Introduction
Les limites du concept de marqueur discursif : discours et syntaxe

Introduction
The limits of the concept of discourse marker: discourse and syntax

Amália Mendes

Universidade de Lisboa (Lisboa, Portugal)
mendes@edu.ulisboa.pt
https://orcid.org/0000-0001-6815-2674

Pierre Lejeune

Universidade de Lisboa (Lisboa, Portugal)
lejeune@letras.ulisboa.pt
https://orcid.org/0000-0002-2495-9043

Fabrizio Macagno

Universidade de Lisboa (Lisboa, Portugal)
fmacagno@edu.ulisboa.pt
https://orcid.org/0000-0003-0712-421X

Reçu le 1/7/2026, accepté le 30/7/2026, publié le 20/11/2026

Creative Commons Attribution 4.0 International (*CC BY 4.0*)
© 2026 Amália Mendes, Pierre Lejeune, Fabrizio Macagno

Pour citer cet article
Mendes, Amália, Pierre Lejeune, Fabrizio Macagno 2026. Introduction. Les limites du concept de marqueur discursif : discours et syntaxe. *Studia linguistica romanica* 2026.17, 1-14. https://doi.org/10.25364/19.2026.17.1.

Résumé
Cette introduction aux cinq articles rassemblés dans ce numéro thématique propose une réflexion sur le concept de marqueur discursif, sur sa délimitation et sur la distinction entre marqueur discursif et particule modale, incluant des perspectives plus englobantes qui analysent la grammaticalisation de constructions syntaxiques. Outre l'étude des marqueurs discursifs ayant des fonctions interpersonnelles, modales et connectives, deux analyses portent également sur des éléments syntaxiques qui ne peuvent pas encore être strictement classés comme des marqueurs discursifs, mais dont la grammaticalisation a été évaluée à l'aide de corpus: le marqueur des questions orientées *a que* en espagnol, et les constructions prédicatives avec des noyaux de type *shell nouns* en français et en roumain.

Mots-clés
Analyse contrastive, grammaticalisation, marqueur discursif, particule modale.

Abstract
The introduction to the five articles featured in the thematic issue reflects on the concept and definition of discourse marker, on the distinction between discourse markers and modal particles and on broader perspectives that analyze the grammaticalization of syntactic constructions. In addition to the study of discourse markers with interpersonal, modal and connective functions, two papers focus on syntactic elements that cannot yet be strictly classified as discourse markers, but whose grammaticalization is examined using corpora: the marker of biased questions *a que* in Spanish and predicative constructions with shell nouns in French and Romanian.

Keywords
Contrastive analysis, discourse marker, grammaticalization, modal particle.

DOI : https://doi.org/10.25364/19.2026.17.1 ISSN : 2663-9815

Sommaire

1 Marqueurs discursifs : diversité d'approches et délimitation du concept....4
2 Marqueurs discursifs dans une perspective contrastive....6
3 Marqueurs discursifs et grammaticalisation....8
Abréviations et références bibliographiques....12

DOI : https://doi.org/10.25364/19.2026.17.1 ISSN : 2663-9815

1 Marqueurs discursifs : diversité d'approches et délimitation du concept

[1] Les marqueurs discursifs (MD), objet de ce numéro thématique, constituent une catégorie vaste, caractérisée par une certaine variation terminologique et des divergences quant à la délimitation des frontières du concept (Pons Bordería 1998 ; Cuenca & Marín 2009 ; Zufferey & Degand 2024). Cette catégorie englobe les particules discursives, qui remplissent une fonction à la fois structurelle et interpersonnelle, en signalant la structure du discours, notamment le début et la fin d'un tour de parole, les pauses, la gestion de l'information, ainsi que la communication avec l'interlocuteur grâce à des valeurs modales d'accord, de désaccord, de distanciation ou d'orientation de la perspective. La catégorie des MD peut également inclure les connecteurs, qui ont une fonction au niveau propositionnel, signalant une relation discursive (ou relation de cohérence) entre des segments discursifs et contribuant, de ce fait, à la cohérence du discours.

