

Radimský, Jan

## I composti N+N attributivi in diacronia: le famiglie N-modello, N-base e N-limite

*Études romanes de Brno*. 2024, vol. 45, iss. 3, pp. 10-34

ISSN 2336-4416 (online)

Stable URL (DOI): <https://doi.org/10.5817/ERB2024-3-2>

Stable URL (handle): <https://hdl.handle.net/11222.digilib/digilib.80982>

License: [CC BY-SA 4.0 International](https://creativecommons.org/licenses/by-sa/4.0/)

Access Date: 03. 01. 2025

Version: 20241231

Terms of use: Digital Library of the Faculty of Arts, Masaryk University provides access to digitized documents strictly for personal use, unless otherwise specified.

---

## I composti N+N attributivi in diacronia: le famiglie *N-modello*, *N-base* e *N-limite*

### Attributive N+N Compounds in Diachrony: the *N-modello*, *N-base* and *N-limite* Families

JAN RADIMSKÝ [radimsky@ff.jcu.cz]

Jihočeská univerzita v Českých Budějovicích, Repubblica Ceca

---

#### SOMMARIO

Questo articolo esamina la nascita e lo sviluppo diacronico dei composti attributivi-appositivi Nome+Nome (N+N) in italiano, concentrandosi specificamente su tre famiglie semi-schematiche: *N-modello*, *N-base* e *N-limite*. Utilizzando dati diacronici tratti dal corpus di Google Books, lo studio analizza l'emergere e la successiva evoluzione di produttività di questi composti a partire dalla fine del XIX secolo fino ai giorni nostri. I risultati rivelano le prime innovazioni importanti (le *leader words*) che hanno segnato la nascita dei rispettivi pattern e identificano tendenze significative e punti di svolta importanti nella proliferazione di queste tre famiglie di composti che hanno svolto un ruolo cruciale nella formazione del modello più ampio Attributivo-Appositivo (ATAP) in italiano. Lo studio è inquadrato nelle teorie della *Morfologia delle Costruzioni* e della *Morfologia Relazionale* e offre spunti sull'interazione tra specificità lessicale e schemi di costruzione astratti.

#### PAROLE CHIAVE

composti attributivi-appositivi; ATAP; composti Nome+Nome; composizione; diacronia; italiano; *Morfologia delle costruzioni*; *Morfologia relazionale*

#### ABSTRACT

This article investigates the origin and diachronic development of Attributive-Appositive Noun+Noun (N+N) compounds in Italian, focusing on three semi-schematic families: *N-modello*, *N-base*, and *N-limite*. Drawing on diachronic data from the Google Books corpus, the study examines these compounds' emergence and subsequent productivity evolution from the late 19<sup>th</sup> century to the present day. The findings highlight the key innovations (*leader words*) that marked the formation of these patterns and identify significant trends and turning points in the proliferation of these three families of compounds, which have played a crucial role in shaping the broader Attributive-Appositive (ATAP) pattern in Italian. The study is framed within *Construction Morphology* and *Relational Morphology* theories, offering insights into the interaction between lexical specificity and abstract constructional patterns.

**KEYWORDS**

attributive-appositive compounds; ATAP; noun+noun compounds; compounding; diachrony; italian; Construction Morphology; Relational Morphology

RICEVUTO 2023-10-30; ACCETTATO 2024-02-16

## 1. Introduzione

Il presente articolo si inserisce nel quadro di una ricerca più ampia il cui scopo è quello di rintracciare la nascita e lo sviluppo dei composti Nome + Nome (NN) italiani del tipo *lettore modello*, *trasporto latte* o *banca dati*, facendo riferimento a big data estratti da Google books. I primi risultati riguardanti il tipo *trasporto latte*, conosciuto comunemente come *pattern verbal-nexus subordinativo*, sono stati pubblicati in Radimský (2023a). Questo articolo fa invece seguito allo studio di Radimský (2023b), che presentava i risultati globali dell'analisi del primo campione di composti del tipo *lettore modello*, conosciuto comunemente come *pattern attributivo-appositivo* (ATAP).

È ben noto che i composti ATAP si raggruppano tendenzialmente intorno a famiglie di modificatori N2 lessicalmente specificati come *N-chiave* o *N-modello*. Nella prospettiva di Rainer (2016: 2714), si potrebbe ipotizzare che il ruolo delle singole famiglie semi-specificate sia decisivo per l'eventuale emergere di pattern più astratti. Oppure, sulla base di Bauer (2017: 74), si potrebbe perfino ipotizzare che i veri e propri pattern, con produttività regolare, risiedano più in queste famiglie semi-specificate che non nelle costruzioni più astratte di livello superiore, come sarebbe appunto il pattern NN ATAP. Comunque sia, risulta evidente che uno studio qualitativo delle famiglie semi-specificate (o semi-schematiche, per usare i termini della Morfologia delle Costruzioni) di composti NN italiani sembra indispensabile per capire meglio la dinamica del fenomeno.

In questo articolo ci si focalizza sullo sviluppo diacronico di tre famiglie semi-specificate di composti attributivi, ossia *N-modello*, *N-base* e *N-limite*. Oltre ad appartenere ancor oggi alle più produttive in termini di *type frequency*, queste famiglie sono decisamente quelle con radici più antiche, cioè quelle che –almeno stando ai dati finora noti– presentano la più alta *type frequency* già nei testi ottocenteschi, sebbene nel corso del Novecento siano state raggiunte o superate da alcune famiglie più recenti (come ad es. *N-chiave*, *N-pilota*, *N-fantasma* o *N-cardine*, per citare quelle più importanti per quanto riguarda i valori di *type frequency*).

Il testo è organizzato come segue: le Sezioni 2–4 riprendono, in una maniera più sintetica e allo stesso tempo aggiornata, le basi teoriche ed empiriche esposte già in Radimský (2023b); più in particolare, si fornirà la definizione dei composti ATAP e una breve panoramica degli studi precedenti sulla loro evoluzione (Sezione 2), i dati che stanno alla base della presente ricerca (Sezione 3) e il profilo diacronico del pattern ATAP in base al campione più recente (Sezione 4). Lo studio dettagliato delle tre famiglie semi-schematiche *N-modello*, *N-base* e *N-limite*, insieme alle relative premesse metodologiche, è oggetto della Sezione 5.

Dal punto di vista teorico, l'articolo e la ricerca sottostante s'inseriscono nel quadro teorico della *Morfologia delle Costruzioni* (CxM; cf. Booij 2010, 2016) e della *Relational Morphology* (Jackendoff e Audring 2020), prendendo ispirazione dalla *Diachronic Construction Grammar* (Hilpert 2021)

## 2. Delimitazione e evoluzione dei composti ATAP

Il gruppo dei composti ATAP, concepito nell'influente classificazione dei composti di Bisetto e Scalise (2005; Scalise e Bisetto 2009) come uno dei tre principali modelli di composizione, insieme ai composti coordinativi e subordinativi, ha dato luogo a stimolanti discussioni che sottolineano come questa classe di composti sia particolarmente difficile da definire su basi formali (cfr. Baroni, Guevara e Pirrelli 2009; Grandi 2009; Grandi, Nissim e Tamburini 2011; Arcodia, Grandi e Montermini 2009). In breve, Scalise e Bisetto (2009) sostengono che i composti ATAP presentano una relazione attributiva tra la testa e il suo modificatore –quest'ultimo esprimerebbe una proprietà o qualità della testa–, ma non forniscono criteri chiari su come verificare l'esistenza di questo tipo di relazione nel caso delle strutture NN, oltre al fatto che il modificatore “è spesso da interpretare metaforicamente” (Scalise e Bisetto 2009: 51–52).

Senza entrare nei dettagli di questo dibattito (cfr. Radimský 2015: 92–112 per una panoramica dettagliata), delimitiamo qui il gruppo dei NN ATAP italiani su basi formali utilizzando il test trasformatore introdotto da Radimský (2015: 92–102; cfr. anche Arnaud 2003: 12–13 e Fradin 2009 per il francese). Il test è basato sull'assunto che la relazione attributiva sia tipicamente quella esistente tra un aggettivo qualificativo e il rispettivo nome-testa e che tale relazione possa essere sistematicamente parafrasata da un'espressione predicativa tramite un verbo copulare. Poiché anche i sostantivi possono figurare come complementi predicativi del soggetto retti dal verbo copulare, lo stesso test può essere utilizzato per individuare una relazione attributiva nelle strutture NN. Pertanto, i composti NN attributivi (AT), in cui compare un modificatore astratto, possono essere trasformati in un'espressione predicativa direttamente (1), mentre i composti NN appositivi (AP), in cui compare un modificatore concreto usato metaforicamente, possono essere trasformati in un'espressione predicativa almeno indirettamente, usando l'operatore *come* (2):

(1) *luogo simbolo* → *questo luogo è un simbolo*

(2) *pesce palla* → *questo pesce è come una palla*

Non riteniamo opportuno adoperare una distinzione netta tra i composti attributivi (1) e quelli appositivi (2), visto che nella pratica tale distinzione può sembrare a volte artificiale, per non dire superflua. In effetti, sarebbe una questione puramente accademica chiedersi se gli NN con i modificatori come *chiave* o *lampo* (ad es. *parola chiave*, *treno lampo*) siano piuttosto appositivi (dato che l'interpretazione del modificatore è metaforica rispetto al significato concreto primario di quel sostantivo) o attributivi (dato che l'interpretazione del modificatore si basa su uno dei suoi significati che già di per sé è metaforico). Tratteremo perciò di seguito tutti i composti ATAP come un insieme omogeneo.

Infine, è opportuno ricordare che, secondo alcuni autori, un sottogruppo di modificatori che figurano nei composti ATAP (come appunto *chiave* o *lampo* citati sopra) dovrebbe essere analizzato piuttosto come costituito da aggettivi invariabili (Grandi, Nissim e Tamburini 2011; Grandi 2009; ma si veda Thornton 2004: 528–530 e Radimský 2015: 22–29). Poiché questi elementi sono principalmente sostantivi, con caratteristiche sintattiche (ma non morfologiche) tipiche degli aggettivi qualificativi, il dibattito sembra inconcludente se condotto su basi formali. Riteniamo quindi che la scelta di classificare come nomi questi modificatori in posizione postnominale e le rispettive strutture complesse come composti NN ATAP sia una delle due opzioni metodologiche possibili, il cui particolare vantaggio è che permette di descrivere gli ATAP NN come parte integrante dei NN italiani, senza ulteriori suddivisioni.

