Imperatives, Discourse Markers, and Grammaticalization É vida, olha…: Imperatives as Discourse Markers and Grammaticalization Paths in Romance

Benjamin Fagard

É vida, olha…: Imperatives as Discourse Markers and Grammaticalization Paths in Romance

A diachronic corpus study

Fagard Benjamin

CNRS – ENS and University Paris 3 (France)

In this paper, we investigate the evolution from imperatives to discourse markers in Romance, with a corpus-based approach. We focus on the case of items coming from verbs meaning "to look", in a semasiological perspective: Spanish and Catalan *mira*, Portuguese *olha*, Italian *guarda*, French *regarde*, Romanian *uite*. We show that they all share many uses, among which turn-taking, introduction of reported speech, hesitation phenomenon, topic-shifting and modalization, except for French *regarde*. We then establish (against Waltereit 2002) that the development of these uses *is* the result of a process of grammaticalization, from lexical, clause-internal uses to uses as discourse markers, in a cline which tends to confirm the predictions made by Brinton and Traugott (2005). The lesser grammaticalization of French *regarde* could seem unimportant, but is in apparent contradiction with the now well-established fact that French is, of all Romance languages, the most grammaticalized. We try, in conclusion, to address this paradox: is French not so grammaticalized after all, or is this just an exception to the rule?

Keywords: grammaticalization, discourse markers, corpus study, spoken language, Catalan/French/Italian/Portuguese/Romance/Romanian/Spanish

0. Introduction

What are you supposed to look at when your great-aunt tells you *É vida, olha…* ("Look, that's life…")?[1] Nothing, of course. There have been many studies on the grammaticalization of the imperative of verbs meaning "to look" or "to observe" into a discourse marker (hereafter DM) used to draw the addressee's attention not on an object but on discourse itself, and later on into a turn-taking device (for instance Brinton 2001, Dostie 2004, Gallardo and Marín Jordà 2005, Marín and Cuenca 2000, Marín Jordà 2005, Pons Bordería 1998, Rossari 2006, Waltereit and Detges 2007). It is found in various languages: English *look*, Dutch *kijk*, Spanish, Galician and Catalan *mira*, Portuguese *olha* and Galician *olla*, Italian *guarda*, Romanian *uite*, Polish *patrz*, etc. We propose here to investigate specifically the development of imperatives of such verbs in French, Spanish, Catalan, Portuguese, Italian and Romanian: *regarde, mira, olha, guarda* and *uite*.

Three main issues are at stake: the status of DMs in general and the uses of the DMs we are investigating in particular; their development from (lexical)

imperatives to DMs, and finally the place of French with respect to the other Romance languages,[2] since French *regarde* has a more restricted range of uses than the other items.[3] In the first part of this paper, Section 1 ("Theory and method"), we present a working definition of DMs, and list the corpora we used for the present research. In the next part, Section 2 ("The discourse marker uses of Romance *look*-forms"), we define and illustrate the main uses of the items under study, and show that only French *regarde* clearly differs from the others, having few DM uses. In Section 3 ("How they come about: the evolution of *look*-forms in Romance"), we address the question of the development of these DMs, with a diachronic corpus study of some of the items we chose to investigate, and argue against Waltereit (2002) that they do constitute a case of grammaticalization, albeit not a prototypical one. These results bring to light a paradox: French is usually taken to be the most grammaticalized Romance language, and DM uses of Romance *look*-forms seem to be the result of a process of grammaticalization, yet the French look-form is the *least* grammaticalized. We will discuss this paradox in Section 4 ("Discussion: what about French?").

1. Theory and method

1.1 A working definition of discourse markers

As Traugott (2007a:141) points out, there are various definitions of DMs. Following mainly Schiffrin (1987:31, 323, 328), Hansen (1996:111), Waltereit and Detges (2007:63) and Traugott (2007a:144), we adopt here a broad definition of DMs as elements which operate at the discourse level. This implies that their meaning is non-propositional, and that they are functionally extra-sentential: this feature of DMs, in turn, is made apparent by their relatively free distribution (i.e. syntactic position), specific prosodic contours, and variable scope. It also makes DMs a useful pragmatic tool for discourse signaling and negotiation of ongoing discourse-related interaction. A final feature, which is of particular interest for the diachronic study of DMs, is that they tend to cluster with other DMs or marks of subjectivity. This is the case for all Romance *look*-forms, which often cluster with or without repetition; examples 1 and 2 below illustrate this point for Italian and Catalan (see also Marín and Cuenca, 2000:225-226 for various examples of clusters in Spanish and Catalan: *Mira, oye; Miri, escolti'm; Sí, a veure, miri,* etc.):

(1) *no guarda allora guarda se se io potessi te lo te lo regalerei volentieri io*
     "no look now look if I could I would gladly give it to you" (LIP, FB14-280A)

(2) *Sí, doncs mira, este estiu he estat a... a Sicília*
     "Yes, so look, this summer I went to… to Sicily"
     (CCCUB, COD, València)

These features amount to the delimitation of a specific paradigm, which can receive various denominations besides discourse markers, for instance *discourse connectives* (Erman and Kotsinas 1993:79), *discourse organizers* (Pons

Bordería, 1998:215) or *discourse signals* (Lamiroy and Swiggers 1991:123). We will stick, for now, to the usual denomination of *discourse marker*.

According to this definition and to their common denomination as *discourse* entities, DMs arising from imperatives of perception verbs display a wide range of discourse-related functions: they are used to structure discourse, call for the addressee's attention, emphasize thematic progression and enable or facilitate turn-taking. The examples given by Brinton (2001:180) for English *look*-forms illustrate some of these uses; she explains that *look*-forms have a global meaning of the type "'pay attention', 'heed me', or 'listen (up)'".

1.2 Literature and corpora

Under the hypothesis that the evolution of English *look*-forms is the result of a universal linguistic trend (in the sense of Haspelmath 1997), we can assume that the functions of their equivalents in other languages are globally quite similar. In order to test this hypothesis, we studied the DM functions of Romance *look*-forms, mainly on the basis of the existing literature listed in the introduction. We then compared and completed the results of this preliminary study with a corpus study, using databases of Modern Spoken Romance; wherever possible, we checked our 'non-native' intuitions with native speakers.[4] We give below some details on the corpora and databases we used: first, the name and the number of words or tokens, which may indicate, not the total number of words of each database, but the size of the portions we used. Second, the number of occurrences of the main *look*-form for each Romance language, including lexical and DM occurrences.

