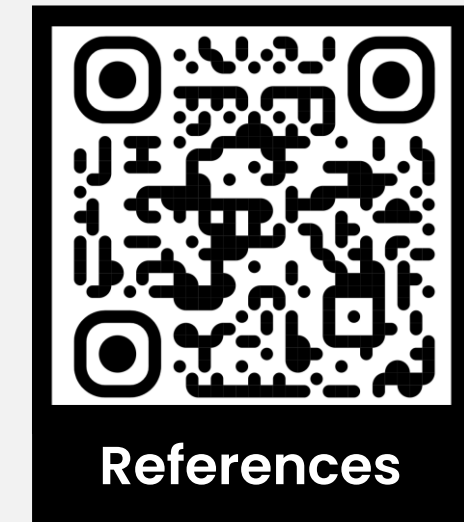




Evaluative (**half-**)prefixes in Italian: the expression of approximation through morphemes of half quantity

Flavio Pisciotta
University of Salerno
fpisciotta@unisa.it



References

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 *Constructionicon*.

DATA

We extracted a random sample of 12.426 occurrences of *semi-* and *mezzo-* (bonded, debonded, and hyphenated) from the **Standard written Italian corpus CORIS** (Rossini Favretti et al. 2002) – corresponding to 40% of their total occurrences in the corpus. After filtering out the false positives, **11.253 occurrences** were left:

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

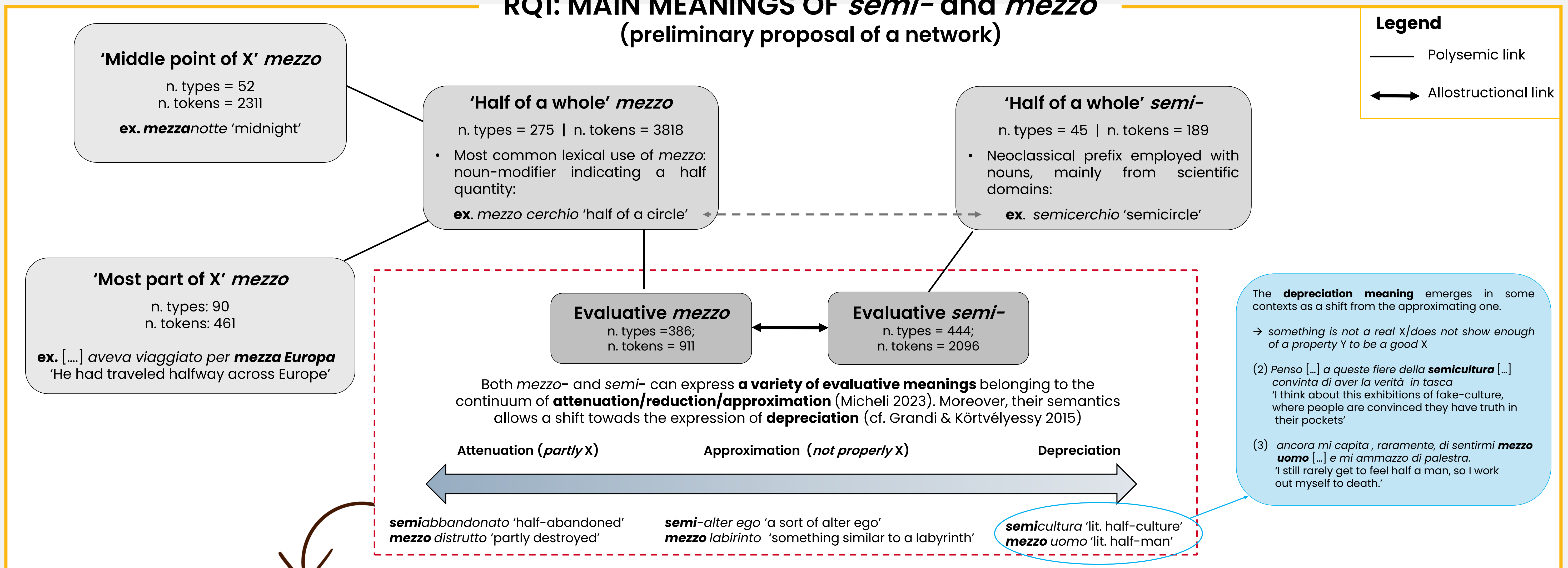
RESEARCH QUESTIONS...

- 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?

