*Francesca Masini & Jacopo Di Donato*

# Non-prototypicality by (discontinuous) reduplication: The N-*non*-N construction in Italian

**Abstract**: This paper contributes to the research on the morphological expression of approximation by analysing the discontinuous reduplication pattern N-*non*-N in Italian, giving rise to complex nominals (e.g., *sapone non sapone*, lit. soap NEG soap, 'non-soap/soap-free cleanser'). The analysis is based on a dataset of 4609 tokens and 692 types extracted from corpora and annotated for orthographical, phonological and categorical parameters. Given its unpredictable formal and semantic properties, N-*non*-N is analysed as a semi-schematic and productive construction, which is hypothesized to have emerged microdiachronically from the entrenchment and reanalysis of the expression *tessuto non tessuto* 'nonwoven fabric'. We claim that the N-*non*-N construction bears a general function of 'non-prototypicality' and produces expressions with complex and specific meanings that are generated in context, by deviating from the 'prototypical' N concept in various ways. Despite its productivity, very few N-*non*-N expressions are conventionalised, making this device similar to other approximative strategies and to evaluative morphology in general.

**Keywords**: evaluative morphology, approximation, non-prototypicality, discontinuous reduplication, negation, Construction Morphology, Italian

## 1. Introduction[*]

In this paper, we contribute to the research on the morphological expression of approximation by analysing a discontinuous reduplication pattern in Italian that we claim to convey a value of 'non-prototypicality'.

Reduplication is known for having a variety of functions cross-linguistically, including approximation (cf., among others, Moravcsik 1978; Inkelas 2014; Mattiola & Barotto 2023). Some examples (retrieved through the ListTyp database) follow:

---

(1) a. *leʼleng* 'black' > *leʼleng-leʼleng* 'blackish' (Makasar; Jukes 2006: 105)

b. *ĭŋgwē* 'white' > *ĭŋgwē ĭŋgwē* 'whiteish, rather white' (Maʼdi; Blackings & Fabb 2003: 105)

Italian is generally considered as a language with no productive reduplication (cf. WALS: https://wals.info). However, some studies, either implying or advocating for a unified approach to reduplication and repetition, show that reduplication is well attested in this language, performing a variety of functions[1] (for an overview, see Mauri & Masini 2022, who talk about 'replication'). This is true of both 'canonical' and 'non-canonical' reduplication (Stolz 2018), including so-called 'discontinuous reduplication' (Mattiola & Masini 2022), which may be defined as a kind of reduplication "where other morphological material may appear between the reduplicant and the base" (Velupillai 2012: 101).

The pattern under scrutiny here is precisely a case of 'discontinuous reduplication' displaying a reduplicated noun N with an interposing negation adverb (*non*) that gives rise to a complex nominal:[2]

(2) a. *sapone    non    sapone*
    soap    NEG    soap
    'non-soap/soap-free cleanser, synthetic detergents (syndets)'

b. *Per una maggiore azione dermoprotettiva, si consiglia il '**sapone non sapone**' perché privo di sostanze alcaline che possono causare irritazioni e inaridimento cutanei.* [itTenTen16]

'For a better skin protection, we recommend a **soap-free cleanser**, as it is free of alkaline chemicals, which can cause skin irritation and dryness'

(3) a. *bomboniera    non    bomboniera*
    party_favour    NEG    party_favour
    'alternative/unconventional party favour'

---

[1] Since we are talking here about 'non-prototypicality', it is worth mentioning that nominal (continuous) reduplication in Italian may convey quite the opposite value, namely prototypicality/authenticity (e.g., *caffè caffè*, lit. coffee coffee, 'genuine coffee') or contrastive focus, as Gomeshi et al. (2004) call it (see Mauri & Masini 2022: 132–134 for a brief discussion and examples).

[2] Italian *non* mainly corresponds to English *not*, but it also appears in formations like *non credente* 'non-believer' or *nonsenso* 'nonsense', where it corresponds to English *non*. For simplicity, we decided to gloss *non* as NEG throughout the paper.

    b.   *Ci sono però anche delle proposte alternative e tra queste, ad esempio, le "**bomboniere non bomboniere**" costituite da oggettini più particolari e originali come il prodotto biologico, il barattolino di miele, il prodotto tipico, ecc. per chi ama l'anticonformismo "naturale".* [itTenTen16]

        'There are, however, alternatives, like, for example, **unconventional party favours**, made of tinier, more unique and original items such as an organic product, a small honey jar, a regional product etc., for those who love a "natural" non-conformism'

The meaning denoted by the N-*non*-N expressions in (2)–(3) can be holistically schematized as 'non-prototypical N'. *Sapone non sapone* in (2) is a cleanser/detergent (thus serving the same function as soap) which is however not a 'proper' soap in that it is not made through saponification (i.e., the combination of natural oils/fats and an alkaline product). *Bomboniere non bomboniere* (3), instead, refers to party favours that depart from canonical objects (typically containing sugared almonds) usually given at weddings and other ceremonies.

We claim that this N-*non*-N pattern – which was identified by Mattiola & Masini (2022) but still lacks a proper description in the literature – has become a specific device used to convey 'non-prototypicality' in Italian.

The paper is organized as follows. Section 2 describes the corpus-based methodology used to study the N-*non*-N pattern, including annotation parameters (orthographical, phonological, categorial). Section 3 illustrates the results in quantitative terms. In Section 4 we discuss the results and we provide a theoretical account in terms of Construction Grammar (Goldberg 1995, 2006) and Construction Morphology (Booij 2010), not without speculating on the why and how this construction emerged. Section 5 contains some concluding remarks.

## 2. Methodology

In this paper, we offer an in-depth analysis of the N-*non*-N pattern based on naturally occurring data from *CORIS* (http://corpora.dslo.unibo.it/TCORIS) and *itTenTen16* (accessed through the SketchEngine: https://www.sketchengine.eu), two corpora of written contemporary Italian. The data were extracted automatically through a query and then manually revised.

The final dataset contains 4609 tokens and 692 types. The dataset was annotated with the following parameters in order to identify possible constraints and tendencies:

- **Reduplicated form**: the exact reduplicated word form.
- **Reduplicated lemma**: the lemmatized version of the reduplicated form.

ORTHOGRAPHICAL AND PHONOLOGICAL PROPERTIES

- **Hyphen [Orthography]**: the presence ('yes') or absence ('no') of two hyphens separating the three components of the pattern.[3] We also annotated as 'hybrid' the cases that display only one hyphen.
- **Scare quotes [Orthography]**: the presence ('yes') or absence ('no') of scare quotes surrounding the pattern. We did not consider as scare quotes quotation marks signalling true quotations or proper names, such as titles of books, programs, etc.
- **Syllables**: the number of syllables of the reduplicated item.

