2. An innovative approach to the study of verbs at the syntax-semantics interface

ELISABETTA JEŽEK

Keywords: *verb class; argument structure; lexical semantics; semantic composition; argument incorporation*

In this contribution we start by critically evaluating the pros and cons of syntactically-oriented studies that were published from the ’90s onward, with the goal of pinning down the semantic properties of verbs. First, we highlight the strengths of this approach. Second, we address its two major weaknesses: the inability to clearly distinguish arguments from adjuncts and with the shortcomings in accounting for the optionality in the syntactic realization of arguments. We argue that these weaknesses may be overcome by adopting a semantic-oriented approach from the start. Specifically, we propose that, unlike adjuncts, arguments are complements that play a role in influencing the meaning that the verb acquires in its contexts of use, and that syntactic optionality is licensed by semantic factors, in particular the degree to which an argument is incorporated in the verb on which it depends.

Original scientific paper

1. Introduction

Syntax-based approaches are an established model in theoretical linguistics for the study of verbs at the syntax-semantics interface. These approaches generally start by identifying subcategorization frames and their alternations for each verb, and then analyse them with the goal of identifying correlations between the syntactic behaviour and the semantic properties (Levin 1993). Despite the results that have been achieved, these approaches have shown not to be without problems. In this contribution we propose a complementary approach to the study of verbs at the syntax-semantic interface, in which the analysis is performed starting from the semantics of the verbs and that of their arguments. After reviewing the nature of the meaning of verbs and its primary constituents (section 2), we critically

evaluate the syntactically-oriented studies that have been systematically pursued from the ’90s onward, with the goal of pinning down the semantic properties of verbs (section 3). First, we highlight the strengths of this approach. Second, we address its two major weaknesses: the inability to clearly distinguish arguments from adjuncts and the difficulty to account for the optionality in the syntactic realization of arguments (section 4). We argue that these weaknesses may be overcome by adopting a semantic-oriented approach from the start (section 5). In particular, we propose that, unlike adjuncts, arguments are complements that play a role in influencing the meaning that the verb acquires in the context of use, and that syntactic optionality is licensed by semantic factors, in particular the degree to which an argument is incorporated in the verb on which it depends.

2. The meaning of verbs

The denotation of a verb includes at least three dimensions which, together, constitute different aspects of the same object, namely the event that the verb expresses.[1] These are: the temporal structure of the event, the participants of the event (the arguments or valency of the verb), and the inherent nature of the event. Together, these components constitute the three main dimensions of analysis that have been pursued in linguistic studies – jointly or individually – to classify verbs. We review them briefly in the next subsections.

2.1. The temporal structure of the event

The event expressed by the verb is placed in time and, if it is endowed with duration, it has an internal structure that can be broken down into phases or intervals. The analysis of these phases is generally implemented through the use of linguistic tests such as the adverbial modification introduced by the expressions ‘in/at *x* time’ and ‘for *x* time’. The applicability of the first would reveal that the event expressed by the verb includes a telic or culminative component; that of the second would exclude it. Based on this test, it is possible to distinguish verbs expressing *states* (*possess*, *remain*),

[1] In the following, we use the term *event* to refer to what Emmon Bach (1986: 6) called *eventualities*, i.e. static and dynamic situations.

processes (*work*, *sleep*, *walk*), all admitting 'for x time' but not 'in x time'; punctual or culminative events (*find*, *arrive*, *burst*) that admit 'at x time' but not 'for x time'; incremental or gradual events (degree achievements, incremental theme- or multi-scalar verbs: *increase*, *cool down*), which admit both adverbials. The event receives a specific temporal encoding when it becomes the denotation of a verb (i.e. its meaning). In principle, there may be a mismatch between an event "as it is observed in the world" and the way it is encoded in the meaning of verbs, although this is not the norm.[2]

2.2. Event participants

The event expressed by the verb usually has participants, i.e. animate/inanimate, concrete/abstract entities (including places and time periods) that hold a relation of *participation* with the event itself, playing different roles: the agent, the experiencer, the patient, the receiver, the origin, the destination, the instrument, etc. The linguistic encoding of the event makes some of these participants grammatically relevant: these are the participants called *arguments* or *actants*, which, contrary to other complements (the so-called *adjuncts* or *circumstantials*; for the latter, cf. Tesniere 1959), must be mandatorily expressed in the syntax because they complete the meaning of the verb. Depending on the number of arguments it is possible to distinguish zero-argument (*snow*, *flash*), one-argument (*sleep*), two-argument (*abolish*, *inhabit*, *dwell*), up to three-argument (*give*, *dedicate*) verbs. The existence of verbs with four arguments (*translate*) is a matter of controversy (Levin & Rappaport 2005). When adopting the notion of argument, an interesting question arises: which aspect(s) of the event may become an argument? Typically, participants or properties of participants in the event (including other events) may become arguments (such as the intentionality of the agent, the causes of the event, instruments, psychological states). Moreover, in the linguistic encoding of the event, some participants may be *incorporated* in the verb semantics, whereas other participants may be located in the background, an issue to which we return later.

Besides bearing a role, which is assigned to them by the verb, the arguments display semantic properties, in particular they belong to a

---

[2] For the inherent temporal properties of the event expressed by verbs, often referred to as *Aktionsart* or *Actionality*, see the taxonomies of Vendler (1967), Dowty (1979), and subsequent ones inspired by those.

semantic class (Physical object, Human, Institution, Animal, Artifact, Abstract entity etc.). This class must satisfy the *selectional restriction* (Chomsky 1965) or *preference* (Wilks 1975) that the verb requires: for example, the It. verb *abitare* 'inhabit, live in/at' only admits Animates in subject position, preferably Humans and less typically Animals, and rules out Institutions, as the example in (1) shows:

(1) a. *Luca abita a Milano.*
'Luca lives in Milan.'

b. **L'ONU abita a New York.*
'The UN lives in New York.'