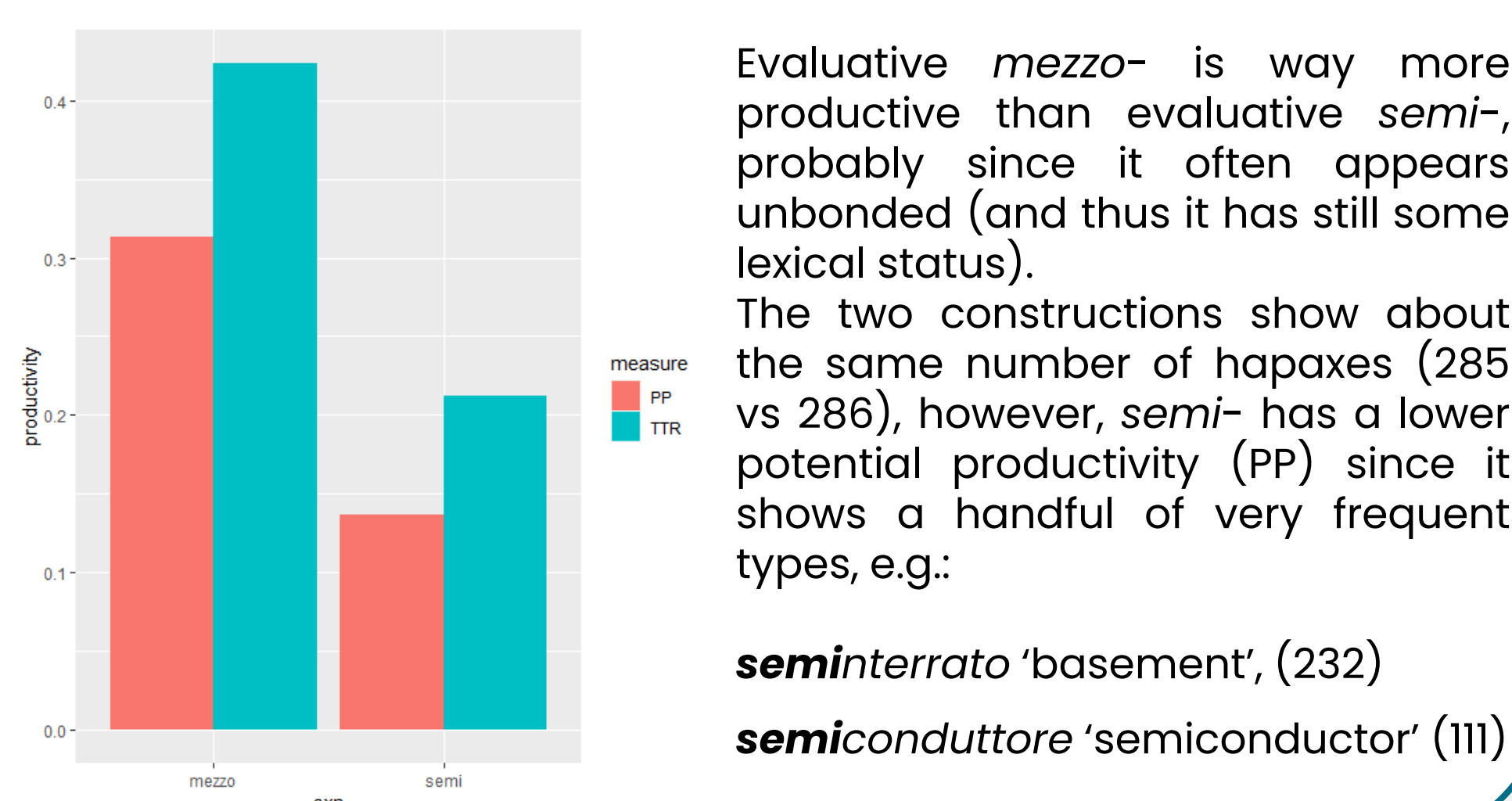
...AND HOW WE ADDRESS THEM

- 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.

