Constructions 17 (2025)
doi 10.24338/cons-737

Construction Grammar meets Graph theory: a network analysis of the relational behavior of Italian psych-predicates

Flavio Pisciotta[1]
[1] University of Salerno

Abstract

The paper provides a discussion of how to model constructional networks in graph-theoretic terms. We discuss the connection between graph theory and constructionist theory, and offer a linguistic and constructional interpretation for some major graph metrics. We illustrate the proposal by applying it to a network of constructions employed to create denominal psychological predicates in Italian. In particular, we employ graph metrics to asses to what extent synthetic and analytic constructions can be said to belong to the same paradigm of denominal verb formation. The results show the complex nature of the relationship between synthetic and analytic patterns: on the one hand, it seems that the constructions in the network belong to more than one paradigm, based on their structural features; on the other, such paradigms seem to be tightly linked in some domains, such as the causative one, which constitutes a bridge between constructions showing different levels of structural complexity. Overall, the analysis shows the advantages of developing a sound methodology to assess the relational behavior of constructions: in fact, it allows us to coherently describe networks of constructions, but also to operationalize concepts such as prototypicality and paradigmaticity in such networks. Nonetheless, it also shows the necessity of a careful theoretical and methodological discussion regarding the nature of objects in Construction Grammar.

1 Introduction

In Construction Grammar (CxG), linguistic knowledge is understood to consist of two basic types of objects: constructions, defined as pairings of form and function, and links, which represent formal and functional relationships between constructions (Goldberg 1995, 2006; Hoffmann & Trousdale 2013). In fact, a central tenet in CxG is that linguistic knowledge is not configured as a mere list of constructions, but as a network (Diessel 2019, 2023; Sommerer & Van De Velde 2025). Constructions can be linked both vertically and horizontally, based on their level of schematicity: vertical links express a relation of inheritance of some features

Corresponding author(s): Flavio Pisciotta, University of Salerno, fpisciotta@unisa.it

between constructions at different levels of schematicity, while horizontal links are generally employed to express semantic and formal relatedness between constructions at the same level of schematicity (Cappelle 2006; Booij & Masini 2015; Audring 2019; Ungerer 2024). For instance, in (1a), a vertical link connects a schematic construction, representing Noun + Noun compounding, with a semi-specified construction that instantiates it, namely the Noun + *-ache* compounding pattern (*headache*, *toothache*, etc.). Instead, in (1b), a horizontal link connect two semi-specified constructions that are often employed to create series of antonymous adjectives (*thoughtful* vs. *thoughtless*, *colourful* vs. *colourless*, etc.).

(1) a. Vertical link
Noun + Noun → Noun + *ache*
b. Horizontal link
Noun-*ful* ↔ Noun-*less*

Given the centrality of the notion of network for constructionist theory, it is highly surprising that research has largely overlooked the potential use of graph theory, and in particular the branch of network analysis, as a descriptive and analytical tool in CxG (Desagulier 2021). As a matter of fact, there has been a growing interest in the use of graph theory for linguistic analysis in corpus linguistics and psycholinguistic modeling (Gries & Ellis 2015; Banisch et al. 2016; Desagulier 2017, 2020; Shadrova 2022).

Recently, it has been suggested that the constructional network can be formalized in graph-theoretic terms, by considering constructions as graph nodes and links between them as edges (Sommerer & Van De Velde 2025), and some constructionist studies have started to incorporate network analysis. However, there are currently no structured proposals on how to exploit tools coming from this field and on how to interpret the relevant metrics, and, most importantly, rarely constructions as a whole have been modeled as graph nodes. In fact, most of the studies concentrate on construction fillers (i.e., lexical items), connecting them by virtue of their co-occurrence in specific constructional patterns (Sommerer & Baumann 2021; Morin et al. 2024), or by virtue of the semantic relations between them (Ellis et al. 2013, 2014).

Our aim is thus to propose a first structured account of constructional networks as graphs. We will take as an example a network of semi-schematic constructions which are employed to create denominal psychological predicates in Italian, comprising both synthetic and analytic patterns. We will show the advantages of employing network analysis for analyzing this small network, by reviewing some major graph measures and concepts and providing a linguistic framework for their interpretation. In particular, we will employ graph metrics and network analysis to assess to what extent analytic and synthetic constructions can be analyzed as exponents of a common paradigm of denominal verb formation. We first introduce the network in Section 2. Then, we discuss the graph modeling choices, and a proposal for the intepretation of the metrics employed and how they relate to the research questions (Section 3.1). We then introduce the dataset (Section 4). In Section 5 we present the results and discuss what our findings show us about the *paradigmhood* of the network under scrutiny. Finally, we address to what extent CxG can benefit from network analysis, and the limitations of our model (Section 6).

2 Case study: a "mixed" paradigm of psych-predicates?

In Italian, it is possible to create predicates from psych-nouns by making use of a variety of light verbs and derivational patterns.

Derivational patterns include conversion (2a), parasynthesis (i.e., the simultaneous application of prefixation and conversion) (2b), and suffixation (2c). Moreover, when they express causative semantics, it is possible to create anticausative synthetic predicates by employing the clitic *si* (2d).

(2) a. Noun-*ire*
gioia 'joy' > ***gioire*** 'rejoyce'
b. *in*-Noun-*ire*
fastidio 'annoyance' > ***infastidire*** 'annoy'
c. Noun-*izzare*
simpatia 'sympathy' > *simpat**izzare*** 'sympathize'
d. *in*-Noun-*ir-si*
fastidio 'annoyance' > ***infastidir-si*** 'get annoyed'

Instead, light verbs constructions are analytic constructions in which the noun is used non-referentially, and plays a major role in predication (Butt 2003; Ježek 2004). Such analytic predicates are generally created by using very frequent and polysemous verbs, as *avere* 'have' (3a), *prendere* 'take' (3b), *mettere* 'put' (3c), and many others. Notably, while studies on light verbs constructions have mainly concentrated on [Verb Noun] patterns (3a-3c), it is possible to find light verbs showing [Verb Preposition Noun] patterns (3d).

(3) a. *avere* Noun 'have N'
vergogna 'shame' > ***avere*** *vergogna* 'be ashamed'
b. *prendere* Noun 'take N'
coraggio 'anguish' > ***prendere*** *coraggio* 'take courage'
c. *mettere* Noun 'put N'
angoscia 'anguish' > ***mettere*** *angoscia* 'distress'
d. *essere in* Noun 'be in N'
ansia 'anxiety' > ***essere in*** *ansia* 'be anxious'

As noted by previous studies (Pisciotta & Masini 2025), psychological predicates, whether formed through synthetic or analytic processes, tend to fall into one of three different semantic classes: stative (2a-3a), inchoative (2b -3b), and causative (2c-3c). As shown by typological literature (Croft 1991; Talmy 2000), these three classes represent a stable contrast among the types of events that can be lexicalized in the verbal domain. This allows for describing the same event from different perspectives: the presence of a state, the process of entering that state, and the causal mechanism that brings about entering that state.

The presence of such a semantic contrast can be seen as a hint towards the paradigmatic nature of psych-predicate formation, at least from the semantic point of view (Pisciotta & Masini 2025): in fact, research carried out on derivational paradigms argue that paradigms are structured around stable semantic contrasts (e.g., Štekauer 2014; Bonami & Strnadová 2019). Under this perspective, the generalization made by the paradigm is a semantic one: by applying the modelization carried out on "action networks" (Roché 2017; Fradin 2020) to our case, we argue that speakers generally know that the existence of a psychological state (denoted by a noun) implies the existence of events of feeling that psychological state, of getting to feel that psychological state, and of causing somebody to feel that psychological state.

We could imagine such paradigmatic system in a cross-table fashion (Table 1), where each cell is determined by the semantic contrast between the event types, and is filled by the patterns that actualize said event type for a specific psych-noun.

noun	**stative predicate** (feel N)	**inchoative predicate** (begin to feel N)	**causative predicate** (cause to feel N)
paura 'fear'	***avere*** *paura*	***prendere*** *paura*, ***im***_paur_***irsi***	***fare*** *paura*, ***mettere*** *paura*, ***im***_paur_***ire***
amore 'love'	***provare*** *amore*	***inn***_amor_***arsi***	-
coraggio 'courage'	***provare*** *coraggio*	***prendere*** *coraggio*	***fare*** *coraggio*, ***dare*** *coraggio*, ***in***_coraggi_***are***

Table 1: Cross-table modeling of the paradigm of psych-predicates in Italian

As we can see in Table 1, both synthetic and analytic predicates serve as exponents in the cells of this paradigm, often overlapping within the same series (e.g., we have both synthetic and analytic causative predicates derived from *paura* 'fear').

This does not come as unexpected, as scholars have highlighted that multiword expressions can be exponents in inflectional and derivational paradigms, either filling gaps in complementary distribution with synthetic forms, or competing for the expression of the same meaning (Ackerman & Stump 2004; Masini 2019; Cetnarowska 2021). This is particularly true from a constructionist perspective, as the existence of paradigms comprising structurally different constructions is granted by the assumption of the syntax-lexicon continuum. Moreover, the non-canonical features shown by such a paradigm should not be a concern: in fact, the presence of multiple exponents actualizing a cell for the same noun, as well as the presence of gaps (e.g., no causative predicate is formed by using *amore* 'love'), are not uncommon for non-inflectional paradigms (Bonami & Strnadová

2019; Melloni & Dal Maso 2022).

Nonetheless, modeling the psych-predicates as a cross-table stands on the *a priori* assumption that the multiplicity of synthetic and analytic predicates found belong to the same paradigm. Instead, the degree of parigm participation of (synthetic and analytic) word-formation patterns should be assessed by taking into account not only the semantic predictability expressed by our set of constructions (i.e., the limited set of event types that can be actualized), but also how recurrently and regularly pairs of patterns actualize related meanings (Fradin 2018; Hathout & Namer 2019). In other terms, in order to generalize a paradigm, we need a frequent alternation between patterns in expressing a semantic opposition, and some degree of predictability and motivation in this relationship (even though non-inflectional paradigms often show lower predictability, cf. Melloni & Dal Maso 2022). We maintain that formalizing a paradigm as a cross-tables does not let us easily approach such problems.

In fact, a cross-table perspective can help us check which patterns are the most frequent and the association between specific patterns and semantic values, but it fails to let us capture relavant information, the most important being which of the patterns take part together in a relevant number of paradigmatic series (i.e., if they share the same nominal bases), how frequently, and if we can find some regularities in our system.

Moreover, it does not inform us on the properties and internal structure of such a network/paradigm: for instance, we do not know which semantic values nor which specific form-function mappings are at the core of psych-predicate formation[1]. This is because cross-tables characterize grammatical paradigms as closed symmetric systems, while in reality they are often asymmetrical and organized around a basic member (e.g., singular number is more basic than plural in nominal paradigms). This can be easily accounted for by conceptualizing the paradigm as a set of relations between its members (Bybee 1985; Diessel 2023).

We propose that a network-based perspective can help us address these problems in a theoretically consistent fashion. As a matter of fact, usage-based and constructional research has shown that paradigms can be more fruitfully modeled as networks (Bybee 1985; Booij 2010; Smirnova 2021; Leino 2022; Diessel 2023). In particular, we have a complex situation including a variable number of exponents for the same cell, and thus it seems more informative to see how actually the network gets shaped by the relationships between individual constructions, and to what extent its behavior resembles a paradigm's one.

2.1 Hypothesis and research questions

We start by putting forward the hypothesis that synthetic and analytic constructions belong to the same paradigmatic family. This, in turn, requires to first verify if constructions employed to create psych-predicates actually project a paradigm-

[1] Note that this is not only a matter of frequency, because a very frequent pattern could theoretically be only loosely connected to the rest of the paradigm, and occupy a niche on its own.

like network. Based on the literature presented in this section, in order to be considered a paradigm in a narrow sense, the network projected by our constructions should show:

1. **frequent and recurrent alternations**: thus, patterns should either a) form a tightly interconnected network, that is, share a high number of fillers, or b) form a network shaped as a collection of tightly interconnected clusters;
2. a tendency towards the expression of relationships of **semantic contrast** through said alternations;
3. **regular alternations**: thus, patterns should, to some extent, be able to motivate and predict their mutual relationship.

Clearly, paradigm participation can be seen as a gradient phenomenon, especially with respect to predictability. However, in order to define a network of constructions as a paradigm in a narrow sense, we require at least points 1 and 2 to be features of our network. Moreover, the more regular and predictable the relationships are (point 3), the clearer the network's paradigmatic nature becomes.

If we hypothesize that analytic and synthetic constructions are part of a common paradigm, these kinds of relations should hold not only between items at the same level of complexity, but also between word and multiword patterns.

