

Chapter 6

Coding Semantic Relations

6.1 A Preliminary Note

Before approaching the question of the coding of semantic relations (that is, their association with the constituents of the sentence), some considerations are in order about what, exactly, is associated with what.

We usually speak of “coding the Agent as the subject”, and equivalent expressions, but this is a convenient abbreviation for a more indirect relation. In a sentence like

[1] Mary helped Jim.

we talk about *Mary* being the Agent, *Jim* the Patient, and in most cases (not all) no inconvenience ensues. But in fact the role of Agent is not associated with the NP *Mary*—this association happens at a cognitive level, that is, as part of the construction of the mental landscape. The subject in [1] evokes a schema, MARY; by reference to the verb, *help*, plus the syntactic function of *Mary*, the role Agent is associated with that schema. In other words, we understand that Mary (the person, not the phrase or the word) is the Agent of the denoted action, HELP.

In this example, the usefulness of the distinction for the analysis may not be too evident. But this is so only because the relations between sentence constituents and schemata in sentences like [1] are pretty direct. In other cases the distinction is crucial, and this will become evident when we approach complex and reciprocal constructions. As will be seen in more detail in Sect. 6.6 below, in

[2] Pete moved the chair away from the table.

we have two events: the movement of the chair away from the table, and the action of Pete, who touches the chair (not the table) and does something to it. That is, the chair appears twice, although we only have one occurrence of the NP *the chair*; here we are dealing not with the linguistic unit *the chair*, but with the corresponding schema. In such cases the distinction made above is essential for a correct analysis

of what we understand when we hear [2]: one, Pete did something to the chair; two, the chair moved away from the table. This can only be expressed if we attach semantic roles to the schemata; since the schema CHAIR appears twice (in Event 1 and again in Event 2), it can, and does, receive two separate semantic roles—here, Patient and Theme. I return to this point below, and it will hopefully become clear enough.

6.2 Coding Semantic Roles

As we know, besides the problem of defining and delimiting semantic roles, we must consider the question of coding them as constituents of a sentence. But before approaching the details of the coding system, let us reflect a little on the place of this system within the general workings of the language. This is a major point, and needs some elaboration.

The coding system aims at making explicit the association between the constituents of a sentence and their respective CSRs.¹ The constituents provide the identity (for instance, the reference) of each participant, and several mechanisms allow us to associate a CSR to each of the participants. Or, otherwise said: the system has as its aim filling in (binding) the variables of a schema evoked by the verb.² The system has then a unique aim, and a unique result, namely, to establish **constituent/CSR** pairs, exhaustively covering all relevant constituents of the sentence.

To detail: faced with the sentence

[3] The dog bit the boy.

the language user must build a mental landscape containing, among other things, the following pieces of information: (a) the sentence reports an event; (b) the event occurred in some unspecified location; (c) the event occurred at some point in the past; (d) the dog, not the boy, used his teeth; (e) the boy underwent contact with the dog's teeth; (f) the boy was probably hurt, etc.

Part of this information comes directly from the schema BITE, evoked by the verb form *bit*. This schema refers to an event, more precisely an action, definable as “grasp with the teeth”; in general it is understood that some force is exerted. The schema BITE contains a set of variables, each of them defining the relation of the event with one of its participants.³ But there is some information that the schema BITE cannot provide. For instance, [3] leaves it clear that the event occurred in the

¹ Keeping in mind the qualification seen in Sect. 6.1: this association is mediated by the schemata involved.

² In certain cases, by a nonverbal element, as in the case of light verbs; see Sect. 6.5.

³ “Participant” is to be taken in a broad meaning, sometimes including circumstances of location, time, etc., and even identity of reference (as in *that blond woman is Susan*).

past, but this cannot come from the schema, which is neutral in that respect; this piece of information comes from the verb form.

We still lack two crucial pieces of information, that the dog used his teeth and that the boy took the biting, not vice versa: (d) and (e) in the above list. Again, the schema alone does not help. What does help is the syntactic structure of the sentence: in [3] the verb *bite* occurs in a diathesis where the subject is an Agent and the object is a Patient. This is part of the process of variable-binding in the schema which, as we know, includes a tooth-using entity and an entity that takes the biting. The function of the diathesis is to give us indication about how to fill in the variables of the schema, and this indication is provided by mechanisms that code semantic roles as complements. These mechanisms must be understood as part of a general process of construction of the mental landscape, i.e., of utterance comprehension. It is important to keep this in mind mainly because, contrary to common belief, language users employ several different resources in order to carry out that task.

6.3 Role-Coding Mechanisms

In example [3] it is the valency of the verb, *bite*, that provides the hints for relating the constituents *the dog* and *the boy* with their semantic roles, respectively Agent and Patient: that is, semantic roles are here coded according to the properties of a governing word, which in this case is the verb. This situation is normally understood as typical, perhaps the only possible one—an idea implicitly accepted in many studies in the area. An example is Baker's (1988) proposal, named UTAH, which only considers the possibility of assigning a semantic role to an item by another item or constituent. See also the possibilities of syntax-semantics mapping mentioned in Haegeman (1991, p. 61), Levin and Hovav (2005), Goldberg (2006, p. 39), Matthews (2007), where only the governing item is mentioned as a determining element. In general, the possibility is not considered that in some cases the verb⁴ may have nothing to do with the complement's semantic role.

In what follows I call into question this general belief, and suggest that several different factors may be involved in semantic role coding, although coding controlled by a governing term is possibly the most frequent situation. Research has already shown several cases in which semantic role coding depends on other factors. Limiting ourselves to sentential structures, the possible situations are the following:

1. Cases in which the semantic role is syntactically coded by reference to the valency of the main verb (this is the classic case, universally recognized).

⁴ As opposed to other elements such as the evoked schemata, and the inherent semantic potential of prepositions and phrases.

2. Cases in which the semantic role is determined by a preposition, independently of the verb of the sentence.
3. Cases in which a (nonprepositional) phrase has inherent semantic role, independently of the syntactic context.
4. Cases in which the main verb is light, and semantic roles are syntactically coded, partly, on the basis of a nonverbal complement.
5. Cases in which semantic role coding results from a linking process, which associates certain CSRs preferentially with certain syntactic functions.
6. Cases in which the constituent depends, semantically, on more than one schema evoked by the main verb. These are semantically complex constructions, where a single clause must be analysed semantically as a complex of propositions.
7. Cases in which some complements receive their semantic roles—or rather their CSRs—by default, following a principle of giving preference to core CSRs over peripheral ones.

While the existence of some of these mechanisms is little news, I am not aware of an approach considering them all, as a set of mutually complementing devices. And some of them are not, to my knowledge, as much as mentioned in the literature.

The resulting picture is much more complex than generally admitted. These seven⁵ coding mechanisms form a battery of resources to be used according to the occasion. Some of them exclude others, as is the case with cases 2 and 3 (cases of semantic transparency of constituents). But other mechanisms are in competition; thus, there are cases in which the semantic role can, in principle, be coded through mechanisms 1 or 3. In these cases both resources are probably available to the language user, who will activate one or the other according to the convenience of the moment. Role-coding is, then, an opportunistic process, capable of adapting its strategy to the context. The process itself is, evidently, part of the user's performance, but the mechanisms that make its functioning possible belong to language knowledge; they are elements of competence, and must be described as part of the structure of the language.

Levin (1993) voices an opinion that seems to be adopted, without much criticism, by other researchers:

[...] the behavior of a verb, particularly with respect to the expression and interpretation of its arguments, is to a large extent determined by its meaning. Thus verb behavior can be used effectively to probe for linguistically relevant pertinent aspects of verb meaning.