For French, we used the Valibel, PFC, and Clapi corpora, which contain a total of 506 occurrences of *regarde, rega'de, r'garde* for approximately 1,6 million words[5]. For Spanish, we used the spoken sections of the Crea database, which contains a total of 1,373 occurrences of *mira* for 2,8 million words (European Spanish only, though we did include an example of South American Spanish). For Catalan, we used the on-line version of the Corpus de Català Contemporani de la Universitat de Barcelona (CCCUB) database, which contains 117 occurrences of *mira* for 80,000 words. For Romanian, we used the CORV and Ruxandoiu corpora, which contain 65 occurrences of *uite* for 100,000 words. For Italian, we used the LIP corpus, which contains 341 occurrences of *guarda* for 500,000 words. Finally, for Portuguese, we used the spoken section of the Corpus do Português (European and Brazilian Portuguese), which contains 609 occurrences of *olha* for approximately 2 million words.[6]

The corpus thus constituted is more or less homogeneous in terms of text type, since most of these databases include both spontaneous face-to-face conversation and more formal types of spoken interactions such as TV shows, debates, lectures and so on. However, as can be deduced from the figures given above, the subcorpora are hardly comparable in terms of size. For this reason, although we annotated a selection of the occurrences detailed above according to formal, morpho-syntactic and semantico-pragmatic criteria,[7] we thought it wiser not to include frequencies or relative frequencies of the different uses described in Section 2 for the different Romance languages. This will have to wait for further research, on the basis of a more homogeneous corpus.

Note that Section 3 also includes corpus data, from a diachronic corpus study on two databases containing Medieval texts, the New Amsterdam Corpus for French (3 million words) and the Opera del Vocabolario Italiano for Italian (21 million words), and two diachronic databases, the Corde database for Spanish (250 million words) and the Frantext database for French (more than 250 million words).

2.   The DM uses of Romance *look*-forms

We present here the main DM uses of Modern Romance *look*-forms.[8] Since these forms are not related, there is no specific reason to expect them to have the same uses, except as the result of a universal (cognitive) trend, as said above. Resemblances resulting from areal phenomena are not expected either, at least for Romanian. However, these forms *do* share many characteristics. They all present morphological variation, usually singular vs. plural and informal vs. formal, the latter actually being a form of subjunctive rather than imperative, compensating for the absence of third person imperative form. We find for instance in Spanish *mira* (singular informal), *mire* (singular formal) and *mirad* (plural informal), in Catalan *mira* (singular informal)*, miri* (singular formal) and *mirin* (plural formal), in Romanian *uite* (singular formal) and *uitaţi* (plural informal), etc. Distributionally, they can all appear in sentence-initial, sentence-medial and sentence-final position. From a functional point of view, they all have variable scope, except maybe European French *regarde*. Finally, as we will now see, they all have similar uses, and all seem to share with English *look*-forms a large enough range of uses to identify them clearly as DMs.

Non-DM uses are also found for all the items under study. This is the case for lexical uses in which the imperative marks a request to look at something, as well as for more abstract uses in which the addressee is requested not to look at an object but to take something into consideration. For instance, all these markers can introduce an example. We could describe it as the result of a slight metonymic change from "look at this (concrete) and take it as a good instance of what I say" to "look at this (abstract) and take it as a good instance of what I say". Example 3 below illustrates this point for French *regarde*:

(3)   *(A) chacun chaque // que je vais dire chaque province // ont leur accent regarde à Liège // Louisa euh / elle le sait bien elle a été |- longtemps (B) mm -| assez à Liège*
"(A) each-one each // that I am-going to-say each region // have (sic) their accent <u>look</u> in Liege // Louisa hum / she knows that well she stayed |- a-long-time (B) hum -| long-enough in Liege"
(Valibel, accPH1r)

In such contexts, Dostie (2004:112) and Rossari (2006: ex. 2) take *regarde* to be a DM. It is, syntactically, between clause-internal and clause-external use: on the one hand, the imperative still has a complement (here *à Liège* "in Liege"); on the other hand, it is not syntactically subordinated to the main verb (*ont* 'have'), and is hardly modifiable (*regardons* 'let us look' would seem quite strange).[9] At any rate, we also found clearly DM uses, for all items under study. In those uses, which we believe to be more or less direct developments of the

lexical (or propositional) meaning of *look* (a call for attention), these items call the addressee's attention:

- on the topic (topic shifts, subtopics, topic continuity): in these textual uses, the DM could be glossed as "talking about A, pay attention to what I have to say about B!"

- on the speaker's evaluation of what is said (mark of surprise, of disapproval, of reinforcement): in these subjective or expressive uses, the DM could be glossed as "talking about A, pay attention to what I think of it!"

- on the speaker's discourse in general (turn-taking): in these intersubjective uses, the DM could be glossed as "(while you talk about A), pay attention (to what I have to say!)". Two other uses seem to us directly derived from this one, as we will see below: the introduction of reported speech, and hesitation phenomena.

This hierarchization of DM uses reflects Traugott and Dasher's (2002:170) subjectification cline, going from propositional to textual, expressive and finally intersubjective meanings.

2.1 Topic issues

In this function, the DM introduces subtopics or topic shifts, thus acting as a continuative or moving marker, "sending a signal from the speaker to the listener that the different parts of the utterance are to be connected, or rather, that the discourse is moved forward, e.g. from one event, state or mode to the next" (Erman and Kotsinas 1993:90). It is found for Spanish *mira* (Gallardo and Marín Jordà 2005:65, 77, Pons Bordería 1998:214, 225), Italian *guarda* (Manili 1985:311, Waltereit 2002:990), Catalan *mira* (Marín Jordà 2005:63, Marín and Cuenca 2000:233), Portuguese (see ex. 4 below) and Quebec French *regarde* (Dostie 2004:114). However, we found no instance of this use for European French *regarde*.

(4) *(A) Pois. Eu tenho, eu tenho tudo pronto, filha se quiseres,*
(B) Vamos, vamos embora!
(A) Pois, vê lá, tu é que sabes. <u>Olha</u> eu não, não forço porque sei, eu sei que é (…)
(B) …
(A) Hã?
"(A) Well. I've got, I've got everything ready, young lady, if you want
(B) Let's go, let's go now!
(A) Well, look here, you're the one who knows. <u>Look</u> I don't, I don't press you on because I know, I know that it's…
(B) …
(A) Huh?"
(Corpus do Português, 19Or:Pt:CRPC, 1383)

For Romanian *uite*, we found no example, but it seems that such uses are not completely ruled out; a suitable context would be the following:

(5) *Mie nu îmi plac maşinile. <u>Uite</u>, motocicletele sunt mai ieftine*
"I don't like cars. <u>Look</u>, motorcycles are cheaper"
(*not* a corpus example; A. Ştefănescu, personal communication)

Pons Bordería (1998:226) explains that this use is linked to problems of discourse planification, resulting from the speed of oral language as described by Chafe (1982). According to Brinton and Traugott (2005:137), quoting Quirk et al. (1985:1114-1115), it is because "these comment clauses may […] be used to claim the hearer's attention"; Waltereit (Ibid:990) provides a similar explanation.

