

# Annales Universitatis Paedagogicae Cracoviensis

Studia de Cultura 9(3) 2017

ISSN 2083-7275

DOI 10.24917/20837275.9.3.11

**Valentina Piunno**

Università Roma Tre

**Federica Cominetti**

Università di Firenze

## Formati di parola inediti: dati emersi dal dizionario *CombiNet*<sup>1</sup>

### 1. I formati di parola nel dizionario *CombiNet*<sup>2</sup>

Il presente lavoro nasce nell'ambito di studi di un progetto di ricerca nazionale che ha come scopo lo studio della combinatoria lessicale in italiano e la realizzazione di un dizionario di combinazioni di parole basato su corpora, *CombiNet*. Il dizionario *CombiNet* è un database lessicografico disponibile in rete, che si propone come strumento di rappresentazione della combinatoria del lessico italiano.

La prima fase della ricerca dedicata alla costituzione del dizionario ha portato alla luce una lista dei tipi combinatori più comuni dell'italiano (Piunno *et al.* 2013), raccolta sulla base di alcuni recenti studi lessicologici dedicati alle strutture sintagmatiche dell'italiano (Voghera 1994, Voghera 2004, Simone & Masini 2007, Masini 2012, Piunno 2013). I diversi tipi sintagmatici sono stati successivamente estratti dai corpora *La Repubblica* e *Paisà* secondo specifici indici statistici che ne individuassero la maggiore forza di associazione (Passaro – Lenci 2016, Castagnoli *et al.* 2015, Castagnoli *et al.* 2016). A tale scopo, sono stati identificati i "formati di parola", ovvero le configurazioni sintattico-semantiche associate ai diversi profili combinatori, che sono stati successivamente distinti a seconda della "categoria di uscita" (la funzione che una combinazione di parole assume in contesto sintagmatico). Il layout lessicografico realizzato per *CombiNet* è stato creato per rappresentare contemporaneamente l'informazione sintagmatica, semantica e funzionale di una combinazione di parole, contribuendo a dar conto della stretta interconnessione tra la pratica lessicografica e l'indagine lessicologica (Piunno 2016b). In particolare, la struttura del layout lessicografico di *CombiNet* include diversi tipi di informazioni, tra cui:

---

<sup>1</sup> Ai soli fini concorsuali si segnala che i paragrafi 1, 2, 2.2, 2.3 sono da attribuirsi a Valentina Piunno e i paragrafi 2.1, 2.1.1, 2.1.2, 2.1.3, 2.1.4 sono a cura di Federica Cominetti. Il paragrafo 3 è opera di entrambi gli autori.

<sup>2</sup> *CombiNet* è un dizionario combinatorio della lingua italiana, creato dalla collaborazione tra l'Università Roma Tre, l'Università di Pisa e l'Università di Bologna nell'ambito del Progetto di Ricerca Nazionale 2010–2011 (n. 20105B3HE8) "*CombiNet* – Word Combinations in Italian: theoretical and descriptive analysis, computational models, lexicographic layout and creation of a dictionary", finanziato dal Ministero dell'Educazione, Università e Ricerca. Sito internet, con accesso limitato: [www.combinet.fileli.unipi.it](http://www.combinet.fileli.unipi.it).

- a) il tipo sintagmatico, la cui rappresentazione varia al variare della natura (es. collocazione, polirematica, espressione idiomatica, proverbio);
- b) la categoria di uscita o funzione dell'unità combinatoria (es. nominale, aggettivale, avverbiale).
- c) le configurazioni sintagmatiche associate alle combinazioni, vale a dire la mera sequenza di parti del discorso (es. [AGG + NOME], [NOME + PREP + NOME], ecc.).
- d) i *formati di parola*, vale a dire le configurazioni sintattico-semantiche che le combinazioni di una lingua realizzano.

