

# From visual perception to evidentiality: A functional empirical approach to *se ve que* in Spanish

Marta Albelda Marco <sup>a,\*</sup>, Marlies Jansegers <sup>b</sup>

<sup>a</sup> *University of Valencia, Spain*

<sup>b</sup> *Ghent University, Belgium*

Received 15 July 2018; received in revised form 24 January 2019; accepted 24 January 2019

Available online 30 January 2019



## Abstract

The Spanish sequence *se ve (que)* presents intricate functional polysemy, including constructionalization as an evidential. The present paper investigates its different formal-functional combinations and degrees of specialization as an evidential construction. The following questions were addressed: (1) How many different senses can be distinguished in the sequence *se ve (que)* and what are their respective frequencies? (2) How do these senses correlate with the morphosyntactic behavior of the sequence *se ve (que)*? (3) Which senses of *se ve (que)* are more closely related to each other, and how does the evidential construction relate to this polysemous network? The semantic and formal affinities of *se ve (que)* were studied through the Behavioral Profiles method developed by Gries and Divjak (2009), which provides an empirical, systematic and verifiable approach to studying lexical phenomena. Its application to a pragmatic phenomenon is a new departure. The results show seven senses of *se ve (que)*, ranging from the lexical value of direct physical perception to the more abstract and evidential value of 'source of information'. According to the corpus analysis, the closest senses to the evidential pole are indirect physical perception and cognitive perception. These all introduce an inflected verb clause, possess propositional scope and are morphosyntactically frozen..

© 2019 Elsevier B.V. All rights reserved.

**Keywords:** *Se ve (que)*; visual perception; evidentiality; Behavioral Profile; empirical approach

## 1. Introduction

The Spanish structure *se ve que* (literally 'it is seen (that)') presents wide-ranging polysemy nowadays. One of the meanings that stands out is its evidential value. The evidential possibilities of verbs of sight in various typologically non-evidential languages are well known. Physical sight serves as a source of evidence for both visual evidentiality (direct, what is seen by the physical eyes) and inferential circumstantial evidentiality (indirect). Taking a broad approach to evidentiality (Anderson, 1986), under certain semantic and syntactic circumstances—particularly the requirement that it introduce propositional scope (Boye, 2010)—the verb *ver* [to see] can be considered to express evidentiality, both with inflectional behavior (Example 1) and when it is more fixed syntactically (Example 2):

\* Corresponding author at: Departamento de Filología Espan ola, Facultat de Filologia, Traduccio i Comunicacio, Avenida Blasco Ib n-ez, 32, 46010 Valencia, Spain.

E-mail address: [marta.albelda@uv.es](mailto:marta.albelda@uv.es) (M. Albelda Marco).

(1)  
 U: el año pasado// íbamos a ir de cena (...) y estaba cerrado el Rausell↑ / y estaba cerrado el Leo↑ / y estaba cerrado todos// fuimos a un sitio de PUTA MIERDA/ no comimos casi (...)  
 E: tienen que poner más cantidad↑ / que os pongan patatas bravas por ejemplo/ que están buenisimas/// ¿no? (...)  
 U: es una cena que nació en el restaurante Osaka de Primado Reig↑  
 E: ha ido degenerando/ por lo que **veo**/ ¿no?  
 (Val.Es.Co. 2.0, Conversation 37, lines 150–182)

U: last year// we were going out to supper (...) and the Rausell was closed↑ / and the Leo was closed↑ / and they were all closed// we went to a SERIOUSLY SHIT place/ we ate almost nothing (...)  
 E: they should give you more quantity↑ / they could give you patatas bravas for instance/ they taste great/// right? (...)  
 U: it's a supper they invented at the Osaka restaurant on Primado Reig↑  
 E: it's been going downhill// as far as I can **see**/ right?

In (1), in speaker E's last turn, *veo* [I see] is placed within a larger segment (*por lo que veo*, [as far as I can **see**]) and has propositional scope over *ha ido degenerando* [it's been going downhill]. This use of *veo* is compositional and presents inflectional mobility, as the speaker could have said *por lo que vemos* [as far as we can see] or *por lo que vi* [as far as I could see], for instance. It expresses inferential evidentiality<sup>1</sup>: from the information provided by U, E infers that the quality of the restaurant has gone down.

In the following example (2), the adverbial phrase *por lo visto* ([apparently], which derives from *ver* [see] and is syntactically fixed) again expresses evidentiality, this time reportative:

(2)  
 B: bueno/ el servicio militar fue un hecho un tanto traumático para mí// yo esperaba quedarme en Valencia/ (...) / mi familia tenía/ un militar amigo/ y él/ **por lo visto**/ les había dicho que si me tocaba por aquí/ que estaría con él/ y podía estar relativamente bien  
 (PRESEEA Valencia, interview ESA23, lines 78–82)

B: well/ my military service was quite a traumatic episode for me// I was expecting to stay in Valencia/ (...) my family had/ a friend in the army/ and he/ **apparently**/ had told them that if I drew somewhere here/ I would be with him/ and could be relatively okay

The present paper focuses on another linguistic construction, *se ve que*, which also derives from the verb *ver* and presents a series of peculiarities. This sequence of the form *se* + the verb *ver* in the third person of the present tense + (optional) the particle *que*, displays different syntactical combinations that appear to correspond to semantic differences. According to Albelda (2016, 2018), when constructionalized it expresses evidential value and means 'apparently.' This occurs with both its formal variants: (i) the construction *se ve que* in utterance-initial position, introducing a phonically and syntactically integrated proposition (e.g. Example 3), and (ii) the parenthetical construction *se ve* in a syntactically and prosodically autonomous position (in medial or final position, as in Example 4). The latter can be considered an extension of the former, arising from its syntactic dislocation in specific contexts. In Example (3), *se ve que* does not relate to physical vision but expresses evidentiality as its core meaning (i.e. it expresses the way in which the information was acquired):

(3)  
 C: Laura y el pequeño↑ Santi↑ tuvieron un accidente de moto (...) el chiquito↑ / entro en la Uvi↑ lo tuvieron en la Uvi en principio porque se le había caído el casco/ **se ve que** no llevaba bien atado el casco↓ le salió despedido↑ y entonces tenían los dos una pierna rota  
 (Val.Es.Co. 2002, MT.97.A1: lines 40–46)

C: Laura and little↑ Santi↑ had a motorbike accident (...) the little guy↑ went into the ICU↑ they had him there at first because his helmet had fallen off/ **apparently** he didn't have the helmet well fastened↓ it flew off↑ and then both of them had a broken leg

Speaker C narrates an accident suffered by a little boy and introduces the reason why the boy's helmet came off. She was not present at the accident scene and now she not only narrates the episode but also points to the type of source of

<sup>1</sup> See Section 4.3 for an explanation of the types of evidentiality.

her information: she knows that the boy did not fasten his helmet correctly because she had previously heard it from someone else, or because she could surmise what happened, given the circumstances of the accident.

In (4), parenthetical *se ve* is used with evidential value:

- (4)  
 P: el niño se abrazó a su madre↑/ acercó a la cara así↑/ así// y no la desapegó  
 C: (LAUGHING) ¡ay qué bonito! (...)  
 P: y hacía así con los ojitos↑ (gesture)// pero claro/ él no veía→ **se ve**↓ bien a su madre/ y se apegó a su madre/ y luego el ayudante del cirujano nos decía/ allí dentro os hacía así aaaa (gesture) que lo sacaran fuera (Val.Es.Co. 2002, G.68.B1: lines 146–155)  
 P: the kid hugged his mother↑/ moved his face closer like this↑/ like this// and he didn't move it from there  
 C: (LAUGHING) oh, how cute!  
 P: and he went like this with his little eyes↑ (gesture)// but of course/ he didn't see→ **apparently**↓ his mother well/ and he held fast to his mother/ and then the surgeon's assistant told us/ there inside he went like this aaaa (gesture) to get the boy away

In (4), speaker P narrates the behavior of a child during a surgical procedure. In this spontaneous conversation, P disrupts the sequence 'verb + complements' and inserts the parenthetical construction *se ve*. The independence of the construction is demonstrated by the fact that it is surrounded by pauses and displays falling intonation. *Se ve* is in medial position within the utterance, and has propositional scope, i.e. scope over the entire proposition *él no veía bien a su madre*.

As well as these two more constructionalized combinations, the intricate polysemy of *se ve que* presents other formal and semantic alternatives. Among its more compositional possibilities, it can be expected to express the same values as any inflected form of the verb *ver*: According to the scholarly literature (Willems, 1983; Hanegreefs, 2008; Fernández-Jaén, 2012, and so on), *ver* expresses physical, visual perception values, both direct (*desde la cocina se ve la calle* [from the kitchen you see the street]) and indirect (*Veo que llevas unas botas nuevas* [I see that you're wearing new boots]), and also cognitive perception values (*Vi que su vida se iba a pique* [I saw that his life was going to ruin]); see below (Section 2.1). Consequently, the question arises as to how these compositional uses in the sequence *se + ve + que* with lexical meaning relate to its more abstract, grammatical meaning, expressing evidentiality.

Hence, the main objective of the present paper is to study the degree of specialization of the sequence *se ve que* as an evidential construction. Related to this objective, it also aims to explore which non-evidential meanings of *se ve que* are closest to the evidential meanings. These objectives lead to the following three research questions:

1. How many different senses can be distinguished in the sequence *se ve que* and what are their respective frequencies?
2. How are these senses correlated with the morphosyntactic behavior of the sequence *se ve que*?
3. Based on questions (1) and (2), which senses of *se ve que* are more closely related to each other, and how does the evidential construction relate to this polysemous network?

To answer these research questions, occurrences of *se ve* and *se ve que* in spontaneous conversations and semi-directed interviews in European Spanish were analyzed. The corpora total around 1 700 000 words, and 230 occurrences were found. The Behavioral Profile (BP) method developed by Gries and Divjak (2009) was used to analyze these units and explore their formal and functional profiles, providing a quantitative, systematic and verifiable approach to the study of their polysemy. Through the analysis of a series of features involved in their use (argument structure, position of the perceived stimulus within the utterance, degrees of formal inflection or fixing of the verb, the semantics of complements), the BP approach also revealed the distribution of their senses and forms.

The structure of this paper is as follows: Section 2 provides the theoretical background on visual perception verbs and their evidential possibilities and on the expression of evidentiality in the Spanish construction *se ve que*. According to the WALs (De Haan, 2013a, 2013b), Spanish is a non-evidential language. It only has a few discourse markers and constructionalized forms with core evidential meaning (*dizque*, *por lo visto*, *al parecer*, Estellés and Albelda, 2014; Albelda, 2015; De la Mora and Maldonado, 2015; Kotwica, 2015, 2017; González-Ramos, 2016; Estellés, 2018). Accordingly, reference will be made to the constructionalization of the sequence *se ve que*. Section 3 presents the method and data used in the Behavioral Profile. Section 4 presents and discusses the BP results, leading to a further refinement of the description of the polysemy patterns of *se ve* and *se ve que*. This is followed by an explanation of the similarities and differences between the different clusters of combinations of *se ve* and *se ve que*. Section 5 summarizes the main contributions of the BP analysis and answers the research questions.