Finally, arguments require a specific syntactic realization (Subject, Direct Object, Prepositional Complement, etc.) which together constitutes the subcategorization frame for the verb, for example, for *abitare* the subcategorization frame of (1) is <Subj> *abita* <PP>.

2.3. The inherent meaning

The inherent nature of the event, once it is linguistically encoded, allows one to identify verb classes such as motion verbs (*go*, *walk*), verbs of manner (*slip*), verbs of perception (*feel*, *hear*), verbs of cognition (*understand*, *grasp*), or, on a more abstract level, verbs indicating reciprocity, reflexivity, and so forth. Several attempts to define a list of verb classes on semantic grounds have been attempted.[3] For example, in the WordNet Project (Fellbaum 1998), a list of 15 classes for verbs (called *supersenses*) has been defined, in order to provide the collaborators working on the project with a broad initial classification for lexicon items. The proposed classes are listed in Table 1, where the first column reports the name of the class and the second column a set of English examples:

[3] Batinić Angster (*this volume*, Chapter 4) for example classifies Croatian verbs taking an infinitive in different semantic classes showing how the semantics of a verb affects its syntactic behaviour.

Table 1. Supersenses for Verbs in WordNet (Fellbaum 1998)

Verb Class	Examples
Body	grooming, dressing and bodily care
Change	size, temperature change, intensifying
Cognition	thinking, judging, analysing, doubting
Communication	telling, asking, ordering, singing (animal sounds)
Competition	fighting, athletic activities
Consumption	eating and drinking
Contact	touching, hitting, tying, digging
Creation	sewing, baking, painting, performing (cooking)
Emotion	feeling
Motion	walking, flying, swimming
Perception	seeing, hearing, feeling
Possession	buying, selling, owning
Social	political and social activities and events
Stative	being, having, spatial relations
Weather	raining, snowing, thundering

The proposed classes raise a number of issues from a linguistic perspective: what is the use of the supersenses? What do they tell us about the syntactic behaviour of verbs, the inferences they give rise to, the participants they involve, the manner in which the event takes place, the purpose of the event? These questions together lead to a broader question: on what grounds is it possible to group verbs in semantic classes?

The operation of identifying semantic classes, that is, groupings of verbs with similar meanings, and organise them in a hierarchy that distinguishes for example between types of motion verbs, types of perception verbs, and so on, is intuitively simple but technically complex (Ježek 2016: 128). In general, the difficulties encountered can be said to derive primarily from the fact that the meaning of verbs consists of a bundle of features with different semantic prominence. Consider for instance the case of verbs describing a change of position, like *sit*: although the action of sitting involves the motion of the person performing the action (a motion we may characterize as "internal"), this is not the

prominent feature in the meaning of the verb, and it would appear odd to classify *sit* as a verb of motion on a par with *enter* and *exit* (it is, in fact, categorized as a 'verb of assuming a position' in Levin 1993). But how is semantic prominence to be defined? Several scholars contend that prominent features are those that are transparent in the syntax, i.e. those that influence a verb syntactic behaviour (Levin & Rappaport Hovav 2005: Chapter 1); nevertheless, linking an observed syntactic behaviour to the appropriate semantic component is a notoriously arduous task.

Another problem is that verbs encode only some aspects of the event they denote while presupposing others. This distinction between denoted and presupposed information is by no means easy to make, although there are clear cases. *Arrive*, for example, presupposes motion but encodes, in fact, the result/effect of such motion, consisting in the fact that the person or thing arrived is located in a place which differs from the one it was located in before the arriving event took place. On this ground, it is reasonable to question whether *arrive* should be considered a verb of change of location rather than a verb of motion, and which criteria are eligible to distinguish between the two.

3. Traditional approaches: from syntax to semantics

The traditional classification of verbs rests on the distinction between transitive verbs and intransitive verbs, i.e. about the ability of a verb to be accompanied or not by a direct object. A more refined classification is based on the observation of the syntactic alternations allowed by the verbs, i.e. of the range of syntactic configurations that each verb can present. This may or may not include the number of arguments (cf. section 2.2). For example, in Italian, some verbs only allow transitive uses (*abolire* 'abolish', *indossare* 'wear', *affittare* 'rent'), others only allow intransitive uses (*camminare* 'walk', *russare* 'snore', *arrivare* 'arrive', *cadere* 'fall'), and others allow both (*affondare* 'sink', *guarire* 'heal'). Looking into those alternations, and following the influential work of Levin (1993), in Ježek (2003) we proposed to classify Italian verbs based on the presence or absence of four different syntactic realizations: transitive use (TR), intransitive use with auxiliary *to have* (INTR AV, unergative), intransitive use with auxiliary *to be* (INTR ES, unaccusative), intransitive use with pronominal marker *si* (and *to be* as

auxiliary, INTR PRON).[4] The result is a list of 15 classes, which exhausts all the possible combinations of these four syntactic realizations, as shown in Table 2, which provides an example of a verb for each class.

