

A COMPARATIVE CORPUS-BASED STUDY OF EUROPEAN PORTUGUESE DISCOURSE MARKERS *BOM* AND *BEM* AND FRENCH *BON* AND *BIEN*¹

Fátima SILVA², Fátima OLIVEIRA³, Françoise BACQUELAINE⁴

Article history: Received 2 August 2023; Revised 16 November 2023; Accepted 17 November 2023; Available online 20 December 2023; Available print 31 December 2023.

©2023 Studia UBB Philologia. Published by Babeş-Bolyai University.



This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License

- ¹ This research was supported by Centro de Linguística da Universidade do Porto, under the FCT Funding Programme - UIDB/00022/2020 (Foundation for Science and Technology).
- ² **Fátima SILVA** is an Associate Professor at the Faculty of Arts and Humanities of the University of Porto (FLUP), she teaches linguistics at the undergraduate, master, and doctoral level, as well as the Project Seminar and collaborates in the teachers training for the master's degree in Portuguese as a Second/Foreign Language. She is a Centre of Linguistics member, integrating the semantics group. Her research areas are text linguistics, lexical semantics, the application of linguistics to teaching and learning Portuguese as a non-native language, and the training of Portuguese teachers as a non-native language. She has collaborated on various research projects and served on scientific committees of conferences and journals in her research areas. She is the author and co-author of several articles published in national and international journals, conference proceedings, and books. E-mail: mhenri@letras.up.pt
- ³ **Fátima OLIVEIRA** is a full professor at the Faculty of Arts and Humanities and a researcher at the Centre of Linguistics of the University of Porto where she coordinates the semantics group. Her research is mainly on sentence and discourse semantics, particularly tense, aspect, modality, genericity, and indefinites, but also on some topics on the interface with syntax and pragmatics. She has coordinated or participated in national or international research projects, some in interface with other linguistics areas and other scientific domains. She has been a member of various conferences and journals' scientific committees in different linguistic areas. She is the editor of some books and author or co-author of several books chapters and papers published in national and international books, journals, and conference proceedings. E-mail: moliv@letras.up.pt
- ⁴ **Françoise BACQUELAINE** is an Assistant Professor at the Faculty of Arts and Humanities of University of Porto (Portugal) and a researcher at the Centre of Linguistics of the University of Porto. She teaches French linguistics, creating and exploiting corpora in language sciences, and general and specialised translation between French and Portuguese. She obtained a degree in Germanic Philology at the University of Liège (Belgium) and a Master's degree in Terminology and Translation at the University of Porto, where she also defended a PhD thesis in Language Sciences on human translation and machine translation of the Portuguese universal quantifier *cada* in four multi-word units to be translated *en bloc* into French and English. Her publications include translation, terminology, corpora, phraseology and universal quantification studies. She is also interested in translation didactics and discourse markers. E-mail: franba@letras.up.pt

ABSTRACT. *A comparative corpus-based study of European Portuguese discourse markers bom and bem and French bon and bien.* According to many authors, Discourse Markers (DM) serve as signals or triggers guiding the process of interpretation (Fraser 2006, Aijmer 2013, Maschler and Schiffrin 2015, among many others), thus having a more procedural than conceptual meaning. This study aims to investigate and specify the structural positions and the semantic-discursive functions of the DMs 'bem' and 'bom' in European Portuguese and of 'bon' and 'bien' in French, then to compare these positions and functions. The study relies on two European oral corpora for each language. The methodology is quantitative and qualitative. The analysis focuses primarily on the structural and modal levels, following mainly Oliveira and Silva's (2020) proposal for the study of 'bem' and 'bom' and the proposal of Peltier and Ranson (2020) for 'bon'. Results show that isolated DM 'bom' and 'bem' are equally frequent in the *C-ORAL-ROM* and that 'bem' is much more frequent than 'bom' in the local corpus *Fala Bracarense*. They also show that 'bien' is very rare as an isolated DM in both French corpora, while 'bon' is much more frequent than 'bom', 'bem' and 'bien'. Regarding structural positions, these DMs occur mainly as turn medial, while structural and modal functions are more differentiated depending on the DM, the corpus and the language. Thus, the study shows that although these European French and Portuguese DMs share the same etymology, they differ in usage.

Keywords: *Discourse markers, European Portuguese, French, oral corpora, semantic-discursive functions*

REZUMAT. *Un studiu comparativ de corpus al marcatorilor discursivi bom și bem în portugheza europeană, respectiv bon și bien în franceză.* Potrivit multor autori, marcatorii discursivi (MD) au rolul de semnale sau declanșatori care ghidează procesul de interpretare (Fraser 2006, Aijmer 2013, Maschler and Schiffrin 2015, printre mulți alții), având astfel o semnificație mai mult procedurală decât conceptuală. Acest studiu își propune să investigheze și să precizeze pozițiile structurale și funcțiile semantico-discursive ale MD „bem” și „bom” în portugheza europeană, respectiv „bon” și „bien” în franceză, iar apoi să compare aceste poziții și funcții. Studiul folosește două corpusuri orale europene pentru fiecare limbă. Metodologia este cantitativă și calitativă. Analiza se concentrează în primul rând asupra nivelurilor structural și modal, urmând în principal propunerea lui Oliveira and Silva (2020) pentru studiul lui „bem” și „bom” și propunerea lui Peltier and Ranson (2020) pentru „bon”. Rezultatele arată că MD „bom” și „bem” au aceeași frecvență în *C-ORAL-ROM* și că „bem” este mult mai frecvent decât „bom” în corpusul local *Fala Bracarense*. De asemenea, analiza arată că „bien” este foarte rar ca MD izolat în ambele corpusuri, în timp ce „bon” este mult mai frecvent decât „bom”, „bem” și „bien”. În ceea ce privește pozițiile structurale, MD apar în principal cu rolul de turn-taking, în timp ce funcțiile structurale și modale sunt mai diferențiate în funcție

de MD, de corpus și de limbă. Astfel, studiul arată că, deși acești MD din franceză și portugheza europeană au etimologie comună, prezintă întrebunițări diferite.

Cuvinte-cheie: *Marcatori discursivi, portugheză europeană, franceză, corpusuri orale, funcții semantico-discursive*

1. Introduction

Discourse Markers (DMs) are broadly considered as expressions whose primary function is guiding the process of interpretation (Fraser 2006; Aijmer 2013; Maschler and Schiffrin 2015, among many others), thus having a more procedural than conceptual meaning. As a non-homogeneous class in terms of its composition, encompassing elements from various grammatical categories, which have typically undergone processes of grammaticalisation (such as adverbs, adjectives, conjunctions, and verbal phrases), these markers share common features, namely being invariant and polyfunctional, having variable mobility within discourse, and contributing to its structuring and management. Some DMs primarily establish connections between two discourse segments, occurring in both written text and spoken discourse, contributing to their cohesion and coherence. In contrast, others, predominantly used in spoken discourse, operate at the level of structural and interactional management.

The DMs 'bem' and 'bom', in European Portuguese (EP), and 'bien' and 'bon', in French⁵, fall within this latter type, also referred to by some authors as conversational markers (Schiffrin 1987; Urbano 2003; Rodrigues 1998; Lefevre 2011) or pragmatic connectors (Cuenca 2013). Categorised in their origin as adverbs ('bem', 'bien') and adjectives ('bom', 'bon'), they retain their prototypical usage and can take on other uses, notably as discourse markers in spoken discourse.