2.2 Evaluative use

In this second type of use, the DM emphasizes the speaker's point of view. It can convey various attitudes, from surprise to disapproval, or simply put emphasis on what the speaker says. The DM is no longer clause-internal, and can therefore appear as a pro-sentence, sometimes with repetition (e.g. Italian *guarda guarda* "look look") or cumulation with other DMs (e.g. Spanish *mira, oye* "look, listen"). Such uses are found for all Romance *look*-forms: Italian *guarda* then conveys surprise, admiration (Manili 1989:8, Waltereit 2002:992); it can also express disapproval, or emphasis, as in example 6 below:

(6) *Claudia non potrei mai <u>guarda</u> fare del sesso con te*
"Claudia I could never <u>look</u> have sex with you"
(Lip, NA1-246B)

Spanish and Catalan *mira* have similar uses, in which *mira* functions as a false solicitation for the addressee's attention, according to Pons Bordería (1998:221-223 and personal communication; see also Marín and Cuenca 2000:217, ex. iv, footnote 5 and Ibid:233, Marín Jordà 2005:61, 66). Portuguese *olha* and Romanian *uite* also have such uses, as shown in examples 7 and 8 below for Romanian:

(7) *A: da' eu zic↓ faceţi şi voi↓ că n:u vă strică. eu eu fac masaj aşa cum pot↑ cu degetele↓*
D: da' <u>Uite</u>↓ cu ăsta faci cu MULT mai bine.
A: păi SIgur că fac mai bine.
"(A) but I say do it too because it won't do you any harm. I do a massage as I can with [my] fingers
(D) but <u>Look</u> this way you do it MUCH better
(A) well SUre (that) I do it better"
(Ruxandoiu, dialog 'aparatul de masaj')

(8) *B: e ALtu.*
C: da↓ <u>uite</u> ĂSta este care trebuie.
"(B) it's ANOther one
(C) yes <u>look</u> THIs one is the one we need"
(Ruxandoiu, Ibid)

Quebec French *regarde* can be used for emphasis, as in Dostie's (2004:117) example 9 below:

(9) *On ne me fait pas assez confiance, c'est tout! <u>Regarde</u>…! C'est loin d'être bête ma suggestion*

"People don't trust me enough, that's all! <u>Look</u>...! My idea is far from being silly"

In European French, these uses are rather infrequent and probably not yet conventionalized, but example 10 below seems to illustrate the possibility of partially integrated instances. Here, *regarde* is not really part of the main sentence (here *tu es déjà en deuxième*), and it is in initial (thus peripheral) position, just like *dis* (lit. "say, tell me") which clearly functions as a DM in this excerpt.

(10) *(A) c'est c'est combien d'années toi que tu es parti là quatre ans ou |- c'est <B> ouais -| ah oui quatre ans / <u>regarde</u> un peu dis tu es déjà en deuxième / ça ça passe vite quand même*
"(A) it's it's how-many years you that you have-been gone there four years or |- it's <B> yes -| oh yes four years / <u>look</u> now say you're already in second grade / it [time] it goes by don't you think"
(Valibel, famES1r)

However, it is part of a (lower-status) syntagm, *regarde un peu* (lit. "look a bit"). *Regarde un peu*, along with *regarde donc* and other similar constructions, has clearly not yet reached the lexicalized state of English *look-a-here* or *lookit*, but might be said to have reached the preceding stage of 'fixed formula', along with other constructions such as *regarde donc* "do look, please look"[10] (Brinton 2001:190, table I).

2.3 Turn-taking / Interruption

This use is attested for most Romance *look*-forms, but not for (at least European) French *regarde*. It is found for Italian *guarda* (Manili 1985:305-6, Waltereit 2002:991, 995), for Spanish *mira* (Pons Bordería 1998:214, Gallardo and Marín Jordà 2005:67, Marín and Cuenca 2000:219, 231), for Catalan *mira* (Marín and Cuenca 2000:232), Romanian *uite* and Portuguese *olha*. Examples 11 and 12 below illustrate this phenomenon for Catalan and Romanian:

(11) *(A) la tieta no se porta bé eh*
*(A) [no s'ha acabat_*
(A) la sopa\]
(B) [<u>mira</u>
(B) digue-li que xx] s'acabi la sopeta com tu
"(A) auntie doesn't behave eh
(A) [she didn't finish
(A) her soup]
(B) [<u>look</u>
(B) tell her to xx] finish her soup like you"
(CCCUB, COC, Conversa 4)

(12) *A: da' nu e decît una din# din străinătate din ucraina.# aşa că:*
B: da chiar↓ noi nu putem să trimi[tem#
A: [<u>uite</u>
+ B: pur şi simplu aşa la:
B: /brucsel/↓ la ăştia.
*A: a trimis domnu c*** ieri la viena (...)*

"(A) but there's only one from from abroad from Ucraine. so…
(B) oh sure we cannot send [it]
(A) <u>look</u>
+ (B) just like that to…
(B) Brussels to that guy
(A) mister C*** sent [it] to Vienna yesterday"
(Ruxandoiu, dialog 'pe stradă')

According to Brinton (2001:181), this use of *look*-forms is due to the fact that they function as 'alerts' (Stenström 1994) – alerts being, in Stenström's terms, "excellent interruptors" (Ibid:74).

Example 13 below illustrates the fact that it is not possible for *regarde* to have such a use in (European) French:[11] even in turn-initial position, *regarde* seems to retain a lexical meaning, since B is showing A that the S does appear.

(13) *(A) que quand je dois taper trois fois normalement j'ai S*
(B) ouais
(A) et ici ça le fait pas
(B) ben si
(A) non |- écoute <(B)> <u>regarde</u> -| le S donc |- normalement <(B)> oui oui -| je fais / un deux trois quatre hein
(B) ouais / voilà |- tu as un S
"(A) that when I have-to hit three times [on the key] normally I get an S
(B) yeah
(A) and this-time it doesn't work
(B) yes it does
(A) no |- listen <(B)> <u>look</u> -| the S now |- normally <(B)> yes yes –| I do / one two three four huh
(B) yeah / there |- you have an S"
(Valibel, famES1r)

2.4 Hesitation phenomenon

A similar DM use of these imperatives is as a hesitation phenomenon or a hedge. We consider this use to be derived from the turn-taking use of these DMs, with a pragmatic meaning shift from "look what I have to say" to "look, I have something to say (even if I don't yet know exactly what)". We found such instances in our corpora for Spanish, Portuguese, Catalan, Italian and Romanian, illustrated by examples 14 to 16:

(14) *E eu! Eu também bem gostava... <u>Olha</u>, agora não, mas nalgum tempo…*
"Me too! I liked [that] too… <u>Look</u>, now I don't [anymore], but there was a time…"
(Corpus do Português, 19Or:Pt:Cordial, OUT08)

(15) *non lo so <u>guarda</u> ieri ieri ho avuto una difficoltà a parlare questo sì*
"I don't know <u>look</u> yesterday yesterday I had trouble talking that's true"
(LIP corpus, R A 8 220 E)

(16) *A: (…) n-am prins trenul de la şase.*
B: CE tren.