We will assess this two-step hypothesis by formalizing the constructional space as a network, in continuity with constructional and usage-based research, and by employing graph metrics as a quantitative-descriptive tool to investigate the features of such network. We will approach the issue in a descriptive fashion, as no research so far has been carried out in this direction. Thus, there is no benchmark for our metrics. We carry out our analysis by:

RQ1: describing the type of network projected by our patterns: that is, the degree and the nature of the connections between constructions employed to derive psych-predicates;

RQ2: finding the most basic member(s) of the network: that is, finding which constructions are mostly available in paradigmatic relations with other constructions, and if they all pertain to some formal and/or semantic types;

RQ3: assessing how predictable the relationship between groups of constructions, i.e, if there are more predictable "subparadigms" of constructions is, and if they comprise constructions at different levels of complexity.

In the following Section, we show how such research questions can be addressed in a graph-theoretic framework.

3 Graph modeling

3.1 Representation choices

In mathematical terms, a graph is a structured set of objects where some pairs of the objects are connected. Objects and relationships between objects are the fundamental units of a graph, and are formalized as *nodes* (or *vertices*) and *edges* (Trudeau 1994). Nodes are represented as dots, connected by lines or arrows representing edges (Figure 1).

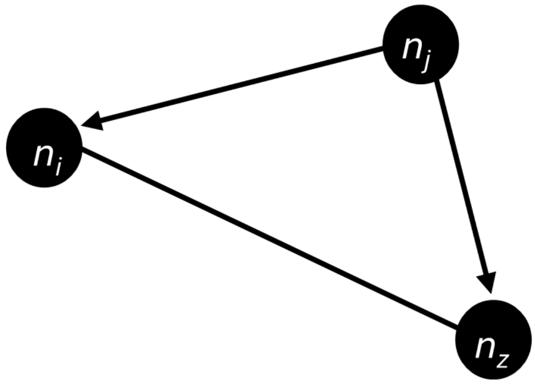

Figure 1: Representation of a mixed graph, containing both directed and undirected edges.

Graphs can include different types of edges: directed edges (arrows), that describe a path starting from a node *j* and culminating on a second node *i*, or undirected edges (lines), that do not describe any direction in the relationship. It is also possible to assign weights to edges, to represent their relevance in the network: for example, when representing air traffic, weights inform us about how frequently a specific airpath is travelled.

In the network we want to model, we propose that the kind of relations holding between the members of our paradigm members are horizontal, bidirectional links, since the paradigm contains constructions that cannot share any formal feature inherited from a common 'mother' schema. Such horizontal links express both quasi-synonymy between the semantics of the patterns (i.e., linking patterns belonging to the same cell in the paradigm) and semantic contrast (i.e., linking patterns belonging to different cells). Our graph should thus contain constructions as nodes, and horizontal links as undirected edges. While formalizing constructions as nodes is quite uncontroversial, we need to choose how to operationalize edges that link constructions. This is not a trivial operation, since linguistic networks are generally quite abstract: the definition of nodes and links is not explicit in the data, but relies on some modeling, and rarely a linguistic network represents an actual process taking place (as instead happens in, e.g., inter-bank money transfers) (Araújo & Banisch 2016). Moreover, constructions themselves can be related on multiple layers: links could represent the formal as well as the semantic relatedness of the patterns.

However, if we consider that horizontal links represent generalizations about the similarity between pairs of constructions, it is clear that such generalizations can be drawn only if the constructions involved can be somehow explicitly recognized by the speakers (Audring 2019). If we get back to (1b), we see that the paradigmatic relation between antonymous suffixes *-less* and *-ful* (cf. 1b) can emerge since they: 1) show a regular semantic opposition and 2) can be often employed with the same bases to express such opposition (4). This is coherent with the approach delineated in Section 2, as the regular and recurrent use with the same fillers to express a regular semantic relation represents the criteria defined to assess the degree of paradigmatic cohesion of the network..

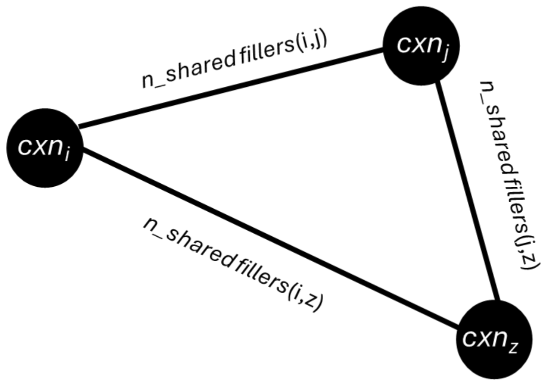

Figure 2: Network of three constructions formalized as a weighted, undirected graph.

(4) Noun-*ful* ↔ Noun-*less*
| |
*colour**ful*** ↔ *colour**less***
*pain**ful*** ↔ *pain**less***
*thought**ful*** ↔ *thought**less***

Since we know that our constructions are always, by design, linked on the semantic level (they all have a quasi-synonymy or a contrast relation with all of the other constructions), we can factor out relations on the semantic level, and we can instead turn our attention to the number of shared fillers (cf. the approach taken in Van De Velde & Fonteyn 2017) to operationalize the amount of perceived relatedness between pairs of patterns. In other terms, we assume that the more two patterns are used with the same fillers, the more likely speakers will generalize a relation between the two patterns. We can translate this assumption in graph-theoretic terms by describing our constructional network as an undirected weighted graph (Figure 2), where:

- **nodes** represent semi-schematic constructions;
- **weighted edges** represent the number of bases shared by two constructions (nodes).

In order to build the graph, we use the *NetworkX* package (Hagberg et al. 2008) in Python 3.12.

3.2 Network-analytic methods and constructional interpretation

3.2.1 Graph metrics

Following the tripartite structure of our research questions (Section 2), we divided the metrics we calculated into three macro-groups: metrics related to **network properties**, node **centrality measures**, and a **predictability measure**. Such metrics were calculated by employing the packages *NetworkX* and *igraph* (Csardi & Nepusz 2006) in Python 3.12. In this section, we explain the selected metrics and their possible constructional interpretation; we summarize them in Table 2.

We mostly employed weighted versions of the metrics. This is a relevant point, since the distinction between weighted and unweighted measures can be very compelling in constructionist terms. In fact, unweighted measures assign to all the constructional links the same value, while weighted ones take into account the number of connections underlying the constructional link (in our case, the number of shared fillers). In this sense, they open a window on phenomena pertaining to a lower level of abstraction, since they assign different values based on the number of links between fully-specified instances of the two constructions. In our case, since we assume that strong paradigmatic generalizations rely on the number of actually encountered doublets, weighted measures seem to be more suited. Nonetheless, we cannot exclude that, in some cases, we could be interested only in the presence/absence of a connection between two semi-specified schemas.

Network properties help us address our first reasearch question (RQ1). They allow us to calculate how many connections there are in our network, to what extent constructions tend to cluster together, and also which types of constructions tend to be linked. In particular, we checked for:

- **Average degree**: the average number of connections per node in the graph. When unweighted, it represents the average number of constructions a construction is linked to; in its weighted version it tells us, on average, how many fillers constructions share. Average degree is calculated as:

$$\langle k \rangle = \frac{1}{N} \sum_{i=1}^{N} k_i \tag{5}$$

  where N is the number of nodes in the graph, and k_i is the degree of the node i, i.e., the number of edges starting from that node (Estrada & Knight 2015). In the weighted version, instead, k_i is calculated as the sum of the weights of its incident edges:

$$k_i = \sum_{j \in \mathcal{N}(i)} w_{ij} \tag{6}$$

  where $\mathcal{N}(i)$ is the set denoting the neighbors of the node i.
- **Transitivity [between 0 and 1]**: it is an unweighted measure of clustering. It measures the probability that two nodes adjacent to a third node are connected in our network, i.e., to what extent all of our constructions tend to be linked together. It is calculated as the ratio between the number of triangles (i.e., closed triplets of nodes where each node is connected to the other two) and the number of possible triangles in the graph, where possible triangles are identified by the number of triads (i.e., triplets of nodes formed by a node and two of its neighbors):

$$T = 3 \times \frac{\#\,\text{triangles}}{\#\,\text{triads}} \tag{7}$$

  Since it is calculated as a coefficient, the closer the value to 1, the higher the closure of the network structure.

- **Average clustering [between 0 and 1]**: it is another clustering metric. It measures, on average, to what extent constructions tend to cluster in the network. We employed this clustering measure as well for two reasons: firstly, it also has a weighted version, unlike transitivity; secondly, studies have shown that it may diverge from transitivity for some types of networks (Estrada 2016). In its unweighted version, it is the average of the local clustering coefficients, which correspond, in turn, to the number of closed triplets (triangles) involving the node i (denoted t_i), normalized by the maximum number of possible connections between the neighbors of that node (Watts & Strogatz 1998):

$$\overline{C} = \frac{1}{N} \sum_{i=1}^{N} \frac{2t_i}{k_i(k_i - 1)} \tag{8}$$

In the weighted version, the local clustering coefficient is multiplied by the geometric mean of the edge weights in the triad formed by the node i and its two neighbors j, k (Saramäki et al. 2007):

$$\overline{C}_w = \frac{1}{N} \sum_{i=1}^{N} \frac{1}{k_i(k_i - 1)} \sum_{j,k \in \mathcal{N}(i)} (w_{ij} w_{ik} w_{jk})^{1/3} \tag{9}$$

The weighted version assigns more importance to "heavy" triangles, and thus, a high score means that groups of constructions that give rise to many low-level alternations contribute more to the connectedness of the network.
- **Attribute assortativity [between -1 and 1]**: This measure, also referred to as homophily, informs us about the type of constructions that tend to be interconnected. Specifically, it quantifies the extent to which nodes in the graph connect preferentially to other nodes with similar attributes (Newman 2002, 2003). Attributes are discrete categories, which correspond, in our case, to constructional properties, such as structural and semantic ones. We calculate attribute assortativity by first computing the mixing matrix $(e_{ij})_{i,j=1}^{C}$, where e_{ij} denotes the fraction of edges in the network that connect nodes of type i to nodes of type j, and C is the number of discrete categories. Then, assortativity is calculated as:

$$r = \frac{\sum_{i=1}^{C} e_{ii} - \sum_{i=1}^{C} a_i b_i}{1 - \sum_{i=1}^{C} a_i b_i} \quad \text{where} \quad a_i = \sum_{j=1}^{C} e_{ij}, \quad b_i = \sum_{j=1}^{C} e_{ji} \tag{10}$$

In the formula, $\sum_{i=1}^{C} e_{ii}$ corresponds to the sum of the diagonal entries in the matrix, which gives the proportion of edges that link nodes of the same type. Instead, a_i and b_i are the marginal probabilities, calculated from the mixing matrix, that each of the two ends of an edge are attached, respectively, to a node of type i, and to a node of type j[2].
In the weighted version, the mixing matrix is computed based on the total weights of edges connecting node types, rather than the count of edges.

[2] In our case, $a_i b_i$ can be rewritten as ${a_i}^2$, since the mixing matrix is symmetrical in undirected graphs, and thus $a_i = b_i$ (Newman 2003).

In linguistic terms, a high weighted assortativity (or assortative mixing, with a score close to 1) tells us that constructions tend to link to constructions sharing similar features; a score of 0 indicates no correlation, while low assortativity (disassortative mixing, close to -1) shows a tendency to find links between different types of constructions. As we mentioned, we will calculate assortativity based on the structural type of construction (synthetic vs. analytic construction), but also based on the event type expressed. This will provide us with two pieces of information, the first being the integration of different structural types of constructions in the paradigm, and the second the type of semantic relations mostly found in the network (semantic similarity vs. semantic contrast). If contrast relations prevail over similarity ones, we can assume that the network is mostly shaped as a paradigm (i.e., by semantic contrast relations).

The presence of clusters was also assessed by finding the **maximal cliques** and the **Louvain communities**. Maximal cliques of a node are all the largest complete subgraphs containing that node - where "complete" means that all the nodes (constructions) in the subgraph are interconnected. Thus, maximal cliques represent clusters of constructions whose connection does not depend on the rest of the network. Instead, Louvain algorithm performs a partitioning of the network that maximizes modularity for each community of nodes. Modularity measures how densely nodes are connected in a community, as compared to what would be expected in a random network with the same degree. Thus, communities represent clusters of constructions that create dense connections, that is, that frequently and regularly share fillers. Both cliques and communities were calculated using the algorithms provided in *NetworkX*.