While some studies in EP focus on the DM 'bem' separately, such as those by Lopes (2004) and Valentim (2008), or analyse both 'bem' and 'bom' within the context of a global discourse marker analysis, as seen in Cabarrão et al. (2018), systematic comparative research on these two markers remains limited. However, there are at least two studies in EP that offer a comprehensive comparison, namely Morozova (2019) and Oliveira and Silva (2020), and several more studies in Brazilian Portuguese (BP), including works by Risso (1999) and Gorski (2020). Regarding French, there has been mainly research

⁵ Throughout the text, the translation into English of the markers and segments does not have a contrasting objective, but rather, it is done to facilitate the reader's comprehension.

on 'bon' and its correlates (a.o. Peltier and Ranson 2020), while literature on the DM 'bien' in isolation is significantly reduced (Moline 2012). On the other hand, 'bem' and 'bom' have already been compared to English DM 'well' (Morozova 2019, 2020; Oliveira and Silva 2020). There has also been work comparing 'bem' and 'bien' (Lejeune and Valentim 2015); however, in this case, the focus is not exactly on the analysis of these two expressions as conversational markers but primarily follows Péroz's (1992) analysis of 'bien' and aims to describe their most significant values.

To our knowledge, no research has yet compared these four DMs. This study aims to contribute to filling that gap by investigating the semantic-discursive functions of the DMs 'bem' and 'bom' and comparing them with 'bon' and 'bien' as single units. Therefore, we do not consider the contexts in which these DMs occur in combination with other connectives or discourse markers in French, such as *eh bien, et bien, bon ben, mais bon, euh bon, parce que bon, et puis bon* (Beeching 2007, 91; Peltier and Ranson 2020, 13-17), and in EP, namely *ah bem, ah bom, mas bem, pois bem, ora bem* (Cabarrão et al. 2018; Ponce de Léon and Duarte 2021).

To achieve the intended goal, the study relies on two oral corpora for each language, manually annotated. It adopts a quantitative and qualitative methodology, mainly following Oliveira and Silva's (2020) proposal to analyse 'bem' and 'bom', and Peltier and Ranson's (2020) for 'bon', although also resorting to the contribution of other authors.

The paper has the following structure. In Section 2, we present a summary of related work on these DMs; in Section 3, we describe the corpus; in Section 4, we explain the methodology; in Section 5, we present the categorisation scheme; and, in Section 6, we give the data analysis and discuss the results, concluding with some final remarks in Section 7.

2. Related work

The analysis of 'bem', 'bom', 'bien' and 'bon' as DMs has been done within the scope of various theoretical frameworks, resulting in different proposals and analysis models with varying degrees of proximity or dissimilarity. For this work, we provide a very brief overview of the characterisation of these markers, emphasising the studies that have supported the classification scheme and the approach adopted in the analysis of DMs in this study. We begin with the Portuguese DMs and then move on to the French ones.

For EP, Lopes (2004) refers to 'bem' as a DM in a study in which she describes the polyfunctionality of 'bem' and points to its following discourse or pragmatic functions: disagreement/non-acceptance, turn initiating, topic change, and discourse initiator. Though its relevance is evident, this analysis remains somewhat underdeveloped.

Valentim (2008) extensively analyses the pragmatic functions of the DM 'bem' within dialogues, highlighting a strong correlation between these functions and the diverse semantic values they acquire, determined by their relationship to either the speaker or the ongoing discourse. Concisely, the prime functions include: rectification through reformulation, expression of a certain disagreement, attenuation of disagreement, signaling the reception of a message with an associated appreciation (positive or negative), the introduction of a response to a question, comment on prior discourse or intervention, with the possibility of expressing uncertainty, signalling the opening of a conversation or its pre-conclusion, a change of topic and thematic continuity.

Also focusing on spoken language, Morozova (2019) and Oliveira and Silva (2020) compare 'bem' and 'bom' as DMs⁶, examining their performance regarding discourse organisation and functions performed by each. In both studies, these discourse markers were compared to the English discourse marker 'well,' leading to the consistent conclusion that 'bem' is closer to 'well' than 'bom,' although they may not always be equivalent.

Morozova (2019) examines the use of 'bem' and 'bom' analysing the role of discourse markers in the textual organisation of stand-up comedy in Portugal and the United States. The author categorises these expressions into non-discursive and discursive uses, further differentiating the latter into textual and interactional functions. Textual functions refer to the relationships between text segments, while interactional functions are related to the pragmatic orientation of discourse. The study results showed that both discourse markers lack interactional functions. 'Bem' exhibited seven textual functions, including signalling conclusion, topic shift, discourse continuity, topic initiation, shift between direct and indirect speech, and repair of a previous discourse, while 'bom' displayed only three functions: signalling delay, introducing a new topic, and continuing the current topic. Furthermore, she examines their positions within the macro-textual organisation and concludes that 'bom' is primarily used in openings, while 'bem' is predominantly found in the body of the text, although it also appears in openings and closings.

⁶ While the study focuses on the European variety of Portuguese, it's worth noting that several Brazilian Portuguese studies analyse 'bem' and 'bom' as DMs, namely those of Gorski (2020) and Risso (1999). Gorski (2020), within a functionalist framework and based on the analysis of dialogical contexts in sociolinguistic interviews, conducted a comparative study of 'bem' and 'bom,' aiming to establish contextual patterns and analyse the grammaticalisation processes. Risso (1999) adopted an interactive textual approach and identified a set of functions frequently performed by these discourse markers: opening markers used by speakers to gain time, intra-topic opening markers, which impact the ongoing information, citation, topic resumption, and concession contributing to managing different viewpoints between two interlocutors.

Oliveira and Silva's (2020) qualitative study, in turn, utilised samples from various discourse genres and corpora to identify usage patterns and specify the semantic-discursive functions of 'bom' and 'bem'. Following Cuenca's (2008) proposal for 'well' and taking into account the data, DMs 'bom' and 'bem' were analysed according to a two-level scheme – structural and modal. The first was composed of five functions – beginning of interaction, beginning of turn, preclosing of interaction, topic change and pause, all related to discourse organisation. The second, corresponding to the management of interaction and interpersonal relation regulation, was divided into four sections: mitigation, disagreement (partial), reformulation, and request for clarification or specification. The study results enabled the establishment of some significant similarities but also differences between the two DMs. Summing up, 'bem' is more frequent than 'bom', it occurs in all structural and modal functions, it may establish a relationship not only with what was said previously (as a reaction to the preceding segment) but also with what will be said next, creating the expectation for a new statement, and it occurs frequently in contexts where some negativity may be expected. On its part, 'bom' presents less structural flexibility, it usually relates to what has been previously said and may exhibit a reaction to it.

Concerning French, 'bon' as a DM is relatively common, but 'bien' seems to be much less used. This is reflected in the number of studies on the first one and the near absence of studies on the latter, as we have already pointed out. The literature on 'bon' is quite diverse and the authors propose different interpretations. Some distinguish between interjective use, which is turn initial and occurs at the end of a conversational sub-sequence, and "discourse marking use 'proper'" which is turn-internal (Hansen 1998, 253-4, 257). Others, like Gilbert (2019), consider that most extrasyntactic units occur in a turn-medial position between two complete syntactic structures.

Based on the idea that 'bon' signals to the interlocutor a reorientation or acceptance and on the mentions of this DM functions in literature, Peltier and Ranson (2020) list nine textual functions, which are organised on two main macro-functions – opening and continuation – and a third macro-function designed as other textual functions. Besides, the authors also propose two attitudinal functions – contrast and acceptance or resignation.

In this study, the macro-function 'opening' is divided into three functions: new topic, topic taking, and new point of view. So, 'bon' is used to introduce a new topic of conversation. However, Barnes (1995, 815) assigns to 'bon' a function of closing the previous speech and calling the interlocutor's attention. Brémond (2003, 74) has a similar position. She also considers that 'bon' marks a stage and enables a global move backwards but that even instances where

'bon' is used to signify the impending closure of an interaction primarily denote the shift towards a new activity, creating the potential for a break in the exchange (Brémond 2004, 9). Jaz (2004, 4) also proposes that 'bon' conveys a conventional implicature that the agent believes or desires that the process in progress is terminated.