A: (uite↓) nu ştiu. trebuia să plecăm amîndoi.
"(A) I missed the six o'clock train
(B) WHAT train
(A) look I don't know. We had to leave, both-of-us"
(Ruxandoiu, dialog 'povestirea unui vis')

For Italian *guarda*, Waltereit defines these uses as "a mere hesitation phenomenon", on account of its turn-medial position (2002:993-4). For Spanish *mira*, Pons Bordería explains that it is associated to no conversational value and that its hedging function reflects very clearly the speaker's insecurity concerning the content of his message (1998:223-4). In the above examples, we consider the *look*-form to be a hedge on account of its context, which indicates that the speaker is not sure of what he asserts, or does not know what to say. This is manifest in the contradiction between *gostava* and *agora não*, and in the presence of a pause (ex. 14), typical of such uses (Manili 1989:4-5); also, in the expressions *non lo so*, *nu ştiu* (ex. 15 and 16). In such contexts, the DM is used to draw the addressee's attention to the fact that the speaker does not know the answer to a given question (which could be the case in ex. 16), and sometimes seems to be used as a simple filler (Manili 1985:308).

2.5 Introducing reported speech

In this final use, the DM introduces a sequence of reported speech, usually after a verb meaning "to say". It seems to be attested for all languages (except European French[12]): it is reported for Spanish *mira* (Pons Bordería 1998:222, Gallardo and Marín Jordà 2005:71) and Italian *guarda* (Waltereit 2002:993; see also our ex. 17 below). It is also attested in Portuguese, as shown by example 18 below, as well as for Catalan and Romanian.

(17) *questo Pierluigi me lo dice perché io gli ho detto <u>guarda</u> per mettergli un po' so' sincera per mettergli proprio paura ho detto <u>guarda</u> il discorso è questo non dovete sottovalutare nulla*
"the reason Pierluigi tells me that is because I told him <u>look</u> to frighten him a little I'm honest to give him a good scare I said <u>look</u> here's the story you shouldn't underestimate anything"
(Lip, RA9-71A)

(18) *E então o meu marido até diz assim: "<u>olha</u>, falamos já tudo antes da filha se sentar à mesa"*
"And then my husband even said this-way: '<u>look</u>, let's talk it over already before the girl comes to the table'"
(Corpus do Português, 19Or:Pt:CRPC, 356)

According to Waltereit (Ibid), "it is not by accident that *guarda* appears frequently at the beginning of quotations. Quotations are, by their very nature, introductions of a different viewpoint. They are therefore very often topic shifts". However, it could be directly derived from the clause-initial DM, used as a turn-taking or topic shifting device, and reproduced in the reported speech as a marker of spoken language. In Renzi et al.'s (1994:226) example, which we reproduce below, as well as in other cases, it is not clear, despite the quotation marks, whether the *look*-form introduces reported speech or is part of the reported speech itself:

(19) *Mi ha detto: "Guardi, così non si va più avanti"*
      "He told me: 'Look, we can't go on this way'"

3. How they come about: The evolution of *look*-forms in Romance

3.1 Synchronic perspective: A grammaticalization process?

From a synchronic perspective, we could describe the formation of these DMs as the result of a grammaticalization process, though the *grammaticalized* status of these DMs has been subject to debate. This can be explained, in our view, by two elements: on the one hand, it is only recently that DMs have been included in the grammar of languages, and it is not yet a consensual issue. In fact, DMs in general are often denied grammatical status, because, for instance, they are not obligatory (Waltereit, 2002), and subject to much interpersonal (Manili 1989:2), topical and regional (Dostie 2009:202) variation. However, Fraser maintains they are "part of the grammar of a language" (1988:32; see also Traugott and Dasher 2002:11).
      On the other hand, as we will now see, the 'usual' grammaticalization criteria (see Hopper and Traugott 2003 for a broad view on grammaticalization, Lehmann 1995 for a stricter definition) are not all met by the DMs we are interested in here, or even DMs in general. Some are met only regionally or for isolated forms, like phonological attrition and morphological bonding. Others are met in all or nearly all Romance languages, like semantic bleaching, syntactic decategorialization, syntagmatic fixation, and paradigmaticalization. Finally, others yet are not met but have recently been challenged as indicative of grammaticalization, such as scope reduction.

3.1.1 Phonological attrition

While the intonational status of Romance *look*-forms marks their parenthetical nature (at least in some cases, see Marín and Cuenca 2000:217), there is little evidence of phonological attrition, be it in Catalan (Marín Jordà 2005:57), Spanish (Pons Bordería 1998:216), Quebec French (Dostie 2004:121), European French or Portuguese. The only cases of phonological reduction appear locally, as in regional *ma* or *mi* (Marín and Cuenca 2000:230). This puts the Romance *look*-forms apart from the English ones, for which phonological attrition is well attested (Brinton and Traugott 2005:138, Brinton 2001:193); besides, it is not a feature of Romance languages in general but of these forms in particular, as shows the existence of reduced forms for Spanish *oye* "hear!" to *ye, yé* or *oy* (Pons Bordería 1998:217, Marín and Cuenca 2000:230-231) and Quebec French *coudon*, from *écoute donc* "do listen!" (Dostie 2004:121).

3.1.2 Morphology

The morphology of these forms also seems to indicate a lesser degree of grammaticalization. There is still partial agreement between the imperative

forms and the addressee; this is true for all languages, which display a variety of forms: *regarde, regardez* (Dostie 2004:121), *mira, miri, mirin* (Marín Jordà 2005:53), *mira, mire, mirad* (Pons Bordería 1998), *guarda, guardi, guardate* (Waltereit 2002, Bazzanella 1994), *olha, olhe* (and rarely *olhai*) in Portuguese; in Romanian, the forms *uite, uita* and *uitaţi* are specific to the DM use, in that they present no reflexive pronoun, as the two examples below illustrate. Whereas the first example is purely lexical, the second one is a DM use of *uite*:

(20)   *şi aşa e↑ uită-te la mi[ne*
         "and that's it look at me"
         (Ruxandoiu, dialog 'colege de serviciu')

(21)   *<_F CE face> el noaptea. CE FAce. uite↓ m-a m-a-ntrebat şi adina că*
         nu-i venea să creadă. <_J adina mă-nţelegi>
         "WHAT he does at night. WHAt he DOes. look Adina has- she asked
         me too because she couldn't believe it. Adina understands me."
         (Ruxandoiu, dialog 'bîrfă')

Even minor changes in form (singular vs. plural, formal vs. informal) can be accompanied by slight functional variations (Renzi et al. 1994:230, Bazzanella 1994:146, 154 for *guarda*); this could indicate that some forms are more grammaticalized than others. There are, besides, a few signs that there is sometimes a loss of agreement, with mostly 2nd person uses, as in Spanish (Pons Bordería 1998:216, Gallardo and Marín Jordà 2005:64, footnote 12), and formal 2nd person in Catalan – 87.5% of occurrences in the study of Marín Jordà (2005:53), with examples of discordance such as *mirin, escolti'm* "look.imp.2pl, listen.imp.2pl.pol (to) me" (Ibid). In Italian, there seems to be a beginning of specialization (with *guardi* indicating attenuation more often than *guarda*, see Manili 1989:7). Of course, the situation is different for English, which does not present such variations, even for lexical verbs.