(Levin 1993, p. 1)

We have to agree that the determination of the verb's arguments, that is, the particular list of semantic roles associated with each verb, is a function of its meaning. But in what regards the way each semantic role is coded in syntactic structure, close examination of the data reveals so many particularities that we have to take Levin's statement only as a vague and general working hypothesis, not as a fact of the language. Yet Malchukov *et al.* (internet) say that Levin's book

⁵ That is, the seven already identified; there may be others. I have focused on factors having some relation to grammar.

shows that a semantic classification of verbs can be achieved through applying syntactic diagnostics.

(Malchukov *et al.*, internet, p. 1)

and Taylor (2002) voices the hope that semantics can always be taken as a reliable guide to morphosyntax:

The expectation [. . .] is that the syntactic (and morphological) facts of a language will be *motivated* by semantic aspects and that they can be exhaustively described by means of symbolic structures.

(Taylor 2002, p. 29; italics in the original)⁶

In this book it is shown, to the contrary, that the symbolic relation between semantic relations and their syntactic expression varies considerably for verbs of similar meaning—which, however, does not preclude the existence of clear tendencies and hierarchies defining preferential association between semantic roles and syntactic functions. In other words, the facts are complex, and in no way can we say that syntactic features are a reliable index of semantic classes.

As we saw, semantic roles are but a resource for the coding of CSRs, which are ingredients of comprehension, that is, of the final understanding which the receptor achieves after processing the perceived signal. There are indications that semantic roles do not always appear as a necessary step in this process; in certain cases the CSR is mapped onto a constituent by direct reference to the schema. In such cases, it must be stressed that the objective of the whole process is attained: the linking between a constituent and a CSR is established, and eventually integrated into the final mental landscape that makes up the understanding of the utterance.

I proceed now to a preliminary explicitation of each of the mechanisms in the above list. Some of them will have to be studied in more detail, which will be the topic of later chapters.

6.4 Coding by Verb Valency

This is the classic case normally recognized in the literature: a semantic role is coded as a particular constituent on the basis of its syntactic function and the valency of the main verb of the sentence.

The syntactic functions of the NP are **subject** and **object**. I call **object** any nonsubject NP, since as far as I can see there is no reason to distinguish syntactically other functions of the NP in the sentence.⁷ What we have in the sentence, then,

⁶ Taylor's position cannot be taken as valid for cognitive linguists in general: see, for instance, Langacker (1987b, p. 53), where a more moderate view is expressed.

⁷ For semantic role coding it is enough to tell subjects from any other NPs; see Chap. 2 above, and discussion in Perini (2008, section 4.1, 2008). For the special case of so-called **predicatives**, see Perini and Fulgêncio (2011). **Object** is, actually, a convenience term, not to repeat “nonsubject NP” all the time; in the notation of the diatheses what appears is just **NP**.

syntactically, is a subject, one or more objects (that is, all other NPs), and other terms referred to either in a concrete manner, through mention of the preposition (*in* NP, *with* NP, etc.) or through the form class (Adjective Phrase, Adverbial Phrase). This means that we can, in some cases, have more than one object in a single clause; this is no inconvenient because their semantic potential, and sometimes their ordering, is sufficient to distinguish them for the purposes of the description.

This coding mechanism depends, naturally, on a list of verbs and valencies (a valency dictionary), which is part of the structure of the language. Each verb in the language is associated with a certain number of diatheses, and the set of diatheses of a verb constitutes its valency. There is some insistence in the literature on defining valency as merely the statement of the *number* of complements of a verb, without mention of their syntactic and semantic representation; this is behind designations like **monovalent**, **bivalent**, and **trivalent** verbs. This conception of valency (which comes from Tesnière 1959) is of little interest, and leaves too much to be expressed. For instance, some authors speak of bivalent verbs—but what is a bivalent verb? A verb like *drink* does occur with two complements (*the cat drank the milk*), but also with only one complement (*her husband drinks*); *look* also has two complements, but not two NPs as *drink*, because one of the complements is a prepositional phrase (*I looked for you*)—and so on. Calling both *drink* and *look* “bivalent” is not very informative, and besides is incorrect, because *drink* can also appear with only one complement.⁸ For these reasons, some linguists nowadays consider not only the number of complements in each diathesis (not each *verb*), but also the much more important issues of their syntactic functions and semantic roles.

The opposition **subject/object** is dominant in semantic role coding in cases like [3] The dog bit the boy.

We have here an Agent and a Patient; and the identification of each depends on the syntactic function (subject or object) of the constituents *the dog* and *the boy*. The verb *bite* has in its valency a diathesis defined as composed of **subject Agent** and **object Patient**; and, since [3] admits the analysis as **subject+verb+object**, and the verb is *bite*, semantic roles are coded with basis on that diathesis: the subject is Agent, the object is Patient.

Now consider the sentence

[4] Barbara is our best teacher.

Syntactically, there is no reason to analyze *our best teacher* as anything but an object. What distinguishes this constituent from the object of [3] are semantic

⁸Herbst and Schüller (2008, p. 137) classify verbs according to their *maximum* number of complements, and thus escape this objection. But again, one may ask what is the special relevance of the diathesis that shows most complements?

factors, which led grammarians to devise the “syntactic” function of predicative. But it is not an autonomous syntactic function, as shown in Chaps. 2 and 10.⁹

The mechanism of role-coding by verb valency is crucial for all cases where the complement is an NP, since NPs are never thematically transparent. But verb valency can also be decisive for other types of constituents, notably prepositional phrases when the preposition is particularly opaque. Thus, with *gostar* ‘like’ the Stimulus is preceded by the preposition *de*: *ele gosta de cerveja* ‘he likes beer’; but with *apanhar* ‘be spanked’ the same preposition introduces the Agent: *ele apanhou de uma vizinha* ‘he was spanked by a neighbor’; and with *se aproximar* ‘come near’ *de* introduces a Goal: *ele se aproximou de mim* ‘he came near me’.

6.5 Light Verb Constructions

The examples seen above suggest that semantic role coding is a deal between the verb and its complements. The verb provides the specification of the event or state and the complements specify the participants—what Tesnière (1959) called the **actants**. This appears in a passage we have already seen:

grammatically relevant facets of a verb’s meaning are represented by a list of labels identifying the role that each of the verb’s arguments plays in the event it denotes.
(Levin and Hovav 2005, p. 35)

This is the general idea, which clearly underlies usual formulations, in particular the proposed lists of possible semantic roles, and seems to be in the basis of the study of valency. This is no doubt adequate in many cases, possibly most; but there are important exceptions. In the classic case, the verb specifies the type of event or state denoted, whereas complements and adjuncts provide the participants and circumstances involved in this event or state. With so-called **light verbs**, however, the specification of the event seems to be shared between the verb and a complement.

6.5.1 Semantic Functions of the Verb in the Sentence

The verb performs several important semantic functions in the sentence. Among these functions, we can select the following three as specially relevant for our discussion:

Lexical semantic function: to specify the nature of an event or state, often in minute detail.

⁹This claim about the predicative is valid for Portuguese. I give English examples here, but I would not like to be held accountable for the details in every case.

Role-finding function: to identify a set of CSRs associated to the meaning of the verb.

Symbolic function: to assign each of these CSRs to one of the complements.

The symbolic function characterizes the syntactic contexts in which a verb can occur. The verb *eat* performs these three functions in the following way:

Lexical semantic function: *eat* denotes an action consisting in putting something solid in the mouth and swallowing it for feeding purposes.

