ON THE GRAMMATICALIZATION OF THE DEVERBAL EPISTEMIC PRAGMATIC MARKER *SABES*. A STUDY IN RECENT LANGUAGE CHANGE

RENATA ENGHELS[1]

Abstract. This article traces the recent evolution of the linguistic item *sabes* 'you know' in contemporary Spanish. An in-depth diachronic corpus analysis based on a corpus of four trend samples of oral data from the seventies, eighties, nineties and the 21st century, shows that the contemporary use of *sabes* is the result of a grammaticalization process implying a series of functional and formal changes. This process has considerably accelerated in the 21st century, and the data demonstrate a gradual shift from the primarily conceptual interpersonal meaning of *sabes* to a more pragmatic and discourse-oriented one. This functional evolution correlates with a higher degree of fossilization of the form, and a more varied positional profile.

Keywords: grammaticalization, epistemic pragmatic marker, Spanish, recent language change.

1. INTRODUCTION

Spoken (informal) language is characterized by the omnipresence of pragmatic markers shaping numerous discourse-related functional principles: (a) they can act upon the relationship between speaker and hearer and perform an 'interpersonal function', (b) they can serve to negotiate the speaker's own knowledge, beliefs, opinions and feelings, and therefore adopt a 'modal function', and (c) they can optimize the structure and course of discourse, performing a mainly textual function (Cuenca 2013)[2].

[1] Ghent University, Renata.Enghels@UGent.be.

[2] The very concept of 'marker' has been subject of a crossfire of diverging theories, and lacks a univocal terminology, definition and classification. In this paper, following Brinton (1996, 2008) and Aijmer (2013), among others, the notion of 'pragmatic marker' is used to refer to highly multifunctional linguistic elements which have a procedural or conceptual core meaning. The notion of pragmatic marker is preferred over alternative, also widespread, notions such as 'discourse marker' (Fraser 1999) or 'discourse particle' (Fischer 2006), because of its more explicit reference to the wide range of pragmatic functions an element like *sabes* can fulfill. As is shown in this paper, although *sabes* can adopt a 'text-monitoring' (and thus discourse-structuring) function (Erman 2001), it can also fulfill a wide range of interpersonal (Brinton 2008) and modal (Cuenca 2013) functions. This is why I prefer to recur to a notion that more explicitly accounts for this multifunctionality, both on the textual and interpersonal domain.

RRL, **LXIII**, *4*, p. 341–360, 2018

BDD-A28574 © 2018 Editura Academiei
Provided by Diacronia.ro for IP 216.73.217.142 (2026-11-14 23:09:52 UTC)

The lexical class of cognitive verbs constitutes a privileged source of 'epistemic' deverbal pragmatic markers, as has been shown by numerous studies on *I think* (Thompson and Mulac 1991) and its cognates in Spanish (Posio 2014) and in other languages (Dendale and Van Bogaert 2007). Indeed, the inherent semantics of cognitive verbs, including subjectivity, epistemicity and evidentiality, makes them particularly convenient as discourse strategies regulating interpersonal relationships between speakers (Nuyts 2001). Following the Matrix Clause hypothesis (Thompson and Mulac 1991), when used parenthetically, cognitive verbs no longer appear as main verbs in a transitive construction with a subordinate *that*-clause (or other complement), but take an independent position and gain scope over the entire sentence (e.g., *I think that there will be enough for all > There will be, I think, enough for all*). From a functional point of view, the cognitive verb forms no longer fulfill a predicate function with propositional meaning, but perform several discourse-related and interpersonal functions. They function as epistemic adverbs that modify the main clause, instead of being the nucleus of the main sentence themselves. As a consequence, they can generally be omitted without affecting the grammaticality of a clause. In (1a-b) the speaker no longer describes himself as being involved in an act of cognition, but instead refers to his or her own belief or involvement in discourse, and more particularly how, in his or her view, the information expressed by the host clause is to be interpreted (i.e. in (1a), the fact that they are poisonous, and in (1b), the fact that the interlocutor is still in possession of the speaker's book). These 'attitudinal meanings' arise as a result of a process of 'pragmatic strengthening'.

(1) a. They are, *I think*, poisonous. (Brinton 1996: 252)
b. Tere, y si te acuerdas tráeme. . . un libro que tienes <name> mío, *creo*, de Italo Calvino. ("Tere, and if you remember to bring me … one of my books that you have <name>, *I believe*, about Italo Calvino.") (Posio 2014: 10)

Similar analyses have been proposed for the – much less studied, though – second person parenthetical expressions (Dostie and de Sève 1999; House 2009). As their form suggests, these expressions are oriented toward the addressee, and essentially convey an interpersonal function. In both (2a-b) the epistemic pragmatic markers *you know* and *t'sais* call upon the attention of the interlocutor, trying to control his or her reaction and/or verify his or her comprehension of the message being transferred.

(2) a. At at least at the very smallest scale we can do something… but you're quite right *you know* in larger scale you know that's true *you know* because eh majority of eh […]. (House 2009: 180)
b. Ça fait que je commence à faire les affaires à ma manière, comme je veux, *t'sais*?
("That means that I start doing things my way, like I want it, *y'know*?)
(Dostie and de Sève 1999: 15)

Bearing in mind that, in general terms, pragmatic markers are found in all spoken languages (Ameka 1992), but that from a cross-linguistic viewpoint numerous pragmatic and formal differences have been described (Borreguero Zuloaga and Gómez Jordana 2015;

BDD-A28574 © 2018 Editura Academiei
Provided by Diacronia.ro for IP 216.73.217.142 (2026-11-14 23:09:52 UTC)

Fleischmann and Yaguello 2004), the conclusions on *you know* cannot be extended as such to cognate elements in other languages. As a consequence, a separate study needs to be set up for Spanish altogether. The present article aims to fill this gap in the literature by providing a formal and functional description of the Spanish pragmatic marker *sabes* ('you know'), as in example (3), a marker whose functioning has been poorly described so far (see references in sections 3 and 4 below).

(3) Ahora las compramos las pollas ya, pa poner, *¿sabes?,* ahora ya las gallinas que tenemos, las vamos matando, como vamos a traer nuevas. ("Now we are buying chickens, yeah, to lay eggs, *you know*, the chickens we have at the moment, we are killing them, because we are going to have new ones.") (Coser, 1997)

As is well known, the status of pragmatic marker is generally ascribed to linguistic elements which have gone through a process of grammaticalization. As a reminder, there is a vast debate on the very concept of grammaticalization, concerned with a narrow vs. wide interpretation of the base definition. This definition has been formulated by Hopper and Traugott (2003: 18) as "[t]he change whereby lexical items and constructions come in certain linguistic contexts to serve grammatical functions, and, once grammaticalized, continue to develop new grammatical functions". In its narrow interpretation, the definition only applies to changes from the lexicon to grammar; a broader definition concerns extensions toward discourse-related functions (Traugott 1995). This, of course, refers to the development of pragmatic markers from lexical elements, a linguistic change also described in terms of pragmaticalization (e.g., Dostie 2004; Erman and Kotsinas 1993). Indeed, the main reason why some authors prefer to speak of 'pragmaticalization' rather than 'grammaticalization' in the realm of pragmatic markers, is that the latter (in its narrowest definition) implies the idea of syntactic fixation and scope reduction, whereas the former implies positional mobility and expansion. But, as has been argued by Traugott (2003a), the terminological debate relates to how one defines 'grammar', and a wider conceptualization of grammar which goes beyond the morphosyntactic level, and includes pragmatic functions, makes the notion of 'pragmaticalization', to some extent, redundant. As a consequence, I agree with Company (2006) and Diewald (2011), among others, who use the term grammaticalization as a kind of umbrella term to refer to any change, independently of the classes involved or the direction of the change.