## 2. Theoretical framework

### 2.1. The polysemy of the visual perception verb *ver*

The Spanish verb *ver* (to see) derives directly from the Latin verb *video*, which, in turn, comes from the Indo-European root *weid-*. Interestingly, this root means ‘vision’ (Grossmann and Tutin, 2010:282) but is related firstly to knowledge and only secondly to visual perception (Sweetser, 1990:33; Grossmann and Tutin, 2010:282). Lexically, *ver* is an experiencer-based (subject-oriented) verb (Viberg, 1983; Whitt, 2011), since it requires a perceiver to experience the physical (Example 5) or mental vision (Example 6), and indicates the object or situation perceived:

- (5) **He visto** un pájaro en el árbol  
 I **have seen** a bird in the tree
- (6) **Veo** que te han educado muy bien (after polite behavior by a child)  
 I **see** you have been raised very well

The rich bibliography dedicated to the phenomenon of visual perception has characterized *ver* as a highly polysemic verb. The sense expressed by each use largely depends on the semantics of its complements and on its syntactic combinations. Different authors have attempted to organize its values in different ways, whether through opposites (Willems, 1983) or through continuum scales (Kirsner and Thompson, 1976; Enghels, 2007; Hanegreefs, 2008; Fernández-Jaén, 2012). The classifications differ, depending on each author. The main values recorded by many scholars—particularly Willems (1983), Rodríguez-Espin-eira (2000), Enghels (2007), Hanegreefs (2008) and Fernández-Jaén (2012)—are presented below. In the present study, the analysis (see Section 4) is based on a continuum of meanings, as the borders between one value and another are not clear and many uses combine physical and abstract values.

The most widespread semantic distinction points to three main values: (a) direct physical perception, (b) indirect physical perception and (c) cognitive perception. In direct physical perception (through the eyes), in principle the perceived stimulus is a physical entity, although some authors also include virtual images (*Veo a un niño moreno; Se ve una imagen en el cielo*; [I see a dark child; An image is seen in the sky]). Indirect physical perception involves inferential processes in the speaker’s mind based on visual stimuli. Semantic indirectness is introduced by grammatical indirectness, as in these cases the verb introduces a sentence complement (Fernández-Jaén, 2012:294; *Veo que te has cortado el pelo* [I see that you have cut your hair]). Some authors have considered cognitive perception a type of indirect perception. It includes cases in which the perception is purely mental and the reasoning is not based on a visual stimulus but on the speaker’s prior knowledge or assumptions (*En esa explicación se ve que España es machista; Veo la ternura de su alma* [In this explanation, Spain is seen to be sexist; I see the gentleness of her soul]).

Some authors also distinguish a fourth type: evaluative perception (Rodríguez-Espin-eira, 2000; Hanegreefs, 2008). In this case, a subjective qualifier is attributed to the perceived stimulus (*Veo que tienen actitudes muy cobardes* [I see that they have very cowardly attitudes]). Based on this classification of the values of the verb *ver*, Sections 3.2 and 4.2 propose a more detailed version for the specific case of the sequence *se ve que*, in keeping with the corpus data.

In short, *ver* displays a complex polysemic profile extending from direct, physical perception senses through indirect physical perception toward more abstract uses of cognitive and evaluative meaning. However, the polysemous potential of the verb goes beyond these lexical meanings and extends to an even more abstract, constructionalized sense, such as evidential uses. Indeed, the contemporary Spanish dictionary *DRAE 2014* includes a definition that is closely related to the evidential meaning: “Dicho de una cosa: parecer o evidenciarse. Ex.: *Se ve que tendremos elecciones pronto*” [Referring to something: to seem or to become evident. *It seems we will be having elections soon*]. The example given by the *DRAE 2014* coincides with the construction under study here. These particular evidential uses will be discussed in the next section.

### 2.2. Evidentiality in Spanish: evidential meaning in the construction *se ve que*

Spanish is a typically non-evidential language in that it has no obligatory grammatical expression of this meaning (Aikhenvald, 2004; De Haan, 2013a, 2013b). From a functional point of view, however, evidentiality is considered to be a phenomenon that goes beyond grammar and must be regarded as a universal semantic category (Cornillie, 2007; Boye, 2010; Diewald and Smirnova, 2010; Albelda, 2015, 2016; Alonso-Almeida, 2015; Izquierdo, 2016; Kotwica, 2017; Estellés, 2018). From this point of view, polyfunctional forms and constructions are considered capable of expressing evidentiality in Spanish, as in some particular uses of the imperfect, perfect and future tenses, verbs of perception, modal

verbs, impersonal pronouns, and so on (Fernández, 2013; Henneman, 2013; Escandell Vidal, 2014; García-Negroni and Libenson, 2014; Rodríguez-Rosique, 2015; García-Negroni, 2016).

These polyfunctional forms sometimes undergo grammaticalization and constructionalization (Traugott and Trousdale, 2013), acquiring a core primary evidential meaning. In addition, according to Boye (2010), they introduce propositional scope. This is the case of discourse markers that nowadays have evidentiality as their central value, such as *dizque* [it is said that], *por lo visto* [apparently] or *al parecer* [it seems]. Historically, they all derive from compositional uses with a semantic basis in verbs of saying or perception. Indeed, based on the core meaning of perceptive experience of these markers, several scholars consider that the process of seeing is closely linked to the expression of evidentiality (Vogeleer, 1994; Cornillie, 2007; Grossmann and Tutin, 2010; Whitt, 2011; Fernández-Jaén, 2012; Hassler, 2010; Fernández, 2013; Kotwica, 2017). The physical experience of vision is certainly considered a type of ‘source of information’ within the field of evidentiality (Anderson, 1986; Willett, 1988; De Haan, 2003; Aikhenvald, 2004). Not only vision verbs but also other perception verbs can evolve into evidential markers. Nevertheless, vision is held to be a stronger or more straightforward way to express evidence (Matlock, 1989), given its higher ranking on the sensorial scale (Viberg, 1983).

The present paper argues that this constructionalization process resulting in the evidential value is still ongoing in the case of *se ve que*. A construction is a conventional form-meaning pairing (Fillmore et al., 1998; Goldberg, 2006) where both form and meaning are treated with equal weight and where *meaning* can be understood in the broad sense of *function* to include not only (lexical) meaning but also discourse function, information structure and other pragmatic phenomena. Hence, in accordance with this definition, the evidential meaning of *se ve que* as exemplified above in (3) is not fully predictable from its component parts, that is to say that it is not the sum of its individual parts, since this would literally yield ‘it is seen that.’

Moreover, the fact that evidential *se ve* can be used parenthetically is a clear sign of how far the fixing process has advanced and of its autonomy as a different construction to the compositional uses (see example 4 above). Parenthetical uses coexist with initial uses integrated into the utterance, and are still not very frequent. In the corpus study by Albelda (2016) they comprise 5% of the total evidential uses of this construction. Indeed, the fact that formally it presents two variants (*se ve que* in initial position and *se ve* parenthetically), together with its polysemous nature, which includes compositional values, demonstrates that *se ve que* is not fully fixed. In other words, the constructionalization process is still ongoing (Albelda, 2016, 2018; Jansegers and Albelda, 2018). Consequently, it is not yet possible to speak of exclusive, unequivocal specialization of a form and a significance.

### 3. Data and method

Within the field of linguistics the study of meaning in the broad sense of *function*—including not only (lexical) meaning but also discourse function, information structure, and other pragmatic phenomena—has a long tradition, characterized by a cyclic pattern of interest. Only in the past three decades has the study of semantics and pragmatics regained center stage in linguistics.

Simultaneously, one of the major changes that the discipline of linguistics has experienced in the past few decades is arguably the significant surge of empirical data and methods: both for describing and explaining concrete linguistic phenomena and for the development of more general theories, the use of quantitative corpus-driven methods has proven useful and become more widespread than ever before. This empirical turn—as Geeraerts (2006) termed it—has also begun to carve a path in the realm of semantics and pragmatics (cf. among others Biber et al., 1998; Gries and Stefanowitsch, 2006; González-Márquez et al., 2007; Jucker et al., 2009; Glynn and Fischer, 2010; Hidalgo, 2014; Figueras and Cabedo, 2018). However, as several authors have pointed out (e.g. Geeraerts, 2010:64; Glynn, 2010:240), applying empirical, quantitative methods to the study of meaning is not straightforward. Indeed, how can meaning (an intrinsically qualitative and non-observable relation in the mind) be investigated by quantitative methods? Therefore, in order to address semantics and pragmatics from an empirical, quantitative point of view, the main challenge lies in how to operationalize meaning.

The answer to this apparently somewhat tricky question lies in an important assumption that lies at the heart of nearly all corpus-based studies and is captured in the so-called ‘distributional hypothesis’, namely that distributional similarity reflects functional or semantic similarity. This idea is embodied in Firth’s (1957:11) famous dictum that “you shall know a word by the company it keeps” (cf. also Bolinger, 1968; Hanks, 1996). From this perspective, then, meaning is operationalized by ‘context of use’ and frequency of co-occurrence.

One recent type of inclusive, fine-grained analysis that is based on this distributional hypothesis is the Behavioral Profile approach (BP). Being corpus-based, the BP approach builds on the idea that corpus data provide distributional frequencies and that distributional similarity reflects functional or semantic similarity. This method has proved useful for the analysis of different phenomena in lexical semantics (Divjak and Gries, 2006; Jansegers et al., 2015, among others).



However, no pragmatic application of this method has previously been employed. In the present study, the BP approach is applied to the pragmatic phenomenon of the Spanish evidential marker *se ve que*. Moreover, following Jansegers and Gries (2017), the statistical exploration of the BP analyses uses Multidimensional Scaling (MDS)-based semantic maps to complement the classic hierarchical cluster analysis (HAC).

In a similar way to previous studies using BP, the following four-step procedure was adopted (cf. Gries and Divjak, 2009; Gries, 2010a):

- (i) Retrieval of all instances of the sequence *se ve que* in context, in the form of a concordance (see Section 3.1).
- (ii) Manual analysis and annotation of a large set of properties of each match of the sequence in the concordance. These properties are termed *ID tags* (Atkins 1987) and include morphological, syntactic, semantic, and other characteristics (see Section 3.2).
- (iii) Generation of a co-occurrence table specifying which ID tag level is attested and how often for each sense, i.e. the conversion of this table into vectors (see Section 3.3).
- (iv) Evaluation of the table through exploratory and other statistical techniques (see Section 3.4).

### 3.1. Corpus retrieval

The data used to investigate *se ve* and *se ve que* came from a sizeable corpus retrieved from colloquial conversations and semi-formal interviews in bilingual Catalan-Spanish areas (Valencia, Castellón and Palma de Mallorca) and in Granada (Andalusia). The selection of these geographical regions was based on the findings of Jansegers and Albelda (2018), which show that these are the Spanish-speaking areas with the most abundant use of the evidential construction *se ve que*. A total of 230 instances of *se ve* and *se ve que* were retrieved and their polysemy and formal behavior were studied.

Table 1 summarizes the type of corpus retrieved and the total number of words.

