</i>
STRUMENTO	<i>affettapane, battigote, battisacco, cacciascudo, cangiacolore, cassapiedi, castigabriccone, castigapazzi, covriceffo, curaorecchie, drizzacrine, fattibello, foraterra, grattacacio, guardabanco, guardacuore, guardanappa, mesciroba, parabordone, paramosche, parapetto, parasole, rizzacrino, stracciafoglio, salvadanaio</i>
EVENTO	<i>batticuore, battilana, cantafavola, crepacuore, guazzabuglio, leccamuffo, squarciasacco, stracciasacco</i>
PERSONA	<i>baciapolvere, battipetto, battisfancello, beccalite, beccamorto, cacapensieri, cacciapensieri, cavadenti, chiudiborsa, guardacorpo, guardamorto, guardapasso, guastamondo, lavaceci, lavascodelle, leccascodelle, mettifuoco, pappaculo, pascibietola, pelacane, piantamalanno, picchiapetto, placamarito, portapesi, raffacane, sbiecaciglio, scannadio, scialacquavino, squarciavela, squassacapo, tagliaborse, volgicuore, vuotacessi</i>
LUOGO	<i>guardaroba, guardaspensa, tornafine</i>
PARTE DEL CORPO	<i>paracuore, pascipeco</i>
PIANTA	<i>girasole, mirasole, mordigallina, tornasole</i>
TEMPO	<i>coprifuoco</i>

**TABELLA 4.** Composti VN dell'italiano antico classificati in base al tipo semantico di output

Per quanto riguarda i composti VN con semantica eventiva, in due casi (*squarciasacco* e *stracciasacco*) questi occorrono soltanto all'interno di una locuzione avverbiale: in particolare, *a stracciasacco* e *a squarciasacco*, aventi entrambi il significato di 'in modo ostile'. Va inoltre notato che, in questo tipo di forme, il rapporto tra l'elemento ver-



**FIGURA 3.** Distribuzione dei composti VN attestati nel periodo 1376–1525 rispetto al tipo semantico di output

bale e il nome può non essere quello di Verbo-Complemento oggetto, come nel caso di *batticuore*, interpretabile come ‘battito del cuore’, in cui *cuore* rappresenta il soggetto.

Riassumendo, in italiano antico la composizione verbo-nominale è utilizzata prevalentemente per la creazione di nomi di strumento, referenti umani, eventi, ma non solo: la variabilità di significati osservata nel campione di composti estratti dal lemmario del TLIO mette in luce che questa Costruzione appare fin dalle Origini relativamente produttiva e ‘flessibile’ dal punto di vista semantico.

Per quanto riguarda il periodo compreso tra il 1376 e il 1525, il sottocorpus 2 attesta 58 composti VN, di cui 37 hapax. Dal punto di vista semantico, questi si distribuiscono tra i diversi tipi semantici come illustrato dal grafico nella Figura 3.

Il quadro emerso dall’analisi semantica dei composti attestati in italiano antico è sostanzialmente confermato dai dati estratti dal secondo sottocorpus del CoDIt\_com e riportati nella Figura 3: la composizione verbo-nominale è usata anche in questo periodo prevalentemente per la creazione di nomi che designano persone, strumenti, eventi.

Riassumendo, si può dunque affermare che il periodo che va dalle Origini al primo Cinquecento rappresenta un arco di tempo abbastanza omogeneo per quanto riguarda la presenza e la struttura interna dei composti VN. Tutte le possibilità strutturali sembrano già essere presenti (in particolare, la coesistenza dell’output semantico di agente/persona e strumento, che di solito viene fatta risalire ai periodi più recenti).

## 4.3 DAL CINQUECENTO AL PRIMO NOVECENTO

Come già osservato, l'analisi diacronica si basa sui sottocorpora 3-4-5 del CoDit\_com, dai quali sono stati estratti 282 composti VN. Un primo quadro generale della diffusione di questo tipo di forme si ha osservando la frequenza dei type e dei token nei singoli sottocorpora, cfr. Figura 4.