In large terms the parameters describing the process of grammaticalization and those defining the category of pragmatic markers coincide, and relate to (1) a process of semantic bleaching accompanied by 'pragmatic strengthening' (Sweetser 1988; Traugott 1988) of the form; (2) a fixation of the form or construction and reduction of its syntactic capacities; and (3) a widening of the scope and higher degree of autonomy of the form. As will be demonstrated in this article, these criteria have been applied differently in the discussion on the status of *sabes,* generating conflicting conclusions at times.

Indeed, whereas the historical reconstruction has been completed for *you know*, the status of *sabes* is still under debate. It was last referred to by Kluge (2011: 306) as follows: "el proceso de pragmaticalización de *sabes* se ha intensificado mucho en los últimos años y acaba de integrarse – o actualmente está integrándose – en el elenco de los marcadores del discurso del español peninsular." [the process of pragmaticalization of *sabes* has considerably accelerated over the last years, and it has recently become a member of – or is

BDD-A28574 © 2018 Editura Academiei
Provided by Diacronia.ro for IP 216.73.217.142 (2026-11-14 23:09:52 UTC)

actually becoming a member of – the inventory of discourse markers of peninsular Spanish]. This statement encapsulates the starting point of the present study. If it is true that the grammaticalization process of *sabes* as a pragmatic marker has only recently been completed or is still ongoing, it constitutes an excellent case study to analyze recent language change.

To recap, in line with the previously mentioned challenges and research questions, this case study aims to contribute to the study of pragmatic makers in at least three ways:

1. Most studies on pragmatic markers derived from cognitive verbs focus on 1st person expressions but the conclusions that come out of these studies cannot, as such, be transposed to 2nd person (much less studied) expressions. Indeed, from a functional perspective, 2nd person pragmatic markers convey a different basic meaning, and this might have an influence on their historical development and current use.
2. At a methodological level, the study illustrates how processes of recent language change can be studied in a corpus of spoken Spanish.
3. Although the article does not explicitly pursue a contrastive goal, it wants to contribute to a better understanding of whether, from a cross-linguistic viewpoint, pragmatic markers have followed similar or different paths of development. In particular, the results of previous studies on *you know* and cognate expressions in other languages can, in future research, be compared with the conclusions on *sabes* exposed here.

The outline of this paper is as follows. Section 2 outlines the methodology and argues in favor of studying the diachrony of the late 20th and early 21st centuries in order to get insight into the process of recent language change. Besides reviewing the literature on the grammaticalization process of epistemic markers, sections 3 and 4 provide the results of a diachronic corpus analysis, and zoom in on the functional and formal evolution of the epistemic marker, respectively. Section 5 concludes this article by formulating an answer to the question of whether *sabes* became entrenched in present-day Spanish as a pragmatic marker.

2. CORPUS AND METHOD

This article starts from the idea that the end of the 20th century and the beginning of the 21st century can be studied as locus of (recent) language change.[3] Moreover, the

[3] A key issue that has not been discussed yet is the evolution of *sabes* from a cognitive predicate with epistemic meaning to its pragmatic marker use. A description of the earlier developments of the pragmatic marker (in the 19th c. and early 20th c.) is provided in a previous study (Azofra Sierra and Enghels 2017). Despite of the data problem (no spontaneous oral conversational data are available for Spanish before the 70s decade, so the description of the earlier development was based on tokens of *sabes* in a selection of literary works, known to 'reflect' the current spontaneous language), that study has led to two main conclusions: (1) *sabes* was already used as a PM in earlier decades, but alternated frequently with that of formally more complex interpersonal epistemic expressions such as *como ya sabes* 'as you already know', *ya sabes tú* 'you already know', and expressions of *sabes* followed by a subordinate *that*-clause; (2) from a semantic viewpoint, its meaning was mostly interpersonal (thus, closer to the full lexical meaning of the verbal form). Brinton (2008) goes even further back in time, and examines the validity of the matrix clause hypothesis for the development of similar expressions in English. She concludes that the historical sources of *you know* are more varied (including adverbial and relative clauses), that the formal

BDD-A28574 © 2018 Editura Academiei
Provided by Diacronia.ro for IP 216.73.217.142 (2026-11-14 23:09:52 UTC)

analysis of recent (and current) linguistic changes (partly) solves the problem of finding representative historical data of interactional speech and talk (Traugott 2014: 78). These resources can facilitate a more accurate perception of linguistic change in spoken discourse, and allow for a more precise phasing of its evolution. Therefore, following Pons Bordería (2014), I resort to a 'micro-diachronic' analysis of the formal and functional evolution of *sabes* in four decades, starting with the period for which the first spoken corpus for peninsular Spanish is available, namely the seventies.

For this study, a specific database has been designed containing tokens from eight spoken corpora of peninsular Spanish.[4] However, given that short-term patterns of change can only been detected when differences in the corpus cannot be attributed to factors of influence other than the time variable, it is important to hold a good overview of the kind of data that is included. To this end, the main distinctive properties of the corpora are compared in Table 1.[5]

Corpus	decade	genre	generation
Habla culta de Madrid	70s	semi-directed interviews	all
CREA oral	80s, 90s	spontaneous conversation	no information
CORLEC	90s	spontaneous conversation	all
Val.Es.Co	80s, 90s	spontaneous conversation	all
Coser	90s, >2000	semi-directed interviews	Gen4
PRESEEA	>2000	semi-directed interviews	all
C-Oral-Rom	>2000	spontaneous conversation	all
COLAm	>2000	spontaneous conversation	Gen2

Table 1.
Overview and comparison of consulted corpora

Indeed, a comparison of the more detailed properties of these corpora shows, in fact, that the dataset is rather heterogeneous with respect to three parameters[6]:

(a) the corpora represent four micro-diachronic time periods (70s, 80s, 90s, > 2000). Note that for the 70s and 80s, I only have access to a restricted set of spoken language data, as opposed to the more widely documented 90s and 2000s decades. This restriction will be taken into account during the quantitative study.

(b) Five corpora (CREA oral, CORLEC, Val.Es.Co, C-Oral-Rom, and COLAm) contain spontaneous conversational data, whereas the other three (Habla Culta, Coser and PRESEEA) come from semi-directed interviews. The possible impact of this genre difference on our results will be tested throughout the article.

developments are more complex than has been assumed, and that the matrix clause hypothesis alone does not explain the observed data. A similar historical analysis of the sources of epistemic pragmatic markers in Spanish is left for future research.

[4] It is highly probable that other dialects of Spanish resort to other lexical elements to express similar functions, such as *¿cachái?* in Chilean. For practical reasons, however, this diatopic factor does not serve as a research parameter in this article.

[5] See Enghels, Vanderschueren and Bouzouita (2015) for a more detailed description of the features of each corpus.

[6] All corpora are equally distributed across male and female speakers, except for CREA Oral where no information is provided on the gender of the speakers.

BDD-A28574 © 2018 Editura Academiei
Provided by Diacronia.ro for IP 216.73.217.142 (2026-11-14 23:09:52 UTC)

(c) With regard to generational variation, most corpora are equally distributed among three generations (adolescents of Gen2=13-25, adults of Gen3=26-55, and elderly of Gen4= ≥56), except for Coser which is a corpus of exclusively elderly speakers (Gen4), and COLAm (Gen2) which is a young speak corpus. As a consequence, the sample of the 21st c. contains many data for Gen2. However, besides the fact that this predominance is partially cancelled out by the sufficient amount of tokens selected from other corpora, this sociolinguistic attribute can also have important explanatory power. Teenage talk (COLAm) is indeed said to play an important role in linguistic change (Zimmerman 2002), and can be a good indicator of how the use of a linguistic phenomenon has gradually spread over time. In an apparent time analysis (that is, the synchronic study of linguistic differences due to age, Cameron 2011), youth speech is considered to represent the most recent language stage.