Turning to centrality metrics, we use them to address the second research question (RQ2), i.e., we want to find constructions that lie at the core of the network. We calculated two weighted **centrality measures**:

- **Degree centrality**: ranks nodes based on the number of connections they have with other nodes. Constructions with a high degree will be the most prototypical and basic ones in the network, since they will be the ones used most frequently and typically in paradigmatic choice with the other constructions. Degree centrality is calculated for each node i as the sum of the weights of its incident edges (neighbors):

$$C_D^w(i) = \sum_{j \in \mathcal{N}(i)} w_{ij} \tag{11}$$

- **Betweenness centrality**: ranks nodes based on the number of shortest paths (between other nodes) that pass through them. In our case, shortest paths are formed by stronger links between constructions. Weighted betweenness is calculated for each node i as the sum of the fraction of shortest paths[3] between the nodes s, t that pass through i:

[3] Note that the formula employed in *NetworkX* treats weights as distances (Brandes 2001): thus, heavier edges will be considered as longer paths. However, this goes against the fact that a heavier connection between constructions makes them closer in our model (cf. Section

$$C_B^w(i) = \sum_{s \neq i \neq t} \frac{\sigma_{st}^w(i)}{\sigma_{st}^w} \tag{12}$$

It lets us isolate critical nodes, that are, nodes which are most influential for the flow of information of the graph. Such critical nodes should correspond to articulation points, i.e., nodes whose removal disconnects the network (Ausiello et al. 2013). Thus, constructions with a high betweenness centrality are the ones most often employed in alternations that link subdomains of the network that would be otherwise separate. In other words, it means that we can find the constructions that contribute most to attracting and linking other constructions to the network.

Finally, we exploit a **predictability** measure to address how regular are relations in the proposed paradigm of psych-predicates (RQ3). Thus, for this last measure we only take into account links that describe transitions from an event type to another.

In the literature, predictability between paradigm exponents has been addressed by employing extensions of Shannon's Entropy, which represents the amount of information we need to capture a distribution (and thus, it is a measure of uncertainty). In particular, studies have employed measures such as conditional entropy (Ackerman et al. 2009; Ackerman & Malouf 2013) or variations, such as (joint) implicative entropy (Bonami & Beniamine 2016): such measures quantify how much the distribution of a variable depends on the distribution of one (or more) variable(s). In other words, they formalize the ease in predicting which pattern will actualize a cell by knowing the exponent(s) of another (or other) cell(s). Since we assume a network perspective on paradigms, we decided to rely on a different entropy-based measure, **node diversity** (Eagle et al. 2010), which tells us the amount of uncertainty in determining our path when starting from a specific node in the network. In other terms, it can be seen as a form of conditional entropy $H(X|Y)$ given a fixed value (i.e., construction represented by the node) for Y.

Diversity can be easily calculated by using the *igraph* package. First, the node's Shannon's Entropy is computed, based on the weights of its incident edges:

$$H_i = -\sum_{j=1}^{k_i} p_{ij} \log(p_{ij}) \tag{13}$$

where p_{ij} is the probability of transitioning from i to j, according to the weights of the connection between them:

$$p_{ij} = \frac{w_{ij}}{\sum_{j'=1}^{k_i} w_{ij'}} \tag{14}$$

Then, the entropy value gets normalized by the log of the degree of that node (i.e., the number of connections starting from that node):

3.2.2). In order to account for that, in this case we substituted weights with their reciprocal (1/*weight*).

$$D_i = \frac{H_i}{\log(k_i)} \tag{15}$$

In our case, we were not interested in every possible connection starting from a node, but only in the amount of uncertainty in transitioning from a construction expressing an event type to another. Since we use a weighted measure, such uncertainty is obviously informed by the frequency of shared fillers, and not only by the number of connections at the semi-schematic level: in case of two equally frequent options, uncertainty will be higher than in case of stronger tendencies towards one of the two constructions. Thus, we employed the following procedure: for each $construction_i$, we created two subgraphs in which $construction_i$ is linked only to constructions expressing, respectively, the two contrasting event types in the paradigm. Then, we calculated the diversity of $construction_i$ in each of the two subgraphs. As an example, in the case of stative ***provare*** N, $subgraph_1$ contains ***provare*** N and its inchoative neighbors, while $subgraph_2$ contains ***provare*** N and its causative neighbors. For each subgraph, we calculated the diversity of ***provare*** N. In this way, we have two scores for ***provare*** N: the first represents the uncertainty in predicting the pattern employed to create the inchoative predicate if we assume that the stative one is expressed by ***provare*** N; the second will give us the same information, but for predicting the causative construction employed. Since diversity is a normalized score, when it is close to 0 it means that, by knowing that $construction_i$ is used with a noun, we can easily predict the corresponding $construction_j$ expressing another event type. Instead, if the score is close to 1, it will reflect high uncertainty (which corresponds to a high diversity in terms of connections).

After this overview, we can summarize the methods employed and their constructional interpretation in Table 2.

3.2.2 Visualization

Coherently with the use of weighted metrics, we used the *spring_layout* function in *NetworkX* to position the nodes in the graph visualization. In fact, the function employs the Fruchterman-Reingold force-directed algorithm to calculate the positions (Fruchterman & Reingold 1991). This algorithm treats nodes as repelling objects, and edges as springs that attract the nodes they connect. Thus, once having reached the equilibrium, nodes that are not connected by any edge are drawn further apart, and instead, nodes that are connected by an edge are positioned closer, accordingly to the weight assigned to the edges. This also results in the fact that the more a node is connected to other nodes (and the heavier are the links), the closer it will be to the center of the graph. In linguistic terms, it means that:

- the higher is the number of fillers shared by two constructions, the closer the two constructions will be in the visualization.
- constructions positioned at the center of the network are the ones that share fillers with most of the other constructions.

Method	***Group***	***Constructional Interpretation***
Average degree	*Network properties*	Average number of links each semi-schematic construction projects.
Average degree (Weighted)	*Network properties*	Average number of low-level alternations in which each semi-schematic construction is involved.
Transitivity & Average clustering [from 0 to 1]	*Network properties*	Indicate the tendency to project a closed network, that is, a score of 1 tells us that all pairs of linked constructions are both connected to another construction.
Average clustering (Weighted) [from 0 to 1]	*Network properties*	Indicates if closed clusters are formed by semi-schematic constructions that give rise to many low-level alternations.
Attribute Assortativity [from -1 to 1]	*Network properties*	Measures homophily at a higher level of schematicity, i.e., tells us if semi-schematic constructions tend to be connected with similar constructions (with respect to formal, structural, semantic, etc. features).
Attribute Assortativity (Weighted) [from -1 to 1]	*Network properties*	Measures homophily at a lower level of schematicity, i.e., tells us if fully-specified constructions tend to be connected with similar constructions (with respect to formal, structural, semantic, etc. features).
Maximal cliques	*Network properties*	Represent "closed" clusters of semi-schematic constructions, i.e., clusters in which every node is directly connected to all others, independently of the rest of the network.
Louvain communities	*Network properties*	Represent clusters of semi-schematic constructions that systematically share fillers.
Degree Centrality (Weighted)	*Centrality*	Ranks semi-schematic constructions based on how connected they are, i.e., the higher the number of alternations in which a construction is involved, the more central a construction will be.
Betweeness Centrality (Weighted)	*Centrality*	Ranks semi-schematic constructions based on how often they give rise to alternations that connect otherwise separate subdomains of the network.
Node diversity (Weighted) [from 0 to 1]	*Predictability*	Tells us the amount of uncertainty in predicting which semi-schematic constructions can be used in paradigmatic alternation with a given construction.

Table 2: Constructional interpretation of the selected methods.

4 Data collection and preparation

For the creation of the dataset, we began by compiling a list of psych-nouns. To this end, we drew on ItEM(Passaro et al. 2015), an Italian emotive lexicon[4]. This resource was selected not only for its extensive lexical coverage, but also because it enables a targeted selection of nouns relevant to our study. In fact, ItEM includes distributional similarity scores between Italian lemmas and Plutchik's (1980) eight basic emotion terms (joy, trust, fear, surprise, sadness, disgust, anger, and anticipation). This feature was originally designed to provide emotional orientation scores for sentiment analysis and related tasks. Nonetheless, it also makes it possible to distinguish emotion-related words from non-emotive vocabulary. Specifically, we assume that psych-nouns will generally yield higher cosine similarity scores to basic emotion terms than to nouns unrelated to the psych-domain, thus providing a principled basis for a preliminary selection of lemmas.

We pre-selected lemmas yielding a cosine similarity of at least 0.45 with basic emotion terms. We then manually selected only the lemmas expressing a psychological state and ended up with a list of 199 nouns, further enriched by comparing it with the list of 153 emotion terms provided in Zammuner (1998). By cross-comparing the two lists, we reached the number of 217 nouns. For the sake of simplicity, we filtered out deadjectival and deverbal nouns, by relying on the data contained in the Italian dictionary GRADIT (*Grande dizionario italiano dell'uso*; De Mauro 2007). The final list contains 86 nouns.

The next step was then to create a dataset containing all the types created from the obtained noun list. Thus, for each noun, we collected all the derived synthetic and analytic predicates, and annotated them for the type of event expressed:

- **stative**: X feels N
- **inchoative**: X begins/gets to feel N
- **causative**: X causes Y to feel N

In particular, for the synthetic ones, we collected for each noun the corresponding denominal verbs we found in GRADIT, filtering out the ones marked as obsolete or literary-only. Instead, for the analytic constructions, we restricted our search to 10 patterns selected from the literature (Ježek 2004; Pompei & Piunno 2023):

- *essere in* N 'be in N', *avere* N 'have N', *provare* N 'feel N', and *sentire* N 'feel N' for the stative meaning;
- *prendere/si* N 'take N', *farsi* N 'do oneself N', *andare in* N 'go in N' for the inchoative meaning;
- *fare N* 'do N', *mettere* N 'put N', and *dare* N 'give N' for the causative meaning.

We then checked for the occurrence of all these patterns filled by the 86 nouns from our list in an Italian Web-crawled corpus, namely itWaC small (Baroni et al. 2009) (~78 Mw). We discarded filled patterns with frequency lower than 5, as well as the ones that are attested but do not actually correspond to psychological meanings. We performed this annotation based on introspective judgements. The predicates were then coded for the type of construction.

4 https://github.com/Unipisa/ItEM/tree/master/output.

The resulting dataset of predicates contains 272 predicates (Table 3), created by employing 19 constructions. Said constructions are the result of the mapping between 14 patterns[5] and three semantic structures (corresponding to the event types described above):

- **statives**: conversion (N-*a/ire*), suffixation (N-*izzare*), *avere* N, *essere in* N, *provare* N, *sentire* N, anticausative converted (N-*a/irsi*) and parasynthetic verbs (*ad-/in-/s*-N-*a/irsi*);
- **inchoatives**: anticausative converted (N-*a/irsi*) and parasynthetic verbs (*ad-/in-/s*-N-*a/irsi*), *prendere* N, *farsi* N, *andare in* N;
- **causatives**: conversion (N-*are/-ire*), suffixation (N-*izzare*), parasynthesis (*ad-/in-/ s*-N-*are/-ire*), *fare* N, *dare* N, *mettere* N.

	analytic	**synthetic**	***total***
stative	121	16	137
inchoative	14	27	41
causative	63	31	94
total	198	74	272

Table 3: Number of predicates crossed by event type and complexity.

In order to build the graph, we crossed all the patterns in our dataset and annotated, for each pair of patterns, how many nominal fillers they share, obtaining 107 pairs (at least one shared base). Then, we set a threshold of shared bases for "relatedness", since only 1 shared filler would not be enough to assume a link between two patterns, and could as well be due to chance, especially because we collected our data based on a closed set of possible fillers. Thus, we set a threshold, based on the median number of shared fillers between the pairs of constructions in the dataset (median = 3): we assumed that there is a connection when two constructions share an above-average number of fillers. The resulting dataset contains 15 constructions, forming 44 pairs (Table 4).

cxn1	cxn2	n_fillers
avere N	*provare* N	34
avere N	*sentire* N	22
provare N	*sentire* N	22
avere N	*dare* N	20
provare N	*fare* N	19
provare N	*dare* N	18
avere N	*fare* N	16
provare N	*mettere* N	13
avere N	*mettere* N	13
sentire N	*fare* N	12

Table 4: First 10 pairs of constructions by number of shared bases.

[5] Note that some patterns can be mapped on more than one semantic structure: for instance, conversion can be employed to create both stative and causative predicates.

5 Results and discussion

The network is plotted in Figure 3. As we can see, setting the threshold to 4 led to the exclusion of four patterns: *-izzare* suffixation and anticausative parasynthetic verbs (statives), *farsi* N (inchoative), and *-izzare* suffixation (causative). In the following subsections, we illustrate and discuss the results of the metrics, divided in macro-groups. Finally, in Section 5.4 we address our hypothesis, by building on pieces of evidence provided by our results.

5.1 Network properties (RQ1)

Table 5 shows the results for the network metrics. As we mentioned, network properties help us analyze how interconnected constructions in the network are, and which kind of constructions tend to be interconnected (RQ1). Given the variety of network measures we have calculated, we present them in two separate groups: one includes metrics about the connectedness of constructions (Section 5.1.1), the other explores which kind of relations emerge, and which types of constructions tend to be linked (Section 5.1.2).

Subgroup	Metrics	Score
Quantity of connections	Average degree	5.87
Quantity of connections	Average Degree (Weighted)	15.47
Quantity of connections	Transitivity	0.68
Quantity of connections	Average Clustering	0.65
Quantity of connections	Average Clustering (Weighted)	0.19
Nature of connections	Attribute assortativity - event type	-0.11
Nature of connections	Attribute assortativity - event type (Weighted)	-0.28
Nature of connections	Attribute assortativity - complexity	-0.12
Nature of connections	Attribute assortativity - complexity (Weighted)	-0.21

Table 5: Scores for the network metrics.