Come sottolinea Rainer (2021), le nostre conoscenze attuali sull'evoluzione e sulla proliferazione moderna dei composti romanzi di tipo NN sono ancora estremamente limitate. L'autore sostiene che il modello non mostra alcuna continuità rispetto ai composti latini, ma deriva piuttosto da una varietà di costruzioni sintattiche eterogenee il cui numero sembra estremamente basso in italiano, almeno fino alla fine dell'Ottocento.

Più concretamente, in un ampio studio diacronico dei composti italiani basato sul corpus Codit dell'italiano letterario, Micheli (2020: 91–93, 145, 152–155) ha rintracciato solo NN 3 ATAP nell'italiano antico (*pescespada*, *pesceporco*, *arcamensa*) e 15 NN ATAP nell'italiano medio (*pescecane*, *pescecalvo*, *pescesega*, *pesceluna*, *formicaleone*, *grillotalpa*, *melarancia*, *cavolfiore*, *cavolrapa*, *casciofiore*, *cartastraccio*, *erbaspada*, *ceralacca*, *gommalacca*). Si noti che un terzo di questo piccolo campione è rappresentato dalla famiglia *pesce+N* che costituisce una vera e propria costruzione isolata, influenzata dalle denominazioni di pesci esistenti anche in altre lingue europee (cfr. ad es. per *pesce spada*: *xiphias gladius* in latino, *pez espada* in spagnolo, *peix espasa* in catalano, *swordfish* in inglese, *Schwertfisch* in tedesco, etc.). In italiano, non c'è una norma stabile per la grafia dei composti NN e lo studio di Micheli (2020) tiene conto solo di quelli scritti attaccati o separati da un trattino. Se venissero estratti anche i composti NN nella forma giustapposta (cioè scritti come due parole separate), nel corpus Codit si potrebbe trovare qualche altro esempio per completare il campione (ad esempio, *edizione principe* – editio princeps). Tuttavia, questo non toglie che la presenza di strutture NN ATAP nella lingua letteraria prima della metà del XX secolo sia stata estremamente bassa.

Se in questo studio ci concentriamo in particolare su 3 famiglie ATAP il cui sviluppo si può ben documentare già dalla metà dell'Ottocento, cioè quasi un secolo prima, non è solo perché queste particolari famiglie rappresentano una sorta di eccezione che conferma la regola, essendo una specie di avanguardia del futuro modello ATAP. Ciò è anche dovuto al fatto che i dati usati per l'estrazione degli Google n-grams (si veda oltre) non si limitano solo al linguaggio letterario, ma includono anche altre varietà stilistiche importanti come i linguaggi specialistici, quello giornalistico e amministrativo. Secondo Terreni (2005: 534), i composti NN ATAP<sup>1</sup> si riscontrano nel corso del Novecento in vari registri stilistici quali la saggistica accademica (critica letteraria, saggi su temi sociologici e politici), nel linguaggio della pubblicità o quello dei cartelli e servono anche a formare denominazioni di oggetti concreti (arredamento, prodotti destinati alla cura del corpo, etc.), mentre l'habitat naturale dei composti NN subordinativi sarebbe

1 In Terreni (2005), i composti NN ATAP vengono chiamati “il tipo *notizia-curiosità*” ossia “tipo B”, mentre i composti NN subordinativi sono etichettati come “il tipo *lista-nozze*” o “tipo A”.

piuttosto il linguaggio burocratico-amministrativo. Come vedremo più avanti esaminando gli esempi concreti, è soprattutto dal linguaggio specialistico che traggono origine molti importanti composti NN ATAP, che costituiscono le radici dell'evoluzione delle rispettive tre famiglie studiate in questo articolo.

### 3. Dati

Dal punto di vista empirico, la presente ricerca si basa su big data diacronici estratti dal corpus Google books sotto forma di liste di frequenza grezze, che sono state pubblicate da Google come terza versione dei Google n-grams italiani<sup>2</sup>. La dimensione del corpus Google books che sta alla base è di 120.410.089.963 *tokens*. I dati per l'estrazione dei composti NN provengono da bigrammi e trigrammi pretrattati utilizzando la procedura descritta in dettaglio da Radimský (2022). L'intero dataset comprende, in forma grezza, 19.319.372 bigrammi (*types*), di cui 2.723.566 sono effettivamente potenziali candidati di tipo NN, poiché potenzialmente corrispondono alle forme dei sostantivi italiani secondo il dizionario MorphIt (Zanchetta e Baroni 2005). Dato che questa lista contiene un tasso estremamente elevato di falsi positivi, è stato necessario effettuare un ulteriore filtraggio manuale degli NN ATAP.

I composti ATAP hanno la caratteristica quella di raggrupparsi intorno a famiglie di modificatori N2 lessicalmente specificati: alcuni autori ipotizzano addirittura che i modificatori nominali nei composti ATAP costituiscano un insieme relativamente limitato di sostantivi (Baroni, Guevara e Pirrelli 2009). Per questo motivo, l'estrazione manuale è consistita nell'identificazione di potenziali modificatori nella posizione N2, seguita da un controllo manuale sistematico di tutte le teste corrispondenti nella posizione N1. Tra i potenziali modificatori N2 –identificati da Radimský (2015) nel campione dei 1.800 composti NN ATAP contemporanei, ma pure estratti dal dizionario Zingarelli (2022)<sup>3</sup>– ci siamo concentrati per ora soprattutto (anche se non esclusivamente) sugli N2 più significativi, cioè quelli con le *type frequencies* più elevate. Fino a questo punto è stato possibile controllare e filtrare famiglie di nomi-testa (N1) che corrispondono a 56 diversi N2. Poiché l'analisi dei dati sta rivelando una potenziale presenza di famiglie basate su un nome-testa (N1) lessicalmente specificato tra i composti NN ATAP –oltre alla famiglia *pesce-N* discussa sopra– nel prossimo futuro si prevede un controllo più sistematico di queste famiglie, cioè un'estrazione più sistematica basata sui nomi-testa lessicalmente specificati.

Per ottenere la massima accuratezza nell'estrazione dei dati, la maggior parte dei composti è stata ricontrollata manualmente su un campione di occorrenze in Google Books e molti falsi positivi sono stati eliminati. Per lo stesso motivo si sono scelte come unità di base parole piuttosto che lemmi, perché risulta che, in virtù dell'ambiguità morfologica, le forme flesse sono una fonte frequente di falsi positivi nei testi reali. Poiché la maggior parte dei modificatori negli NN ATAP italiani sono invariabili (Radimský 2015: 49–50; Radimský 2018), la distinzione parola/lemma ha una rilevanza relativamente limitata nel caso degli N2.

2 <https://storage.googleapis.com/books/ngrams/books/datasetsv3.html>

3 La lista dei modificatori N2 di Radimský (2015) è stata completata da nomi estratti dalla versione elettronica dello Zingarelli (2022) che hanno come parte del discorso alternativa: “in funzione di agg. inv.”, oppure per cui il dizionario indica la nota d'uso “postposto a un sost.” o “postposto al sost.”.

L'attuale campione comprende 2.541 NN ATAP (*types* non lemmatizzati) che includono 56 N2 (nomi modificatori) diversi e 1.389 N1 (nomi testa) diversi.

Per visualizzare le tendenze diacroniche nei dati tramite linee di regressione si utilizzerà lo stimatore di Theil-Sen, integrato dal test di Mann-Kendall per la verifica della significatività delle tendenze stesse (implementazione Python di Hussain e Mahmud 2019). Questi metodi non parametrici basati sui ranghi sono adatti a testare qualsiasi forma di dipendenza (non solo quella lineare), non assumono una distribuzione normale degli errori e non sono sensibili agli outlier, il che li rende particolarmente adatti all'identificazione delle tendenze nell'uso delle parole nei corpora diacronici (cfr. Kovář e Herman 2013).

Purtroppo qualsiasi analisi basata sui dati tratti da Google n-grams soffrirà necessariamente di alcuni problemi che non possono essere evitati: i corpora Google books sottostanti non sono bilanciati per genere, l'accesso completo ai testi originali non è sempre possibile, la riedizione di fonti più vecchie può perturbare l'accuratezza della datazione e i dati negli n-grams contengono una parte di falsi positivi che è impossibile eliminare, per citare almeno gli svantaggi più importanti. Nonostante questo, i dati tratti da Google n-grams rappresentano la fonte più ricca e accurata per l'analisi diacronica della lingua scritta degli ultimi due secoli attualmente disponibile.

#### 4. Profilo diacronico del pattern ATAP

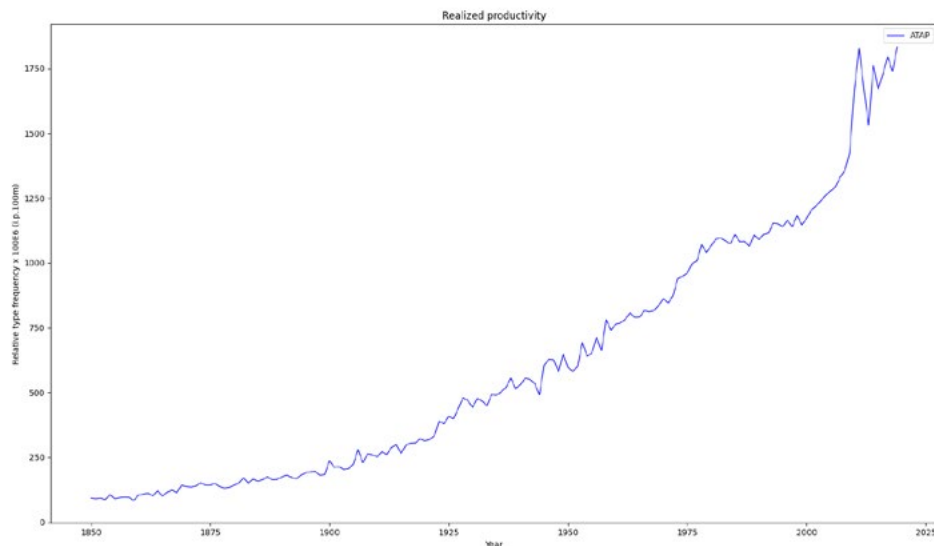
Nei termini della Morfologia delle Costruzioni (CxM), l'evoluzione diacronica dei composti ATAP può essere analizzata a vari livelli di astrazione, ossia su quello di Costruzioni sostanziali (composti o *types* individuali), quello di Costruzioni semischematiche (famiglie di composti NN basate sullo stesso N1 o N2) e quello dell'intera Costruzione ATAP schematica. In questa sezione presenteremo brevemente una panoramica globale dell'intero campione dei 2.541 NN ATAP, che corrisponde alla costruzione schematica (3), per passare poi, nel paragrafo successivo, a una descrizione più dettagliata di tre importanti famiglie semischematiche (*N-modello*, *N-base*, *N-limite*), soffermandoci anche, laddove opportuno, sui composti individuali più significativi per la storia di ogni famiglia.