[2] Les cinq articles de ce numéro thématique, qui proposent différentes approches de l'étude des MD, sont représentatifs de la diversité conceptuelle qui caractérise ce domaine et offrent au lecteur un aperçu des perspectives théoriques et méthodologiques mobilisées dans l'analyse de cette catégorie au sens large. Ces articles sont issus de communications présentées lors du colloque *Discourse markers in Romance langages (DISROM 8)*, organisée par l'Université de Lisbonne du 19 au 21 juin 2024. Cette sélection d'articles vise à présenter au lecteur un continuum incluant des éléments plus ou moins prototypiques de la catégorie des MD, commençant par le cas des MD lexicaux aux fonctions structurelles et interactionnelles (Silva, Bacquelaine & Oliveira) et des MD aux fonctions connectives (Müller & Baptista). Ce dernier article introduit également la discussion de la délimitation difficile entre syntaxe et discours. Deux articles ont été sélectionnés pour apporter une réflexion sur des constructions qui acquièrent une fonction de MD au cours d'un processus de grammaticalisation (Herrero Ruiz de Loizaga ; Petraş). Enfin, l'article de Duarte discute de la différence entre MD et particule modale, deux catégories qui partagent des fonctions pragmatiques.

[3] L'article de Silva, Bacquelaine & Oliveira traite d'un élément présentant les caractéristiques prototypiques d'un MD (*eh bien* et ses équivalents fonctionnels en portugais), dont les fonctions s'inscrivent au niveau de l'organisation structurelle et de la gestion interactionnelle. Müller & Baptista se concentrent sur des connecteurs de type adverbial ayant une fonction au niveau propositionnel, assurant la connexion entre les segments textuels. Leur article soulève d'ailleurs la question de la délimitation des fonctions de connecteur des adverbes, en s'appuyant sur la distinction entre modification externe et modification interne. L'article de Duarte présente, dans le cadre de l'analyse de *lá* 'là', une réflexion sur la distinction entre MD et particule modale. Bien qu'elles partagent avec les MD certaines valeurs au niveau des attitudes et des actes de parole, les particules modales s'en distinguent par le fait qu'elles ne peuvent occuper la position initiale d'une phrase ou d'un énoncé et qu'elles sont syntaxiquement intégrées (Coniglio 2008 ; Diewald 2013).

DOI : https://doi.org/10.25364/19.2026.17.1 ISSN : 2663-9815

En portugais, un petit ensemble d'éléments, incluant *lá*, a été considéré comme présentant des usages qui réunissent les propriétés de cette catégorie (Franco 1988 ; Meisnitzer 2012 ; Favaro 2025).

[4] Les articles de Herrero Ruiz de Loizaga et de Petraş vont au-delà des catégorisations habituelles des MD et se concentrent sur des constructions syntaxiques ayant des fonctions discursives : chez Herrero Ruiz de Loizaga, un marqueur de questions orientées ; chez Petraş, des constructions prédicatives dont le noyau est constitué de *shell nouns* qui introduisent, par des valeurs modales, une prédication constituant l'information principale de l'énoncé. La disponibilité croissante de corpus oraux et de corpus parallèles, y compris de corpus annotés sur le plan discursif, a facilité l'identification de séquences lexicales et de constructions syntaxiques, plus ou moins fixes, qui remplissent des fonctions de liaison ou véhiculent des valeurs interpersonnelles et modales, pouvant simultanément conserver une interprétation non idiomatique de leurs éléments constitutifs.

[5] Le fait que nous soyons en présence d'éléments ambigus, entre une fonction discursive et une fonction syntaxique, comme c'est le cas des adverbes complexes, implique une discussion des propriétés syntaxiques et pragmatiques de ces unités, le degré de fixité étant d'ailleurs un facteur fondamental pour identifier des fonctions de MD dans l'article de Müller & Baptista. Herrero Ruiz de Loizaga adopte une perspective diachronique dans l'analyse des marqueurs syntaxiques des questions orientées et examine la progression de la grammaticalisation de la construction depuis le 16e siècle jusqu'à l'époque moderne, un parcours de grammaticalisation également abordé par Petraş dans le cas des constructions prédicatives avec des *shell nouns*.

