One Rule To Rule Them All: Logical Structures
For Spanish Non Reflexive se Sentences

Carlos González Vergara
Pontificia Universidad Católica de Chile

In Spanish, the particle *se* appears as a component in a large amount of different syntactic constructions that express different and varied meanings. Contreras (2004), for instance, proposes that there are up to thirteen kinds of *se*. She postulates that this particle can take different values as varied as the ones that we can see in (1). The Spanish traditional names for these sentences are shown in parentheses.\(^1\)

(1) a. Pedro *se* lava. (reflexive *se*)
   Pedro refl wash.3sg
   'Pedro washes himself.'

   b. Pedro y María *se* observan. (reciprocal *se*)
   Pedro and María refl watch.3pl
   'Pedro and María watch each other.'

   c. Se firmó el acuerdo de paz. (passive *se*)
   refl signed.3sg the treaty of peace
   'The peace treaty was signed.'

---

\(^1\)This paper is based on my doctoral dissertation, that I developed under the direction of Dr. Ricardo Mairal. I want to express my gratitude to him, Robert Van Valin, Valeria Bel-lloro and María Inés Prado for their support and generosity.

\(^2\)The abbreviations used in the paper are as follows: 3 = third person, acc = accusative, agx = agreement index, asp = aspect, crea = Corpus de Referencia del Español Actual ('Modern Spanish Reference Corpus'), hra = highest ranking argument, mod = modality, nom = nominative, np = nominal phrase, nuc = nucleus, pl = plural, pp = prepositional phrase, pred = predicate, psa = privileged syntactic argument, refl = reflexive, s = sentence, sg = singular, und = undergoer, v = verb.
d.  \( \text{Se aplaudió a los artistas.} \)  
    (impersonal se)  
    \( \text{REFL applauded.3sg to the artists} \)  
    'The artists were applauded.'

e.  \( \text{Pedro se durmió.} \)  
    (aspectual se)  
    \( \text{Pedro refl slept.3sg} \)  
    'Pedro got sleep.'

f.  \( \text{Pedro se murió.} \)  
    (diaphasic se)  
    \( \text{Pedro refl died.3sg} \)  
    'Pedro died.'

g.  \( \text{Pedro se fue de su casa.} \)  
    (lexical se)  
    \( \text{Pedro refl went.3sg from his home} \)  
    'Pedro went away from his home.'

h.  \( \text{Pedro se bebió un vaso de vino.} \)  
    (affective se)  
    \( \text{Pedro refl drank.3sg a glass of wine} \)  
    'Pedro drank up a glass of wine.'

i.  \( \text{Pedro se arrepintió.} \)  
    (morphological se)  
    \( \text{Pedro refl repented.3sg} \)  
    'Pedro repented.'

j.  \( \text{Pedro se enfermó.} \)  
    (dialectal se)  
    \( \text{Pedro refl sicked.3sg} \)  
    'Pedro got sick.'

k.  \( \text{Érase una vez un rey.} \)  
    (narrative se)  
    \( \text{was.3sg-refl one time a king} \)  
    'Once upon a time there was a king.'

Some of the categories proposed by Contreras have sociolinguistic or stylistic meaning. For instance, the \( \text{se} \) in \( \text{enfermarse} \) ('to get sick') is named “dialectal \( \text{se} \)” because this word is used in just some areas of America; other Spanish speakers prefer the word \( \text{enfermar} \) (without \( \text{se} \)) that expresses the same meaning. In the word
morirse (‘to die’), the particle se is named “diafasic se” because this word without se (morir) seems more formal. Finally, se in the expression érase una vez is named “narrative se” because this construction evokes the literary gender of fairy tales, as an equivalent of “once upon a time”. We will not be concerned with this sociolinguistic or stylistic uses in this text, but the detailed classification put forward by Contreras is important because it shows us the complexity of this field of study in Spanish.

One of the most widely known classifications of the meanings of se usually distinguishes six kinds of se sentences (adapted from Hernández 1966):

a) Reflexive and reciprocal se sentences. As we can see in (2), this kind of sentence can have both a reflexive reading (‘my friends hit themselves’) or a reciprocal reading (‘my friends hit each other’).