The tokens were selected from these corpora through a lexical query for the string *sabes*. This generated a total sample of 5,304 tokens. These second-person instances of *saber* had then to be manually sorted in order to discard full lexical uses of the verb, as in (4).

(4) ¿Qué cara pones tú cuando *sabes* que tu mujer le gusta alguien, o que tiene un amante […]? ("How would you react when you got to *know* that your wife liked someone, or that she had a lover.") (CREA oral, 1996)

The final selection resulted in a sample of 2,379 cases of *sabes* (corresponding to 44.9% of all instances of *sabes*), distributed among the corpora as indicated in Table 2[7].

corpus	decade	# total tokens *sabes*	# *sabes* PM
Habla Culta	70s	90	27
CREA oral (1)	80s	178	29
CREA oral (2)	90s	1129	206
CORLEC	90s	140	64
Val.Es.Co (1)	80s	17	8
Val.Es.Co (2)	90s	102	40
Coser (1)	90s	657	333
Coser (2)	> 2000	747	309
C-Oral-Rom	> 2000	360	171
PRESEEA	> 2000	165	93
COLAm	> 2000	1719	1099
total		5,304	2,379

Table 2.
Quantitative overview of data extraction

These tokens were then sorted and annotated for a series of properties, all operationalizing the diagnostics of grammaticalization discussed in the literature (Company

[7] As opposed to English (e.g., Van Bogaert 2011), the syntax of Spanish cognitive verbs does not present *that*-deletion. As a consequence, the distinction between the full lexical use of *sabes* as a main verb and its use as a pragmatic marker could be established in a unequivocal manner. This was possible by simply observing the syntax of the verbal form, which included a subordinate clause (main verb use), or behaved parenthetically (pragmatic marker).

BDD-A28574 © 2018 Editura Academiei
Provided by Diacronia.ro for IP 216.73.217.142 (2026-11-14 23:09:52 UTC)

(2006); Brinton (1996), (2008); and other references cited throughout this article), and which are further specified in sections 3 and 4 below.

3. FUNCTIONAL EVOLUTION OF *SABES*

3.1. Previous accounts

Boyero Rodríguez (2002), Cuenca (2013), and López Serena and Borreguero Zuloaga (2010) do not hesitate to include *sabes* and other related second person deverbal markers (such as *entiendes* 'do you understand'*, crees* 'do you think') in the category of pragmatic markers. They are integrated by these authors into the categories of conversational and interactional markers that draw the interlocutor's attention, control his or her reactions, and verify his or her comprehension of the message. In his *Diccionario de Partículas*, Santos (2003) defines *sabes* as a "[p]alabra fática con que se busca la atención, participación y complicidad del interlocutor más que la averiguación de si sabe o no sabe el contenido pertinente" [phatic word one uses to search for the attention, participation and complicity of the interlocutor more than the verification of whether (s)he knows the relevant content or not]. A similar definition can be found in Fuentes' (2009) dictionary, where *sabes* is described as an interactive contact-establishing marker with different values: affirmative, continuative and emphatic. Thus based on its functional behavior, *sabes* has straightforwardly been defined in the literature as a highly grammaticalized and multifunctional marker.

3.2. Empirical verification

Also in the corpus, *sabes* shows a wide variety of functions, corresponding to varying degrees of semantic bleaching. Following the terminology and classification of Brinton (2008) and Cuenca (2013), these can be organized into three macro functions reflecting an (1) interpersonal, (2) modal, and (3) textual use of the marker[8]. Indeed, even though, as a

[8] An alternative option would have been to classify the tokens of *sabes* according to their mainly 'subjective' or 'intersubjective' meanings. These functions have indeed been related to semantic change, and have been defined respectively as processes whereby "meanings come explicitly to index and acknowledge [the speaker/writer]'s attitude toward [the addressee/reader] in the here and now of the speech event" (Traugott and Dasher 2002: 31), and "a mechanism whereby meanings become more centered on the addressee" (Traugott 2003b: 129). These meaning changes occur in specialized contexts "via the coding or conventionalization of invited inferences" (Brinton 2008: 53) (see also Traugott 2010). However, the form *sabes* in itself already expresses addressee-oriented (and intersubjective) meanings which are thus not the outcome of a process of change. Moreover, given the nature of our data, it is impossible to verify whether these meanings have occurred through conventionalization of invited inferences. As a consequence, the notions of (inter)subjectivity appear to be too theoretically loaded for the purpose of this study. Therefore, the more 'neutral' terminology of Brinton (1996, 2008), who distinguishes between interpersonal and textual meanings, is applied. However, the classification in this article is to some extent more fine-grained than Brinton's, given that it distinguishes between interpersonal and modal meanings, for reasons that will be made clear during the analysis.

BDD-A28574 © 2018 Editura Academiei
Provided by Diacronia.ro for IP 216.73.217.142 (2026-11-14 23:09:52 UTC)

consequence of its form, all functions of *sabes* center around the interpersonal one, other functional types can be distinguished which, to a variable extent, deviate from this prototypical meaning. Moreover, in order to avoid an intuition-based classification, a series of contextual parameters and paraphrase tests have been applied during the analysis.

In its first use, when *sabes* conveys an epistemic interpersonal function, it is an interactional marker mainly directed toward the interlocutor, with whom the speaker wants to share responsibility on the correctness of information being communicated. The interpersonal function coincides with the presence of contextual marks of interaction such as direct references to the addressee (e.g., *para ti* in (4b) or *¿Te interesa eso?* in (4d)), or affirmative or negative answer on behalf of the addressee (e.g., *una vena importante* (4a) or <assent> in (4b)). The shared knowledge between speaker and addressee can be general encyclopedic knowledge (e.g., about the human body in (4a)), or can refer to information provided by the deictic context (e.g., the size of the clothes demonstrated by the seller in (4b)), or the anaphoric context (e.g., the previous mention that the person could be an artist in (4c)). In these contexts *sabes* conveys a primary topical function, but it can also be used to introduce new information (Kluge 2011: 337), presented as possibly shared between the interlocutors (e.g., the reference to the price of the object in Portugal in (4d)). In general terms, these functions stay rather close to the etymological epistemic meaning of the verbal form *sabes,* and can therefore be defined as being the most remote from the semantically bleached and grammaticalized pole.

(4) a. ABU: […] pues le han le le ha dado una / embolia / y le ha cogido la parte / izquierda / y la vena esa / la vena aorta / o / la vena que tenemos // *¿sabes?* (“well (s)he had an embolism which had him (her) on the left side, and that important vein, the aorta vein, or the vein that we have… *you know*?”)
MAM: una vena importante // (“an important vein”) (C-Oral-Rom 2001)

b. <H2> Pues mira, eh... nos queda. (“well look, eh… we still have”)
<H1> No
<H2> Una cuarenta la que... la que más... es una cuarenta y de poca talla. (“a forty the one that… the one that most… it is a forty and a small one.”)
<H1> Sí (“yes”)
<H2> ¿*sabes*? Y camisas para ti pues sí (“*You see*? And shirts for you, well of course.”) (CORLEC 1991)

c. AMA: pero no sé si tiene que ver algo con / todo esto //no sé // igual es que es de arte // (“but I don’t know if it has something to do with all that. I don’t know. Perhaps it has something to do with art.”)
SAR: <seguro> // [certain]
AMA: ¿*sabes*? que igual es un escultor / o <algo> // (“*You know*, perhaps he’s a sculptor or something like that”)
SAR: <assent> (C-Oral-Rom 2001)

d. H1> Muy bonito... Que es un violetero; yo no sabía que era un violetero. Pero lo es. (“Very nice… It is a vase with violets; I didn’t know it was a vase with violets. But it is.”)
<H3> ¿Te interesa eso? (“Are you interested?”)
<H1> -Si [sic] es precioso, pap, para adornar! (“yes, it is gorgeous, dad, as a decoration”)