It should be noted that some occurrences were excluded from the frequency counts, and thus from the subsequent analysis: (i) uses of *se ve* as a pronominal verb (*verse*) in the pseudo-copular construction (e.g. *el cultivo de hortalizas se ve afectado por los efectos de la lluvia*, NGLE 2009, § 38.5 [the cultivation of vegetables is affected by the effects of rain]), where the clitic *se* is part of the morphological structure of the verb and occurs in combination with the participle form used as an adjective; and (ii) forms of *se ve* in a verb tense other than the present and in other persons besides the third singular (*se verían*, *se verán*, etc.), since the research objective was to study the degree of specialization of *se ve que* in its constructionalization as an evidential.

### 3.2. Annotation of parameters in the corpus analysis

In order to apply the BP approach to the sequences *se ve* and *se ve que*, the study endeavored to determine all the morphosyntactic and functional features of their uses in the corpus. The features of *se ve que* itself and of the accompanying complements were both studied. Thus, all occurrences were annotated for a wide range of properties (called ID tags). This method yielded a total of 10 different ID tags and around 50 ID tag levels. In line with the basic tenets

Table 1  
Corpus composition.

Genre and register	Corpus <sup>a</sup>	No. of words
Colloquial conversations	Val.Es.Co 2002	100 000
	Val.Es.co 2.0	250 000
	Cogila Granada	36 000
	COJEM Palma Mallorca	120 000
Semi-formal interviews	PRESEEA Valencia	420 000
	PRESEEA Castellón	575 000
	Habla culta de Granada	168 000
Total		1 669 000

<sup>a</sup> The corpora are shown in the reference list as follows: Val.Es.Co. (Briz et al., 2002), Val.Es.Co. 2.0 (Cabedo and Pons, on line), Cogila Granada (Barros et al., 2012), COJEM Palma de Mallorca (Méndez, 2015), PRESEEA-Valencia (Go´mez-Molina, 2001–2007), PRESEEA-Castellón (Blas-Arroyo, 2010), *Habla culta de Granada* (Salvador, 2007).

of a corpus-driven approach, this is not a pre-established list of variables but an open class that grew and expanded throughout the study. A comprehensive overview of the variables used may be found here below.

The fundamental parameter, used as the main criterion in the analysis, is the sense of the sequences *se ve* and *se ve que*. Since the sense is an essential part of the analysis, this aspect merits some comment. It should be noted that for the semantic analysis, we resorted to a very fine-grained annotation of the different possible senses of the verb. This was mainly based on the existing literature (see Section 2.1), which indicates four values of *ver*: direct physical, indirect physical, cognitive and evaluative perception. However, the corpus data analysis required a more detailed classification, as the same sense was found to cover divergent formal behavior patterns. Consequently, seven senses were established for data analysis purposes.

Starting from the hypothesis that these formal features determine “*les caractéristiques sémantico-conceptuelles des VdP*” [VdP: verbs of perception] (Enghels, 2007:6), some of the senses originally proposed in the bibliography have been divided. Accordingly, as well as adding ‘evidentiality’ to the four original values, a new sense termed ‘direct physical valorization’ was split off from ‘evaluative perception’ (Rodríguez-Espin-eira, 2000; Hanegreefs, 2008, see Section 2.1), thereby differentiating two ways of evaluating the perceived stimulus:

- perceived stimuli categorized through physical qualities (‘direct physical valorization’): *Se ve un pueblo bastante grande* [You see quite a large town],
- perceived stimuli categorized through axiological features (‘evaluative perception’): *Se ve un juego muy aburrido* [It looks like a very boring game].

A value half-way between ‘indirect physical perception’ and ‘cognitive perception’ (Fernández-Jaén, 2012, see Section 2.1) was also added, namely ‘deduction as a result of physical perception’. This value applies when the syntax is not indirect (what is introduced is not a proposition but a non-propositional stimulus) but the perceived stimulus (abstract or metaphorical) is deduced from sensory data: *se ve mucha miseria por arriba del puente* [going over the bridge you see a lot of misery]; *se ve la diferencia de una calle a otra* [you see the difference from one street to another].

In addition to the parameter of ‘sense’, used as the main semantic criterion for the cluster analysis, the additional semantic/functional variables taken into account are the following:

- The semantic nature of the perceived stimulus: physical/concrete entity, abstract entity, concrete metaphorical entity and event/situation.
- Presence/absence of qualification of the perceived stimulus: physical property qualifiers, axiological/evaluative qualifiers. In addition, there are two kinds of syntactic relations between the qualification and the perceived stimulus (RAE, NGLÉ 2009 §13.1 m, §37.1p): direct/internal modification of the stimulus (*Se ve una puerta giratoria* [you see a revolving door]) and indirect/external modification, in this case a predicative complement (*El video se ve demasiado fuerte* [the video looks too heavy]).
- The semantics of the adjuncts: quantity, frequency, physical space, metaphorical space, physical manner, metaphorical manner, pronominal manner and time.

Besides these semantic variables, the formal variables taken into account are the following:

- The formal variant of the sequence under study: *se ve* in singular or plural (*se ve/se ven*); *se ve que* and *se ve co’mo*. Table 2 shows the frequencies of each of these formal variants:
- Potential inflection of the verb form *ver*. This variable distinguishes between the occurrences that despite appearing in singular have the potential to be pluralized and those that do not, because they introduce an inflected verb clause with

Table 2  
Distribution of formal variants of *se ve* and *se ve que*.

Formal variants	#
<i>Se ven</i>	9
<i>Se ve</i> (non-propositional)	59
<i>Se ve</i> (propositional)	10
<i>Se ve que</i>	150
<i>Se ve co’mo</i>	2
Total	230

- que*. For instance, *Se ve la cocina* [The kitchen is seen] allows pluralization (*Se ven las cocinas*), but *Se ve que se le ha caído el casco* [It seems that his helmet has fallen off] is inflectionally immobilized and pluralization is not possible (\**Se ven que...*). In the latter cases the only option for agreement is third person singular since the construction is impersonal.
- Syntactic autonomy: phonically and syntactically integrated into the utterance (dependency), or parenthetical, that is to say appearing between pauses and having positional and syntactic mobility (autonomy). Parenthetical uses always present propositional scope, and they only occur with the sequence *se ve*.
  - Position of the perceived stimulus in relation to *se ve que*. A flexible stimulus position in the sentence can be considered to favor constructionalization of this sequence. The corpus exhibits four levels:
    - Perceived stimulus placed before the verb: *el ambiente latino se ve enseguida* [the latino atmosphere is seen straight away].
    - Perceived stimulus placed after the verb: *se ve que me gustan las cosas difíciles* [it seems I like difficult things]
    - Perceived stimulus placed both before and after the verb: a deictic pronoun referring to the stimulus is placed before *se ve que* and the stimulus itself after it: *eso se ve que lo pone la comunidad autónoma* [that it seems that the region provides it].
    - Interrupted: *se ve que* interrupts the syntactic sequence that acts as the perceived stimulus: *En Estados Unidos la justicia se ve que es más rápida que aquí* [In the USA justice it seems is faster than here]; *hay algunos que se dedican/ se ve/ a exportar* [there are some who devote themselves/ it seems/ to exporting].
  - Presence/absence of adjuncts: whether any were present, how many, and their grammatical form (adverb, NP, PP).
  - Scope: This variable was included in the analysis as propositional scope turns out to be one of the fundamental requirements for expressing evidentiality in languages such as Spanish (Boye, 2010). It is a formal-functional variable, straddling the border between formal and semantic/functional parameters. Consequently, a distinction was made between uses that introduce propositional scope and those that introduce states of affairs rather than propositions.

All these variables contributed to the Behavioral Profile analysis, which made it possible to explore the degree of (dis)similarity both within and between the seven senses of the sequence *se ve* and *se ve que*. Of particular interest was the position of the evidential value in relation to the other values in the semantic network.

Table 3 presents a summary of ID tags and their levels:

### 3.3. Statistical analysis

The third step consists in converting these data into a co-occurrence table that provides the relative frequency of co-occurrence of each sense of *se ve que* (columns in Table 4) with each ID tag level (rows in Table 4). As illustrated in Table 4, the percentages of ID tag levels within each ID tag should add up to 1, since all the levels in a tag represent the analysis of all the occurrences for this tag. Each column represents a set of co-occurrence percentages for one sense of the verb (direct physical perception, indirect physical perception, cognitive perception, etc.). This vector of co-occurrence percentages is called a 'Behavioral Profile.' The procedure was performed with Gries's (2010b) BehavioralProfiles 1.01 script written for the R programming language.

Table 3  
Summary of ID tags and ID tag levels.

ID TAG	ID TAG LEVEL
Formal variant	<i>se ve, se ve que, se ve co'mo, se ven</i>
Potential inflection of <i>ver</i>	Third person singular and plural; just third person singular
Syntactic autonomy	Parenthetical, integrated
Position of perceived stimulus	Anteposition, postposition, anteposition + postposition, interrupted
Form of adjunct	NP, adv, PP, n/a
Syntactic relation between qualification and stimulus	Direct/internal modification, indirect/external modification
Scope	propositional, non-propositional
Sense	Direct physical perception, direct physical valorization, indirect physical perception, deduction as result of physical perception, cognitive perception, evaluative perception, evidential
Semantics of perceived stimulus	Physical/concrete entity, abstract entity, concrete metaphorical entity, event/situation, n/a
Semantic qualification of perceived stimulus	Physical qualifiers, axiological/evaluative qualifiers.
Semantics of adjunct	Quantity, frequency, physical space, metaphorical space, physical manner, metaphorical manner, pronominal manner, time, n/a



Table 4  
Examples of BP vectors.

ID tag	ID tag level	<i>direct physical perception</i>	<i>indirect physical perception</i>	<i>cognitive perception</i>	<i>evaluative perception</i>	...
Type of evidential	reportative	0.30	0.36	0.29	0.55	...
	circumstantial	0.35	0.40	0.53	0.30	...
	conjectural	0.01	0.01	0.00	0.02	...
	n/a	0.34	0.23	0.18	0.13	...
Potential mobility	sing./ plural	0.18	0.41	0.24	0.41	...
	just sing.	0.82	0.59	0.76	0.59	...
...	...	...	...	...	...	...

This table only serves as a means to illustrate the conversion of the qualitative formal-functional variables into quantitative BP vectors, which is a crucial intermediate step between corpus annotation and the statistical exploration of the data. The next section will discuss this in more detail.

### 3.4. Exploration through statistical techniques

The next step is to explore Table 4 by means of statistical techniques. In this case, both a hierarchical agglomerative cluster analysis (HAC) and a multidimensional scaling analysis (MDS) were performed. HAC is an exploratory technique that aims to identify and represent (dis)similarity relations between items in the form of a hierarchical tree diagram (dendrogram), composed of several clusters that are characterized by high within-cluster similarity and low between-cluster similarity. MDS is an exploratory dimensionality-reduction technique that seeks to condense a large number of dimensions in a multivariate data set into a smaller number of dimensions—typically two or three (cf. Wheeler, 2005; Hilpert, 2011, 2013:66–74; Levshina, 2011; Croft and Timm, 2013). The basic assumption underlying this technique is that (dis)similarities between the entities under study can be represented as spatial distances. Applied to the concrete example above, this means that the senses in Table 4 can be transformed into points on a two-dimensional plane such that the distances between these points reflect the (dis)similarities between the senses as accurately as possible.