[6] Plusieurs langues font l'objet des études réunies dans ce numéro thématique, et trois articles adoptent une perspective contrastive. Herrero Ruiz de Loizaga se concentre spécifiquement sur l'espagnol, tandis que Müller & Baptista s'intéressent au portugais, comparant le portugais européen et le portugais du Brésil. Les autres articles adoptent une perspective contrastive, une approche souvent utilisée pour mettre en évidence des variations subtiles dans les valeurs pragmatiques d'un MD. Dans le cas de Duarte, l'analyse part de la langue portugaise pour observer les équivalents fonctionnels de *lá* en français, en espagnol et en italien. Silva, Bacquelaine & Oliveira empruntent le chemin inverse : elles partent de l'expression *eh bien* en français et observent ses équivalents fonctionnels, en analysant en détail trois MD du portugais.

[7] Les données utilisées pour les analyses dans ces cinq articles proviennent toutes de corpus, le choix des corpus étant adapté aux objectifs de chaque étude. Dans le cas d'analyses contrastives, on utilise des corpus de discours oral en portugais, en français et en roumain (Duarte ; Silva, Bacquelaine & Oliveira ; Petraş) ainsi que des corpus de discours oral formel transcrit, comme les transcriptions des séances du Parlement portugais et le corpus parallèle *ParlE*, avec alignement des sessions du Parlement européen dans les langues portugaise et française. Duarte

DOI : https://doi.org/10.25364/19.2026.17.1 ISSN : 2663-9815

utilise également une autre source de données parallèles, en analysant les traductions en français, en espagnol et en italien du roman classique de la littérature portugaise *Os Maias* d'Eça de Queirós. Les œuvres littéraires constituent aussi une source d'information diachronique dans l'article de Herrero Ruiz de Loizaga, qui utilise un corpus d'œuvres littéraires allant du 16e siècle à l'époque moderne.

[8] Outre le thème général du colloque, cette édition proposait une thématique spécifique, portant plus précisément sur les interactions entre les MD et leur contexte explicite/implicite, en dépassant la distinction traditionnelle entre leurs fonctions textuelles et interpersonnelles, ainsi que sur la fonction d'ajustement subjectif des MD. En effet, certains des travaux présentés dans ce numéro thématique abordent des aspects de la dimension subjective des MD. Duarte, en analysant la particule modale *lá*, met en évidence une valeur de distanciation du locuteur. Silva, Bacquelaine & Oliveira mentionnent les valeurs évaluatives du marqueur modal *pois bem*, l'un des équivalents fonctionnels de *eh bien*. L'orientation de la question adressée à l'interlocuteur par le biais d'un marqueur syntaxique, thème de l'article de Herrero Ruiz de Loizaga, explore cette dimension d'ajustement subjectif au cours de l'interaction.

2 Marqueurs discursifs dans une perspective contrastive

[9] Dans son article, Duarte reprend et développe des analyses antérieures de l'adverbe *lá* du portugais européen, en s'appuyant sur le corpus *C-Oral-Rom* ainsi que sur l'original de l'œuvre littéraire *Os Maias* de José M. Eça de Queirós (Eça de Queirós 1888) et sur ses traductions en français, en espagnol et en italien (Duarte 1989, 2009 ; Duarte & Marques 2018 ; Marques & Duarte 2015, 2017, 2019). L'approche contrastive permet, selon l'auteure, de mieux cerner les valeurs de *lá* car le travail de traduction montre des « valeurs pragmatiques subtiles, difficiles à cerner pour ces particules ».