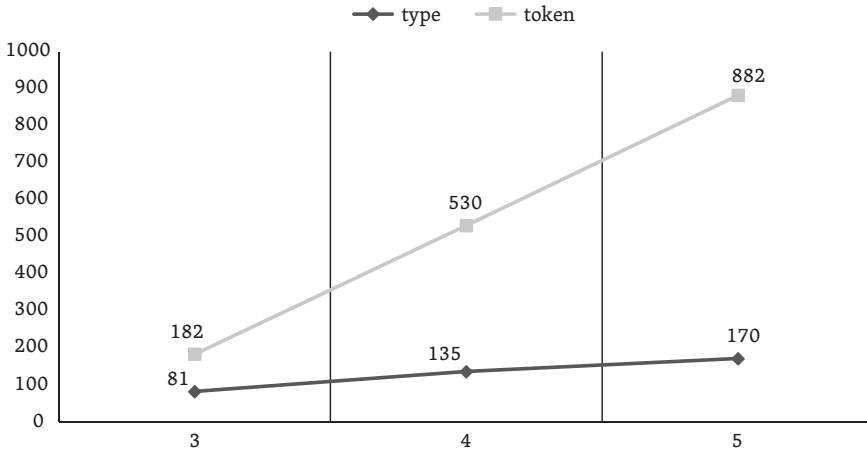


FIGURA 4. Frequenza dei type e dei token dei composti VN in diacronia per il periodo dal 1533 al 1947

Come illustrato dal grafico nella Figura 4, il numero dei type, ossia dei nuovi composti VN attestati nel corso dei secoli, cresce stabilmente nel corso dei secoli: il sottocorpus più recente contiene infatti quasi il triplo dei composti attestati nel secondo sottocorpus (periodo 1376–1532). Differente è invece l'andamento della frequenza dei token: in questo caso, si osserva una più rapida crescita in corrispondenza del terzo sottocorpus (che copre, vale la pena di ricordarlo, il periodo dal 1533 al 1691), a partire dal quale la frequenza dei composti VN aumenta considerevolmente. In particolare, è interessante osservare quali sono, per ciascun sottocorpus, i tre composti più frequentemente attestati (Tabella 5).

3	4	5
<i>guardaroba</i> (25)	<i>passaporto</i> (88)	<i>parapetto</i> (55)
<i>passatempo</i> (25)	<i>passatempo</i> (47)	<i>marciapiede</i> (53)
<i>guardacamera</i> (7)	<i>portalettere</i> (19)	<i>passaporto</i> (37)

TABELLA 5. Composti VN più frequenti nei sottocorpora 2–5

Come mostra la Tabella 5, i composti più frequenti nei tre periodi considerati presentano sempre un elemento verbale bisillabico, appartenente alla coniugazione in *-are*. Dal punto di vista semantico, si tratta prevalentemente di nomi indicanti oggetti concreti, strumenti: *parapetto*, *guardaroba*, *passaporto*; *marciapiede* ha invece semantica



locativa, mentre *passatempo* è un composto eventivo. Soltanto la forma *portalettere* indica un referente umano: come mostrano gli esempi in (3), essa è attestata sia in funzione di modificatore (3a) sia in quella di sostantivo (3b).

(3) a. Insomma, chi favella ogni ora per bocca del compagno è simile a uno corriere *portalettere* che, senza averci mai nulla del suo, è continuo ingombrato di tutti gli interessi d'altri. (Aretino, *Lettere sull'arte*)

b. I latini scrivevano bene spesso in greco del loro. Così fa molte volte Cicerone nelle epistole ad Attico (forse anche nelle altre); dove forse per non essere inteso dal *portalettere*, la qual gente, com'egli dice, soleva alleviare la fatica e la noia del viaggio leggendo le lettere che portava; [...] (Leopardi, *Zibaldone*)

Tuttavia, per completare il quadro, si può aggiungere anche la testimonianza di un testo cinquecentesco non presente nel corpus CoDIt\_com ma incluso nella LIZ 4.0, e cioè *Piazza universale di tutti i mestieri* di Tommaso Garzoni, in cui il composto *passaporto*, presto diventato di lettura esclusivamente strumentale, compare anche come nome di persona:

c. I *passaporti* erano detti hippaggia, ovvero pontones, secondo Apuleio. (...) L'ufficio de' quali è sempre stato di passare i viandanti, riscuoter le gabelle debite a loro, impedire il transito de' fuorusciti o d'altra gente sospetta, guardare le robbe che passano, usar gran diligenza intorno ai contrabandi, e non far trarre le persone, come oggidì s'usa da molti, stentarli nel passo, chiederli più dell'ordinario, non voler'essentare quelli che son privilegiati. (Garzoni, *Piazza universale... Disc. 144, 3*).