Role-finding function: *eat* associates with an “agent” and a “patient”.

Symbolic function: *eat* codes the “agent” as the subject, and the “patient” as the object.¹⁰

The lexical semantic function is what is usually known as the “meaning” of the verb. This function is carried out without the help of elements external to the verb; that is, the lexical semantic function associates directly with the verbal lexeme, not with its complements. The lexical semantic and role-finding functions are a direct consequence of the fact that *eat* evokes the schema EAT, and have to do with the schema rather than the verb. The symbolic function, however, crucially depends on the identity of the lexical item *eat*, and cannot be derived from the schema.

6.5.2 Exceptional Verbs

According to the usual notion, the verb is a sort of semantic axis of the sentence, and the complements¹¹ provide the reference for the occupants of semantic roles. Now we will examine some verbs for which this arrangement seems to fail.

One of these verbs is *undergo*, as in

[5] The player underwent a surgery.

Here we have a subject with the feature **affected**, which we can analyze as a Patient. The subject receives its semantic role in this sentence in the normal way, based on the symbolic function of *undergo*.

But the object, *a surgery*, does not fit into any of the usual semantic roles. Instead, this constituent seems to have **lexical semantic function**, that is, it specifies the kind of event evoked by the sentence—a function normally performed

¹⁰ In the transitive construction: *o rato comeu o queijo* ‘the mouse ate the cheese’. In the passive, coding is different, but this is only an apparent exception, because in a sentence like *o queijo foi comido pelo rato* ‘the cheese was eaten by the mouse’ there is no occurrence of the verb *comer* ‘eat’, but of a nominal related to *comer* by derivation; I have shown this for Portuguese in Perini (2008, 2010). For English, of course, the analysis may be different.

¹¹ Sometimes confused, terminologically, with arguments—for instance in the quote from Levin and Hovav seen above. I prefer to reserve the term **complement** for a morphosyntactically realized term of the sentence; an **argument** is a semantic ingredient.

by the verb, not by a complement. It seems, then, that the event specification in [5] is split, and the verb takes charge of the information that there is a Patient involved; but the kind of event (a surgery) is informed by the object.

The situation is very different in

[6] A lion ate my goat.

where the object contributes nothing to specify the kind of event, which is totally contained in the verb *ate*.

In [5], the verb has the role-finding and symbolic functions, but it is deficient in what respects the lexical semantic function. This can be seen clearly by comparing [5] with other sentences with the same verb:

[5] The player underwent a surgery.

[7] The player underwent torture.

[8] The project underwent a reformulation.

Each of these sentences expresses a different kind of event, although the verb is the same. The event is specified by the object; as a matter of fact, *undergo* always occurs with an event-denoting object:

[9] * The player underwent the coach.

[10] * The old lady underwent a foot. (OK ... *foot treatment*, which is an event)

The verb *undergo*, then, although not totally devoid of lexical meaning, is somewhat impoverished as to that function.

6.5.3 Light Verbs

Other verbs show impoverished lexical meaning. For instance, turning to Portuguese examples,

[11] Eu dava um passeio toda manhã.

I gave a walk every morning

‘I took a walk every morning’

In this case the meaning of the sequence **verb+object** may also be expressed by a specialized verb, with full lexical meaning:

[12] Eu passeava toda manhã.

‘I walked every morning’

The semantic analysis of [12] is identical to [11] as far as the arguments are concerned: in both cases we have an Agent and an action, and that is all (the time adjunct does not interest us here). In [12] this is coded in the regular way, only with subject and verb; but in [11] there is in addition an object, which serves to complete the lexical meaning of *dar* (here, corresponding to English *take*). Verbs that behave like Portuguese *dar* and English *undergo* are called **light verbs**, and are defined like

A verb with little or no semantic content of its own which combines with a (usually indefinite) direct object noun or NP which itself expresses a verbal meaning.
(Trask 1992, p. 160)

The lexical semantic aspect is adequately expressed in Trask's definition; but we still need to examine the consequences of the semantic poverty of light verbs on the definition of semantic roles.

Scher (2003) defines the light verb construction by using the following criteria:

- (I) The main verb is semantically vague;
- (II) the complement (represented by a nominal) has as its head an action noun, generally a deverbal one, which really predicates about events;
- (III) there is, in general, a paraphrase between the light verb construction and the simple verb that corresponds to the nominal head.

(Scher 2003, p. 205)

This is basically Trask's definition, plus the possibility of a paraphrase as a test. The paraphrase test, to be sure, functions very precariously, because it is subject to lexical unpredictability (see below). As for the specification of the event, Scher notes, correctly, that it does not depend only on the verb:

In all cases, [...] the nominal element seems to be really responsible for denoting the event expressed by the sentence.

(Scher 2003, p. 206)

Semantic role coding in these cases does not work in the ordinary way. According to Scher, with light verbs

The verb [...] does not seem to be responsible, or at least not the only responsible factor, in the association of thematic roles with the argument NPs of sentences. The real responsible factors [...] seem to be the nominal elements [...] which, together with the light verb [...] compose complex predicates.

(Scher 2003, p. 208)

I agree with Scher in the above points, to which we return below.

Trask's definition, stating that a light verb has "little or no semantic content", suggests the possibility that some have little, and others have none. In fact, a quick perusal seems to show that some verbs are "lighter" than others. Thus, *dar* in [11] assigns the role of Agent to the subject, and that is apparently all (besides expressing tense, aspect, person, etc.). Now, in

[13] *Eu levei uma surra da Mathilde.*

I took a beating from Mathilde

the verb *levar* 'take' does more than mark the subject as Patient, because you cannot *levar* a kiss, or some help; the restriction seems to be that the event must be not only unpleasant, but even harmful to the patient. One cannot say

[14] * Eu levei uma cirurgia do doutor Júlio.

I took a surgery by dr Júlio

although it is OK if we substitute *um tapa* ‘a slap’ for the surgery.

Duarte *et al.* (internet) sought to specify the semantic and syntactic contribution of each component in the sequences of light verb plus nominal; they summarize their position as follows:

[...] both light verb and derived noun seem to contribute to the properties of the complex predicate, in such a way that argument structure and attribution of thematic roles are determined by both constituents through the combination of their thematic structures.

(Duarte *et al.*, internet, p. 1)

Something like the combination of thematic structures does in fact happen. We have now to specify how this combination takes place. Duarte *et al.* attempt to state the combination, but their analysis is open to certain objections. They point out, correctly, that

C: light verbs, but not auxiliary verbs, impose restrictions on the semantic selection of the subject.

(p. 2)¹²

In fact, light verbs not only determine selectional restrictions, but also assign semantic roles to certain constituents (e.g., *undergo* has always an affected subject); in this they differ from auxiliaries, which are neutral as to these factors.

Another distinctive feature of light verbs, according to Duarte *et al.*, is

A: possibility of paraphrasing the sequence <light V + deverbal N> with a main verb, morphologically related to the noun.

(p. 2)

There are many cases in which this paraphrase is possible, although it does not always result in identical meanings for the two sentences. But in other cases the verb is unquestionably light and there is no main verb to base a paraphrase. This depends on the presence in the lexicon of a suitable item, which is largely accidental: *dar uma varrida* ‘to give a sweeping’ is an (approximate) synonym of *varrer* ‘sweep’, but *sofrer uma cirurgia* ‘undergo a surgery’, *dar uma gafe* ‘commit [lit. ‘give’] a social blunder’, *levar uma bronca* ‘take a scolding’ have no synonymous verbs. At most, the existence of such a paraphrase may be one of the indications that a verb is light; it cannot be a criterion for identification.