CATEGORIAL PROPERTIES

- **Category of reduplicated form**: we searched for patterns including items tagged as nouns in the two corpora, and we indeed found mostly nouns ('N'), but a minority of occurrences turned out to be verbs (e.g., *vedo non vedo*, lit. I.see NEG I.see), which were annotated accordingly as 'V' plus TAM information, resulting in two tags ('V_1SG.PRES.IND' and 'V_2SG.PRES.IND'). The output category of the whole construction, instead, is noun for all the examples.[4]
- **Ambiguity**: a number of occurrences turned out to display categorial ambiguity, since in these cases the reduplicated element (the copy) is interpretable not just as a noun but also as a modifier: either an adjective (e.g., *dolce non dolce*, lit. sweet$_N$ NEG sweet$_{N/ADJ}$, where the first *dolce* is a noun but the second *dolce* might be an adjective), a past participle (e.g., *bollito non bollito*, lit. boiled_meat$_N$ NEG boiled_meat$_N$/boiled$_{ADJ}$) or a present participle (e.g., *cantanti non cantanti* lit. singers$_N$ NEG singers$_N$/singing$_{ADJ}$). We marked these cases as

---

[3] The corpora used for data extraction treated hyphens differently in terms of tokenization. We took this difference into consideration and elaborated different queries in order to capture examples with hyphens in both corpora.

[4] An anonymous reviewer raised the question whether a V-*non*-V pattern yielding verbs (instead of nouns) exists in Italian. To the best of our knowledge, the answer is no.

ambiguous ('yes'), whereas those that unequivocally display a noun in both positions were annotated as unambiguous ('no').

– **Category of copy**: when the example is ambiguous (see previous parameter), we annotated the other possible category of the copy, hence 'ADJ' (adjective), 'PART_PAST' (past participle) or 'PART_PRES' (present participle).

## 3. Results

### 3.1 Lexical distribution and productivity

As already said, the final dataset consists in 4609 tokens for 692 types. The lexical distribution is uneven since we have a very high number of types with low frequency and very few cases with an extremely high token frequency: *tessuto non tessuto* 'nonwoven fabric' [lit. fabric NEG fabric/woven] occurs 1966 times, amounting to approx. 42% of the total dataset, followed by *vedo non vedo* 'translucent/see-through effect' [lit. see.1SG.PRES NEG see.1SG.PRES] with 508 tokens, and *luogo non luogo* 'non(-anthropological) place' [place NEG place] with 193 tokens. See Fig. 1, displaying a classic Zipfian distribution for types with fq. 10 or higher.
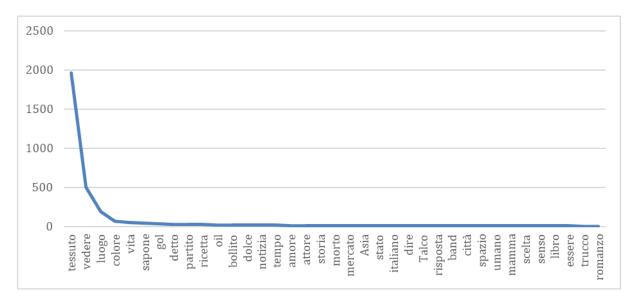


**Fig. 1**: Lexical distribution of types with fq. 10 or higher

Hapaxes amount to 402 (approx. 58% of the total types), speaking in favour of the productivity of the construction, which is indeed able to generate new expressions.

Productivity measures like type/token ratio (TTR = 0.149) and potential productivity (PP = 0.087)[5] are not very meaningful taken in isolation, but they will be applied to different subsets of our data in Section 4.4. They might also suffer from the unbalance caused by the few very frequent items mentioned above – *tessuto non tessuto* (4a), *vedo non vedo* (4b) and (secondarily) *luogo non luogo* (4c) – which are clearly the 'leaders' in our dataset.

(4) a. *Questo ad evitare, nel tempo, possibili intasamenti. Sopra le pietre va posto uno strato di **tessuto non tessuto**, che ha il compito di impedire il passaggio di terra o di materiale organico, che altrimenti intaserebbe il drenaggio stesso* [itTenTen16]

'This is to avoid, over time, possible clogging. A layer of **unwoven fabric** should be placed over the stones, which has the task of preventing the passage of soil or organic material, which would otherwise clog the drainage itself'

b. *I body in pizzo giocano con un intrigante **vedo non vedo**; i più romantici coordinati, arricchiti da raffinati dettagli, sono sempre un must.* [itTenTen16]

'The lace bodysuits play with an intriguing **see-through effect**; the most romantic coordinates, enriched by refined details, are always a must.'

c. *La storia è semplice, narrata attraverso un lungo flashback da un 'cantastorie' mentre fa la fila in un "**luogo non luogo**" contemporaneo come l'ufficio postale.* [itTenTen16]

'The story is simple, told through a long flashback by a 'storyteller' as he lines up in a contemporary "**non(-anthropological) place**" like the post office.'

It is interesting to note that these three expressions are representative of the three types we identified in terms of categories in Section 2: *tessuto non tessuto* is an 'ambiguous' case, since *tessuto* means both 'fabric$_N$' and 'woven' (past participle of *tessere* 'to weave'), hence the whole expression might be interpreted either as a 'true' reduplication mediated by *non*, or as a noun followed by a negated modifier ('unwoven fabric') (cf. Sections 3.3 and 4.4); *vedo non vedo* contains a reduplicated verbal form instead of a noun; and *luogo non luogo* would be an example of the truly reduplicative construction we aim to investigate here, where the second *luogo* is unequivocally a noun.

---

[5] For a discussion and illustration of type/token ratio (TTR) and potential productivity (PP) as productivity scores see Baayen & Lieber (1991) and Baayen (2009).

3.2 Orthographical and phonological properties

As for orthographic properties, we annotated the presence of hyphens within the pattern and scare quotes around it, in order to understand if orthography plays any role in marking certain kinds of expressions.

The vast majority of examples in our dataset does not display hyphens (81%; see Fig. 2) and is not enclosed in scare quotes (88%; see Fig. 3). Hybrid cases are clearly very marginal: in the vast majority of cases, the single hyphen appears before *non*, whereas very few examples display the hyphen after *non* (95 vs. 8 occurrences).

However, the percentage of hyphens and scare quotes increases if we consider only hapaxes, which are more likely to be new formations with a supposedly more creative semantics: approximately half of the hapaxes display at least one hyphen ('yes' or 'hybrid') and approx. 32% of the hapaxes are enclosed in scare quotes.
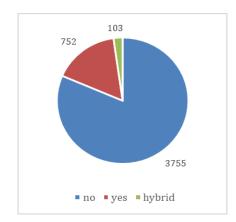


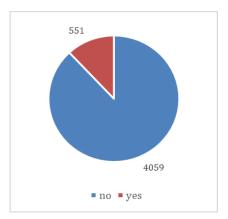**Fig. 2**: Presence of hyphen(s) based on tokens



**Fig. 3**: Presence of scare quotes based on tokens

Note that one and the same item may occur with and without hyphens (cf. (5a) vs (5b)), with and without scare quotes (cf. (6a) vs (6b), where the two instances refer exactly to the same referent).

