DOI: 10.20396/joss.v14i00.20417

SPACES, PARALLELISM, AND THE NOMINAL CLAUSE

PERINI, Mário Alberto[1*]

[1] Professor Emeritus, Universidade Federal de Minas Gerais – ORCID: https://orcid.org/0000-0001-7946-5406

Abstract: The rules of a language interact with world knowledge in order to generate the meaning in certain verbless sequences, when the latter are conjoined to clauses. In such cases the cognitive representation corresponding to the verbless sequence is completed through the insertion of a copy of the schema evoked in the clause, so as to generate a well-formed cognitive structure. This happens by effect of a rule of schema retrieval, which is part of the structure of the language, but is also subject to grammatical, semantic, and symbolic conditions, which instantiate a degree of interference between cognitive and formal spaces. This process is triggered by the presence of certain coordinating conjunctions, in particular *e* 'and', and is subject to conditions of semantic parallelism between the two members of the construction. The phenomenon is studied in relation to the data of standard Brazilian Portuguese.

Keywords: Cognitive grammar; schemas; syntax; Portuguese; coordination.

[*] Corresponding author: marioperini7@gmail.com

1 Introduction

In my recent work (Perini, 2019; 2021; 2024) I have examined several cases of what are usually termed grammatical phenomena, and come to the conclusion that they are often the product of the interaction between grammatical rules, on one hand, and principles and constraints based on world knowledge on the other. In this article I study some factors determining the occurrence, structure, and interpretation of verbless clauses in Portuguese, and show that they also undergo the effect of such interaction between language and cognition[1].

Here I am chiefly concerned with the description of the facts. I could not avoid theoretical observations, while attempting to faithfully represent the data as they can be observed; they may call for further theoretical interpretation, but this is beyond the aims of the present paper.

2 Verbless clauses

A sentence like

[1] Ana adora passas, e Joaquim amendoim.
'Ana loves raisins, and Joaquim [loves] peanuts'

means that Joaquim *loves* peanuts, although there is no occurrence of the verb *adora* 'loves' in the second half of the construction. Some kind of retrieval takes place, so that one ends up understanding that the thematic relation of Joaquim to peanuts is the same as the relation of Ana to raisins: Ana and Joaquim are Experiencers, raisins and peanuts are Stimuli.

What is retrieved in the second half of [1] cannot be the *verb*,[2] unless one is willing to pay for some useless complication in the analysis. For instance, in

[2] Eu adoro passas, e Joaquim amendoim.
'I love raisins, and Joaquim [loves] peanuts'

the verb in the first half of the construction, *adoro*, is inflected for the first person singular, which cannot occur with *Joaquim*, a third person NP, as its subject. If we insist on the verb-retrieval analysis, we must add that person-number inflection is in some way disregarded for retrieval purposes. This complication can be avoided if we assume that in the second half of [1] what we retrieve is the schema LOVE, which occurs in cognitive space and not in the sentence proper[3].

Another reason for retrieving the schema is that it saves the need to postulate the presence of a verb in the second half of the construction, only in order to explain why we understand that Joaquim *loves* peanuts. The schema, LOVE, is "already there", as clearly shown by speakers' introspection, and the real problem is to identify it, which is done by reference to the schema evoked in the first half. The latter is evoked by a verb, but this verb is overt in the structure, and no abstract element need be postulated.

The proposed solution frees us from the need to postulate an underlying verb, and then delete it again, in order to explain the fact that in [1] we understand "Joaquim loves peanuts".

[1] Cognition, of course, includes the knowledge of a language; but here I use this term as short for "nonlinguistic knowledge".
[2] I refer to the *halves* of the construction, since the second one has no verb and is not a clause, rigorously speaking.
[3] I distinguish three *spaces*, namely the grammatical, the cognitive and the symbolic, which relates the other two; in this I follow Langacker (1991) – see quote in section 3.2.5 below.

When one tries to construct a cognitive representation (**CR**) based on the information provided in [1], the first half presents no problem, but the second half, made up of two NPs in sequence, cannot, at first sight, provide a well-formed CR. A rule of the language then allows us to construct this CR on the model of the first one, with a repetition of the first schema, LOVE; a process whereby we retrieve information, not words or syntactic structure. The problem of the form of the second verb in [2] disappears, since the schema is not constrained by verb agreement, which is a rule of the language and does not apply to cognitive representations.