5.1.1 Number and strength of the connections in the network

The network of psych-predicates has an average degree of 5.87, meaning that each construction is linked on average to six other constructions. Such connections give rise on average to approx. 15 lower-level alternations for each semi-schematic construction. In other words, the predicates created by each semi-schematic constructions enter in 15 alternations.

However, the raw numbers of paradigmatic relations are not very useful in absence of a standard. Instead, what we get from clustering metrics is more easily interpretable. Both transitivity and unweighted average clustering show a similar score, which points at a very high ratio of triangles, both globally (transitivity) and locally (average clustering). This means that if we encounter two constructions sharing fillers, we have an approx. 65% probability that they will be both be linked to another constructions. A high clustering score points at a high cohesion of the network: generally, constructions in our network tend to be intercon-

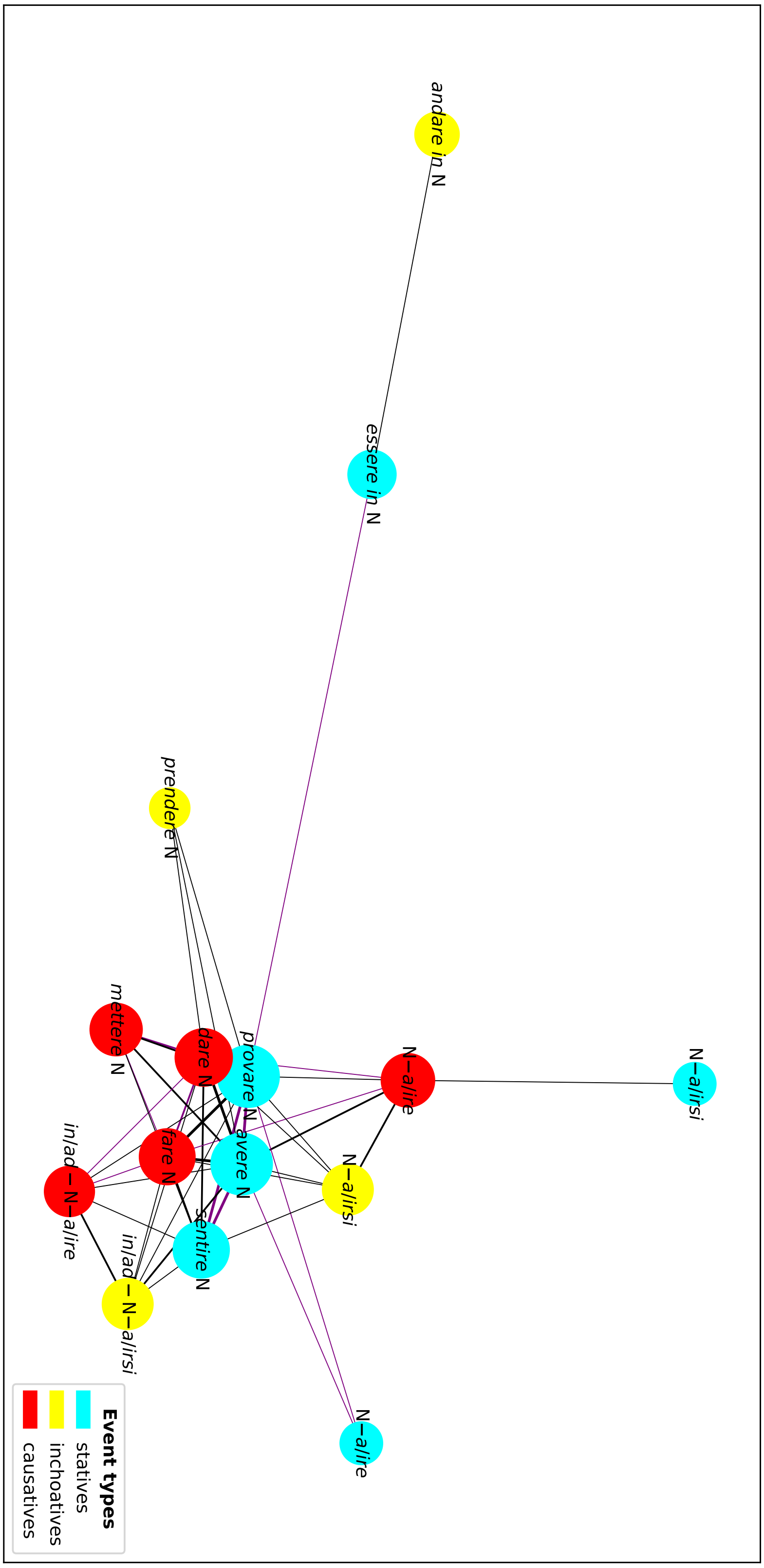

Figure 3: The network of denominal psych-predicates.

nected. This can also be seen by examining the maximal cliques found (Table 6). In fact, while we do have smaller cliques comprising two or three constructions (e.g., *essere in* N and *andare in* N, causative and anticausative converted verbs), almost half of the cliques comprise 6 constructions (40% of the total number of constructions).

Clique	**Weight sum**
provare N, *essere in* N	4
N-*a/irsi*(stative), N-*a/ire*(causative)	5
andare in N, *essere in* N	6
provare N, *avere* N, N-*a/ire*(stative)	45
provare N, *avere* N, *dare* N, *prendere* N	86
provare N, *avere* N, *dare* N, *fare* N, N-*a/ire*(causative), N-*a/irsi*(inchoative)	186
provare N, *avere* N, *dare* N, *fare* N, *sentire* N, N-*a/irsi*(inchoative)	217
provare N, *avere* N, *dare* N, *fare* N, *sentire* N, *mettere* N	235
provare N, *avere* N, *dare* N, *fare* N, *in/ad*-N-*a/ire*(causative), *in/ad*-N-*a/irsi*(inchoative)	265

Table 6: Maximal cliques found in the network.

A dense presence of links between constructions can be observed when we look at Figure 3: it seems that almost all of the constructions occupy the core of the graph, while the periphery of the network, formed by one-to-one relations, is way less populated. This situation is nicely captured by the extraction of Louvain communities (Table 7). The network is split in three clusters: on the one hand, we have two peripheric relations between structurally similar constructions, namely: 1) *andare in* N and *essere in* N, which share the same [V Prep N] structure, and 2) causative conversion and its anticausative (stative and inchoative) counterparts; on the other hand, there is the cluster of interconnected constructions at the core of the network.

Community	**Weight sum**
andare in N, *essere in* N	6
N-*a/ire*(causative), N-*a/irsi*(inchoative), N-*a/irsi*(stative)	16
fare N, *avere* N, *provare* N, *in/ad*-N-*a/ire*(causative), *prendere* N, *dare* N, *mettere* N, *sentire* N, *in/ad*-N-*a/irsi*, N-*a/ire*(stative)	342

Table 7: Louvain communities in the network.

However, the amount of links between the central cluster of constructions does not automatically mean strong interconnectedness, and this is shown by weighted average clustering: in fact, the low score seems to reveal that many triangular connections found are not very strong. This means that even though the network is closely interconnected, many of the links in our network correspond to weaker generalizations[6]. Thus, getting back to Louvain communities, we could hypothesize that so many constructions were placed together in the same community because most of them do not share enough fillers to form separate clusters. In this sense, a densely connected network does not automatically imply tight connections in that network.

5.1.2 Nature of the relationships in the network

After this overview on the number and strength of connections in our network, we get to the **assortativity scores**, which tell us more about the qualitative tendencies in the relational behavior of constructions. As mentioned, we calculate assortativity based on both event types and construction types.

As for **event types**, the unweighted score shows a low negative score, and thus we could say that quasi-synonymy and contrast relationships between semi-schematic constructions have a comparable frequency in our network. However, the weighted score reveals a moderate tendency towards contrast relations. This means that most of the low-level alternations in our dataset are aimed at expressing different event types, and thus predicate alternations show a tendency towards the expression of paradigmatic contrast, rather than the expression of quasi-synonyms.

The relevance of contrast relations in our network can also be seen by plotting separate graphs for links between same and different event types (Figure 4)[7]. In fact, we clearly see that almost all of the constructions are interconnected by links expressing semantic contrast (apart from stative conversion that constituites an isolated node, and the two linked constructions *andare in* N and *essere in* N). Instead, semantic similarity creates separate networks of stative and causative predicates, but totally disconnects inchoative constructions that become isolated nodes. This means that links expressing paradigm-like, contrast relationships play a pivotal role in shaping the domain of denominal predicates expression. Moreover, the expression of semantic contrast seems to motivate the inclusion in the network of inchoative predicates. Nevertheless, the moderate assortativity scores tell us that there is still a good number of synonymous constructions, representing potential rivals for the expression of the same meaning. However, the creation of rival constructions is mainly found in the analytic domain. As we see by calculating assortativity separately for analytic and synthetic constructions, synthetic constructions are almost only employed to create paradigmatic alternations, while the situation is more “mixed” in the analytic one (Table 8).

We also calculated assortativity by **construction types** (with respect to their structural complexity). In this case too, the scores go from lowly to moderately negative. This means that there is a low-moderate tendency in projecting links

[6] Obviously, this has to be considered in relation with very productive alternations in our dataset, such as the one between *avere* N and *provare* N, which are used interchangeably with 34 fillers.

[7] In this case we employed a shell layout, which does not take into account edge weights for positioning nodes. We made this choice to make the visualization clearer.

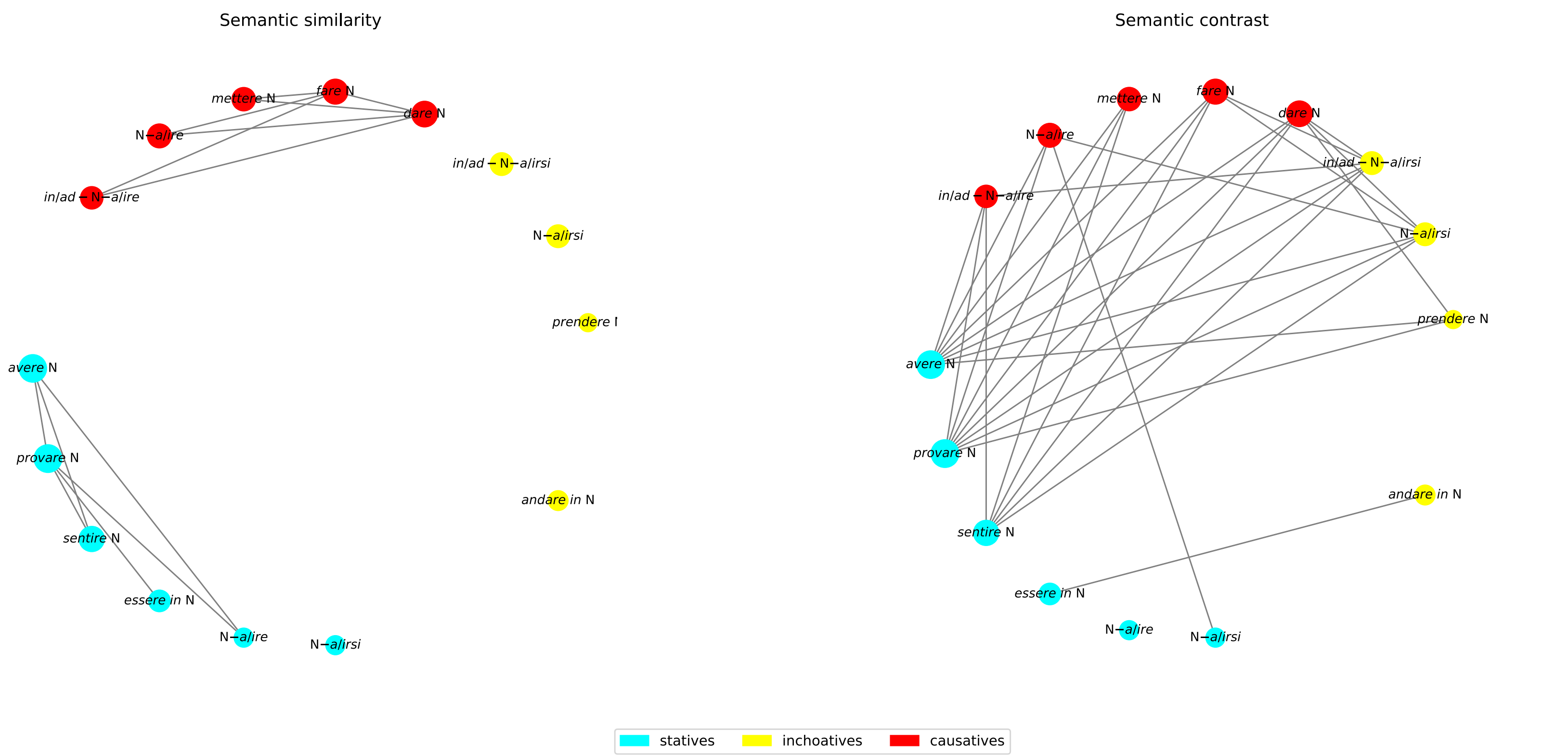

Figure 4: Subgraphs comprising links between constructions expressing the same event type (semantic similarity) and different event types (semantic contrast).