Table 2. Italian verb classes and syntactic alternations (Ježek 2003)

Class	Verb	TR	INTR AV	INTR ES	INTR PRON
1	*abolire* 'abolish'	x	–	–	–
2	*camminare* 'walk'	–	x	–	–
3	*cadere* 'fall'	–	–	x	–
4	*pentirsi* 'repent'	–	–	–	x
5	*squillare* 'ring'	–	x	x	–
6	*approfittare* 'take advantage'	–	x	–	x
7	*ammuffire* 'get mouldy'	–	–	x	x
8	*sedimentare* 'settle'	–	x	x	x
9	*mangiare* 'eat'	x	x	–	–
10	*affondare* 'sink'	x	–	x	–
11	*alzare* 'raise'	x	–	–	x
12	*ingiallire* 'yellow'	x	–	x	x
13	*continuare* 'continue'	x	x	x	–
14	*chiudere* 'close'	x	x	–	x
15	*bruciare* 'burn'	x	x	x	x

To understand the table, consider the verb *abolire* 'to abolish' in (2), which presents only a transitive use and therefore falls into class 1, and the verb *russare* 'to snore' in (3) that exhibits only an intransitive use with the auxiliary *avere*, and falls into class 2:

(2) *abolire* TR *Il governo ha abolito la pena di morte.*
'The Government has abolished the death penalty.'
INTR AV **Il governo ha abolito.* *'The Government has abolished.'
INTR ES **Il governo è abolito.* *'The Government is abolished.'[5]
INTR PRON **Il governo si è abolito.* *'The Government PRON is abolished.'

[4] The classification was resumed, among others, in Salvi & Vanelli (2004).

[5] Note that 'The government is abolished' is grammatical in the stative reading (as in 'This may mean that the government is abolished') but not in the dynamic reading, whereas the latter is required by intransitive uses.

(3) *russare* TR **Il raffreddore ha russato Gianni.* *'The cold snored Gianni.'
INTR AV *Gianni ha russato.* 'Gianni has snored.'
INTR ES **Gianni è russato.* *'Gianni is snored.'
INTR PRON **Gianni si è russato.* *'Gianni PRON is snored.'

The case of *bruciare* 'to burn' is more complex, as it admits all four uses, as shown in (4) (class 15):

(4) *bruciare* TR *L'incendio ha bruciato la casa.* 'The fire has burned the house.'
INTR AV *La casa ha bruciato per ore.* 'The house has burned for hours.'
INTR ES *La casa è bruciata.* 'The house is burned.'[6]
INTR PRON *La casa si è bruciata.* 'The house PRON is burned.'[7]

The classification we performed based on a large sample of Italian verbs (cf. Ježek 2003 for details on the dataset) shows that the properties mostly represented in the syntax of the verbs (particularly, in relation to the syntactic alternations that the verbs exhibit) are those related to the *event type*, that is, the aspectual or actional properties (process vs. state vs. change of state, cf. section 2.1), *agency*, and the presence or absence of *causality*. For example, TR verbs (that is, verbs of class 1 that only admit a transitive use) tend to be lexically agentive (*abolire* 'abolish'), i.e. they do not occur with unintentional subjects.[8] Verbs that alternate between TR and INTR AV (class 9), on the other hand, mainly denote processes (*mangiare* 'eat', *bere* 'drink')

---

[6] See the previous note concerning stative vs. dynamic readings.

[7] It is well known that the pronominal marker *si* may occur in Italian in several constructions with different semantic values along the parameter of voice (reflexive, middle and passive respectively). In the current study we consider only intransitive constructions, where *si* marks spontaneity/lack of causalness, event completion, or both (middle uses). We exclude reflexive constructions in which *si* signals coreference with the subject (*Luca si è vestito* 'Luca got dressed') and passive constructions (*si vedono le stelle* 'the stars are seen', *si abolirono le tasse* 'taxes were abolished'). For an overview of pronominal uses in Italian see Ježek (2003: 130–141). On the passive reinterpretation of the *si* + V construction in Italian, see Cennamo (2001).

[8] Note that most verbs that are generally classified as agentive, in certain contexts allow for non-agentive uses, such as *ferire* 'injure' (involuntarily). This is not the case of TR verbs: unlike *abolire*, *ferire* does not belong to this latter class as it admits also an INTR PRON use (*ferirsi*).

instead of states or happenings. Finally, verbs that alternate between TR and INTR ES and/or PRON (classes 10 to 12) describe predominantly changes of state (*ingiallire* 'become yellow').

4. Problems with syntax-first approaches

As stated above, the syntactic alternation model represents a refined approach with respect to the model of analysis based exclusively on transitivity/intransitivity. If we also add the number of arguments, we obtain an effective procedure to analyse the linguistic encoding of the participants in the event expressed by the verb. However, there are at least two weaknesses, which are connected to each other. The first is the well-known difficulty in drawing a distinction between arguments and adjuncts with respect to a specific verb. The second is the optionality of some arguments, that questions the very notion of argument, which is grounded on the idea that arguments are those complements which require mandatory syntactic realization. The first of these problems is exemplified in (5) where the adverbial 'in front of the hotel' has an uncertain status (argument or adjunct) with respect to the verb *arrive*, while the second is exemplified in (6), in which the direct object of the verb *drink*, expressed in (6a), is absent in (6b):[9]

(5) *Gianni rilesse quel che aveva scritto fin quando il taxi non arrivò davanti all'albergo.*
'Gianni reread what he had written until the taxi arrived in front of the hotel.'

(6) a. *I ragazzi bevono birra al pub.*
'The boys are drinking beer in the pub.'
b. *Il vecchio si riposò, bevve e fu assalito da uno strano pensiero.*
'The old man rested, drank, and was haunted by a strange thought.'