As for the text functions of 'bon' grouped under the heading of continuation, which include sub-theme, result, supplement and additive element, Peltier and Ranson (2020) consider that they mark a reorientation within an already established topic.

Regarding the DM position, 'bon' is mostly found in the middle of a sentence, but, more than half of the occurrences that introduce a new theme are at the beginning of the speaking turn.

As for 'bien' as a discourse marker, Moline (2012) acknowledges that it can be a DM, but she does not propose any other information regarding structural position or function.

3. Corpus

In this study, we used the Portuguese sub-corpus from C-Oral-Rom: Integrated Reference Corpora for spoken Romance Languages (Bacelar do Nascimento et al. 2005), a multilingual corpus of Spanish, Portuguese, French, and Italian, made available with transcription and alignment using the EXmaRaLDA program (cf. Mendes 2016)⁷, and PSFB: Perfil Sociolinguístico da Fala Bracarense (Sociolinguistic Profile of Braga Speech) (Marques and Aguiar 2014; Mendes 2016), on the one hand, and two subcorpora (the French subcorpus C-ORAL-ROM⁸ and CLAPI: Corpus de Langues Parlées en Interaction) from C.E.F.C.: Corpus d'Etude pour le Français Contemporain (Study Corpus for Contemporary French) resulting from the Orféo (Outils et Recherches sur le Français Écrit et Oral) project (Benzitoun et al. 2016; Debaisieux and Benzitoun 2020; ATILF 2021), on the other hand.

Although they present characteristics that distinguish them from each other, notably concerning their size, construction objectives, production context, and genres they encompass, these corpora have in common the fact that they are oral, mainly interactional, and therefore constitute a significant source of occurrences of the discourse markers under analysis. Additionally, at least two of the corpora are comparable since they are in the scope of the same project, C-ORAL-ROM: the EP and FR sections of C-ORAL-ROM present identical designs, criteria, and transcription procedures.

⁷ From now on COR_EP.

⁸ From now on COR_FR.

The criteria for the corpora selection were: i) access to the complete corpus or sub-corpus; ii) availability of orthographic transcription of the text and audio files; iii) occurrence of various genres; iv) representation of language varieties.

COR-EP and COR_FR consist of spontaneous or semi-spontaneous speech produced by speakers with diverse sociolinguistic profiles, predominating standard Portuguese and French. They present a differentiated structural organisation (monologue, dialogue, and conversation), distribution according to the context (formal or informal), the domain of use (familiar, public, media, and natural context), and modality (face-to-face, telephone, media). COR_EP is 29:51:00 long, while COR_FR is 22:17:09 long, each containing approximately 300,000 words.

The PSFB corpus, also with transcription and alignment using the EXmaRaLDA program and made available in its entirety, focuses on the variety of Portuguese spoken in the city of Braga. It comprises a total of 81:51:00 spontaneous speech and consists of 80 semi-planned face-to-face interviews conducted with speakers of both sexes, with varying levels of education and age groups. It belongs to the public domain, presenting a semi-formal and informal register domain.

Finally, CLAPI “is a multimedia database containing corpora that have been recorded in various real situations, such as workplace, institutional, private, commercial, medical and educational interactions” (CLAPI, <http://clapi.icar.cnrs.fr>). Here, we used the subset of the database made available to the Orféo project. It consists of 17 recordings for 16:45:41 and approximately 170,000 words.

4. Methodology

French data were extracted using the tools available on Orféo’s platform. The concordance search (KWIC) was used to extract all occurrences of ‘bien’ and ‘bon’; then the occurrences of ‘bien’ and ‘bon’ as DM were manually extracted, and, finally, we retained the occurrences in which ‘bien’ and ‘bon’ occur as DM as single units. To do so, we used the automatic transcriptions available. These automatic transcriptions do not include any punctuation marks or any indication of pauses and are of low quality when turns of speech overlap or the recording is poor. To fill these gaps and improve the quality of the transcriptions, we listened to the audio files available on Orféo’s platform. The Portuguese data were obtained using the keywords ‘bem’ and ‘bom’ for each audio and transcription on the EXMARaLDA application. Then, there was manual extraction of the occurrences of ‘bem’, ‘bom’, ‘bien’ and ‘bon’ as DMs, followed by the extraction of these DMs occurring as single units, with sufficient context for their interpretation.

The data were classified according to a categorisation scheme that took into account previous analyses, in particular those of Cuenca (2008), Oliveira and Silva (2020), and Peltier and Ranson (2020), but was adapted considering the data from the corpora. The final version of this scheme resulted from the initial annotation by the authors of this article of a 10% sample of the selected occurrences, followed by an analysis of the inter-agreement among the three annotators, with a discussion of divergent classification proposals and consideration of specific cases not occurring in other studies. Two of the three annotators handled the remaining part of the data, with the third annotator's involvement whenever issues with the classification arose.

The analysis was based on the collection of quantitative data regarding the structural positions, structural functions, and modal functions of the markers under study, and the qualitative analysis of these different functions in each sub-corpus. Then, the uses of the DMs were compared within each language, followed by a cross-linguistic comparison of the results in Portuguese and French.

5. Categorisation scheme

The categorisation scheme and chosen tags are a revised proposal based on Oliveira and Silva (2020), that follows Cuenca (2008), driven by three main reasons: i) the necessity to account for the emerging data from the corpora and the correlations derived from them in the analysis of discourse markers in EP and French; ii) reorganisation of the levels of analysis, and iii) an expansion of functions for each level, with special consideration given to Beeching's (2011) and Peltier and Ranson's (2020) proposals, along with the contributions of Valentim (2008) and Morozova (2019).

In contrast to the two-level analysis (structural and modal) proposed by Oliveira and Silva (2020), this study adopts a three-level analysis, which involves additional subdivision of the first level into structural positions and functions. Although the focus remains on discourse structuring and organisational management levels, which can have an interactive or monologue configuration, this partition allows for distinguishing the position of the DM in the structure and its thematic-informational function. Consequently, position and function do not have to be in complementary distribution; they can co-occur because they belong to different levels.

As Cuenca (2008) asserts for the discourse marker 'well', we can consider that structural positions and functions may be intertwined with modal aspects to convey different meanings. In fact, the structural functions differ from modal functions in that the former operate at the textual level, primarily presenting a

metadiscursive or discursive frame value, while the latter function at an interpersonal or interactional level, having a qualifying function concerning what is said and the speaker's emotions and attitudes towards it.

5.1. *Structural positions*

Concerning the structural level, we considered five structural positions for both groups of DMs – beginning of interaction ((1)), beginning of turn ((2)), turn medial ((3)), turn final ((4)) and (pre-ending) an interaction or an exchange ((5))⁹¹⁰.

- (1) MAR: *bon* alors raconte-moi ton week-end
 CHA: *bon* alors non le week-end euh je suis rentrée chez moi (COR_FR, ffamdl01)
 MAR: *So* tell me about your weekend
 CHA: *So*, no, at the weekend, um, I went home

- (2) Ent [v] •• O sol aumenta a produção de se/ de serotonina.
 Fal21 [v] •• *Bem*, não sei o que é, mas calculo que isso seja uma coisa boa [PSFB, 21H2D]
 Ent [v] •• The sun increases the production of se/ serotonin.
 Fal21 [v] •• *Well*, I don't know what it is, but I suppose that's a good thing.

- (3) PAT: [...] je serai dans le sud pendant les vacances mais en allemagne c' est plus quand même
 JUD: ouais ça doit b / *bon* ça sera pas le même budget en Allemagne [CLAPI, aperitif_glasgow]
 PAT: [...] I will be in the south during the holidays, but in Germany, it's more expensive though
 JUD: *Well*, it must b / be good, but the budget won't be the same in Germany.