(5)    a.    *Poi l'attrice ammette: Sì, potrei lavorare ancora con lui, perché Claudio è un **attore-non-attore**, sovverte continuamente le regole, ha un modo anarchico di intendere il mestiere.* [itTenTen16]

          'Then the actress admits: Yes, I could still work with him, because Claudio is an **unconventional actor** [lit. **actor-NEG-actor**], he constantly subverts the rules, he has an anarchic way of understanding the job.'

b. […] *non c'era certo bisogno del film di Antonio Albanese, non avevamo neanche bisogno di **attori non attori**, quindi poco professionisti come ormai da troppo tempo si hanno sugli schermi italiani* [itTenTen16]

'[…] there was certainly no need for Antonio Albanese's film, we didn't even need **non-actors** [lit. **actors** NEG **actors**], namely unprofessional actors as we have had on Italian screens for too long now'

(6)  a. *Interverranno lo chef Massimo Bottura, presidente del comitato dei Maestri di Cucina e ideatore di metodi di cottura innovativi, di particolare interesse per il mondo scientifico (famoso il suo "**bollito non bollito**" e l'uso in cucina di macchinari per la distillazione sottovuoto per concentrare le essenze)* […]. [itTenTen16]

'Speakers include chef Massimo Bottura, president of the Committee of the Masters of Cuisine and creator of innovative cooking methods, of particular interest to the scientific world (famous is his "**unboiled boiled meat**" [lit. boiled_meat NEG boiled_(meat)] and the use in the kitchen of machinery for vacuum distillation to concentrate the essences) […]' [itTenTen16]

b. *Il **bollito non bollito** di Massimo Bottura, ad esempio, pur essendo la quintessenza della semplicità […] richiede l'uso di procedimenti non proprio alla portata di tutti.* [itTenTen16]

'Massimo Bottura's **unboiled boiled meat** [lit. boiled_meat NEG boiled_(meat), for example, while being the quintessence of simplicity […] requires the use of procedures not really within everyone's reach.'

Another feature we checked is the number of syllables of the reduplicated form: the syllabic structure may shed light on instantiation preferences, based on the phonological weight of the word to be reduplicated. Our data show that trisyllabic and bisyllabic words are definitely more used in this construction (93% of the tokens, 78% of the types). Much less attested are monosyllabic words (mainly borrowings from English like *gol non gol* or *film non film*) and longer bases of 4 (e.g., *gravidanza-non-gravidanza*, lit. pregnancy NEG pregnancy), 5 (e.g., *appuntamento non appuntamento*, lit. date NEG date) and 6 (e.g., *ideologia-non-ideologia*, lit. ideology NEG ideology) syllables (see Fig. 4).
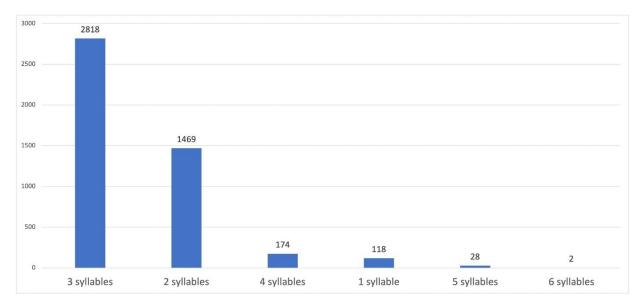
**Fig. 4**: Number of syllables

The preference for short bases may just reflect a general tendency in the Italian lexicon, so we confronted our data with those that can be retrieved from the NVdB (*Nuovo Vocabolario di Base*)[6]. The sum of trisyllabic and bisyllabic nouns in the NVdB equals 46%. The difference with our dataset is quite marked, so there is some ground to state that the N-*non*-N construction has indeed a preference for bi/trisyllabic, hence 'short' words. This may have something to do with the prosody associated with the construction, an aspect we cannot investigate here given that we are dealing with written languages, but that sounds promising.

## 3.3 Categorial properties

As anticipated in Section 2, reduplicated items are normally nouns (and noun is also the output category of all our examples), but in some cases (11% of tokens, 1% of types) we retrieved verbal forms (obviously due to wrong POS tagging). In the latter case, basically all examples are 1st person singular indicative forms (like *vedo non vedo*, cf. (4b)). As regards the two instances of V_2SG.PRES.IND (*vedi non vedi*, lit. see.2SG.PRES NEG see.2SG.PRES), they are probably (meant to be) variants of *vedo non vedo*, or simply typos. See the figures in Tab. 1.

---

[6] We are indebted to Claudio Iacobini, who kindly gave us his personal annotated version of NVdB, that made the counting easy and fast.

**Tab. 1**: Category of the reduplicated form

| Category | Subcategory | Tokens | | Types | |
|---|---|---|---|---|---|
| N | | | 4093 (89%) | | 685 (99%) |
| V | V_1SG.PRES.IND | 514 | 516 (11%) | 6 | 7 (1%) |
| | V_2SG.PRES.IND | 2 | | 1 | |
| Total | | | 4609 | | 692 |

Half of the tokens display categorial ambiguity, meaning that the copy is interpretable either as a noun, identical to the first noun (thus yielding a N-*non*-N reduplicative structure), or as an adjective/participle that modifies a head noun (thus yielding a modifying structure rather than a reduplicative one). The high percentage of ambiguous cases is due to the very high frequency of *tessuto non tessuto*, which, as already observed in Section 3.1, is an ambiguous case, like *etica non etica* (see also (7b) below):

| REDUPLICATIVE STRUCTURE | | | | | MODIFYING STRUCTURE | | |
|---|---|---|---|---|---|---|---|
| *tessuto*$_N$ | *non* | *tessuto*$_N$ | | vs. | *tessuto*$_N$ | *non* | *tessuto*$_{PART\_PAST}$ |
| fabric | NEG | fabric | | | fabric | NEG | woven |
| *etica*$_N$ | *non* | *etica*$_N$ | | | *etica*$_N$ | *non* | *etica*$_{ADJ}$ |
| ethics | NEG | ethics | | | ethics | NEG | ethic(al) |

**Fig. 5**: Reduplicative structure vs. modifying structure

However, if one turns to types, the impact of ambiguous cases lowers considerably to only 21% of the cases, as illustrated in Tab. 2.

**Tab. 2**: Detail of ambiguous cases

| Ambiguity | Subcategory | Tokens | | Types | |
|---|---|---|---|---|---|
| no | | | 2225 (48%) | | 545 (79%) |
| yes | ADJ | 256 | 2384 (52%) | 110 | 147 (21%) |
| | PART_PAST | 2111 | | 29 | |
| | PART_PRES | 17 | | 8 | |
| Total | | | 4609 | | 692 |

When the copy is unambiguous, it means that the reduplicated item is unequivocally a noun. When the copy is ambiguous, instead, the word can be a noun but also an adjective/participle. As Tab. 2 shows, if we look at tokens, we have a predominance of past participles (7a) (again, because of *tessuto non tessuto*), whereas if we consider types, in most cases we have a N/ADJ ambiguity (7b). Present participles, instead, are less frequent (7c).