## 2. I formati di parola inediti

Il dizionario include i diversi tipi sintagmatici identificati nei numerosi lavori recentemente pubblicati nell'ambito degli studi dedicati al lessico, tra cui in particolare sono stati selezionati:

- a. Collocazioni e combinazioni preferenziali (es. *abbattere una casa, assente ingiustificato*);
- b. Verbi supporto con complemento (es. *stare fresco, mettere fretta, fare una corsa*);
- c. Parole sintagmatiche
  - nomi sintagmatici (es. *punto di vista, macchina da scrivere*);
  - verbi complessi (es. *tirare su, avere a cuore, risultare facile, attraversare incolume*);
  - aggettivi sintagmatici (es. *in codice, a senso unico, in bianco e nero, all'acqua di rose*);
  - avverbi sintagmatici (es. *a rate, a colpo sicuro, a mezza bocca, volente o nolente*);
  - preposizioni sintagmatiche (es. *per mezzo di, al fine di, assieme a*);
  - congiunzioni sintagmatiche (es. *dopo che, fintanto che, salvo che*).
- d. Costruzioni binomiali o trinomiali (es. *mordi e fuggi, gratta e vinci, aglio olio e peperoncino*).
- e. Espressioni idiomatiche (es. *prendere il toro per le corna*).
- f. Interiezioni (es. *grazie al cielo!*).
- g. Proverbi (es. *can che abbaia non morde*).

I diversi tipi sintagmatici possono essere quindi rappresentati da molteplici configurazioni morfosintattiche e semantiche (*formati di parola*), che costituiscono un problema di ordine lessicologico, lessicografico e computazionale. Nel corso della stesura delle voci del dizionario, sono emersi formati di parola che non hanno ancora ricevuto attenzione in letteratura, e che per tale motivo abbiamo definito "inediti". Tra questi, abbiamo formati nuovi dal punto di vista della configurazione sintagmatica (cfr. § 2.1), dal punto di vista semantico (cfr. § 2.2), e dal punto di vista della funzione sintattica (cfr. § 2.3).

## 2.1. Formati di parola inediti dal punto di vista sintattico

In questa sezione intendiamo presentare i formati di parola non individuati dalla letteratura, che sono emersi nel corso dell'elaborazione del dizionario *CombiNet*.

Nell'individuazione di nuovi formati sono state rilevate difficoltà legate alle diverse caratteristiche di annotazione dei corpora. Ad esempio, in alcuni tagset, possessivi, numerali e indefiniti non sono annotati come sottocategorie di aggettivi, ma come categorie a sé. In tal senso, la ricerca PREP AGG NOME non restituirà la combinazione *a sua volta*, e AGG NOME non restituirà *primo ministro*. Come ulteriore esempio, si consideri che nel corpus La Repubblica tutte le varie funzioni dell'elemento *che* sono sottospecificate in un tag molto vago (CHE, per l'appunto).

### 2.1.1. Categoria nominale

Per quanto riguarda la categoria nominale, tra i casi più interessanti si riportano alcune polirematiche con testa preposizionale (*fuori programma, pro forma, fuori legge*) e con testa participiale (*facente funzioni, messa in opera*). Si segnala anche la struttura a testa nominale seguita da due aggettivi (*senso unico alternato*, ben rappresentata anche da termini specialistici come *servizio sanitario nazionale* e *Fondo Monetario Internazionale*).

Per quanto riguarda le combinazioni, ve ne sono diverse caratterizzate da testa nominale e struttura sintagmatica piuttosto articolata, come NOME AGG PREP (DET) NOME (*olio extravergine di oliva, persona informata sui fatti*), NOME PREP (DET) NOME AGG (*causa di forza maggiore, avvocato delle cause perse*), NOME PREP (DET) AGG NOME (*crisi del settimo anno*), AGG NOME PREP (DET) NOME (*primo giorno di scuola*).