Structural complexity	Assortativity - event types	Assortativity (Weighted) - event types
Analytic cxns	-0.12	-0.36
Synthetic cxns	-0.63	-0.73

Table 8: Assortativity scores for networks including respectively analytic and synthetic constructions only.

between synthetic and analytic constructions, rather than between constructions of the same structural type, which seems to go in the direction of a mixed network.

However, this behavior is not equally true for all the type of connections in our network, nor for all the semantic domains. Getting back to the distinction between semantic similarity and contrast, we can calculate assortativity separately for the subgraphs including paradigm-like relations on one side, and quasi-synonymy on the other. By looking at unweighted scores, in contrast relations we find no tendencies in any direction, and instead the presence of links across different structural types is more significantly present when expressing the same event type.

Nonetheless, such a tendency is mitigated by taking into account the number of shared fillers: this probably means that some of the frequent synthetic/analytic relations among quasi-synonymous constructions involve a relatively low number of fillers, while in the case of semantic contrast, such relations, though rarer, involve more fillers.

Semantic relations	Assortativity - complexity level	Assortativity (Weighted) - complexity level
Semantic similarity	-0.29	-0.24
Semantic contrast	-0.08	-0.23

Table 9: Assortativity scores for networks including similarity and contrast relations only.

This tendency is not to be attributed to all the semantic subdomains: in fact, the attraction between synthetic and analytic constructions is mostly found in the causative domain (Table 10), while it is quite weak in the stative domain (especially by looking at weighted scores). This supports the observation made by Pisciotta & Masini (2025) that the causative domain seems to be the most fruitful one to study competition between analytic and synthetic constructions.

Semantic sub-domains	Assortativity - complexity level	Assortativity (Weighted) - complexity level
Stative cxns	-0.20	-0.12
Inchoative cxns	0	0
Causative cxns	-0.4	-0.46

Table 10: Assortativity scores for networks describing the three event types (semantic subdomains).

5.1.3 Summary

The results of our network metrics unveiled the following properties:

- **Frequency of connections:** the predicate-formation schemas in our dataset are densely interconnected, and there is a high chance that pairs of constructions share fillers. Consequently, most of the constructions belong together in a core cluster, where we find both synthetic and analytic constructions; nonetheless, most of the connections do not involve many shared fillers. On the one hand, this behavior suggests that most of the schemas stand as an option for predicate formation and combine quite freely with psych-nouns; on the other hand, stable alternations seem to be lacking, at least at the core of the network.
- **Semantic contrast:** the network seems to be shaped more by relations of semantic contrast, rather than by relations of quasi-synonymy; this means that predicate formation, at least in the psych-domain, tends to be exploited to express different perspectives on psych-events. This is particularly true for synthetic predicates, while the creation of series of synonymous predicates is more frequent in the analytic domain. Moreover, contrast relations motivate the inclusion in the network of inchoative constructions, that tend to not share fillers, and seem to be more dependent on the existence of stative and causative constructions.
- **Synthetic/analytic coexistence:** in our network, analytic and synthetic patterns often share the same fillers; however, there is generally no systematic tendency to create relationships between constructions at different levels of complexity. In particular, contrast relationships (which shape paradigmatic series) do not tend to involve different structural types; instead, it is more common to find synthetic/analytic alternation for the expression of the same meaning, in particular in the causative domain.

5.2 Centrality of constructions (RQ2)

Centrality metrics help us uncover which constructions stand at the core of the network, and which of them have a pivotal role in shaping it (RQ2). Table 11 shows the ranking by degree centrality of the constructions in our dataset.

We can see that the most central constructions are two stative, analytic constructions, namely *provare* N and *avere* N. This result is coherent with the visualization provided in Figure 3, by using the Fruchtermann-Reingold algorithm. What does it mean that these two constructions are the most “central” in our network?

From a graph perspective, a high degree points at a high number of connections between *provare* N and *avere* N and the other constructions in the network. Moreover, since we are taking into account the weighted degree, it means that fully-specified constructions that instantiate both *provare* N and *avere* N enter in 152 alternations, respectively. This means that these two constructions are the ones that participate most often in alternations in the network. As shown in Table 11, degree centrality appears to strongly correlate with the number of types that

Construction	Event type	n. types	Degree (weighted)	Degree (unweighted)
avere N	stative	43	152	11
provare N	stative	45	152	12
dare N	causative	26	103	10
sentire N	stative	23	89	8
fare N	causative	22	85	9
mettere N	causative	15	52	5
in/ad-N-*a/irsi*	inchoative	13	48	6
N-*a/irsi*	inchoative	13	45	6
N-*a/ire*	causative	17	44	6
in/ad-N-*a/ire*	causative	12	44	6
prendere N	inchoative	5	14	3
N-*a/ire*	stative	6	11	2
essere in N	stative	10	10	2
andare in N	inchoative	7	6	1
N-*a/irsi*	stative	6	5	1

Table 11: Constructions ranked by weighted degree centrality.

instantiate a given construction (which, in our case, corresponds to the number of fillers). Theoretically speaking, degree centrality should not strictly depend on the number of types associated with a construction. For instance, it is conceivable that *avere* N, although instantiated by 43 types, is used with highly idiosyncratic fillers, which would not be widely shared with other constructions. We could rather say that a high weighted degree centrality depends both on the number of fillers associated with a construction and on how commonly those fillers are shared across other constructions in the network.

An example is offered by the two inchoative constructions *andare in* N and *prendere* N. Although they occur with a similar number of fillers, *andare in* N tends to select psych-nouns that are generally not found in other constructions, except for *essere in* N. Said nouns generally express "extreme" psychological states, often anxiety or fear-related, in which the speaker has a low dominance (16a). In contrast, *prendere* N, despite being instantiated by a small number of types, stands in paradigmatic relation with more constructions (*provare*, *avere*, and *dare* N), since it is used with fillers that are more frequent and central within the domain of psychological nouns (16b). This discrepancy results in the fact that *andare in* N's degree strictly depends on his number of types (since it only alternates with one constructions), while *prendere* N's degree is almost three times higher than its type frequency, a ratio closer to more central constructions.

(16) a. *andare in panico* 'panic' (lit. 'go into panic')
andare in paranoia 'become paranoid' (lit. 'go into paranoia')
andare in estasi 'go into raptures' (lit. 'go into ecstasy')
b. *prendere paura* 'become afraid' (lit. 'take fear')
prendere coraggio 'gain courage' (lit. 'take courage')
prendere interesse 'take interest'

That being said, in a network as densely interconnected as ours, we can expect that the greater the number of fillers found with a construction, the more likely it is that many of them will be shared with other constructions. Therefore, in our case, constructions with high degree centrality can be considered the most prototypical in the network for two reasons: 1) they are the most frequently and possibly most default constructions for lexicalizing a predicate from psychological nouns; and 2) they are (almost) always present as paradigmatic alternatives, regardless of which other construction is considered. The prototypicaliy of stative constructions, at least in the light verb domain, is not unexpected: studies on light verbs constructions have highlighted that stative light verbs are the most unmarked option when used with state nouns (such as psychological ones) (Ježek 2004; Mastrofini 2004; Pompei & Piunno 2023). In fact, they preserve both the actional properties of the nominal base (differently from inchoatives), and its argument structure (differently from causatives). Thus, statives analytic constructions are found at the core of the network as they constitute the most prototypical option to create a predicate in the psychological domain. Instead, inchoative constructions seem to be the relatively less central in the network, as the ranks 1-6 only include stative and causative constructions. This is in line with the observation that inchoative constructions are less interconnected than other constructions in the network, and stand as more marginal in the expression of psychological events.

Ranks 1-6 in Table 11 also show us that the most central constructions seem to be the analytic ones. In particular, [Verb Noun] constructions (exception made for *prendere* N) are more central than synthetic constructions and, notably, of [Verb in_{Prep} Noun] analytic constructions. Thus, even though there is indeed a tendency to have links between constructions at different levels of complexity (Section 5.1), the core of the network mainly includes constructions sharing the same structure. That being said, we find strong differences in the synthetic domain: causative/inchoative[8] parasynthesis and conversion have a weighted degree which is not dissimilar from *mettere* N, and seem to be part of the core, although as peripheric members; intead, stative synthetic constructions seem to be way less integrated in the network, and show scores that pattern with [Verb in_{Prep} Noun] light verb constructions and *prendere* N.

Betweenness centrality provides us with a slightly different picture (Table 12): the most central elements are both the constructions at the core of the network (*avere* N and *provare* N), but also constructions that "attract" isolated nodes in the network (Figure 3), namely *essere in* N, and causative conversion (N-*a/ire*). Notably, *provare* N, *essere in* N and causative conversion act as articulation points of the network: by removing them, the structure of the network would be dismantled in several subgraphs. Instead, all the other nodes show almost no betweenness: this means that they do not have a fundamental role in the connection

8 The higher degree of inchoative anticausative constructions could seem strange, as they are created from corresponding causative predicates (Section 2). This behavior is due to the obsolescence of former causative verbs, that used to take part in anticausative alternations, such as *innamorare* 'make fall in love'. Such verbs only exist as anticausatives in contemporary Italian: *innamorarsi* 'fall in love'.

of different pieces of the network (and thus do not have privileged one-to-one connections with any of the constructions).

Construction	**Event type**	**Betweenness (weighted)**	**Betweenness (unweighted)**
avere N	stative	45.5	0.11
provare N	stative	29.5	0.38
essere in N	stative	13	0.14
N-*a/ire*	causative	13	0.14
dare N	causative	1	0.06
sentire N	stative	0	0.01
N-*a/ire*	stative	0	0
N-*a/irsi*	stative	0	0
andare in N	inchoative	0	0
prendere N	inchoative	0	0
N-*a/irsi*	inchoative	0	0.004
in/ad-N-*a/irsi*	inchoative	0	0
fare N	causative	0	0.03
mettere N	causative	0	0
in/ad-N-*a/ire*	causative	0	0

Table 12: Constructions ranked by weighted betweenness centrality. Articulation points calculated in *igraph* are highlighted in yellow.

Thus, constructions with high betweenness can be seen as "core" in structuring the network. In particular, in our case they seem to hold together the three Louvain communities we found (Section 5.1): *avere* N and *provare* N are the most central constructions in the network (both for their degree and betweenness), standing inside the most populated community (the central core); moreover, they are both connected with causative conversion, which in turn is the articulation point connecting the cluster comprising anticausative converted constructions; finally, *provare* N is the only construction of the core connected with *essere in* N, that participates in the separate cluster with isolated *andare in* N. These results suggest a tripartite architecture of the network, and both *essere in* N and causative conversion are transit points that connect the core (which is mainly formed by [Verb Noun] analytic constructions) with two clusters that comprise, respectively, converted constructions (synthetic) and [Verb in_{Prep} Noun] constructions.

Overall, thus, both the centrality metrics we employed highlight that different structural types of constructions, though connected, seem to occupy different spots in the network, and seem to have a partially differing behavior. This does not only depend on the level of complexity, but also on the analytic pattern employed. Nonetheless, it should be noted that parasynthesis is still quite integrated in the [Verb Noun] analytic network, and, differently from conversion, does not project a subparadigm on its own.

5.3 Predictability of paradigmatic relationships (RQ3)

In order to evaluate the third point of our hypothesis, that is, the relationship of predictability between exponents of different event types, we calculated node diversity. We averaged diversity scores of nodes belonging to the same event type to give an overview of predictability relations in a possible paradigm of psych-predicates. Disaggregated scores for each node are provided in Appendix A.1. We show the scores in Table 13. The table can be read as follows: row names are "predictors" (since we know which exponents actualize those meanings), while column names are "predictees" (the event types whose exponent we want to predict, if there is one). The scores represent the amount of uncertainty in the prediction[9].

	> **stative**	> **inchoative**	> **causative**
stative >		0.23	0.71
inchoative >	0.56		0.62
causative >	0.89	0.24	

Table 13: Diversity averaged by event types of the starting nodes.

