

Author: Renata Enghels
University/Affiliation: Ghent University
Email address: Renata.Enghels@UGent.be

How word order frequencies reveal cognitive schemes: a Romance case study

1. Introduction: word order in infinitive complements

This poster presents the results from an ongoing corpus analysis of word order in Romance infinitive complements (InfC). In the languages considered, namely Spanish, French and Portuguese, this complement type appears with two main verb classes, namely perception verbs (PVs) and causative verbs (CVs). The InfC is a subordinated complement type containing a nominal constituent (NP₂) responsible for the process represented by the infinitive, as illustrated by the following sentences:

Spanish:

- (1a) [...] *esperando ver [entrar a un doctor joven y atrevido que le diría, sencillamente: “Vamos”]*_{InfC}. (SOL: Palomino A., 1971)
'[...] waiting to see a young and impudent doctor who would simply say to him: “let’s go”.'
- (1b) *Pero era sólo la lluvia que hacía [crujir las ramas secas del acebuche]*_{InfC}. (CREA: Maqua J., 1992)
'But it was just the rain that made the dead branches of the oleaster crack.'

French:

- (2a) [...] *c’est la coutume de voir [le pouvoir échapper à ses détenteurs légaux]*_{InfC}. (LM: 12/2/1994)
'[...] it is customary to see the power escape from its legal rulers.'
- (2b) *Laisse [pousser tes cheveux]*_{InfC} [...]. (Frantext: Weyergans F., 1981)
'Let your hair grow [...].'

Portuguese:

- (3a) [...] *ouvi [um médico dizer para outro que me deviam fazer um TAC [...]]*_{InfC} (CDP: O Público, 1994)
'[...] I heard one doctor say to another one that they had to make me a CAT scan.'
- (3b) *E a porta deixa [passar Cajango e a mulher]*_{InfC}, *um ao lado do outro, [...]*. (CDP: Amaral M., 1992)
'And the door lets Canjango and his wife pass through, side by side [...].'

In Germanic languages such as English and Dutch, the position of the subordinated nominal NP₂ is fixed, always appearing before the infinitive, as showed by the above translations. In Romance languages however, it varies: it can occur before or after the infinitive. This observation leads to the question of why these different word orders exist and by which parameters they are determined.

A comparative/contrastive method is used, since it allows investigating whether the observed correlations are language specific or whether they can be linked to cross-linguistically (though not necessarily universally) valid cognitive schemes. To be more

precise, a quantitative analysis of real discourse examples in three different languages will show that:

- (a) word order in the Romance InfC is largely determined by the semantics of the main verb;
- (b) this semantics has an impact on the relationships between the main participants (that is, the main subject and the subject of the infinitive complement) of the situation represented by the sentence;
- (c) word order can reveal different cognitive schemes or ‘dynamicity configurations’.

2. Data collection

Most previous analyses of the syntax of InfCs after PVs (eg. Rodríguez Espiñeira 2000) and CVs (eg. Treviño 1994) lack any empirical foundation, and do not accurately distinguish different verb types (Danell 1979). Therefore, in order to achieve the above-stated goals, the present study builds on a large corpus containing 5732 sentences with InfCs. The category of the PVs is divided between visual (*ver/mirar, voir/regarder, ver/olhar*) and auditory PVs (*oír/escuchar, entendre/écouter, ouvir/escutar*); the class of CVs contains *make-verbs (hacer, faire, fazer)* which can be referred to as verbs indicating ‘positive causation’ as well as *let-verbs (dejar/laisser/deixar)* or verbs of ‘negative causation’ (cf. Soares da Silva 1997). The sentences represent different media of language use (fiction, newspaper articles, etc.) and are taken from electronic databases, namely the *Corpus de Referencia del Español Actual* (CREA, over 150 million words) and the *Corpus del Español* (CDE, 100 million words) for Spanish, *Frantext* (210 million words) and *Le Monde* (1994, 1997-1998) for French and the *Corpus do Português* (CDP, 45 million words) for Portuguese. As table 1 shows the study on PVs has been completed, whereas the conclusions formulated for the CVs are based on a pilot study. Consider the distribution throughout the three languages and the different verb classes:

Table 1. corpus distribution

	PV + InfC			CV + InfC		
	‘see’	‘hear’	total	‘make’	‘let’	total
Spanish	1181	693	1874	100	285	385
French	1700	419	2119	100	290	390
Portuguese	388	376	764	100	100	200

These cases were manually annotated with the following variables: [main verb], [position NP₂], [animatedness NP₁], [animatedness NP₂] and [transitivity Inf]. As to the animatedness of NP₁ and NP₂, we distinguished between human, animate, inanimate self-controlled bodies (the wind, a car,...), inanimate non dynamic and abstract entities. The first three categories were considered to be instances of dynamic participants, whereas the remaining two classes represent non dynamic entities. Finally, the infinitive can also imply different degrees of dynamicity: a semantically transitive verb (such as *eat*) represents a transfer of energy and is thus highly dynamic; an unergative intransitive verb (such as *dance*) represents an emission of energy and is also considered to be dynamic, unlike unaccusative verbs (such as *fall*) which represent a reception of energy by their subject and are thus less dynamic.