##### (3) ATAP NN construction

$$[N_i N_j] N_k \leftrightarrow [N_{i\text{-head}} \text{È (COME) UN(A)} N_{j\text{-non-head}}]_k$$

La panoramica globale è visualizzata nei grafici 1 e 2, che presentano rispettivamente la *type frequency (realized productivity)* e la *family type frequency* dell'intera costruzione ATAP a partire dagli anni '50 dell'Ottocento fino al 2019, secondo i principi metodologici presentati già in Radimský (2023b), ma prendendo in considerazione dati aggiornati e più ampi.

Dal punto di vista tecnico, il Grafico 1 presenta in diacronia il numero di tipi (*type frequency*) della Costruzione ATAP, cioè il numero di composti (*types*) per ogni anno. Poiché la dimensione del corpus sottostante varia in diacronia, la *type frequency* è espressa in termini relativi come il numero di tipi diviso per la dimensione del corpus nel rispettivo anno, il che corrisponde alla *realized productivity* di Baayen (2009). Il risultato è stato moltiplicato per la costante  $10^8$  in modo da avvicinarsi intuitivamente all'ordine di grandezza dei dati originali.



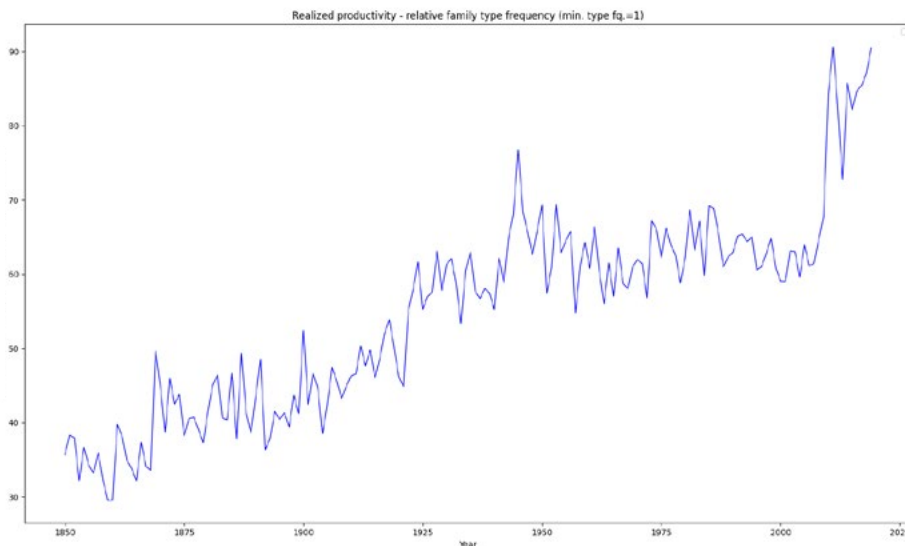
**Grafico 1.** *Realized productivity* dei NN ATAP NNs in diacronia  
( $V/N \times 10^8$ , dove  $V$  è il numero di *types* e  $N$  è la dimensione del corpus nel rispettivo anno)

La *family type frequency*, visualizzata nel Grafico 2, fornisce una prospettiva diversa sullo stesso fenomeno. Questa misura è ispirata dal fatto che la *realized productivity* (ossia la *type frequency relativa*) globale cela il ruolo delle singole famiglie semispecificate, cioè non informa sul fatto se l'aumento della *type frequency* globale sia dovuto piuttosto a un numero crescente di varie costruzioni di livello inferiore oppure all'aumento della *type frequency* di una o di poche costruzioni semi-specificate. Visto che i composti ATAP si raggruppano soprattutto intorno a famiglie basate sullo stesso N2, un'immagine più accurata della *realized productivity* può essere ottenuta chiedendo quante diverse famiglie N2 sono presenti in un determinato momento. Infatti, solo un numero crescente di famiglie N2 comporterebbe una copertura (*coverage* in termini nella CxM) più uniforme della costruzione ATAP (3), così da rafforzare la sua rappresentazione mentale. In termini tecnici, la *family type frequency* nel Grafico 2 dà dunque il numero relativo di famiglie N2 presenti in un determinato anno. In altre parole, essa esprime pure la *realized productivity* di Baayen (2009), ma tutti i composti (*types*) con lo stesso N2 sono contati come uno solo.

La curva del Grafico 1 mostra un aumento di tipo esponenziale della *realized productivity* a partire dall'inizio del Novecento in poi. Due momenti particolari in cui la curva comincia a crescere più rapidamente si possono identificare a occhio nudo: il primo è situato all'inizio del dopoguerra (dunque negli anni '50) e il secondo intorno al 2000. Questi due punti di svolta permettono dunque di identificare tre tappe diverse nell'evoluzione della *realized productivity*.

Il Grafico 2 segna le stesse tre fasi nell'evoluzione degli NN ATAP ma, a differenza del grafico precedente, esse non corrispondono a un'evoluzione sempre più rapida del pattern. Nel primo periodo, che va dagli anni '50 dell'Ottocento fino agli anni '50 del Novecento, la *realized productivity* lentamente crescente di NN ATAP, osservata nel Grafico 1, è dovuta a un numero





**Grafico 2.** *Family type frequency* di famiglie basate sui N2

crescente di famiglie N2 indicato nel Grafico 3. In altri termini, questo è il periodo in cui emergono per la prima volta quasi tutte le famiglie N2 studiate. Nel secondo periodo (1950–2000), il rapido aumento della *realized productivity* degli NN ATAP (Grafico 1) è imputabile a un numero relativamente limitato di famiglie, poiché la *family type frequency* nel Grafico 2 non mostra alcuna tendenza ( $p=0,482$  secondo il test di Mann-Kendall) e il numero globale di famiglie N2 è persino inferiore a quello degli anni Quaranta. Pertanto, il modello ATAP in quanto tale non si espande, lo fanno solo alcune famiglie N2. La svolta successiva si verifica dopo il 2000, quando la *Family type frequency* nel Grafico 2 ricomincia a crescere.

## 5. Profilo diacronico delle famiglie N2 scelte

### 5.1 Premesse metodologiche

Una visione più dettagliata e articolata dello sviluppo dei composti ATAP può essere fornita dall'analisi diacronica delle singole famiglie semi-schematiche, cioè quelle basate su un modificatore N2 lessicalmente specificato. Già dai primi risultati si evince che ogni famiglia ha una sua storia particolare che non è necessariamente una copia fedele dell'evoluzione generale del pattern e che merita di essere raccontata.

Dal punto di vista metodologico, l'analisi di ogni famiglia partirà dal confronto della sua descrizione semantica in due grandi dizionari di lingua, GDLI (Grande dizionario della lingua italiana) di Salvatore Battaglia (1961–2002) e Zingarelli (2022). Il primo ha il merito di essere la fonte lessicografica più dettagliata sull'italiano novecentesco, mentre il secondo, particolarmente

affidabile, abbastanza dettagliato e soprattutto costantemente aggiornato, rappresenta la fonte lessicografica più recente di dati affidabili che si possa ottenere.

Di seguito, un'analisi quantitativa della *realized productivity* della famiglia in questione permetterà di tracciarne le origini e di individuare i punti di svolta importanti che segnano i confini dei rispettivi *clusters* diacronici. I punti di svolta importanti possono essere stimati, con un certo margine di precisione, leggendo la curva della *realized productivity* a occhio nudo e la loro validità si conferma (o meno) sulla base del fatto che il Mann-Kendall test indica un cambiamento di tendenza per il periodo successivo rispetto a quello precedente<sup>4</sup>.

Da ultimo, si realizza un'analisi qualitativa basata sull'identificazione delle *leader words*, cioè dei composti con la più alta *token frequency* relativa in varie *tranches* diacroniche. Le *tranches* sono state scelte in maniera uniforme per tutte le famiglie prendendo in conto sia la qualità e il carattere dei dati sottostanti, sia alcune tappe storiche importanti per l'evoluzione della società. In particolare, si tratterà dei seguenti sei periodi: 1850–1930, 1930–1945, 1945–1960, 1960–1980, 1980–2000 e 2000–2019. Il primo periodo (1850–1930) copre una fase particolarmente lunga, ma non va ulteriormente suddiviso, perché i composti ATAP sono ancora rari, mentre i falsi positivi risultano più frequenti rispetto a periodi successivi. Segue poi il periodo fascista con la Seconda guerra mondiale (1930–1945) e il lungo Dopoguerra, diviso in quattro fasi di lunghezza simile (1945–1960, 1960–1980, 1980–2000, 2000–2019). Tutto sommato, non sembra importante che – a parte l'anno chiave 1945, che rappresenta da sé un vero e proprio punto di svolta – i confini delle fasi siano “arrotondati” in maniera un po' arbitraria: dato che le frequenze delle *leader words* sono calcolate in maniera cumulativa per i rispettivi periodi, uno spostamento del confine di 2–3 anni non produce in generale nessun cambiamento sui primi posti della classifica. L'analisi qualitativa delle *leader words* in vari periodi può gettare nuova luce sull'evoluzione interna del pattern e sulla sua dinamica, compresa quella semantica.

## 5.2 La famiglia *N-modello*

Il GDLI indica tre diverse situazioni in cui il nome *modello* può servire da modificatore (1–3). Le strutture (4) e (5) sono più complesse e corrispondono a sottoschemi specifici che non rientrano pienamente nel pattern NN ATAP, perciò non sono state prese in considerazione nell'estrazione dei dati.