In the following sections, we examine two solutions we propose for solving these two problems. Our proposal highlights the role of semantics in determining the distinction between arguments and adjuncts and in licensing syntactic optionality.

---

[9] All examples are derived and adapted from the ItTenTen 2020 corpus (about 13 billion words; Jakubíček *et al.* 2013), queried through the Sketch Engine online platform (https://app.sketchengine.eu; Kilgarriff *et al.* 2004).

5. Innovative Approach: Starting from the Semantics

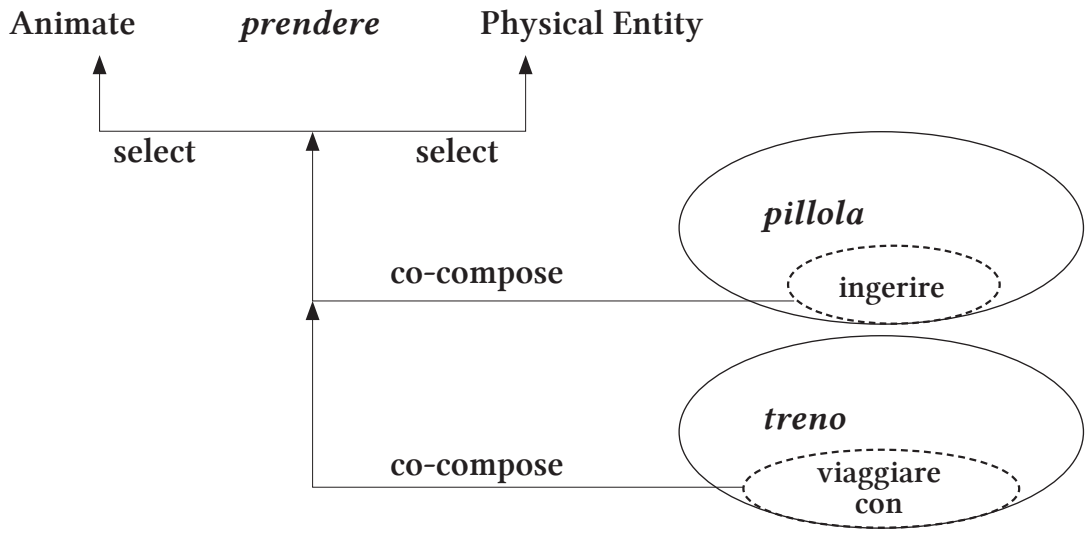
In contrast with the common practice that tackles the problems highlighted in section 4 starting from a syntactic perspective, we propose that those phenomena are better understood if the analysis is performed starting from the semantic level, particularly from the assumptions that, i.) the meanings of the arguments play an active role in the semantic composition with verb (contrary to adjuncts), and that ii.) arguments may be incorporated at different degrees into the inherent meaning of the verb. According to this proposal, the syntactic level is secondary and dependent on the semantic one. We first review our semantic approach, and then apply it to the problems raised in section 4.

5.1. Enriched composition in semantics

Our proposal departs from the traditional semantic account, according to which polysemy is a list of pre-defined senses stored in each lexical entry, and embraces a contextual approach, in which words have a core meaning that is flexible and varies from occurrence to occurrence as a function of the interaction with the other words they combine with, and of the situation of utterance (Recanati 2012). Accordingly, our approach departs from the principle of compositionality outlined in formal semantics (Partee 1995: 153), which claims that the meaning of a complex linguistic expression is determined by the sum of the meanings of its component parts, and the way they are put together. In contrast to this view, we assume that the meanings of words not only add up but also influence each other when they combine: from a theoretical point of view, we move from composition seen as an additive operation, to *enriched composition* (Jackendoff 1997) and *co-composition* (Pustejovsky 2012). Specifically, according to the principle of co-composition, the meaning of a complex expression is the result of the mutual adjustments that take place between the meanings of the combining words. Under this line of reasoning, one can argue that in the combination of a verb with its argument, the meaning of the verb is co-dependent on the meaning of its arguments, and that the variation in the semantics of arguments co-causes a variation in the verbal meaning. An example applied to the light verb *prendere* 'to take' is provided in Figure 1, where we graphically represent the co-composition with two different direct object arguments, i.e. *pillola* 'pill' and *treno* 'the train', resulting in two different

meanings for the verb ('ingest' and 'travel with' respectively). The example is based on the assumption that the meanings of nouns, represented by the oval shapes in the figure, include in their sublexical information the typical action in which the denoted entity is used.[10] When the verb selects its arguments (the direct object Physical Entity in our case), these actions (*ingest* for *pill* and *travel with* for *train*) are activated by the co-composition mechanism, and fused with the meaning of the verb. This operation results in two meanings for the verb *prendere* dependent on the meaning of the argument, namely *ingerire* 'ingest' and *viaggiare con* 'travel with'.

Figure 1. Co-composition of *prendere* + direct object



Co-composition is a unifying principle that applies to different syntactic roles (Subjects, Direct Objects, etc.) and to a wide range of semantic phenomena, including the shift from a literal to a figurative meaning of verbs, as in the examples below, that can be all accounted for in co-composition terms as illustrated in Figure 1.[11]

[10] This idea has been formalized in different ways in the semantic literature: see the notion of *hidden event* and *telic quale* proposed in Pustejovsky (1995).

[11] In (7) we provide a synonym for each sense that the verb acquires in combination with the different arguments.