[10] Dans cet article, l'auteure aborde la différence entre les MD et les particules modales, une distinction particulièrement pertinente dans le cas de l'adverbe *lá* (Franco 1988 ; Meisnitzer & Müller 2024 ; Müller 2024 ; Favaro 2025, entre autres). Elle estime que la classification de *lá* en tant que particule modale est la plus appropriée, car cet élément remplit une fonction au niveau de l'acte discursif, révélant l'attitude du locuteur vis-à-vis du contenu propositionnel de l'énoncé, et non une fonction de liaison entre les éléments du discours ; en outre, selon l'auteure, *lá* apparaît surtout à l'intérieur et pas au début d'une séquence et n'est pas séparé du reste de l'énoncé par des pauses, contrairement aux MD. Cependant, l'auteure ne manque pas de signaler des cas où *lá* ne semble pas correspondre aux définitions d'une particule modale, comme le fait qu'il modifie la polarité de l'énoncé lorsqu'il est placé après le verbe. Elle souligne également le fait que certaines constructions avec *lá* sont devenues de vrais MD (par exemple, *sei lá*).

[11] L'analyse des emplois pragmatiques de *lá* montre que les contextes conservent la valeur de distance qui caractérise la valeur déictique de l'adverbe.

DOI : https://doi.org/10.25364/19.2026.17.1 ISSN : 2663-9815

Quand *lá* suit un verbe à l'impératif, la valeur décrite est une atténuation, et parfois un renforcement, d'un acte directif. Dans d'autres contextes, il a une valeur négative, de distanciation entre le locuteur et ce dont il parle. Antéposé au verbe, dans des actes assertifs, il indique que la situation fait partie du *common ground* et est prévisible, bien qu'elle puisse se réaliser avec difficulté, ce qui amène Duarte à considérer que *lá* est une particule modale. Lorsque *lá* précède un syntagme nominal, il sert à thématiser une information et à signaler un changement de sujet. Enfin, il a également une valeur de négation, que Duarte décrit plutôt comme une valeur épistémique dubitative, suggérant la distanciation du locuteur à l'égard de ce dont il parle. L'analyse de Duarte met en évidence diverses valeurs situées sur un continuum allant des valeurs locatives aux valeurs modales, avec des cas ambigus ou hybrides difficiles à classer définitivement.

[12] Duarte montre que les traductions utilisent trois stratégies : l'omission de *lá*, la traduction avec des marqueurs déictiques, interprétant erronément *lá* comme un adverbe de lieu, ou d'autres solutions recourant à des équivalents dans d'autres langues ou à des paraphrases, comme *allora* en italien pour rendre la valeur de situation prévisible, bien que parfois difficile à réaliser. L'occurrence de *lá* après un verbe à l'impératif est la valeur la plus fréquente dans le corpus oral et dans *Os Maias*, mais il existe des différences que l'auteure attribue au genre textuel ou aux différentes époques représentées par ces données : la valeur de négation métalinguistique est très fréquente dans le corpus oral, tandis que la valeur de *common ground* est plus fréquente dans *Os Maias*.

[13] Dans *Marqueurs discursifs en contraste.* Eh bien *et ses équivalents fonctionnels en portugais européen* bem*,* ora bem *et* pois bem, Silva, Bacquelaine & Oliveira s'intéressent dans une perspective contrastive aux emplois du marqueur discursif français *eh bien* et de trois de ses équivalents fonctionnels en portugais, tous formés à partir de l'adverbe *bem* : *bem*, *ora bem* et *pois bem*. S'appuyant sur les descriptions existantes de *eh bien*, les auteures relèvent la polyvalence dans le discours oral de ce marqueur, parfois décrit dans la littérature comme *marqueur conversationnel* (Schiffrin 1987 ; Lefeuvre 2012), qui non seulement intervient dans l'organisation structurelle et la gestion interactionnelle du discours, mais exprime aussi des attitudes modales et évaluatives du locuteur.

[14] Afin de mieux saisir les équivalences fonctionnelles entre *eh bien* et les trois marqueurs portugais, les auteures ont construit un sous-corpus du corpus *ParlE*, qui contient les transcriptions intégrales des séances plénières du Parlement européen. Pour ce corpus, les auteures ont sélectionné les séances datant d'avant 2011, date à laquelle ont cessé les traductions directes entre les langues de l'Union européenne pour passer à un système de traductions indirectes ou l'anglais joue le rôle de langue-pivot. Le corpus est aussi limité aux orateurs francophones, à partir de la version française de *ParlE*. Les 200 occurrences de la séquence *eh bien* de ce sous-corpus ont été alignées avec leur contexte sur les séquences équivalentes de la version portugaise, où, dans plus de la moitié des cas, *eh bien* a été traduit par *pois bem*.