### 2.1.2. Categoria aggettivale

Tra le strutture di aggettivo polirematico con testa preposizionale inedite si segnalano sequenze che arrivano a comprendere fino a cinque posizioni, come PREP NOME PREP AGG NOME (*nell'esercizio delle sue funzioni*), PREP AGG CONG AGG NOME (*di sana e robusta costituzione*) e PREP AGG NOME CONG NOME (*di ogni ordine e grado*). Di grande interesse sono anche gli aggettivi polirematici a testa nominale, in particolare con struttura NOME PREP NOME (*chiavi in mano, bocca a bocca, porta a porta*) o NOME NOME (*fronte retro*). Piuttosto produttivo sembra essere il profilo combinatorio a testa avverbiale AVV PREP (DET) NOME (*avanti con l'età, avanti con gli anni, su di giri, fuori di testa*). Le sequenze combinatorie inedite includono articolate strutture a testa aggettivale: AGG PREP AGG NOME (*degnò di miglior causa, unico nel suo genere*), AGG PREP NOME (*sano di mente, facile all'ira, fatto in casa, informato sui fatti*).

### 2.1.3. Categoria avverbiale

La letteratura riconosce l'esistenza di un vasto catalogo di combinazioni con funzione avverbiale. Le locuzioni avverbiali, diversamente da altre espressioni multiparola, sono infatti presenti in tutti i dizionari e le grammatiche "tradizionali". La ricerca per la realizzazione del dizionario *CombiNet* ha in ogni caso messo in evidenza l'esistenza di alcune sequenze combinatorie con funzione avverbiale inedite. In particolare, mostrano una certa diffusione locuzioni avverbiali a testa nominale del

tipo DET NOME PREP PRO o DET NOME PREP AVV (*una volta per tutte, una volta per sempre*). Sono inoltre state individuate varie sequenze a testa preposizionale: PREP AGG PREP (DET) NOME (*nella peggiore delle ipotesi, nel migliore dei casi*), PREP AGG NOME PREP NOME (*nel vero senso della parola*). Si segnala inoltre la struttura PREP AGG (*da grande*). Infine, è stata individuata una locuzione a testa avverbiale, con struttura AVV PREP DET NOME (*fino alla fine*).

#### 2.1.4. Categoria verbale

La categoria verbale è quella per cui sono state individuate più sequenze combinatorie inedite. L'innovazione principale riguarda sequenze dal grado di cristallizzazione diverso, introdotte dalla negazione NEG, come NEG VERBO PREP AGG (*non chiedere di meglio*), NEG VERBO AGG (*non sembrare vero*), NEG VERBO (DET) NOME (*non avere età, non conoscere crisi, non aprire bocca, non capire un tubo*). Vi sono poi molte sequenze a testa verbale: VERBO DET NOME PREP NOME (*cambiare le carte in tavola*), VERBO (DET) NOME AGG (*avere le spalle larghe, avere le mani bucate, sapere il fatto proprio, rimangiarsi la parola data*), VERBO (DET) AGG NOME (*cercare la propria strada, dire l'ultima parola, dare il buon esempio*), VERBO (DET) NOME PREP NOME (*tenere le mani a posto, avere voce in capitolo*), VERBO PREP AGG (*prendere in pieno*), VERBO PREP AGG NOME (*respirare a pieni polmoni, contare sulle proprie forze*), VERBO PREP DET NOME PREP NOME (*stare con le mani in mano*), VERBO PREP AGG NOME (*andare a buon fine, passare in secondo piano, fare un certo effetto*), VERBO PREP NOM AGG (*vincere a mani basse, passare alle maniere forti*), VERBO CONG AGG (*tornare come nuovo*), VERBO CONG AVV (*tornare come prima*).

Alcuni formati inediti si individuano anche per i verbi supporto: VERBO DET AGG NOME (*fare il grande passo, fare un ottimo lavoro*), VERBO DET NOME PREP NOME (*dare un colpo di telefono*).