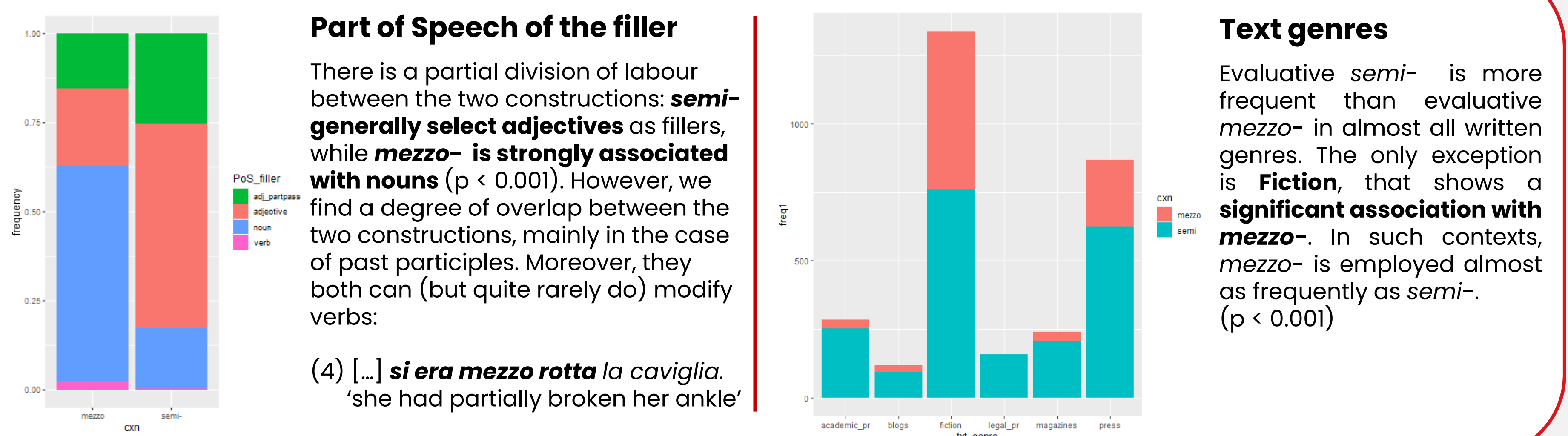
RQ1: MAIN MEANINGS OF *semi-* and *mezzo* (preliminary proposal of a network)