DOI : https://doi.org/10.25364/19.2026.17.1 ISSN : 2663-9815

[15] Dans la version française du corpus *ParlE*, les auteures ont effectué un relevé systématique des emplois structurels/interactionnels (conclusion, conséquence, élaboration, initiateur de réponse, reprise de sujet et remplissage) et modaux (accord, atténuation, changement de point de vue, contraste, désaccord, digression, évaluation, opposition inattendue et reformulation) de *eh bien*. Elles ont étudié les équivalents de *eh bien* dans la version portugaise de *ParlE* et aussi, à titre complémentaire, dans deux autres corpus : le corpus *C-Oral-Rom*, constitué notamment d'échanges informels spontanés, d'extraits de médias et de conversations téléphoniques, dans ses versions française et portugaise, et le sous-corpus de transcriptions des séances du Parlement portugais, inclus dans le corpus *CRPC*. L'analyse des emplois spécifiques des trois marqueurs portugais dans ces corpus montre que *pois bem* remplit diverses fonctions structurelles (initiateur de réponse, conséquence, élaboration, conclusion, remplissage, reprise de sujet) et modales, où l'évaluation prédomine nettement, tandis qu'*ora bem* remplit des fonctions typiques du discours oral spontané et dialogué telles que le remplissage et l'emploi en position indépendante au début d'un tour de parole, où il manifeste une réponse à valeur concordante avec l'interlocuteur.

[16] Ce travail sur des corpus complémentaires amène les auteures à s'interroger sur la relation entre la fréquence des marqueurs portugais et les genres discursifs où ils se trouvent. On note en particulier que *pois bem* est de loin le marqueur le plus fréquent dans les transcriptions de séances parlementaires (corpus *ParlE* et *CRPC*), alors qu'il est complètement absent du *C-Oral-Rom*.

3 Marqueurs discursifs et grammaticalisation

[17] Herrero Ruiz de Loizaga aborde la question des marqueurs de questions orientées, c'est-à-dire des questions qui ne se contentent pas de solliciter une réponse de la part du destinataire, mais qui indiquent également le type de réponse attendue (Giannakidou & Mari 2025 ; Krifka 2015). L'analyse des marqueurs des questions orientées dans les langues romanes constitue un domaine relativement peu exploré de l'analyse linguistique (Gómez González & Silvano 2022 ; Silvano, Gómez González & Cordeiro 2022). Alors que le latin disposait de marqueurs spécifiques pour ce type de questions, les langues romanes ont développé leurs propres particules syntaxiques, dont l'étude permet une compréhension plus approfondie de l'interface pragmatique-syntaxe. Au cours de son histoire, l'espagnol a développé différents types de particules signalant l'orientation de la réponse attendue ou tenue pour acquise par le locuteur – qu'elle soit caractérisée par la même polarité que la question (question co-orientée ou question orientée positivement) ou par la polarité opposée (question anti-orientée ou question orientée négativement).

[18] Herrero Ruiz de Loizaga concentre son analyse sur un marqueur spécifique des questions biaisées, l'expression espagnole *a que*, qui indique que la question attend une réponse dans le même sens. L'explication de cet effet pragmatique est reconstituée à travers son évolution diachronique. Il est démontré que ce

DOI : https://doi.org/10.25364/19.2026.17.1 ISSN : 2663-9815

marqueur trouve son origine dans des constructions visant à lancer des défis (*¿Qué te apuestas a que ... ?*), qui sont par la suite devenues elliptiques. À travers une analyse de corpus de textes allant des œuvres littéraires des 16e et 17e siècles à l'époque moderne et contemporaine, l'article retrace comment l'introduction fréquente de questions appelant une réponse orientée dans le même sens a évolué pour devenir un marqueur de questions biaisées.