D: light verbs keep part of the meaning of their homonyms.

(p. 3)

The same example shows that this is not true: we do not find in *dar um passeio* ‘take a walk’ any of the fundamental semantic ingredients that characterize *dar* ‘give’ as a full verb. This seems to be a typical situation.

E: the light verb external argument controls the event denoted by the derived noun.

¹²I keep the authors’ alphabetic references (A, B, C, etc).

(p. 3)

This again works in some cases, not in others. As we saw, *dar* as a light verb assigns the subject (“external argument”) the role of Agent, that is, controller. But this property does not extend to all cases: with *sofrer* ‘undergo’ the subject is Patient, and has no control on the event.

F: the light verb is sensitive to the argument structure of the deverbal noun it combines with.

(p. 4)

One of the examples given is *ter*, which as a full verb “combines preferably with deverbal nouns which do not select a *source* argument [...] except when the *source* argument denotes an embedded event with a culmination point or a resulting state”. The examples given seem to involve “goal”, not “source”: * *O João teve uma promessa à Maria* ‘*João had a promise to Maria’.¹³

But *ter* does not make this requirement: in *o João teve algum auxílio do governo* ‘João had some help from the government’ the deverbal noun is *auxílio* ‘help’, which does select a “source”. The passage is not very clear, but from what I understand of it I have to disagree with this criterion; the feature is occasional, not systematic as Duarte *et al.* suggest.

Part of the objections above will be removed if we define a light verb as only those that co-occur with a nominal having a cognate verb which provides the paraphrase. In this case, *dar* will be light in *o Pedro deu uma corrida* ‘Pedro ran [lit. gave a run]’, since there is a cognate verb *correr* ‘run’, but *sofrer* will not (by this criterion) be light in *o jogador sofreu uma cirurgia* ‘the player underwent a surgery’, where there is no synonymous full verb. But if we adopt this solution we will not be speaking of light verbs, but rather of V+NP sequences, or complex predicates as Duarte *et al.* call them. A complex predicate would be defined in terms of a nongrammatical factor, the existence in the lexicon of an adequate verb that paraphrases it. The light verb phenomenon, which is here described in terms of the presence of a constituent (other than the verb) characterizing the denoted event, would have no place in their analysis. Besides, tying the definition of light verb to a morphological connection subject to lexical accidents will split the category, leaving outside verbs that behave, syntactically and semantically, like those that are considered “light”.

To summarize, the attempt to show parallelism between light verbs and their full homonyms does not work. Duarte *et al.* deny that light verbs are functional or auxiliary elements. They are right in this point, as certainly a light verb is not the same as an auxiliary; and it is not a main verb either, at least not of the common variety, which is responsible for specifying the denoted event—an important semantic feature that can be used as a basis for a better definition.

¹³ Underlining as in the original; the authors probably refer to the constituent *à Maria* ‘to Maria’, which is a Goal.

Which brings us to the question, What is a light verb? A light verb be defined as a third category, separate from full verbs and auxiliaries, characterized by the following features:

- (a) The verb selects a constituent with the syntactic function of object,¹⁴ whose semantic function has to do with the specification of the denoted event. This semantic function can be represented as a semantic role, called **Event Specification (EvSpec)**.
- (b) The symbolic function of the verb is limited to defining the semantic role of the subject.¹⁵
- (c) The semantics of the verb is impoverished in comparison with its full homonym (if any), and with full verbs in general.¹⁶

Features (b) and (c) require further research, because they were arrived at by the examination of a small number of verbs, and may have to be reformulated; but the general idea is tolerably clear. Furthermore, (c) is pretty vague, but its presence in all cases seems to me unquestionable, and as we saw is normally recognized in the literature.

6.5.4 *Event Participants*

Semantically, a sequence of **light verb+object** functions as a simple verb. In

[16] *Eu levei uma surra.*

I took a beating

where the Agent is omitted, we have an event with only one overt participant (*eu* ‘I’); and in

[13] *Eu levei uma surra da Mathilde.*

I took a beating from Mathilde

we have two participants: the Patient, *eu* ‘I’, and the Agent, *da Mathilde* ‘from Mathilde’—in spite of the additional presence of an object NP, *uma surra* ‘a beating’, which does not contribute to the list of participants, but specifies the event. That this is the situation is clear from sentences like

[17] *Eu apanhei.* ‘I took a beating’

[18] *Eu apanhei da Mathilde.* ‘I took a beating from Mathilde’

where the participants are the same as in [16] and [13], respectively. Whatever the function of *uma surra* ‘a beating’ in [16] and [13], it is different from the function of

¹⁴ I.e., a nonsubject NP.

¹⁵ The verb sometimes defines some other features such as “event unpleasant to the Patient”—but this is part of the lexical semantic function.

¹⁶ I speak of ‘homonyms’, but of course it is just another function of the same verb.

the phrases that add a participant to the resulting mental landscape. Another example are the sentences

[19] *Eu espirrei.*

I sneezed

[20] *Eu dei um espirro.*

I gave a sneeze

which are very close synonyms, although one of them has an object that does not appear in the other. These examples support Scher's (2003) claim that the semantic function of this constituent is to identify the event we are speaking about; we will express this property through the semantic role **EvSpec**.

6.5.5 *Event Specification as a Semantic Role*

Let us compare the ways symbolic coding works in sentences with and without light verbs. We saw that in the sentence

[21] I sneezed.

there is only one expressed argument (the Agent, expressed by the subject); and the event is totally expressed by the verb, which also assigns the subject its semantic role.¹⁷ And in

[5] The player underwent a surgery.

we find again only one expressed argument (the Patient, expressed by the subject *the player*). As for the expression of the event (in the case of *undergo*), it is shared by the verb and the object, as follows:

The verb is responsible for number, aspect, tense, and person; and it assigns semantic role to the subject (here, Patient).

The object specifies the essential part of the lexical semantic function, including details of the event: in [5], a surgery.

The CSR of the object *a surgery* must be something like “event specification”. Although the semantic function of this CSR looks a bit different from the more usual ones we are accustomed to (“patient”, “agent”, “source” . . .), the grammatical mechanism is similar: the semantics of this constituent elaborates the semantics of the verb. If “event specification” is a CSR (and, eventually, a semantic role), it is because the complement it is attached to restricts the meaning extension of the verb—and this is the function of complements in general. Just as *eat* has a more schematic reference than *eat a cookie*, (light) *take* is more schematic than *take a walk*. Syntactically, the situation is analogous in the two cases, and semantically the

¹⁷ We might then say that EvSpec is a semantic role of the verb in these cases.

difference is in the degree of schematicity of the verb. The difference between cases of light verbs and of sequences of nonlight verb plus object is that the semantic role of the verb is much more reduced with a light verb, and elaboration is left to the object to a greater extent.¹⁸

It is true that *a walk* in *take a walk* is not properly to speak a participant of the event, because it denotes the event itself. But, in any case, participants expressed by semantic roles cannot be understood only as dramatis personae, because we accept semantic roles like Quality, Location, and also identity of reference (which I note as α Ref),¹⁹ in sentences like

[22] That blond woman is the director.

The semantic role of both *that blond woman* and *the director* is α Ref because this sentence asserts the referential identity of these two elements. At least in the present state of research, I think we must accept a pretty wide variety of semantic relations as CSRs and semantic roles.