## 2.2. Formati semantici inediti

Molte combinazioni di parole emerse attraverso l'estrazione per le voci *CombiNet* rappresentano strutture sintagmatiche in grado di veicolare informazioni di tipo qualitativo o quantitativo. Tra queste abbiamo i *classificatori*, elementi nominali che forniscono informazioni semantiche a proposito del nome a cui sono sintatticamente uniti (1) e i *quantificatori sintagmatici* (2), strutture fisse in grado di attribuire una cornice quantitativa al nome che segue (Simone 2006)<sup>3</sup>:

(1) [NOME PREP]<sub>CLASSIFICATORE</sub><sup>4</sup> es. *briciola di (pane), cucchiaino di (olio), mazzo di (fiori, chiavi)*

(2) [DET NOME<<sub>massa/collettivo</sub>>PREP]<sub>QUANTIFICATORE</sub><sup>5</sup> es. *un mondo di, un mare di, una folla di*

<sup>3</sup> Cfr. l'analisi dei nomi leggeri svolta in Simone – Masini (2014).

<sup>4</sup> In questa rappresentazione il carattere maiuscolletto in pedice indica la funzione associata al sintagma.

<sup>5</sup> In questa rappresentazione il carattere in corsivo, in pedice tra le parentesi graffe, evidenzia le restrizioni semantiche associate ad una porzione del sintagma.

Vale la pena notare che gli esempi in (2) tendono perlopiù ad esprimere il concetto di numerosità attraverso l'intensificazione. D'altronde, la classe dei quantificatori è spesso associata a quella degli intensificatori, data la condivisione dell'informazione semantica relativa alla scalarità. A tal proposito, la raccolta dei dati per *CombiNet* ha mostrato non solo una vasta quantità di esempi combinatori, ma anche una ricca varietà di *formati di intensificazione*, tra cui in particolare si notano i seguenti, che non hanno ancora ricevuto sufficiente attenzione in letteratura:

a) *formati nominali*NOME AGG<sub>INTENS</sub>es. *freddo glaciale, evento spaventoso*AGG<sub>INTENS</sub> NOMEes. *autentico fallimento*AGG<sub>INTENS</sub> AGG<sub>sostantivizzato</sub><sup>6</sup>es. *vero idiota, totale imbecille*NOME [PREP NOME]<sub>INTENS</sub>es. *casa da sogno, vestito da urlo*b) *formati aggettivali*AGG AGG<sub>INTENS</sub>es. *pieno zeppo*AGG [PREP NOME]<sub>INTENS</sub>es. *bella da matti*AGG [PREP VERBO<sub>infinito</sub>]<sub>INTENS</sub>es. *bello da morire*c) *formati verbali*VERBO PREP NOME<sub>plurale</sub>es. *riempire di calci, tempestare di domande*

NEG VERBO DET NOME (AGG)

es. *non capire un cavolo, non importare un fico secco*d) *formati avverbiali*

DET NOME

es. *(piacere) un sacco, (dispiacere) un casino*

Tra questi, alcuni sono rappresentazioni di pattern collocazionali (ovvero sequenze di parole frequentemente associate tra loro ma non lessicalizzate, es. *tempestare di domande*), altri sono combinazioni di parole pienamente lessicalizzate (es. *non capire un cavolo*), altri ancora costituiscono vere e proprie costruzioni stabili e produttive, spesso riempite solo parzialmente di materiale lessicale (cfr. Piuanno 2017).

Appartengono al primo e al secondo tipo le cosiddette *costruzioni aggettivali superlative* (Berlanda 2013), vale a dire configurazioni sintagmatiche costituite da coppie di aggettivi sintatticamente giustapposti e semanticamente legati da un rapporto di intensificazione. La configurazione che rappresenta tali costruzioni è la seguente:

(3) a. [AGG<sub>INTENSIFICATO</sub> + AGG<sub>INTENSIFICATORE</sub>]b. *ubriaco perso, comico nato, bugiardo incallito*c. *stanco morto, bagnato fradicio, innamorato pazzo, caro arrabbiato, pieno zeppo*