3.1.3 Semantics

The meaning of *look*-forms is quite clearly bleached, in most uses, as noted by various studies (Manili 1985:308, Marín Jordà 2005:49, Pons Bordería 1998:214, Waltereit and Detges 2007:78); this is also the case for English (Brinton 2001:193). Marín Jordà notes, besides, that the desemanticization of *look*-forms is more important than that of similar *hear*-forms (2005:50), which also have DM uses.

In some instances, including in some examples we gave in Section 2, the original meaning is retained, or at least some features of the original meaning (Dostie 2004:112, Rossari 2006: Section 6, Pons Bordería 1998:215): Brinton, for instance, writes that "It should be noted that all of the *look*-forms retain literal (perceptual) meaning" (2001:181). This can be explained inside the framework of grammaticalization by the phenomenon of *persistence* (Hopper & Traugott 2003:96).

3.1.4 Paradigm

In all Romance languages, these forms are part of a larger paradigm. This paradigm includes other imperative DMs which have similar semantic, morphosyntactic and discursive features (like Catalan *a veure* "let's see", Marín and Cuenca 2000:216), or similar functions but a different (etymological) meaning such as "listen" (like Italian *senti*, Spanish *oye*, French *écoute*, see Manili 1985:305, Pons Bordería 1998, Rodríguez Somolinos 2003) or "say", "tell me" (like French *dis(-moi)*, Italian *dimmi*…). It also includes other verbal forms such as "you see": French *tu vois*, Romanian *vezi*, Italian *vedi*… Finally, DMs of adverbial origin can have similar functions, like *alors, bon, ben* "then, well" in French (Hansen 1996).

3.1.5 Syntax

The syntactic features of Romance *look*-forms also clearly indicate that they have undergone or are undergoing a process of grammaticalization. In most instances, these forms do not accept objects or subjects, thus rejecting an analysis as centre of a verb phrase; this is true for Catalan (Marín Jordà 2005:52), Spanish (Pons Bordería 1998:216-217), and French (Rossari 2006: Section 4.2), as well as Italian, Portuguese and Romanian: all rarely take overt subjects, and only tolerate postposed vocatives. For instance, Marín Jordà notes the presence, in 10.5% of the occurrences of *miri* in her corpus, of postposed vocatives (2005:56). In our corpora, we found no DM occurrence of the type *tu guarda*, etc., although such constructions do exist (Manili 1989:8). Indeed, the few occurrences of *tú mira*, for instance (only two in the whole CREA/spoken corpus), are not proper DM uses, like the example 22 below:

(22) *yo pasé pensando que estaba solo, pero como no había luz yo pensé: bueno, estoy solo. Entonces cuando escucho una voz que tú, mira, ven acá. ¿Para qué? No, yo no te conozco. Como era en la oscuridad, bueno.*
"I walked on, thinking I was alone, but it was dark so I thought: OK, I'm alone. So when I hear (sic) a voice that [said] you, look, come over here. What for? No, I don't know you. Since I was in the dark, you know."
(Crea, CSHC-87, Entrevista 19, Venezuela)

This is quite similar to what has been observed for English (Brinton 2001:193). There are also tense constraints, as Rossari shows for French *regarde* (Ibid and Section 2.2).

The *fixation* parameter (or syntagmatic variation) is not as problematic, in our view, as it seems that there is a clear tendency for these DMs to occupy definite positions, in particular clause-initial and clause-final. It is true that they can also appear in clause-medial position; however, in this case, they are most often clearly parenthetic, and can also be the result of "false starts" as in example 23 below:

(23) *Hace poco mira, hace poco, Jordi, me decía me decía Ana Belén una cosa muy graciosa…*
"Recently look, recently, Jordi, Ana Belén was telling me was telling

me something very nice"
(ex. 12a in Marín and Cuenca 2000:227)

Besides, in different positions, they also tend to have different pragmatic meanings, as pointed out by Lamiroy and Swiggers (1991:139). In our corpus, the main correspondences which can be observed between position and meaning in DM uses of Romance *look*-forms are the following. Turn-initial uses generally correspond to turn-taking and interruption. Turn-medial, clause-initial uses (such as false starts, ex. 23) are generally associated to topic-shift and subtopic introduction. Finally, turn-medial, clause-medial uses are associated to various uses: evaluative and hedge uses as well as, in specific contexts (most often after a verb of speech), the introduction of reported speech.

3.1.6 Other grammaticalization hallmarks

The criterion of *scope reduction* (Lehmann 1995) seems more problematic. DMs do not show signs of this phenomenon, as Brinton and Traugott (2005:138) note. However, as they write, "the notion of scope reduction has been challenged in grammaticalization generally (Tabor and Traugott 1998) and in the case of the grammaticalization of discourse markers specifically" (Ibid).

To sum up on grammaticalization, we could say, in a prototypical view of grammaticalization (as suggests Pons Bordería 1998:218), that the DM uses of Romance *look*-forms constitute a borderline case of grammaticalization. Similar views are defended by Marín Jordà (2005:58), Brinton and Traugott (2005:138), Dostie and Pusch (2007) and Traugott (2007:151); and, for other DMs, by Rodríguez Somolinos (2003:73, 81); Lamiroy and Swiggers (1991:121) speak of "functional 'displacement'" with a "diachronic 'emptying'" of their meaning.

3.2 A brief history of imperative DMs

The evolution of English *look*-forms, as reconstructed on the basis of written corpora by Brinton (2001), seems to confirm the grammaticalization hypothesis. She explains (there and in Brinton and Traugott 2005:138) that these forms start out as free matrix clauses in Middle English, later develop pragmatic uses, in which the imperative "gives a pragmatic instruction to the hearer to 'be careful'", and finally what we take to be proper DM uses in Modern English, in which the imperative "conveys the speaker attitude of impatience" (Ibid). These imperatives thus evolve from free clauses to adjunct uses and finally extra-sentential uses.

In this evolution, the importance of the frequent initial and parenthetic use of DMs has been pointed out by several studies (see Fraser 1996), in particular studies of imperatives of perception verbs, like Brinton (2001:178, 183), Marín Jordà (2005:51, 54); indeed, this seems to be an important feature of their evolution (Traugott 2007a:151). Is this true for Romance imperative DMs as well?