(7)  a.  *Tisbe: quali sono le tue scelte faticose, scomode e costose? Non si capiscono dal tuo post, anzi, sembra che la tua sia quasi una* **scelta non scelta**, *tant'è che dici che ci sei nata così...* [itTenTen16]

'Tisbe: what are your strenuous, uncomfortable and expensive choices? You can't tell from your post, in fact, it sounds like yours is almost a '**non-choice**' [lit. **choice NEG choice/chosen**], so much so that you say you were born with it...'

b.  *Berlusconi ha una sua coerenza e trasparenza. Può non piacere o disgustare. Ha una sua* **etica non etica**, *però tutti la conosciamo. Io temo più quelli che non conosco che quelli che conosco bene.* [itTenTen16]

'Berlusconi has his own consistency and transparency. One may dislike him or feel disgust towards him. He has his own **unethical/'fake' ethics** [lit. **ethics NEG ethics/ethical**], but we all know it. I fear more those I do not know than those I know well.'

c.  [...] *quando parlavo con i ragazzi ad "X Factor" dicevo sempre che la verità non è stonata.* [...] *L'emozione non stona... Quanti* **cantanti non cantanti** *abbiamo che ci fanno venire la pelle d'oca?"* [itTenTen16]

'[...] when I was talking to the guys at "X Factor" I always said that truth is not out of tune. [...] Emotion is not out of tune... How many **not-proper-singers** [lit. **singers NEG singers/singing.PL**] do we have that give us goosebumps?"'

## 4. Analysis and discussion

The quantitative data illustrated in Section 3 show that the N-*non*-N pattern is a productive construction in contemporary Italian that is able to create new expressions. The lexical distribution of our dataset points to a situation where we have very few expressions that are more or less established in the lexicon (especially *tessuto non tessuto*, which is basically the only one recorded in dictionaries, cf. Section 4.2) and a very large array of new or occasional expressions formed according to this template, including proper names (like *Battiato non Battiato*, the name of a 1996 tribute album for singer-songwriter Franco Battiato).

N-*non*-N is employed typically with 'short' nouns (bi/trisyllabic). Even though we did not annotate semantic properties of nouns, we can report that the pattern can host a variety of semantic classes of nouns: both concrete (e.g., *libro non libro*, lit. book NEG book) and abstract (e.g., *bellezza non bellezza*, lit. beauty NEG beauty), both count (e.g., *preti non preti*, lit. priests NEG priests) and mass (e.g., *latte non latte*, lit. milk NEG milk). Sometimes hyphens are used to link the constituents, and the whole expressions are sometimes "scare-quoted".

The item that is reduplicated or copied is a noun in 99% of the cases in terms of types (coherently with our query on corpora), but we also found examples with a verbal form, mostly belonging to the *vedo non vedo* type. Both cases are 'not strictly predictable' in terms of formal properties and semantics. In the following we will focus on the nominal cases, which are the core of our present investigation.[7]

## 4.1 Unpredictable formal properties

As regards form, in Italian nouns are not found in this pattern in 'normal' syntax: we do find nouns followed by negation and another noun, but either (i) in juxtapositions with a contrastive function (like in (8a–b)), or (ii) in asyndetic coordination with additive function, like in (8c).[8] In both cases, the two parts are typically mediated by a comma before *non* in written language (which would plausibly correspond to a pause in spoken language), and most importantly identify two different referents.

(8)  a.  *Fa l'attore, non il cantante*
         'He's an actor, not a singer'

     b.  *Producono macchine, non biciclette*
         'They produce cars, not bikes'

     c.  *La manifestazione è aperta a scacchisti, non scacchisti, grandi e piccini.*
         'The event is open to chess players, non-chess players, young and old alike.'

---

[7] The verbal ones are even less predictable, by the way, since two identical verbal forms give rise to a complex nominal, thus displaying exocentricity.

[8] During the manual revision of the data automatically extracted from corpora, we identified instances of what seem to be other (peculiar) constructions. One is exemplified by (i), which is not approximation but rather a concise way to express that the 'question' is about using or not using béchamel (another way to express the same concept would be *besciamella sì besciamella no*, lit. béchamel yes béchamel no). Another one is exemplified by (ii), where *allenamento non allenamento* is detached from the main clause and emphasizes that the content of the main clause itself ('tomorrow evening he must be in Bologna') cannot be changed, *allenamento* 'training' being irrelevant (this construction would however sound better with a disjunction: *allenamento o non allenamento*, lit. training or NEG training).

(i)  *Sulla questione **besciamella non besciamella**, io ce la metto ma io poi ovviamente non faccio testo* [itTenTen16]
     'About the question '**should-we-use-béchamel-or-not**' [lit. **béchamel NEG béchamel**], I use it, but my opinion is not worth tuppence'

(ii) [...] ***allenamento non allenamento** lui domani sera anche sul tardi deve essere a Bologna* [...] [itTenTen16]
     '**training or not** [lit. **training NEG training**], tomorrow evening, even late (evening), he must be in Bologna'

Differently from these cases, our N-*non*-N pattern forms a whole unit (sometimes marked by the use of hyphens or scare quotes, to keep the multiword expression together) with a unique referent, there is no pause or comma within the expression, and, most importantly, the second noun *must* be identical to the first. The latter property keeps it apart especially from cases like (8a–b). Indeed, contrastive juxtaposition with identical nouns would result in a contradictory, inappropriate utterance (#*Producono macchine, non macchine* 'They produce cars, not cars'), unless some (prosodic or gestural) emphasis is added to one of the two nouns, thus making them different, not identical items (e.g., *Producono MACCHINE, non macchine*).

Another peculiar formal property is that the second noun in the N-*non*-N pattern must be bare (e.g., *sapone non sapone* vs. \**sapone non il/un/del sapone*, lit. soap NEG the/a/some soap), directly attached to *non*, even when the first noun is not: this would not be the case in contrastive juxtapositions, where the definiteness of the second noun mirrors the definiteness of the first noun (cf. \**Fa l'attore, non Ø cantante*, with no article *il* 'the' before *cantante* 'singer').

Finally, the N-*non*-N sequence cannot be interrupted, not only by determiners, as just illustrated, but by any other element (e.g., *sapone non sapone* vs. \**sapone profumato non sapone*, lit. soap scented NEG soap).

## 4.2 Unpredictable semantic properties

As regards semantics, N-*non*-N expressions are also unpredictable because they convey approximation of a concept. What is approximated is the meaning of N, whose core properties are called into questions. Let us take the following examples.

(9)   a.   *Casa nuova, vita nuova e quind… [sic] inviti a cena! Ovviamente sono impazientissima di tornare a cucinare e fare un po' di cene serie, nel mentre però, quando ci vuole ci vuole, e così ho fatto una* **cena non cena**, *insomma un invito a pane e salame, che poi proprio proprio pane e salami* [sic] *non è stato.* [itTenTen16]

'New house, new life and therefore…. dinner invitations! Obviously I'm eager to get back to cooking and make some serious dinners, but in the meantime, when you need it you need it, so I made a 'non-dinner' [lit. **dinner NEG dinner**], in short, a bread-and-salami invitation, which was not so much bread-and-salami after all.'

b. *Zarazà: tradizione e modernità si mescolano per un mix affascinante e imperdibile. Uno dei piatti tipici è la famosa "**carbonara non carbonara**".* [itTenTen16]

'Zarazà: tradition and modernity intertwine for a fascinating and unmissable mix. One of the typical dishes is the famous "**carbonara NEG carbonara**".'

c. *Il riso si cuocerà nel tempo indicato sulla confezione, generalmente richiede 40 minuti di cottura. Il metodo della **cottura non cottura** è utile per sfruttare il tempo impegnato altrove per cuocere il riso.* [itTenTen16]