The outcome of the hierarchical cluster analysis led to a dendrogram and to a set of multidimensional scaling maps. The most relevant of these for the status of evidential *se ve que* will be shown and discussed in Section 4.3 (see Figs. 2–3).

In the next section, the results of the corpus analysis are presented and explored through the above-mentioned statistical techniques (HAC and MDS graphs). These results make it possible to address the three research questions, namely:

- (1) How many different senses can be distinguished in the sequence *se ve que* and what are their respective frequencies? (Section 4.1)
- (2) How are these senses correlated with the morphosyntactic behavior of the sequence *se ve que*? (Section 4.2)
- (3) Based on research questions (1) and (2), which senses of *se ve que* are more closely related to each other, and how does the evidential construction relate to this polysemous network? (Section 4.3)

## 4. Results and discussion

### 4.1. Frequencies of senses of *se ve* and *se ve que*

As explained in Section 3.2, the corpus was analyzed for seven values of *se ve* and *se ve que*. They were distributed as shown in Table 5.

This shows that over half the uses (64.8%) expressed the value of ‘evidentiality’. However, as evidential constructions can only introduce propositional scope (Sections 1 and 2.2), the same quantitative data needed to be reexamined using the parameter of scope (propositional vs non-propositional). Interestingly, of the 162 cases of propositional uses (‘indirect

Table 5  
Distribution of senses of *se ve* and *se ve que* in the corpus.

Direct physical valorization	7	3,0
Indirect physical perception	6	2,6
Deduction as a result of physical perception	12	5,2
Cognitive perception	11	4,8
Evaluative perception	19	8,3
Evidentiality	149	64,8
<b>TOTAL</b>	<b>230</b>	<b>100</b>

Table 6  
Distribution of senses of *se ve* and *se ve que* with propositional scope.

Senses	#	%
Indirect physical perception	6	3,7
Cognitive perception	7	4,3
Evidentiality	149	92
<b>Total</b>	<b>162</b>	<b>100</b>

physical perception', 'cognitive perception' and 'evidentiality') of *se ve que*, 149 (92%) were evidential (see Table 6). This reveals a strong tendency for this sequence to appear with evidential value.

#### 4.2. Correlation between morphosyntactic and functional variables in the different senses of *se ve que*

In order to answer the second research question, this section examines the behavior of each of the seven senses of the sequences *se ve* and *se ve que*, based on the correlation between formal and functional traits revealed by the corpus. Each offers different possibilities for recurring combinations of the variables described in Section 3.2. Here below, the syntactic behavior and argument structure related to each sense will be described in detail. Because of their high frequency in some cases, adjuncts will also be taken into account in this description.

##### (i) Direct physical perception

The perception expressed by *se ve* is physical, and therefore objective. *Se ve* introduces a non-propositional stimulus, in principle a physical entity. For this value, the syntactic structure in the corpus is:

Verb <sub>SE VE</sub> + Perceived stimulus <sub>physical entity</sub> + Internal descriptive qualifier <sub>[optional]</sub> + Adjunct <sub>[optional]</sub>
--

The perceived stimulus is obligatory (not only here but for all the senses of *se ve*), as otherwise the act of perceiving could not be satisfied. In Example (7), this stimulus is *las gallinas y los conejos* [the hens and the rabbits]; in (8), *un moreno y un niño rubio* [a dark one and a fair child]. In this category, which has evident compositional value, the verb can be considered a reflexive passive and expresses syntactic agreement between the verb and the stimulus (pluralization is possible: *se ve/ se ven*). Consequently, the stimulus functions as the syntactic subject.

(7)  
nosotros la casita la tenemos en lo alto del pueblo/ de allí **se ven las gallinas y los conejos**  
(Preseea Castellón, Interview 13, p. 337)

we have our house at the top of the village/ from there **you see the hens and the rabbits**

(8)

Cuando ponen un anuncio así de nin~os pequen~os, anunciando, yo que sé, un jabo~n para la ropa, porque lo manchan comiendo chocolate... eso sí me he fijado yo, **se ve un moreno y un nin~o rubio también**

(*Habla culta de Granada*, Interview XI, p. 153)

When they put on an advertisement like that with small children, advertising say a soap for clothes, because they stain them eating chocolate... I've certainly noticed this myself, **you see a dark one and a fair child too**

In this group, qualifiers of the perceived stimulus are optional. They are internal modifiers of the noun phrase (the perceived stimulus), so they modify it directly (RAE, *NGLLE 2009* §13.1m, §13.1n). Example (8) has one—*rubio*—but (7) does not. Of the total occurrences of this category in the corpus (26 out of 230), only 19% contained a stimulus qualifier. The corpus revealed that, semantically, these elements (adjectives or prepositional phrases) designate physical properties (Dixon, 2004:4) referring to objects perceived through the senses (Demonte, 1999, §3.4.2.2). They are objective and descriptive, and do not assess the stimulus subjectively.

An optional adjunct expressing quantity, place, manner or time may also be used in these cases. In example (7), *de allí* [from there] is an adjunct of place. Adjuncts were more numerous than qualifiers in the corpus studied: out of the total occurrences of physical perception, 42% presented an adjunct.

## (ii) Direct physical valorization

This category presents very similar characteristics to the previous one. The perceived stimulus for *se ve* is obligatory, non-propositional and, in principle, physical: it is perceived with the eyes. Unlike the previous category, however, all the stimuli are qualified, although in this case only indirectly: syntactically, the qualifier is a predicative complement of the stimulus (RAE, *NGLLE 2009* §37.1p), so the noun is qualified through the verb *ver* and the evaluation already involves a certain degree of subjective mediation. The corpus only showed two types of semantic property for the qualifiers: quantity (Example 9 below) and size (Example 10) (Dixon, 2004:4, 5; Demonte, 1999:§3.4.2.2).

Optionally, adjuncts can also be used (see Example 11, which contains two adjuncts). Adjuncts of quantity can be considered external evaluators of the perceived stimulus. In the present corpus, all but one of the examples of this category (7 in total) were classified as adverbs of quantity. The only case that had no adjunct of quantity contained an adjective of quantity (predicative complement). Consequently, the evaluation function of the qualifier is obligatory (\*), but the qualifier can be replaced by an adjunct of quantity.

<p>Verb<sub>SE VE</sub> + Perceived stimulus<sub>physical entity</sub> + External qualifier<sub>[quantity, dimension]</sub>* + Adjunct<sub>[optional: quantity, space]</sub>*</p>
---

(9)

**las pulgas no se ven muchas/** pero las cucarachas son→

(Val.Es.Co. 2.0, Conversation 13: line 195)

**fleas you don't see many/** but the roaches are→

(10)

es un pueblloo quee tienee unos doscientos habitantes// es muy extenso// (...) **se ve un pueblo bastante grande/** pero al mismo tiempo/ con muy pocas personas/ claro

(PRESEEA Valencia, interview B16, lines 111–115)

it's a villaage thaat haas around two hundred inhabitants// it's very spread-out// (...) **you see quite a large village/** but at the same time/ with very few people/ of course

(11)

**allí además pan de este/ se ve muy poco/** todo el pan es de molde

(Val.Es.Co. 2002, H38.A1: line 465)

**there also bread like this/ you see very little/** all the bread is tin loaves

**(iii) Indirect physical perception**

In this category, as in the previous one, the starting point is an external physical (ocular) stimulus, but in this case the perception is indirect. Its indirectness is correlated with its form, as the perceived stimulus is now propositional (Example 12). An adjunct of place or time may appear (in this corpus, they were found in one third of this category).

Verb<sub>SE VE</sub> + Perceived stimulus<sub>PROPOSITIONAL</sub> + Adjunct<sub>[optional: space, time]</sub>

(12)  
 (...) los hombres sí que estudiaban algunos/ pero mujeres poquísimas poquísimas/ que a lo mejor de cuatrocientas de un aula/ además ahora **se ve** en las orlas **que a lo mejor hay dos o tres mujeres**  
 (Preseea Castello'n, Interview 67, p. 1325)

(...) the men did study some of them/ but women very very few/ maybe out of four hundred in a lecture hall/ also **you see** now in the class photographs **that maybe there are two or three women**

Even though these are propositional uses, all the occurrences were phonically and syntactically integrated into the utterance, with no parenthetical examples (these were only found for the evidential sense). In these uses *se ve que* inflectionally agrees with the third person singular since syntactically this is an impersonal construction.

**(iv) Deduction as a result of physical perception**

In this category the stimuli are again non-propositional but here they are abstract entities, unlike those of groups (i) and (ii). In these uses the sequence *se ve* can be accompanied by an adjunct. In the corpus, adjuncts were employed in half of the occurrences of this sense (6 out of 12 cases) and in almost all of them the adjunct was one of physical place (Example 13) or metaphorical place (Example 14), although it could also be one of frequency or time). In accordance with the corpus data, the formal-semantic scheme of this category is as follows:

Verb<sub>SE VE</sub> + Perceived stimulus<sub>abstract entity</sub> + Adjunct<sub>[optional]</sub>

In Example 13, physically seeing people in cardboard boxes in the old riverbed leads the speaker to conclude (in B's second turn) that *mucha miseria* [much misery] is seen.

(13)  
 B: si no toda esa gente no estaría ahí bajo del río de la forma que están entre cartones/ ¿eh? (...) yo lo estoy viendo cada día  
 A: paseando [por arriba del puente se ve]  
 B: [cada día/ si pasas por arriba]/ se ve/ **y se ve mucha miseria ahí** ¿eh?// MUCHA MUCHA miseria// date cuenta con el frío que hace  
 (PRESEEA Valencia, interview B20, lines 629–632)

B: otherwise all those people wouldn't be down there in the river the way they are among cardboard boxes/ eh? (...) I'm seeing that every day  
 A: going [over the bridge you see it]  
 B: [every day/ if you go over]/ you see it/ **and you see a lot of misery there**, eh?// A WHOLE LOAD of misery// just think when it's this cold

The perceived stimuli in Examples (13) and (14) are abstract, since *miseria* ([misery], Ex. 13), *diferencia* and *contraste* ([difference] and [contrast], Ex. 14) are qualities, relationships or effects of other elements, but are not themselves perceived through the eyes. In (14) the adjunct is metaphorical because what *ahí* [there] refers to is diversification between “very posh people” and “more ordinary people”.

(14)

el barrio del Carmen (...) está más diversificado/ hay sitios que va gente muy pija (...) y gente más normal/ **ahí sí que se ve la diferencia**/ de una calle a otra o de una zona a otra// yo es que tampoco es que salga mucho aquí pero/ hay sitios/ más normales y sitios de mucho lujo/ muy caros/ yy **se ve el contraste**

(PRESEEA Valencia, interview ESA14, lines 160–165)

The Carmen district (...) is more diversified/ there are places where verry posh people go (...) and more ordinary people/ **there definitely you see the difference**/ from one street to another or one area to another// me it's not that I go out much here but/ there are places/ that are more ordinary and veery luxurious places/ very expensive/ and **you see the contrast**

As explained in Section 3.2, other authors have classified these cases as indirect perception (Willems, 1983; Hanegreefs, 2008; Fernández-Jaén, 2012), since they express a mental conclusion based on the perception of a visual stimulus. This group differs from the previous one (iii, indirect physical perception) in that the perceived stimuli are not propositional. It differs from the following category in continuing to have a physical reference, which the cognitive perception sense does not.