3.2.1 History of DMs: the case of Romance

On the basis of the literature and the observation of synchronic, spoken corpora, we proposed in Section 2 a similar cline of grammaticalization and (inter)subjectification for the development of Romance *look*-forms, hypothesizing that their evolution went from propositional to textual, expressive and finally intersubjective uses.

In order to check this hypothesis, we conducted a limited diachronic corpus study, analysing occurrences of Spanish *mira*, Italian *guarda* and French *regarde* in Medieval texts. This was done with the following databases: NCA, Frantext, Corde and OVI. Of course, our results are only indicative of these items' evolution, rather than proof: the fact that DMs are spoken elements *par excellence* makes it hard to use written texts to study their development. We did, however, observe an evolution which closely resembles at least the first steps of our 'cline'. The initial use is, of course, lexical; we do not need to illustrate this use. The first step of the evolution is the development of metaphorical uses, still intrasentential, illustrated by examples 24 and 25 below:

(24) *¿Non bastan tus propios males e particulares penas, que con las glorias ajenas sientas tormentos mortales? mira que todos yguales en este mundo venimos, e assí mesmo morimos, mas bivimos desiguales.*
"Are your own aches and pains not enough, that you must endure mortal torment through others' glories? look that we all arrive as equals to this world, and perish in the same way, but our lives are not equal"
(Corde, *Cancionero castellano*, 15th c., 283)

(25) *regarde chu ki est presens et chu ki est a uenir.*
"look at what is present and what is to come"
(NCA, *Gloses wallonnes*, 1290, 201)

The second step, illustrated in the examples 26 and 27 below, is the use of *look*-forms with a meaning close to English "beware!".

(26) *car regarde a ton afaire si te retien de mal a faire dont tu mes t'ame en grant dolor*
"do take care of yourself and beware not to do any wrong by which you might put your soul in great pains"
(NCA, *Verite*, 1450, l.1181)

(27) *E llamaron al portero del castillo e dixeron: - Cata aquí presos que traemos al rey Artur; e liévagelos e mira que no pierdas ninguno dellos, ca te dezimos, cierto, que tu señor nunca tan grand plazer uvo ni vio como éste.*
"And they called the guard at the door of the castle and told him: - See here the prisoners we bring to the King Arthur; and take them to him and look [to it] that you do not lose any of them, because we tell you, for sure, that your lord never had or saw anything so pleasant."
(Corde, *El baladro del sabio Merlín con sus profecías*, 15th c., Fol. 60 R°)

(28) *Se tu voli biasimare o riprendere altrui, guarda che tu non sia magagnato di quello vizio medesimo*
"If you want to blame or criticize others, look [to it] that you not be

devoured by the same vice"
(Ovi, *Tesoro di Brunetto Latini volgarizzato*, 1300, L. 7 cap. 13)

Non-literal reduplicated uses appear later:

(29) *¿Y asimismo, que hasta agora siempre has estado dudoso de la*
alcanzar y tenías sofrimiento; agora que te certifico el fin de tu penar,
¿quieres poner fin a tu vida? <u>mira, mira</u>, que está Celestina de tu parte,
y que aunque todo te faltase lo que en un enamorado se requiere, te
vendería por el más acabado galán del mundo (...)
"And still, although until now you have always doubted your success
and suffered; now that I assure you the end of your pains has come, you
want to put an end to your life? <u>Look, look</u>, that Celestina is on your
side, and that even if you lacked all that is required in a lover, she
would claim you to be the most perfect gentleman on earth"
(Corde, Fernando de La Celestina Rojas, *Tragicomedia de Calisto y*
Melibea, 1499-1502, 234)

Finally, extrasentential uses only arise very late, in the 19th century for Spanish
mira:

(30) *No, mujer, no me altero... pero a lo menos que oiga el que yo la*
propongo, que le conozca y le trate, y después... <u>mira</u>, Bernardo a la
hora esta debe haber llegado ya de Barcelona; habrá consagrado los
primeros instantes a sus parientes; pero de un momento a otro le
tendremos aquí, y es preciso recibirle como a quien viene a ser mi
yerno
"No, woman, I am not angry… but hear at least what I have to offer
you, get to know him, and after that… <u>look</u>, Bernardo by now must
have arrived already from Barcelona; he will have devoted his first
hours to his parents; but any minute we will have him here, and we
have to honor him like one who will become my son-in-law"
(Corde, Mariano José de Larra, *No más mostrador*, 1831, 18)

The evolution of Romance *look*-forms, as we (tentatively) reconstructed it on
the basis of written corpora, thus also seems to confirm the grammaticalization
hypothesis. Two other arguments can be called upon here. The first one is that a
feature often associated (though not consensually) with grammaticalization is
(inter)subjectification. The subjective and intersubjective uses we saw in Section
2 could therefore constitute an argument for the grammaticalization hypothesis,
if we accept the claim that the two do go hand in hand. The second one is that
we can show, as Himmelmann suggests (2004:33), that "semantic-pragmatic
context expansion is accompanied by syntactic context and host-class
expansion". This is certainly true for our DMs, which start out with a semantic-
pragmatic context expansion from contexts in which the speaker calls the
addressee's attention on something to contexts in which he simultaneously
organizes discourse (while still calling the addressee's attention); this expansion
goes hand in hand with the diversification of syntactic contexts (clause-internal
uses to clause-initial, clause-medial and clause-final, extrasentential uses) and
host-class expansion from *look at an object* to *look at this (object/event)* to *look*
at this (object/event/discourse).

4. Discussion: What about French?

Our corpus is not homogeneous enough in terms of size for us to draw conclusions on the specificities of each marker: We established, rather, the *possibility* for each marker to have the different functions we illustrated in Section 2. We have shown, from this point of view, that French *regarde* has a much more limited range of uses, and lacks in particular a series of DM uses related to modalization and turn-taking. Other uses seem present only regionally (i.e. for Quebec French) or in certain contexts, constituting a mark of intersubjectivity rather than intersubjectivation (see Traugott 2007b:298-9, 2010).[13]

These elements indicate a lesser degree of grammaticalization for French *regarde*, with respect to its Romance counterparts. This seems to go against recent observations that French is more grammaticalized than all other Romance languages, in most if not all areas of language: from articles to quantifiers and auxiliaries (see for instance Carlier 2007, De Mulder 2001, Herslund 2005, Marchello-Nizia 2006:53, 111-112, 139). This apparent contradiction could be explained in one of three ways. First, it could simply constitute a counter-example to the general evolutive trend of Romance languages. Second, it might mean that this trend suffers partial exceptions, with elements which started to grammaticalize earlier in French (see for instance the existence of *voici* and *voilà* in Old and Middle French, Oppermann 2006) but stopped on the way and were replaced by yet other elements. Finally, it could also constitute not a counter-example but another proof that there is such a trend. This hypothesis is more convincing yet if we take into account Dostie's sketch of the history of *regarde*: she suggests (Dostie 2004:117 sqq.) that *regarde* did not acquire discourse uses in Quebec French, but already had them before the 16th century, and were thus imported with French into the 'New World'. This would mean that the discourse uses of *regarde* later decreased in European French, as has been shown for other DMs (Dostie 2009:204,207), owing to a continued renewal of the paradigm of DMs in this variety of French.[14]