3. Discussion

A first quantitative comparison between the three languages shows that for both verb classes – perceptive and causative – the highest number of preverbal NP₂s can be found in Portuguese (PV: 83,7%, CV: 24,5%) and French (PV: 73% - CV: 16%), whereas NP₂ is most frequently postverbal in Spanish (PV: 74,4% - CV: 97%). The following figures illustrate these tendencies:

Figure 1. position NP₂ – perception verbs

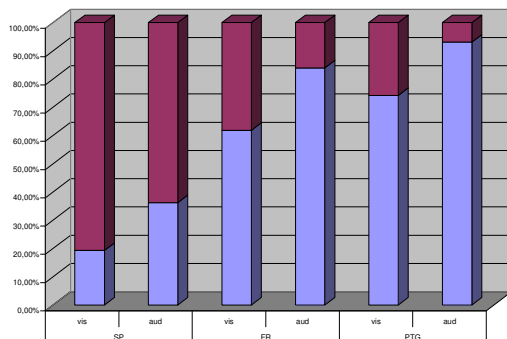
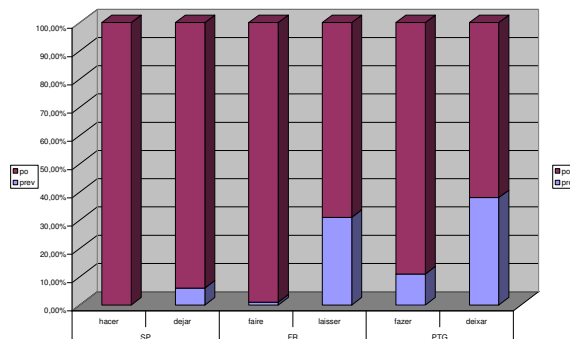


Figure 2. position NP₂ – causative verbs



However, besides these differences the statistical data point towards some striking analogies between the three languages and between the two verb classes. As the above figures show, in Spanish, French and Portuguese:

- auditory PVs more frequently select preverbal NP₂s than visual PVs;
- negative CVs more frequently select preverbal NP₂s than positive CVs.

The main goal of this investigation is to explain these correspondences. It will be argued that what seems to (partly) determine word order in the InfC is the semantics of the main constituents of this complements, which for its part, depends on the extralinguistic cognitive properties of the perception or causation modality. To be more precise, a thorough corpus analysis will allow us to establish following correlations:

- (a) a dynamic NP₂ is more frequently preverbal than a non dynamic NP₂ and NP₂ is more frequently placed before a highly dynamic infinitive, whereas NP₂ mostly appears behind less dynamic unaccusative infinitives;
- (b) auditory PVs and negative CVs more frequently select highly dynamic InfCs (with anteposed NP₂s) than visual PVs and positive CVs which more frequently select less dynamic InfCs (with postposed NP₂s).

To conclude it will be demonstrated that these correlations between the main verb type and the dynamicity of InfC depend on the extralinguistic conceptual properties of the perception modalities and the modalities of causative acting.

Firstly, the stimulus of auditory perception needs to produce some noise in order to be heard, whereas the stimulus of visual perception can but does not have to be implicated in an activity. This extralinguistic difference between the two perception modalities explains why in the three languages the auditory PVs mostly opt for dynamic (human, animate non-human and self-controlled) NP₂s and dynamic (transitive and unergative) infinitives, whereas the visual PVs more easily allow less dynamic (inanimate non dynamic) NP₂s and less dynamic (unaccusative) infinitives. Secondly, in positive causation events ('to make') the subordinated caused event is mostly dependent

on the main event and non-dynamic, since its occurrence depends on the causative act of the main participant. On the contrary, the subordinated event of negative causation processes is more autonomous and dynamic, since the causer NP₁ opposes to a process that tends to occur any way.

These different dynamicity configurations will allow us to explain the similarities between the three languages on the one hand and the differences within the verb classes on the other one. To put it another way, word order tendencies in the InfC in three Romance languages will be shown to reveal more cross-linguistically valid cognitive schemes.

4. References

Corpus

- [CDE] Davies, Mark, Brigham Young University: *Corpus del Español*, <http://www.corpusdelespanol.org/>
- [CDP] Davies, Mark, Brigham Young University / Michael J. Ferreira, Georgetown University: *Corpus do Português*, <http://www.corpusdoportugues.org/>
- [CREA] Real Academia Española: *Corpus de Referencia del Español Actual*, <http://www.rae.es/>
- [FRANTEXT]
Base Textuelle, <http://www.frantext.fr>
- [LM] *Le Monde*, Cd-rom, 1994, 1997-1998.

Cited references

- Danell, Karl Johan 1979. *Remarques sur la construction dite causative: faire (laisser, voir, entendre, sentir) + infinitif*. Stockholm: Almqvist and Wiksell international.
- Enghels, Renata 2007. *Les modalités de perception visuelle et auditive: différences conceptuelles et répercussions sémantico-syntaxiques en espagnol et en français*. Beihefte zur Zeitschrift für romanische Philologie. Tübingen: Niemeyer.
- Rodríguez Espiñeira, María José 2000. Percepción directa e indirecta en español. Diferencias semánticas y formales, *Verba* 27: 33–85.
- Silva, A. Soares da. 1997. *A semântica de deixar: uma contribuição para a abordagem cognitiva em semântica lexical*. Lisboa: FCG/MCT.
- Treviño, Esthela 1994. *Las causativas del español con complemento infinitivo*. México: El Colegio de México.