[19] *A que* a progressivement évolué : d'abord utilisé pour introduire un pari, il a ensuite servi à signaler que l'action qui suit – ou éventuellement celle qui est omise – est un pari, un défi ou une action que les interlocuteurs pourraient ou devraient accomplir, et il finit par servir à mettre l'accent sur un élément du discours. L'attente fréquente d'une réponse positive à un défi a conduit à la conventionnalisation de cette inférence, ce qui a transformé l'expression en un marqueur de question biaisée. Ce marqueur interrogatif a récemment évolué vers les expressions *¿a que sí/no ?*, qui, en Espagne, sont utilisées pour demander une confirmation ou un accord avec l'affirmation suivante du locuteur et sous-entendent que l'idée exprimée est partagée. Cette analyse diachronique permet d'expliquer la coexistence de différents usages de *a que*. L'acquisition de ses nouvelles fonctions, en tant que marqueur de demande de confirmation ou marqueur de confirmation, n'a pas entraîné la disparition de ses usages les plus anciens, à savoir ceux de défi, de menace ou de marqueur visant à retenir l'attention. L'usage de *a que* en tant que marqueur de question biaisée est limité à l'Espagne, et pratiquement absent dans le reste du monde hispanophone. Sa fréquence normalisée en tant qu'expression est bien plus élevée en Espagne que dans le reste des pays hispanophones (près de quatre fois) ; de plus, le pourcentage relatif de son utilisation pour introduire une question biaisée co-orientée est nettement plus élevé en Espagne qu'en Amérique latine.

[20] Alors que les contributions de Duarte, de Silva, Bacquelaine & Oliveira, ainsi que celle de Herrero Ruiz de Loizaga, analysent des MD ayant des fonctions pragmatiques, Müller & Baptista s'occupent de MD ayant une fonction connective. Ils examinent, dans le cadre du modèle théorique du Lexique-Grammaire (Gros 1981), ces connecteurs, qu'ils appellent des adverbes conjonctifs composés en portugais, entendus comme éléments composés de plus d'un élément et qui établissent des relations discursives entre des segments textuels (Molinier & Levrier 2000). Cette étude s'inscrit dans une tentative plus vaste de décrire les adverbes composés du portugais du Brésil et du Portugal (Müller, Mamede & Baptista 2023, 2024).

[21] À partir d'un inventaire préalable de 3700 expressions formées de plusieurs mots, couvrant le portugais brésilien et le portugais européen, les auteurs ont obtenu une liste d'environ 155 adverbes conjonctifs composés distincts et de 41 adverbes simples se terminant par *-mente* (équivalent de *-ment* en français).

[22] D'un point de vue syntaxique, ces adverbes ne peuvent pas apparaître en position initiale absolue du discours et ils fonctionnent comme des modificateurs externes à la phrase, qu'ils modifient dans son ensemble, contrairement aux modi-

DOI : https://doi.org/10.25364/19.2026.17.1 ISSN : 2663-9815

ficateurs internes, qui agissent sur un constituant de celle-ci. Ainsi, les adverbes conjonctifs ne sont pas affectés par la négation.

[23] Müller & Baptista présentent une classification sémantique spécifique des adverbes conjonctifs, en s'appuyant sur les propositions de Molinier & Levrier (2000) et de Lopes & Carrilho (2020). Les auteurs ont analysé ces adverbes dans le corpus oral du portugais du Brésil *Roda Viva* et annoté sémantiquement leurs contextes, d'abord automatiquement, puis en révisant manuellement les cas ambigus. Le travail de Müller & Baptista permet d'obtenir un inventaire détaillé des adverbes conjonctifs composés du portugais, accompagné d'une analyse de leurs propriétés formelles et de leurs fonctions discursives, ainsi que d'une classification sémantique. Un sous-corpus de plus de 1700 énoncés, contenant 2467 séquences candidates, a été annoté et désambiguïsé manuellement en cas de lectures multiples potentielles.