Tornando all'esempio in (3a), esso è particolarmente rilevante perché testimonia l'uso di un composto VN con funzione di modificatore, assente nella fase antica e qui attestato in un testo della prima metà del Cinquecento. In realtà, analizzando più nel dettaglio la categoria di output dei composti attestati nel CoDIt\_com, si possono collocare già nel periodo 1376–1532 (secondo sottocorpus) le prime attestazioni di composti VN con valore attributivo, come mostrato dai seguenti esempi:

(4) a. Questo Vescovo *lavaceci*, vogliendo ammaestrare nel vizio della gola, riprende 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 (4) mostrano l'uso di composti VN in funzione di modificatori di sostantivi: in particolare, in (4a) e (4b), i due composti *lavaceci* e *vuotacessi* si rife-





riscono a referenti umani; in (4c), il composto *rubacuori* è riferito agli occhi di una donna. Va inoltre notato che i primi due composti occorrono all'interno del corpus anche come sostantivi, mentre il terzo esclusivamente come modificatore.

I dati estratti dal CoDIt\_com ci permettono di osservare che l'uso in funzione attributiva dei composti VN rimane comunque raro nel corso di tutta la storia della lingua italiana: dei 285 composti attestati nei sottocorpora del CoDIt\_com, soltanto 12 occorrono anche come modificatori, nella maggior parte dei casi riferiti a referenti umani: *castraporci*, *lavaceci*, *mangiapolenta*, *perdiggiorni*, *portalettere*, *rompicollo*, *rubacuori*, *scanapagnotte*, *scansafatiche*, *storcileggi*, *tagliapane*, *vuotacessi*. Si noti che alcuni di questi composti occorrono esclusivamente in funzione di modificatore: si tratta di 'medico *castraporci*', 'occhi *rubacuori*', 'dottore *storcileggi*', 'coltelli *tagliapane*'. Questi dati apportano un'importante rettifica all'analisi di Štichauer (2015: 141-142; 2016: 127-128) che, in base ai dati tratti esclusivamente dalla LIZ 4.0, fa risalire il primo esempio dell'uso modificatore/aggettivale al Settecento.<sup>22</sup> In realtà, come si vede dagli esempi appena citati in (4), la funzione aggettivale può essere considerata come *disponibile* sin dal periodo dell'italiano antico.

Nel complesso, i tre sottocorpora del CoDIt\_com considerati attestano 135 basi verbali diverse, di cui le più produttive risultano essere *porta-* (18 type), *guarda-* (15 type), *batti-* (15 type), *para-* (12 type), *mangia-* (10 type). Si tratta di elementi verbali con struttura prevalentemente bisillabica (fanno eccezione soltanto 16 forme su 135):<sup>23</sup> tale restrizione fonologica, già riscontrata in italiano antico, è quindi presente senza soluzione di continuità nel corso dei secoli. Per quanto riguarda la coniugazione dell'elemento verbale, i dati diacronici confermano una netta tendenza a selezionare basi appartenenti alla classe flessiva in *-are*, in particolare, 244 composti (type) su 282 totali; sono invece 34 i composti con verbi appartenenti alla seconda coniugazione e 4 quelli con verbi della terza coniugazione.

Tale quadro combacia, anche se non del tutto, con la testimonianza delle fonti lessicografiche. Messe da parte le quattro edizioni del *Vocabolario degli Accademici della Crusca*, in cui la registrazione dei composti VN risulta essere particolarmente scarsa (cfr. Štichauer 2015: 134-136), sia il dizionario di John Florio (nell'edizione del 1611), che il *Tommaseo* offrono parecchi esempi di composti VN con le basi verbali più frequenti. Ad esempio, Florio registra 5 esempi con *porta-*, lo stesso numero di composti con *salva-*, 6 composti con *mangia-*, 7 con *batti-* e ben 12 esempi con la base *guarda-*. La stessa proporzione, con numeri di *type* notevolmente aumentati, la si riscontra anche nel *Tommaseo*. Infatti, anche nel *Tommaseo*, la base più frequente, *guarda-*, offre 30 composti, seguita poi da *porta-* con 29 composti, 27 con *mangia-*, e infine *para-* e *batti-* che così rispecchiano la situazione descritta in base ai corpora.

