

Evaluative (half-)prefixes in Italian: the expression of approximation through morphemes of half quantity

Flavio Pisciotta
University of Salerno
fpisciotta@unisa.it

INTRODUCTION

The semantic domain of half-quantity is expressed in mainly two ways in Italian: through the prefix semi- (e.g., semicilindro 'semicylinder'), and through adjectival modification of nouns with mezzo 'half' (e.g., mezzo cilindro 'half-cylinder').

Interestingly, it has been independently noted in the literature that both exhibit emergent evaluative functions, ranging from attenuation to approximation (Benigni 2023; Micheli 2023):

(1) a. *un clima semi-estivo* 'a near-summery climate' b. *un mezzo manicomio* 'something similar to an asylum'

In addition, it has been noted that in these uses *mezzo* often behaves like a prefixoid, also appearing orthographically bonded (Masini & Micheli 2020).

AIMS

We want to carry out a preliminary quantitative analysis of the behaviour of two coexisting strategies (*mezzo*-and *semi*-) that come from the **same semantic domain** (halfness), and whose extensions seem to go towards the same direction: the expression of evaluative meanings.

We will do so from a **Construction Grammar perspective**, looking at how this constructions could be linked in the *Constructicon*.

DATA

We extracted a random sample of 12.426 occurrences of semi- and mezzo- (bonded, debonded, and hyphened) from the **Standard written Italian corpus CORIS** (Rossini Favretti et al. 2002) – corresponding to 40% of their total occurrences in the corpus.

References

Linguistics

After filtering out the false positives, **11.253 occurences** were left:

	bonded	hyphenated	unbonded	Total
mezzo	2225	16	6335	<i>8576</i>
semi-	2392	281	4	2677
Total	4617	297	6339	11253

RESEARCH QUESTIONS...

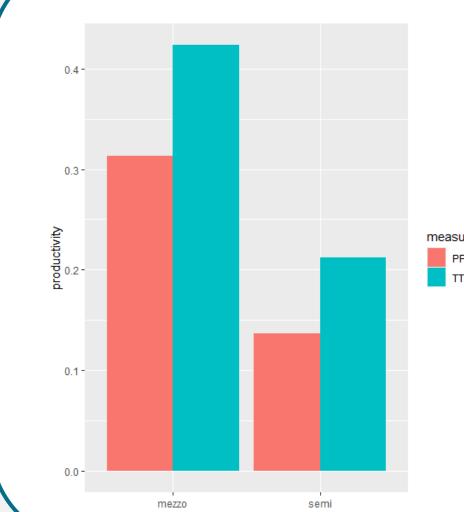
- **RQ1 –** Which meanings are expressed by *mezzo* and *semi*-? How are they mutually related?
- **RQ2 –** How productive are evaluative *mezzo-* and *semi-*?
- **RQ3 –** Do evaluative *mezzo-* and *semi-* show differences in their linguistic and "extra-linguistic" behaviour?

— ...AND HOW WE ADDRESS THEM-

- **M1 –** Semantic annotation of our dataset, starting from lexicographic definitions of *mezzo* and *semi* (cf. GRADIT, De Mauro 2007)
- M2 We calculate type-token ratio (TTR) and potential productivity (PP) (Baayen 2009) of the two constructions.
- **M3 –** We look at possible preferences on the fillers of the slots (Part of Speech, lemma) and the distribution across different text genres.