BDD-A28574 © 2018 Editura Academiei
Provided by Diacronia.ro for IP 216.73.217.142 (2026-11-14 23:09:52 UTC)

<H2> ¿*Sabes*? En Portugal valían... los ratoncitos como el vuestro y eso... Valían cada uno cinco mil <simult neo> pelas... ("*You know*? In Portugal they would cost … the little mice like yours and that … they would cost each one five thousand pesetas…")
<H3> -¡Ah, sí! ("Ah, indeed!") (CORLEC 1991)

The interpersonal macro function also groups cases in which *sabes* is used to control the interaction between speakers: (a) in a phatic function, it verifies whether the interlocutor is handling the flood of information (e.g., the fact that the speaker is an 'anti-alcoholic' although (s)he has a bottle of alcohol in his/her possession in (5a)); (b) when adopting a conative function, *sabes* asks for collaboration or an intervention on behalf of the interlocutor (e.g., the addressee has to eliminate certain racist recordings from the corpus in (5b)). These two uses share the same ground, namely *sabes* still has an interpersonal function directed toward the interlocutor (as can be derived from the interrogative (5a) and affirmative (5b) answers by the addressees). However, instead of referring to shared knowledge, it conveys shared discourse collaboration, thus pointing towards a higher level of semantic bleaching of the verb form.

(5) a. E: yo qué sé yo soy antialcohol / lo que pasa que la botella de Jotabé me la regalaron ("I, how would I know, I'm anti-alcoholic. What happened was that the bottle of Jotabé they gave it to me as a present")
L:¿antialcohol? ("Antialcoholic?")
E: no lo que pasa que me la regalaron la Jotabé me la regalaron ¿*sabes*? ("no what happened was that they gave me la Jotabé as a present they gave it to me as a present, *you know*?")
G: ¿antialcohol qué quiere decir? ("antialcoholic what does that mean?") (COLAm)

b. JOA: / hay que hay que censurarlo / *¿sabes?* no puede aparecer ("it needs to be censored, *you know*, it must not appear")
ALV: ¿que hay que censurarlo? ("it needs to be censored?")
JOA: sí sí sí ("yes yes yes") (C-Oral-Rom 2002)

Secondly, *sabes* is also often used by the speaker to reinforce the expression of his/her own emotions, opinions, or beliefs, conveying thus a modal function (Cuenca 2013). Moreover, *sabes* allows the speaker to reveal a personal interpretation of the message by emphasizing some part of the information (e.g., expressing the need for more time to do the job in (6a)), or by expressing (dis)content (e.g., with the explanation given by speaker E in (6b)). In the corpus, this function coincides with the presence of exclamation marks, recurrent references to the 1st person singular (as in (6a)), semantically loaded positive or negative words (as for instance *consuelo de tontos* in (6b)), or the use of superlative degree. Given that the speaker is not looking for an explicit reaction on behalf of the addressee, the interpersonal meaning is backgrounded. In these contexts, *sabes* thus partly loses the form-meaning connection (a second person form is used to refer to a first person's interpretation), its use therefore being more 'bleached' than in the former purely interpersonal cases.

(6) a. I: mmm no es que no me guste es que// eeh // si tuviese más tiempo // *¿sabes?*

BDD-A28574 © 2018 Editura Academiei
Provided by Diacronia.ro for IP 216.73.217.142 (2026-11-14 23:09:52 UTC)

pues sí me metería […]. (“mmm it is not that I don’t like it it is more that eeh if I had more time *you know*”) (PRESEEA 2001)

b. E: si no fuera bueno no lo habrían copiado (“If it weren’t good they wouldn’t have copied it”)
I: sí / exacto me dice porque si es un es un trabajo bueno por eso te lo han copiado / si no fuese bueno no te lo hubiesen copiado y yo sí pero eso es consuelo de tontos ¿*sabes*? (“indeed, (s)he tells me it is a good work therefore they have copied you, if it weren’t good they wouldn’t have copied your work, and I was like okay but that is a ridiculous comfort *you know*”) (PRESEEA 2001)

The third macro function comprises cases of textual *sabes*, in which the marker is used as a discourse connector coordinating the course of the conversation and maintaining coherence. In the corpus, the textual function coincides with the following diagnostics: presence of repeated words or parts of the utterance, reformulation or specification of parts of an utterance (e.g., 7a, 7b), the use of vague words, or repeated pauses (e.g., 7c, 7d)). In concrete, the marker *sabes* allows the speaker to hint at the relevance of upcoming information, to introduce a new discourse unit that specifies or explains a previous one (7a), or to reorient the aforementioned information (7b). It can also be used as a gap-filler which provides the speaker with some extra time to look for words, or to plan the upcoming message, avoiding any silences (Cortes Rodríguez 1999; Molina 2005). In example (7c), speaker *I* is clearly looking for the right words, (s)he hesitates (indicated by <vacilación> in the transcription) and the phrase is characterized by a distorted syntax. It goes without saying that here the marker loses most of the connection with the semantic meaning of its form, and is further bleached. It can, in fact, easily be replaced by another ‘filler’ such as *uhm*. Finally, some cases are observed in which *sabes* merely functions as an affirmative particle, repeating a previously given positive answer (7d).

(7) a. MALCC2G03: yo nunca tengo cuantas veces pienso cuando tengo los libros *sabes* cuando tengo el libro entero digo voy a sacar la mitad y lo metes *sabes* en el altavoz […] (“I never have how many times do I think when I have the books *you know* when I have the entire book I’m going to take half of it and you put it *you know* on speaker”) (COLAm)

b. MAESB2J01: es que joder cuando yo cuando yo me voy de campamento no es que *sabes*/ igual cogemos un autobús que sale a las doce de la noche y llegas allí a las cinco […] (“It is that damn I when I go camping it is not that *you know* we just take a bus that leaves at 12 o’clock at night and you get there at 5”) (COLAm)

c. I: nada pues dejen / vayan dejando todo encima de le <vacilación/> de la cinta esta donde se dejan los productos (“fine so leave it all on top of the <hesitation> the tape the one where the products are left on”)
E: sí (“yes”)
I: para que te los cobren / nos dijeron / m <vacilación/> *¿sabes?* dijeron que fuésemos dejando todas las cosas de valor (“in order for them to charge you they told us <hesitation> *you know* they told us that we had to leave behind all valuable things”) (PRESEEA 2001)

BDD-A28574 © 2018 Editura Academiei
Provided by Diacronia.ro for IP 216.73.217.142 (2026-11-14 23:09:52 UTC)

d. I1: Pues ya la casca cae abajo, que es lo | la piel de la uva pa que lo entendamos, [HS:E1 [Asent]] cae abajo de la tenaja y el… la… y el vino queda a flote por encima. ("So the skin falls down, that is the skin of the grape for the sake of clarity, falls off the cutter and the… the.. and the wine stays floating on top")
E1: Ya, ya. [Asent] ("yes, yes")
I1: ¿*sabes*? ("*you know*")
E1: Y ¿qué es vino tinto el que tenéis o...? ("and is it red wine that you have or…") (Coser 1995)

The panorama that has just been drawn clearly supports the idea that, on a semantic and pragmatic basis, *sabes* can be defined as a grammaticalized marker: apart from its clear epistemic and interpersonal uses, it conveys a series of extended functions, both modal and discourse-organizing ones, and it can even be used as a 'meaningless' (in the semantic meaning of the term, not the pragmatic one) gap-filler. As a consequence, the functional diagnostic of grammaticalization is confirmed by the data observed in the corpus. However, the question arises whether the distribution of these three macro functions remains constant throughout the corpus, or whether a functional evolution – or even a shift – can be observed in the last decades.