Escludendo queste basi particolarmente frequenti, i restanti composti si distribuiscono tra ben 130 basi diverse: questo dato induce a ritenere che la produttività

22 "(...) e ce lo mandò a bordo il vascello da guerra **guardacoste**, che sta sempre ancorato a quattro miglia di distanza" (Francesco Algarotti, *Viaggi di Russia*, Lettera 3,7).

23 In particolare, si tratta delle seguenti basi verbali: *arruffa-*, *fa-*, *scrutina-*, *affitta-*, *stuzzica-*, *schiccar-*, *asciuga-*, *appiccica-*, *spaventa-*, *acchiappa-*, *castiga-*, *cavalca-*, *strascina-*, *imbratta-*, *scavezza-*, *acconcia-*.



mostrata da questo tipo di composti riguarda l'intera Costruzione astratta [VN] oltre che 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 che trasporta un oggetto oppure uno strumento che contiene un oggetto; i due significati della base verbale *porta-* sono esemplificati in (5).

- (5) 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 convivere*)
- b. [...] sul piano di bardiglio, una boccia d'acqua col bicchiere capovolto, un *portacenero*, un *portorologio* (Pirandello, *Maschere nude*)

Il composto *portabaldacchini*, esemplificato in (5a), 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 *portacenero* e *portorologio*, si tratta di oggetti che contengono qualcosa.<sup>24</sup> Come osservato nel precedente paragrafo, 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 (6).

- (6) 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 (6) mostrano come il composto sia utilizzato per indicare uno strumento che contiene fogli o documenti, una 'cartellina'. Dal suo ingresso nel lessico dell'italiano, il composto ha avuto notevole fortuna in termini di frequenza (esso occorre 13 volte nel quarto sotto-corpus e 27 nel quinto) e può aver certamente influito nella creazione e diffusione del significato '(strumento) che contiene x' del sotto-schema [*porta-x*].<sup>25</sup> Questo caso dimostra come un prestito/

<sup>24</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 VN (cfr. De Mauro 2009, s.v. *portare*).

<sup>25</sup> A comprova di questa ipotesi, si possono osservare alcuni esempi di composti *porta-N* registrati da vari dizionari specialistici o anche bilingui. Ad esempio, Štichauer (2015: 137-139)

calco possa contribuire alla creazione e alla fortuna di un particolare sotto-schema che costituisce un'innovazione dal punto di vista della forma o, come in questo caso, della semantica.

Dal punto di vista semantico, la Tabella 6 illustra la distribuzione dei composti VN attestati nei sottocorpora 3-4-5 del CoDIt\_com rispetto al tipo semantico di output.

Tipi semantici	Composti (type)
PERSONA	147
STRUMENTO	70
EVENTO	41
PIANTA	8
CONTENUTO INFORMATIVO	6
ANIMALE	5
LUOGO	4
TEMPO	2
PARTE DEL CORPO	1

**TABELLA 6.** Distribuzione dei composti VN attestati nel CoDIt\_com (sottocorpora 3-4-5) rispetto al tipo semantico di output

I composti VN attestati nel CoDIt\_com sono prevalentemente utilizzati per indicare persone, strumenti ed eventi; più rara, ma comunque significativa, è la presenza di forme che designano animali, contenuti informativi, parti del corpo, piante e periodi di tempo.

Focalizzando l'attenzione sul piano diacronico, è interessante osservare la produttività dei tre tipi maggiormente attestati (persona, strumento, evento) nel corso dei secoli: il grafico nella Figura 5 ne illustra il numero dei type nei sottocorpora 3-4-5.

Come illustrato dal grafico riportato nella Figura 5, 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 composti VN indicanti un referente umano a partire dal periodo successivo al terzo sotto-corpus (dal 1691 in poi); gli altri due tipi semantici aumentano in termini di frequenza dei type in modo più graduale.

I composti VN agentivi consistono prevalentemente in appellativi dispregiativi di persone, identificate sulla base di un'azione che le caratterizza: si tratta di occasionalismi che, nella maggior parte dei casi, compaiono una sola volta (nel sottocorpus 4, ad es., circa il 67% sono hapax). Sono infatti più rare le denominazioni stabili che identificano individui sulla base del mestiere che svolgono, come *falegname* (17 occ.), *beccamorto* (15 occ.) e *guardaboschi* (6 occ.). La tendenza a utilizzare la composizione

si sofferma sul caso interessante del dizionario bilingue francese-italiano di Francesco Alberti di Villanova (del 1834) e, in particolare, del *Dizionario militare italiano* di Giuseppe Grassi (del 1833), in cui si riscontrano esempi appartenenti, ad esempio, all'ambito della chirurgia con questo significato di "contenitore" (ad es., *porta-ago*).