The question is: what elements are used in Modern European French for the modalization and turn-taking functions displayed by *guarda*, *mira* and the like? We would have liked to provide here a definitive answer to this question, but will have to restrain ourselves, for now, to propose a few possible lines of inquiry. The most obvious answer would be *tu vois*, which can in some contexts alternate with *regarde* (Bolly and Degand 2009); however, it is also far from having the range of uses of Romance *look*-forms. Other DMs which should be taken into consideration here are the elements *bon, ben, beh*,[15] *eh bien* and *alors*, which Hansen (1996:115, 18, 124, 140) describes as introducing topic changes and/or marking refutation. Rodríguez Somolinos (2003:74, 78) and Dostie (1998:89, 92) show that *écoute* "listen!" may be used to express disapproval. As for turn-taking devices, we analysed the Valibel corpus, and found that three paradigms are very frequent as first elements of a turn: *oui* "yes" and the like (*ouais, mouais,* etc.: approximately 3000 occurrences); *hm* and the like (*euh, oh, ah*, etc.: approximately 5000 occurrences) and finally *ben* and *bon* "well, good" (approximately 800 occurrences). Their relative frequency at the beginning of turns is 8 (for *ben* and *oui*) to 17 (for *hm*) times higher than it is inside a turn. This is not true of other grammatical elements, such as the definite article (half as frequent in turn-initial position) or common nouns (*français* "French", ten

times less frequent in turn-initial position). Further research could show whether such elements are used by French speakers as equivalents of the other Romance *look*-forms, and whether this has ever been the case – providing we find a suitable corpus.

References

Bazzanella, C. 1994. *Le facce del parlare. Un approccio pragmatico all'italiano parlato*. Firenze: La Nuova Italia.

Beeching, K., Armstrong, N., Mougeon, R., Gadet, F. and Treffers-Daller, J. (eds). 2009. *Sociolinguistic Variation in Contemporary French*. Amsterdam/New York: John Benjamins.

Bisang, W., Himmelmann, N.P. and Wiemer, B. (eds). 2004. *What Makes Grammaticalization? A Look from its Fringes and its Components*. Berlin/New York: Mouton de Gruyter.

Bolly, C. and Degand, L. 2009. "Frequency Effects on the Evolution of Discourse Markers in Spoken vs. Written French." Paper presented at the Corpus Linguistics Conference, Liverpool University, 20-23 July 2009.

Brinton, L.J. and Traugott, E.C. 2005. *Lexicalization and Language Change*. Cambridge: Cambridge University Press.

Brinton, L.J. 2001. "From Matrix Clause to Pragmatic Marker: The History of *Look*-Forms". *Journal of Historical Pragmatics* 2(2):177-199.

Carlier, A. 2007. "From Preposition to Article. The Grammaticalization of the French Partitive". *Studies in Language* 31(1):1-49.

Chafe, W. 1982. "Integration and Involvement in Speaking, Writing, and Oral Literature". Tannen 1982. 35-53.

Davidse, K., Vandelanotte, L. and Cuyckens, H. (eds). 2010. *Subjectification, intersubjectification and grammaticalization*. Berlin/New York: Mouton de Gruyter.

De Mulder, W. 2001. "La linguistique diachronique: grammaticalisation et sémantique du prototype". *Langages* 130:8-32.

Dostie, G. and Pusch, C. 2007. "Introduction". *Langue française* 154:3-12.

Dostie, G. 1998. "Deux marqueurs discursifs issus de verbes de perception: de *écouter/regarder* à *écoute/regarde*". *Cahiers de lexicologie* 73/2:125-146.

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

Dostie, G. 2009. "Discourse Markers and Regional Variation in French. A Lexico-Semantic Approach". Beeching et al. 2009. 201-214.

Drescher, M. and Frank-Job, B. (eds). 2006. *Les marqueurs discursifs dans les langues romanes. Approches théoriques et méthodologiques*. Frankfurt am Main: Peter Lang.

Fleischman, S. and Waugh, L.R. (eds). 1991. *Discourse-Pragmatics and the Verb: the Evidence from Romance*. London: Routledge.

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

Fraser, B. 1988. "Types of English discourse markers". *Acta Linguistica Hungarica* 38:19-33.

Fraser, B. 1996. "Pragmatic markers". *Pragmatics* 6(2):167-190.

Gallardo Paúls, B. and Marín Jordà, M.J. 2005. "Marcadores discursivos procedentes de verbos perceptivos en el discurso afásico". *Revista de investigación lingüística* 8:53-94.

Giacalone-Ramat, A. and Hopper, P.J. (eds). 1998. *The Limits of Grammaticalization*. Amsterdam: Benjamins.

Hansen, M.-B.M. 1996. "Some common discourse particles in spoken French". Hansen and Skytte 1996. 105-146.

Hansen, M.-B.M. and Skytte, G. (eds). 1996. *Le discours: cohérence et connexion*. Copenhagen: Etudes Romanes/Museum Tusculanum Press.

Haspelmath, M. 1997. *From space to time: Temporal adverbials in the world's languages*. Munich/Newcastle: Lincom Europa.

Herslund, M. 2005. "Lingue endocentriche e lingue esocentriche: aspetti storici del lessico". Korzen and Marello 2005. 19-30.

Himmelmann, N.P. 2004. "Lexicalization and grammaticization: Opposite or orthogonal?". Bisang, Himmelmann and Wiemer 2004, 21-44.

Hopper, P.J. and Traugott, E.C. 2003 [1993]. *Grammaticalization*. Cambridge: Cambridge University Press.

Korzen, I. and Marello, C. (eds). 2005. *Tipologia linguistica e società*. Firenze: Franco Cesati.

Lamiroy, B. and Swiggers, P. 1991. "Imperatives as discourse signals". Fleischman and Waugh 1991, 121-146.

Lehmann, C. 1995 [1982]. *Thoughts on grammaticalization*. Munich: Lincom Europa.

Manili, P. 1985. "Sintassi di connettivi di origine verbale". *Gli annali della università per stranieri* 6:305-314.

Manili, P. 1989. "Relazioni tra aspetti morfologici e aspetti testuali di alcuni connettivi di origine verbale". *Parallela* 4:157-169 [= 1989, Benucci, 20 p.; references are to this version].