Nelle costruzioni esemplificate in (3), l'aggettivo in seconda posizione attribuisce un valore scalare al primo aggettivo, elevandolo al massimo grado. Si può

<sup>6</sup> In questa rappresentazione il carattere minuscolo in pedice indica le restrizioni morfologiche associate ad una porzione del sintagma.

osservare che mentre le sequenze in (3b) rappresentano pattern produttivi, suscettibili di variazione lessicale (il primo aggettivo può essere sostituito con una serie di aggettivi anche semanticamente non affini), quelle in (3c) sono pienamente lessicalizzate, poiché invariabili. Inoltre, come si evince dagli esempi, i due aggettivi non sono necessariamente connessi a livello semantico: mentre *fradicio* condivide alcuni tratti semantici con *bagnato*, *arrabbiato* non ha alcuna connessione semantica con l'aggettivo *caro*. Un fenomeno simile si può riscontrare negli *intensificatori sintagmatici* (cfr. Benigni 2017), sintagmi preposizionali con funzione aggettivale o avverbiale, espressioni iperboliche utilizzate dal parlante per veicolare un significato rafforzativo o peggiorativo:

- (4) a. AGG|NOME<sub>INTENSIFICATO</sub> + Sintagma Preposizionale<sub>INTENSIFICATORE</sub>  
 b. *pazzo da legare, bello da morire, giornata da favola, casa da sogno*

A differenza degli esempi in (3), in questo caso le costruzioni sono specificate lessicalmente solamente in forma parziale: tranne alcuni rari esempi pienamente lessicalizzati, le sequenze in questione possono essere infatti riempite di materiale lessicale variabile, anche se spesso semanticamente omogeneo.

Se un intensificatore è una risorsa linguistica che ha la funzione di rendere scalare una proprietà (Bolinger 1972: 17), l'intensificazione può operare su una scala che può avere dei valori sia positivi sia negativi (Labov 1984: 44). Il caso della negazione usata come intensificatore rappresenta un fenomeno interessante, che varrebbe la pena indagare anche da un punto di vista interlinguistico. La negazione può operare come intensificatore in diversi tipi di formati di parola, tra cui emergono in particolare quelli verbali, spesso costruzioni discontinue con doppia negazione:

- (5) *non capire un tubo, non vedere un bel niente*

Queste costruzioni sono spesso altamente lessicalizzate, e sono caratterizzate da: a) negazione obbligatoria (la sequenza *\*vedere un tubo* non è ammessa, almeno nel suo significato idiomatico), b) presenza di un elemento nominale che ha subito un processo di grammaticalizzazione, assumendo la funzione di intensificatore.

Le strutture verbali obbligatoriamente negate rappresentano un fenomeno molto interessante nell'ambito degli studi sul lessico polirematico. Pur essendo presente la negazione obbligatoria (una sua omissione comporterebbe una lettura compositiva delle sequenze), molte combinazioni verbali non sempre esprimono una vera polarità negativa, e per tale motivo possono essere distinte in a) vere negative (es. *non vedere un tubo*), la cui negazione ha portata verbale, e b) false negative (es. *non vedere l'ora*), in cui la negazione non ha alcuna vera polarità negativa. In tali costruzioni, più una negazione è "falsa", maggiore sarà il grado di lessicalizzazione dell'intera sequenza (Piunno 2017).

### 2.3. Formati multifunzione inediti

Lo spoglio dei dati ha messo in evidenza un formato di parola particolare, la cui proprietà peculiare è la multifunzionalità. I *formati multifunzione* possono essere rappresentati da diversi tipi di configurazioni sintagmatiche, che hanno per output

sintagmi di diversa natura (perlopiù con funzione aggettivale e avverbiale). Si tratta di combinazioni di parole identiche a) nella forma (in quanto condividono una stessa configurazione sintagmatica e selezione lessicale) e b) nel significato (dato il medesimo contenuto semantico), che si differenziano per la funzione che ricoprono nel contesto sintattico. Si osservino gli esempi registrati nella tabella seguente:

- (6) Esempio: *mano nella mano, gambe all'aria*  
 Struttura: NOME PREP DET NOME  
 Funzione: a. Aggettivale: (*coppia*) *mano nella mano*, (*bambino*) *gambe all'aria*  
 b. Avverbiale: (*camminare*) *mano nella mano*, (*scivolare*) *gambe all'aria*
- (7) Esempio *chiavi in mano*  
 Struttura: NOME PREP NOME  
 Funzione: a. Aggettivale: (*automobile*) *chiavi in mano*  
 b. Avverbiale: (*vendere*) *chiavi in mano*
- (8) Esempio *fuori onda*  
 Struttura: PREP NOME  
 Funzione: a. Aggettivale: (*programma*) *fuori onda*  
 b. Avverbiale: (*commentare*) *fuori onda*  
 c. Nominale: (*mandare*) *un fuori onda*

Tutte le sequenze in (6)–(8) possono svolgere una duplice funzione – avverbiale e aggettivale –, a seconda del contesto sintattico in cui sono impiegate; la combinazione in (8) può anche essere sostantivizzata e impiegata in funzione nominale. In tutti i casi, le combinazioni rimangono immutate dal punto di vista semantico e sintattico. Nonostante i formati multifunzione siano diversi e piuttosto numerosi dal punto di vista dei tipi e degli esemplari che li rappresentano, sono tuttora poco studiati in letteratura (a parte l'ambito dei cosiddetti *modificatori misti*, pattern preposizionali con duplice funzione aggettivale e avverbiale, cfr. Piunno 2016a). Vale la pena notare inoltre che, mentre gli esempi in (6) sono rappresentativi di una configurazione sintagmatica molto frequentemente usata in funzione avverbiale e quelli in (8) si configurano secondo lo schema morfosintattico più produttivo per la formazione di costruzioni avverbiali e aggettivali (SPrep, cfr. Voghera 1994, 2004; Piunno 2013, 2015), gli esempi in (7) sono “anomali”. Questo perché la configurazione [Nome Preposizione Nome] è quella che più frequentemente realizza nomi sintagmatici e ha uno scarso impiego – e per conseguenza una limitata produttività – nella creazione di aggettivali o avverbiali.

### 3. Riflessioni conclusive

La stesura delle voci del dizionario *CombiNet* ha prodotto vari ordini di considerazioni, di natura metodologica e teorica.

In primo luogo, l'individuazione delle diverse categorie, dei formati di parola e dei fenomeni combinatori deve necessariamente fare riferimento ai dati reali,



anche allo scopo di individuare i diversi gradi di produttività che caratterizzano formati di parola diversi.

In secondo luogo, l'ampliamento dell'elenco dei formati di parola conduce a una riflessione di portata più ampia a livello teorico (Simone *et al.* 2015): il lessico dell'italiano è costituito da entità polirematiche e da combinazioni di parole in proporzione molto superiore a quanto la struttura tradizionale dei dizionari lasci intendere. In generale, tale osservazione porta ad auspicare la diffusione di modelli diversi di lessico e dizionario.