Thus, the valency of *undergo* includes the diathesis

VSubj>Patient V NP>EvSpec

Semantic roles must be lexico-grammatically relevant, as this is the base of the distinction between them and CSRs. Now, the semantic relation I am calling EvSpec must appear in the valency of certain verbs, and not in the valency of other verbs, thus subclassifying them: the object of *eat* is a Patient, the object of *undergo* is an EvSpec. Consequently it makes sense, for purposes of valency description, to treat EvSpec as a semantic role. As a bonus, this provides us with a short semantic definition of **light verb**: it is a verb that accepts, in at least one of its diatheses, a complement with the semantic role EvSpec.²⁰

The sharing of tasks between a light verb and its object has been observed by some authors—for instance, Scher (2003), mentioned above. Jun (2003) proposes a system of semantic unification between the complement and the verb:

the C[ognitive] S[tructure] of the nominal, instead of serving as a semantic argument of the light verb, is unified with the CS of the verb as a whole. The composite CS has an argument

¹⁸ The difficulty in using a light verb without an object certainly comes from its excessively schematic meaning: it violates Grice's (1975) maxim of quantity.

¹⁹ α Ref does not identify with Quality. To quote Lemaréchal (1989, p. 31), we have to distinguish between [a] "propositions where the predicate introduces a supplementary attribute", as in *Pete is a doctor*, or *Pete is tall*, and [b] "propositions where the predicate constitutes already the designation of a substance previously identified and where the predication amounts to indicating that this substance is identical with the one designated by the subject", as in *Pete is the doctor*. The former are here expressed by the role Quality, the latter by α Ref.

²⁰ Syntactically, a light verb is probably regular. An object like *a walk* in *I took a walk* cannot be resumed by a pronoun (*I took a walk; and *I took it yesterday morning*), but this can be readily described by stating that a phrase with the semantic role EvSpec cannot be resumed by *it* (probably because it is not referential). It is the same, for instance, with NPs with the semantic role Measure, as in *he weighs 160 pounds; *he weighs them since 2008*.

structure that reflects the common arguments of the verb and the nominal, while allowing room for the nominal to include extra material not present in the light verb.

(Jun 2003; *apud* Culicover and Jackendoff 2005, p. 223)

Jun's analysis is possibly motivated by the observation of cases of equivalence between a sequence of **light verb+object** and a full verb, practically synonymous with the sequence. I tend to feel, however, that marking the complement as EvSpec is enough to describe the facts in these cases.

6.5.6 *Assertion and Presupposition*

Before closing this section, I want to raise an idea which may deserve future consideration, relative to the way the semantics of nominals with light verbs is represented. In

[5] The player underwent a surgery.

we have a light verb and an object denoting the event in question. In this sentence the occurrence of the event is **asserted**. The sentence has as its central function to inform that there has been a surgery, and to add some limiting elements: the patient, the time, etc. This seems to be the case with the other examples of light verbs we have seen, always complemented by an object asserting the occurrence of the event. In the case of [5] this can be readily tested by negation: if we negate the verb, there was no surgery.

But these same NPs can also occur with verbs expressing the event, that is, nonlight verbs, and in these cases the occurrence of the event is not asserted, but **presupposed**. For instance,

[23] Dr House finished the surgery.

This sentence presupposes the occurrence of the surgery, and adds an assertion to it (Dr House finished it). If we negate the verb, the surgery is still understood as occurring, which shows that its occurrence is presupposed:

[24] Dr House did not finish the surgery.

It would be interesting to find out if this opposition between assertion and presupposition is systematic for light verbs. If so, it may be interpreted as a consequence of the lack of assertive content of light verbs, which are limited to giving information about tense, aspect, etc., plus eventual secondary information like "the event is not pleasant to the Patient" and the like. If this is a general fact (which remains to be checked), we will be able to derive from it an additional semantic criterion to characterize light verbs.

Light verb constructions represent a further complicating element to the semantic role assignment system. And, as we will see, that is not yet the end of the story.

6.6 Complex Constructions

6.6.1 One Clause, Two Events

In some cases a simple sentence must be semantically analyzed as a complex of propositions, and these cases require a partial reformulation of the concept of semantic role coding. An example is found in Jackendoff (1972), and involves verbs like *buy*, *sell*, and *trade*. His example is

[...] the verb *trade*, which takes a direct object, an optional phrase with *to*, and an obligatory phrase with *for*.

(2.48) Esau traded his birthright (to Jacob) for a mess of pottage.

This sentence describes two related actions. The first is the change of hands of the birthright from Esau to Jacob. The direct object is Theme, the subject is Source, and the *to*-object is Goal. Also there is what I will call the *secondary action*, the changing of hands of the mess of pottage in the other direction. In this action, the *for*-phrase is Secondary Theme, the subject is Secondary Goal, and the *to*-phrase is Secondary Source.

(Jackendoff 1972, p. 35; his numbering)

In such cases, naturally, one cannot simply assign double semantic roles to the complements: Esau is not simply Source and Goal, but Source of the birthright and Goal of the mess of pottage. In the notation here adopted the analysis of the sentence will be as shown in [26] below. For graphic reasons, I had to put the semantic roles *under* the syntactic form-class and functional symbols, instead of just connecting them with ‘>’ as done for simplex constructions.

[25] Esau traded his birthright to Jacob for a mess of pottage.

[26]	Syntax:	VSubj	V	NP	to+NP	for+NP
	Event 1:	<i>Source</i>		<i>Theme</i>	<i>Goal</i>	
		<i>Agent</i>				
	Event 2:	<i>Goal</i>			<i>Source</i>	<i>Theme</i>

The subject, *Esau*, is also marked as Agent, since the sentence leaves it clear that the trading was a result of Esau’s initiative.²¹ The two events are simultaneous and interdependent—otherwise some complements would lack a semantic role, which we know is not allowed. In the formulation given, each complement has at least one semantic role.

Without this dissociation into two events it is not possible to assign semantic roles to all constituents in [25] in an intuitively satisfactory way. This sort of analysis leads to a reformulation of some aspects of the traditional perspective. In a way, role-coding in this case follows the traditional model in that semantic roles are assigned as a result of the valency of the main verb (*trade*). But we cannot establish a simple relation between the constituents of the sentence and the semantic roles: we must define two concomitant events, and each of them defines a set of

²¹ There may be a better solution than marking this constituent as Agent in the diathesis; see Sect. 8.2.1 below.

semantic roles and maps them onto the appropriate constituents. In this particular case, the duality of events correlates with the existence of two Themes: the birthright and the mess of pottage, which both undergo motion,²² in opposite directions.

In cognitive space, what happens is the following: the constructional meaning with the verb *trade* entails assigning the role Source to the schema ESAU, Theme to BIRTHRIGHT, Goal to JACOB. And also (by effect of Event 2) Goal to ESAU, Source to JACOB, and Theme to MESS.OF.POTTAGE. As seen, two schemata, ESAU and JACOB, receive two roles each, but since these roles correspond to different events, no confusion arises. It is as if the receiver learned, first, that ESAU is the Source of the BIRTHRIGHT; then, she goes on to complete the landscape by learning that ESAU is also the Goal of the MESS.OF.POTTAGE, and so on, until the mental landscape intended by the speaker is complete.

Of course, this description of the process works only if we refer to schemata, not to sentence constituents. Here the distinction made in Sect. 6.1, that semantic roles are attached to schemata, not directly to constituents, become essential if we are to represent all the relations obtaining in [25] in a clear way (we shall see other examples further on).