(2) Mis amigos se golpearon.

my friends refl hit.3pl

‘My friends hit themselves’ / ‘My friends hit each other.’

b) Passive reflexive sentences. As we can see in (3), a sentence of this kind has a passive meaning, not a reflexive one. The traditional name “passive reflexive” is only motivated by the presence of se and opposes to the “periphrastic passive” built with the verb ser (‘to be’).

(3) Se construyeron muchas escuelas.

refl built.3pl many schools

‘Many schools were built.’

c) Impersonal reflexive sentences. As can be seen in the translation of the example in (4), this kind of sentence expresses actually the same meaning of the passive reflexive. The traditional label “impersonal reflexive” is motivated by the lack of agreement between the SN Pedro and the personal morpheme in the verb.

(4) Se acusó a Pedro.

refl accused.3sg to Pedro

‘Pedro was accused.’
d) “Interest-se” sentences. The *se* in this kind of sentences is traditionally said to be a dative of interest; it is to say an expletive marker, that can be erased from the sentence without affecting its meaning. It is a common opinion, however, that this marker gives to the sentence a special expressive feeling (Bello 1847; Lenz 1935; Gili y Gaya 1943; Alarcos Llorach 1970). The “interest *se*” sentences are usually constructed with verbs of creation, consumption or movement, as can be seen in (5).

(5)  a. *Pedro se bebió una cerveza.*
    *Pedro refl drank-3sg a beer*

    ‘Pedro drank up a beer.’

    b. *Pedro se fue.*
    *Pedro refl go.3sg.past*

    ‘Pedro went away.’

e) “Intrinsic-*se*” sentences. This is a fairly broad and vague category. In traditional Spanish studies this kind of sentences is also known as “middle sentences”, because their meaning is neither active nor passive, but instead *se* expresses a shade of participation in the action (Gili y Gaya, 1943; Roca-Pons 1960; Seco 1972). Here we can find examples of “psychological” intrinsic-*se* sentence (6a), and “physical” intrinsic-*se* sentences with animate (6b) and inanimate arguments (6c).

(6)  a. *Pedro se enojó.*
    *Pedro refl angered.3sg*

    ‘Pedro got angry.’

    b. *Pedro se levantó.*
    *Pedro refl rose.3sg*

    ‘Pedro got up.’

    c. *La puerta se cerró.*
    *the door refl closed-3sg*

    ‘The door closed.’
f) Middle sentences. These sentences, as we can see in (7), describe a property of the argument, and are usually constructed with a manner adverb as fácilmente (’easily’).

(7) Esta puerta se cierra fácilmente.
this door refl close-3sg easily
’This door closes easily.’

Approaches like these have undoubtedly a high degree of descriptive adequacy, but they are not able to explain why so many grammatical constructions with so varied meanings can be built with the same se particle. Moreover, they do not explain how this grammatical element combines with the rest of the sentence so as to generate these meanings.

Likewise, it is a very common intuition among Spanish scholars that, in spite of its apparent diversity, at heart there is only one kind of se in Spanish. This is the position that we can find, for instance, in Martín Zorraquino (1979), Otero (1999) and Mendikoetxea (1999).

In this paper, I will focus on the passive-reflexive, impersonal-reflexives, middle and intrinsic-se sentences. I will try to show how the se that appears in these constructions can be explained based on the Role and Reference Grammar theoretical framework as the morphological manifestation of one simple lexical rule.

My proposal is based on the explanation that Centineo (1995), Van Valin and LaPolla (1997) and Bentley (2004) postulate for similar structures in Italian.

I will leave for a future work my account of how this rule or a closely related one can explain the presence of se in the “interest-se sentences” and reflexive and reciprocal Spanish sentences.

Crucially, I propose that the se that appears in the sentences named “intrinsic-se sentences”, “passive-reflexive se sentences”, “impersonal-reflexive sentences” and “middle sentences” corresponds to the morphological manifestation of one lexical phenomenon.