- (4) N + *modello* + N

*Vuoi una borsa da mare? le scarpe anni quaranta? un abito nuovo modello zingara?* (Arbasino, GDLI)

- (5) N + *ultimo modello*

*Per farsi prendere in considerazione da lei, una bella Maserati ultimo modello contava molto di più che aver costruito il Partenone.* (Buzzati, GDLI)

4 Il tentativo di applicare a questo fine il *variability-based neighbour clustering* proposto da Gries e Hilpert (2008) non ha finora prodotto risultati che siano intuitivamente sensati per quanto riguarda la forma delle curve (si è usata l'implementazione per R disponibile qui: <https://rdrr.io/github/browndw/vnc/>).

La struttura (6) con i vari sottotipi semantici corrisponde invece a un esempio tipico del pattern NN ATAP. Le varie sfumature semantiche (6a-6g) sono così vicine l'una all'altra da poter essere raggruppate sotto una definizione sintetica, come quella fornita dallo Zingarelli (7). I dati degli Google n-grams indicano che, nonostante questa vicinanza, le sfumature (6e-g) che mettono in rilievo l'innovazione tecnologica e/o organizzativa nell'ambito dell'industria e dell'agricoltura, sono state particolarmente rilevanti per il periodo dell'Ottocento e del primo Novecento, come si vedrà più avanti.

(6) N + *modello*. Con valore attributivo invar.

(a) Che, per la perfezione morale, per le virtù, per le capacità intellettuali e professionali, per le doti, per il comportamento, è indicato o proposto come esempio o è fatto oggetto di grande ammirazione [...]

*“Amico mio, ti predico che sarai un padre di famiglia modello; hai pensato a tutto!”*, disse l'ungherese ridendo. (Verga, GDLI)

(b) Che rivela compiutamente le caratteristiche e le peculiarità culturali, ambientali, etniche di un paese o di un popolo.

*Egli parla correttamente molte lingue, ha le abitudini e i costumi di tutti i paesi che ha visitato; in Italia è italiano, in Inghilterra è inglese, e in America è americano modello.* (Tarchetti, GDLI)

(c) Che soddisfa pienamente determinate esigenze espressive e tecniche ed è ritenuto paradigmatico.

*Che cosa rispondere ai precettisti e regolisti dei generi letterari della lingua-modello e simili?* (Croce, GDLI)

(d) Tipico, emblematico di una mentalità, di una tendenza, di un ambiente.

*Ford scrittore può ricordare Franklin. È un moralista impegnato a scrivere una vita modello, una vita rappresentativa.* (Gobetti, GDLI)

(e) Realizzato, costruito o condotto in maniera razionale e progredita, tale da soddisfare pienamente le esigenze previste. - Anche: che attua la sperimentazione di nuove tecnologie (un impianto, una struttura produttiva).

*Guglielmo visitò successivamente la stalla-modello, il fienile-modello, la concimaia-modello, la colombaia-modello.* (Cassola, GDLI)

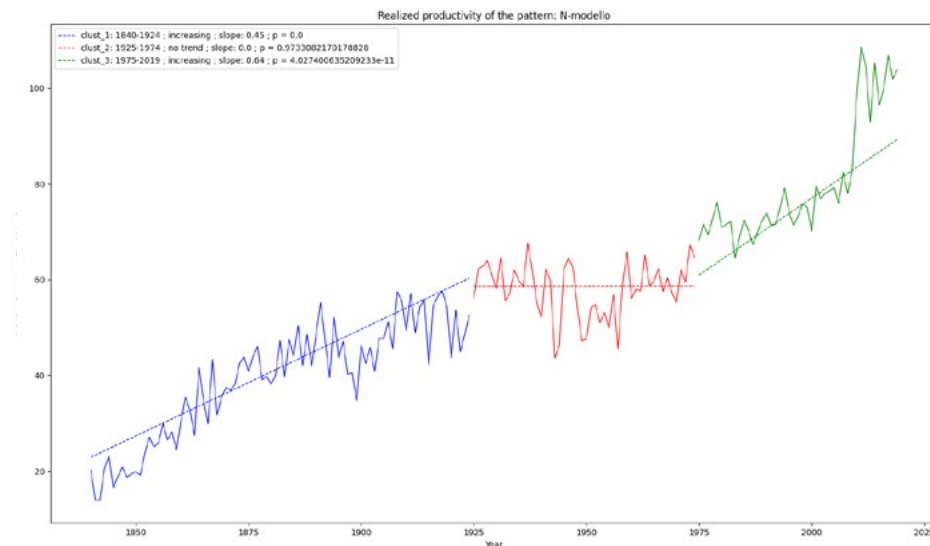
(f) Che segue criteri molto avanzati; assai efficienti (un'istituzione) bene organizzato e amministrato (un paese).

*A Firenze abbondano le circostanze i mezzi le suppellettili per crescervi una facoltà filologica modello.* (Carducci, GDLI)

(g) Altamente selezionato; ben allevato (il bestiame).

*Nelle praterie va pascolando un bestiame modello e ruzzano bambini eleganti.* (Montano, GDLI) (GDLI, voce *modello*; è stato riprodotto solo un esempio per ogni sottotipo)

(7) In funzione di agg. inv. (postposto a un sost.) detto di chi (o di ciò che) è perfetto nel suo genere e quindi degno di essere imitato: *soldato, impiegato, studente modello; sposa modello; scuola modello; podere, orto, frutteto modello.* (Zingarelli, 2022, voce *modello*)



**Grafico 3.** *Realized productivity (type frequency)* della famiglia *N-modello*

Dal punto di vista diacronico, la famiglia *N-modello* appartiene a quelle più antiche, essendo ben presente in italiano già alla fine dell'Ottocento, come si evince inoltre dalla lista degli autori citati nel GDLI (Tarchetti, Carducci, Verga, etc.), parzialmente riprodotta negli esempi (6a-g). Allo stesso tempo è anche una delle più produttive in termini di *type frequency*. Finora sono stati individuati 173 NN attributivi che rientrano regolarmente nello schema ATAP (6–7); il filtraggio di ulteriori dati viene purtroppo perturbato dalla presenza di altri pattern, sia sintattici (apposizioni) sia morfologici (4–5), che risultano spesso difficili da distinguere nelle liste di frequenza. Il Grafico 3 riproduce la panoramica diacronica della *realized productivity* della famiglia *N-modello* a partire dai dati finora filtrati.

La curva nel Grafico 3 segue un'evoluzione in almeno tre tappe con due punti di svolta situati approssimativamente nella metà degli anni '20 e nella metà degli anni '70 del Novecento. Il test di Mann-Kendall conferma questa periodizzazione indicando una tendenza crescente per il primo e il terzo periodo, mentre nessun trend è rilevato per il periodo intermedio (1925–1975). Il forte aumento nella parte destra del grafico corrisponde all'anno 2010.

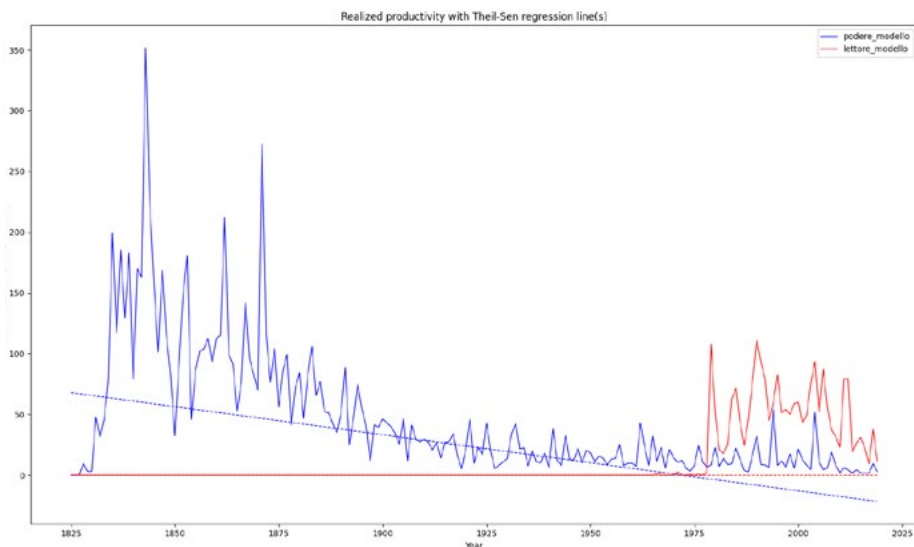
Gli esempi di N1 più significativi in termini di *token frequency* (le *leader words*) della famiglia *N-modello* di diverse fasi diacroniche sono riprodotte nella Tabella 1.

Periodo LWs	<1930	1930–1945	1945–1960	1960–1980	1980–2000	2000+
	podere	azienda	aziende	alberi	lettore	lettore
	poteri	podere	azienda	aziende	città	agente
	scuola	scuola	podere	podere	azienda	studente
	statuto	aziende	fattorie	azienda	testo	legge
	stabilimento	alberi	città	città	autore	studentessa
	istituto					
	scuole					
	paese					
	città					
	famiglia					
	asili					
	marito					
	ospedale					
	fattoria					

**Tabella 1.** *Leader words* della famiglia *N-modello*  
(dati non lemmatizzati, ordinati per *rank* decrescente)

Come si evince da questi esempi, praticamente tutte le sfumature di significato (6a–g) sono presenti già nella prima fase. Risalendo indietro nel tempo, la prima *leader word* significativa è senz'altro il NN *podere modello*, comune già negli anni '40 dell'Ottocento e usato in documenti di vario tipo, incluso ad es. il *Dizionario analitico di diritto e di economia industriale e commerciale* di Angelo Melano di Portula (1843)<sup>5</sup>. Si noti che le versioni francesi delle tre prime *leader words*, cioè *ferme modèle*, *écoles modèle* ed *établissement modèle*, compaiono nello stesso periodo (gli anni '40-'80 dell'Ottocento) e nella stessa tipologia di testi secondo il Google n-gram viewer. La rispettiva sfumatura di significato (6e) assieme a quelle più vicine (6f–g) occupano il primo *rank* degli esempi fino alla fine degli anni '70 del Novecento, cioè fino alla fine del secondo periodo visualizzato nel Grafico 3. Un'altra *leader word* significativa – *lettore modello* – appare nel 1979 con la pubblicazione di *Lector in fabula* di Umberto Eco. A partire da questo momento, *lettore modello* occuperà il primo rank della famiglia *N-modello* e, vent'anni dopo, altri NN con la stessa sfumatura di significato (6d), il cui componente N1 è [+um], appaiono tra le combinazioni più frequenti (*agente*, *studente*, *autore*). La concorrenza diacronica tra le due *leader words* è visualizzata nel Grafico 4.