- (4) SOP: [...] tu vois à Aix tu connais l' Unplugged c' est ben d' ailleurs c'est à côté du Sunset
 ANT: ouais donc *bon* [COR_FR, ffamcv02]
 SOP: [...] You know in Aix, there's the Unplugged, it's good; by the way, it's next to the Sunset.
 ANT: Yeah, well, alright then.

⁹ These positions do not necessarily overlap with the DM syntactic position within the text flow.

¹⁰ For each described function, we offer an illustrative example chosen from one of the sub-corpora without providing illustrations for all sub-corpora and discourse markers.

- (5) Ent [v] •• *Bem*, •• terminava por aqui. Agradeço, mais uma vez, muito por por nos ter ajudado, por nos ter dado esta... [PSFB, 21H2D]
Ent [v] •• *Well*, •• that's it for now. Thank you, once again, very much for helping us, for giving us this...

5.2. Structural functions

For structural functions, we considered the following: topic initiation, topic change, topic recovery, elaboration, conclusion, response initiator, and filler.

The topic initiation function pertains to situations where the marker introduces a discursive topic. It occurs when the discourse marker appears at the onset of the interaction, aligning with the concept identified by Estellés and Pons Bordería (2014) as the 'absolute initial position', as in (6), and thus playing a foreground role. However, we have also considered this function in cases where, although not in the absolute initial position, it was only preceded by another word or expression that does not introduce the topic.

- (6) Ent [v] •• *Bem*, então, muito obrigada, realmente por teres aceiteado dar-nos esta entrevista, por ter dado um tempinho. • • ((hesitação)) Eu queria começar por perguntar, tu sempre viveste na Sé, correto?
Int [v] •• *Well* [v] •• So, thank you very much for accepting to give us this interview, for taking the time. • • ((hesitation)) I wanted to start by asking, you have always lived in the Sé, right?

When introducing a change of topic or subtopic, these DMs signal a transition to a new informational node, with either a complete change from the previous topic ((7)) or the initiation of a new subtopic derived from the ongoing topic ((8)). This change can be textually supported either explicitly, as in ((9)) with the segment 'changing the subject' following the DM, or implicitly ((7)).

- (7) Ent [v] •• Eu, por acaso, tenho um bocado de medo de agulhas, é por isso que me faz um bocado de aflição
Fal43 [v] Não não não. Por acaso, não... É um sistema que não • • que não me incomodou demasiado.
Ent [v] Hum, hum. ••• ((hesitação)) *Bem*, eu queria perguntar-lhe o que é que o levou a seguir a vida religiosa? [PSFB, 43H4D]
Ent [v] •• I, actually, have a bit of a fear of needles; that's why it makes me a bit anxious.
Fal43 [v] No, no, no. Actually, no... It's a procedure that didn't • • didn't bother me too much.
Ent [v] Hmm, hmm. ••• ((hesitação)) *Well*, I wanted to ask you what led you to pursue religious life?

- (8) la nature existait bien avant nous et lorsque nous aurons disparu j'espère pour elle qu'elle existera toujours c'est-à-dire que nous ne l'aurons pas trop abîmée / *bien* // la haute montagne est un élément essentiel de cette nature [COR_FR, ffamnm20]
Nature existed long before us, and when we are gone, I hope for its sake that it will still exist, meaning that we haven't damaged it too much / *well* // the high mountains are an essential element of this nature.
- (9) Fal21 [v] •• Não conosco, mas com outras pessoas que acabaram por acabar essas brincadeiras. •• Mas, no geral, •• foi foi positivo.
Ent [v] Foi positivo. ••• *Bem, mudando de assunto.* ••• ((hesitação))
Imagine, por exemplo, ••• que ganhava o Euromilhões. [PSFB, 21H2D]
Fal21 [v] •• Not with us, but with other people who eventually stopped those games. ••• But, overall, •• it was positive.
Ent [v] It was positive. ••• *Well, changing the subject.* ••• ((hesitation))
Imagine, for instance, ••• winning the Euromillions.

With a topic recovery, the DM indicates that the speaker intends to revisit a topic introduced earlier, which may be more or less distant from the moment of resumption. While this revisiting may signal the conclusion of the current topic to return to another previously present in the discourse, it does not represent a topic change. The DM instructs that what follows should be understood as a return to a topic previously introduced, initiating a backward movement from which the discourse proceeds ((10)).

- (10) SOP: c' est l' histoire d' Hansel et Gretel
FRA: donc c' est la sorcière qui est dans l' histoire d' Hansel et Gretel // *bien* // donc Quenotte veut sortir du livre mais elle n' y arrive pas donc on va voir maintenant qu' est-ce qui va se passer [COR_FR, fpubcv07]
SOP: It's the story of Hansel and Gretel.
FRA: So, the witch is part of the story of Hansel and Gretel. *Well*, Quenotte wants to escape from the book, but she can't manage it. So, let's see what will happen next.

Elaboration, in turn, occurs when the DM signals that the speaker is expanding, developing, or delving into a topic or idea previously presented in the discourse, indicating topic continuity. In that sense, it functions as a foreground and background textual reference, as illustrated in (11), where LUR uses 'bem' to signal she will continue the topic she's talking about.

- (11) LUR: hhh / entrou-me no [//] a tua mãe + &ah / bem / a minha não me larga / e quando me perde de vista ... pegaram-se as duas porque a minha mãe me perdeu-me de vista / e não saía de onde estava sem me ver //
FBA: ai meu Deus //
LUR: *bem* / mas o melhor foi a tua mãe entrar / no [//] no Mike and Spencer / e diz-me logo assim // eu não posso com grandes superfícies [COR_PT, ptelpv01]
LUR: hhh / your mother entered my [//] + &ah / well / mine won't leave me alone / and when she loses sight of me... the two of them started arguing because my mother lost sight of me / and she wouldn't leave where she was without seeing me //
FBA: oh my God //
LUR: *well* / but the best part was your mother entering / at [//] Mike and Spencer / and she tells me right away // I can't handle big stores

Conclusion is a structural function that draws its functionality from its core meaning. It introduces a segment assumed to be the culmination of preceding content, as exemplified in (12), pointing backwards and signifying that what follows is either the endpoint or a summary of the topic previously addressed by the speaker.

- (12) CAT: chorei / baba e ranho / nessa noite / quando estava deitada / porque aquele ano / que tinha sido tão bom / tinha acabado // e podia ser agora que / por mudar de dia
SAN: hhh / because that's the way things are / aren't they//
CAT: *well* / it was a very good year // [COR_PT, pfmadl32Hf]
CAT: I cried / drool and snot / that night / when I was in bed / because that year / that had been so good / was over // and it could be now that / because of changing days

The function response initiator (cf. Cuenca 2008) is primarily noticeable in interactions. However, it may also occur in monologues, indicating that there is a follow-up to the speaker's stimulus by the interlocutor. In this sense, it acts as a preface to a response, not only signifying that the interlocutor is willing to react to the stimulus launched, with potential modal nuances but also providing time to organise what they are going to say (Peltier and Ranson 2020: 6). In (13), the DM introduces the speaker's reaction to the interlocutor's stimulus, expressed in the form of a tag question, but also an additional modal value of mitigation of the illocutionary force of the question by initiating with the topic contextualisation.