This lexical phenomenon modifies the logical structure of the sentence, diminishing the actor’s importance and privileging the undergoer when it is present. In other words, this phenomenon intervenes in the logical structure with the purpose of not expressing syntactically the natural argument hierarchy.

The proposed phenomenon can be articulated as the lexical rule in (8). Here we also can see how this lexical rule is expressed for different kinds of predicates according to their Aktionsarten.
(8) Given any kind of logical structure, unspecify the argument $x$ of the predicate.

a. States: $\text{pred}'(x, y) \leftrightarrow \text{pred}'(\emptyset, y)$

b. Activities: $\text{do}'(x, [\text{pred}'(x, y)]) \leftrightarrow \text{do}'(\emptyset, [\text{pred}'(\emptyset, y)])$

c. Active accomplishments:
   c’. $\text{do}'(x, [\text{pred}_1'(x, y)]) \& \text{INGR}\text{pred}_2'(y) \leftrightarrow \text{do}'(\emptyset, [\text{pred}_1'(\emptyset, y)])$
   & $\text{INGR}\text{pred}_2'(y)$

   c’’. $\text{do}'(x, [\text{pred}'(x)]) \& \text{INGR}\text{be-LOC}'(y, x) \leftrightarrow \text{do}'(\emptyset, [\text{pred}'(\emptyset)])$
   & $\text{INGR}\text{be-LOC}'(y, \emptyset)$

d. Accomplishments and achievements: $\text{BECOME/INGR}\text{pred}'(x, y) \leftrightarrow \text{BECOME/INGR}\text{pred}'(\emptyset, y)$

e. Semelfactives:
   e’. $\text{SEML}\text{pred}'(x, y) \leftrightarrow \text{SEML}\text{pred}'(\emptyset, y)$
   e’’. $\text{SEML}\text{do}'(x, [\text{pred}'(x, y)]) \leftrightarrow \text{SEML}\text{do}'(\emptyset, [\text{pred}'(\emptyset, y)])$

f. Causatives:
   f’. $[\text{do}'(x, \emptyset)] \text{CAUSE} [\text{(BECOME/INGR)}\text{pred}'(y)] \leftrightarrow [\text{do}'(\emptyset, \emptyset)] \text{CAUSE} [\text{(BECOME/INGR)}\text{pred}'(y)]$
   f’’. $[\text{do}'(x, \emptyset)] \text{CAUSE} [\text{do}'(y, [\text{pred}'(y)])] \leftrightarrow [\text{do}'(\emptyset, \emptyset)] \text{CAUSE} [\text{do}'(y, [\text{pred}'(y)])]$

The application of the rule in (8) can have different syntactic consequences, and these consequences correspond to the varied traditional $se$ sentences. All of them, nevertheless, have in common that the semantic argument that originally would have been expressed as privileged syntactic argument (PSA) of the sentence is not available to be selected after the application of the rule.

For instance, we can have a transitive “non-$se$” sentence with a causative logical structure, as in (9).

(9) Pedro ensució la camisa.
Pedro stained.3sg the shirt
‘Pedro stained the shirt.’

$[\text{do}'(\text{Pedro, } \emptyset)] \text{CAUSE} [\text{BECOME dirty}'(\text{camisa})]$

If we apply the rule on (8) to this logical structure, we get one where the highest ranking argument has been unspecified, as can be seen in (10). This exclusive logical structure can materialize as four kinds of sentences: an intrinsic-$se$ sentence (10a), a passive-reflexive sentence (10b), an impersonal-reflexive sentence (in this
case, in order to obtain this meaning, we have to change the inanimate argument \textit{camisa} for a human argument like ‘Juan’ (10c), and —finally— a middle sentence (10d).

(10)[\text{do’}(Ø,Ø)] \text{CAUSE} \text{[BECOME dirty’(camisa)]}

a. \textit{La camisa} se ensució.
   \text{the shirt} \text{refl} \text{stained.3sg}
   ‘The shirt got dirty.’

b. \textit{Se ensució la camisa}.
   \text{refl} \text{stained.3sg} \text{the shirt}
   ‘The shirt was stained.’

c. \textit{Se ensució a Juan}.
   \text{refl} \text{stained.3sg} \text{to Juan}
   ‘Juan was stained.’

d. \textit{La camisa se ensucia fácilmente}.
   \text{the shirt} \text{refl} \text{stained.3sg} \text{easily}
   ‘The shirt gets dirty easily.’