#### (v) Cognitive perception

In cognitive perception, as mentioned in the previous paragraph, the perception process happens entirely in the mind. As it includes two possibilities, two patterns are available for constructing this sense:

- (a) mental appreciation of an abstract reality (non-propositional stimulus), as in a metaphorical view (Example 15)
- (b) the result of an inference the speaker draws based on his or her knowledge of the world and on logical reasoning (Example 16). In this case, there is a clause with an inflected verb form.

- (a) Verb<sub>SE VE</sub> + Perceived stimulus<sub>abstract entity</sub> + Adjunct<sub>[optional: metaphorical space, frequency, manner, quantity]</sub>  
 (b) Verb<sub>SE VE</sub> + Perceived stimulus<sub>PROPOSITIONAL</sub> + Adjunct<sub>[optional: metaphorical space, manner]</sub>

In (15), *ver* involves mental reception of a non-propositional abstract stimulus (*una carrera universitaria* [a university degree course]). It is accompanied by adjuncts of metaphorical place (*en los círculos en que te mueves* [in the circles you move in]), frequency (*siempre* [always]) and quantity (*mucho más* [much more]).

(15)

A: ¿y tú piensas/ que sería mejor/ estudiar una carrera universitaria/ o una formación profesional?

B: sí/ yo creo que las dos cosas/ (...) ahora están preocupándose más por la formación profesional// y el típico mito de que los que no sirven para formación profesional/ creo que ya no existe casi/ aunque bueno siempre los padres y los familiares/ y en los círculos en que te mueves pues/ siempre **una carrera universitaria se ve mucho más**

(PRESEEA Valencia, interview MED14, lines 285–299)

A: and do you think/ it would be better/ to study for a university degree/ or vocational training?

B: yes/ I think both/ (...) now people are paying more attention to vocational training// and the typical myth that those who are no good go for vocational training/ I think practically no longer exists/ although of course parents and families always/ and in the circles you move in well/ **you see university degrees much more** always

In (16) the speaker expresses a conclusion reached through the ideas expressed by his or her interlocutor, not through physical visual stimuli. This example also uses an adjunct:

(16)

A: él me decía que los empresarios contratan a chicas porque siempre quedan mejor/ a la gente le gusta más/ yy yo le decía *hombre pues si van a comprar mujeres/ yo creo que a una mujer antes le gustará un chico que no una chica*/ no pero/ él me lo rebatía/ diciendo que no/ que una chica siempre lo hace más aseado

B: ¡qué va!! es una tontería enorme// **ahí se ve que España es machista a tope** y de ahí no va a salir nunca/ la verdad (PRESEEA Castellón, Interview 26, p. 543)

A: he told me that employers take on girls because they always make a better impression/ people like that more/ and I said to him/ *man well if women go shopping/ I think a woman will like a guy better than a girl* no but/ he countered that/ saying no/ a girl always does things more neatly

B: no way!! that's complete rubbish// **there you see that Spain is totally sexist** and is never going to get away from it/ honestly

The corpus results identified 11 cases of 'cognitive perception' (4.8% of the total corpus). The number of occurrences of the two patterns was similar: 5 type (a), non-propositional, and 6 type (b), propositional. At least one adjunct was found in almost all cases in both groups.

### (vi) Evaluative perception

In 'evaluative perception,' a quality is attributed subjectively to the perceived stimulus. As Fernández-Jaén (2012:360–361) points out, "this is a highly axiological and modalized process in which the verb of perception is reanalyzed until it becomes a verb of propositional attitude like *juzgar* or *considerar*" [to judge; to consider]. The syntactic pattern for this value is as follows:

Verb<sub>SE VE</sub> + Perceived stimulus<sub>abstract, physical entity</sub> + External axiological qualifier\* + Adjunct<sub>[optional: quantity, frequency, space, manner]</sub>

Some authors include propositional uses of the perceived stimulus in the 'evaluative perception' category (for example, Fernández-Jaén, 2012:361–363). However, at least in the present corpus study of the sequences *se ve* and *se ve que*, all the uses recorded (19 out of 230, 8.3% of the total corpus) were non-propositional (as also found by Willems and Defrancq, 2000:16). The perceived stimuli can be abstract (*el pasado* in Example 17, including events or situations: *juego* in Example 18), or physical (like *video* in Example 19):

(17)

B: lo importante de Venecia es su **pasado** que **se ve maravilloso**  
(PRESEEA Castellón, Interview 33, p. 708)

B: the important thing about Venice is its **past/** which **looks wonderful**

As the perceived stimulus is qualified by a predicative complement, it is qualified through the verb (RAE, *NGLE* 2009 §37.1p). The qualifiers are axiological, evaluative adjectives (Demonte, 1999:§3.4.2.2; Dixon, 2004:4; *good, bad, lovely, atrocious, nice, ugly, horrible*, etc.). They are affective adjectives and the perceiver is heavily involved in the subjective judgment (Demonte, 1982). This can be seen in Example (18): the qualifier *aburrido* [boring] is A's opinion, as the other interlocutors think differently and express their disagreement:

(18)

B: yo sé jugar che// pero a mí no me gusta

C: es divertido (...)

A: de todas formas **se ve un juego aburrido**

B: no es verdad

(Val.Es.Co. 2.0, Conversation 37, lines 183–191)

B: I know how to play man// but I don't like it

C: it's fun (...)

A: anyway **it looks like a boring game**

B: that's not true

In (19) the adjective *fuerte* (strong, heavy) is used metaphorically and expresses an evaluation of the video:

(19)

A: la canción la he oído poco pero la verdad es que sí ↓ sí que creo que tiene marcha de ella → // pero el vídeo ↑ no lo he visto/ o sea creo que vi el comienzo un poco/ pero **[se ve demasiado fuerte]**



B: [el vídeo ↑ es que le sobran ↑] / le sobran ↑ loos tres primeros minutos  
(Val.Es.Co. 2.0, Conversation 1, lines 413–421)

A: the song I haven't heard much but the truth is that yes ↓ yes I think it's got her liveliness → // but the video ↑ I haven't seen it/ actually I think I saw the beginning a bit/ **buut [it looks too heavy]**

B: [the video ↑ it could do without ↑] / without ↑ thee first three minutes

In this type of perception the qualifier is obligatory, in principle, as it is through this element that the evaluation is applied. However, the qualifier can be replaced by an adjunct of manner (which is why the asterisk \* has been included in the pattern). In the present corpus this only occurs in two of the 19 cases; one is (20):

(20)

la ensen~anza desde fuera se ve muy bien  
(Corpus COJEM- Palma, C4, line 122)  
from the outside the teaching seems very good

### (vii) Evidentiality

As noted in Section 4.1, evidentiality is the most frequent value for the sequences *se ve* and *se ve que* in this corpus: 149 occurrences out of a total of 230. It is the most abstract meaning and completely grammatical in nature. It expresses the source of the information, so semantically there is a qualitative leap between the previous senses and this one, which can no longer be said to involve any type of perception. Another decisive formal difference is that the sequence is highly fixed, so it can already be considered quasi-constructuralized (Traugott and Trousdale, 2013). In addition, another basic semantic and formal requirement for expressing evidential value in non-evidential languages such as Spanish (see Section 2.2) is that evidentials must have scope over propositions (Boye, 2010). Hence, this is the only value in the corpus that allows parenthetical uses. The two possible patterns for expressing evidentiality are:

- |  |
|--|
| <p>(a) Parenthetical construction<sub>SE VE + Proposition</sub><br/>(b) Non-parenthetical construction<sub>SE VE QUE + Proposition</sub></p> |
|--|

Two different patterns have been distinguished, based on the idea that the parenthetical use is becoming more frequent in Spanish. Nevertheless, it should be stressed that this is one and the same construction (*se ve que*), although it presents an incidental, right-dislocated position (*se ve*, losing the complementizer *que*) in specific contexts.

Compared to the previous patterns, this one contains a construction instead of a verb. Inflectionally, the verb agrees with the third person singular (its only option), as in the other values found in this corpus that also introduce propositional scope ('cognitive perception' and 'indirect physical perception'), and it reaches the highest level of constructionalization in its parenthetical uses (Cuenca and Marín, 2012; Rossari, 2012). The evidential value patterns present another important difference compared to all the previous categories: none of the cases in the corpus contain adjuncts. This is another clear sign that *se ve que* is losing its value as a verb and can already be considered a construction.

The corpus records 10 parenthetical uses among the 149 evidentials, as in Example (21). Here, the discourse is interrupted by *se ve*, which constitutes a phonic group on its own and acts as a parenthetical:

(21)

P: y ahora te mandan si quieres ↑ / chorizo de allí yy queso de allí/ y eso lo mandan allí/ hay algunos que se dedican ↓ **se ve** ↓ a exportar// y venden en los bares  
(Val.Es.Co. 2002, PG119.A1: lines 241–242)

P: and now they send you if you want ↑ / chorizo from there and cheese from there/ and they send it there/ there are some people that dedicate themselves ↓ **apparently** ↓ to export/// and they sell in the bars

Another parenthetical evidential can be seen in Example (4) above, and an evidential that is part of the utterance (non-parenthetical) in Example (3) above (see Section 1).

Following this characterization of the seven values, the main formal and semantic resemblances and differences between them can now be highlighted here below in order to interpret the different clustering solutions that emerged from the BP approach (Fig. 1) and the multidimensional scaling map (MDS) (Fig. 2). The aim is to recognize which of the compositional values of this sequence (the first six senses) are closest to its evidential value, and which traits they share.

4.3. Assessing the degree of cluster (dis)similarity: on the status of evidential *se ve que* in the network

So far, the different senses have been described separately, through sets of morphosyntactic and functional parameters that distinguish them from each other. However, based on the definition of polysemy (and on what distinguishes polysemy from homonymy), all of these senses are expected to be related to each other in a particular way, some more closely than others. This section presents the hierarchical cluster analysis (HAC) results that led to the dendrogram in Fig. 1 and to the multidimensional scaling map (MDS) of senses of *se ve que* (see Fig. 2). It examines precisely which senses of *se ve que* are more closely related in the dendrogram and in the MDS map, and how the evidential construction relates to this polysemous network.