To this end, Table 3 details the distribution of the three macro functions across the four time periods under scrutiny.

decade	corpus	interpersonal		modal		textual		total	
		#	%	#	%	#	%	#	%
70s		17	**63**	10	37	0	0	27	100
	HC	17	**63**	10	37	0	0	27	100
80s		22	59.5	14	37.8	1	2.7	37	100
	CREA	16	55.2	12	41.4	1	3.4	29	100
	VSC	6	75	2	25	0	0	8	100
90s		311	48.4	198	30.8	134	**20.8**	643	100
	COS	189	56.8	75	22.5	69	**20.7**	333	100
	CREA	65	31.5	99	48.1	42	**20.4**	206	100
	VSC	31	77.5	7	17.5	2	**5**	40	100
	CRLC	26	40.6	17	26.6	21	**32.8**	64	100
>2000		490	29.3	682	40.8	500	**29.9**	1672	100
	COS	171	55.3	100	32.4	38	**12.3**	309	100
	PRES	29	31.2	32	34.4	32	**34.4**	93	100
	C-O-R	39	22.8	55	32.2	77	**45**	171	100
	COLAm	251	22.8	495	45	353	**32.1**	1099	100

Table 3.
Functional evolution of sabes: *macro functions*

In statistical terms, a significant correlation can be observed between the period of occurrence of *sabes* and its semantic-pragmatic profile (meaning the quantitative distribution and prevalence of its different functions) (X^2= 97.06, $p < 0.001$; Cramer's V= 0.01438, yet pointing toward a weak correlation)[9]. Indeed, when we compare the use of

[9] The Chi-square test has been calculated on the basis of the means per decade.

BDD-A28574 © 2018 Editura Academiei
Provided by Diacronia.ro for IP 216.73.217.142 (2026-11-14 23:09:52 UTC)

sabes in the seventies corpus with that in the contemporary corpus, it immediately becomes clear that an important shift has taken place. Albeit limited in extension, the seventies corpus displays a majority of cases (n=17, 63%) in which *sabes* fulfills an interpersonal function, thus remaining closest to its original hearer-oriented meaning (see examples 4 and 5 above). Also, in a non-negligible number of examples (n=10, 37%), *sabes* is used by the speaker to strengthen the modal and argumentative power of the message (see examples 6 above). Similar proportions can be found in the eighties corpus, which, however, also attests an isolated textual, and more particularly gap-filling, use of *sabes* (8)[10].

(8) ¿Y tú te acuestas aún con tu marido? ("And do you still sleep with your husband")
No, bueno, alguna vez, *sabes*, es que, no sé, se pone pesadito y así... ("no, well, sometimes, *you know*, it's that, I don't know, he's becoming rather heavy and so…") (CREA oral 1989)

In the nineties corpus, and especially the 21st-century corpus, the frequency of the interpersonal uses decline, in favor of a remarkable development of the discourse-related uses of the marker. However, two corpora do not confirm this general trend, namely Val.Es.Co and Coser. In the former case, this could perhaps be due to a dialectal factor, the corpus being representative of Valencian Spanish, as opposed to the other corpora which mainly include data from Madrid (e.g. PRESEEA, COLAm), or from a wider range of varieties (e.g. CREA Oral, CORLEC). The Coser corpus also constitutes a particular case, given that its speakers all pertain to the fourth generation category. In an apparent time analysis, these speakers represent an earlier language stage (Cameron 2011), which could then explain the higher degree of conformity to the interpersonal function in this corpus.

Despite these individual differences between the corpora, it appears that, on average, in present-day Spanish almost one third (n=500, 29.9%) of the uses of *sabes* are triggered by a discourse-internal organizational and textual need (see examples 7 above). The use of the personal modal function seems to have remained more or less stable over time, and ranges between approximately 30% and 40%.

As an interim conclusion, it can thus be stated that the empirical data confirm the gradual and ongoing bleaching process in which the marker *sabes* is involved: the verb form is increasingly used to fulfill more various textual and modal functions. We will now return to the discussion on the pragmatic marker status of *sabes*, because what has caused more debate in the literature is the question to what extent it meets the formal parameters of grammaticalization.

4. FORMAL EVOLUTION: FOSSILIZATION AND WIDENING OF SCOPE

4.1. Previous accounts

Thompson and Mulac (1991: 324) initially tested the behavior of epistemic parentheticals against the narrow definition of grammaticalization, which led to the conclusion that they do not constitute "a 'textbook case' study in grammatization": they do not satisfy the criteria of syntactic fixation nor narrowing of scope. However, within a broader definition of grammaticalization, including discourse-related phenomena, these criteria have been considered as positive indicators of the phenomenon under consideration (see section 1).

[10] Example (8) dates from 1989, and can be considered as a transition to the nineties. Of course, the temporal axis should also to be seen as a continuum rather than a categorical variable.

BDD-A28574 © 2018 Editura Academiei
Provided by Diacronia.ro for IP 216.73.217.142 (2026-11-14 23:09:52 UTC)

This then led authors such as House (2009), Macaulay (2002), and Vincent et al. (2009) to conclude that *you know* has been fully grammaticalized into a pragmatic marker.

For Spanish *sabes*, Martín Zorraquino and Portolés (1999: 4186) accept its grammaticalized status on the basis of its pragmatic meaning. Given that it marks relationships between participants in a communication act, it is classified together with other second person verbal forms such as *ves/verás* ('you (will) see'), *entiendes* ('do you understand'), *escucha* ('listen') or *fíjate* ('think about it') as 'illuminators of variability', meaning that the speaker changes his or her behavior in function of his or her relationship with the interlocutor. However, these second person markers are said not to have completely grammaticalized because they do not fulfil Hopper's (1991: 30) 'freezing' condition. This principle states that from a morphological point of view, grammaticalized forms are invariable and do not inflect for gender, number, time or aspect. The particles cited by Martín Zorraquino and Portolés (1999: 4186) deviate from this condition as they

(a) combine with different complement types (*me entiendes, ya ves, tú ya me entiendes, fíjate bien, ya sabes*, etc.);
(b) admit certain temporal alternations (*ves, verás*);
(c) can be negated (*ves/¿no ves?, sabes/¿no sabes?, entiendes /¿no me entiendes*?);
(d) and can be used with an assertive or interrogative modality (*ves/¿ves?, sabes/¿sabes?, me entiendes, ¿me entiendes?*).

Criterion (b) does not hold for *sabes*, which is indistinctively used in the present indicative form. It constitutes a morphological minimal pair only with the third person singular courtesy form *sabe*, which is very rare. Its formal meaning contrasts sharply with the informal contexts in which these markers are typically used (9).

(9) <H5>Mire, *sabe*, es que es un problema porque ya me la han arreglado pero... se ha vuelto a estropear. ("Look, *you know*, that is a problem because they have given it to me as a present but… now it is broken again") (CORLEC 1991)

Criterion (d) is hard to verify on the basis of transcriptions and would require a thorough prosodic analysis of oral data, which, unfortunately, falls beyond the scope of this article.

This leaves us with counterarguments (a) and (c), which actually both come down to the same question, namely whether *sabes* has formally fossilized as a pragmatic marker, or whether the form still alternates with morphologically related epistemic expressions.