5 Link diretto GB.



**Grafico 4.** *Relative token frequency delle due leading words più significative podere modello e lettore modello*

Il confronto dei Grafici 3 e 4 permette di avanzare l'ipotesi che il termine *lettore modello*, coniato da Umberto Eco nel 1979, potrebbe aver causato un risveglio della famiglia *N-modello*, tradottosi nell'aumento della sua *realized productivity*, che vede coinvolgere più spesso nuovi N1 riferiti a persone e sfumature di significato del tipo (3a-d).

### 5.3 La famiglia *N-base*

Secondo i dizionari, i significati del modificatore *base* sono meno variegati rispetto al modificatore *modello* trattato nella sezione precedente. Il GDLI fornisce come sinonimi due aggettivi qualificativi, ossia *fondamentale* e *inalienabile* (8); lo Zingarelli, oltre a *fondamentale*, aggiunge *essenziale* e *principale* (9).

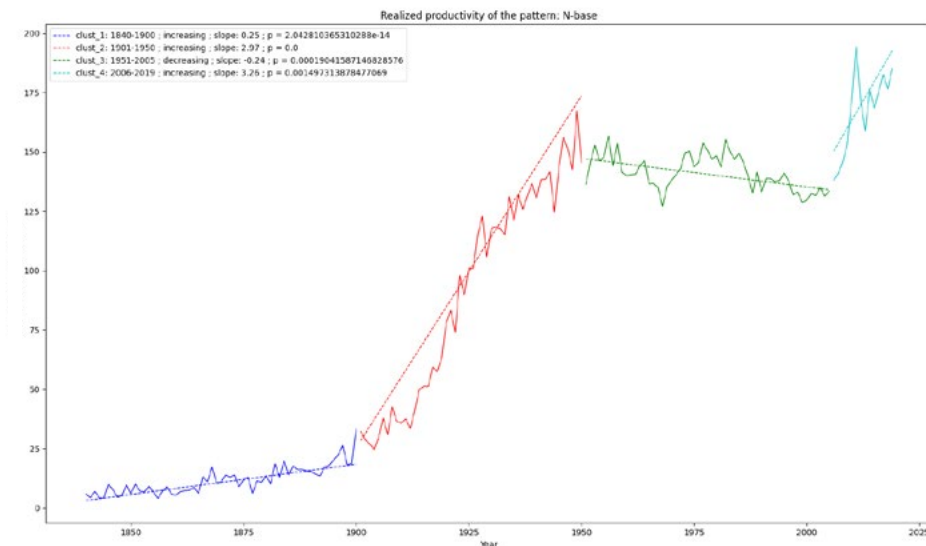
- (8) Con valore di agg. (invar. anche al plur.) : fondamentale, inalienabile (stipendio o stipendi base; indice base; problema base, costo base, ecc.).

*Le sue serate al teatro e al caffè... costituivano l'argomento base.* (Palazzeschi, GDLI)

*Vidi come le piacentine facevano la loro polenta con i fagioli... Così la razione base viene corretta secondo i desideri delle squadre, mediante accomodamenti particolari con il padrone.*

(Gadda, GDLI) (GDLI, voce *base*; solo alcuni esempi sono stati riprodotti)

- (9) In funzione di agg. inv. (posposto a un sost.) fondamentale, essenziale, principale: salario base; problema base; testo base; campo base. (Zingarelli, 2022, voce *base*)



**Grafico 5.** *Realized productivity (type frequency) della famiglia N-base*

Anche la famiglia *N-base* affonda le sue radici già nell'Ottocento e, con l'andar del tempo, ha acquisito un'alta *type frequency* che si è affermata nell'uso: finora sono stati individuati 252 NN attributivi nei Google n-grams. Il Grafico 5 riproduce la panoramica diacronica della *realized productivity* della famiglia *N-base* a partire dai dati finora filtrati.

A occhio nudo si possono distinguere quattro tappe importanti nell'evoluzione diacronica della *realized productivity*. I primi due periodi (1825–1900 e 1901–1950) mostrano una tendenza crescente secondo il test di Mann-Kendall, ma nel secondo periodo la pendenza della linea di regressione di Theil-Senn è molto più forte. Il terzo periodo (1951–2005) è segnato invece da una tendenza decrescente, seguito da un risveglio drammatico nell'ultimo periodo dopo il 2006. Nell'interpretazione di questa curva va sottolineato che, se la *type frequency* fosse calcolata in maniera cumulativa per i rispettivi periodi, la maggior parte dei types (più di 200, cioè i 4/5 del campione) è probabilmente apparsa per la prima volta, almeno come *hapax legomena*, già nel primo periodo, cioè nell'Ottocento. Risulta però difficile verificare questi conti con certezza, perché il tasso di falsi positivi nei dati è di norma assai più alto per i testi più antichi.

Gli esempi di N1 più significativi in termini di *token frequency*, ossia le *leader words* della famiglia *N-base* di diverse fasi diacroniche sono riprodotte nella Tabella 2.

Periodo	< 1930	1930–1945	1945–1960	1960–1980	1980–2000	2000+
LWs	punti	prezzo	punti	paga	prezzo	campo
	punto	punti	prezzo	prezzo	testo	prezzo
	prezzo	paga	punto	testo	importo	testo
	curva	metallo	paga	contributi	paga	principi
	lunghezza	stipendio	pena	prezzi	anno	elementi
	forma					
	gruppo					
	linea					
	salario					
	prezzi					
	campo					

**Tabella 2.** *Leader words* della famiglia *N-base* (dati non lemmatizzati, ordinati per *rank* decrescente)

Le *leader words* della prima fase sono *punti* o *punto base* che, assieme ad alcuni altri (*curva*, *lunghezza*, *forma*, *gruppo*, *linea*), appartengono al vocabolario della matematica, fisica e/o geometria. Si trovano sulla lista, inoltre, due termini dell'ambito dell'economia (*prezzo/prezzi* e *salario*). La presenza di questi termini nei testi scientifici può essere rintracciata e verificata nei Google books, il che permette di avanzare l'ipotesi che la famiglia *N-base* trae effettivamente le sue origini da questi ambiti specializzati nel periodo ottocentesco.

Nelle fasi successive, un'altro termine importante sarà *campo base*, relativo all'alpinismo. Assente ancora nei testi ottocenteschi (o meglio, presente in questi dati antichi solo occasionalmente come un falso positivo), comincia ad apparire nei primi decenni del Novecento, ad es. in *La Stampa sportiva* (1914)<sup>6</sup>, nel *Bollettino del Club alpino italiano* (1925)<sup>7</sup> e nella *Gazzetta ufficiale del Regno d'Italia* (1910)<sup>8</sup>. Dopo il 2000, *campo base* diventa *leader word* della famiglia *N-base* in termini di *token frequency*, riflettendo probabilmente il fatto che in quel periodo l'alpinismo sta diventando uno sport di massa. La concorrenza diacronica tra le tre *leader words* è visualizzata nel Grafico 6.

Come nel caso della famiglia *N-modello* (cfr. *podere modello*), anche per *N-base* la *leader word* più antica, cioè *punti base*, viene emarginata con l'andar del tempo senza scomparire del tutto, mentre appaiono nuove *leader words*.

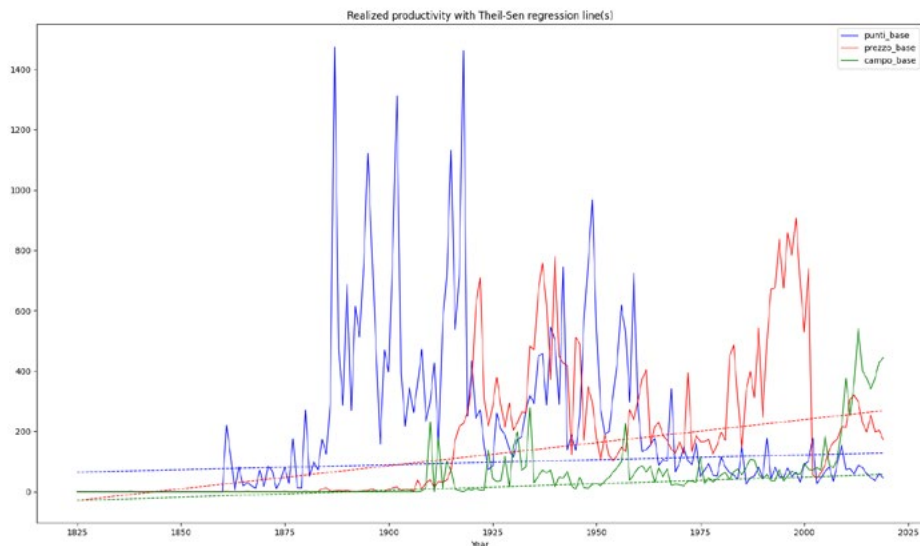
La particolarità della famiglia *N-base*, rispetto ad altre famiglie ATAP, risiede nel fatto che questa costruzione entra in concorrenza diretta con la costruzione preposizionale *N-di-base*, almeno per alcuni N1: così, ad esempio, *prezzo base* e *prezzo di base* risultano costrutti sinonimi. In generale, la rivalità tra il pattern *N-N* e *N-di-N* varia molto rispetto alla tipologia del composto. Dal punto di vista quantitativo, Baroni, Guevara e Pirrelli (2009) riportano sul loro corpus una presenza sistematica delle varianti *N-di-N* per il 90,2% dei composti NN argomentali (come ad es. *protezione persone*) e per il 75,42% dei composti NN *grounding* (come ad es. *centro città*),

6 Link diretto GB.

7 Link diretto GB.

8 Link diretto GB.





**Grafico 6.** Relative token frequency delle tre leader words più significative: *punti base*, *prezzo base*, *campo base*

mentre nel caso dei composti attributivi le rispettive varianti N-di-N sono più rare (il 31,71%)<sup>9</sup>. Un'analisi più dettagliata condotta sui dati francesi (Radimský 2020) indica che, dal punto di vista teorico, non c'è da aspettarsi che la relazione attributiva presente all'interno di un NN (ad es. *parola chiave*) si possa esprimere mediante la costruzione N-di-N (*\*parola di chiave*); l'analisi empirica poi dimostra che se una coppia di questo tipo è effettivamente presente nei dati francesi, le due costruzioni – salve poche eccezioni<sup>10</sup> – non sono sinonime. Si noti che lo stesso fenomeno si può osservare in italiano, ad es. *problema chiave*<sup>11</sup> (*problema importante*) non è sinonimo di *problema di chiave*<sup>12</sup>; si può ipotizzare che molte delle varianti N-di-N dei composti NN attributivi registrati nello studio di Baroni, Guevara e Pirrelli (2009) rientrino in questa categoria (per l'italiano cfr. anche la discussione in Terreni 2005). La famiglia *N-base* italiana sembra invece rappresentare un'interessante eccezione, un vero e proprio caso di rivalità, e costituisce così un'istanza di *pattern competition* (Masini 2019: 290).