Fig. 2 shows the result of an MDS analysis:

Both the dendrogram (Fig. 1) and the MDS map (Fig. 2) present a clear image of how the different senses of *se ve que* cluster together. As indicated by the two red boxes in Fig. 1 (and by the spatial distances and proximities in Fig. 2), two significant clusters can be distinguished. The first cluster groups four senses together: Direct physical perception, Evaluative perception, Direct physical valorization, and Deduction as a result of physical perception. The second large

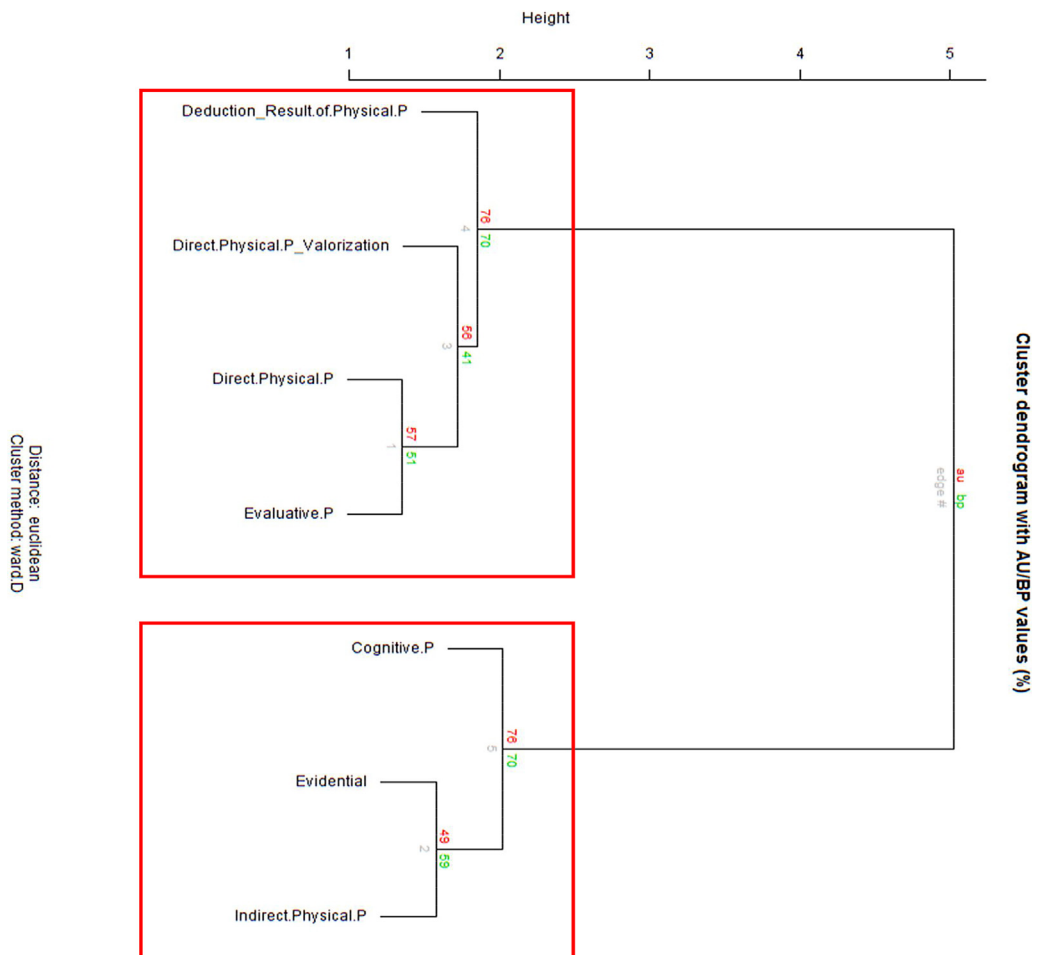


Fig. 1. Cluster dendrogram of *se ve* and *se ve que* senses.

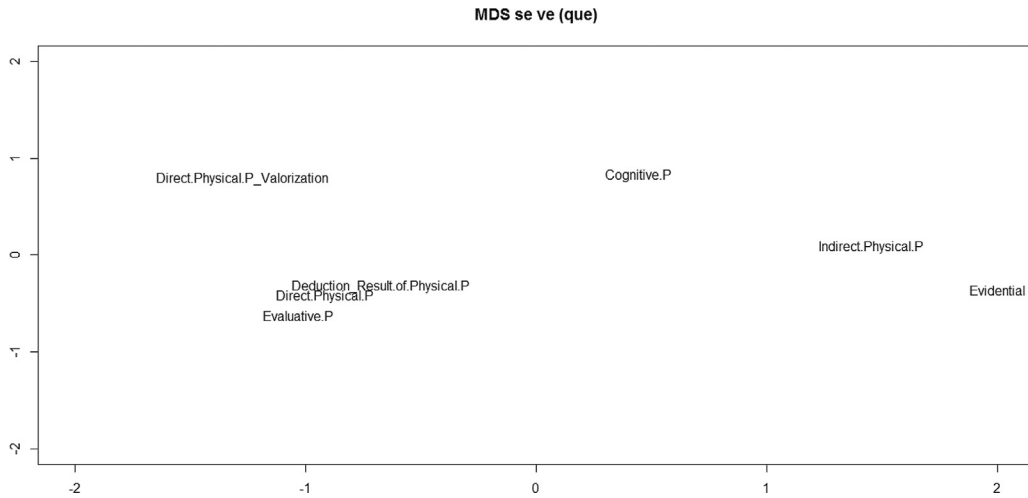


Fig. 2. Multidimensional scaling map of senses of *se ve* and *se ve que*.

cluster comprises Evidentiality, Indirect physical perception and Cognitive perception. In the remainder of this paper, these clusters will be referred to as CLUSTER1 and CLUSTER2, respectively.

The most discriminating variable leading to this bipartite division turned out to be scope. The senses included in CLUSTER1 introduce a state of affairs, that is to say, physical or abstract entities, but not a proposition. In contrast, in CLUSTER2 the scope is propositional in the great majority of cases (Evidentiality, Indirect physical perception, and 64% of Cognitive perception). This bipartite division of the data can be visualized clearly by performing a MDS analysis of the individual examples in order to explore the relationships between them in a low-dimensional space. Fig. 3 plots the two levels of the ‘scope’ variable (propositional and non- propositional) according to the positions of the examples showing these levels in the MDS solution.

In addition to scope, two other features also contribute to this clustering solution: the potential inflection of the verb form and the formal variants of *se ve que*. Whereas in CLUSTER1 the verb can be used both in singular and in plural (*se ve* and *ven*), CLUSTER2 is restricted to the singular form, which agrees with the syntactic impersonal construction, although two formal variants are used: *se ve* (without the option to pluralize) and *se ve que*. Consequently, it comes as no surprise that the evidential sense appears in CLUSTER2, since as argued above (Section 4.2, vii), one of the essential features of constructionalization of this evidential marker is precisely its morphosyntactic immobilization. In other words, it can be said that the evidential construction evolves from the impersonal constructions in Cluster 2.

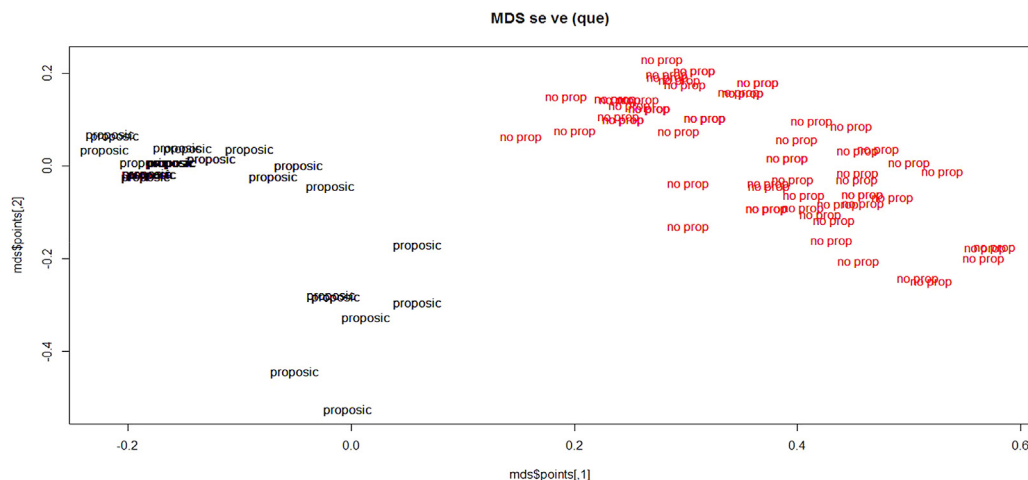


Fig. 3. Multidimensional scaling map of the ‘scope’ variable.

In addition to the above, it will also be seen that CLUSTER2 presents less capacity for receiving arguments than CLUSTER1. In the four values in CLUSTER1, a stimulus acts as the syntactic subject of the reflexive passive form of the verb *ver*. Moreover, in three of these four values the stimulus has a qualifier (direct or indirect/predicative modifier). Furthermore, they all present the option of including adjuncts. In contrast, practically no arguments appear in CLUSTER2 and it is only possible to speak of a syntactic subject in the non-propositional uses of ‘cognitive perception’ (considering the construction a reflexive passive: *en los círculos en que te mueves, una carrera universitaria se ve mucho más*—see Example (15)). In propositional uses, *se ve que* impersonalizes the sentence and it is no longer possible to speak of a syntactic subject, only of a direct object at sentence level (Maldonado, 1996, 1999). Also, qualifiers are not used and there is less possibility of encountering adjuncts. In ‘evidentiality’ uses no adjuncts were found, nor were any recorded in 67% of the cases of ‘indirect physical perception’ or 18% of the cases of ‘cognitive perception’. In short, ‘evidentiality’ is the starkest of the three values in its use of complements: its constructionalization has reduced its ability to receive arguments.

Having noted the differences between the two large clusters, it is time to discuss the (dis)similarities within each cluster. CLUSTER1 is large and contains two sub-clusters. One is ‘deduction as a result of physical perception’ (SUBCLUSTER1.1) and the other covers three senses: ‘direct physical valorization’, ‘direct physical perception’, and ‘evaluative perception’ (SUBCLUSTER1.2). The distinction is based on two features of ‘deduction as a result of physical perception’, which unlike the other sub-cluster (a) does not allow qualifiers of the stimulus, but only indicates the existence of a cognitive stimulus arising from physical stimuli; and (b) has cognitive (abstract) stimuli, not physical ones. These differences are important because, as seen in the MDS plot (Fig. 2), ‘deduction as a result of physical perception’ is the value within CLUSTER1 (on the left-hand side of the MDS plot) that lies closest to CLUSTER2 (on the right-hand side of the MDS plot).

In SUBCLUSTER1.2, the distinction between the three values lies in the semantics of the adjective qualifying the perceived stimulus. As seen in Section 4.2, in ‘direct physical perception’ the qualifiers are descriptive and objective; in ‘evaluative perception’ they are axiological and subjective, and in ‘direct physical valorization’ they are of quantity or size. Also, ‘direct physical perception’ and ‘evaluative perception’ appear to differ from ‘direct physical valorization’ in that adjuncts were found in almost all cases of the latter (6 out of 7) but in under half the cases of ‘direct physical perception’ and ‘evaluative perception’.<sup>2</sup>

Turning to CLUSTER2, what is most striking is that the closest meaning to ‘evidentiality’ is ‘indirect physical perception’. This can be seen in both the dendrogram and the MDS plot (on the right-hand side of Fig. 2). Sentence-level clauses are obligatory in both these values, so there is no longer any perceived stimulus. Formal indirectness entails semantic indirectness and, consequently, mental mediation by the speaker. Both categories express more abstract values, even though ‘indirect physical perception’ arises out of visual stimuli. This also occurs in a specific type of evidentiality expressed by *se ve que*. In fact, delving deeper into the cases recorded in the corpus, evidentiality can be classified into three sub-types (as proposed by Squartini, 2008):

- Circumstantial Inferential: 28%
- Generic and Conjectural Inferential: 46%
- Reportative: 26%

The difference between these types of evidentiality is the relationship between the speaker’s involvement in the reasoning and the presence of external evidence (Squartini, 2008). In circumstantial inference evidentials, the speaker’s inference is based on external sensory evidence. In generic and conjectural inferences, “all external evidence is missing, the speaker being solely responsible for the reasoning process” (Squartini, 2008:925). Hence, circumstantial inferential evidentiality is semantically the closest to ‘indirect physical perception’, as both base the inference on perceptual stimuli.