## Bibliografia

- Berlanda S. 2013. Constructional Intensifying Adjectives in Italian, Proceedings of the 9th Workshop on Multiword Expressions (MWE 2013), Atlanta, Georgia, 13–14 June 2013: 132–137.
- Benigni V. 2017. Una festa da paura! Mi sono divertito da morire! Gli intensificatori iperbolici dell'italiano e la loro resa in russo, *Studia de Cultura "Gli orizzonti dell'italianistica: tradizione, attualità e sfide di ricerca"*, IX(1): 5–18.
- Bolinger D. 1972. *Degree Words*, The Hague, Mouton de Gruyter.
- Castagnoli S., Lebani G.E., Lenci A., Masini F., Nissim M., Piunno V. 2015. Towards a corpus-based online dictionary of Italian Word Combinations Automatic Knowledge Acquisition for Lexicography, COST ENeL WG3 meeting, Herstmonceux Castle, 13 August 2015. [http://www.elexicography.eu/wp-content/uploads/2015/10/Paper\\_Castagnoli\\_et\\_al\\_ENeL\\_final.pdf](http://www.elexicography.eu/wp-content/uploads/2015/10/Paper_Castagnoli_et_al_ENeL_final.pdf)
- Castagnoli S., Lebani G.E., Lenci A., Masini F., Nissim M., Passaro L.C. 2016. POS-patterns or Syntax? Comparing methods for extracting Word Combinations, [in:] *Computerised and Corpus-based Approaches to Phraseology: Monolingual and Multilingual Perspectives (Full papers) – Fraseología computacional y basada en corpus: perspectivas monolingües y multilingües (Trabajos completos)*, G. Corpas Pastor, Geneva, Switzerland, Tradulex: 101–114.
- Labov W. 1984. *Intensity, [in:] Meaning, Form, and Use in Context: Linguistic Applications*. D. Schiffrin, Washington D.C., Georgetown University Press: 43–70.
- Masini F. 2009. *Combinazioni di parole e parole sintagmatiche, [in:] Spazi linguistici. Studi in onore di Raffaele Simone*, E. Lombardi Vallauri & L. Mereu, Roma, Bulzoni: 191–209.
- Masini F. 2012. *Parole sintagmatiche in italiano*, Cesena–Roma, Caissa Italia.
- Passaro L.C., Lenci A. 2016. *Extracting Terms with EXTra*, Paper presented at EUROPHRAS 2015 – Computerised and Corpus-based Approaches to Phraseology: Monolingual and Multilingual Perspectives. Malaga, Spain, 29.06.–1.07.2015.
- Piunno V. 2013. *Modificatori sintagmatici con funzione aggettivale e avverbiale*, Tesi di dottorato, Roma, Università degli Studi Roma Tre.
- Piunno V. 2015. *Sintagmi Preposizionali come Costruzioni Aggettivali*, *Studi e Saggi di Linguistica*, 53, 1, s. 65–98.
- Piunno V. 2016a. *Multiword Modifiers in Romance languages. Semantic formats and syntactic templates*, *Yearbook of Phraseology*, Berlin: Mouton de Gruyter.
- Piunno V. 2016b. *CombiNet. A Corpus-based Online Database of Italian Word Combinations*, *GLOBALLEX 2016: Lexicographic Resources for Human Language Technology*, LREC2016, Portorož, 24 maggio 2016.



- Piunno V. 2017. Combinazioni di parole parzialmente riempite, Formati e rappresentazione lessicografica, Studia de Cultura “Gli orizzonti dell’italianistica: tradizione, attualità e sfide di ricerca”, IX(3): 103–113.
- Piunno V. 2017. Negated Multiword Expressions. Types, properties and degrees of lexicalization, [in:] *Konventionalisierung und Variation / Conventionalization and variation*, N. Filatkina, – S. Stumpf, Frankfurt am Main, Peter Lang.
- Piunno V., Simone R., Mereu L., Pompei A. 2013. Tipi di combinazioni di parole, Rapporto tecnico, Roma, TRIPLE-Università Roma Tre.
- Simone R. 2006. Nominales sintagmáticos y no-sintagmáticos, [in:] *Estructuras léxicas y estructuras del léxico*, E. De Miguel *et al.*, Frankfurt am Main, Peter Lang, s. 225–246.
- Simone R. – Masini F. 2007. Support nouns and verbal features: a case study from Italian, [in:] *Verbes et classes sémantiques*, A. Grezka – F. Martin-Berthet, Special issue of *Verbum*.
- Simone R. – Masini F. 2014. *Word Classes*, Amsterdam & Philadelphia, John Benjamins.
- Simone R., Piunno V., Cominetti F. 2015. Le categorie lessicali preesistono ai dati o originano dai dati?, Relazione presentata al XLIX Congresso internazionale di Studi della SLI: Tipologia e ‘dintorni’. Il metodo tipologico alla intersezione di piani d’analisi, Malta 24–26 settembre 2015.
- Voghera M. 1994. Lessemi complessi: percorsi di lessicalizzazione a confronto, *Lingua e Stile*, 29. 2: 185–214.
- Voghera M. 2004. Polirematiche, [in:] *La formazione delle parole in italiano*, M. Grossmann, – F. Rainer, Tübingen, Max Niemeyer Verlag: 56–69.