RQ2: PRODUCTIVITY



RQ3: LINGUISTIC AND CONTEXTUAL DISTRIBUTION



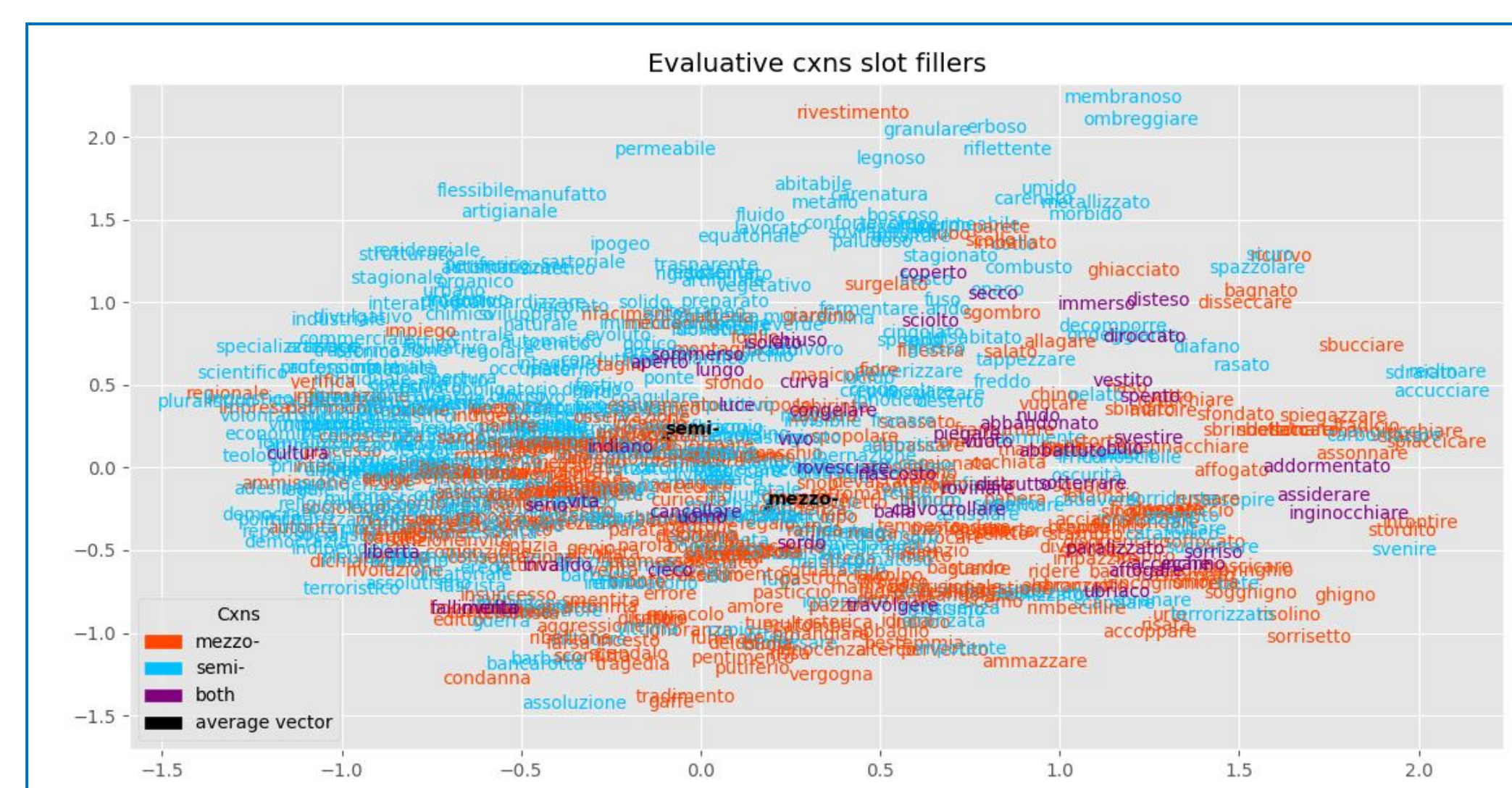
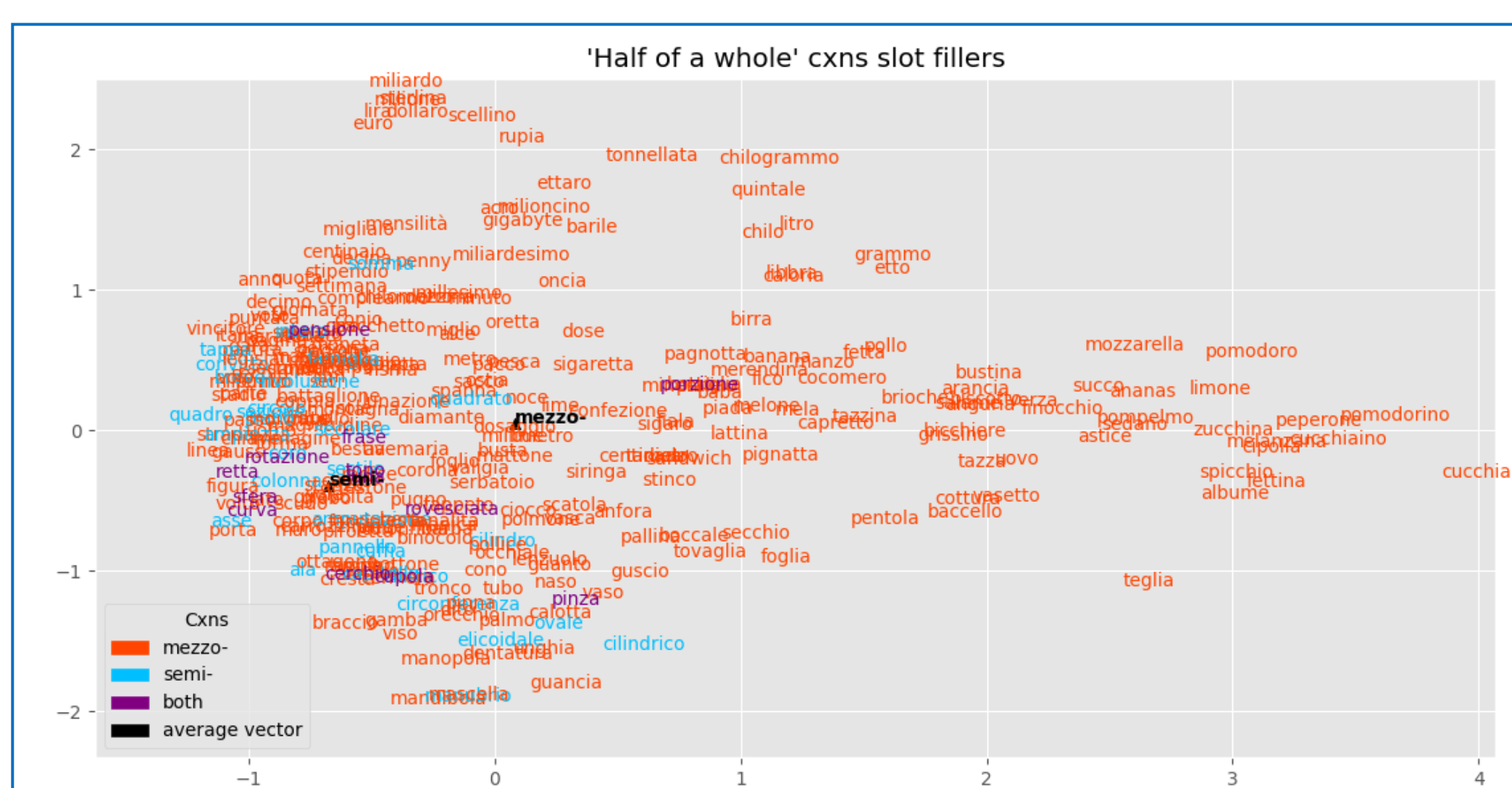
FUTURE STEPS IN THE ANALYSIS: ONGOING «ALLOSTRUCTIONAL» ATTRACTION?

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.