[24] Dans *Marqueurs discursifs construits à partir de* shell nouns *en français et en roumain*, Petraş s'intéresse dans une perspective contrastive à des suites semi-figées du français et du roumain comme *le fait est que* / *faptul este că*, *le problème est que* / *problema este că*, *l'idée est que* / *ideea este că*, *la vérité est que* / *adevărul este că*, envisagées comme constructions, c'est-à-dire, selon la perspective des grammaires des constructions, des unités qui associent d'une manière conventionnelle une forme et un sens et dont certains aspects formels ou sémantiques ne sont pas strictement prévisibles à partir des éléments constitutifs de cette construction ou d'une autre construction identifiée (Carlier & Prévost 2021). Le type de construction étudiée est du type [N *être que* + P conjonctive], où le contenu de la proposition conjonctive est l'information principale de la phrase, tandis que la proposition principale à copule joue un rôle extraprédicatif de marqueur discursif focalisant ou connectif.

[25] Dans un premier temps, l'article passe en revue la littérature consacrée aux noms qu'on retrouve dans la construction, à leurs dénominations – *shell nouns* (Schmid 2000), *noms sous-spécifiés* (Legallois 2006) – et à leurs fonctions discursives à travers la construction étudiée (focalisation de l'élément à droite de *que*, fonction cohésive provenant de la résolution dans le contexte gauche de l'incomplétude sémantique du *shell noun*). Ensuite, partant du principe que le sémantisme (même s'il reste assez vague) propre à chaque *shell noun* contribue à expliquer les valeurs spécifiques (spécification ou focalisation ; types de relations discursives entre le contexte gauche et le contexte droit) des expressions dans lesquels il entre, et en s'appuyant sur des corpus synchroniques et diachroniques (*French Web 2023* ; *Romanian Web 2021* ; *Frantext*), l'auteure passe en revue l'emploi en discours d'une sélection de variantes – en français et en roumain – de la construction de base incluant les *shell nouns* les plus typiques (en français *fait*, *problème*, *idée*, *vérité*, *point*, *chose*, *truc*).

[26] Elle remarque par exemple que les expressions du type *le fait est que* ou *le problème est que* sont celles qui ont atteint le stade de grammaticalisation le plus

DOI : https://doi.org/10.25364/19.2026.17.1 ISSN : 2663-9815

avancé. Concernant les relations discursives qui promeuvent la construction au rang de marqueur pragmatique, *le fait est que* se caractérise par une polyfonctionnalité qui recouvre notamment certaines valeurs de *de fait* (confirmation) et de *en fait* (opposition, expansion, restriction) ; *le problème est que*, du fait du caractère évaluatif de *problème*, recatégorise – comme problème – la situation présentée dans la P conjonctive, *l'idée est que* véhicule souvent une valeur modale boulique (contexte de projets), *la vérité est que* favorise les relations d'opposition (apparence vs. réalité), *le point est que* et son équivalent roumain semblent être des calques de l'anglais.

[27] Le fonctionnement des expressions roumaines est souvent très semblable à celui des expressions françaises qui recourent à des noms comparables. Une exception : on ne retrouve pas dans l'expression roumaine *fapt este că* la valeur contrastive de *le fait est que*. Si enfin on compare les expressions roumaines et françaises à leurs équivalents anglais, on constate que seule cette dernière langue permet régulièrement des emplois sans conjonction (par exemple *the fact is* en incise).

[28] Les articles de ce numéro, bien que présentant des études de cas spécifiques, contribuent de façon plus générale à la circonscription des MD, une catégorie qui, par ses diverses fonctions sémantiques, textuelles et pragmatiques, inclut une large diversité de catégories morphosyntaxiques et est enrichie par des processus de grammaticalisation de séquences syntaxiques.

DOI : https://doi.org/10.25364/19.2026.17.1 ISSN : 2663-9815

Abréviations et références bibliographiques

Carlier & Prévost 2021 = Anne Carlier, Sophie Prévost 2021. Constructions, constructionnalisation et changement linguistique. *Langue française* 209, 9-22.