RQ1: MAIN MEANINGS OF semi-and mezzo Legend (preliminary proposal of a network) Polysemic link 'Middle point of X' mezzo n. types = 52Allostructional link 'Half of a whole' *mezzo* 'Half of a whole' *semi*n. tokens = 2311 ex. mezzanotte 'midnight' n. types = 275 | n. tokens = 3818 n. types = 45 | n. tokens = 189 Most common lexical use of *mezzo*: Neoclassical prefix employed with nouns, mainly from scientific noun-modifier indicating a half quantity: domains: ex. mezzo cerchio 'half of a circle' ex. semicerchio 'semicircle' 'Most part of X' *mezzo* The **depreciation meaning** emerges in some contexts as a shift from the approximating one. n. types: 90 **Evaluative** *mezzo* Evaluative *semi*n. tokens: 461 n. types = 444; n. types =386; → something is not a real X/does not show enough n. tokens = 911 of a property Y to be a good X n. tokens = 2096] aveva viaggiato per **mezza Europa** (2) Penso [...] a queste fiere della **semicultura** [...] 'He had traveled halfway across Europe' Both *mezzo*- and *semi*- can express **a variety of evaluative meanings** belonging to the convinta di aver la verità in tasca 'I think about this exhibitions of fake-culture, continuum of attenuation/reduction/approximation (Micheli 2023). Moreover, their semantics where people are convinced they have truth in allows a shift towads the expression of depreciation (cf. Grandi & Körtvélyessy 2015) their pockets' (3) ancora mi capita , raramente, di sentirmi **mezzo** Attenuation (partly X) Approximation (not properly X) Depreciation uomo [...] e mi ammazzo di palestra. 'I still rarely get to feel half a man, so I work out myself to death.' semiabbandonato 'half-abandoned' **semi**-alter ego 'a sort of alter ego' semicultura 'lit. half-culture' mezzo labirinto 'something similar to a labyrinth' mezzo distrutto 'partly destroyed' mezzo uomo 'lit. half-man'

RQ2: PRODUCTIVITY

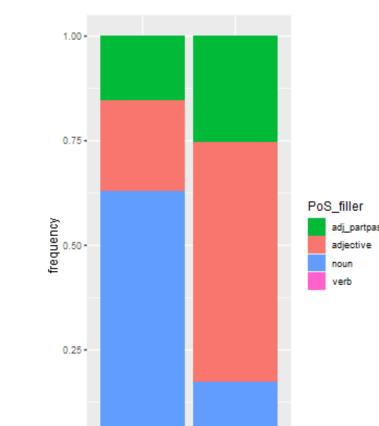


Evaluative *mezzo*- is way more productive than evaluative *semi*-, probably since it often appears unbonded (and thus it has still some lexical status).

The two constructions show about the same number of hapaxes (285 vs 286), however, *semi*- has a lower potential productivity (PP) since it shows a handful of very frequent types, e.g.:

seminterrato 'basement', (232)
semiconduttore 'semiconductor' (111)

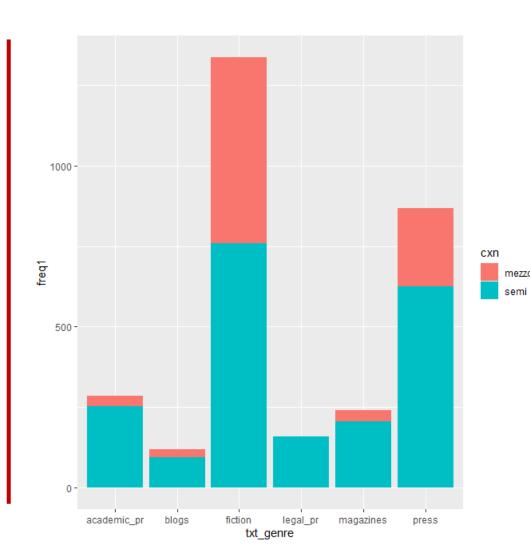




There is a partial division of labour between the two constructions: **semi-generally select adjectives** as fillers, while **mezzo- is strongly associated with nouns** (p < 0.001). However, we find a degree of overlap between the two constructions, mainly in the case of past participles. Moreover, they both can (but quite rarely do) modify

Part of Speech of the filler

(4) [...] **si era mezzo rotta** la caviglia. 'she had partially broken her ankle'



Text genres

Evaluative semi- is more frequent than evaluative mezzo- in almost all written genres. The only exception is **Fiction**, that shows a **significant association with** mezzo-. In such contexts, mezzo- is employed almost as frequently as semi-. (p < 0.001)