The least predictable relationship seems to be the one between causative and stative constructions, and vice versa. In fact, by looking at the unaveraged scores (Table 16 in Appendix A.1), we find that the only node showing no diversity in the stative/causative relationship is anticausative conversion, which only predicts causative conversion in the causative domain. An effect of such alternation is found also when trying to predict the corresponding stative construction when the causative one is conversion; however, the score (0.66) tells us that converted verbs are often in paradigmatic relation with other constructions outside of the anticausative alternation. An example is the predicate network for the noun *interesse* 'interest': starting from the stative predicate formed by anticausative conversion, we can easily guess the corresponding causative predicate, and we find no alternatives in the causative domain (17a). Nonetheless, if we were to predict which construction can be used to express a stative event by knowing that conversion is employed in the causative domain, we would find at least three alternatives (17b):

(17) *interesse* 'interest'

a. N-*a/irsi* (stative) > causative predicate
interessarsi 'be interested' > *interessare* 'interest, concern'

b. N-*a/ire* (causative) > stative predicate
interessare 'interest, concern' > *interessarsi*, *provare interesse*, *avere interesse* 'be interested'

9 When a node has no connections to any other node expressing a specific event type, NaN is returned.

Instead, the most predictable relationships in the paradigm link statives and causatives with their inchoative counterparts. This tendency can be explained by two facts: 1) there are less competing constructions in the inchoative domain than in other event types, and 2) they seem to be expressed only secundarily in the paradigm, as they depend more, formally and/or semantically, on the other two cells. While this depends to a large extent on the anticausative alternation, and thus to the more regular/paradigm-like behavior of synthetic verbs (18a), we also find some relatively strong relations in the analytic domain that seem to depend on transparent lexical relations between light verbs. Two examples are *andare in* N > *essere in* N , that exploit an opposition in dynamicity in the spatial domain (GO vs. BE) to express the semantic opposition between inchoativity and stativity (18b), and *prendere* N > *dare* N, which exploit two different perspectives on change-of-possession events (TAKE vs. GIVE) to express the alternation between inchoativity and causation (cf. Nicoletti 2025) (18c).

(18) a. *calmare* ↔ *calmarsi*
'calm' ↔ 'calm down'
b. *essere in ansia* ↔ *andare in ansia*
'be anxious' ↔ 'get anxious'
c. *prendere coraggio* ↔ *dare coraggio*
'take courage' ↔ 'give courage'

These relations hold bidirectionally; however they are not as predictable when starting from inchoative patterns, since constructional competition is more developed in both the causative and the stative domain, and thus there are several possible paradigmatic alternatives. Nonetheless, diversity in $inchoative > stative$ and $inchoative > causative$ relations is still lower than between statives and causatives, and vice versa.

As suggested by the behavior of inchoative constructions, the most predictable relationships seem to arise between structurally similar constructions. This goes in the direction of different paradigmatic structures for constructions showing different levels of complexity. However, we should check if diversity actually gets lower by splitting graphs based on the analytic/synthetic distinction, or if this phenomenon only pertains to the relation involving inchoative constructions as *predictees*. In Tables 14-15 we show averaged diversity scores for the subgraphs that contain, respectively, synthetic constructions and analytic constructions only[10].

	> stative	**> inchoative**	**> causative**
stative >		NaN	0
inchoative >	NaN		0
causative >	0	0	

Table 14: Diversity averaged for event types (synthetic constructions subgraph).

10 See Tables 17-18 in Appendix (A.1) for disaggregated scores.

	> **stative**	> **inchoative**	> **causative**
stative >		0	0.97
inchoative >	0.49		0
causative >	0.95	0	

Table 15: Diversity averaged for event types (analytic constructions subgraph).

As we see in Table 14, the synthetic subgraph shows no diversity at all: in the case of relationships involving a causative construction (both as *predictor* and as *predictee*), there is a fully predictable relationship, provided by the anticausative alternation; instead, there are no links between constructions expressing stative and inchoative meanings. This suggests that synthetic constructions are structured as a strongly regular paradigm, which revolves around causative predicates, which play a pivotal role in determining their inchoative and stative counterparts. In fact, the two constructions employed to create causative verbs (i.e., parasynthesis and conversion) seem to stand in complementary distribution. The tendency of conversion and parasynthesis not to appear with the same bases has been already noted in Iacobini (2004), and is a consequence of the resolution of rivalry between competing predicates created from the same nouns (Iacobini & De Rosa 2024). This avoids the possibility of finding alternatives when transitioning from the inchoative/stative domain to the causative one, and vice versa, since inchoative or stative predicates are based on the causative one. However, not all synthetic constructions are part of the paradigm: stative conversion (N-*a/ire*) has no connections with other synthetic constructions[11], and thus is part of the network only due to its links with stative analytic constructions.

Turning to the analytic paradigm (Table 15), in this case too we find some positive effects in splitting the paradigms: also in this case relations involving inchoative exponents become more predictable. This is due to the complementary distribution between inchoative constructions *prendere* N and *andare in* N: in the stative domain, *avere* N and *provare* N are linked only to the former, while *essere in* N is systematically linked to *andare in* N. Instead, in the causative domain the only construction showing a consistent association with inchoatives is *dare* N, which is only connected with *prendere* N. The association between *essere in* N and *andare in* N lowers the score for $stative > incoative$ prediction; instead, *prendere* N can predict two different stative outcomes: *avere* N and *provare* N. Conversely, the relationship between statives and causatives seems to become even more unpredictable than in the case of a "mixed" network. This suggests that, at least in this domain, there are no advantages in considering analytic constructions only: unpredictable relationships in the network are not to be attributed to the interactions between different types of constructions in the the psych-predicate network; instead, unpredictability seems to be a feature of the analytic domain.

[11] Note that stative converted verbs can be part of paradigms at a lower-level of abstraction. For instance, in the case of the noun *schifo* 'disgust', both stative and causative converted verbs are created: *schifare* means both 'to feel disgust' and 'disgust'. However, this kind of relation is not found systematically, since it only applies to two nouns in the dataset, and thus this relationship was not included in our network.

We could ask to what extent unpredictability represents a problem for speakers in the analytic domain. In fact, by looking at the data, most of the nouns can be employed to form stative predicates both with *provare* N and with *avere* N. Thus, high diversity in this case does not prevent speakers from guessing which exponent will be employed to create stative predicates, since both the options generally provide acceptable results. Instead, when looking at the causative domain, overlaps and divergences between *fare* N, *dare* N, and *mettere* N are way less regular and predictable. Thus, predictions in the causative domain, at least when based on the stative exponents, are actually hardly generalized by the paradigmatic structure.

Summing up, our results show that there are indeed some regular and predictable paradigmatic niches, which include constructions sharing common structural features. Moreover, by analyzing analytic and synthetic constructions separately, we find diverging behaviors: synthetic verbs show systematic and regular associations between constructions, with a strict division of labor between word-formation schemas; similarly, [Verb $\textit{in}_{\text{Prep}}$ Noun] analytic constructions show a regular behaviour, since they are employed with very specific fillers; finally [Verb Noun] analytic constructions include some regular series, but overall, seem to be the most unpredictable, especially when looking at the causative domain.

5.4 How many paradigms of psych-predicates?

The results of the analysis reveal a complex picture, but they nonetheless allow us to formulate a number of observations.

First of all, our network appears to be densely populated with relations between constructions, which, however, are in most cases not particularly strong. These relations may involve either synonymy or semantic contrast, with the latter in particular accounting for the majority of connections among constructions in the network. In this sense, we might say that the semantic structure of our network resembles that of a paradigm: starting from a set of nouns, we observe a tendency to form predicates that can express alternative perspectives on the same psych-event. However, this does not happen with equal frequency for all event types: in fact, our analysis clearly highlighted the asymmetrical nature of the network, as we observed that inchoative semantics is generally less developed and indeed only secondarily expressed in this domain, whereas stative and causative meanings appear to be more prototypical for psych-predicates in Italian. This is not a problem for the hypothesis, since asymmetries in paradigms are expected (Section 2).

However, if we depart from the purely semantic assessment of the paradigm, and we take into account how frequently and regularly elements are in paradigmatic relation, things get more complex. Firstly, as we mentioned, most of the connections are not particularly strong; nonetheless, there are some clusters of constructions that stand out, at least when compared to the average behavior of comparable sub-networks. In particular, these clusters are characterized by the structural similarity of constructions: not only do we find an opposition between synthetic and analytic constructions, but also between different analytic schemas, namely [Verb Noun] and [Verb *in*$_{\text{Prep}}$ Noun]. We thus wonder whether this network can be seen as a tripartite structure instead of one coherent paradigm.

In fact, while it is true that synthetic and analytic constructions do share fillers, and, for instance, parasynthetic verbs are part of the cluster populated by analytic constructions, some other facts point in the direction of a modular structure of the network. Firstly, different structural types have different behaviors as for their connectedness and centrality in the network: [Verb Noun] constructions tend to have many connections between them and with other members, and instead synthetic constructions seem to be tangential to the core of the network, since they project less developed networks. Secondly, the network does not grant high predictability among paradigm exponents, apart from some regular alternations, that, again, seem to hold only between structurally similar constructions. While we admit the possibility of poor predictability in a non-inflectional paradigm, we also should prioritize splits that grant predictability in some subsets of the network. In other terms, if some subset behaves as a canonical paradigm, we are more inclined to assume that such a subset forms a paradigm on its own. As a matter of fact, when we split the network according to the level of complexity of constructions, we achieve more coherent results: on the one hand, synthetic constructions and [Verb *in*$_{\text{Prep}}$ Noun] have strong and predictable relations, while [Verb Noun] have a generally less predictable behaviour. Finally, we should also note that sub-networks of structurally different constructions form series that cover different portions of the semantic paradigm of event types: synthetic constructions mainly actualize the causative-inchoative alternation, and less often stative meanings; [Verb Noun] constructions mainly express stative and causative events, and only secundarily inchoative events; the two [Verb *in*$_{\text{Prep}}$ Noun] express a stative-inchoative alternation. The only common trait seems to be the secundary role of inchoative events in the psych-event domain.

Summing up, our network of structurally different constructions does not seem to form a paradigm in a narrow sense, but only in the semantic sense. In fact, if we lump together different structures we find no uniformity in their behaviour and in their integration in the network, and no regular connections between them; moreover, there is no clear complementary distribution between analytic and synthetic predicates, which could point at the differential exploitation of different strategies with certain noun subclasses. Thus, it seems more plausible to assume that different types of construction form separate paradigms.

Nonetheless, we should not forget that these separate paradigms overlap in some points in the network; some of them are articulation points, but we also find a consistent interplay between synthetic causative constructions and analytic causative constructions. Said synonymy relations are clearly recognized by speakers at a low level of abstraction, as shown by the relationship of paraphrasability between analytic and synthetic doublets in lexicographic sources (19):

(19) a. *impaurire* ∼ *mettere paura* 'frighten' (Treccani[12])
b. *infastidire* ∼ *dare fastidio* 'annoy' (Repubblica[13])

12 https://www.treccani.it/vocabolario/impaurire$_{(}Sinonimi-e-Contrari)/$

13 https://dizionari.repubblica.it/Italiano/I/infastidire.html

Moreover, both in the case of parasynthesis and conversion, relationships with analytic constructions seem to be quite systematic in the network, and thus should not be overshadowed. In fact, such paradigm-external relations allow the network to be kept together, with a moderate tendency for different constructional types to be linked together, as testified by assortativity scores.

Thus, our network is, in a sense, "modular": on the one hand, we have a strong core of (mainly) stative and causative analytic constructions, which can be considered as a paradigm on their own; on the other, there are some structurally diverse constructions that get attracted into the paradigm by means of their synonymy or regular paradigmatic contrast with some of the more integrated elements. In particular, synthetic constructions are found at the core of the network in the case of causative predicates, that definitely seem to be the actual field in which we find the most consistent coexistence and interplay between analytic and synthetic strategies. Thus, the network of denominal psych-predicates is a complex network of paradigms, which comprise synonymous and contrasting constructions; such paradigms get to be linked together because of their mapping with a shared set of semantic values, and thanks to the relatively strong relationship of paraphrasability between constructions at different levels of complexity.

6 Theoretical considerations

Our analysis of the results showed that graph metrics can be a useful tool to analyze relations between constructions, and to describe their behavior, if complemented with a coherent linguistic interpretation. In the following sections, we will discuss to what extent a graph-theoretic perspective can provide insights for constructional theory, along with the limitations and (sometimes necessary) simplifications of our model.

6.1 Putting relations center-stage in CxG

We believe that graph-theoretic modeling of constructional networks can be a valuable resource in a construction grammarian's toolkit, for both theoretical and methodological considerations.

From a theoretical standpoint, graph theory seems to be naturally compatible with CxG, starting from the definition of its architecture. This is evident if we compare the mathematical definition of graphs (Section 3.1) with Goldberg's (1995) definition of basic units of constructional knowledge: both constructions/nodes and links/edges are treated as basic objects in the model, which can carry attributes/weights that concur to define the structure of the Constructicon/set of nodes.

However, even though relations between constructions are of primary importance in CxG, much of constructionist research has focused primarily on the description and analysis of the nodes, following a traditional descriptive or inferential approach to linguistic phenomena (though with different assumptions). On par, many of the statistical models employed actually help us discerning the behavior of different constructions by contrasting them. However, contrasting constructional behavior is only one of the (indirect) ways to assess paradigmatic relations (here intended in a broad sense) between linguistic signs. Under this approach, one generally tries to deduct network properties starting from its nodes' internal properties.