'The rice will cook in the time indicated on the package, generally requiring 40 minutes of cooking time. The 'non-cooking' [lit. **cooking NEG cooking**] method is useful for using the time committed elsewhere to cook the rice.'

d. *La legge che prometteva la chiusura definitiva degli ospedali psichiatrici giudiziari, quella specie di '**carcere non carcere**' per i malati di mente che si fossero macchiati di un crimine, è entrata definitivamente in vigore lo scorso 31 marzo.* [CORIS]

'The law that promised the definitive closure of forensic psychiatric hospitals, that sort of 'non-jail' [lit. **jail NEG jail**] for the mentally ill who are guilty of a crime, finally came into force on March 31.'

e. *Noi ci mettiamo il parco, la voglia e la medaglia. Voi i polpacci. Torna la **corsa non corsa**, che fa contenti tutti: mamme, papà, bambini, musicisti e scansafatiche. 6 chilometri di festa che altro non sono che una scusa per passare una domenica nel parco e ricordarsi che se si suda tutti assieme ci si sente meglio.* [itTenTen16]

'We'll throw in the park, the desire and the medal. You put your calves. Back to the 'non-running race' [lit. **race NEG race**], which makes everyone happy: moms, dads, kids, musicians and slackers. 6 kilometers of celebration that are nothing but an excuse to spend a Sunday in the park and remember that if you sweat all together you feel better.'

In (9a), a *cena non cena* (lit. dinner NEG dinner) identifies a kind of dinner which is not a full-fledged dinner, not apt for inviting people at home, which is an occasion that requires a 'proper' dinner, not a bread-and-salami frugal meal. *Carbonara non carbonara* in (9b) is the name of a special dish included in the menu of the Zarazà restaurant in Frascati: it's a 'carbonara' (a typical Roman recipe for pasta with eggs and bacon) with no eggs, so in this case a key ingredient is missing. In (9c) we find another food-related example: *cottura non cottura* here is an 'alternative' cooking method for rice that consists in letting it boil for a few minutes, turning off the flame (without uncovering the pot) and then letting it quietly cook by itself. In (9d) *carcere non carcere* is the expression the writer chooses for defining forensic psychiatric hospitals, comparing their status to that of a jail while at the same time

recognizing they are not officially jails, as also hinted at by the approximator *specie di* 'sort of' that introduces the expression (a 'redundancy' that we find in a few other examples). Finally, the event described in (9e) (the 'Magnolia Run' in Milan) is not a proper race since the true, ultimate goal is doing physical activity and spending quality time together rather than competing and winning.

As is clear from these examples, the semantics emerging from these expressions can be quite complex and specific, and is not the result of the sum of its parts: a literal reading of the parts would actually produce a contradiction; this triggers an unliteral interpretation which is not always easy to predict and that heavily relies on context. This is also testified by the fact that one and the same expression (not established in the lexicon) may carry (slightly) different meanings in different contexts. Take *canzone non canzone* in the following two examples:

(10)   a.   *Alla 57° edizione del Festival di Sanremo presenta "Ti regalerò una rosa". Una **canzone non canzone** in forma di lettera lacerante e commovente, microstoria di quel microuniverso della follia che tanto lo appassiona.* [itTenTen16]

'At the 57th edition of the Sanremo Festival he presented "Ti regalerò una rosa". An unconventional song [lit. **song NEG song**] in the form of a lacerating and moving letter, a micro-story of that micro-universe of madness that fascinates him so much.'

b.   *Un susseguirsi di brani non pianificati, **canzoni-non-canzoni** dilatate e contorte, come indagini radicali sulle possibilità del suono e del linguaggio.* [itTenTen16]

'A succession of unplanned pieces, dilated and twisted 'non-songs' [lit. **songs-NEG-songs**], like radical investigations into the possibilities of sound and language.'

In (10a) the approximation concerns the fact that "Ti regalerò una rosa" is not a canonical Sanremo (the traditional Italian Music Festival) song in terms of lyrics, since it resembles a 'lacerating and moving letter'. In (10b) the same term is used to identify songs which are unplanned, dilated and twisted. So, in both cases we have some sort of non-canonical song, but the reason why it's non-canonical is different, pointing to different classes of objects.

Variation in context may be displayed also by expressions that do have a more established meaning: *colore non colore*, for instance, occurs 73 times in our dataset and in the vast majority of cases it refers to either 'white', 'black' or 'grey'. However, we also find more creative uses like (11), where the writer is reviewing a pink-like (cosmetic) illuminating powder, by describing it as very neutral and delicate, almost uncolored.

(11)    *La cialda è molto rosa, ma questo highlighter è molto meno rosa di ciò che sembra, è un **colore non colore**, stupendo, e la sua luminosità è pari quasi al caro Mary Lou citato prima.* [itTenTen16]

        'The wafer is very pink, but this highlighter is much less pink than it looks, it's a '**non-color**' [lit. color NEG color], gorgeous, and its brightness is almost equal to the dear Mary Lou mentioned earlier.'

So, the kind of semantic operation the N-*non*-N pattern performs on N is one of approximation. More specifically, we claim it conveys 'non-prototypicality' by building a concept which is a deviating or peripheral instance of the category N, for some reason to be reconstructed by the reader/interlocutor. This 'deviating reason' is not always transparent and easy to get, and it may change from context to context (like for other approximating markers). Possible deviating reasons are:[9]

(i)    the lack of one or more salient/defining properties of N (see *cereale non cereale*, lit. cereal NEG cereal, referring to food plants like *grano saraceno* 'buckwheat', which – despite the name – is not closely related to wheat and is not a cereal);[10]

(ii)    the lack of one or more salient ingredients/parts of N (see the *carbonara* without eggs example in (9b));

(iii)    the lack of a shared function/purpose with N (see the *corsa non corsa* example in (9e));

(iv)    the lack of the expected/usual processes or factors involved to bring N into existence, despite having a similar function (see *sapone non sapone*, example (2), Section 1).

These reasons are reminiscent of Pustejovsky's (1995: 76) "qualia structure", which "specifies four essential aspects of a word's meaning (or *qualia*)", inspired by Moravcsik's (1975) interpretation of Aristotle's notion of modes of explanations. Deviating reason (i) would refer to Pustejovsky's[11] Formal quale, "which distinguishes the object within a larger domain" (think of properties like shape, magnitude, color, etc.); (ii) hints at the Constitutive quale, which specifies the "relation between an object and its constituents" or parts; (iii) is reminiscent of the Telic quale, which coincides with the "[p]urpose or function of the object"; and finally (iv) calls for the Agentive quale, namely the "[f]actors involved

---

[9] The list that follows is the result of a cursory observation of our data and is not meant to be exhaustive. However, we believe it might pave the way to a more fine-grained typology of approximative values (cf. Masini, Norde & Van Goethem 2023).

[10] Interestingly, these plants are called *pseudocereals* in English.

[11] All qualia quotations in the rest of this paragraph are from Pustejovsky (1995: 85–86).

in the origin or 'bringing about' of an object" (e.g., distinguishing between natural kinds and artifacts).

The presence of one or more 'deviations' might end up challenging the status of the N-*non*-N expression as a 'legitimate' type/instance of N: is a *sapone non sapone* a kind of *sapone*? The answer is no if we consider a soap an object that is obtained through saponification only (strict reading). The answer may be yes if we intend soap, in a broader sense, as a cleaning detergent for the body (loose reading).