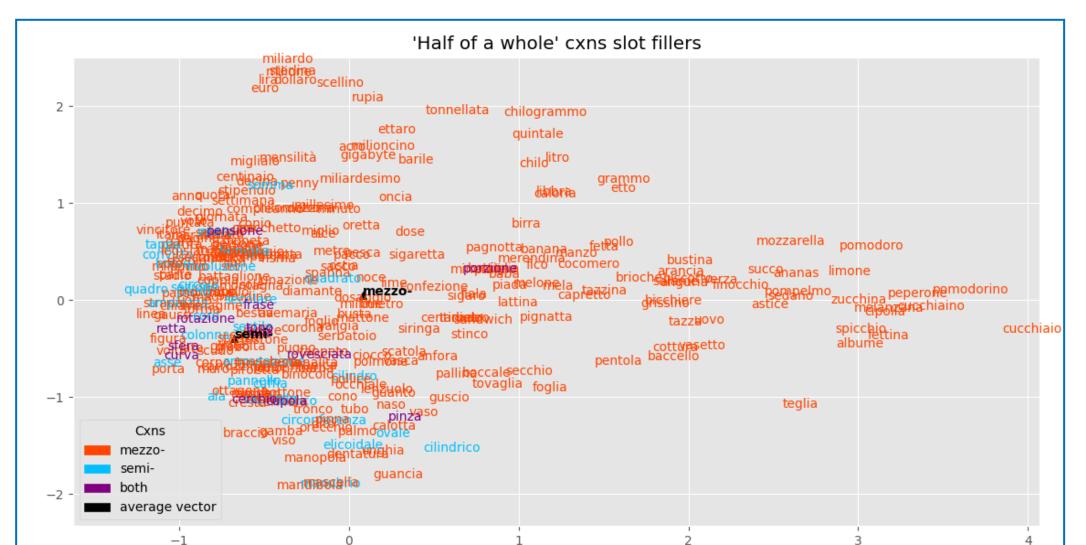
FUTURE STEPS IN THE ANALYSIS: ONGOING «ALLOSTRUCTIONAL» ATTRACTION?

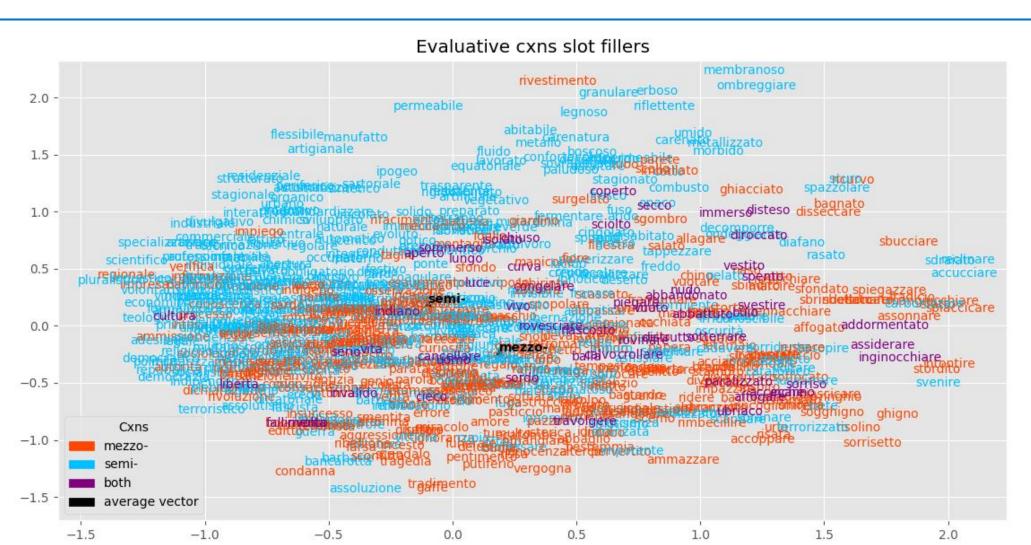
While a finer-grained semantic classification of the lexical fillers is still in progress, we noticed that **evaluative** semi- and mezzo- share many fillers when compared to their "basic" counterparts, i.e., "half of a whole" constructions → mainly past participles with adjectival function, but also nouns and adjectives.

Question: Are *semi*- and *mezzo*- getting "more semantically similar"?

Hypothesis: This could be a case of attraction between allostructions: allostructions tend not only to diverge semantically, but also to converge (De Smet et al. 2018).

As an exploratory analysis, we visualized the semantic distribution of the fillers by applying PCA to their word embeddings (the distributional model was trained on Italian web corpus itWaC, Baroni et al. 2009). PCA was performed in Python, using SciKit-Learn library. Then we compared it to the distribution of "half of a whole" cxns:





We also calculated the cosine similarity between average vectors of the fillers for both the couples of cxns (plotted in black): evaluative ones proved to be closer in the

distributional space (0.93 vs 0.83), and thus, their fillers tend on average to be more similar.