- (13) PAU [v] senhor professor / além de ficarmos todos muito zangados / com a Samsung / o que é que o estado vai fazer / para recuperar / &eh / pelo menos uma parte daquilo que investiu / não é ?
 CAS [v] *bom* / o [//] como sabe / este contrato> [//] &e / este investimento tem [//] é um investimento de regime contratual // há um contrato subscrito / entre o estado português e [/] e a empresa / e as casas-mãe / cada uma delas // [COR_PT, pnatbu02]
 PAU [v] Dear professor / besides all of us getting very upset / with Samsung / what is the government going to do / to recover / &eh / at least a part of what it invested / isn't it?
 CAS [v] *well* / as [//] you know / this contract> [//] &e / this investment has [//] is an investment under a contractual framework // there is a signed contract / between the Portuguese government and [/] and the company / and their parent companies / each one of them //

These DMs can also be used to pause discourse, frequently employed to organise forthcoming speech. To some extent, they correspond to the hesitation typically associated with a brief interruption of information flow (see (14)).

- (14) Ent [v] •• Hum, hum. ••• *Bem*. ••• ((hesitação)) •• Olha, como é que vais passar o Natal •• este ano? [PSFB, 48M1B]
 Ent [v] •• Hmm, hmm. ••• *Well*. ••• ((hesitation)) •• Look, how are you going to spend Christmas •• this year?

5.3. Modal functions

Moving on to modal functions, we encounter the following in the corpora: partial or total agreement or disagreement, mitigation, reformulation, request for clarification or specification, point of view change, and digression.

The DMs can signal partial or total agreement or disagreement. In cases where the speaker disagrees or does not accept what was previously stated or implied by the interlocutor, the DM can introduce a softening of this disagreement on the speaker's part. (15), (16) and (17) illustrate respectively a case of total agreement, partial agreement, and disagreement.

- (15) Fal77 [v] •• Ela não tem quase uma/ •• um ((hesitação)) uma coisita
 Ent [v] Uma pecita?
 Fal77 [v] É.
 Ent [v] Não tem, não tem. Nada.
 Fal77 [v] Não tem, não. Tudo pelo liso. •• Ela ela ela só diz: - Ó mamã, eu limpo o pó em menos de nada porque é...
 Ent [v] *Bem* isso é verdade. ((risos)) [PSFB, 77M3D]

Fal77 [v] She hardly has a... a little thing (hesitates)
Ent [v] A little thing?
Fal77 [v] Yes.
Ent [v] She doesn't have, she doesn't have. Nothing.
Fal77 [v] She doesn't have, she doesn't. Everything for the smooth •• She, she, she just says: 'Oh, Mom, I can dust in no time because it is...
Ent [v] *Well*, that's true. (laughs)

- (16) Ent [v] •• Verdade. •• Mas essa questão de de as pessoas se preocuparem só em desenrascarem-se, parece-me um bocado egoísta. •• Quer dizer, primeiro as pessoas preocupam-se em em desenrascarem-se ela e não se preocupam
Fal44 [v] Bem.
Ent [v] com o resto.
Fal44 [v] •• *Bem*, isso isso tem muita razão no que diz. •• Mas, se vamos a pensar assim, •• então, nessa altura, •• até no caso, •• por exemplo, de •• de amizade, de conhecimentos •• é tudo só para mim, nada para os outros. [PSFB, 44H4D]
Ent [v] •• True. •• But this issue of of people only worrying about getting by seems a bit selfish to me. •• It means that people first worry about getting by and don't worry about
Fal44 [v] Well
Ent [v] the rest.
Fal44 [v] •• *Well*, that makes a lot of sense in what you're saying. •• But if we think like that, •• then, at that point, •• even in the case, •• for example, of •• of friendship, acquaintances, •• it's all just for me, nothing for others."

- (17) BAP [v] boa noite // Georges Méliès / ilusionista / prestidigitador / [...] / dizia / o seguinte // " o cinema não é para ganhar dinheiro // o cinema / é para ajudar o homem / a sonhar " // *bom* // como se sabe / não é bem assim // [...] [COR_PT, pmdin03]
BAP [v] good evening // Georges Méliès / illusionist / magician/ [...] / used to say // "cinema is not about making money // cinema / is about helping people / to dream" // *well* // as we know / that's not quite true.

The modal function of mitigation does not necessarily imply a disagreement with the speakers' interlocutor. It may also indicate mitigation regarding negation, refusal, or objection, where the DM attenuates the illocutionary force of what was previously said (Schiffrin 1987; Oliveira and Silva 2020; a.o.). Example (18) illustrates this function: the MD mitigates the illocutionary force of the first assertion in which it is stated that there are no different genres of music.

(18) il n' y a pas de différents genres musicaux à mon sens il y a les musiques qu'on aime et celles qu' on n' aime pas c' est tout / *bien* // on peut aimer aussi bien on peut aimer euh de l' opéra on peut aimer Johnny Hallyday on peut aimer Georges Brassens on peut aimer les entendre à différents moments [COR_FR, ffammn16]

In my opinion, there are no different musical genres; there are the music we like and the ones we don't, that's it. *Well*, you can like both opera, Johnny Hallyday, Georges Brassens, and enjoy listening to them at different times.

Although it is infrequent, there are cases of reformulation introduced by the DMs 'bom' and 'bem'. It may involve repairing, correcting, or altering a previous statement, potentially affecting its content or the very act of enunciation (cf. Oliveira and Silva 2020, 216), and, therefore, conveying the speaker's point of view about it. In (19), the speaker reformulates the previous segment ('todos os dias') by adding additional information regarding the quantifier 'todos'.

(19) LAL [v] pois / o problema é esse //
 MOT [v] e [/] e eu já tenho as
 LAL [v] não é?
 MOT [v] cassetes
 MOT [v] cá em casa e não sei quê // agora vou ter que começar aos serões
 // todos os dias // *bom* / também não são muitos //
 LAL [v] pois //
 MOT [v] são para aí seis ou sete // [COR_PT, ptelpv09]
 LAL [v] well / that's the problem //
 MOT [v] and [/] and I already have the
 LAL [v] isn't it?
 MOT [v] tapes
 MOT [v] at home and all that // now I'll have to start in the evenings //
 every day // *well* / they're not many either //
 LAL [v] right //
 MOT [v] there are about six or seven //

Oliveira and Silva (2020, 216) point out a modal function in which the DM signals the introduction of a request for clarification or specification, typically expressed by an interrogative utterance. In (20), JOS introduces a request for clarification using 'bon', relating NAT intervention.

(20) NAT: ma mère une fois elle a fait ça euh pour pour regarder où est-ce qu'
 il y avait une fuite de gaz tu sais elle a mis une allumette dans le four [...]
 JOS: // *bon* tu as pas fait ça toi?
 MAT: non j' ai pas fait ça [...] [C-Oral-Rom > ffamcv05]

NAT: my mother once did that uh to look where there was a gas leak you know she put a match in the oven [...].
JOS: // *Well*, you didn't do that, did you?
MAT: no, I didn't [...].

Another occurring modal function is the point of view change, which consists of the fact that the DM signals the beginning of reported speech, introducing a new point of view, either the speaker's (see 21) or someone else's, while reproducing the quotation in an inexact manner (cf. Peletier and Ranson 2020). Using the analysed DMs may offer a means to justify a decision or create a more vivid ambience within the discourse.

(21) PAT: [...] on corrigeait le contrôle et le prof il dit *bon* vous voulez qu' on corrige en français ou en allemand moi je dis ben en français enfin [...] [CLAPI, aperitif_glasgow]
PAT: [...] We were correcting the test, and the teacher said well do you want to correct it in French or in German? I said *well* in French of course...

The modal function of digression (22) typically corresponds to a type of parenthesis or interruption related to the current topic, textually signalled. Here, the speaker can offer additional information through a comment, evaluation, or expression of emotion related to what is being discussed (cf. Peltier and Ranson 2020, 8).