(7) a. *divorare la pasta* 'devour the pasta (=eat)' | *divorare un romanzo* 'devour a novel (=read)'
b. *aprire il libro* 'open the book (=unfold)' | *aprire un dibattito* 'open a debate (=begin)'

In conclusion, by adopting the view proposed in this section, the perspective of analysis of verbs and their arguments is reversed: instead of claiming that the number and type of arguments varies according to the meaning of the verb, as in the traditional account, we claim that the meaning of the verb varies according to the semantics of the argument(s).[12] In the next sections, we take this background into account and apply it to the analysis of the two issues raised in section 4.

5.2. The argument-adjunct distinction

In section 5.1 we moved away from the traditional view according to which, in the semantic composition, a word behaves either as an active functor or as a passive argument, and embraced the idea that the arguments can both complete and influence the meaning of the verb they combine with. This move has important consequences on the way one can account for the argument-adjunct distinction.[13] In particular, we can assume that the ability of complements to co-compose with the verb does not apply to the adjuncts, which are instead *inactive* on verbal semantics and never play a role in determining the sense that the verb assumes in its context of use (Ježek 2021). According to this line of reasoning, the verb *partire* 'to leave' has a single argument in the example in (8a) (the subject *la macchina* 'the car') not for syntactic reasons but because that single argument is sufficient for the verb to take on the meaning of 'to start (off)'. As it happens, that single argument is essential to ensure that the verb acquires this meaning and not another possible meaning it could acquire (e.g. 'go away, move away from a place'). Note that in (8a), the adverbial expressions that follow the

[12] A reviewer asked an interesting question whether arguments – which influence the verb meaning – also influence the selection of the verb. In this regard, we refer the reader to a different contribution where the selection of verbs by nouns is discussed using the notion of collocational meaning (Ježek 2014).

[13] For a survey of tests for argumenthood applied to Croatian, see Birtić & Brač (*this volume*, Chapter 9).

verb are instead adjuncts, as they neither complete nor modify the meaning of the verb, which is assigned to the verb by the subject argument alone. In example (8b), instead, the verb *partire* 'to start' takes two arguments, as long as both the subject *dolore* 'pain' and the expression indicating the source (*dal braccio* 'from the arm' / *a livello del collo* 'at the neck level') are necessary to complete the meaning that the verb takes on in the context ('originate from'). In other words, according to our account, the unacceptability of the expression *'the pain starts' is the result of a semantic inconsistency and not a syntactic one, or rather, the syntactic inconsistency is due to semantic factors:

(8) a. *La macchina non parte più / mai al primo colpo.* (1 argument)
'The car no longer starts / never starts the first time.'
b. *Il dolore parte dal braccio / a livello del collo.* (2 arguments)
'The pain starts from the arm / at the neck level.'

The assumption that the distinction between arguments and adjuncts originates in the semantic composition of the verb and its complements, and that only the arguments act as functors in this process, is fully supported by the empirical observation that the valence of the verb varies according to its meaning. When one attempts to draw the distinction between arguments and adjuncts, what must be observed, therefore, is not the verb but the meaning it acquires (that is, the event that it expresses) in the specific context of use.

We applied the approach described above to identify verb senses and to tell apart arguments from adjuncts in the construction of the T-PAS resource (Ježek *et al.* 2014), an inventory of argument structures for Italian verbs focused on the semantics of arguments.[14] In T-PAS, for each verb meaning, a specific typed predicate-argument structure (T-PAS, informally called *pattern*) is provided. An example of pattern for the verb *guidare* 'to drive' in its 'operate' sense is [Human] *guida* [Road Vehicle]. Arguments are defined in terms of semantic classes notated between square brackets, called *semantic types*. Patterns are corpus-derived, i.e. they are acquired through manual clustering and annotation of corpus instances, following the Corpus Pattern Analysis (CPA) methodology (Hanks 2013). Several studies

[14] The T-PAS project was developed at the Department of Humanities of University of Pavia, with the technical support of Lexical Computing Ltd. The resource can be freely accessed and downloaded at https://tpas.unipv.eu.

performed on T-PAS data[15] have shown that co-composition is a useful principle for discriminating verb senses and in telling apart arguments from adjuncts. Also, T-PAS data allow to define semantic classes for verbs based on the semantics of the arguments (subject, direct object, prepositional complement, also in combination). A list of four classes is reported in Table 3: for example, **vehicle_verbs** are verbs that take Vehicles as Subject or Object, and so forth.

Table 3. Verb classes based on the semantic class of their arguments

CLASS	EXAMPLES
vehicle_verbs	*accelerare*, *frenare*, *parcheggiare*, *guidare*, *svoltare*, *rallentare*, *sterzare*, *sbandare*
hair_verbs	*annodare*, *pettinare*, *sistemare*, *spazzolare*, *spettinare*, *tagliare*, *tingere*, *tirare*
money_verbs	*accettare*, *accreditare*, *attirare*, *cedere*, *corrispondere*, *depositare*, *distribuire*, *emettere*, *guadagnare*, *incassare*, *pagare*, *prestare*, *riscattare*, *scommettere*, *vincere*
plants_verbs	*fiorire*, *crescere*, *germinare*, *germogliare*, *riprodursi*, *nascere*, *sbocciare*, *appassire*

Note that the verbs in the classes may vary considerably as regards their semantic spectrum, and consequently their distribution. For example, in the case of **hair_verbs**, *tirare* 'pull' has a broad semantic spectrum and one might arguably claim that it is unusual or far less frequent with an argument like *hair* than with other arguments. This is confirmed by corpus analysis, according to which the typicality score of the combination of *capelli* 'hair' with *tirare*, calculated through the Sketch Engine platform[16] is lower (7.4) than that of *capelli* with the narrow spectrum verb *pettinare* 'comb' (9.0). At the same time, *tirare* is more typical for *hair* than *annodare* 'tie' (4.6), and less typical for *hair* than *tagliare* (9.4). We conclude that each class has members that are more or less prototypical for the class and that the typicality score

[15] Currently, T-PAS contains 1160 analyzed verbs, 5529 patterns and ca. 200,000 annotated corpus instances. The references to the studies performed using T-PAS data are listed at https://tpas.unipv.it/publications/.