4.2. Fossilization: Empirical verification

In order to answer the above-mentioned question, the relative frequency of use of *sabes* needs to be compared to the one of other related expressions in the corpus, such as *ya sabes*, *tú sabes*, *no sabes*, *sabes qué*, *como tú sabes*, *como sabes*, *no lo sabes*, *lo sabes*, *ya lo sabes*, *tú no sabes*, *tú lo sabes*, *sabes tú*, *ya sabes tú*, *sabes una cosa*, *tú sabes* + NP, *tú ya sabes*, *lo sabes tú*, *como sabes tú*, *ya tú sabes*, *te lo sabes*, *no sabes tú*, *no lo sabes tú*, *tú ya sabes algo*, *ya sabe usted*, *sabes ya*, etc. These expressions fulfill similar functions as have been previously described for *sabes* (supra section 3), ranging from a purely interpersonal (10a) to a clearly textual (affirmative) function (10c), over a modal emphasizing one (10b)[11].

[11] Remind that the outset of this article has been defined by a semasiological research question, namely the recent evolution of *sabes* as a pragmatic marker, in its turn, inspired by Kluge's hypothesis. A comparison of *sabes* with the behavior of other morphologically and functionally related expressions would certainly constitute an interesting topic, but would lead us too far for this study.

BDD-A28574 © 2018 Editura Academiei
Provided by Diacronia.ro for IP 216.73.217.142 (2026-11-14 23:09:52 UTC)

(10) a. Como árbitro un australiano nos dio muy bien por el pobrecito mi balón, todos le dan patadón. Pero *tú sabes*, Alberto, y te paso otra vez el balón. ("As a referee an Australian gave us good money for my poor ball, they all give him hard kicks. But *you know*, Alberto, I pass the ball to you once again") (CREA Oral, 1986)

b. I5: Como se nota que no tienes nuera. ("One can see very well that you don't have a daughter in law")
I1: Y yo sa-, y yo sabes que he sío nuera, y mi suegra, me ha querío más que a sus hijas, *lo sabes*. ("And I you know I have been a daughter in law, and my mother in law loved me more than her own daughters, *you know that*.")

c. JOA: chaval / pero que me voy a hacer un / profesional ya // ("guy, but I'm going to be a professional")
ALV: sí sí // <pues *ya sabes* / ¡hala!> // ("yes yes well *you already know*, let's go") (C-Oral-Rom 2002)

The data presented in Table 4 show a steady increase in the relative frequency of use of the bare form *sabes*, from 37.5% in the seventies corpus to 79.5% in the most recent sample. Note that each individual corpus confirms this tendency. In other words, although the form still alternates with other morphologically related epistemic expressions, the use of the fossilized form has become more widespread in present-day Spanish in comparison with earlier time periods.

decade	corpus	# sabes PM	# related expressions	total	relative freq. *sabes*
70s		27	45	72	37.5
	HC	27	45	72	37.5
80s		37	43	80	46.3
	CREA	29	32	61	47.5
	VSC	8	11	19	42.1
90s		643	378	1021	63
	COS	333	183	516	64.5
	CREA	206	153	359	57.4
	VSC	40	23	63	63.5
	CRLC	64	19	83	77.1
>2000		1672	430	2102	79.5
	COS	309	194	503	61.4
	PRES	93	25	118	78.8
	C-O-R	171	44	215	79.5
	COLAm	1099	167	1266	86.8

Table 4.
Absolute and relative frequencies of sabes *vs. related epistemic expressions*

As additional formal proof of grammaticalization, pragmatic markers are said to often be phonetically reduced (Brinton 1996). In the corpus, several cases of phonetic reduction have been attested, all in the present-day teenage talk corpus COLAm[12].

[12] This could mean that the phonetic reduction of *sabes* is a very recent phenomenon, or it could be explained by different transcription conventions of the corpora. In order to clarify this issue, a systematic prosodic analysis of the audios is required.

BDD-A28574 © 2018 Editura Academiei
Provided by Diacronia.ro for IP 216.73.217.142 (2026-11-14 23:09:52 UTC)

(11) MALCC2J03: eeh. Sub cuarenta y nueve o o algo de eso *saes*/ sii música surfera de ésta <risa/> ("eeh … Sub forty-nine or or something like that *you know* / yes that surf music <laugh>") (COLAm)

4.3. Positional variation: Empirical verification

A second parameter to take into account concerning the formal behavior of *sabes* is its position, both with respect to the turns or interventions of different speakers, and within a particular utterance itself.[13] If it has acquired the full status of pragmatic marker in present-day Spanish, *sabes* is indeed expected to display a high degree of positional variation, and we will observe an increasing tendency toward more positional mobility over the four time periods (e.g., Company 2006).

At the higher level of the different turns which make up a conversation, four possible positions can be distinguished, as detailed by Table 5: turn-initial (e.g., 4b *supra*), turn-internal (5b), turn-final (4a), and an independent position (7d).

decade	corpus	turn-initial		turn-internal		turn-final		independent		total	
		#	%	#	%	#	%	#	%	#	%
70s		1	3.7	16	59.3	10	37	0	0	27	100
	HC	1	3.7	16	59.3	10	37	0	0	27	100
80s		2	5.4	19	51.4	16	43.2	0	0	37	100
	CREA	0	0	17	58.6	12	41.4	0	0	29	100
	VSC	2	25	2	25	4	50	0	0	8	100
90s		26	4.1	432	68.4	168	26.6	6	0.9	632	100
	COS	12	3.6	245	73.6	71	21.3	5	1.5	333	100
	CREA	0	0	132	67.7	63	32.3	0	0	195	100
	VSC	6	15	17	42.5	17	42.5	0	0	40	100
	CRLC	8	12.5	38	59.4	17	26.6	1	1.6	64	100
>2000		111	6.6	1125	67.3	411	24.6	25	1.5	1672	100
	COS	16	5.2	239	77.3	50	16.2	4	1.3	309	100
	PRES	14	15.1	50	53.8	24	25.8	5	5.4	93	100
	C-O-R	21	12.3	115	67.3	32	18.7	3	1.8	171	100
	COLAm	60	5.5	721	65.6	305	27.8	3	1.2	1099	100

Table 5.
Formal evolution of sabes*: position in turn*[14]

[13] The various ways in which the position of markers in discourse can be defined have given rise to many different theories. For the present analysis, I rely on an economized version of the Val.Es.Co model of discourse segmentation (e.g., Briz and Bordería 2010). This model organizes the flow of talk into well-defined discourse units at different levels, namely, the intervention or turn as a whole, the act and subact. With regard to those discourse units, a pragmatic marker can be placed at the beginning, in the middle or at the end; or, it can appear in an independent position. In this study, I will not distinguish between the levels of the act and subact, but only between the turn, and the lower level of turn-internal positions.

[14] Unfortunately, the high number of low cell frequencies does not allow us to perform a Chi-square test. The same problem goes for Table 6. Note that the total number of tokens is reduced to 2,368 because some ambiguous cases (mostly from the CREA corpus) were excluded.

BDD-A28574 © 2018 Editura Academiei
Provided by Diacronia.ro for IP 216.73.217.142 (2026-11-14 23:09:52 UTC)

All periods show a tremendous prevalence of the turn-internal position, whose frequency has even increased over time. In contrast, the turn-final position seems to become more reduced. The nineties and 21st-century corpora also increasingly attest cases in which *sabes* in itself constitutes an independent turn[15]. It therefore seems fair to say that a shift can be observed in the positional behavior of *sabes*, slowly moving away from its more 'interpersonal' turn-final position (explicitly addressing the interlocutor and leaving him the opportunity to (re)act), and gradually taking up more discourse-internal, independent positions. This evolution confirms the tendencies observed in Beeching and Detges (2014) and in other works that examine this form-function relationship in more detail, such as Traugott (2012).