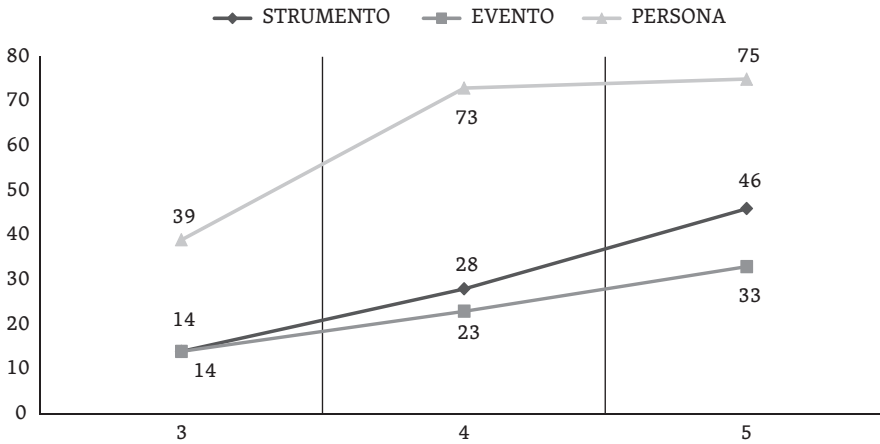


FIGURA 5. Frequenza dei type dei composti VN con tipo semantico persona, evento e strumento

verbo-nominale per appellativi estemporanei e con connotazione negativa, già riscontrata in italiano antico, trova quindi conferma anche nei dati diacronici.

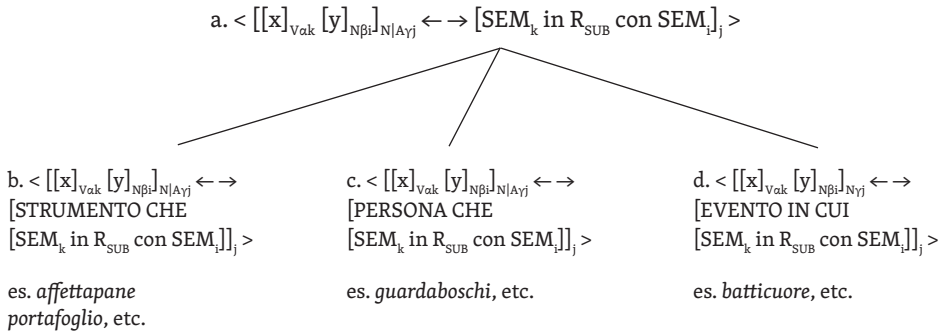
Per quanto riguarda i composti VN con semantica strumentale, questi registrano un incremento nel numero dei type, a partire dal terzo sottocorpus, che si mantiene costante fino al sottocorpus più recente.

Prendendo in esame l'ultimo sottocorpus, si può notare che la composizione agentiva è nettamente prevalente rispetto a quella strumentale: ai due tipi più frequenti si affianca il tipo eventivo, meno sfruttato ma comunque sempre presente nel corso dei secoli. Esempi di questo tipo, tra i più frequenti nel sottocorpus 5, sono i seguenti composti: *passatempo* (37), *battimano* (21), *guazzabuglio* (13), *rompicollo* (17), *crepacuore* (17).

Va infine notato che un discreto gruppo di composti VN occorre soltanto (o prevalentemente) all'interno di locuzioni avverbiali: ad es., l'ultimo sottocorpus registra *ad armacollo* (13), *a squarciagola* (15), *a bruciapelo* (13), *a crepapelletta* (9), *a rompicollo* (5 occ.). La natura di tali composti tuttavia appare diversa: dal momento che ormai occorrono soltanto all'interno della locuzione avverbiale, infatti, essi potrebbero essere considerati formazioni che rappresentano una Costruzione diversa, ossia  $[a + VN]_{AVV}$ , che dovremmo collocare fuori dallo schema illustrato nella Figura 6.<sup>26</sup>

In ottica costruzionista, quanto finora osservato sulla semantica dei composti VN può essere rappresentato attraverso una Gerarchia in schemi e sottoschemi (illustrata nella Figura 6), che rende conto dei composti VN attestati nel CoDIt\_com.