In the case of reportative evidentials, as noted by Estellés (2018) and Izquierdo (2016), in Spanish constructions that express evidentiality the evidential values reported are really inferences based on auditory stimuli or oral data. If this proposal is accepted, it will be seen that over half the evidential uses found in this corpus were generated through perceptible stimuli (visual or auditory).

It is also interesting to note that in principle, Spanish does not express direct evidentiality through grammaticalized constructions (Cornillie, 2007; Estellés and Albelda, 2014; Albelda, 2015; Kotwica, 2017). The MDS plot (Fig. 2) shows this clearly, as ‘evidentiality’ occupies the farthest position.

<sup>2</sup> Another way, although less important, in which ‘direct physical valorization’ can be considered to differ from ‘direct physical perception’ and ‘evaluative perception’ is that it presents greater mobility of the perceived stimulus in relation to the verb (before, after, interrupting, and before & after were recorded in this corpus analysis). In both ‘direct physical perception’ and ‘evaluative perception’, the perceived stimulus only appeared before or after the verb.

Lastly, what brings ‘cognitive perception’ closer semantically to ‘evidentiality’ is that both involve the intellect and express mental (therefore abstract) values.

In conclusion, ‘evidentiality’ shares with ‘indirect physical perception’ and ‘cognitive perception’ not only the formal question of propositional scope, but also that *se ve que* can express both circumstantial inferential and reportative evidentiality based on physical stimuli (like ‘indirect physical perception’), on the one hand and, on the other, generic and conjectural evidence of a purely mental nature (like ‘cognitive perception’).

Finally, this cluster solution seems to point toward the existence of two quite different constructions. On the one hand, this outcome seems to suggest that CLUSTER1 represents the root visual perception meaning of *se ve* and *se ve que* along with the reflexive passive construction. On the other hand, CLUSTER2 seems to correspond to an impersonal construction where the subject is systematically a schematic human that considers and evaluates the content of a sentential object. This syntactic impersonal construction of *se ve que* in CLUSTER2 not only gives coherence to all the constructions in this cluster but also makes it possible to account for the emergence of the evidential construction. In other words, the evidential construction may evolve from or be an extension of the impersonal construction. This is an interesting line of research that merits future study.

## 5. Concluding remarks

This characterization of the different formal-functional combinations of the sequence *se ve que*, together with the study of its behavioral profile, has demonstrated its high degree of specialization as an evidential. The three research questions raised in the introduction reflect the general objective of this study on the specialization of *se ve que* as an evidential, as discussed below.

First, seven senses of the sequences *se ve* and *se ve que* were distinguished in the corpus, with sufficiently different behavior between each. Of these, ‘evidentiality’ presented high frequency, a factor that is considered to favor specialization of *se ve que* as an evidential. Out of the total occurrences of this sequence in the corpus, both propositional and non-propositional, ‘evidentiality’ accounted for 64.8%. Considering only the occurrences with propositional scope (a requirement for expressing evidentiality), the evidential value of *se ve que* accounted for 92% of the total.

Second, in-depth study of the morphosyntactic and functional correlates of the seven senses of *se ve que* clearly shows the state of constructionalization of the ‘evidentiality’ value. Accordingly, ‘direct physical perception’ and ‘evidentiality’ were found at opposite ends of a scale ranging from greater to lesser inflectional and argument mobility. As mentioned in Section 4.3, in evidential uses the possible arguments are fewer; this also happens to a lesser extent in the nearest compositional value, ‘indirect physical perception’, as clearly seen in the dendrogram and the MDS plot.

Third, regarding the degree of (dis)similarity between the different senses of *se ve que*, particularly ‘evidentiality’, the statistical analysis has shown that the main factor separating CLUSTER1 from CLUSTER2 (which includes evidentiality) is the formal indirectness of the latter, which accompanies the introduction of propositional scope. This formal indirectness is correlated with semantic indirectness. For this reason, the CLUSTER2 values either move in the area of cognitive and more abstract values (‘cognitive perception’, generic and conjectural ‘evidentiality’) or, if they present a physical value, do so indirectly (‘indirect physical perception’, and circumstantial and —falsely— reportative ‘evidentiality’). Accordingly, thanks to the analysis of evidentiality sub-types (Section 4.3), it can be seen more clearly that the ‘evidential’ sense shares similar traits (distribution on the graphics, theoretical features) to ‘indirect physical perception’ and ‘cognitive perception’, and they are all placed in the same cluster.

In short, this study suggests that the evidential construction *se ve que* presents a high degree of specialization, as shown by the frequency of its use. It also shows that the position of ‘evidentiality’ in the clustering solution resulting from the BP graph demonstrates its formal and semantic closeness to two other senses in the same cluster: ‘indirect physical perception’ and ‘cognitive perception’. These three senses share propositional scope and morphosyntactic verbal agreement in the third person singular. Indeed, the introduction of an inflected verb clause automatically entails perception becoming more abstract and prevents an exclusively physical perception. This, in turn, is a factor that favors constructionalization, as formal indirectness leads to the emergence of more abstract senses, such as evidentiality.

## Funding

This work was supported by the Spanish Ministry of Economy and Competitiveness, MINECO (Project FFI2016-75249-P, La atenuación pragmática en su variación genérica: géneros discursivos escritos y orales en el español de España y América).

## Acknowledgements

The authors are grateful to Mary Georgina Hardinge for translation and English editing assistance.

## References

- Aikhenvald, A., 2004. *Evidentiality*. Oxford University Press, Oxford.
- Albelda, M., 2015. *Evidentiality in non-evidential languages: are there evidentials in Spanish?* J. Pragmat. 85, 135–137.
- Albelda, M., 2016. *La expresio´n de la evidencialidad en la construccio´n se ve (que)*. Span. Context 13 (2), 237–262.
- Albelda, M., 2018. *¿Atenuacio´n del compromiso del hablante?: el caso de los evidenciales por lo visto y se ve que*”, Rilce. Rev. Filol. Hisp. 34 (3), 1179–1214.
- Alonso-Almeida, F., 2015. *The functions of seem and parecer in early medical writing*. Discourse Stud. 17 (2), 121–140.
- Anderson, L.B., 1986. *Evidentials, paths of change and mental maps: typologically regular asymmetries*. In: Chafe, W., Nichols, J. (Eds.), *Evidentiality: The Linguistic Coding of Epistemology*. Ablex Publishing Corp., Norwood, NJ, pp. 273–312.
- Barros, P., Barros, M.J., Lo´pez, M.P., Morales, J., 2012. *Cogila: espan˜ol oral conversacional*. Universidad de Granada, Granada.
- Biber, D., Conrad, S., Reppen, R., 1998. *Corpus Linguistics. Investigating Language Structure and Use*. Cambridge University Press, Cambridge.
- Blas-Arroyo, J.L. (coord.), 2010. *Corpus sociolingüístico de Castello´n de la Plana y su área metropolitana*. Publicacions Universitat Jaume I, Castello´n.
- Bolinger, D., 1968. *Entailment and the meaning of structures*. Glossa 2, 119–127.
- Boye, K., 2010. *Evidence for what? Evidentiality and scope*. STUF-Language Typology and Universals Sprachtypologie und Universalienforschung 63 (4), 290–307.
- Briz, A., Val.Es.Co. Group, 2002. *Corpus de conversaciones coloquiales*. Arco/Libros, Madrid.
- Cabedo, A., Pons, S. (Eds.), 2002. *Corpus Val.Es.Co 2.0*. Retrieved from: <http://www.valesco.es> (online).
- Cornillie, B., 2007. *Evidentiality and Epistemic Modality in Spanish (Semi-) Auxiliaries. A Cognitive-Functional Approach*. Mouton de Gruyter, Berlin.
- Croft, W.A., Timm, J., 2013. *Using Optimal Classification for Multidimensional Scaling Analysis of Linguistic Data*.
- Cuenca, M.-J., Marín, M.-J., 2012. *Discourse markers and modality in spoken Catalan: the case of (és) clar*. J. Pragmat. 44 (15), 2211–2225.
- De Haan, F., 2003. *Visual Evidentiality and Its Origins*. Unpublished manuscript, University of Arizona.
- De Haan, F., 2013a. *Semantic distinctions of evidentiality*. In: Dryer, M.S., Haspelmath, M. (Eds.), *The World Atlas of Language Structures Online*. Max Planck, Leipzig Accessible at: <http://wals.info/chapter/77>
- De Haan, F., 2013b. *Coding of evidentiality*. In: Dryer, M.S., Haspelmath, M. (Eds.), *The World Atlas of Language Structures Online*. Max Planck Institute for Evolutionary Anthropology, Leipzig Accessible at: <http://wals.info/chapter/78>
- De la Mora, J., Maldonado, R., 2015. *Dizque: epistemics blurring evidentials in Mexican Spanish*. J. Pragmat. 85, 168–180.
- Demonte, V., 1982. *El falso problema de la posicio´n del adjetivo. Dos análisis semánticos*. BRAE LXII 453–485.
- Demonte, V., 1999. *El adjetivo. Clases y usos. La posicio´n del adjetivo en el sintagma nominal*. In: Bosque, I., Demonte, D. (Eds.), *Gramática descriptiva de la lengua espan˜ola*. Espasa/Calpe, Madrid, pp. 129–215.
- Diewald, G., Smirnova, E., 2010. *Introduction. Evidentiality in European Languages: the lexical-grammatical distinction*. In: Diewald, G., Smirnova, E. (Eds.), *The Linguistic Realization of Evidentiality in European Languages*. Mouton de Gruyter, Berlin/New York, pp. 1–14.
- Divjak, D.S., Gries, S.Th., 2006. *Ways of trying in Russian: clustering behavioral profiles*. Corpus Linguist. Linguist. Theory 2 (1), 23–60.
- Dixon, R.M.W., 2004. *Adjective classes in typological perspective*. In: Dixon, R.M.W., Aikhenvald, A.Y. (Eds.), *Adjective Classes. A Cross-linguistic Typology*. Oxford University Press, pp. 1–49.
- Enghels, R., 2007. *Les modalités de perception visuelle et auditive. Différences conceptuelles et répercussions sémantico-syntaxiques en espagnol et en français*. Niemeyer, Tübing.
- Escandell Vidal, M.V., 2014. *Evidential futures: the case of Spanish*. In: De Brabanter, P., Kissine, M., Sharifzadeh, S. (Eds.), *Future Tense vs. Future Time: An Introduction*. Oxford University Press, Oxford, pp. 219–246.
- Estellés, M., 2018. *What genres tell us about evidentials and vice versa: a study of al parecer in Spanish parliamentary debates*. Pragmat. Soc. 9 (3), 401–427.
- Estellés, M., Albelda, M., 2014. *Intonation, evidentials and politeness in Spanish. A corpus analysis*. J. Politeness Res. 10 (1), 29–62.
- Fernández, S.S., 2013. *Impersonality in Spanish personal pronouns*. In: Jeppesen, K., Lindschouw, J. (Eds.), *Deixis and Pronouns in Romance Languages*, pp. 87–108.
- Fernández-Jaén, J., 2012. *Semántica diacrónica cognitiva de los verbos de percepcio´n física del espan˜ol* (Ph.D. thesis). Universidad de Alicante.
- Fillmore, C., Kay, P., O’Connor, M.C., 1998. *Regularity and idiomacity in grammatical constructions: the case of Let Alone*. Language 84 (3), 501–538.
- Firth, J.R., 1957. *Papers in Linguistics, 1934–1951*. Oxford University Press, London/New York.
- Figueras, C., Cabedo, A. (Eds.), 2018. *Perspectives on Evidentiality in Spanish: Explorations across Genres*. John Benjamins, Amsterdam.
- García-Negróni, M.M., 2016. *Polifonía, evidencialidad citativa y tiempos verbales. Acerca de los usos citativos del futuro morfológico y del futuro perifrástico*. In: González, R., Izquierdo, D., Loureda, O. (Eds.), *La evidencialidad en espan˜ol: teoría y descripcio´n*. Eunsa, Pamplona, pp. 279–302.
- García-Negróni, M.M., Libenson, M., 2014. *“Esto/eso, que X/de que X” en contraste. Del significado evidencial perceptivo al significado evidencial citativo*. Estudios de lingüística. ELUA 28, 235–250.
- Geeraerts, D., 2006. *Methodology in cognitive linguistics*. In: Kristiansen, G., Achard, M., Dirven, R., Ruiz, F. (Eds.), *Cognitive Linguistics: Current Applications and Future Perspectives*. Mouton de Gruyter, Berlin & New York, pp. 21–50.
- Geeraerts, D., 2010. *The doctor and the semantician*. In: Glynn, D., Fischer, K. (Eds.), *Quantitative Methods in Cognitive Semantics: Corpus-driven Approaches*. Mouton de Gruyter, Berlin & New York, pp. 63–78.