Given these global figures, it is necessary to have a closer look at the distribution of *sabes* in turn-internal position, presented in Table 6. At the turn-internal level, *sabes* also displays very complex behavior as it occupies five different positions: it can appear at the beginning (5b) or the end (11) of an utterance; it can be placed in between two utterances as a kind of connector (e.g., in 6a *sabes* connects the protasis and apodosis of a conditional sentence); and, it can be placed in an utterance-medial position, distorting its syntax to a lower (7a) or higher (7b) degree.

decade	corpus	initial		final		between		medial		interruption		total	
		#	%	#	%	#	%	#	%	#	%	#	%
70s		0	0	16	100	0	0	0	0	0	0	16	100
	HC	0	0	16	100	0	0	0	0	0	0	16	100
80s		0	0	14	73.7	2	10.5	3	15.8	0	0	19	100
	CREA	0	0	13	76.5	1	5.9	3	17.6	0	0	17	100
	VSC	0	0	1	50	1	50	0	0	0	0	2	100
90s		8	1.9	312	72.2	22	5.1	54	12.5	36	8.3	432	100
	COS	0	0	182	74.3	16	6.5	31	12.7	16	6.5	245	100
	CREA	6	4.5	91	68.9	4	3	16	12.2	15	11.4	132	100
	VSC	1	5.9	12	70.6	2	11.8	0	0	2	11.8	17	100
	CRLC	1	2.6	27	71.1	0	0	7	18.4	3	7.9	38	100
>2000		60	5.3	627	55.7	103	9.1	119	10.6	217	19.3	1126	100
	COS	3	1.3	154	64.4	45	18.8	29	12.2	8	3.3	239	100
	PRES	6	12	28	56	5	10	3	6	8	16	50	100
	C-O-R	9	7.8	46	40	13	11.3	13	11.3	34	29.6	115	100
	COLAm	41	5.8	399	55.3	40	5.5	74	10.3	167	23.1	721	100

Table 6.
Formal evolution of sabes: *turn-internal position*

[15] An attentive reviewer noticed that the increasing use of *sabes* in independent position could be due to different transcription conventions of the corpora. However, on a closer inspection of the relevant cases as well as the transcription conventions of the individual corpora, this factor does not appear to have had an influence on the classification of the position of *sabes*. Moreover, note that all four 21st c. corpora present some cases of independent turn, what shows that it constitutes a rather spread-out phenomenon, not restricted to a particular corpus.

BDD-A28574 © 2018 Editura Academiei
Provided by Diacronia.ro for IP 216.73.217.142 (2026-11-14 23:09:52 UTC)

Three facts cannot be left unnoticed. First, later time periods display a higher degree of positional variation of *sabes* within the middle field (one position taken in the seventies corpus vs. three in the eighties vs. five in the two most recent time periods). Second, similar to what was observed at the level of the turn-alternating structure, the utterance final position has gradually declined from being the only option in the seventies corpus, to 55.7% in present-day Spanish. Finally, the corpora gradually display more cases in which *sabes* is used to interrupt an utterance and start a new one (almost 20% in the 21st-century sample). In conclusion, the diachronically increasing tendency of more positional variation of *sabes* can be confirmed.

5. CONCLUSION

The starting point of this study was the observation of different opinions in the literature on the degree of grammaticalization reached by *sabes* and, by extension, other second person deverbal markers. At stake was the hypothesis of Kluge (2011), claiming that this linguistic item has been integrated into the class of Spanish markers very recently. In order to examine this hypothesis, the recent evolution of *sabes* was reconstructed in detail, and the stage of grammaticalization reached by *sabes* was carefully examined through a recent diachronic corpus analysis, quantitatively and qualitatively tracking its behavior during four contemporary time periods.

It was found that the grammaticalization process of *sabes* indeed seems to have accelerated in the last decade of the 20th century and at the beginning of the 21st century. This grammaticalization goes hand in hand with some gradual changes situated at both the functional and the formal level. First, the corpus study has shown that, in recent decades, there has been a gradual functional shift in the use of *sabes* from the primarily interpersonal domain to a mainly discourse-oriented and textual one. The data from the nineties and 21st-century corpus show that *sabes* has increasingly acquired a more procedural function (Wilson 2011), and that, more than calling upon the knowledge of the interlocutor, *sabes* serves as a discourse-organizing device, and as a kind of organizational frame marker. Secondly, this 'discoursivization' process is formally marked by its changing positional profile, characterized by increasing complexity. The particle moves away from the prototypical interpersonal right-periphery position, and at the same time displays a more diversified range both in turn-medial and utterance-medial positions.

However, since all macro functions are frequently used today, *sabes* is still highly multifunctional in present-day Spanish, both at the semantic-pragmatic and morphosyntactic level. Its use also alternates with related but morphosyntactically more complex epistemic expressions. However, it is clear that this phenomenon of 'differentiation' does not conflict with the evolution of individual linguistic items, which as the above analysis has shown, points at the fully acquired pragmatic marker status of *sabes*. The way in which *sabes* is related to these similar expressions within a more complex network of epistemic parentheticals is subject for future research.

BDD-A28574 © 2018 Editura Academiei
Provided by Diacronia.ro for IP 216.73.217.142 (2026-11-14 23:09:52 UTC)

REFERENCES

Aijmer, K., 2013, *Understanding Pragmatic Markers. A Variational Pragmatic Approach*, Edinburgh, Edinburgh University Press.

Ameka, F., 1992, "Interjections: the universal yet neglected part of speech", *Journal of Pragmatics*, 18, 101–118.

Azofra Sierra, E., R. Enghels, 2017, "El proceso de gramaticalización del marcador epistémico deverbal *sabes*", *Iberoromania*, 85, 1–25.

Beeching, K., D. Ulrich, 2014, "Introduction", in: K. Beeching, U. Detges (eds), *Discourse functions at the left and right periphery: crosslinguistic investigations of language use and language change,* Leiden: Brill, 1–23.

Borreguero Zuloaga, M., S. Gómez Jordana, 2015, *Marcadores discursivos en las lenguas románicas. Un enfoque contrastivo*, Limoges, Lambert Lucas.

Boyero Rodríguez, M.J., 2002, *Los marcadores conversacionales que intervienen en el desarrollo del diálogo*, Salamanca, Publicaciones Universidad Pontificia.

Brinton, Laurel J., 1996, *Pragmatic Markers in English: Grammaticalization and Discourse Functions*. Berlin/New York, Mouton de Gruyter.

Brinton Laurel J., 2008, *The Comment Clause in English. Syntactic Origins and Pragmatic Development*, Cambridge, Cambridge University Press.

Briz Gómez, A., S. Pons Bordería, 2010, "Unidades, marcadores discursivos y posición", in: Ó. Loureda, E. Acín (eds), *Los estudios de marcadores del discurso en español, hoy*, Madrid, Arco Libros, 327–358.

Cameron, R., 2011, "Age, aging and sociolinguistics", in: M. Díaz Campos (ed.), *The Handbook of Hispanic Sociolinguistics*, Oxford, Willey-Blackwell, 207–229.

Company, C., 2006, "Zero in syntax, ten in pragmatics", in: A. Athanasiodou, C. Canakis, B. Cornillie (eds), *Subjectification: various paths to subjectivity*, Berlin/New York: Mouton de Gruyter, 375–398.

Cortés Rodríguez, L., 1999, *Sobre conectores, expletivos y muletillas en el español hablado*, Málaga, Editorial Librería Ágora.

Cuenca, M.J., 2013, "The fuzzy boundaries between discourse marking and modal marking", in: L. Degand, B. Cornillie, P. Pietrandrea (eds), *Discourse markers and modal particles. Categorization and description*, Amsterdam, John Benjamins, 191–216.