<sup>26</sup> Si noti che il composto *rompicollo* è attestato sia nella locuzione *a rompicollo* 'a rotto di collo', sia come sostantivo con semantica agentiva, con il significato di 'ragazzaccio, un vagabondo, un vero rompicollo... — Calunnie! Tutte calunnie!'. Inoltre, vale anche la pena di aggiungere che *rompicollo* è attestato nel dizionario di John Florio con il significato di luogo, 'a breake-neck place'.



**FIGURA 6.** Gerarchia dei composti VN

La Gerarchia rappresentata dalla Figura 6 è costituita dallo schema più astratto dei composti VN in (a), in cui non sono segnalate restrizioni sul tipo semantico di output: questo schema è esemplificato da tre sotto-schemi in (b), (c), (d), la cui semantica è espressa dalle perifrasi poste a destra delle costruzioni. Questi sotto-schemi sono emersi nell'analisi come più produttivi in termini di frequenza sia dei type sia dei token. Gli altri composti, in cui non si è ravvisata una significativa regolarità in termini formali e semantici, possono essere considerati esemplificazioni dello schema più astratto in (a).

L'analisi ricavata dai dati estratti dal CoDIt\_com ha messo in luce una crescente produttività dei tre principali tipi semantici di output (persona, strumento, evento); accanto alla semantica agentiva e strumentale, infatti, la composizione verbo-nominale è costantemente sfruttata fino alla fase contemporanea anche per la creazione di nomi che indicano eventi (si pensi ad esempio a un 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). Dal punto di vista morfosintattico, l'analisi ha confermato la netta preferenza, riscontrata già nella fase antica, a selezionare basi verbali bisillabiche e appartenenti alla coniugazione in *-are*. Infine, nel periodo dal Cinquecento in poi, i composti VN vengono prevalentemente utilizzati in funzione di sostantivi, ma non mancano casi di usi attributivi.

## 5. CONCLUSIONI

In questo articolo, ci siamo fissati un obiettivo ambizioso, quello di offrire una sintesi del processo di formazione di parole che è la composizione verbo-nominale lungo un asse diacronico molto esteso. Infatti, abbiamo tentato di ripercorrere la storia dei composti VN sin dalle Origini fino alle soglie del Novecento per dimostrare come il processo sia sempre stato *disponibile*, ben pronto, per così dire, a essere sfruttato per diversi *output* semantici (referenti umani che svolgono azioni sia tipiche, professionali sia estemporanee, nomi di strumento, nomi di evento ecc.), così come categoriali. Di fatto, la funzione aggettivale dei composti VN che, come abbiamo visto, rappresenta oggi un uso estremamente diffuso, trova le sue radici in diacronia, in cui si



possono scorgere esempi sicuri già nel periodo precinquecentesco (si pensi all'esempio di *rubacuori*).

Inoltre, la nostra indagine conferma che, pur essendoci un numero cospicuo di basi verbali coinvolte, la crescente produttività dei composti VN sembra essere dovuta a quella che abbiamo chiamato “(verbal) base type frequency”, e cioè l'alta frequenza di un numero ridotto di basi verbali quali *guarda-*, *mangia-*, *batti-*, *salva-*, *para-* e *porta-*. Quest'ultima base risulta essere particolarmente interessante anche per il fatto che mostra come uno schema o sotto-schema, come quello legato alla semantica di “contenitore”, che ha verosimilmente origine dal prestito dal francese *porte-feuille*, possa diventare un modello per ulteriori estensioni. Le singole Costruzioni, come quelle in Figura 6, istanziano un modello astratto, il quale però altro non è che una nostra generalizzazione elaborata a partire da esemplari concreti. Vorremmo, a questo proposito, riportare l'osservazione di Bauer (2017: 74) che, in base all'esempio dei composti inglesi NN, afferma che “[it] seems that (...) it is not the pattern N + N of compounding which is productive, but patterns with individual lexemes within that.” Infatti, anche nel caso dei composti VN italiani, la loro produttività, seppur trattata come una proprietà globale dell'intero processo, parte dai singoli esempi con determinate basi verbali destinate ad avere, come abbiamo visto, una grande fortuna nella storia linguistica dell'italiano.