Graph modeling of constructional networks can help us fill this analytical (and theoretical) gap, in that it puts centre-stage the relational nature of linguistic knowledge (Jackendoff & Audring 2020; Aronoff & Sims 2023; Budel et al. 2023), as compared to other corpus-based methods (Shadrova 2022). In fact, while it is true that for graph modeling we sometimes rely on construction-internal information to define what makes two constructions linked (in our case the selected fillers), once the theoretical assumptions on the model are drawn, what we receive from network analysis is mostly relational/construction-external information.

If we look at our case, the measures we took into account provided us with a description of how, how frequently, and which constructions are connected, and so on. In this, we did not directly receive any information on construction-internal properties, apart from some quantitative data on slot fillers, and we were able to operationalize information that is meaningful only in relative terms, such as the notion of prototypicality in paradigms. By taking this approach, we traveled the opposite path with respect to the traditional one, as we started from a description of the relational behavior of constructions, and tried, in our discussion, to find construction-internal, e.g., formal, semantic, frequency related, features that could explain our network's topology.

It is clear that we do not want to imply any primacy of relations over constructions, as graph properties are often the *explanandum* and not the *explanans*. Nonetheless, relational information should be taken into account seriously and employed jointly with construction-internal information to better inform our knowledge about a specific phenomenon. A possible way to do this is to evaluate information on constructional network's structure against psycholinguistic and diachronic models to check whether deduction based on experimental and descriptive studies are confirmed. In particular, there are many phenomena that have been attributed to the nature of links between constructions, such as strenghtening (Hilpert 2021), attraction (De Smet et al. 2018), contamination (Pijpops & Van de Velde 2016), and so on, which could be assessed by designing an actual model of the constructional network analyzed. As an example, our interpretation of the metrics suggests that inspecting articulation points and betweenness in a graph-based representation could prove to be very useful to assess phenomena of attraction and constructional change in diachronic analyses, since these concept can be used to isolate constructions acting as bridges in wider networks.

The advantages of having a model of the constructional network come, however, with the burden of a deep theoretical discussion to validate such model. In fact, while construction-internal phenomena rely on quite solid empirical foundations (also because of the wide literature on the matter), construction-external relations are part of our modeling of language knowledge, i.e., they are not explicitly marked in the data, as they are mostly paradigmatic in nature. This makes our representation and analysis of the network more abstract and less statistically powerful, and the interpretation potentially less solid and meaningful. For this reason, a detailed and motivated reflection on the correspondence between measures and linguistic interpretation is needed. In this paper, we tried to move some steps in that direction, keeping in mind that the interpretation of network measures is variable, based on the nature of the phenomenon. Nonetheless, there are some shortcomings in our model, which let emerge some points for future discussions.

6.2 Limitations of the model

Though the model we proposed provided us with promising results, there definitely is room for improvement, and, in particular, theoretical discussion is needed to refine structural and representational aspects of the model. In fact, there are a number of factors that were not taken into account, which could have shaped the network in more realistic terms. Some of them, however, come at the cost of a higher complexity in computational and interpretational terms.

Firstly, when operationalizing the edges, we should remember that not all types have the same frequency, and thus we should probably assign weights to different bases, as some of them could be more salient than others as fillers. However, we should be careful when weighting fillers by their token frequency: firstly, we do not have clear and reliable frequency thresholds for salience; moreover, very frequent types are generally assumed as being stored separately from their mother nodes, and thus are more likely not to be processed by speakers as types of said mother nodes (Barðdal 2008). This would raise the question whether frequent types really strengthen paradigmatic generalizations or not.

Another related observation is that, as noted by Cappelle et al. (2023), shared bases might tell us only part of the story. Even though fillers seem an intuitive and objective way to look for relatedness between constructions, we should probably craft the model by taking into account richer information, such as the semantic similarity of bases that are not shared.

A third point is the absence of vertical links in our network: in the case of light verbs constructions, we should have included two abstract nodes, namely [V N] and [V Prep N]. This, however, would have led to complex theoretical and technical choices. In fact, vertical links could be seen as directed ones (as they clearly imply a direction from a construction to another), but it is unclear whether horizontal and vertical links should be assigned similar weights, and how strength in vertical links could be operationalized when we do not have information on all the daughter nodes. Moreover, this would require different techniques, since the presence of directed and undirected edges in a graph would qualify it as a mixed graph, and the construction of such types of graph are generally not supported in Python libraries.

A fourth and final point is the possible presence of more than one link between two constructions: for instance, the formal similarity between some of the constructions could influence their closeness in the constructional network, as they would be recognized as “more similar” by speakers. This could be captured via another horizontal link, thus redefining our graph as a multigraph (i.e., a graph where is permetted to have multiple edges between two nodes). Obviously, the way in which such a formal similarity should be expressed has to be discussed case by case, depending on the variables taken into account.

Thus, much work has yet to be done in assessing and validating graphs as more realistic representations of constructional networks. However, we believe that the possible advantages of employing graph theory make the use of such a formalism appealing for constructional research.

7 Conclusions

In this contribution, we proposed a structured account of constructional networks in graph-theoretic terms, and applied it to analyze a network of Italian psych predicates. In particular, we employed graph metrics and network analysis to assess whether analytic and synthectic constructions employed to create denominal psych-predicates belong in the same paradigm, defined as a set of recurring and regular contrast relations. The results reveal that structurally different types of constructions show different types of relational behaviors, and, while there is some tendency to create links across the syntax-lexicon continuum, coherent and regular contrast relations seem to hold mainly between constructions at the same level of complexity (synthetic/analytic). Graph formalization proved to be a useful tool to evaluate (and partially refute) our initial hypothesis, as it provided a set of clearly defined measures to evaluate the way in which constructions are interconnected.

In light of our analysis, we believe that CxG would benefit from the use of network analysis for a number of reasons. Firstly, graphs are a useful descriptive tool for visualization, and are flexible enough to encode the type of information needed, according to different theoretical considerations and modeling choices. Moreover, it seems a natural option for CxG, as it constitutes a theoretically compliant and informed model, and since it emphasizes the relational and paradigmatic nature of linguistic knowledge. However, research is needed to capture which aspects of the constructional network should be encoded in a graph-based model (e.g., what should be encoded in a link), and the cognitive plausibility of such a modeling. We believe, however, that graph modeling can be readily adapted to models resulting from both psycholinguistic and diachronic investigations, and can therefore be directly evaluated against experimental evidence.

In conclusion, we hope that our study will foster research in the direction of graph modeling in CxG, and that research will turn its attention towards reaching a partial, though adequate, model of linguistic knowledge.

Acknowledgements

I would like to thank Jenny Audring for posing me a question about my data on psych-predicates, which I couldn't answer at the time. That question led me to explore network representations and graph theory to provide an adequate description of my dataset, and was the starting point for the work presented here. I am also grateful to Federico Pisciotta for carefully checking the mathematics in the paper. Finally, I am thankful to the editors and the anonymous reviewers for their insightful comments and suggestions, which helped refine and significantly expand the scope of the article. The responsibility for all the remaining errors is mine.

Conflict of interest statement

The authors declare none.

Data availability statement

Data and analysis scripts for the present paper are available at https://doi.org/10.5281/zenodo.16729050.

License

References

Ackerman, Farrell, James P. Blevins & Robert Malouf. 2009. *Parts and wholes: Implicative patterns in inflectional paradigms* 54–82. Oxford: Oxford University Press. doi:10.1093/acprof:oso/9780199547548.003.0003.

Ackerman, Farrell & Robert Malouf. 2013. Morphological organization: The low conditional entropy conjecture. *Language* 89(3). 429–464. http://www.jstor.org/stable/24671935.

Ackerman, Farrell & Gregory Stump. 2004. Paradigms and periphrastic expression: A study in realization-based lexicalism. In Louisa Sadler & Andrew Spencer (eds.), *Projecting morphology*, 111–157. Stanford: CSLI Publications.

Araújo, Tanya & Sven Banisch. 2016. Multidimensional Analysis of Linguistic Networks. In Alexander Mehler, Andy Lücking, Sven Banisch, Philippe Blanchard & Barbara Job (eds.), *Towards a Theoretical Framework for Analyzing Complex Linguistic Networks*, 107–131. Berlin/Heidelberg: Springer. doi:10.1007/978-3-662-47238-5_5.

Aronoff, Mark & Andrea D. Sims. 2023. The relational nature of morphology. In Davide Crepaldi (ed.), *Linguistic morphology in the mind and brain*, 7–25. London: Routledge. doi:10.4324/9781003159759-2.

Audring, Jenny. 2019. Mothers or sisters? The encoding of morphological knowledge. *Word Structure* 12(3). 274–296. doi:10.3366/word.2019.0150.

Ausiello, Giorgio, Donatella Firmani & Luigi Laura. 2013. The (betweenness) centrality of critical nodes and network cores. In *2013 9th International Wireless Communications and Mobile Computing Conference (IWCMC)*, 90–95. doi:10.1109/IWCMC.2013.6583540.

Banisch, Sven, Philippe Blanchard, Barbara Job, Andy Lücking & Alexander Mehler (eds.). 2016. *Towards a Theoretical Framework for Analyzing Complex Linguistic Networks*. Berlin/Heidelberg: Springer. doi:10.1007/978-3-662-47238-5.

Baroni, Marco, Silvia Bernardini, Adriano Ferraresi & Eros Zanchetta. 2009. The WaCky wide web: a collection of very large linguistically processed web-crawled corpora. *Language Resources and Evaluation* 43(3). 209–226. doi:10.1007/s10579-009-9081-4.

Barðdal, Johanna. 2008. *Productivity: Evidence from Case and Argument Structure in Icelandic*. Amsterdam & Philadelphia: John Benjamins Publishing Company. doi:10.1075/cal.8.

Bonami, Olivier & Sacha Beniamine. 2016. Joint predictiveness in inflectional paradigms. *Word Structure* 9(2). 156–182. doi:10.3366/word.2016.0092.

Bonami, Olivier & Jana Strnadová. 2019. Paradigm structure and predictability in derivational morphology. *Morphology* 29(2). 167–197. doi:10.1007/s11525-018-9322-6.

Booij, Geert. 2010. *Construction morphology*. London, England: Oxford University Press.

Booij, Geert & Francesca Masini. 2015. The role of second order schemas in the construction of complex words. In Laurie Bauer, Lívia Körtvélyessy & Pavol Štekauer (eds.), *Semantics of complex words*, 47–66. Cham: Springer International Publishing. doi:10.1007/978-3-319-14102-2_4.

Brandes, Ulrik. 2001. A faster algorithm for betweenness centrality. *The Journal of Mathematical Sociology* 25(2). 163–177. doi:10.1080/0022250x.2001.9990249.

Budel, Gabriel, Ying Jin, Piet Van Mieghem & Maksim Kitsak. 2023. Topological properties and organizing principles of semantic networks. *Scientific Reports* 13(1). doi:10.1038/s41598-023-37294-8.

Butt, Miriam. 2003. The light verb jungle. *Harvard Working Papers in Linguistics, Papers from the GSAS/Dudley House Workshop on Light Verbs* 9. 1–49.

Bybee, Joan L. 1985. *Morphology: A Study of the Relation between Meaning and Form*. Amsterdam: John Benjamins Publishing Company. doi:10.1075/tsl.9.

Cappelle, Bert. 2006. Particle placement and the case for "allostructions". *Constructions SV* 1(7). doi:10.24338/CONS-381.

Cappelle, Bert, Robert Daugs & Stefan Hartmann. 2023. The english privative prefixes near-, pseudo- and quasi-: Approximation and "disproximation". *Zeitschrift für Wortbildung / Journal of Word Formation* 7(1). doi:10.21248/zwjw.2023.1.35.

Cetnarowska, Bożena. 2021. Derivational paradigms or paradigms of function? Competition between Polish affixal formations, morphological compounds and phrasal lexemes. Presented at the Second Workshop on Paradigmatic Word Formation Modelling (ParadigMo II), 3–4 June 2021, Bordeaux, France.

Croft, William. 1991. *Syntactic categories and grammatical relations: The cognitive organization of information*. Chicago: University of Chicago Press.

Csardi, Gabor & Tamas Nepusz. 2006. The igraph software package for complex network research. *InterJournal* 1695. https://igraph.org.

De Mauro, Tullio (ed.). 2007. *Grande dizionario italiano dell'uso*. Torino: UTET.

De Smet, Hendrik, Frauke D'hoedt, Lauren Fonteyn & Kristel Van Goethem. 2018. The changing functions of competing forms: Attraction and differentiation. *Cognitive Linguistics* 29(2). 197–234. doi:10.1515/cog-2016-0025.

Desagulier, Guillaume. 2017. *Corpus Linguistics and Statistics with R: Introduction to Quantitative Methods in Linguistics*. Cham: Springer International Publishing. doi:10.1007/978-3-319-64572-8.

Desagulier, Guillaume. 2020. Graph theory and corpus linguistics. *Around the word* doi:10.58079/N4UZ. https://corpling.hypotheses.org/3265.