[16] The typicality score in the Sketch Engine platform is based on *logDice*, which is a metrics that indicates how strong the collocation of two words is. The higher the score, the stronger the collocation. A low score means that the words in the collocation also frequently combine with many other words.

can help define the internal organization of the classes. It has been noted that the classes in Table 3 show similarities with the *frames* (Fillmore 1982/2006) contained in the FrameNet repository (Fillmore *et al.* 2002),[17] although the methodologies adopted in the two resources differ. For example, the frames listed in FrameNet, although they are stipulated with the help of corpus data annotation, are not induced by manual clustering of corpus concordances based on the semantics of arguments, as are the patterns identified in T-PAS. The study of the similarities and differences between semantic patterns and cognitive frames represents an interesting direction of future research that could valuably contribute to our understanding of how semantic classes for verbs are organized and can be identified.

5.3. Types of arguments and compositional operations

As referenced in section 2.2, before influencing the meaning of verbs, arguments complete it, providing an essential semantic component of the linguistically encoded event, namely the participants: as a result, they are syntactically mandatory. Consider for example the case of the direct object of the verb *sbarrare* 'to bar' in (9), which is mandatory in all the occurrences of the verb (see the agrammaticality in (9b)):

(9) a. *Luca ha sbarrato accuratamente tutte le porte.*
'Luca barred all doors.'
b. **Luca ha sbarrato.*
*'Luca barred.'

It has been noted, however, that «sometimes an argument is obligatorily left out of the surface structure because it is subsumed as a part of the meaning of the predicate» (Fillmore 1969: 119). According to Pustejovsky (1995), this is a particular type of argument (*shadow* argument) that can be expressed only making it more specific. For example, the shadow argument *con il telefono* 'with the phone' can co-occur with the verb *telefonare* 'to phone' only if the type of phone is specified, as in (10b), and not in (10a):

(10) a. **Luca ha telefonato a Luisa con il telefono.*

[17] The FrameNet resource can be accessed at https://framenet.icsi.berkeley.edu/. I am thankful to the participants in the Workshop on Argument Structure held at Roma Tre University in September 2022 for pointing out this aspect.

*'Luca phoned Luisa with the phone.'
b. *Luca ha telefonato a Luisa con il suo telefono portatile.*
'Luca phoned Luisa with his mobile phone.'

Interestingly, shadow arguments do not occur only with denominal verbs such as *telefonare* 'to phone' but also with verbs involving body parts, such as *camminare* 'to walk (*by foot)', *vedere* 'to see (*with the eyes)', and others (Ježek 2018):

(11) a. **Luca stava camminando a piedi.*
*'Luca was walking by foot.'
b. *Luca stava camminando a piedi scalzi.*
'Luca was walking barefoot.'

(12) a. **Luca ha visto una stella cadente con gli occhi.*
*'Luca saw a shooting star with his eyes.'
b. *Luca ha visto una stella cadente con i propri occhi, a occhio nudo.*
'Luca saw a shooting star with his own eyes, with the naked eye.'

It was also noted that there exist arguments, called *hidden*, that are not expressed at all in the syntax of the verb. An example is provided by the arguments of verbs of implicit creation (Bisetto & Melloni 2007; Ježek 2014), that is, verbs that denote the creation of something through the representation of a source (Ježek & Pustejovsky 2019), such as *fotografare* 'photograph', *copiare* 'copy', and *dipingere* 'paint'. Implicit creation verbs denote the coming into being of a new entity as a result of the event itself, which, unlike explicit creation verbs such as *costruire* 'to build', does not appear superficially as the direct object of the verb; instead, the direct object expresses the source of the representation (*representation source theme* in Dowty's 1979 terminology). This is the case of the direct object in (13): the created entity (the photograph) is not expressed nor expressible as an argument of the verb in the syntax.

(13) *Luisa ha fotografato la targa.* (=created entity → photograph)
'Luisa photographed the license plate.'

The resulting entity, although not expressed, can be referred to indirectly via anaphoric expressions. In (14), for example, the clitic pronoun establishes an anaphoric relation with the unexpressed created entity, the resulting photograph:

(14) *Luisa ha fotografato la targa e l'ha inviata a Luca.*[18]
'Luisa photographed the licence plate and sent it to Luca.'

The anaphoric test supports the hypothesis that, with these verbs, the participant that expresses the result of the action is an argument, but it is *hidden*, not expressible in the syntax of the verb.

Finally, it is well known that some arguments show optionality in the syntax, that is, they can be dropped or omitted (cf. section 4). In Ježek (2018), we argue that syntactic optionality is licensed by either *pragmatic* or *semantic* factors. A pragmatic omission occurs when an argument remains unexpressed because the referent is known to the interlocutors. In this case, the saturation of the valency of the verb takes place through the retrieval of this information from the situational context. This is the case of the example in (15), in which the dots in the brackets highlight the information that must be reconstructed for the expressions to be semantically complete and interpretable:

(15) *Gianni iniziò (…) alle 8.30 e finì (…) alle 15.*
'Gianni started (…) at 8.30 and finished (…) at 3pm.'