In Figure 1, we can see how the linking process works for the original transitive sentence in (9). In the logical structure, we can observe the basic causative structure. The highest ranking argument (\textit{Pedro}, the first argument of the activity predicate) takes the actor macrorole, and the argument \textit{camisa} takes the undergoer macrorole. Then, the actor is selected as PSA of the sentence and takes nominative case. The undergoer takes accusative case. We proceed with the selection of the syntactic template, that critically shows two positions in the core and has a nucleus with an AGX node.\textsuperscript{3} The PSA materializes as the first core NP, and as the personal suffix in the AGX node. The undergoer is expressed as the second core argument. The verb in active voice, finally, links to the predicate node.

\textsuperscript{3}The AGX is a notion proposed by Belloro (2004: 43): it is a node dependent of the nucleus, and it receives the agreement specifications of all core argument positions present in the logical structure.
What happens to the linking process if we apply the lexical rule in (8) to this logical structure? The result can be seen in Figure 2. Here, we notice that the highest ranking argument has been unspecified. I mark this with the label HRA=Ø. It is important to see that this label and the dotted line are not really part of the link-
ing. They only represent that the logical structure has suffered a lexical phenomenon. Therefore, we don’t violate the completeness constraint (Van Valin, 2005: 129). The argument *camisa* takes the undergoer macrorole and, because of the unspecification of the argument *x*, it is selected as the PSA of the sentence. This argument takes nominative case and is linked as a core NP and as the personal suffix of the verb. The morpheme *se* appears in the AGX node as the morphological manifestation of the lexical rule. The sentence we obtain (*la camisa se ensució*) has a predicate focus structure and its meaning is ‘the shirt got dirty (by accident or naturally)’. From a traditional point of view, this would be an intrinsic-*se* sentence.

Figure 2. Linking diagram for *la camisa se ensució*
In the following text, taken from CREA ( Corpus de Referencia del Español Actual 'Modern Spanish Reference Corpus') we can see an equivalent intrinsic-se sentence. It describes a scene in the novel Los aires difíciles.\(^4\) In this context, the clause el aire se ensuciaba (‘the air got dirty’) is perceived as a part of a natural process.

\[
(11) \text{Aún no había podido dormirse, pero no dijo nada. [...] Nadie más parecía darse cuenta, pero ella lo veía, lo sentía con tal nitidez que cerraba los ojos cuando los muros empezaban a combarse, a inclinarse entre sí, y el aire se ensuciaba, se enturbiaba en el presentimiento de la polvareda que armarían los cascotes al caer como una lluvia gruesa y mortal sobre sus cabezas. 'She wasn’t asleep yet, but she didn’t say anything. [...] Nobody else seemed to notice, but she could see it, she felt it so sharply that she closed her eyes when she felt the walls wobbling inwards and leaning one against the other, and the air got dirty, it got dirty because she could feel the dust that would rise when the rubble fell like thick and deadly rain over their heads.'}
\]

If we maintain the same logical structure of Figure 2, with the highest ranking argument unspecified, we can also obtain a sentence with a different meaning: se ensució la camisa (‘the shirt was stained (by someone)’), as can be seen in Figure 3. The only difference between this diagram and Figure 2 is the information structure. Se ensució la camisa has a sentence focus structure, contrasting with the intrinsic-se in (10a), where the NP la camisa acts as a topic. This favors a reading similar to ‘someone stained this shirt’, that has been traditionally related to the passive-reflexive sentence (see figure 3, next page).