The item we obtain is generally created on-the-go, giving rise to an ad hoc expression, but some of the items in our dataset became somehow conventionalised, although there is basically no record of these expressions in dictionaries (see Section 4.4).

## 4.3 N-*non*-N as a construction

We believe that the 'not strictly predictable' formal and semantic properties of the N-*non*-N pattern observed in Sections 4.1 and 4.2, respectively, speak in favor of the constructional status of this pattern, whose characteristics may be captured by a semi-schematic construction (where only *non* is lexically specified) that can be sketched as follows:
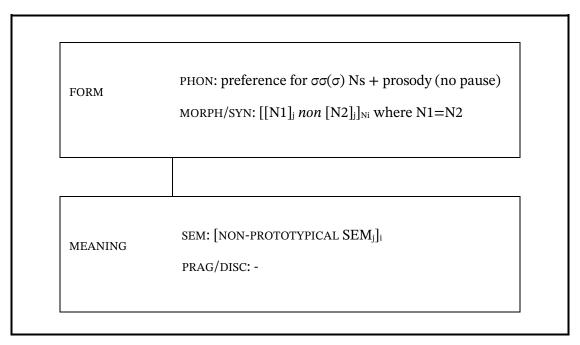
| FORM | PHON: preference for $\sigma\sigma(\sigma)$ Ns + prosody (no pause) <br><br> MORPH/SYN: $[[N1]_j$ *non* $[N2]_j]_{Ni}$ where N1=N2 |
|---|---|
| MEANING | SEM: $[\text{NON-PROTOTYPICAL SEM}_j]_i$ <br><br> PRAG/DISC: - |

**Fig. 6**: The N-*non*-N construction

Note that the semantics of the construction, at this level, is quite abstract: the exact meaning conveyed by specific constructs (i.e., instantiations of this construction) can only be generated in context and can vary from time to time. In other words, the construction

in Fig. 6 does not state 'how' exactly non-prototypical the meaning of N will be (which aspect of meaning, or 'quale', will be 'deviant' with respect to the prototype), leaving a certain degree of flexibility that can be exploited creatively by the speakers.

As sketchily represented in Fig. 7, the maximally abstract construction in Fig. 6 can be instatiated both by lexically specified constructions (i.e. the few stored N-*non*-N expressions with conventionalised semantics, e.g., *tessuto non tessuto*) and by constructs, namely new extemporaneous coinings whose exact meaning is context-generated and may also vary from case to case by virtue of the semantic flexibility of the mother construction.
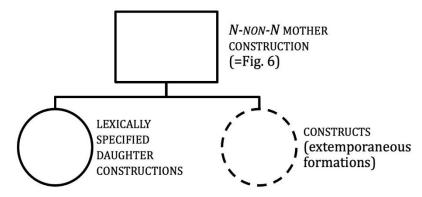


**Fig. 7**: The network of the N-*non*-N construction

Expressions such as *colore non colore*, which have a fairly stable meaning (however vague) but can also take on more extemporaneous meanings (cf. (11)), may find an explanation in the different possibilities of instantiation: whereas the semantically stable *colore non colore* may be a candidate for a lexically specified daughter construction, an extemporaneous *colore non colore* formation like (11) would be a construct, namely a direct, creative instatiation of the mother constructions.

We did not encode any genre/textual-related information into the construction because it is used in a variety of contexts. However, we may observe that it seems to be employed quite effectively to create terms for commercial products. Another (non unrelated) domain that seems to accommodate these formations quite easily is cooking and food. Both kinds of context probably rely on the expressiveness of these formations.

Finally, we decided not to encode any specific connotation. Sometimes, a pejorative connotation emerges (like for *cena non cena* in (9a) or *carcere non carcere* in (9d)), as quite expected in approximation. But in many instances there is no obvious derogatory effect: many expressions are quite neutral (like *carbonara non carbonara* in (9b)) and in some

cases they even convey, on the contrary, positive concepts (like *cottura non cottura* in (9c) or *corsa non corsa* in (9e)).

4.4 What about ambiguous cases?

As we said, a subset of nouns used in our construction (21% of types, including the very frequent *tessuto non tessuto*) are inherently ambiguous, since they can also be interpreted as adjectives/participles. This characteristic makes the analysis of these expressions uncertain: one may claim that these are not cases of reduplication because the second 'noun' is *not* a copy of the first noun but something different, namely an adjective/participle that modifies the first (head) noun. The two possible analyses may be sketched as follows (cf. also Fig. 5, Section 3.3):

(12)  a.  Reduplicative analysis:        N-*non*-N$_{RED}$
      b.  Modifying analysis:            N-*non*-ADJ/PART

According to the modifying analysis, we are in front of a much less peculiar structure (cf. Section 4.1), namely, simply, a N followed by an adjective/participle. Also from a semantic point of view, the construction would gain in predictability and interpretability. If we read *tessuto non tessuto* as unwoven fabric, it becomes much easier to compute the overall meaning: it still conveys 'non prototypicality', in a way, but we do not need to guess the 'deviating reason' because it is explicitly stated by the participle (the object is not brought into existence by weaving); so, the 'non prototypicality' effect would be a by-product. In this respect, it is worth noting that ambiguous cases have a lower percentage of examples displaying hyphens and scare quotes (interpretable as orthographic cues to signal 'special' cases; see Section 3.2) with respect to unambiguous cases: ambiguous examples with hyphens and scare quotes amount to 9.1% and 7.1% (respectively) of total ambiguous cases, whereas unambiguous examples with hyphens and scare quotes amount to 28.6% and 17.1% (respectively) of total unambiguous cases.

It is difficult to find concrete evidence pro or against one of the two analyses in (12), because the items are genuinely ambiguous and both readings may be plausible in many cases. Let us take *etica non etica* (7b) (Section 3.3), repeated here as (13):

(13)    *Berlusconi ha una sua coerenza e trasparenza. Può non piacere o disgustare. Ha una sua **etica non etica**, però tutti la conosciamo. Io temo più quelli che non conosco che quelli che conosco bene.* [itTenTen16]

'Berlusconi has his own consistency and transparency. One may dislike him or feel disgust towards him. He has his own **unethical**/'**fake**' **ethics** [lit. **ethics NEG ethics**/**ethical**], but we all know it. I fear more those I do not know than those I know well.'

We may interpret *etica non etica* as 'unethical ethics' (modifying analysis), namely an ethics that is not morally correct (as it should be), or as a 'fake' or not genuine ethics (under the reduplicative analysis), possibly some sort of opportunistic or personalistic ethics. It is not trivial to pick one interpretation over the other, because they both hint at some (form of) non-prototypical ethics.

In this respect, the case of *tessuto non tessuto* deserves a broader discussion. On the one hand, as we already said, *tessuto non tessuto* is a non-prototypical *tessuto* 'fabric' irrespective of the analysis we choose. What happens if we read the second *tessuto* as a participle is that we get immediate access to *why* it is non-prototypical (and this would be one reason to regard this analysis as preferable, in fact), whereas in the other case it remains more vague and open to interpretation. How speakers actually perceive these expressions, however, is not a trivial question: even the two authors of this paper had opposite perceptions of *tessuto non tessuto* before starting the analysis. So, behavioral data is needed to better understand this point.