#### RIFERIMENTI BIBLIOGRAFICI

- 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, consultabile al sito di [https://www.laurenceanthony.net/research/20130722\\_26\\_cl\\_2013/cl\\_2013\\_paper\\_final.pdf](https://www.laurenceanthony.net/research/20130722_26_cl_2013/cl_2013_paper_final.pdf) [ultimo accesso: 18.05.2019].
- Bauer, B. (2011) Word formation. In: Maiden, M., J. C. Smith and A. Ledgeway (eds) *The Cambridge History of the Romance Languages. Volume I. Structures*, 532–563. Cambridge: Cambridge University Press.
- Bauer, L. (2001) *Morphological productivity*. Cambridge: Cambridge University Press.
- Bauer, L. (2017) *Compounds and compounding*. Cambridge: Cambridge University Press.
- Bisetto, A. (1999) Note sui composti VN dell'italiano. In: Benincà, P., L. Vanelli and A. Mioni (eds), *Fonologia e morfologia dell'italiano e dei dialetti d'Italia: atti del 31. Congresso della Società di linguistica italiana, Padova, 25–27 settembre 1997*, 505–538. Roma: Bulzoni.
- Bisetto, A. (2004) Composizione con elementi italiani. In: Grossmann, M. and F. Rainer (eds) *La formazione delle parole in italiano*, 33–51. Tübingen: Niemeyer.
- Bisetto, A. and S. Scalise (2005) The classification of compounds. *Lingue e linguaggio*, 4/2, 319–332.
- Bloomfield, L. (1933) *Language*. New York: Holt.
- Booij, G. (2005) Compounding and derivation: Evidence for construction morphology. In: Dressler, W. U., F. Rainer, D. Kastovsky and O. Pfeiffer (eds) *Morphology and its demarcations*, 109–132. Amsterdam: John Benjamins.
- Booij, G. (2010) *Construction morphology*. Oxford: Oxford University Press.
- Brucale, L. (2012) Latin compounds. *Probus* 24, 93–117.
- Dardano, M. (2009) *Costruire parole. La morfologia derivativa dell'italiano*. Bologna: il Mulino.
- De Mauro, T. (2009) *Grande dizionario dell'uso della lingua italiana*. Torino: UTET.
- Fabb, N. (1998) Compounding. In: Spencer, A. and A. Zwicky (eds) *The Handbook of*