This solution also places the retrieval of the second LOVE in cognitive space, which is its proper location, since the LOVE connected to Joaquim and peanuts is not formally realized, but is accessible by introspection as part of the meaning of the sentence. It may be noted that the solution is not entirely new, and its basic elements can be found in the literature since the 1970's; here I examine additional examples and develop some aspects that have not been considered so far. This process is one aspect of the general phenomenon of anaphora-solving, and may seem simple at first sight; but there are several complexities involved, as we shall see below.

3 Retrieving information

3.1 The conjunction *e* 'and' as a trigger for retrieval

The first thing to consider is that application of the schema retrieval rule seems to be conditioned to the presence of the conjunction *e* 'and' or other marks of coordination. In our example *e* 'and' is found to conjoin a clause and a sequence of two NPs. Detection of *e* 'and' triggers the rule, and links to the properties of the particle, of which the one that concerns us is that

the particle *e* 'and' conjoins parallel structures

This condition has been stated in the literature thus:

> In a typical coordinate structure, all of the conjoined constituents (the **conjuncts**) are of the same category, and the whole structure is an instance of the same category [...]
>
> [Trask, 1992, p. 63]

which I take to be roughly equivalent to the rule above. But being "an instance of the same category" is too vague an expression; as will be seen below, the notion of parallelism, or belonging to the same category, can be stated in more precise terms.

3.2 Parallelism

3.2.1 Cognitive types

To start with semantic conditions, we may observe that some combinations of cognitive types are forbidden; for instance, the conjoining of an "event" with a "thing"[4]:

[4] * A gata caiu do nono andar e morte imediata.
'the cat fell from the ninth floor and sudden death'

This is not all; there are other constraints, like for instance the impossibility of conjoining qualifying and classifying adjectives:

[4] I do not mark unacceptability in the English glosses, which are not to be taken as examples from the language.

[5] * Um engenheiro simpático e mecânico
'a nice and mechanical engineer'

although these adjectives can co-occur if there is no conjunction:

[6] Um engenheiro mecânico simpático
'a nice mechanical engineer'

3.2.2 Thematic relations

Parallelism also has to do with thematic relations; for example, one can say

[7] João chegou da Austrália, e Maria da China.
'João arrived from Australia, and Maria [arrived] from China'

Here the schema ARRIVE is retrieved in the second half, as expected. But the following does not work:

[8] * João chegou da Austrália, e Maria na China.
'João arrived from Australia, and Maria to China'

Although the semantic roles involved (Source for *da Austrália*, Goal for *na China*) are duly marked by prepositions, the result is unacceptable. Therefore, the requirement of parallelism seems to apply to the semantic roles as well, just as it does in [1]: retrieving the schema from the first half of the construction entails retrieving the same semantic roles for the overt complements. This is certainly a consequence of the fact that the schema is retrieved in its entirety, including its realized variables, and explains why the second half of [1] includes an Experiencer and a Stimulus. But it does not explain why [8] is unacceptable: the schema CHEGAR has as its variables both Source and Goal; but [8] shows that we must have *the same* thematic assignments in both members of the construction: Source in the first half and again Source in the second one, but not Source in the first and Goal in the second.

3.2.3 Order of constituents

Parallelism has also syntactic aspects. For instance, in

[9] * Ana adora passas, e amendoim Joaquim.
'Ana loves raisins, and peanuts Joaquim'

we can retrieve a second occurrence of the schema LOVE, just as was done in [1]; but here the result is not acceptable, although the CR includes all the right components. The problem is that *amendoim* 'peanuts' cannot be the Experiencer; but why cannot just assign this NP the role of Stimulus, which will yield an acceptable result, with *Joaquim* as the Experiencer? The only difference is that *amendoim* 'peanuts' is syntactically parallel to *Ana* in the first member of the coordination. That is, the two members must be assigned their semantic roles in a respective manner: the first constituent in the first half (that is, the subject *Ana*) gets the same role as the first constituent in the second half (that is, *amendoim* 'peanuts'), and so for the second constituent in each half of the construction. This is an example of interference of a formal factor (order of

constituents) on a cognitive phenomenon (semantic role assignment). Let us call this the **condition of respective assignment (CRA)**, a device triggered by the presence of *e* 'and' as a conjoining element between two sequences.