The omissions in (15) have a pragmatic and/or textual nature (Fillmore 1986). The verbs *iniziare* and *finire* still denote a relationship between two participants, even if the second argument is not realized in the syntax, and force the listener to find something in context to saturate their meaning. If an element is not contextually assigned to the omitted argument, the sentence remains semantically incomplete.

A different phenomenon is at play when syntactic optionality is licensed by semantic factors (semantic omission). In this case, an argument may be dropped because it can be predicted by the verb semantics. Specifically, the implicit information that needs to be reconstructed to saturate the verb is the class of entities that the verb typically selects, for example the FOOD

[18] It should be noted that in a different example such as *Luisa ha fotografato il tramonto e l'ha mandato a Luca* 'Luisa photographed the sunset and send it to Luca' the clitic shows agreement with the noun *tramonto* (masculine) and not with the noun *fotografia* (feminine). This may be accounted for by assuming that the clitic refers metonymically to the information (the sunset) transferred through the photograph. In any case, the transfer cannot take place in the absence of a resulting object on which the information is recorded: this holds also for (14). See Ježek & Pustejovsky (2019) for further details.

class in the case of eating in (16a) and the DOCUMENT class in the case of reading in (16b):

(16) a. *Luisa ha mangiato alle 18.00.*
'Luisa ate at 6pm.'
b. *Luca legge spesso in treno.*
'Luca often reads in the train.'

In other words, the target of the omission in (16a) and (16b) is the particular class of objects intended by the verb for that argument, and not a specific referent of the class, as in the case of the pragmatic omission in (15). Only under the condition that the omitted object is understood generically as a class and the focus is on the action being performed, the omission can be interpreted as semantic. If the target of the omission is a specific referent of the class (for example, a pizza for FOOD; a letter for DOCUMENT), then the omission should be interpreted as pragmatic and not as semantic.

5.4. On the incorporation of arguments

After examining the different properties of arguments and providing a semantic account of the variability in their syntactic behaviour in 5.3, we finally propose that the relation between the verb and its arguments can be understood and represented as a relation of greater or lesser incorporation. The notion of incorporation has been used primarily in syntactic studies (Mithun & Corbett 1999). Following a suggestion that dates back to Gruber (1962), in Ježek (2018) we applied it to semantics. According to our proposal, the arguments exhibit different degrees of semantic incorporation with respect to the inherent meaning of the verb on which they depend. This is represented in (17), where the verbal *root* is meant not as the root in morphological terms but as the root in semantic terms (Levin & Rappaport 2005), coinciding with what in section 2 we have called the inherent meaning of verbs (motion, perception, cognition, etc.).

(17) $[[[[V_{root}\ arg_1]\ arg_2]\ arg_3]]$

The most incorporated arguments (arg_1) are the arguments that reside in the root. They can be extracted and expressed in the syntax only if they add new information to the information already introduced by the root, that is, if they are 'informative' (as in (10b)). Partially incorporated arguments

(arg_2), on the other hand, are less built-in in the root, and are therefore more easily projected into the syntax, but they can be omitted, because they are still predictable from the verbal meaning. This is the case of arguments that can be omitted based on semantic factors (s. (16)). Finally, arguments that are not incorporated to any degree (arg_3) cannot be omitted, except for pragmatic reasons (that is, unless the referent they introduce can be reconstructed pragmatically; s. (15)). According to our proposal, only arg_3 are outside the verbal root: this justifies their syntactic mandatoriness in all contexts (as in (9)). Whether the type of argument discussed in relation to creation verbs (cf. (13)) shall be accounted for in terms of incorporation, remains a matter of debate.

6. Concluding observations

After reviewing the three main components of the meaning of verbs, in this contribution we have emphasized in the first place the usefulness of syntactic approaches to the classification of verbs and their arguments at the syntax-semantic interface, especially if they take into account the syntactic alternation patterns in which each verb occurs. Secondly, we have highlighted that these approaches leave two critical issues unsolved, namely the argument-adjunct distinction and the syntactic optionality of arguments. For both we have proposed a solution based on semantic principles. Particularly, we have shown that adopting a semantic perspective as a starting point in analysing the relationship between the verb and its arguments, and assuming that, unlike adjuncts, arguments play an active role in the semantic composition with the verb, contributes to account for phenomena that cannot be explained by the traditional syntax-first approach. Moreover, we have shown that by assuming that arguments may be semantically incorporated to different degrees in the verbal meaning, one can successfully account for their syntactic optionality, when it is not due to a pragmatic factor.

Appendix

Acknowledgments

We would like to thank two reviewers for their insightful comments on an earlier version of the paper.

Abbreviations

arg	argument
INTR AV	intransitive with auxiliary *avere* 'to have'
INTR ES	intransitive with auxiliary *essere* 'to be'
INTR PRON	intransitive with reflexive marking
TR	transitive

References

Bach, E. (1986). The Algebra of Events. *Linguistics and Philosophy*, 9, 5–16.

Bisetto A. & Melloni, C. (2007). Result Nominals: A Lexical-Semantic Investigation. In: Booij, G., Ducceschi, L., Fradin, B., Guevara, E., Ralli, A. & Scalise, S. (eds) *On-line Proceedings of the Fifth Mediterranean Morphology Meeting (MMM5), September 15–18, 2005, Fréjus, France*, 393–412. (Available online: https://geertbooij.files.wordpress.com/2014/02/mmm5-proceedings.pdf).