In (12) we have an example taken from CREA where we can see this meaning. This is a discussion about television programs and their stars. In this context, the phrase se ensucie la personalidad del rival (‘the rival’s personality be stained’) has a meaning where an unspecified person is the instigator of the state of affairs.

\[
(12) \text{¿Por qué preferimos la guerra de imágenes y rechazamos el debate de ideas? Parece que esto se debe a que nuestros publicistas y especialistas en “marketing”, consideran que la gente no vota por programas, sino por personalidades. Por ello, entre más se hunda, se degrade, se ensucie la personalidad del rival más posibilidades tiene el otro de ser aceptado.}
\]

\(^4\) Almudena Grandes, 2002: Los aires difíciles, Barcelona: Tusquets.
'Why do we rather have a war of images and reject debate of ideas? It seems that this is because our publicists and marketing specialists consider that people is not kin on programs but on personalities. This is why the more the rival’s personality gets sunk, degraded and stained, more chances has the other one of being accepted.'

Figure 3. Linking diagram for ‘Se ensució la camisa’
How do we obtain the impersonal-reflexive sentence? I propose that the logical structure for this kind of sentence is, again, the same we have shown for intrinsic-se and passive-reflexive sentences. The only difference in this case is in the semantic properties of the undergoer. In Figure 4, we can see that *Juan* is a human, fully specified argument. For a logical structure like this, the speaker can choose if the argument *Juan* is selected as PSA or not. If it is chosen, the form and meaning of the resulting sentence are reflexive: *se ensució Juan* (‘Juan stained himself’). If the argument *Juan* is not selected as PSA, then it takes accusative case and materializes as a prepositional phrase in the core, the morpheme *se* appears in the AGX node, and the resulting sentence is impersonal-reflexive, whose meaning is equivalent to the passive-reflexive: *se ensució a Juan* (‘Juan was stained (by someone)’). We can observe this linking in Figure 4.

In (13), an example taken from a municipal act,⁵ we can see how this kind of sentence is used. In this case, the clause *se ensucia a los ediles* (‘the mayors get stained’) implies that the state of affairs was caused by someone unspecified although suggested: *la prensa* (‘the press’). It is important to notice that the meaning of this clause is not really different of the meaning of a passive-reflexive. In both cases, we have an unspecified person that instigates the state of affairs.

(13) *En ese marco, quiero que enérgicamente nosotros respondamos, pero no con una mera respuesta de prensa —que creo debemos hacer—, sino de tal forma que estas cosas no vuelvan a pasar, porque cuando se ensucia a los ediles en la prensa lo que queda no es lo que uno repara después, sino la denuncia que fue hecha.*
‘Within this context, I want us to answer energetically, but not with only an answer through the press —that I think we must do—, but also in a way that this things won’t happen again, because when the mayors get stained in the press what remains is not what you can mend afterwards, but what was stated in the first place.’

Let us examine finally how middle sentences are derived. We can start with the same logical structure and specify the presence of two operators: modality (possible) and aspect (imperfect), as well as a new predicate: *easy*. In this case, the highest ranking argument is unspecified and this triggers the presence of *se*. The argument *camisa*, on the other hand, takes the undergoer macrorole and is se-

---

⁵http://www.juntamvd.gub.uy/actas/indexados/sesion_060310.htm
Figure 4. Linking diagram for \textit{Se ensució a Juan}

\textbf{HRA} = \emptyset \\
\textbf{ACTIVE} \\
\textbf{UND} \\
\[\text{do}'(\emptyset, \emptyset)\] \textbf{CAUSE} \[\text{BECOME dirty}'(\text{Juan})\] \\
\ [+human]
lected as PSA. Therefore, it takes nominative case and is linked to the core (and its pronominal features to the AGX). The verb is in active form, and the predicate easy’ is realized as a peripheral adverb. The resulting sentence la camisa se ensucia fácilmente (‘the shirt gets dirty easily’) is a typical middle sentence in Spanish. The linking process for this sentence can be seen in Figure 5.