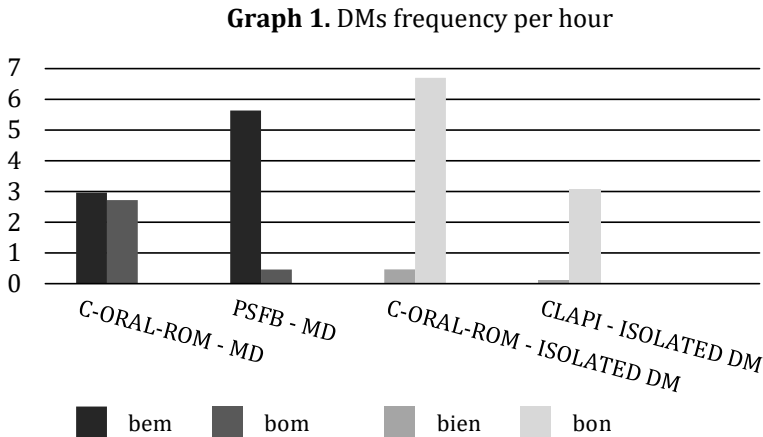
(22) nem sabia muito bem // enfim // *bem* / sou completamente louca / tresloucada // nem disse nada lá em casa // hhh vão achar que eu sou doida // então a minha mãe / RQL [v] que só acha que eu devia era fazer / mestrado / ou concorrer ao CEJ // aquelas coisas de mãe // [COR_PT, pfmadl17]
I didn't even know very well // anyway // *well* / I'm completely crazy / crazy // I didn't even say anything at home // hhh they're going to think I'm crazy // so my mother / RQL [v] who only thinks I should do / master's degree / or apply for CEJ // those mother things //

6. Results and discussion

In this section, we present the primary results of the data analysis, beginning with a global overview of the distribution of the DMs in the four sub-corpora, followed by an analysis of data concerning structural positions and structural and modal functions while engaging in a discussion on the presented results.

6.1. Global results

The global overview of the DMs analysis is presented in Graph 1. The frequency of the DMs is measured per hour and not per word because there was no indication of the number of words in the case of PSFB, but all corpora had their total number of hours.



These data seem to indicate that the standard or regional variant of the sub-corpus and the variety of textual genres in the sub-corpora are not, at least in this case, a decisive factor in justifying the observed tendencies in the two languages. However, since the two EP sub-corpora differ in terms of the variants and the genres they encompass, this difference may be relevant. This hypothesis will need to be further explored to be validated.

6.2. Results of the DMs positions and functions per sub-corpus

Next, we examine the results per sub-corpus, considering the distribution of DMs for structural positions, structural functions, and modal functions.

Table 1 presents the distribution of structural position in each sub-corpus.

Table 1. DMs Structural Positions per sub-corpus (frequency per hour)

Structural Positions	COR_EP		PSFB - EP		COR_FR		CLAPI - FR	
	bem	bom	bem	bom	bien	bon	bien	bon
Beginning of interaction	0,00	0,07	0,04	0,00	0,00	0,04	0,00	0,00
Beginning of turn	1,04	1,07	0,79	0,09	0,04	1,08	0,06	1,79
Turn medial	1,88	1,57	0,98	0,09	0,58	7,63	0,12	3,34
Turn final	0,00	0,00	0,00	0,00	0,00	0,22	0,00	0,18
(Pre-)ending of interaction	0,07	0,00	0,24	0,00	0,00	0,00	0,00	0,06

Starting with the EP sub-corpora, we observe that the most frequent structural position is turn medial, followed by beginning of turn, with approximately equal frequency in the two corpora. Regarding the (pre)ending of interaction position, it only occurs with 'bem,' being more frequent in the PSFB corpus than in the COR_EP. There are no instances of using 'bem' and 'bom' in the turn final position, and those assuming the beginning of interaction position are very scarce. The use of DMs in this position also differs in the two sub-corpora, as 'bem' occurs in the COR_EP, and 'bom' appears in the PSFB.

In French, the most frequent position is turn medial, followed by the beginning of turn position, whose frequency is around half the first. In these positions, 'bon' has the highest frequency, with 'bien' scoring very low. For the other three positions, we observe that 'bien' doesn't occur in any of them. As for 'bon', it occurs in COR_FR and CLAPI with similar frequency in turn final position. At the beginning of the interaction, it only occurs in COR_FR with shallow frequency and in CLAPI with scarce frequency, in a (pre-)ending of interaction position.

Both in EP and French, there is a preference for the positions of turn medial and beginning of turn, although the distribution of the discourse markers shows different choices, with a clear dominance of 'bon' in French and a more balanced use of 'bem' and 'bom' in EP.

Shifting to the structural functions, we present the results obtained in Table 2.

Table 2. DMs Structural Functions per sub-corpus (frequency per hour)

Structural Functions	COR_EP		PSFB - EP		COR_FR		CLAPI - FR	
	bem	bom	bem	bom	bien	bon	bien	bon
Conclusion	0.47	0.23	0.13	0.01	0.18	0.49	0.00	0.78
Elaboration	0.34	0.27	0.11	0.00	0.04	1.79	0.00	0.89
Filler	0.64	0.37	0.11	0.02	0.09	2.65	0.06	0.84
Response initiator	0.47	0.87	0.45	0.09	0.00	0.58	0.00	0.18
Topic change	0.10	0.23	0.34	0.01	0.04	0.22	0.12	0.48
Topic initiation	0.00	0.13	0.07	0.00	0.00	0.04	0.00	0.00
Topic recovery	0.30	0.17	0.02	0.02	0.13	0.18	0.00	0.12

In EP, although the two sub-corpora show similar trends regarding the most and least used structural functions, they exhibit variations in frequency depending on the functions and distinct distribution of 'bem' and 'bom' based on the sub-corpus, with a global wider use of these DMs to express such functions in COR_EP. In PSFB, the use of 'bom' is much less frequent than that of 'bem' for expressing these functions, while in COR_EP, though 'bem' is globally more frequent, both 'bem' and 'bom' demonstrate a more even utilisation than in COR_EP, with fluctuations in frequency based on the functions. In COR_EP, the least represented structural function is topic initiation, while the most represented is response initiator, followed by filler, conclusion, and elaboration functions, and then by topic recovery and topic change. In this context, there is a significant difference in the use of both DMs in four of the functions, with 'bom' representing almost a double frequency of 'bem' in response initiator and topic change whereas 'bem' has the same distance to 'bom' in conclusion, filler, and topic recovery functions. In PSBF, the most frequent functions are response initiator and topic change and the least, topic recovery and topic initiation, which does not occur with 'bom', with conclusion, elaboration and filler having roughly the same frequency. Contrary to COR_EP, 'bom' doesn't occur with the elaboration function nor as topic initiation.

In French, 'bon' is significantly more frequent than 'bien', which only occurs in both sub-corpora in the topic change function. In COR_FR, 'bien' has mainly a conclusion or topic recovery function, also appearing with a topic recovery, filler, elaboration or topic change function. In contrast, its structural functions are limited to filler and topic change in CLAPI. In turn, 'bon' is more frequent as a filler, with a considerable presence in contrast with the closer functions, which are elaboration, followed by conclusion, topic recovery and topic change. Topic initiation is the least used function in the French corpora, only occurring residually with 'bon' in COR_FR.

Comparing the sub-corpora of both languages, we observe that filler and elaboration functions are the most productive in French and that the response initiator function is the most frequent in EP.

Finally, the distribution of the DMs according to their modal functions is presented in Table 3.