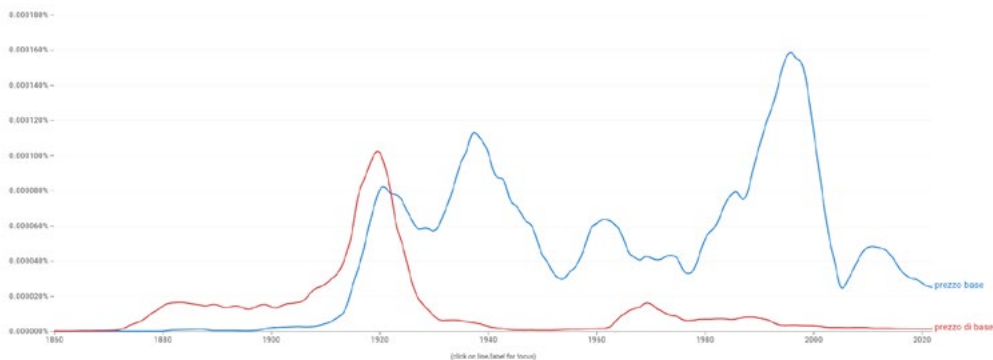
Purtroppo i dati GN che abbiamo a disposizione non ci permettono ancora di analizzare i pattern N-di-N nella loro complessità, un approfondito esame di questa rivalità sarà possibile

9 I composti *argomentali* e *grounding* (Baroni, Guevara e Pirrelli 2009) corrispondono a due sottotipi di composti subordinativi: i primi, chiamati *verbal-nexus* nella terminologia di Scalise e Bisetto (2009), manifestano una relazione di tipo “testa deverbale–argomento” tra i componenti, mentre i secondi no.

10 Ad es. *rôle (de) pivot* – ruolo fondamentale, let. “ruolo (di) perno” (Radimský 2020: 171).

11 Il **problema chiave** per la crescita italiana è che i tre maggiori fattori della domanda puntano nella direzione sbagliata. (Coris)

12 Se dall'esterno funziona e dall'interno no è un **problema di chiave** (prova ad usare un'altra chiave) oppure di cilindro (devi cambiare il cilindro). (Google)



**Grafico 7.** *Relative token frequency – rivalità diacronica di prezzo base vs. prezzo di base (Google n-gram viewer)*

solo in studi successivi. Ci limitiamo perciò a illustrare un caso emblematico di questa rivalità (*prezzo base* vs. *prezzo di base*) tramite il Google n-gram viewer, riportato nel Grafico 7.

Nel caso riportato nel Grafico 7, la variante NN, più recente, vince la competizione già alla fine degli anni '20 del Novecento.

## 5.4 La famiglia *N-limite*

Le descrizioni del valore semantico che il nome *limite* assume nella costruzione *N-limite*, fornite rispettivamente dal GDLI (10) e dallo Zingarelli (11), sembrano a prima vista divergenti. Tuttavia, si noti che i concetti cui ci si riferisce, quali *definitivo*, *ultimo* (10), *insuperabile*, *improbabile*, *ai confini del possibile* (10, 11), derivano tutti dallo stesso significato metaforico del nome *limite* (letteralmente *confine*, *barriera*), descritto dallo Zingarelli (2022) sotto la seconda accezione della rispettiva voce, in maniera seguente: “grado ultimo, valore estremo, da non superare”. Le definizioni (10) e (11) risultano perciò complementari e andrebbero riunite in una sola definizione complessa. Lo si vede nel tentativo di interpretare correttamente i diversi NN presenti nell’uso. Ad esempio, *data limite*, *prezzo limite* o *velocità limite* si riferiscono al concetto di “insuperabile” citato nella definizione (10), *stato limite* al concetto di “ultimo/definitivo” (10), mentre *ipotesi limite* si riferisce piuttosto alla “improbabilità” riportata nella definizione (11). I confini semantici in questione consentono ovviamente un margine interpretativo relativamente largo.

(10) Con valore di apposizione: definitivo, ultimo, insuperabile.

*Nell’età romantica, l’idea del circolo... cedette innanzi all’idea del corso lineare, desunta dal cristianesimo, e del progresso a termine, che si conclude con uno **stato limite** o con l’entrata in un paradiso di progresso indefinito, d’incessante gioia senza dolore.* (Croce, GDLI) (GDLI, voce *limite*; solo alcuni esempi sono stati riprodotti)

- (11) In funzione di agg. inv. (posposto al sost.) [...] 2 fig. detto di ciò che rappresenta il massimo dell'improbabilità, che sta ai confini del possibile: *ipotesi limite*; *si tratta di casi limite* (Zingarelli, 2022, voce *limite*)

Anche nel caso della famiglia *N-limite*, e forse ancor più che nelle due precedenti, risulta evidente che le origini vanno rintracciate soprattutto nel linguaggio specialistico e scientifico, legato a una sorprendente varietà di discipline. Va innanzitutto ricordato che il nome *limite* (al di fuori della costruzione NN) diventa già nell'Ottocento un importante termine matematico, che si riferisce al numero al quale i valori di una funzione o di una successione si mantengono arbitrariamente vicini ("limite di una funzione" o "limite di una successione") (v. Zingarelli 2022, voce *limite*, accezione no. 6). Ma, soprattutto, *limite* fa parte di composti NN terminologici presenti in varie discipline, alcuni dei quali citati anche dai dizionari. Zingarelli nota il termine matematico *punto limite* (12). Dall'ambito della filosofia, GDLI (13) ricorda il termine jaspersiano *situazione limite* (ted. *Grenzsituation*, ingl. *limit-situation*, fr. *situation limite*) e il termine *concetto limite*. Dalla terminologia dell'ottica provengono i termini *angolo limite*, *incidenza limite* e *rifrazione limite*, riportati sempre dallo GDLI (14).

- (12) Mat. nella loc. *punto limite*, nel quale concorrono le rette che in una data rappresentazione sono le immagini di rette parallele. (Zingarelli, 2022, voce *limite*)

- (13) 7. **Filos.** Termine che delimita, determina, definisce un ente finito nell'infinito spaziale, sostanziale o concettuale.

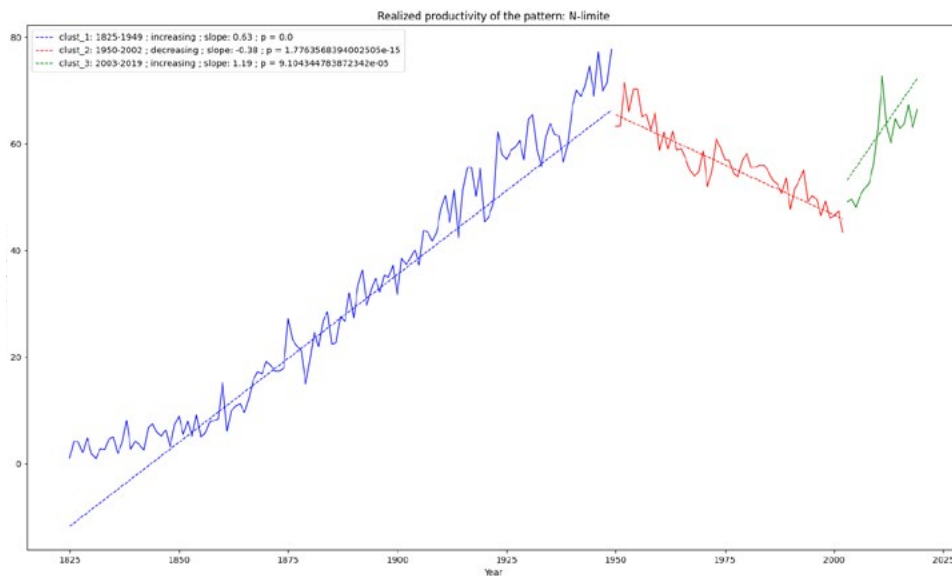
- **Situazioni limite:** nel pensiero del filosofo tedesco Karl Jaspers (1883–1969), quelle in cui l'uomo non può non venire a trovarsi (la colpa, la sofferenza, la morte) e dalle quali non può evadere.
- **Concetto limite:** quello che costituisce un'astrazione e un'esagerazione della realtà, e che ha per lo più il valore di formulazione di un'ipotesi o di escogitazione di un espediente metodologico.

*Bruti e santi sono concetti-limite o astrazioni, e non esistono nella realtà.* (Croce, GDLI) (GDLI, voce *limite*)

- (14) 18. **Ott. Angolo limite:** angolo massimo di rifrazione, corrispondente all'angolo di incidenza di 90°, quando un raggio luminoso passa da un mezzo trasparente a un altro più rifrangente (al contrario, se il raggio passa da un mezzo trasparente a un altro meno rifrangente, il termine indica l'angolo massimo di incidenza corrispondente all'angolo di rifrazione di 90°, oltre il quale non si ha più rifrazione e tutta la luce incidente si trova nel raggio riflesso).