Dendale, P., J. Van Bogaert, 2007, "A semantic description of French lexical evidential markers and the classification of evidentials", *Rivista di linguistica*, 19, 1, 65–89.

Diewald, G., 2011, "Pragmaticalization (defined) as grammaticalization of discourse functions", *Linguistics*, 49, 365–390.

Dostie, G., 2004, *Pragmaticalisation et marqueurs discursifs: analyse sémantique et traitement lexicographique*, Bruxelles, De Boeck Duculot.

Dostie, G., S. de Sève, 1999, "Du savoir à la collaboration. Étude pragma-sémantique et traitement lexicographique de *t'sais*", *Revue de sémantique et pragmatique*, 5, 11–35.

Enghels, R., C. Vanderschueren, M. Bouzouita, 2015, "Panorama de los corpus y textos del español peninsular contemporáneo", in: E. Roegiest, M. Iliescu (eds), *Manuel des anthologies, corpus et textes romans*, 7, Berlin/New York, Mouton de Gruyter, 147–170.

Erman, B., 2001, "Pragmatic markers revisited with a focus on *you know* in adult and adolescent talk", *Journal of Pragmatics*, 33, 1337–1359.

Erman, B., U.-B. Kotsinas, 1993, "Pragmaticalization: The case of *ba'* and *you know*", *Studier i Modernspråkvetenskap*, 10, 76–93.

Fischer, K. (ed.), 2006, *Approaches to Discourse Particles*, Amsterdam, Elsevier.

Fleischmann, S., M. Yaguello, 2004, "Discourse markers across languages. Evidence from English and French", in: C. Lynn Moder, A. Martinovic-Zic (eds), *Discourse Across Languages and Cultures*, Amsterdam, John Benjamins, 129–147.

BDD-A28574 © 2018 Editura Academiei
Provided by Diacronia.ro for IP 216.73.217.142 (2026-11-14 23:09:52 UTC)

Fraser, B., 1999, "What are discourse markers?", *Journal of Pragmatics*, 31, 931–952.
Fuentes Rodríguez, C., 2009, *Diccionario de conectores y operadores del español*, Madrid, Arco Libros.
Hopper, Paul J., 1991, "On some principles of Grammaticization", in: E.C. Traugott, B. Heine (eds), *Approaches to grammaticalization*, Amsterdam, John Benjamins, 17–36.
Hopper, P.J., E. C. Traugott, 2003, *Grammaticalization*, Cambridge, Cambridge University Press.
House, J., 2009, "Subjectivity in English as Lingua Franca discourse: The case of *you know*", *Intercultural Pragmatics*, 6, 2, 171–193.
Kluge, B., 2011, "Camino de un marcador del discurso: una comparación del español *¿sabes?* con el francés *tu sais* y el inglés *you know*", in: H. Aschenberg, Ó. Loureda (eds), *Marcadores del discurso: de la descripción a la definición*, Madrid/Frankfurt, Iberoamericana/Vervuert, 305–341.
López Serena, A., M. Borreguero Zuloaga, 2010, "Los marcadores del discurso y la variación lengua hablada vs. lengua escrita", in: Ó. Loureda, E. Acín (eds), *Los estudios sobre marcadores del discurso en español, hoy*, Madrid, Arco Libros, 415–495.
Macaulay, Ronald. 2002. "You know, it depends." *Journal of Pragmatics*, 34, 749–767.
Martín Zorraquino, M.A., J. Portolés Lázaro, 1999, "Los marcadores del discurso", in: I. Bosque, V. Demonte (eds), *Gramática descriptiva de la lengua española*, vol. III, Madrid, Espasa Calpe, 4051–4214.
Molina Martos, I., 2005, "La moda del ¿*sabes*? en Madrid: un análisis sociolingüístico", in: *Lingüística y literatura. Homenaje a Antonio Quilis*, Madrid, CSIC/UNED/Universidad de Valladolid, 1025–1036.
Nuyts, J., 2001, *Epistemic modality, language and conceptualization: a cognitive pragmatic perspective*, Amsterdam, Benjamins.
Posio, P., 2014, "Subject expression in grammaticalizing constructions. The case of *creo* and *acho* 'I think' in Spanish and Portuguese", *Journal of Pragmatics*, 63, 5–18.
Pons Bordería, S., 2014, "El siglo XX como diacronía: intuición y comprobación en el caso de *o sea*", *RILCE*, 30, 3, 985–1016.
Santos Río, L., 2003, *Diccionario de partículas*, Salamanca, Luso-española de ediciones.
Sweetser, E., 1988, "Grammaticalization and semantic bleaching", *Berkeley Linguistics Society*, 14, 389–405.
Thompson, S.A., A. Mulac, 1991, "The discourse conditions for the use of the complementizer *that* in conversational English", *Journal of pragmatics*, 15, 237–251.
Traugott, E.C., 1988, "Pragmatic strengthening and grammaticalization", in: S. Axmaker, A. Jaisser, H. Singmaster (eds), *Proceedings of the Fourteenth Annual Meeting of the Berkeley Linguistics Society*, Berkeley, Berkeley Linguistics Society, 406–416.
Traugott, E.C., 1995, "The role of discourse markers in a theory of grammaticalization", paper presented at the 12th International Conference on Historical Linguistics, Manchester, UK, August 1995. Available at www.stanford.edu/~traugott/ect-papersonline.html
Traugott, E.C., 2003a, "Constructions in grammaticalization", in: B.D. Joseph, R. D. Janda (eds), *The handbook of historical linguistics*, Oxford, Blackwell, 624–647.
Traugott, E.C., 2003b, "From subjectification to intersubjectification", in: R. Hickey (ed.), *Motives for Language Change*, Cambridge, Cambridge University Press, 124–139.
Traugott, E.C., 2010, "(Inter)subjectivity and (inter)subjectification: A reassessment", in: K. Davidse, L. Vandelanotte, H. Cuyckens (eds), *Subjectification, intersubjectification and grammaticalization*, Berlin, Mouton de Gruyter, 29–71.
Traugott, E.C., 2012, "Intersubjectification and clause periphery", *English Text Construction*, 5, 1, 7–28.
Traugott, E., 2014. "On the function of the epistemic adverbs *surely* and *no doubt* at the Left and Right Peripheries of the Clause", in: K. Beeching, U. Detges (eds), *Discourse functions at the Left and Right Periphery. Crosslinguistic investigations of language use and language change*, Leiden/Boston, Brill, 72–91.
Traugott, E.C., R.B. Dasher, 2002, *Regularity in Semantic Change*, Cambridge. Cambridge University Press.

BDD-A28574 © 2018 Editura Academiei
Provided by Diacronia.ro for IP 216.73.217.142 (2026-11-14 23:09:52 UTC)

Van Bogaert, J., 2011, "*I think* and other complement-taking mental predicates: A case of and for constructional grammaticalization", *Linguistics*, 42, 2, 295–332.
Vincent, S., S. Darbaky, A. Mettouchi, 2009, "The grammaticalization of *you know.* From shared knowledge to control over the co-speaker", *English Text Construction*, 2, 2, 209–227.
Wilson, D., 2011, "The Conceptual-Procedural distinction: past, present and future", in: V. Escandell-Vidall, M. Leonetti, A. Ahern (eds), *Procedural meaning: problems and perspectives*, Bingley, Emerald Group publishing, 3–31.
Zimmermann, K., 2002, "La variedad juvenil y la interacción verbal entre jóvenes", in: F. Rodríguez González (eds), *El lenguaje de los jóvenes*, Barcelona, Ariel, 137–164.

BDD-A28574 © 2018 Editura Academiei
Provided by Diacronia.ro for IP 216.73.217.142 (2026-11-14 23:09:52 UTC)