Table 3. DMs Modal functions per sub-corpus (frequency per hour)

Modal Functions	COR_EP		PSFB - EP		COR_FR		CLAPI - FR	
	bem	bom	bem	bom	bien	bon	bien	bon
Agreement	0,23	0,20	0,07	0,00	0,00	0,31	0,00	0,48
Digression	0,27	0,13	0,07	0,02	0,04	0,63	0,00	0,30
Disagreement	0,13	0,17	0,18	0,00	0,00	0,18	0,00	0,12
Mitigation	0,20	0,13	0,16	0,01	0,04	1,12	0,00	0,60
Point of view change	0,17	0,27	0,26	0,01	0,00	0,67	0,00	0,60
Reformulation	0,00	0,13	0,06	0,00	0,00	0,04	0,00	0,00
Request for clarification or specification	0,10	0,03	0,02	0,00	0,04	0,04	0,00	0,06

In COR_EP, there is a balanced use of both 'DMs,' which is not the case in PSFB. The use of 'bom' is much less frequent than that of 'bem' in PSFB, a situation that is reversed in the French sub-corpora, where 'bien' is residual, and 'bon' is predominantly used. Across all four sub-corpora, the least frequent modal functions are reformulation, only scarcely used with 'bom' in COR_EP, 'bem' in PSFB, and 'bon' in CLAPI, and the request for clarification or specification function, non-existent with 'bom' in PSFB and 'bien' in CLAPI, and very low in the other sub-corpora, particularly in PSFB and both French sub-corpora. The reason for the scarce frequency may be related to the fact that these modal functions are primarily expressed by more specific markers, such as 'that is' and 'or rather' in the case of reformulation. When looking into the distribution of the four DMs in contexts in which they participate in reported speech, expressing a point of view change, we verify that the frequency of 'bon' is quite higher in both French sub-corpora when compared to their frequency in EP, with a similar frequency expressed by 'bem' and 'bom', respectively in COR_EP and PSFB, being almost inexistent with 'bom' in this latter sub-corpus. The same tendency for higher frequency in French than in EP occurs with the digression, the mitigation functions and the agreement functions, which are globally higher in the French sub-corpora than in the Portugues ones.

6.3. Discussion

The analysis of the data and these results allow us to draw some conclusions about the discursive semantic functioning of the markers 'bem,' 'bom,' 'bien,' and 'bon' and to compare their usage in EP and French. Some of these conclusions strongly support the observations about markers in general and specifically for this type of DM.

These DMs are polyfunctional, a property commonly ascribed to DMs in general. In the corpora analysed, we observe they can potentially occupy four different positions in discourse, play seven functions in their structural and thematic organisation, and operate at the modal level with seven pragmatic values.

Even though, due to their specificity and for methodological reasons, we have presented their analysis separately, we assume that these levels and functions are often interconnected and must not be seen as mutually exclusive (cf. Cuenca 2008), which is proved when we correlate different layers of classification for the DMs in a specific context. This situation became more evident with some of the structural and modal functions used in the classification scheme, such as the case of response initiator and filler in structural functions and digression and reformulation in modal functions. In the case of the former, it is possible to associate modal nuances, especially partial agreement or disagreement and mitigation, in various examples. At the same time, hesitations can also convey modal traits, which leads them to be frequently considered part of interactional functions. However, in our corpora, their occurrence was more often linked to the processing of spontaneous oral discourse. Various authors consider the latter textual. Agreeing with their textual organisation and, therefore, with the fact that they can play a structural function, we emphasise, by placing them in the set of modal functions in this analysis, the prominence of the assumption of a position regarding the topic under discussion or the situation (digression) or regarding the reevaluation attitude expressed in reformulation.

The interconnection, more pronounced in certain functions as mentioned, is less conspicuous in functions identified at each level. Some functions are more typically structural (e.g., topic change, topic recovery, elaboration), while others are more distinctly modal (e.g., agreement, disagreement, mitigation). For this reason, Cuenca's (2008) analysis proposal for the marker 'well,' considering it as a radial category with certain core values, seems to us to be perfectly applicable to these markers, which can, in many contexts, be equivalent to 'well.'

In addition to the variety of functions they can perform at the textual and modal levels, overall, the results allow us to highlight this variability in terms of

the behaviour of the markers 'bem' and 'bom' both between themselves and in the two EP sub-corpora, in relation to the DMs 'bien' and 'bon' between themselves and in the two French sub-corpora, as well as crosslinguistically. As previously observed, the most used DM is different in EP and French, being 'bem' in EP and 'bon' in French. While in French, the significant dominance of the DM 'bon' over 'bien' is consistent across both corpora, the situation differs in EP. The PSFB corpus records the same tendency in the opposite direction of the French DMs. Still, the other corpus, COR_EP, exhibits a distinctly different trend by showing a minor discrepancy in the frequency of the DM 'bom' compared to 'bem'. A hypothesis for this difference might result in the specific characteristics of each corpus regarding variants and genres. This hypothesis would not necessarily be incongruent with the fact that the two comparable corpora in French and EP - COR_FR and COR_PE - do not show the same tendency. We know that the DM 'bien' in French is rarely used in isolation, mainly occurring in combination, which does not apply to the Portuguese DM 'bom.' Besides, both corpora in French have similar characteristics. In addition to this overall interlinguistic variation, there is interlinguistic variation in the use and frequency of each pair of markers. For example, while 'bon' and 'bom' often perform the same functions, there are situations where, unlike 'bon', 'bom' never occurs, such as in the turn final position. However, this variation also occurs at an intralinguistic level, where many contexts involve both DMs performing the same function, more noticeably in EP, although potentially with different frequencies (e.g., response initiator, mitigation). There are also contexts where one of the markers appears to specialise in expressing a specific value, as it is the case with 'bon' in French, particularly in most modal functions. Although these interlinguistic and intralinguistic variations need further analysis, they show, as already pointed out by Lejeune and Valentim (2015, 95) about 'bem' and 'bien', that, although similar in their etymology, the place they are assigned to in each language is unique and conforms them to their dynamic internal organisation. In this sense, despite the similarities, equivalence between these pairs of markers is possible in various contexts but still limited, among other factors, by their semantic and discursive functionalities. This is a central aspect to consider in their description.

7. Final remarks

The main goal of this study was to investigate and specify the structural positions and semantic-discursive functions of the DMs 'bem' and 'bom' in European Portuguese, as well as 'bon' and 'bien' in French and to compare them.

As these DMs are used in conversational settings, four oral corpora were selected, two for each language. They constituted the base for the analysis carried

out. After a thorough examination of the data, based on structural position, structural function and modal function, we can summarize the following findings. ‘Bem’ and ‘bien’ do not have the same distribution since ‘bien’ has a very low frequency and ‘bem’ is rather frequent; ‘bon’ is the preferred DM in French in both corpora while ‘bem’ is the most used in EP. The dominant structural positions in both languages are turn medial and beginning of turn. Regarding the structural functions, we observed that in the Portuguese corpora, response initiator is the most frequent and topic initiation the least for both DMs; In French, the most frequent is filler and the least topic initiation. As to modal functions, the less frequent are reformulation and request for clarification or specification and the most frequent is mitigation in French and point of view change in EP. In spite of these general results, we observe a tendency to interlinguistic and intralinguistic variation of the DMs functions and positions.

We believe that one of the reasons for the discrepancies in the use of these DMs may be the text genres and register of each corpus, but this will be analysed in future research.