On the other hand, it is quite important to clarify the status of *tessuto non tessuto* since this specific item is most likely the 'leader expression' that paved the way to the emergence of the abstract N-*non*-N construction, given that it is by far the most frequent and established, and probably the oldest attested. *Tessuto non tessuto* is recorded as a main (multiword) entry in the *Nuovo De Mauro* online dictionary[12] and as a subentry of *tessuto* in the *il Sabatini Coletti* dictionary, but no first attestation date is provided. Both *il Sabatini Coletti* dictionary and *lo Zingarelli* dictionary have a main entry for *non tessuto* (carrying the very same meaning as *tessuto non tessuto*) with 1978 as a date of attestation.[13] In the *tessuto* entry of the Treccani Enciclopedia online,[14] we find that *non tessuto* is also called *tessuto non tessuto* (abbreviated as TNT). Finally, the full string *tessuto non tessuto* is first

---

[12] https://dizionario.internazionale.it/parola/tessuto-non-tessuto (accessed 6 March 2023).

[13] Nuovo De Mauro also has an entry for *non tessuto*, which however refers to the entry for *tessuto non tessuto*.

[14] https://www.treccani.it/enciclopedia/tessuto (accessed 6 March 2023).

attested in 1989 in the diachronic newspaper corpus *Archivio La Stampa*.[15] These pieces of information seem to suggest that there are two equivalent terms: *non tessuto* and *tessuto non tessuto*. The first has an earlier attestation, but the second one is much more used nowadays: a search on *itTenTen16* revealed that the occurrences of *non tessuto* amount to around 400, versus the 1966 occurrences of *tessuto non tessuto*. So, *non tessuto* appears first but is subsequently accompanied by (and superseded, in actual use, by) *tessuto non tessuto*, although most dictionaries do not reflect its primacy.

Note that also *non tessuto* is structurally ambiguous, since it can be perceived as a negated participle ('not woven/unwoven', which is probably the original, intended reading) but also as a noun preceded by prefixal *non* (cf. Iacobini 2004: 143; cf. also Dugas 2017 for the *non*-N construction in French), along the lines of other formations in Italian: the *lo Zingarelli* dictionary, for instance, contains (as subentries) *non luogo* (lit. NEG place), based on French *non-lieu* (a term introduced by French anthropologist Marc Augé in 1992), and *non notizia* (lit. NEG news, namely not a real news, because it's not new information or because it's irrelevant or lacks real content).[16] So, the second *tessuto* may have been perceived as a participle by some speakers (driven by its 'unwoven' semantics) but as a noun by other speakers (by analogy with other formations). This ambiguity is projected onto *tessuto non tessuto*, which emerges later but soon becomes a well-established and frequent term with its own acronym *TNT* (which often appears in fabric's labels). Our hypothesis is that *tessuto non tessuto* was (re)analysed as N-*non*-N by a sufficient number of speakers to eventually lead to the entrenchment of a more abstract N-*non*-N construction that could then be applied to *any* noun. The ambiguity, however, 'persists' with those nouns that *happen* to have an homophonous adjective or particle. In sum, what was initially a structurally ambiguous modifying structure (*tessuto non tessuto*) was reanalysed and then 'went viral'.

In support of this hypothesis, we observe two synchronic facts.

The first is that, if we look at the productivity of ambiguous vs. unambiguous cases, what we get is that the truly productive configuration is the unambiguous one, namely the one

---

[15] http://www.archiviolastampa.it/ (accessed 6 March 2023).

[16] Note that our dataset contains *luogo non luogo* (lit. place NEG place) and *notizia non notizia* (lit. news NEG news), with the same meanings. The two strategies – *non*-N and N-*non*-N – are obviously closely intertwined: their relationship and competition deserves a separate study. In addition, it might be worth looking into the N-*senza*-N pattern (*senza* meaning 'without').

that calls for the N-*non*-N$_{RED}$ analysis. Table 3 shows that TTR and PP are much higher for the subset of our data including unambiguous cases only (even higher if we look at nominal cases only, thus excluding *vedo non vedo*) with respect to the whole dataset and the subset of ambiguous cases. So, synchronically speaking, the truly productive construction is the one that can be unequivocally interpreted as N-*non*-N$_{RED}$ (namely, the one schematized in Figure 6, Section 4.3).

**Tab. 3**: Productivity scores for different subsets of our data

| Productivity scores | Ambiguous cases only | Whole dataset | Unambiguous cases only | Unambiguous cases only, nouns only |
|---|---|---|---|---|
| TTR | 0.06 | 0.14 | 0.24 | 0.31 |
| PP | 0.03 | 0.08 | 0.14 | 0.18 |

The second is that, among ambiguous cases, we find examples where the ambiguity is just virtual, since the actual interpretation that we obtain is not a consequence of the adjectival/participial reading of the second item. In other words, the second item does not contribute its meaning as an adjective/participle. This is what happens in example (7c) (Section 3.3), repeated here as (14) (short version):

(14)     *Quanti **cantanti non cantanti** abbiamo che ci fanno venire la pelle d'oca?* [itTenTen16]

'How many **not-proper-singers** [lit. **singers** NEG **singers/singing.PL**] do we have that give us goosebumps?'

Here the ambiguity lies in the fact that *cantante* in Italian is both 'singer' and 'singing' (present participle). In (14), *cantanti non cantanti* identifies a special, non-prototypical category of singers, but their deviant property is not the fact that they are 'non-singing' (they do sing!). So, the only possible gloss would be "singers NEG singers".

Similar examples are offered in (15): in (15a) *agenti* means both 'agents' and 'acting.PL' (present participle) but *agenti non agenti* identifies non-prototypical agents who are such not because they are 'non-acting' but because they are abusive; in (15b), *gelati* means both 'ice creams' and 'frozen.PL' (past participle) but *gelati non gelati* identifies a category of frozen products different from proper ice creams, not unfrozen products.

(15) a. *Nella filiera del mercato immobiliare, in Italia, gli **agenti non agenti**, ovverosia quelli abusivi, generano annualmente un'evasione fiscale che la Fiaip, Federazione Italiana Agenti Immobiliari Professionali, stima in ben 450 [...].* [itTenTen16]

'In the supply chain of the real estate market in Italy, **non-agents** [lit. **agents NEG agents/acting.PL**], i.e. abusive agents, annually generate a tax evasion that Fiaip, Italian Federation of Professional Real Estate Agents, estimates at 450 [...]'

b. *Via libera insomma a grattachecche, ghiaccioli e sorbetti, tutti "**gelati non gelati**" ipocalorici.* [itTenTen16]

'Green light to *grattachecche*, ice lollies and sorbets, all low-calorie "**ice cream-like products**" [lit. **ice_creams NEG ice_creams/frozen.PL**].'

These data show that, even when the reduplicated item is ambiguous, technically speaking, the N-*non*-ADJ/PART interpretation is not necessarily activated, it may remain latent. In all these examples, it is in fact excluded, to the benefit of the competing reduplicative analysis.

These two observations seem to strengthen the idea of a truly reduplicative construction being quite entrenched and productive. What is the role of the ambiguous cases, then?