In [9] the result is unacceptable also because of the meaning of *amendoim* 'peanuts', which is not eligible as an Experiencer. But even when the result is in principle acceptable the CRA must apply, as in

[10] Ana beliscou Sérgio, e Carolina Breno.
'Ana pinched Sérgio, and Carolina [pinched] Breno'

In principle we might understand that Ana is the pinching person, and so is Breno. But this interpretation is not possible for [10]: role assignment follows a respective order, so that we must understand that both Ana and Carolina did the pinching, and Sérgio and Breno were pinched. This shows that the constraint has a syntactic component; and we have to add the condition of respective assignment (CRA) to the properties of *e* 'and' – or, perhaps better, to the conditions on schema retrieval in the presence of *e* 'and'.

What prevents *Joaquim* to be assigned the role Experiencer in [9], which would make this sentence synonymous with [1], is that it comes last, after *amendoim* – a syntactic feature, to be taken into account when applying the CRA in this example. The CRA is then a symbolic, not semantic, constraint. It resembles the rules that assign semantic roles to terms of the clause, and which are included as diatheses in the valency of the verbs; the difference is that the CRA makes use of different syntactic relations, apparently just the order of constituents (since *Carolina* in [10] is not a subject, there being no verb in this half of the construction), while diatheses may refer to "subject", "object" and the like. For our immediate purposes, then, we can label it a valential rule, to apply to constructions of the form

[11] VSubj V NP e NP NP

After assignment, we have

[12] VSubj>SR_i V NP>SR_j e NP>SR_i NP>SR_j

where 'SR' means "semantic role", and the subscripts, *i* and *j*, refer to specific roles – for example, in [10] *i* is Agent, and *j* is Patient.

3.2.4 Functional preposition requirement

Another case of syntactic interference on the retrieval of the schema can be seen in sentences like

[13] * Ana gosta de passas, e Joaquim amendoim.
'Ana likes [of] raisins, and Joaquim peanuts'

The verb *gostar* 'like' requires a preposition, *de*, before its nonsubject complement, so that **Ana gosta passas* is not acceptable. This is a syntactic requirement, and has nothing to do with the semantic roles involved; in principle, the second half of [13] could be interpreted in a manner parallel to the first half, with *Joaquim* as the Experiencer and *amendoim* 'peanuts' as the Stimulus. Yet the sentence is marked as ill-formed because the (syntactic) requirement of the presence of *de* is violated in the second half; but the requirement depends on the verb, *gostar* 'like', and the second half contains no verb. This seems to indicate the existence of another aspect

of respective assignment: the *form* of the complements in both halves of the constructions must be the same in some respects – in [13], both must be **[NP ... *de* NP]**. And in fact, if we add the right preposition the sentence becomes acceptable:

[14] Ana gosta de passas, e Joaquim de amendoim.
'Ana likes [of] raisins, and Joaquim [of] peanuts'

The preposition *de* in *de passas* '[of] raisins', as well as in *de amendoim* '[of] peanuts' has no semantic content; its repetition functions then only as a sort of reinforcement for the parallelism already established by the CRA. Now, in

[15] Ana viajou sem dinheiro, e Joaquim sem passaporte.
'Ana traveled without money and Joaquim without a passport'

the two complements introduced by *sem* 'without', besides being semantically parallel, are also marked by a meaningful ("predicating") preposition. [15] can be given a purely semantic analysis, unlike [13], where the preposition says nothing about the thematic relation involved. The complements in [15] are not valential, but rather semantically autonomous, so that they can also have different semantic roles, as in

[16] Ana viajou sem documentos, e Joaquim com passaporte.
'Ana traveled without papers and Joaquim with a passport'

In other words, we can explain the presence of *sem* 'without' in the second half of [15] in semantic terms, based on the meaning of *sem* (and also of *com* in the second half of [16]). In [15] there are at least two ways of retrieving the semantic role of the complement in the second half: by CRA or by the meaning of the preposition. But this explanation fails for [13], where the presence of *de* in both halves of the construction is not determined by meaning considerations; the verb, *gostar* 'like', requires a semantically empty preposition *de* in its complement, even in the second half, where no verb is present, and the semantic role of the second prepositional phrase depends entirely on the valency of the verb.