- Morphology*, 66–83. Oxford: Oxford University Press.
- Gather, A. (2001) *Romanische Verb-Nomen Komposita. Wortbildung zwischen Lexikon, Morphologie und Syntax*. Tübingen: Narr.
- Giurescu, A. (1975) *Les mots composés dans les langues romanes*. The Hague-Paris: Mouton.
- Goldberg, A. (1995) *Constructions. A Construction Grammar Approach to Argument Structure*. Chicago: University of Chicago Press.
- Goldberg, A. (2013) Constructionist approaches. In: Hoffmann, T. and G. Trousdale (eds) *The Oxford Handbook of Construction Grammar*, 15–31. Oxford: Oxford University Press.
- Haspelmath, M. and A. D. Sims (2010) *Understanding Morphology. 2nd edition*. London: Hodder Education.
- Ledgeway, A. (2012) *From Latin to Romance. Morphosyntactic Typology and Change*. Oxford: Oxford University Press.
- LIZ 4.0 (2001) *Letteratura italiana Zanichelli*. CD-ROM dei testi della letteratura italiana, a cura di Pasquale Stoppelli ed Eugenio Picchi. Bologna: Zanichelli.
- Magni, E. (2010) From the periphery to the core of Romance [VN] compounds. *Lingue e Linguaggio* IX/1, 3–40.
- Marazzini, C. (2009) *L'ordine delle parole. Storia di vocabolari italiani*. Bologna: il Mulino.
- Marchand, H. (1969) *The categories and types of present-day English word-formation: A synchronic-diachronic approach*. Munich: Beck.
- Masini, F. (2016) *Grammatica delle Costruzioni*. Roma: Carocci.
- Masini, F. and J. Audring (2019) Construction Morphology. In: Audring, J. and F. Masini (eds), *The Oxford Handbook of Morphological Theory*, 365–389. Oxford: Oxford University Press.
- Michele, M. S. (2020) *Composizione italiana in diacronia. Le parole composte dell'italiano nel quadro della Morfologia delle Costruzioni*. Berlin / Boston: De Gruyter.
- Moyna, M. I. (2011) *Compound Words in Spanish. Theory and History*. Amsterdam: John Benjamins.
- Olsen, S. (2001) Copulative compounds: a closer look at the interface between syntax and morphology. In: Booij, G. and J. van Marle (eds) *Yearbook of Morphology 2000*, 279–320. Dordrecht: Springer.
- Pizzoli, L. (2004) *Le grammatiche di italiano per inglesi (1550–1776). Un'analisi linguistica*. Firenze: Accademia della Crusca.
- Radimský, J. (2006) *Les composés italiens actuels*. Paris: Cellule de Recherche en Linguistique.
- Ricca, D. (2005) Al limite tra sintassi e morfologia: i composti aggettivali V-N nell'italiano contemporaneo. In: Grossmann, M. and A. M. Thornton (eds) *La formazione delle parole. Atti del XXVII Congresso Internazionale di Studi della Società di Linguistica Italiana. L'Aquila, 25–27 settembre 2003*, 465–486. Roma: Bulzoni.
- Ricca, D. (2010) Corpus data and theoretical implications. With special reference to Italian V-N compounds. In: Scalise, S. and I. Vogel (eds) *Cross-Disciplinary Issues in Compounding*, 237–254. Amsterdam: John Benjamins.
- Rosenberg, M. (2007) Agent nouns, productivity and diachrony: An analysis of [VN/A]<sub>N/A</sub> compounds and -eur derivations in French. In: Booij, G. et al. (eds), *On-line Proceedings of the Fifth Mediterranean Morphology Meeting (MMM5), Fréjus 15–18 September 2005*, 359–378, consultabile al sito di <http://mmm.linge.unibo.it>.
- Rosenberg, M. (2008) *La formation agentive en français. Les composés [VN/A/Adv/P]N/A et les dérivés V-ant, V-eur et V-oir(e)*. Thèse pour le doctorat, Université de Stockholm.
- Rosenberg, M. (2011) The polysemy of agent nouns: diachronic, synchronic and contrastive evidence from French and Swedish. *STUF — Language Typology and Universals Sprachtypologie und Universalienforschung*, 64, 53–64.
- Scalise, S. (1994) *Morfologia*. Bologna: il Mulino.
- Scalise, S. and A. Bisetto (2008) *La struttura delle parole*. Bologna: il Mulino.
- Scalise, S. and A. Bisetto (2009) The classification of compounds. In: Lieber, R. and P. Štekauer (eds) *The Handbook of Compounding*, 34–53. Oxford: Oxford University Press.



- Scalise, S. and E. Guevara (2006) Exocentric compounding in a typological framework. *Lingue e linguaggio*, 2, 185–206.
- Scalise, S. and E. Guevara (2008) I composti esocentrici in una prospettiva tipologico-comparativa. In: Cresti, E. (ed) *Prospettive nello studio del lessico italiano: atti del 9. congresso SILFI (Firenze, 14–17 giugno 2006)*, 583–590. Firenze: Firenze University Press.
- Spencer, A. (1991) *Morphological theory: An introduction to word structure in generative grammar*. Oxford: Blackwell.
- Štichauer, P. (2015) *La formazione delle parole in diacronia: studi di morfologia derivazionale dell'italiano tra il Cinquecento e l'Ottocento*. Prague: Karolinum Press.
- 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.
- Tesi, R. (2001) *Storia dell'italiano. La formazione della lingua comune dalle origini al Rinascimento*. Roma-Bari: Laterza.
- Tesi, R. (2005) *Storia dell'italiano. La lingua moderna e contemporanea*. Bologna: Zanichelli.
- Tollemache, F. (1945) *Le parole composte nella lingua italiana*. Roma: Roes.
- Il Tommaseo (2004) *Dizionario della lingua italiana di Niccolò Tommaseo e Bernardo Bellini (1857–1879)*. Testo integrale degli otto tomi originali e interrogazione per lemmi e a tutto testo in CD-ROM. Bologna: Zanichelli.

### **M. Silvia Micheli**

Dipartimento di Scienze Umane per la Formazione “Riccardo Massa”  
 Università degli studi di Milano — Bicocca  
 Piazza dell'Ateneo Nuovo 1, 20126 Milano (Italy)  
 ORCID ID: 0000-0002-2085-560X  
 maria.micheli@unimib.it

### **Pavel Štichauer**

Dipartimento di Studi Romanzi  
 Facoltà di Lettere, Università Carlo  
 nám. Jana Palacha 2, 11638 Praha  
 ORCID ID: 0000-0002-6777-1104  
 pavel.stichauer@ff.cuni.cz