- Glynn, D., 2010. Testing the hypothesis: objectivity and verification in usage-based Cognitive Semantics. In: Glynn, D., Fischer, K. (Eds.), *Quantitative Methods in Cognitive Semantics: Corpus-driven Approaches*. Mouton de Gruyter, Berlin & New York, pp. 239–269.
- Glynn, D., Fischer, K. (Eds.), 2010. *Quantitative Methods in Cognitive Semantics: Corpus-Driven Approaches*. Mouton de Gruyter, Berlin & New York.
- Goldberg, A., 2006. *Constructions at Work: the Nature of Generalization in Language*. Oxford University Press.
- Go´mez-Molina, J.R. (Coord.), 2001–2007. *El espan~ol hablado de Valencia. Materiales para el estudio sociolingüístico*, vols. I, II, III. Nivel sociocultural Alto, Medio, Bajo. Universitat de València, Valencia.
- González-Márquez, M., Mittelberg, I., Coulson, S., Spivey, M. (Eds.), 2007. *Methods in Cognitive Linguistics*. John Benjamins, Amsterdam & Philadelphia.
- González-Ramos, E., 2016. Por lo visto y al parecer: evidencialidad y atenuación de compromiso con la veracidad de un contenido enunciado. In: González, R., Izquierdo, D., Loureda, O. (Eds.), *La evidencialidad en espan~ol: teoría y descripción*. Eunsa, Pamplona, pp. 129–151.
- Gries, S.Th., 2010a. Behavioral profiles: a fine-grained and quantitative approach in corpus-based lexical semantics. *Mental Lexicon* 5 (3), 323–346.
- Gries, S.Th., 2010b. Behavioral Profiles 1.01. A Program for R 2.7.1 and Higher.
- Gries, S., Divjak, Th.D., 2009. Behavioral profiles: a corpus-based approach to cognitive semantic analysis. In: Evan, V., Pourcel, S. (Eds.), *New Directions in Cognitive Linguistics*. John Benjamins, Amsterdam, pp. 57–75.
- Gries, S., Stefanowitsch, Th.A. (Eds.), 2006. *Corpora in Cognitive Linguistics: Corpus-Based Approaches to Syntax and Lexis*. Mouton de Gruyter, Berlin & New York.
- Grossmann, F., Tutin, A., 2010. Evidential markers in French scientific writing: the case of French verb voir. In: Diewald, G., Smirnova, E. (Eds.), *Linguistic Realization of Evidentiality in European Languages*. de Gruyter Mouton, Berlin, New York, pp. 279–307.
- Hanegreefs, H., 2008. Los verbos de percepción visual. Un análisis de corpus en un marco cognitivo. Katholieke Universiteit Leuven, Leuven.
- Hanks, P., 1996. Contextual dependency and lexical sets. *Int. J. Corpus Linguist.* 1, 75–98.
- Hassler, G., 2010. Epistemic modality and evidentiality and their determination on a deictic basis. The case of Romance languages. In: Diewald, G., Smirnova, E. (Eds.), *The Linguistic Realization of Evidentiality in European Languages*. Mouton de Gruyter, Berlin, pp. 223–247.
- Henneman, A., 2013. *A Context-sensitive and Functional Approach to Evidentiality in Spanish or Why Evidentiality Needs a Superordinate Category*. Frankfurt am Main, Peter Lang.
- Hidalgo, A., 2014. Special issue “The prosodic expression of linguistic im/politeness in Romance languages”. *J. Politeness Res.* 10 (1).
- Hilpert, M., 2011. Dynamic visualizations of language change. Motion charts on the basis of bivariate and multivariate data from diachronic corpora. *Int. J. Corpus Linguist.* 16 (4), 435–461.
- Hilpert, M., 2013. *Constructional change in English: Developments in Allomorphy, Word Formation, and Syntax*. Cambridge University Press, Cambridge.
- Izquierdo, D., 2016. *Alcance y límites de la evidencialidad. Aspectos teóricos y propuesta de análisis aplicada a un conjunto de adverbios evidencialoides del espan~ol* (Ph.D. Dissertation). Universidad de Navarra, Pamplona.
- Jansegers, M., Albelda, M., 2018. Estudio contrastivo de los evidenciales ‘por lo visto’ y ‘se ve (que)’ en las variantes dialectales del espan~ol. In: Fuentes, C., Messias, A., Martí, M. (Eds.), *Marcadores discursivos en espan~ol y portugués: contrastes, enseñanza y traducción*. (in press).
- Jansegers, M., Gries, S.Th., 2017. Towards a dynamic behavioral profile: a diachronic study of polysemous sentir in Spanish. *Corpus Linguist. Linguist. Theory* 1–42.
- Jansegers, M., Vanderschueren, C., Enghels, R., 2015. The polysemy of the Spanish verb sentir: a Behavioral Profile analysis. *Cognit. Linguist.* 26 (3), 381–421.
- Jucker, A., Schreier, D., Hundt, M. (Eds.), 2009. *Corpora. Pragmatics and Discourse*. Rodopi, Amsterdam.
- Kirsner, R.S., Thompson, S.A., 1976. The role of pragmatic inference in semantics: a study of sensory verb complements in English. *Glossa* 10 (2), 200–240.
- Kotwica, D., 2015. Evidential al parecer: between the physical and the cognitive meaning in Spanish scientific prose of the 18th, 19th and early 20th centuries. *J. Pragmat.* 85, 155–167.
- Kotwica, D., 2017. From seeing to reporting: grammaticalization of evidentiality in Spanish constructions with ver (‘to see’). In: Marín-Arrese, J.I., et al. (Eds.), *Evidentiality and Modality in European Languages. Discourse-pragmatic Perspectives*. Peter Lang, Bern, pp. 87–109.
- Levshina, N., 2011. *Doe wat je niet laten kan: a Usage-based Analysis of Dutch Causative Construction* (Ph.D. dissertation). Katholieke Universiteit Leuven, Leuven.
- Maldonado, R., 1996. Concordancia y selección de sujeto: entre impersonales y pasivas. III Encuentro de Lingüística del Noroeste. Universidad de Sonora, pp. 121–146.
- Maldonado, R., 1999. Espacios mentales y la interpretación del se impersonal. In: Colombo, F. (Ed.), *El centro del la lingüística hispánica y la lengua espan~ola*. UNAM, México, pp. 205–227.
- Matlock, T., 1989. Metaphor and the grammaticalization of evidentials. *Proc. Annu. Meet. Berkeley Linguist. Soc.* 15, 215–225.
- Méndez, B., 2015. *Corpus Oral Juvenil del Espan~ol de Mallorca (COJEM)*. Lingüística en Red, Universidad de Alcalá.
- Real Academia Espan~ola (RAE), 2014. *Diccionario de la Real Academia Espan~ola*, 23ª ed. Espasa Libros, Madrid.
- Real Academia Espan~ola, Asociación de Academias de la Lengua Espan~ola (NGLE), 2009. *Nueva gramática de la lengua espan~ola*. Espasa Libros, Madrid.
- Rodríguez-Espin-eira, M.J., 2000. Percepción directa e indirecta en espan~ol. *Verba* 27, 33–85.
- Rodríguez-Rosique, S., 2015. Distance, evidentiality and counter-argumentation: concessive future in Spanish. *J. Pragmat.* 85, 181–199.
- Rossari, C., 2012. The evidential meaning of modal parentheticals. *J. Pragmat.* 44 (15), 2183–2193.
- Salvador, F. (Coord.), 2007. *El habla culta de granada: Materiales para su estudio*. Editorial Universidad de Granada, Granada.
- Squartini, M., 2008. Lexical vs. grammatical evidentiality in French and Italian. *Linguistics* 46 (5), 917–947.
- Sweetser, E., 1990. *From Etymology to Pragmatics: Metaphorical and Cultural Aspects of Semantic Structure*. Cambridge University Press.
- Traugott, E.C., Trousdale, G., 2013. *Constructionalization and Constructional Changes*. Oxford University Press.
- Viberg, A., 1983. The verbs of perception: a typological study. *Linguistics* 21, 123–162.

- Vogeleer, S., 1994. L'accès perceptuel à l'information: à propos des expressions un homme arrive/on voit arriver un homme. *Langue française* 102, 69–83.
- Wheeler, E.S., 2005. Multidimensional scaling for linguistics. In: Koehler, R., Altmann, G., Piotrowski, R.G. (Eds.), *Quantitative Linguistics: An International Handbook*. Mouton De Gruyter, Berl, pp. 548–553.
- Whitt, R.J., 2011. (Inter)Subjectivity and evidential perception verbs in English and German. *J. Pragmat.* 43, 347–360.
- Willems, D., 1983. Regarde voir: les verbes de perception visuelle et la complémentation verbale. In: Roegiest, E., Tasmowski, L. (Eds.), *Verbe et phrase dans les langues romanes. Mélanges offerts à Louis Mourin (RG 20)*, Gent, pp. 147–158.
- Willems, D., Defrancq, B., 2000. L'attribut de l'object et les verbes de perception. *Langue Française* 127, 6–20.
- Willett, T., 1988. A cross-linguistic survey of the grammaticalization of evidentiality. *Stud. Lang.* 12 (1), 51–97.