- **Incidenza o rifrazione limite:** quelle definite dall'angolo limite nel passaggio di un raggio luminoso da un mezzo trasparente a un altro rispettivamente meno o più rifrangente. (GDLI, voce *limite*)

Per la famiglia *N-limite* abbiamo individuato 99 NN attributivi nei Google n-grams, il Grafico 8 ne riproduce la panoramica diacronica della *realized productivity*.



**Grafico 8.** *Realized productivity (type frequency) della famiglia N-limite*

L'evoluzione diacronica della *realized productivity* di *N-limite* è nella sua forma molto simile a quella della famiglia *N-base* riportata nel Grafico 5, cioè si riscontrano le stesse tendenze (che risultano forse ancora più marcate), segnate da simili confini temporali. La crescita della curva di *N-limite* dalle origini ottocentesche fino al 1949 è più regolare rispetto a quella di *N-base*, senza ulteriori balzi, ed è nel 1949 che la curva raggiunge il suo picco assoluto di produttività. Il declino della produttività della curva *N-limite* nella seconda metà del Novecento (1950–2002) è più pronunciato rispetto a quella di *N-base*, mentre il risveglio della produttività dopo il 2003 segue un profilo molto simile in ambedue i casi.

Gli esempi di *N1* più significativi in termini di *token frequency*, ossia le *leader words* della famiglia *N-limite* di diverse fasi diacroniche, sono riprodotte nella Tabella 3.

Si noti che nel caso di *N-limite*, gli elenchi delle *leader words* per le diverse fasi sono particolarmente stabili. In effetti, *punto limite*, *caso limite* e *valore limite* si riscontrano ai primi posti in tutte le fasi, mentre *età limite* fa un picco eccezionale negli anni '40 del Novecento. Il termine *situazione limite* compare sporadicamente già nei testi del primissimo Novecento, testimoniando la vitalità del pattern *N-limite*: ad esempio, nella *Rivista italiana di sociologia* del 1911, il termine *situazione-limite* viene usato nell'articolo intitolato "Il concetto di progresso", in (15), con riferimento a una meta ideale e astratta da raggiungere (cfr. la definizione del termine *concetto limite* in 13), cioè in un senso diverso rispetto a quello che gli darà negli anni '30 del Novecento il filosofo Karl Jaspers (13) e che farà salire questo NN tra i più frequenti della famiglia *N-limite* a partire dagli anni '80 del Novecento.

(15) È lecito quindi fissare in ogni singolo campo, almeno idealmente, una *situazione-limite*, e constatare se nel corso dei secoli l'umanità vi si sia avvicinata. (Savorgnan 1911: 567)

Periodo LWs	< 1930	1930–1945	1945–1960	1960–1980	1980–2000	2000+
	punto	strato	caso	caso	caso	valori
	caso	caso	strato	casi	casi	caso
	valore	età	casi	strato	valori	valore
	curva	valore	valore	valore	data	casi
	superficie	punto	punto	prezzi	valore	situazioni
	angolo		concetto	punto	situazioni	situazione
	posizione				situazione	punto
	linea				punto	data
	età					
	utilità					

**Tabella 3.** *Leader words* della famiglia *N-limite*\* (dati non lemmatizzati, ordinati per *rank* decrescente)

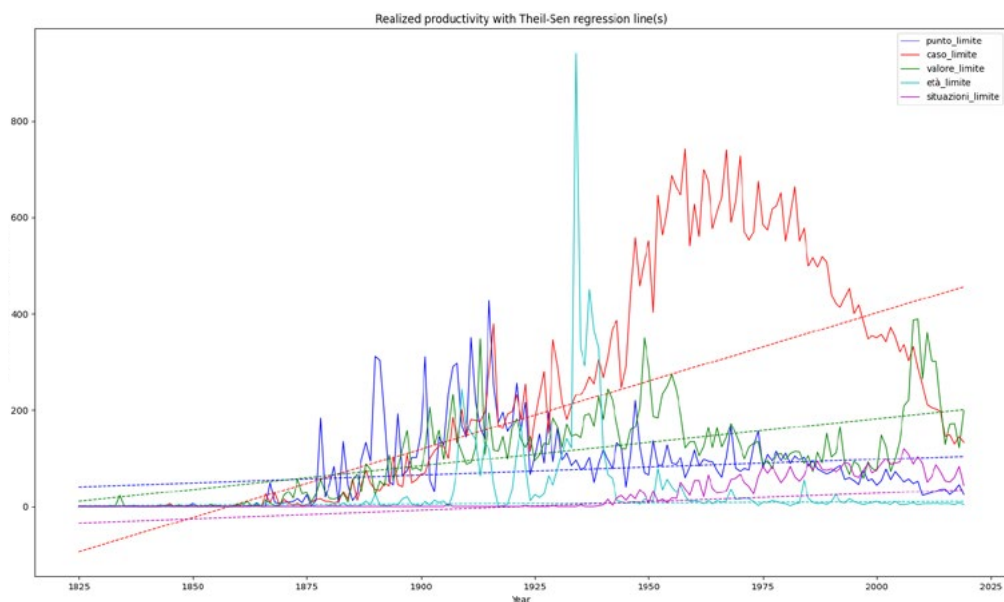
\* Dato che nei dati a partire dal 1945 le due varianti flesse della stessa parola (*caso/casi*) vengono spesso a occupare i primi *rank*, riportiamo più dati, di modo a fornire 5 lemmi per ogni periodo.

Per quanto riguarda le varietà della lingua, anche le origini della famiglia *N-limite* risalgono a registri specialistici di fine Ottocento e inizio Novecento, anche se – a differenza della famiglia *N-base* – ai primi posti della classifica non si tratta esclusivamente di termini specialistici collegabili a una disciplina precisa. La prima *leader word*, *punto limite*, è un termine della matematica (12), le altre come *caso limite*, *valore limite* o *età limite* non sono collegabili a una disciplina concreta, ma un controllo qualitativo dei dati su Google books anteriori al 1950 mostra una chiara prevalenza di testi specialistici. Lo si può esemplificare tramite un brano tratto dall'articolo *Agganciamenti automatici all'esposizione di Milano*, pubblicato nel 1907 sulla rivista *L'ingegneria ferroviaria* (16).

(16) *Agganciamenti automatici all'esposizione di Milano [...] Nel caso limite di un urto eccezionalmente violento, cui non bastasse la reazione del sistema descritto, il bilanciario, giunto in fin di corsa della sua feritoia, trascina anche l'asta per tutta la corsa che ha già ora in servizio (6 cm) ed inflette così la molla al centro del vagone. (IF 1907: 62)*

Tuttavia, in posizioni più basse della lista di frequenza, troviamo già una serie di termini (altamente) specializzati provenienti da un'ampia gamma di discipline, la cui storia risale almeno agli inizi del Novecento (ma spesso anche oltre). Tra questi vi sono termini della matematica e della geometria (*funzione limite*, *minimo limite*, *piano limite*, *cerchio limite*, *insieme limite*, *classe limite*), della fisica (*strato limite*, *carico limite di snervamento*, *curva limite*, *momento limite*), dell'ingegneria ferroviaria (*sagoma limite*), dell'economia (*utilità limite*), dell'arte militare (*linea limite*), della filosofia (*concetto limite*, *forma limite*), della chimica (*titolo limite*, *formula limite*), etc.

La concorrenza diacronica tra le cinque *leader words* è visualizzata nel Grafico 9.



**Grafico 9.** Relative token frequency delle leader words più significativi: *punto limite*, *caso limite*, *valore limite*, *età limite*, *situazioni limite*

Il calo di *punto limite* a partire dagli anni '20 del Novecento può probabilmente spiegarsi con un abbassamento relativo della proporzione dei testi specialistici di matematica nel corpus sottostante. Le ragioni per cui il NN *età limite* (con riferimento di preferenza a nomi [+um]) spicca nei testi pubblicati intorno al 1934 meriterebbero un apposito studio qualitativo, anche se sono molto probabilmente interpretabili in riferimento alla situazione politica e sociale particolare dell'Italia fascista alle soglie della Seconda guerra mondiale. La fortuna di *valore/valori limite* dopo il 2000 è ricollegabile a un risvegliato interesse per la protezione dell'ambiente al livello istituzionale e normativo.

## 6. Conclusioni

L'analisi presentata in questo articolo ha permesso di mettere a confronto l'emergere graduale del pattern compositivo NN ATAP, rappresentato dal campione di 2.541 *types* non lemmatizzati che includono 56 N2 (nomi modificatori) diversi, con lo sviluppo progressivo di tre costruzioni semi-schematiche concrete (*N-modello*, *N-base* e *N-limite*), che hanno in gran parte contribuito alla sua formazione.

Dal punto di vista globale, lo sviluppo del pattern ATAP sembra segnato da due punti di svolta importanti, situati rispettivamente intorno al 1950 e al 2000. Nella prima fase, che comprende l'Ottocento e la prima metà del Novecento, emergono le radici di tutte le famiglie N2 sotto forma di prime attestazioni delle *leader words* rispettive; lo si evince dal fatto che verso la fine degli anni Quaranta, la *relative family type frequency* dell'intero pattern ATAP raggiunge il

suo picco assoluto. È pertanto il periodo di maggiore *innovazione* (nel senso di Traugott e Trousdale 2013), accompagnato dalla *costruzionalizzazione* di alcune famiglie. Il periodo successivo (1950–2000) sembra radicalmente diverso: il numero globale di famiglie N2 diminuisce, il che implica una minore copertura del pattern ATAP, ma la *realized productivity* dei composti ATAP cresce in maniera esponenziale grazie all'apposito sviluppo di produttività di un sottoinsieme limitato di famiglie N2. Nell'ultimo periodo, dopo il 2000, il pattern ATAP riprende la sua vitalità: sia il numero di famiglie N2 che la loro *realized productivity* ricomincia a crescere in maniera esponenziale.

Dal punto di vista quantitativo è interessante osservare che la sorte delle famiglie *N-base*, *N-limite* e, con un certo margine di precisione, anche quella di *N-modello* segue le orme dell'intero pattern ATAP. Per *N-base* e *N-limite*, i periodi del primo sviluppo, seguito da stagnazione o addirittura declino, e la successiva ripresa replicano gli stessi confini temporali, cioè il 1950 e il 2000. Per *N-modello*, invece, la fase della stagnazione inizia già intorno al 1925 e anche la ripresa si verifica prima, alla fine degli anni '70, motivata dalla coniazione del termine *lettore modello* da Umberto Eco, che diventa una *leader word* per una nuova serie di composti il cui primo elemento denota un essere umano. In altre parole, queste tre famiglie hanno giocato un ruolo decisivo nella fase delle origini del pattern ATAP, ma, tranne per *N-modello* nel ventennio 1980–2000, non hanno contribuito alla crescita della sua *realized productivity* nella seconda metà del Novecento.