6.6.2 More Examples

6.6.2.1 Ask and the Complex Construction

Let us now consider the verb *ask* in the meaning of ‘request’. In the sentence [27] Bob asked his father for some money.

the subject *Bob* is the Agent of *ask*; and, since *ask* is a **verbum dicendi**, the Agent is elaborated as the “speaker”. *His father* is the Goal (which elaborates as the “addressee”) of the same verb, ultimately of the schema ASK. But what is the semantic role of *for some money*? Verbs of saying usually include the possibility of a “message”, but this CSR does not seem satisfactory for *some money*, which is not in itself a message; the message would rather be “give me some money”. And, on the other hand, we understand the money as something that goes, or may go, from the father to Bob, that is, a Theme. This semantic ingredient is indispensable in the interpretation of [27].

Let us see how ASK is represented in recent cognitive models. In FrameNet we find the following (for the frame REQUEST, one of whose actualizations is *ask*):

a Speaker asks an Adressee for something [. . .]
[FrameNet, *data*, REQUEST]

²² More elaborately, transfer of possession.

The frame elements (which we may roughly equate with CSRs) are developed in this entry of FrameNet in this manner: the Addressee belongs to the semantic type Sentient; the Message (that is, the thing asked for) belongs to the semantic type Message, and is further specified as “the content of the request”; the Speaker belongs to the semantic type Sentient. FrameNet admits the possibility that the Message be represented as a complement, which would be *for some money* in [27].²³ This is clearly inadequate: (*for*) *some money* is a concrete object, not a message—unlike, for instance, *some questions* in *Bob asked some questions*, which does express a message. Note the clear difference of semantic relation between the verb and the complement in each case: Bob does not want the questions to be transferred to him, as is the case with the money in [27].

What is behind the CSR of *for some money* is a connection with another schema, GIVE. This is recognized by Wierzbicka (1996), who states that speech act verbs may refer to a (second) action which one asks for someone to perform:

[...] the verbs *ask* and *order* describe an attitude that includes the following component:
 (I say:) I want you to do it
 (Wierzbicka 1996, p. 174)

The component “you to do it” is a proposition, and can be expressed as a separate clause:

[28] Bob asked his father to give him some money.

and *some money* here is clearly Theme (but of *give*, not of *ask*).

Wierzbicka’s analysis is more satisfactory than the FrameNet’s. The thematic synonymy of [27] and [28] does not fit into FrameNet’s analysis, according to which *some money* would presumably be Message in [27], but Theme in [28]. As the two sentences are semantically equivalent, the FrameNet analysis reaches a deadlock. The analysis must clearly recognize that, as we saw in Sect. 6.1 above, CSRs are associated with schemata, not verbs, that is, they are relations defined at a conceptual level. In the case of [27], *some money*—or rather MONEY—is the “theme” of GIVE, not of ASK or of the verb *ask*.

Starting from the semantic equivalence of [27] and [28], I propose the following semantic analysis, valid for both sentences:

[29] Event 1:

ASK (BOB>*agent*, FATHER>*addressee*, EVENT 2>*message*)

Event 2:

GIVE (FATHER>*agent+source*, BOB>*goal*, MONEY>*theme*)

This is still a preliminary formula, to be completed below.

²³ One example given in the site is *the offender begged for clemency*.

6.6.2.2 Representing Complex Diatheses

One of the things that make the analysis of

[27] Bob asked his father for some money.

different from the analysis of

[30] Bob gave his son some money.

is that in [27] we have two events—in this case, one is real and the other is virtual, because it is not asserted that the father gave the money to Bob. They are both necessary to the characterization of the semantic structure of the sentence. [27] has a **complex** semantic structure because there is no way to map all the semantic roles directly onto the syntax. This is possible in [30], if we assign Agent to the subject and Goal and Theme to the objects, so that [30] does not need to be represented as a symbolically complex structure.

The semantic structure of a complex construction must then be expressed as more than one event. The diathesis instanced by [27] can be analyzed as

[31]	<i>Bob</i>	<i>asked</i>	<i>his father</i>	<i>for some money</i>	
	Ev.1: <i>Agent</i>	SAY	<i>Goal</i>		<i>Message: [Ev. 2]</i>
	Ev. 2: <i>Goal</i>	GIVE	<i>Source</i>		<i>Theme</i>
			<i>Agent</i>		

In this presentation of the semantics of [27], SAY represents a predicate of linguistic communication (expressed, cumulatively with GIVE, by the verb *ask*). The same semantic representation is valid also for [28], in spite of syntactic differences:

[28] Bob asked his father to give him some money.

Message is one of the semantic roles attached to the schema SAY, besides Agent (“speaker”) and Goal (“addressee”). The element **Ev. 2** has no syntactic form of its own; it is a conventional mark referring to the second event, and reference to a second event is part of the meaning of *ask*.²⁴ This type of semantic analysis derives from Jackendoff (2002, p. 365ff.); here I employ a slightly different formalism.

Complex analyses are regular for sentences with more than one clause, as in

[32] Bob told me his dog died.

Here we must analyze the semantic structure in terms of two events, “Bob told me Ev.2” and “Ev.2=his dog died”. What is exceptional with [27] is the need to use two events to analyze what is syntactically expressed as a single clause; this is attributable to features of the semantics of *ask (for)*, composed of SAY plus GIVE,

²⁴ English *ask* has two meanings: here we have the one that corresponds to Portuguese *pedir* ‘request’; the other meaning corresponds to *perguntar*, and appears in *Bob asked a question*—in this case, the construction is not complex.

each with its accompanying set of variables, basically as expressed by Wierzbicka in the passage previously quoted.

The descriptive character of the system we are elaborating allows us to ignore certain cases in which a complex analysis might be appropriate. For instance, the sentence

[33] The boy broke the bottle.

may, in principle, be analyzed semantically as

[34] Ev. 1: CAUSE (BOY>*agent*, Event 2>*caused.event*)
 Ev. 2: BREAK (BOTTLE>*patient*)

But this decomposition does not interest us because we can map the same semantic relations onto the syntax in a more direct way:

[35] The boy broke the bottle.
 Agent Patient

This simplification is possible because the events do not assign different semantic roles to the same constituent (as happens with *trade* and *ask*), so that the syntactic structure contains constituents directly relatable to all relevant semantic relations. The use of the complex construction in this case, although it is semantically defensible, is unnecessary for effects of the description of these constructions. These examples illustrate the descriptive character of the notation I am using (some might call it utilitarian, or even *ad hoc*). My excuse is that at this moment it is necessary to represent data in the most direct way, so that they can be examined and accumulated in view of further generalization and the eventual elaboration of an adequate valency theory.

At first sight, it may seem that the duality of events we are dealing with might be described in terms of Fauconnier's (1994) mental spaces. However, there is a basic difference, in that Fauconnier insists that mental spaces are built discursively and

result in unique and temporary 'packets' of conceptual structure, constructed for purposes specific to the ongoing discourse.

(*apud* Evans and Green 2006, p. 369)

Mental spaces, so defined, do not coincide with the events of complex structures, which are permanent and codified in the lexicon; that is, a schematic version of [29] is part of the meaning of the verb *ask*, stored in the language user's semantic memory.

6.6.2.3 Kinds of Messages

Let us go back to the semantic representation

[29]

Event 1:

ASK (BOB>*Agent*, FATHER>*Addressee*, EVENT 2>*Message*)

Event 2:

GIVE (FATHER>*Agent+Source*, BOB>*Goal*, MONEY>*Theme*)

We have here one event subordinate to another, which is a very common situation, particularly in compound sentences. But this subordination can occur in several different ways. Taking compound sentences for a moment, we may have

[36] Bob said that his father gave him some money.