### **Formati di parola inediti: dati emersi dal dizionario *CombiNet***

Questo contributo si propone di analizzare alcuni “formati di parola inediti” della lingua italiana, emersi dalla ricerca su corpora finalizzata all’elaborazione del dizionario *CombiNet*. L’analisi dei dati estratti dai corpora per la compilazione delle voci *CombiNet* ha infatti rivelato l’esistenza nel lessico dell’italiano di numerose strutture sintagmatiche e di “formati di parola”, non descritti nella letteratura. Questo contributo intende presentare i formati di parola inediti dell’italiano, discutendone alcune specifiche proprietà (es. la relazione tra i tipi sintagmatici e la funzione “alta” da essi svolta nella frase).

**Parole chiave:** combinazioni di parole, formati di parola, configurazioni sintagmatiche e semantiche, dizionario *CombiNet*

### **Innovative word formats. Data from *CombiNet* dictionary**

This paper aims at analyzing some “word formats” of Italian that emerged from a corpus-based research for the creation of *CombiNet* dictionary. The analysis of such data has revealed the existence in the Italian lexicon of many syntagmatic structures and word formats which have not yet been described in the literature. This contribution intends to present particular Italian word formats and to discuss their specific properties (e.g. the relationship between combinatorial profiles and the function they held in the sentence).

**Keywords:** word combinations, word formats, syntagmatic and semantic patterns, *CombiNet* dictionary

**Formacje nieskategoryzowane według danych ze słownika komputerowego *CombiNet***

Niniejszy artykuł analizuje „formacje nieskategoryzowane” w języku włoskim na podstawie badań korpusowych przeprowadzonych w celu opracowania słownika *CombiNet*. Analiza danych, które posłużyły do uzupełnienia haseł *CombiNet*, wykazała istnienie w słownictwie języka włoskiego licznych struktur syntagmatycznych i metod słowotwórczych nieopisanych dotąd w literaturze naukowej. Niniejszy artykuł ma na celu przedstawienie sposobów powstawania formacji nieskategoryzowanych oraz analizę ich wybranych cech (np. związku typów syntagmatycznych z ich funkcją w zdaniu).

**Słowa kluczowe:** kombinacje słów, formacje nieskategoryzowane, związki syntagmatyczne i semantyczne, słownik *CombiNet*

**Federica Cominetti** è assegnista di ricerca presso l'Università di Firenze, partecipa al progetto di analisi dell'italiano parlato Lablita. Interessi di ricerca: linguistica dei corpora, analisi delle categorie grammaticali in prospettiva cognitiva e comparativa, acquisizione e didattica delle L2, applicazione di modelli computazionali alla lessicologia. Nei suoi lavori riporta spesso dati dal cinese, lingua che la appassiona da molti anni.

**Valentina Piunno** – è assegnista di ricerca presso l'Università degli Studi Roma Tre. Ha conseguito il dottorato in Linguistica Sincronica, diacronica e applicata nel 2013 presso l'Università degli Studi Roma Tre, discutendo una tesi sui “Modificatori sintagmatici con funzione aggettivale e avverbiale”. Ha collaborato al progetto “CombiNet”, occupandosi della creazione di un *tool* lessicografico per la rappresentazione delle combinazioni di parole. I suoi interessi di ricerca si rivolgono principalmente allo studio del lessico (italiano, spagnolo e francese) e alla combinatoria lessicale, alla lessicografia combinatoria, alla semantica lessicale, alla linguistica dei corpora.