3.2.5 Parallelism and space interaction

The conditions seen in 3.2.1 – 3.2.4, which define parallelism in what concerns us, are not all similar; some refer to meaning, some to form, some to the relation between the two. Langacker puts the distinction in these terms:

> [...] I assume that it is reasonable to speak of "phonological space" and "semantic space" as two facets of cognitive organization; jointly they define an abstract, bipolar "symbolic space". Phonological, semantic, and symbolic units are characterized relative to these respective domains. The list is exhaustive: every linguistic unit is phonological, semantic, or symbolic.
>
> [Langacker, 1991, p. 116-117]

I usually prefer to call Langacker's "phonological space" just "formal", or "grammatical", which seems more adequate in a discussion on grammar. In any case, the four factors seen above as conditioning the occurrence of conjoined structures with *e* 'and' can be classified thus:

The requirement of **similarity in cognitive type** (as seen in example [5]) is purely

semantic;

the requirement of **identity in thematic relations** (cf. example [8]) is also purely semantic;

the requirement that **complements with specific thematic roles appear in the same order in both halves of the construction (CRA)** (cf. example [9]) is symbolic; and

the requirement that **functional prepositions be kept in the second half** (cf. example [13] is grammatical.

While each procedure is neatly placeable in one of the three spaces, they interact in order to mark each sentence of the form shown in [11] as acceptable or not. There is no reason to order these conditions among themselves, or in relation to other rules of the language; they can be understood as a (partial) definition of what is acceptable in Portuguese as a sentence made up of two members conjoined by *e* 'and'. As such, these conditions are simultaneous available to the language user to apply whenever the right environment appears, as a system of controls, rather than a set of rules[5].

3.2.6 Syntactic skewness

In spite of the existence of at least one syntactic condition on the application of the retrieval rule (the requirement that functional prepositions be kept in the second half), there must be some tolerance for syntactic skewness; to start with, the second half of [1] must be defined as syntactically well-formed, otherwise it would not occur in this form (NP NP). Here a *semantic* ill-formedness is detected, and gives rise to the need for retrieval; since a rule exists to that effect, it applies, and the result is semantically well-formed. One consequence is that two sequences conjoined by *e* 'and' do not have to be syntactically identical in all levels. Other examples of different structures so conjoined allowed by the syntax of the language are

[17] Detesto banheiro no mato e ter que armar a barraca toda noite.
'I hate outdoor bathrooms and having to set up a tent every evening'

Here we have two NP's, with different internal structures; and

[18] Beth era simpática e a mulher mais rica da cidade.
'Beth was nice and the richest woman in town'

where we have an adjective phrase and an NP, conjoined by *e* 'and'. These examples show that the definition of parallelism still needs work.

3.3 Other conjunctions

The condition that schema retrieval occur in the presence of the conjunction *e* 'and' can be seen as a formal condition, in that it singularizes a lexical item as a trigger for a semantic process. On the other hand, it could be just the result of some semantic ingredient present in *e* 'and', which would mark it as a semantic factor after all. However, the available evidence suggests that *e* 'and'

[5] This is in harmony, I think, with the principles proposed by Jackendoff (2002, p. 198).

is a formal element that triggers schema retrieval in the second half of the construction, not a carrier of a particular semantic feature. This can be seen when we test with some near-synonymous connectives; thus, *assim como* 'as well as', also an additive conjunction, does not allows retrieval of the schema from the first half:

[19] Alice toca oboé, assim como Júlia toca violino.
'Alice plays the oboe, as well as Júlia plays the violin'

[20] * Alice toca oboé, assim como Júlia violino.
'Alice plays the oboe, as well as Júlia violin'

But *nem* 'nor' does allow for retrieval:

[21] Ana não aprecia passas, nem Joaquim amendoim.
'Ana does not like raisins, nor [does] Joaquim peanuts'

If these few examples are typical, we can say that retrieval of the schema in the first half depends on lexical selection, therefore on a formal factor. Now, this seems to work only with *additive* conjunctions, which suggests that there is also a semantic factor at work: retrieval depends on the presence of some (individually marked) conjunctions, all of them with additive content.