WORKS CITED

- Aijmer, Karin. 2013. *Understanding Pragmatic Markers: A Variational Pragmatic Approach*. Edinburgh: Edinburgh University Press.
- (ATILF), Laboratoire d'informatique fondamentale de Marseille - UMR 7279 (LIF), Laboratoire lorrain de recherche en informatique et ses applications - UMR 7503 (Loria), Cognition, Langue, Langages, Ergonomie - UMR 5263 (CLLE), Interactions, corpus, apprentissages et représentations - UMR 5191 (ICAR), Langues, textes, traitements informatiques, cognition - UMR 8094 (Lattice) (2021). CEFC [Corpus]. ORTOLANG (Open Resources and TOols for LANGuage) - www.ortolang.fr, v1.5, <https://hdl.handle.net/11403/cefc-orfeo/v1.5>
- Bacelar do Nascimento, Maria Fernanda, et al. 2005. “The Portuguese Corpus”. In *C-ORAL-ROM: Integrated Reference Corpora for Spoken Romance Languages*, edited by Emanuela Cresti, and Massimo Monegnia, 163-207. Amsterdam: John Benjamins.
- Barnes, Betsy K. 1995. “Discourse Particles in French Conversation: (eh)ben, bon, and enfin”. *The French Review* 68, no. 5: 813–21.
- Beeching, Kate. 2007. “La co-variation des marqueurs discursifs *bon, c'est-à-dire, enfin, hein, quand même, quoi et si vous voulez* : une question d'identité?”. *Langue Française* 154, n^o. 2 : 78-93.
- Benzitoun, Christophe, Jeanne-Marie Debaisieux, and Henri-José Deulofeu. 2016. “Le projet ORFÉO : un corpus d'étude pour le français contemporain”. *Corpus* 15: 91-114.

- Brémond, Capucine. 2003. "Bon, moteur d'action, moteur du discours". *Travaux Interdisciplinaires du Laboratoire Parole et Langage* 22: 65-84.
- Brémond, Capucine. 2004. "La petite marque *bon*, l'indice d'un accord en cours de négociation". *Travaux de linguistique* 48, no. 1: 7-19.
<https://doi.org/10.3917/tl.048.0007>.
- Cabarrão, Vera, et al. 2018. "Cross-domain analysis of discourse markers in European Portuguese". *Dialogue & Discourse* 9, no. 1: 79-106.
- Cuenca, Maria-Josep. 2008. "Pragmatic markers in contrast: the case of *well*". *Journal of Pragmatics* 40, no. 8: 1373-1391.
- Cuenca Maria-Josep. 2013. "The fuzzy boundaries between discourse marking and modal marking". In *Discourse Markers and Modal Particles. Categorization and description*, edited by Liesbeth Degand, Bert Cornillie, and Paola Pietrandrea, 181-216. Amsterdam: John Benjamins.
- Debaisieux, Jeanne-Marie, and Christophe Benzitoun. 2020. "Présentation". *Langages* 219, no. 3: 9-24.
- Estellés, María, and Salvador Pons Bordería. 2014. "Absolute initial position". In *Discourse Segmentation in Romance Languages*, edited by Salvador Pons Bordería, 121-155. Amsterdam: John Benjamins.
- Fraser, Bruce. 2006. "Towards a theory of discourse markers". In *Approaches to Discourse Particles*, edited by Kerstin Fischer, 189-204. Oxford: Elsevier.
- Gilbert, Jane Alexander. 2019. *The Syntactic Environment of the French Discourse Marker Bon*. MA, The University of Georgia.
- Gorski, Edair Maria. 2020. "Espectro funcional de bem e bom no português falado: instâncias de gramaticalização". *Revista da ABRALIN* 19, no. 3, 131-158.
- Hansen, Maj-Britt Mosegaard. 1998. The semantic status of discourse markers'. *Lingua* 104: 235-260.
- Jaez, Jacques. 2004. "Bon : le mot de la fin". Handout of a talk given at Université de Genève - 23 mars 2004 (Version: 18/03/04).
- Lefevre, Florence. 2011. "Bon et quoi à l'oral : marqueurs d'ouverture et de fermeture d'unités syntaxiques à l'oral". *Linx* 64-65 : 223-240.
<https://doi.org/10.4000/linx.1417>.
- Lejeune, Pierre, and Valentim, Helena. 2015. "Étude contrastive de quelques valeurs de *bien* et *bem*". In *Variation, Ajustement, Interprétation*, edited by Daniel Lebaud, and Catherine Paulin, 95-114. Besançon : Presses Universitaires de Franche-Comté.
- Lopes, Ana Cristina Macário. 2004. "A polifuncionalidade de 'bem' no PE contemporâneo". In *Linguagem, cultura e cognição. Estudos de Linguística Cognitiva*, edited by Amadeu Torres and Augusto Soares da Silva, vol. II, 433-458. Coimbra: Almedina.
- Marques, Maria Aldina, and Micaela Aguiar. 2014. "Usos de *portanto* no falar bracarense". In *XXIX Encontro Nacional da Associação Portuguesa de Linguística (Coimbra, 2013). Textos Selecionados*, edited by João Veloso, Anónio Moreno, Fátima Silva, Isabel Falé, and Isabel Pereira, 321-331. Lisboa: APL

- Maschler, Yael, and Deborah Schiffrin. 2015. "Discourse markers: Language, meaning, and context". In *The Handbook of Discourse Analysis*, edited by Deborah Tannen, Heidi E. Hamilton, and Deborah Schiffrin. Second edition. 189-221. Chichester: John Wiley & Sons.
- Mendes, Amália. 2016. "Linguística de Corpus e outros usos do corpus em Linguística". In *Manual de Linguística Portuguesa*, edited by Ana Maria Martins, and Ernestina Carrilho, 224-251. Berlin/Boston: Walter de Gruyter,
- Moline, Estelle. 2012. "Aperçu des emplois de bien en français contemporain". *Travaux de Linguistique* 65, no. 2: 7 -26.
- Morozova, Milana Andreevna. 2019. *The role of discourse markers in text organization of the genre stand-up comedy in Portugal and in the United States*. PhD Dissertation. Nova Lisbon University.
- Morozova, Milana Andreevna. 2020. "Discourse markers in English and European Portuguese translations: establishing functional equivalents and types of omission." *Filologia e Linguística Portuguesa* 22, no. 1: 103-121.
- Oliveira, Fátima, and Fátima Silva. 2020. "Para uma comparação dos marcadores discursivos *bem e bom* em português europeu com *well* em inglês". In *Marcadores discursivos. O português como referência contrastiva* edited by Isabel Margarida Duarte, and Rogelio Ponce de León, 207-226. Frankfurt: Peter Lang.
- Peltier, Joy P. G., and Diana L. Ranson. 2020. "Le marqueur discursif *bon* : ses fonctions et sa position dans le français parlé". In *SHS Web Conf.*, 78 (2020). Article Number 01006, 20 pages. <https://doi.org/10.1051/shsconf/20207801006>.
- Péroz, Pierre. 1992. *Systématique des valeurs de bien en français contemporain*. Genève-Paris: Droz.
- Ponce de León, Rogelio, and Isabel Margarida Duarte. 2021. "Os operadores discursivos *ahora bien / ahora, (que)* e as suas correspondências em traduções literárias para português". *Revista de Estudos Linguísticos da Universidade do Porto*. N.º Especial: 535-553.
- Risso, Mercedes Sanfelice. 1999. "Aspectos textuais-interativos dos marcadores discursivos de abertura *bom, bem, olha, ah*, no português culto falado". In *Gramática do Português Falado*, edited by Maria Helena Moura Neves, vol. 7, 259-296. Campinas: Editora da Unicamp.
- Rodrigues, Isabel. 1998. *Sinais Conversacionais de Alternância de Vez*. Porto: Granito Editores e Livreiros.
- Schiffrin, Deborah. 1987. *Discourse Markers*. Cambridge: Cambridge University Press.
- Urbano, Hudinilson. 2003. "Marcadores conversacionais". In *Análise de textos orais*, edited by Dino Preti. 6ª ed., 93-116. São Paulo: Humanitas.
- Valentim, Helena. 2008. "Dialogical occurrences of *bem* in European Portuguese. Enunciative stability and deformability". *L'analisi Linguistica e Letteraria. Special Issue: Word Meaning in Argumentative Dialogue XVI*: 305-315.