Figure 5. Linking diagram for La camisa se ensucia fácilmente

It is important to notice that the middle sentence in Figure 5 has a predicate focus structure. Combined with the other properties of the construction, this motivates that the sentence is interpreted as an opinion about the topic properties, as we can also see in the example in (14), a text taken from CREA that describes the characteristics of some kind of paints. Here we can observe how the sentence esas pinturas se ensucian rápidamente (‘these paints can get dirty quickly’) is the description of a property of the paints.
(14) Estas pinturas están compuestas por una resina sintética (vinílica o acrílica) que se halla emulsionada en agua. Presentan una alta impermeabilidad, son muy poco transpirables y, aunque se pueden lavar, se ensucian rápidamente.

‘These paints are composed by a synthetic (vinyl or acrylic) resin in the form of an emulsion based in water. They present high impermeability, perspire very little and even though they can be washed, they get dirty quickly.’

The analysis for middle sentences in Spanish put forward in Figure 5 is based on a proposal by Felíu Arquiola (in press). I don’t follow here the analysis of Van Valin and LaPolla (1997: 417), who say that the logical structures of the middle sentences correspond to an attributive state, where the adverb acts as attribute. This is because in Spanish the presence of the adverb, although frequent, is not mandatory in middle sentences. We can see this in examples like la tuberculosis se cura (‘tuberculosis can be cured’) in (15), taken from CREA, whose meaning is also middle.

(15) Según el estudio, la tuberculosis es una enfermedad social y su presencia es un factor negativo en el desarrollo económico de los países del tercer mundo, sin embargo, desde el campo de la salud, es posible revertir la situación actual porque la tuberculosis se cura y el tratamiento actúa como medida de control de la transmisión.

‘According to the research, tuberculosis is a social disease and its presence is a negative factor in the economical development of Third World countries. However, from a medical perspective, is possible to revert the current situation because tuberculosis can be cured and treatment acts as a propagation control measure.’

In conclusion, the passive-reflexive, impersonal-reflexive, middle and intrinsic-se sentences differ from each other in aspects like their information structure, the lexical features of their arguments and the relevance of lexical operators. These properties can be described in terms of different constructional schemata (Van Valin 2005), as can be seen in González Vergara (2006). All these sentences, nevertheless, are originated from no more than one kind of logical structure, that have undergone the application of the lexical rule proposed in (8), that unspecifies the highest ranking argument.
In Table 1 I show an example of these schemata: the proposal for the properties of the Spanish middle construction. In Figure 6, we can see how these properties affect the linking process: (1) the presence of the AGX node in the nucleus, (2) the argument modulation, (3) the PSA modulation, (4) the omission of the actor, (5) the active voice of the verb, (6) the presence of the imperfect aspect operator, (7) the presence of the morpheme se in the AGX node, (8) the presence of the modality operator POSSIBLE, and (9) the predicate focus structure.

<table>
<thead>
<tr>
<th>Construction: Spanish middle sentence</th>
</tr>
</thead>
</table>
| **Syntax**
Template: AGX in nucleus
Argument modulation: central positions are reduced in 1
PSA modulation: variable
- Undergoer argument [-human]: the undergoer argument is selected as PSA
- Undergoer argument [+human]: no argument is selected as PSA

**Morphology**
Verb: active voice, imperfect aspect
Morpheme se in AGX node

**Semantics**
Property interpretation, related to the modality operator POSSIBLE
PSA is not the instigator of the state of affairs, but it is affected by it

**Pragmatics**
Focal structure: predicate focus (default)
Illocutionary force: unspecified

Table 1. Constructional schema for Spanish middle sentences

It is evident that in this paper I have not explained, or merely started to analyze, many complex aspects of the Spanish se constructions and their idiosyncratic meanings. However, I hope I have shown how the Role and Reference Grammar theoretical framework allows us to explain that the occurrence of the marker se in all the kind of sentences presented can be based on a single lexical rule.
Figure 6. Properties of the Spanish middle construction and their influence in the linking

< MOD POSSIBLE < ASP IMPERFECT (easy’ ([do’ (Ø, Ø)]) CAUSE [BECOME dirty’ (camisa)])>>
References

CREA (Corpus de referencia del español actual). http://corpus.rae.es/creanet.html
Roca-Pons, J. 1960. *Introducción a la gramática (con especial referencia a la lengua española).* Barcelona: Teide.