Un'analisi qualitativa delle *leader words* e dei relativi contesti ha chiaramente mostrato che le famiglie *N-base*, *N-limite* e *N-modello* traggono origine dal linguaggio tecnico-scientifico dell'Ottocento e del primo Novecento. È pertanto pienamente lecito supporre che l'ambiente scientifico internazionale abbia avuto un'influenza decisiva sulla loro formazione, cioè che molte *leader words* siano calchi presi da o ispirati a importanti lingue europee, come il francese (cfr. le famiglie francesi *N-base*, *N-limite* e *N-modèle*), il tedesco o l'inglese. Ulteriori ricerche saranno necessarie per mettere in luce i meccanismi di queste influenze reciproche.

## Riferimenti bibliografici

- Arcodia, G. F.; Grandi, N.; & Montermini, F. (2009). Hierarchical NN compounds in a cross-linguistic perspective. *Italian Journal of Linguistics*, 21(1), 11–33.
- Arnaud, P. J. L. (2003). *Les composés timbre-poste* (Travaux du C.R.T.T). Lyon: Presses Universitaires de Lyon.
- Baayen, R. H. (2009). Corpus linguistics in morphology: Morphological productivity. In A. Lüdeling, & M. Kytö (Eds.), *Handbooks of Linguistics and Communication Science* (pp. 899–919). Berlin & New York: Mouton de Gruyter. <https://doi.org/10.1515/9783110213881.2.899>
- Baroni, M.; Guevara, E., & Pirrelli, V. (2009). Sulla tipologia dei composti N+N in italiano: principi categoriali ed evidenza distribuzionale a confronto. In G. Ferrari, R. Benatti, & M. Mosca (a cura di), *Linguistica e modelli tecnologici di ricerca: atti del XL Congresso internazionale di studi della Società di linguistica italiana: Vercelli, 21–23 settembre 2006* (SLI 51) (pp. 73–96). Roma: Bulzoni.
- Bauer, L. (2017). *Compounds and Compounding* (1<sup>st</sup> ed.). Cambridge: Cambridge University Press.
- Bisetto, A.; & Scalise, S. (2005). The Classification of Compounds. *Lingue e linguaggio*, 2, 319–332. <https://doi.org/10.1418/20728>
- Booij, G. (2016). Construction Morphology. In A. Hippisley, & G. Stump (Eds.), *The Cambridge Handbook of Morphology* (1<sup>st</sup> ed.) (pp. 424–448). Cambridge: Cambridge University Press. <https://doi.org/10.1017/9781139814720.016>
- Booij, G. (2010). *Construction morphology*. Oxford: Oxford University Press.
- Fradin, B. (2009). IE, Romance: French. In R. Lieber, & P. Štekauer (Eds.), *The Oxford Handbook of Compounding*. Oxford: Oxford University Press (pp. 417–435). <https://doi.org/10.1093/oxfordhb/9780199695720.013.0022>
- GDLI = Battaglia, S. (1961–2002). *Grande dizionario della lingua italiana*. 21 voll. Torino: UTET. <https://www.gdli.it/>
- Grandi, N. (2009). When Morphology ‘Feeds’ Syntax: Remarks on Noun > Adjective Conversion in Italian Appositive Compounds. In F. Montermini, G. Boyé, & J. Tseng (Eds.), *Selected Proceedings of the 6<sup>th</sup> Décebrettes* (pp. 111–124). Somerville: Cascadilla Proceedings Project.
- 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*, 1(1), 161–176. <https://doi.org/10.1418/34543>
- Gries, T. S.; & Hilpert, M. (2008). The identification of stages in diachronic data: Variability-based neighbour clustering. *Corpora*, 3(1), 59–81. <https://doi.org/10.3366/E1749503208000075>
- Hilpert, M. (2021). *Ten lectures on diachronic construction grammar* (Distinguished lectures in cognitive linguistics vol. 26). Leiden: Brill.
- Hussain, Md.; & Mahmud, I. (2019). pyMannKendall: a python package for non parametric Mann Kendall family of trend tests. *Journal of Open Source Software* 4(39), 1556. <https://doi.org/10.21105/joss.01556>
- Jackendoff, R.; & Audring, J. (2020). *The texture of the lexicon: relational morphology and the parallel architecture* (1st ed.). Oxford: Oxford University Press.
- Kovář, V.; & Herman, O. (2013). *Methods for Detection of Word Usage over Time*. Brno: Tribun EU.



- IF 1907 = L'ingegneria ferroviaria. Bollettino del Collegio nazionale degli ingegneri ferroviari italiani (1907), 4(16): 272.
- Masini, F. (2019). Competition between morphological words and multiword expressions. Competition in inflection and word-formation. In F. Rainer, F. Gardani, W. U. Dressler, & H. C. Luschützky (Eds.), *Competition in Inflection and Word-Formation* (Studies in Morphology 5) (pp. 281–305). Cham: Springer. [https://doi.org/10.1007/978-3-030-02550-2\\_11](https://doi.org/10.1007/978-3-030-02550-2_11)
- Micheli, M. S. (2020). *Composizione italiana in diacronia: Le parole composte dell'italiano nel quadro della Morfologia delle Costruzioni*. Berlin: De Gruyter. <https://doi.org/10.1515/9783110652161-001>
- Radimský, J. (2023a). Where did the Italian Verbal-Nexus N+N compounds come from? In J. Audring, K. T. Frantzi, N. Koutsoukos, G. Markopoulos, & K. Nikolou (Eds.), *Proceedings of the Mediterranean Morphology Meetings, Vol. 13, Comparing typologies* (pp. 71–82). Patras: University of Patras. <https://pasishee.library.upatras.gr/mmm/issue/view/436>
- . (2023b). Tracing back the history of Italian Attributive-Appositive Noun+Noun compounds: First outcomes. *Linguistica Pragensia*, 33(2). <https://doi.org/10.14712/18059635.2023.2.3>
- . (2022). Towards a diachronic analysis of Romance morphology through Google n-grams. *Corpus*, 23. <https://doi.org/10.4000/corpus.7205>
- . (2020). Are French NNs variants of N-PREP-N constructions? A corpus-based study of two competing patterns. *Linguistica Pragensia*, 30(2), 156–186. <https://doi.org/10.14712/18059635.2020.2.4>
- . (2018). Inflection of binominal ATAP compounds in French and Italian: a paradigmatic account. *Lingue e linguaggio*, 17(2), p. 261–272.
- . (2015). *Noun+Noun compounds in Italian: a corpus-based study* (Episteme. Theoria). Vydání 1. České Budějovice: Jihočeská univerzita v Českých Budějovicích.
- Rainer, F. (2021). Compounding: From Latin to Romance. In M. Aronoff (Ed.), *Oxford Research Encyclopedia of Linguistics* (pp. 1–31). Oxford: Oxford University Press. <https://doi.org/10.1093/acrefore/9780199384655.013.691>
- Rainer, F. (2016). Italian. In P. O. Müller, I. Ohnheiser, S. Olsen, & F. Rainer (Eds.), *An International Handbook of the Languages of Europe, Volume 4 Word-Formation* (pp. 2712–2731). Berlin: De Gruyter Mouton. <https://doi.org/10.1515/9783110379082-017>
- Savorgnan, F. (1911). Il concetto di progresso. *Rivista italiana di sociologia*, 15 (sett.-ott.)
- Scalise, S.; & Bisetto, A. (2009). The Classification of Compounds. In R. Lieber, & P. Štekauer (Eds.), *The Oxford Handbook of Compounding* (pp. 34–53). Oxford: Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780199695720.013.0003>
- Terreni, R. (2005). Composti N + N e sintassi: i tipi economici lista nozze e notizia-curiosità. In A. M. Thornton, & M. Grossmann (Eds.), *La formazione delle parole. Atti del XXXVII Congresso internazionale di studi della Società di linguistica italiana, L'Aquila, 25–27 settembre 2003* (SLI 48) (pp. 521–546). Roma: Bulzoni.
- Thornton, A. M. (2004). Conversione. In M. Grossmann, & F. Rainer (a cura di), *La formazione delle parole in italiano* (pp. 499–533). Tübingen: Max Niemeyer Verlag.
- Traugott, E. C.; & Trousdale, G. (2013). *Constructionalization and constructional changes* (Oxford linguistics 6) (1st ed.). Oxford: Oxford University Press.

**I composti N+N attributivi in diacronia: le famiglie *N-modello*, *N-base* e *N-limite***

Zanchetta, E.; & Baroni, M. (2005). Morph-it! A free corpus-based morphological resource for the Italian language. In P. Danielsson (Ed.), *Proceedings from the Corpus Linguistics Conference Series, 1(1)*. Birmingham: University of Birmingham. <http://www.birmingham.ac.uk/research/activity/corpus/publications/conference-archives/2005-conf-e-journal.aspx>

Zingarelli (2022) = N. Zingarelli (2022), *Lo Zingarelli 2022, vocabolario della lingua italiana*. Bologna: Zanichelli.

**Fonti primarie citate da Google Books**

Angelo Melano di Portula, *Dizionario analitico di diritto e di economia industriale e commerciale*, Torino, Giuseppe Pomba e comp., 1843.

Franco Savorgnan: Il concetto di progresso, *Rivista italiana di sociologia*, anno XV, Fascicolo di Settembre-Ottobre 1911, Fratelli Bocca editori, Torino-Milano-Roma, p. 567.

L'ingegneria ferroviaria, *Bollettino del Collegio nazionale degli ingegneri ferroviari italiani*, 1907, Vol. IV, No. 16, Italia: Stab. Tip. G. Civelli, p. 272.



This work can be used in accordance with the Creative Commons BY-SA 4.0 International license terms and conditions (<https://creativecommons.org/licenses/by-sa/4.0/legalcode>). This does not apply to works or elements (such as images or photographs) that are used in the work under a contractual license or exception or limitation to relevant rights.