[37] Bob revealed that his father gave him some money.

[38] Bob asked his father to give him some money.

In all of these cases we have the expression of an event subordinate to another (with variations in complementizer and verb form that do not interest us here). But the semantic relationship between the two events varies: in [36] Event 2 represents an assertion made by Bob, who is sole responsible for it; in [37] the truth of Event 2 is presupposed; and in [38] Event 2 represents the content of Bob's wish, which is not asserted nor presupposed, and does not have to be factual.

These differences result from the semantics of each of the verbs: *say* is a different thing from *reveal* or *ask*. The “message” (defined as “information transmitted to an addressee”) comes in at least three flavors²⁵: **assertion**, **presupposition**, and **wish**. These distinctions must be marked in the structure, so that [29] is reformulated as follows:

[39]

Event 1:

ASK (BOB>Agent, FATHER>Addressee, EVENT 2>Message.wish)

Event 2:

GIVE (FATHER>Agent+Source, BOB>Goal, MONEY>Theme)

[39] makes explicit the semantic relationship between the two events as **Message.wish**—a relationship which I analyze as the semantic role of Event 2. In the other examples, naturally, we have **Message.assertion** and **Message.presupposition**. All these are components of the meaning of the verb in question: one of the semantic features of *reveal* is that its subordinate event is presupposed, but with *say* it is only asserted, and with *ask* it is desired.

These relations can be considered semantic roles because they play a role in the subclassification of verbs; as was observed by Kiparsky and Kiparsky (1971), verbs can be divided into two categories, **factives** and **non-factives**, according to the semantic relation of their complements, respectively presupposed and asserted (here we have a third category, verbs that express a **wish** in their complements).

²⁵ Plus, certainly, several others.

6.6.2.4 Notational Questions

Let us now examine a few examples. In the case of *buy*, both events involve the schema TRANSFER.POSSSESSION. It then becomes possible to represent the diathesis in the following way:

[40]	VSubj	V	NP	from+NP	for+NP
Ev. 1:	<i>Agent</i>	TRANSF.POSS	<i>Theme</i>	<i>Source</i>	
	<i>Goal</i>				
Ev. 2:	<i>Source</i>	TRANSF.POSS	<i>Goal</i>		<i>Theme</i>

This can be expressed as the sentence

[41] Leah bought a car from Jim for \$12,000.

Leah is the Agent (because she initiated the event)²⁶ and the Goal of the car, being also the Source of the money. Jim, Source of the car, is also the Goal of the money. In both events the schema is TRANSFER.POSSSESSION. The distinction into two events allows proper representation of the semantic connections: Leah is Goal in Event 1 (whose Theme is the car), but Source in Event 2 (whose Theme is the money). In this case, the relation between the events is just one of simultaneity and does not involve wishes or messages; both events are simply asserted.

In the case of [27]

[27] Bob asked his father for some money.

two schemata are involved, SAY and GIVE,²⁷ which must be represented separately:

[42]	<i>Bob</i>	<i>asked</i>	<i>his father</i>	<i>for some money</i>	
Ev. 1:	<i>Agent</i>	SAY	<i>Goal</i>		Message.wish: [Ev. 2]
Ev. 2:	<i>Goal</i>	GIVE	<i>Source</i>	<i>Theme</i>	

SAY and GIVE are both ingredients in the semantics of *ask*. Neither SAY nor GIVE has a separate formal representation in this sentence: *ask* at the same time expresses the emission of a message (SAY) and gives information about this message: the WISH that someone GIVE something. By replacing *Bob* by 'VSubj', *asked* by 'V', *his father* by 'NP', and *for some money* by 'for+NP', we obtain diathesis [43], which sentence [27] elaborates. As we saw, the marker 'Ev. 2' has no separate formal representation, and serves as a semantic connector between the two events; and the element 'Message.wish' can be considered the semantic role of Event 2.

²⁶ If Jim were the Agent, we would say *Jim sold a car to Leah for \$12,000*.

²⁷ The latter may be again TRANSFER.POSSSESSION, but I use GIVE as a slightly more elaborated variety (which is the one that appears in the FrameNet entry).

[43]	VSubj	V	NP	<i>for</i> NP	
Ev. 1:	<i>Agent</i>	SAY	<i>Goal</i>		<i>Message.wish</i> : [Ev. 2]
Ev. 2:	<i>Goal</i>	GIVE	<i>Source</i>	<i>Theme</i>	

Diagram [43] represents semantic relations in a straightforward way, placing each semantic role immediately below the constituent that elaborates it—for instance, the Agent of Event 1 is VSubj, the Theme in Event 2 is *for*+NP, etc. Events are partially represented by schemata, here SAY and GIVE. We must also find the ways in which event schemata associate with verbs—and sometimes, as we saw for light verbs, with other constituents—in the appropriate contexts. We cannot go into this question now, because our main problem is how to represent the relations between constituents and verbs. For the time being, we must be content with the intuitive perception that the meaning of *ask* (*for* something) includes the emission of a message (SAY) and the WISH for a TRANSFER OF POSSESSION.

6.6.2.5 Character of the Semantic Representation

It should be clear that saying that *ask* is represented semantically by SAY plus GIVE is not to propose anything in the line of lexical decomposition as used by generative semanticists in the 1970s. This is only a way to represent a complex meaning; and the relation between this meaning and a sentence with *ask* is symbolic, and the two levels of analysis, formal and semantic, are kept carefully apart, as the two faces of the sign.

As noted above, the semantic representation used here for diatheses is somewhat *ad hoc*, in the sense that its details do not necessarily refer to some general theory. In the absence of such a general theory, we have no choice: the aims of the notation are immediate and strictly descriptive, and, up to a point, arbitrary. We are attempting to represent something on the page, and we do not really know how it “is”. This does not entail, obviously, that it is not necessary to eventually develop a theory to reduce the degree of arbitrariness of the description; it means only that (as far as I know) no tolerably adequate theory currently exists.

In [43] we find the element ‘Event 2’, which refers to the second event and has no lexical representation. This is certainly not an “empty category” present in syntactic structure, but a notational means of semantically tying the second event to the first, at the same time making explicit the semantic role of the second event (*Message.wish*). In the model adopted in this book, syntactic structure is made up exclusively of overt units, endowed with phonological representations. These units have also a semantic face, which is part of the data, but the separation of the two levels, form and meaning, must be rigorously kept, lest we vitiate the objectives of the research, which have to do with describing the association between concepts and acoustic images.

The position here adopted differs in crucial ways from Hale and Keyser’s (1993). Hale and Keyser correctly observe that when we decompose simple

sentences as distinct events it becomes frequently easier to identify the semantic roles. This greater ease comes, I think, from the conceptually more basic character of the evoked schemata (for instance, SAY and GIVE as part of the meaning of *ask*). But Hale and Keyser propose a syntactic analysis of the phenomenon, presumably deriving a sentence like [27] from a syntactic structure containing the verb *give*, eventually incorporated with *say* to generate the surface verb *ask*. Hale and Keyser's solution increases the role of the syntactic component (enlarging it with a transformational, or move- α , section) in order to simplify the semantic component; the serious problems caused by this kind of analysis were convincingly shown by Culicover and Jackendoff (2005, chap. 1).²⁸ Furthermore, Hale and Keyser's analysis, which cannot be examined here in detail, seems weakened by the acceptance of some assumptions not duly founded on empirical evidence.