Cennamo, M. (2001). Classi verbali e cambiamento sintattico: la reinterpretazione passiva del costrutto riflessivo. In: Salvi, G. (ed) *Semantica e lessicologia storiche, Atti del XXXII Congresso Internazionale della Società di Linguistica Italiana*, 225–242. Roma: Bulzoni.

Chomsky, N. (1965). *Aspects of the Theory of Syntax*. Cambridge (MA): The MIT Press.

Dowty, D. R. (1979). *Word Meaning and Montague Grammar*. Dordrecht: Reidel.

Fellbaum, C. (ed) (1998). *Wordnet: An Electronic Lexical Database*. Cambridge (MA): The MIT Press.

Fillmore, C. J. (1969). Types of lexical information. In: Kiefer, F. (ed) *Studies in Syntax and Semantics*, 109–137. Dordrecht/ Boston: Reidel.

Fillmore, C. J. (1986). Pragmatically Controlled Zero Anaphora. In: V. Nikiforidou, V., VanClay, M., Niepokuj, M. & Feder, D. (eds) *Proceedings of the XII Annual Meeting of the Berkeley Linguistics Society*, 95–107. Berkeley: Berkeley Linguistic Society.

Fillmore, C. J. (1982/2006). Frame Semantics. *Cognitive Linguistics. Basic Readings*, 34, 373–400.

Fillmore, C. J., Baker, C. F. & Sato, H. (2002). The FrameNet Database and Software Tools. In *Proceedings of the Third International Conference on Language Resources and Evaluation. Volume IV, Las Palmas. LREC.*

Gruber, J. S. (1965). *Studies in lexical relations*. PhD thesis, Massachusetts Institute of Technology.

Hanks, P. (2013). *Lexical Analysis: Norms and Exploitations*. Cambridge (MA): The MIT Press.
Jackendoff, R. (1997). *The Architecture of the Language Faculty*. Cambridge (MA): The MIT Press.
Jakubíček, M., Kilgarriff, A., Kovář, V., Rychlý, P. & Suchomel, V. (2013). The TenTen Corpus Family. In: *7th International Corpus Linguistics Conference CL*, 125–127.
Ježek, E. (2003). *Classi di verbi tra semantica e sintassi*. Pisa: ETS Edizioni.
Ježek, E. (2014). *Classes of Creation Verbs*. In: Simone, R. & Masini, F. (eds) *Word Classes: Nature, Typology, Computational Representations*, 37–50. Amsterdam/Philadelphia: John Benjamins.
Ježek, E. (2014). Esistono le collocazioni? Denotazione vs. significato collocazionale. In: Danler, P. & Konecny C. (eds) *Dall'architettura della lingua italiana all'architettura linguistica dell'Italia*, 409–419. Frankfurt a. M.: Peter Lang.
Ježek, E. (2018). Partecipanti impliciti nella struttura argomentale dei verbi. In: Dallabrida, S. & Cordin, P. (eds.) *La grammatica delle valenze. Spunti teorici, strumenti e applicazioni*, 55–71. Firenze: Franco Cesati.
Ježek, E. (2022). La teoria della struttura argomentale dei verbi: problemi sintattici e proposte semantiche. *Storie e linguaggi. Rivista di studi umanistici*, 8(1), 47–66.
Ježek, E. & Pustejovsky J. (2019). Dynamic Interpretation of Predicate-Argument Structure. *Lingue e Linguaggio*, 18, 179–207.
Kilgarriff, A., P. Rychlý, P. Smrz, & D. Tugwell (2004). The Sketch Engine. In *Proceedings of the Eleventh EURALEX International Congress. Lorient, France.*
Levin, B. (1993). *English Verb Classes and Alternations. A Preliminary Investigation*. Chicago: The University of Chicago Press.
Levin, B. & Rappaport Hovav, M. (2005). *Argument Realization*. Cambridge: Cambridge University Press.
Mithun, M. & Corbett, G. (1999). *The Effect of Noun Incorporation on Argument Structure*. In: Mereu, L. (ed) *Boundaries of Morphology and Syntax*, 49–71. Amsterdam-Philadelphia: John Benjamins.
Partee, B. (1995). Lexical semantics and compositionality. In: Osherson, D. (ed) *Invitation to Cognitive Science (2nd ed.). Part I: Language (editors: Gleitman, L. & Liberman, M.)*, 311–360. Cambridge (MA): The MIT Press.
Pustejovsky, J. (1995). *The Generative Lexicon*. Cambridge (MA): MIT Press.
Pustejovsky, J. (2012). Co-compositionality in grammar. In: Werning, M., Hinzen, W. & Machery, E. (eds) *The Oxford Handbook of Compositionality*, 371–384. Oxford: Oxford University Press.
Recanati, F. (2012). Compositionality, flexibility and context-dependence. In: Werning, M., Hinzen, W. & Machery, E. (eds) *The Oxford Handbook of Compositionality*, 175–191. Oxford: Oxford University Press.
Salvi, G. & Vanelli, L. (2004). *Nuova grammatica italiana*. Bologna: Il Mulino.
Tesnière, L. (1959) *Éléments de syntaxe structural*. Paris: Klincksieck.
Vendler, Z. (1967). Verbs and Times. In: *Linguistics in philosophy*, 97–121. Ithaca (NY): Cornell University Press.
Wilks, Y. (1975). A preferential, pattern-seeking semantics for natural language interference. *Artificial Intelligence*, 6, 1, 53–74.