Marchello-Nizia, C. 2006. *Grammaticalisation et changement linguistique*. Bruxelles: De Boeck.

Marín Jordà, M.J. 2005. "Discurs i gramaticalització: verbs de percepció usats com a marcadors discursius en el debat electoral". *Caplletra* 38:47-71.

Marín, M.J., and Cuenca Ordiñana, M.J. 2000. "Verbos de percepción gramaticalizados como conectores: análisis constrastivo español-catalán". *Revista española de lingüística aplicada* Vol. Extra 1:215-238.

Martín Zorraquino, M.A. and Montolío, E. (eds). 1998. *Marcadores discursivos: teoría y práctica*. Madrid: Arco.

Oppermann, E. 2006. "Les origines du présentatif *voici/voilà* et son évolution jusqu'à la fin du XVIè siècle". *Langue française* 149(1):77-91.

Pons Bordería, S. 1998. "Los apelativos *oye* y *mira* o los límites de la conexión". Martín Zorraquino and Montolío 1998. 213-228.

Quirk, R., Greenbaum, S., Leech, G. and Svartvik, J. 1985. *A comprehensive grammar of the English language*. London: Longman.

Renzi, L., Salvi, G., and Cardinaletti, A. 1994. *Grande grammatica italiana di consultazione* (vol. III). Bologna: Il Mulino.

Rodríguez Somolinos, A. 2003. "Un marqueur discursif du français parlé: *écoute* ou l'appel à la raison". *Thélème, Revista Complutense de Estudios Franceses, Número Extraordinario*:71-83.

Rossari, C. 2006. "Grammaticalization and Persistence Phenomena in Two Hybrid Discourse Markers: *La preuve* and *Regarde*". *Acta linguistica Hafniensia* 38:161-179.

Schiffrin, D. *Discourse markers*. 1987. Cambridge: Cambridge University Press.

Stenström, A.-B. 1994. *An Introduction to Spoken Interaction*. London: Longman.

Tabor, W. and Traugott, E.C. 1998. "Structural Scope Expansion and Grammaticalization". Giacalone-Ramat and Hopper 1998. 229-272.

Tannen, D. (ed). 1982. *Spoken and Written Language: Exploring Orality and Literacy*. Norwood: Ablex.

Traugott, E.C. and Dasher, R.B. 2002. *Regularity in semantic change*. Cambridge: Cambridge University Press.

Traugott, E.C. 2007a. "Discussion article: Discourse markers, modal particles, and contrastive analysis, synchronic and diachronic". *Catalan Journal of Linguistics* 6:139-157.

Traugott, E.C. 2007b. "(Inter)subjectification and unidirectionality". *Journal of Historical Pragmatics* 8(2):295-309.

Traugott, E.C. 2010. "(Inter)subjectivity and (inter)subjectification: a reassessment". Davidse, Vandelanotte and Cuyckens 2010.

Waltereit, R. and Detges, U. 2007. "Different functions, different histories. Modal particles and discourse markers from a diachronic point of view". *Catalan Journal of Linguistics* 6:61-80.

Waltereit, R. 2002. "Imperatives, Interruption in Conversation, and the Rise of Discourse Markers: A Study of Italian *guarda*". *Linguistics* 40(5):987-1010.

Author's address

Benjamin Fagard
Laboratoire Lattice, CNRS
1 rue Maurice Arnoux
F-92120 Montrouge
France

benjamin.fagard@ens.fr

Notes

[1] This is, actually, a true example. All examples are given with a personal translation; here, as in all other translations, we try to follow the structure of the original language, and modify it only to obtain acceptable English translations; for DMs and other idiomatic elements, however, we did take a few liberties. Corpus examples are presented with their original annotations.

[2] Here and in the rest of our paper, we will (for the sake of simplicity) use the term "Romance" as referring to the Romance languages we took into account in our study; of course, what we say is not necessarily true for Sardinian, Galician, Rhaeto-romance and other Romance languages or dialects. Further research will, hopefully, fill the gaps.

[3] This has been noted for French vs. Italian by Rossari (2006) and Waltereit (2002:1008, footnote 6).

[4] For Italian, Spanish, Catalan, Portuguese and French; for Romanian, we relied on native speakers' judgments and benefited in addition from the advice of two colleagues, Alexandru Mardale and Ariadna Ştefănescu, which we here thank. We also would like to thank Salvador Pons Bordería, Sophie Prévost, Fabio Zinelli, Anna Mauri Graells, Laure Sarda and Paola Pietrandrea for their advice, as well as the anonymous reviewers for their very helpful comments.

[5] We also used the CFPQ corpus for Quebec French, with 136 occurrences of *regarde* for 160,000 words (by our count).

[6] As we did not, unfortunately, find the exact size or constitution of this corpus, this is only a rough approximation.

[7] We coded mainly meaning, pragmatic inferences, integrated vs. non-integrated use (presence of a complement, of a complementizer, etc.), internal vs. peripheral position, form and context (presence of (other) DMs), for all Catalan, French and Romanian occurrences, and for a selection of Portuguese, Spanish and Italian occurrences of *look*-forms. For these three languages, we arbitrarily extracted 150 occurrences from the respective subcorpora.

[8] For obvious reasons of space, we cannot illustrate all uses for all languages… We tried, as far as possible, to give examples for those cases that are not illustrated in the literature.

[9] I would like to thank one of the anonymous reviewers for pointing this out to me.

[10] With relative frequencies which remain however very low, even with the slight increase in the 19th century for *regarde donc* (1,6 occurrences for 1 million words, ten times as much as in the 16th century, in the same database).

[11] One anonymous reviewer suggested that uses similar to that in (11) could be found in Quebec French. This would indicate that *regarde* can appear in turn-initial position in this variety of French, and may therefore have a turn-taking use:

(A) *Elle n'a même pas fini sa soupe*
(B) ***Regarde****! dis-lui de la finir si tu veux bien*
(A) She hasn't even eaten up her soup
(B) **Look**! Tell-her to finish it, please.

[12] This use seems to be quite frequent in the CFPQ corpus, in exactly the same contexts: *j'ai dit regarde prends le temps de te faire soigner pis tout ça* "I said <u>look</u> take the time to get this taken care of and all that" (Subcorpus 6, section 12, p. 87, l. 1).

[13] Meaning that the intersubjective reading results only from the context of use: "unless a form-meaning pair has come to code intersubjectivity, we are not seeing intersubjectification (*-ation* being the important item here)"; such items might be said to have "pragmatically" but not "semantically" intersubjectified, "because they do not code semantic intersubjectivity" (Traugott 2010, sections 2 and 4).

[14] As one reviewer kindly pointed out, suggesting that this decrease of *regarde*'s DM uses might be linked to a competition with the DM uses of *écoute* "listen".

[15] Considered by the author to be "a free variant" of *ben* (Ibid:146).