6.6.3 Reciprocals

Complex constructions provide a means to analyze sentences with reciprocal reference, such as

[44] Tiago e Jane se odeiam. 'Tiago and Jane hate each other'

Because in Portuguese the reflexive pronoun (here, *se*) also serves as a reciprocal marker, [44] is ambiguous: it can also mean 'Tiago and Jane hate themselves'. In this reading, [44] can be analyzed as a noncomplex construction, simply by assigning Experiencer to the subject and Stimulus to the reflexive pronoun.²⁹ But in its reciprocal reading 'Tiago hates Jane and Jane hates Tiago', [44] must be analyzed as a complex construction. Here we have two concomitant statements, as is also the case with *sell*, *buy*, *trade*, seen in Sect. 6.5.1; we can, then, state that [44] is the realization of a construction defined as

[45] [NP e NP]_{vSubj} Refl V
 Ev. 1: *Exp* *Stim*
 Ev. 2: *Stim* *Exp*

Or, perhaps better,

[46] [NP e NP]_{vSubj} Refl V
 Ev. 1: R_i R_j
 Ev. 2: R_j R_i

(where R = semantic role)

²⁸ Culicover and Jackendoff (2005) criticize Hale and Keyser's proposal directly on pp. 53ff.

²⁹ The reflexive pronoun is independently marked as being coreferential with the subject: this is its "meaning".

to include all cases of reciprocals, regardless of the semantic roles involved: *Tiago and Jane hate each other* (Experiencer—Stimulus)/*Tiago and Jane kissed each other* (Agent—Patient) etc.

Complex constructions are a solution to cases of “differences between what one could term semantic valency and syntactic valency”, to use Herbst and Schüller’s expression (2008, p. 136). What we have here is a point where the usual rough parallelism between syntax and semantics fails, so that it becomes difficult to relate semantic roles directly to morphosyntactic units. The case mentioned by Herbst and Schüller is the semantic similarity between

[47] Dick met Martha.

and

[48] Dick and Martha met.

In [47], we have two participants and two NPs, and there is no special problem in expressing the coding relations. But [48] means (in terms of semantic roles at least) just the same as [47], and yet only one NP is present.³⁰ Here, again, the solution is to include in the semantic matrix of the verb *meet* the possibility of expressing two simultaneous events, [A *meet* B] + [B *meet* A] following the model of [46]. This possibility sets *meet* apart from other verbs, in that with *meet* reciprocity does not have to be overtly marked with *each other*, as it must be with *pinch*, *hate*, *love*, *kill*, *see*.

Herbst and Schüller (2008), who found the examples and correctly saw the problem, do not develop a practical analysis for the phenomenon, and merely mention a “merger of participants”, which I find somewhat vague, and may or may not be the basis for a convenient analysis. [46] offers us another possibility, which falls under the already known case of complex constructions.

[46] shows a feature that may seem strange: it includes consideration, at the sentence level, of subsentential constituents—that is, the two NPs that compose (with the conjunction *e* ‘and’) a higher NP. That is, in this formula we had to “break” an NP (*Tiago e Jane*) in order to mark its constituent NP’s with the respective roles. It thus violates a general principle of sentence analysis, sometimes named the “A-over-A principle”.³¹ Our purpose here is merely descriptive, and [46] seems sufficiently clear as a representation of the semantics of the relevant reading of [44]. It may be that the A-over-A principle has exceptions, or it may apply to syntax but not to semantics, and so on. In any case, the fact remains that sentence [44] asserts a thematic relation between these subordinate NP’s, and the most direct way to represent this fact seems to be the one given above; for immediate purposes, it will have to do as it is.

³⁰ A complex NP, to be sure: [(Dick)_{NP} and (Martha)_{NP}]_{NP}.

³¹ Chomsky (1964, 1973). The A-over-A Principle was originally formulated as a condition on the application of transformations, whereas here we are dealing with an interpretive (symbolic) phenomenon.

But there may be more to be said about these cases. Note that the complex NP found as the subject in [44] and [48] can be replaced by a single plural, as in

[49] *Essas meninas se odeiam.* ‘these girls hate each other’

and yet the system has no difficulty in decomposing the plural in its elements, assigning them separate semantic roles, which seems to indicate that the process here is fundamentally semantic.

The best way to account for this apparent problem is to consider that when we talk about roles assigned to the constituents of a complex NP we are in fact dealing with roles (CSRs) assigned to variables of the corresponding schemata. In other words, we are not dealing with a purely grammatical phenomenon. It is only in cognitive space that *Sally and Helen* and *these girls* are parallel: in both cases we have more than one person, although morphosyntax shows a composite NP or a plural one, respectively. [43] and [44] must, therefore, be amended in order to show what is really going on here: not an association of semantic roles with NPs, but rather with schemata. And we must find a way to make explicit that plurals involve more than one individual, each with its schema. *Girls* is, cognitively, something like [GIRL + GIRL + . . .], and it is on this structure that the assignment of the roles works.

Here we have a case where just using syntactic units makes it impossible to properly state the phenomenon. And this generalizes for all cases where a plural is parallel with a collective or a composite NP, as against a singular. This appears for instance in

[50] *The couple/they/Sally and Paul/split up.*

[51] * *Sally split up.*

Instead of devising an *ad hoc* feature to help explain the parallel behavior of collectives, plurals and composite NPs, we use their property of referring to more than one entity, and therefore evoking more than one schema. Here the phenomenon mentioned in Sect. 6.1 becomes evident: strictly speaking, the assignment of semantic roles is something that happens to schemata, not to overt sentence constituents.

Consequently, a really rigorous way to state the distribution of semantic roles in a sentence like [49] (in its reciprocal reading)

[49] *Essas meninas se odeiam.* ‘these girls hate each other’

requires a previous analysis of the NP *essas meninas* ‘these girls’ into the corresponding schemata, [GIRL + GIRL + . . .] in the way suggested above. I cannot go into this matter more deeply here, since it calls for a revision of part of our analyses. But a task that must be taken up eventually is the restatement of diatheses (and constructions in general) in terms of the association of semantic roles with schema variables, instead of sentence constituents.

Finally, it should be observed that the reciprocal construction, represented in [46], is not lexically governed, and applies whenever the structural conditions described are met; that is, it is not a diathesis of Portuguese. On the other hand,

the English construction exemplified in [47] is a diathesis, since it only works for some verbs: *Dick and Martha met* has a reciprocal reading, but we do not have **Dick and Martha saw* meaning “Dick saw Martha and Martha saw Dick”.

6.7 Primary and Secondary Events

As already observed by Jackendoff (1972, p. 35), the events that make up a complex construction’s semantic face are not necessarily equal in importance. Jackendoff speaks of *primary and secondary action*, and this is what distinguishes the meaning of *buy* in

[41] Leah bought a car from Jim for \$12,000.

from the meaning of *pay* in

[52] Leah paid \$12,000 to Jim for a car.

Although the scene depicted is the same, the first sentence profiles the transference of the car, and the second profiles the transference of the money. We can represent this in our formulas by stipulating that the numbering of the events (Event 1, Event 2) is significant, only the first event being profiled. As we see, then, profiling can be an integral part of a verb’s valency, and consequently of its meaning: it is only a difference in profiling that distinguishes *buy* from *pay*.

Here we have a notion of profiling that differs a little from the way it is usually defined and exemplified. According to Langacker (1991, p. 9), profiling refers to “an asymmetry in the portrayal of the relational participants”. Here we have a similar asymmetry referring to events, not to individual participants, but I believe the general phenomenon is basically the same, and until further notice I refer to it as **profiling**. It is important because it instances a nonpropositional ingredient—that is, a kind of construal—in the meaning of verbs.