Desagulier, Guillaume. 2021. Capturing horizontal links in a construction network using semantic vector spaces. In *Modelling constructional variation and change, 15-16 November, Universität Zürich, Zürich, Switzerland*, https://shs.hal.science/halshs-03435830/file/Desagulier_Zurich_2021.pdf.

Diessel, Holger. 2019. *The grammar network: How linguistic structure is shaped by language use*. Cambridge: Cambridge University Press. doi:10.1017/9781108671040.

Diessel, Holger. 2023. *The constructicon: Taxonomies and networks*. Cambridge: Cambridge University Press. doi:10.1017/9781009327848.

Eagle, Nathan, Michael Macy & Rob Claxton. 2010. Network diversity and economic development. *Science* 328(5981). 1029–1031. doi:10.1126/science.1186605.

Ellis, Nick C., Matthew Brook O'Donnell & Ute Römer. 2013. Usage-Based Language: Investigating the Latent Structures That Underpin Acquisition. *Language Learning* 63(s1). 25–51. doi:10.1111/j.1467-9922.2012.00736.x.

Ellis, Nick C., Matthew Brook O'Donnell & Ute Römer. 2014. The processing of verb-argument constructions is sensitive to form, function, frequency, contingency and prototypicality. *Cognitive Linguistics* 25(1). 55–98. doi:10.1515/cog-2013-0031.

Estrada, Ernesto. 2016. When local and global clustering of networks diverge. *Linear Algebra and its Applications* 488. 249–263. doi:10.1016/j.laa.2015.09.048.

Estrada, Ernesto & Philip A. Knight. 2015. *A First Course in Network Theory*. Oxford: Oxford University Press.

Fradin, Bernard. 2018. Paradigms and the role of series in derivational morphology. *Lingue e Linguaggio* 17(2). 155–171. doi:10.1418/91863.

Fradin, Bernard. 2020. Characterizing derivational paradigms. In Jesús Fernández-Domínguez, Alexandra Bagasheva & Cristina Lara Clares (eds.), *Paradigmatic relations in word formation*, 49–84. Leiden: Brill.

Fruchterman, Thomas M. J. & Edward M. Reingold. 1991. Graph drawing by force-directed placement. *Software: Practice and Experience* 21(11). 1129–1164. doi:10.1002/spe.4380211102.

Goldberg, Adele. 1995. *Constructions: A construction grammar approach to argument structure*. Chicago: University of Chicago Press.

Goldberg, Adele. 2006. *Constructions at Work: The Nature of Generalization in Language*. Oxford: Oxford University Press. doi:10.1093/acprof:oso/9780199268511.001.0001.

Gries, Stefan Th. & Nick C. Ellis. 2015. Statistical Measures for Usage-Based Linguistics. *Language Learning* 65(S1). 228–255. doi:10.1111/lang.12119.

Hagberg, Aric A., Daniel A. Schult & Pieter J. Swart. 2008. Exploring network structure, dynamics, and function using NetworkX. In Gaël Varoquaux, Travis Vaught & Jarrod Millman (eds.), *Proceedings of the 7th python in science conference*, 11 – 15. Pasadena, CA USA. http://conference.scipy.org/proceedings/SciPy2008/paper_2/.

Hathout, Nabil & Fiammetta Namer. 2019. Paradigms in word formation: what are we up to? *Morphology* 29(2). 153–165. doi:10.1007/s11525-019-09344-3.

Hilpert, Martin. 2021. The upward strengthening hypothesis. In *Ten lectures on diachronic construction grammar*, 229–247. Leiden: Brill. doi:10.1163/9789004446793_010.

Hoffmann, Thomas & Graeme Trousdale (eds.). 2013. *The Oxford Handbook of Construction Grammar*. Oxford: Oxford University Press. doi:10.1093/oxfordhb/9780195396683.001.0001.

Iacobini, Claudio. 2004. Parasintesi. In Maria Grossmann & Franz Rainer (eds.), *La formazione delle parole in italiano*, 165–188. Tübingen: Niemeyer.

Iacobini, Claudio & Maria Pina De Rosa. 2024. A diachronic perspective on competition in denominal verb formation in Italian. In Alexandra Bagasheva, Akiko Nagano & Vincent Renner (eds.), *Competition in word-formation*, 247–274. Amsterdam: John Benjamins Publishing Company. doi:10.1075/la.284.08iac.

Jackendoff, Ray & Jenny Audring. 2020. Relational morphology: A cousin of Construction Grammar. *Frontiers in Psychology* 11. doi:10.3389/fpsyg.2020.02241. http://dx.doi.org/10.3389/fpsyg.2020.02241.

Ježek, Elisabetta. 2004. Types et degrés de verbes supports en italien. *Lingvisticae Investigationes* 27(2). 185–201. doi:10.1075/li.27.2.04jez.

Leino, Jaakko. 2022. Formalizing paradigms in Construction Grammar. In Gabriele Diewald & Katja Politt (eds.), *Paradigms regained: Theoretical and empirical arguments for the reassessment of the notion of paradigm*, 37–65. Berlin: Language Science Press. doi:10.5281/ZENODO.5675843.

Masini, Francesca. 2019. Competition between morphological words and multiword expressions. In Wolfgang U. Dressler Hans Christian Luschützky Franz Rainer, Francesco Gardani (ed.), *Competition in inflection and word-formation*, 281–305. Cham: Springer International Publishing. doi:10.1007/978-3-030-02550-2_11.

Mastrofini, Roberta. 2004. Classi di costruzioni a verbo supporto in italiano: implicazioni semantico- sintattiche nel paradigma v + n. *Studi Italiani di Linguistica Teorica e Applicata* 33. 371–398.

Melloni, Chiara & Serena Dal Maso. 2022. For a topology of derivational paradigms. In Alba E. Ruz, Cristina Fernández-Alcaina & Cristina Lara-Clares (eds.), *Paradigms in Word Formation: Theory and applications*, vol. 225, 21–56. Amsterdam: John Benjamins Publishing Company. doi:10.1075/slcs.225.02mel.

Morin, Cameron, Guillaume Desagulier & Jack Grieve. 2024. A social turn for Construction Grammar: double modals on British Twitter. *English Language and Linguistics* 28(2). 275–303. doi:10.1017/S1360674323000576.

Newman, Mark E. J. 2002. Assortative Mixing in Networks. *Physical Review Letters* 89(20). 208701. doi:10.1103/PhysRevLett.89.208701.

Newman, Mark E. J. 2003. Mixing patterns in networks. *Physical Review E* 67(2). doi:10.1103/physreve.67.026126.

Nicoletti, David. 2025. Constructional semantics of Italian light verb constructions. In Anna Riccio & Jens Fleischhauer (eds.), *Light verbs: Synchronic and diachronic studies*, 199–232. Berlin/Boston: Düsseldorf University Press (De Gruyter). doi:10.1515/9783111388878-008.

Passaro, Lucia C., Laura Pollacci & Alessandro Lenci. 2015. ItEM: A vector space model to bootstrap an italian emotive lexicon. In *Proceedings of the second italian conference on computational linguistics clic-it 2015*, 215–220. Accademia University Press. doi:10.4000/books.aaccademia.1530.

Pijpops, Dirk & Freek Van de Velde. 2016. Constructional contamination: How does it work and how do we measure it? *Folia Linguistica* 50(2). 543–581. doi:10.1515/flin-2016-0020.

Pisciotta, Flavio & Francesca Masini. 2025. A paradigm of psych-predicates. Unraveling the constructional competition between light verb constructions and derived verbs in Italian. In Anna Riccio & Jens Fleischhauer (eds.), *Light verbs: Synchronic and diachronic studies*, 123–176. Berlin/Boston: Düsseldorf University Press (De Gruyter). doi:10.1515/9783111388878-006.

Plutchik, Robert. 1980. *Emotion: A psychoevolutionary synthesis*. New York: Harper Row,.

Pompei, Anna & Valentina Piunno. 2023. Light Verb Constructions in Romance languages. An attempt to explain systematic irregularity. In Anna Pompei, Lunella Mereu & Valentina Piunno (eds.), *Light Verb Constructions as Complex Verbs: Features, Typology and Function*, 99–148. Berlin/Boston: De Gruyter. doi:10.1515/9783110747997-005.

Roché, Michel. 2017. Les familles dérivationnelles: comment ça marche? Unpublished manuscript. https://hal.science/hal-04912269v1/file/Roch%C3%A9%202017%20Familles%20d%C3%A9rivationnelles.pdf.

Saramäki, Jari, Mikko Kivelä, Jukka-Pekka Onnela, Kimmo Kaski & János Kertész. 2007. Generalizations of the clustering coefficient to weighted complex networks. *Physical Review E* 75(2). doi:10.1103/physreve.75.027105.

Shadrova, Anna. 2022. It may be in the structure, not the combinations: Graph metrics as an alternative to statistical measures in corpus-linguistic research. In Tara Andrews, Franziska Diehr, Thomas Efer, Andreas Kuczera & Joris van Zundert (eds.), *Graph technologies in the humanities - proceedings 2020. vienna, austria, february 21-22*, 245–278. CEUR Workshop Proceedings. https://ceur-ws.org/Vol-3110/paper12.pdf.

Smirnova, Elena. 2021. Horizontal links within and between paradigms: The constructional network of reported directives in German. In *Modality and diachronic construction grammar*, 185–217. Amsterdam: John Benjamins Publishing Company. doi:10.1075/cal.32.07smi.

Sommerer, Lotte & Andreas Baumann. 2021. Of absent mothers, strong sisters and peculiar daughters: The constructional network of English NPN constructions. *Cognitive Linguistics* 32(1). 97–131. doi:10.1515/cog-2020-0013.

Sommerer, Lotte & Freek Van De Velde. 2025. Constructional Networks. In Mirjam Fried & Kiki Nikiforidou (eds.), *The Cambridge Handbook of Construction Grammar*, 220–246. Cambridge: Cambridge University Press. doi: 10.1017/9781009049139.010.

Talmy, Leonard. 2000. *Toward a cognitive semantics. typology and process in concept structuring*. Cambridge, MA: MIT Press. doi:10.7551/mitpress/6848.001.0001.

Trudeau, Richard J. 1994. *Introduction to graph theory*. Mineola, NY: Dover Publications.

Ungerer, Tobias. 2024. Vertical and horizontal links in constructional networks: Two sides of the same coin? *Constructions and Frames* 16(1). 30–63. doi: 10.1075/cf.22011.ung.

Van De Velde, Freek & Lauren Fonteyn. 2017. Degeneracy: the evolutionary advantage of the violation of isomorphism. In *50th Annual Meeting of the Societas Linguistica Europaea , 10-13 September, Universität Zürich, Zürich, Switzerland*, https://lirias.kuleuven.be/1821625&lang=en.

Watts, Duncan J. & Steven H. Strogatz. 1998. Collective dynamics of 'small-world' networks. *Nature* 393(6684). 440–442. doi:10.1038/30918.

Zammuner, Vanda Lucia. 1998. Concepts of emotion: "emotionness", and dimensional ratings of Italian Emotion Words. *Cognition and Emotion* 12(2). 243–272. doi:10.1080/026999398379745.

Štekauer, Pavol. 2014. Derivational paradigms. In Rochelle Lieber & Pavol Štekauer (eds.), *The Oxford Handbook of Derivational Morphology*, 354–369. Oxford: Oxford University Press.

A Appendix

A.1 Diversity

construction	**> statives**	**> inchoatives**	**> causatives**
avere N		0	0.95
provare N		0.95	0.96
sentire N		0	0.96
essere in N		0	NaN
N-*a/ire*(stative)		NaN	NaN
N-*a/irsi*(stative)		NaN	0
andare in N	0		NaN
prendere N	0.48		0
N-*a/irsi*(inchoative)	0.89		0.91
in/ad-N-*a/irsi*(inchoative)	0.89		0.97
dare N	0.97	0	
fare N	0.97	0.97	
mettere N	0.93	NaN	
N-*a/ire*(causative)	0.66	0	
in/ad-N-*a/ire*	0.88	0	

Table 16: Node diversity in transitioning on contrasting event types.

construction	**> statives**	**> inchoatives**	**> causatives**
N-*a*/*ire*(stative)		NaN	NaN
N-*a*/*irsi*(stative)		NaN	0
N-*a*/*irsi*(inchoative)	NaN		0
in/*ad*-N-*a*/*irsi*(inchoative)	NaN		0
N-*a*/*ire*(causative)	0	0	
in/*ad*-N-*a*/*ire*	NaN	0	

Table 17: Node diversity (network of synthetic constructions only) in transitioning on contrasting event types.

construction	**> statives**	**> inchoatives**	**> causatives**
avere N		0	0.95
provare N		0	0.97
sentire N		NaN	0.99
essere in N		0	NaN
andare in N	0		NaN
prendere N	0.99		0
dare N	0.96	0	
fare N	0.97	NaN	
mettere N	0.93	NaN	

Table 18: Node diversity (network of analytic constructions only) in transitioning on contrasting event types.