One may wonder whether ambiguous cases could be treated as instantiations of a separate daughter construction of the N-*non*-N construction with a less vague and more transparent semantics and a lower degree of productivity. However, it is very unclear if positing such a sub-construction is actually justified or necessary to account for the synchronic situation we depicted. The ambiguous cases may well be instantiations either of 'normal syntax', where a noun is modified by a homophonous adjective/participle for the sake of expressiveness,[17] or simply instantiations of the posited N-*non*-N construction.

In sum, although ambiguous *tessuto non tessuto* has most likely played a major role in the development of the N-*non*-N construction, synchronically speaking ambiguous cases may just be an epiphenomenon. If this analysis holds, then the emergence of the N-*non*-N construction would be an intriguing case of microdiachronic constructionalisation (given the short time span) due to the entrenchment of an expression whose original, intended structure does not coincide with the structure of the new construction being created.

---

[17] Note that examples like (14)–(15), despite being 'ambiguous', would not be instantions of this 'normal syntax', but of the reduplicative construction, for the semantic reasons given above.

## 5. Conclusions

In this paper, we provided the first thorough description of the N-*non*-N discontinuous reduplication in Italian, based on corpus data. The N-*non*-N construction proved to be a productive and powerful device to approximate nouns in contemporary Italian. More specifically, the construction has a general function of 'non-prototypicality' and can produce expressions with complex and specific meanings that are generated in context, by deviating from the 'prototypical' N concept along different directions. Despite its productivity, very few N-*non*-N expressions are conventionalised, making this device similar to other approximative strategies and to evaluative morphology in general.

Although this study is primarily synchronic in nature, we speculated about the origin of this construction, since the leader expression in our dataset (*tessuto non tessuto*) has an ambiguous structure that allows for both a reduplicative and a modifying interpretation. We hypothesised that the N-*non*-N construction microdiachronically emerged through the entrenchment of *tessuto non tessuto*, which was (re)analysed as a N-*non*-N structure, perhaps also helped by the existence of an independent *non*-N pattern yeilding nouns.

Many questions remain unanswered and deserve further investigation. One is prosody. Our analysis is based on written language, but the tendency to host 'short' nouns may reflect some prosodic requirement that can only be inquired through the analysis of spoken data. The semantics of these expressions, especially in relation to the 'deviations' from the prototype (but also to the ambiguity issue), also needs to be explored more fully, by resorting to speakers' judgements. Finally, a comparison with other constructions conveying approximation in Italian (cf. Masini & Micheli 2020) is in order, to better understand the role of this specific strategy within the larger picture.

## References

Baayen, Harald & Rochelle Lieber. 1991. Productivity and English derivation: A corpus-based study. *Linguistics* 29. 801–843.

Baayen, Harald. 2009. Corpus linguistics in morphology: Morphological productivity. In Anke Lüdeling & Merja Kytö (eds.), *Corpus Linguistics. An International Handbook*, 900–919. Berlin: Mouton de Gruyter.

Blackings, Mairi & Nigel Fabb. 2003. *A Grammar of Ma'di*. Berlin: Mouton de Gruyter.

Booij, Geert. 2010. *Construction Morphology*. Oxford: Oxford University Press.

Dugas, Edwige. 2017. Constraints on the French [*non*-N] construction. In Nikos Koutsoukos, Jenny Audring & Francesca Masini (eds.), *Morphological Variation: Synchrony and Diachrony. Online Proceedings of the Eleventh Mediterranean Morphology Meeting (MMM11) Nicosia, Cyprus, 22–25 June 2017,* 14–25.

Ghomeshi Jila, Ray Jackendoff, Nicole Rosen & Kevin Russell. 2004. Contrastive focus reduplication in English (The salad-salad paper). *Natural Language & Linguistic Theory* 22. 307–357.

Goldberg, Adele. 1995. *Constructions: A Construction Grammar Approach to Argument Structure.* Chicago: University of Chicago Press.

Goldberg, Adele. 2006. *Constructions at Work*. Oxford: Oxford University Press.

Iacobini, Claudio. 2004. Prefissazione. In Maria Grossmann & Franz Rainer (eds.), *La formazione delle parole in italiano*, 97–163. Tübingen: Niemeyer.

Inkelas, Sharon. 2014. Non-concatenative derivation: Reduplication. In Rochelle Lieber & Pavol Štekauer (eds.), *The Oxford Handbook of Derivational Morphology*, 169–189. Oxford: Oxford University Press.

Jukes, Anthony. 2006. *Makassarese (basa Mangkasara'): A Description of an Austronesian Language of South Sulawesi*. PhD dissertation. University of Melbourne.

ListTyp = Masini, Francesca, Simone Mattiola & Stefano Dei Rossi. 2021. *The ListTyp database*. In Elisabetta Fersini, Marco Passarotti & Viviana Patti (eds.), *Proceedings of the Eighth Italian Conference on Computational Linguistics*, paper 17. CEUR Workshop Proceedings: http://ceur-ws.org/Vol-3033/ (accessed 6 March 2023).

Masini, Francesca & Maria Silvia Micheli. 2020. The morphological expression of approximation: the emerging *simil-* construction in Italian. *Word Structure* 13(3). 371–402.

Masini, Francesca, Muriel Norde & Kristel Van Goethem. 2023. Approximation in morphology: A state of the art. *Zeitschrift für Wortbildung/Journal of Word Formation* 7(1). 1–26.

Mattiola, Simone & Alessandra Barotto. 2023. Nominal reduplication in cross-linguistic perspective. From PLURALITY to CHANGE OF REFERENTS' SPECIFICITY. *Studies in Language* 47(1). 135–189.

Mauri, Caterina & Francesca Masini. 2022. Diversity, discourse, diachrony: A converging evidence methodology for grammar emergence. In Miriam Voghera (ed.), *From Speaking to Grammar*, 101–150. Bern: Peter Lang.

Moravcsik, Edith. 1978. Reduplicative constructions. In Joseph H. Greenberg (ed.), *Universals of Human Language. Volume 3: Word Structure*. 297–334. Stanford: Stanford University Press.

Moravcsik, Julius. 1975. Aitia as generative factor in Aristotle's philosophy. *Dialogue* 14(4). 622–638.

Pustejovsky, James. 1995. *The Generative Lexicon*. Cambridge: The MIT Press.

Stolz, Thomas. 2018. (Non-)Canonical reduplication. In Aina Urdze (ed.), *Non-Prototypical Reduplication*, 201–278. Berlin: Mouton de Gruyter.

Velupillai, Viveka. 2012. *An Introduction to Linguistic Typology*. Amsterdam: John Benjamins.

WALS = Dryer, Matthew S. & Martin Haspelmath (eds.). 2013. *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. http://wals.info (accessed 18 July 2021).

Francesca Masini
Alma Mater Studiorum – Università di Bologna
Dipartimento di Lingue, Letterature e Culture Moderne
Via Cartoleria 5
40124 Bologna, Italy
francesca.masini@unibo.it

Jacopo Di Donato
Alma Mater Studiorum – Università di Bologna Alumnus
Via IV Novembre 6
86100 Campobasso, Italy
jacopodidonato@gmail.com