4 The schema retrieval rule

We need then a rule to effect the retrieval of the schema evoked in the first half of construction [1], and copy it in the second half – or, rather, in the CR generated by the second half. The rule, as we have seen, is subject to some constraints, of various orders, some being semantic, some symbolic, and some syntactic. That this is one of the rules of the language, not a cognitive constraint of some sort, is shown by the fact that its details are language-specific. For instance, take the Latin sentence

[22] Caesarem Brutus, Remum Romulus occidit.
Caesar Brutus Remus Romulus killed
'Brutus killed Caesar, and Romulus [killed] Remus'

Here we have much the same as in [1], but in reverse order: the apparently defective first half of the CR is filled in by a replica of the schema evoked by *occidit* 'killed' in the second half; this comes from a grammatical difference between Portuguese and Latin.

Another example of the interaction of grammatical factors is the sentence

[23] Alda comprou muitos livros mas não leu.
'Alda bought many books but did not read [them]'

There is a lacuna in the second half of the construction, which is filled in by a copy of the schema BOOK; the language provides a rule for that (not the same rule that applies in [1], of course). But in Italian this rule does not exist, and it is mandatory to include a pronoun in the second half in order to retrieve the second occurrence of BOOK:

[24] Alda ha comprato tanti libri ma non *li* ha letti.
'Alda bought many books but did not read them'

Li is the pronoun that retrieves the schema BOOK from the first clause, and cannot be omitted. Here again each language makes its own demands on the conditions for retrieving information from other parts of the sentence.

5 Summary

Let us now summarize what happens during the processing of sentence

[1] Ana adora passas, e Joaquim amendoim.
'Ana loves raisins, and Joaquim [loves] peanuts'

First of all,[6] an inspection device detects the occurrence of *e* 'and' joining two structures which are not identical, syntactically or cognitively. In this particular case, the language has a rule to repair the situation in cognitive space, retrieving the schema evoked by the verb and inserting it in the CR corresponding to the second half of the construction (which is the defective one).

The rule is language-specific, and thus not part of world knowledge. It is symbolic, in that it relates a syntactic situation (presence of *e* 'and') with a semantic process (retrieval of a schema). And it is subject to several conditions, some syntactic, some semantic, as seen in 3.2 above. Since in [1] all conditions are properly met, it is acceptable, and generates a well-formed cognitive representation.

The example of construction defined in [11] illustrates some aspects of sentence processing, involving knowledge of the language (syntactic conditions) and also world knowledge (as in the requirement of parallel semantic types and the presence of the same thematic relations in both halves of the construction); and, additionally, symbolic factors, such as the requirement that complements with specific thematic roles appear in the same order in both halves of the construction. The result is a surprisingly complex system that must be put into action in order to ensure that these sequences be acceptable and cognitively well-formed.

REFERENCES

1. Jackendoff, Ray. 2002. *Foundations of language – Brain, meaning, grammar, evolution*. Oxford University Press.

2. Langacker, R. W. *Concept, image, and symbol: The cognitive basis of grammar*. Berlin: Mouton De Gruyter, 1991.

3. Perini, M. A. *Thematic relations: A study in the grammar-cognition interface*. Cham, Switzerland: Springer, 2019.

4.Perini, M. A. *Function and class in linguistic description: The taxonomic foundations of grammar*. London: Palgrave Macmillan, 2021.

5. Perini, M. A. *Linguagem e cognição: Esquemas, gramática e os limites da língua*. Petrópolis: Ed. Vozes, 2024.

6. Trask, R. L. *A Dictionary of grammatical terms in linguistics*. London: Routledge, 1992.

[6] "First" only in the present exposition; as explained above, the process is global: all needed information is simultaneously accessible, without any need for ordering of rules and constraints.