Coniglio 2008 = Marco Coniglio 2008. Modal particles in Italian. *University of Venice Working Papers in Linguistics* 18, 91-129. https://phaidra.cab.unipd.it/detail/o:458871?mycoll=o:458878.

C-Oral-Rom = Emanuela Cresti, Massimo Moneglia (éds.) 2005. *C-Oral-Rom. Integrated reference corpora for spoken Romance languages*. Benjamins.

CRPC = Centro de linguística da Universidade de Lisboa (éd.) 2014. *Corpus de referência do português contemporâneo*. https://clul.ulisboa.pt/recurso/corpus-de-referencia-do-portugues-contemporaneo.

Cuenca & Marín 2009 = Maria J. Cuenca, Maria J. Marín 2009. Co-occurrence of discourse markers in Catalan and Spanish oral narrative. *Journal of Pragmatics* 41.5, 899-914.

Diewald 2013 = Gabriele Diewald 2013. *Same same but different* – Modal particles, discourse markers and the art (and purpose) of categorization. Liesbeth Degand, Bert Cornillie, Paola Pietrandrea (éds.). *Discourse markers and modal particles. Categorization and description*. Benjamins, 19-45.

Duarte 1989 = Isabel M. Duarte 1989. *Alguns operadores de agulhagem comunicativa (na prosa narrativa de Eça de Queirós e José Cardoso Pires)*. Thèse de Master, Université de Porto. https://repositorio-aberto.up.pt/bitstream/10216/56179/2/tesemestisabelduarte000129063.pdf.

Duarte 2009 = Isabel M. Duarte 2009. La dimension modale de *cá* et *lá* en portugais. *Studii şi cercetări lingvistice* 60.2, 179-195. https://www.lingv.ro/SCL%202%202009.pdf.

Duarte & Marques 2018 = Isabel M. Duarte, M. Aldina Marques 2018. Funções discursivas das construções com a partícula *lá*. João Veloso, Joana Guimarães, Purificação Silvano, Rui Sousa-Silva (éds). *A linguística em diálogo. Volume comemorativo dos 40 anos do Centro de Linguística da Universidade do Porto*. Centro de Linguística Universidade do Porto, 195-211.

Eça de Queirós 1888 = José M. Eça de Queirós 1888. *Os Maias*. Livraria Lello & Irmão.

Favaro 2025 = Marco Favaro 2025. The interplay between illocutionary force and information management: On the discourse-pragmatic functions of Italian *un po'* 'a bit' and European Portuguese *lá* 'there'. *Journal of Portuguese Linguistics* 24.1. https://doi.org/10.16995/jpl.15408.

Franco 1988 = António C. Franco 1988. Partículas modais da língua portuguesa. Relances contrastivos com as partículas alemãs. *Línguas e Literatura. Revista da Faculdade de Letras do Porto* 5.1, 137-156. https://ojs.letras.up.pt/index.php/rll/article/view/8439.

Frantext = Analyse et traitement informatique de la langue française (ATILF) (éd.) 1998-. *Base textuelle Frantext*. http://www.frantext.fr.

French Web 2023 = Lexical computing 2023. *French Web 2023*. https://www.sketchengine.eu/frtenten-french-corpus/.

Giannakidou & Mari 2025 = Anastasia Giannakidou, Alda Mari 2024. Modalization and bias in questions. Tue Trinh, Anton Benz, Daniel Goodhue, Kazuko Yatsushiro, Manfred Krifka (éds.). *Biased questions. Experimental results and theoretical modelling*. Language Science Press, 1-36. https://langsci-press.org/catalog/book/487.

Gómez González & Silvano 2022 = Maria de los Ángeles Gómez González, Maria da Purificação Moura Silvano 2022. A functional model for the tag question paradigm: The case of invariable tag questions in English and Portuguese. *Lingua* 272. 103255, 1-29. https://doi.org/10.1016/j.lingua.2022.103255.

DOI : https://doi.org/10.25364/19.2026.17.1 ISSN : 2663-9815

Gross 1981 = Maurice Gross 1981. Les bases empiriques de la notion de prédicat sémantique. *Langages* 63, 7-52. https://www.persee.fr/doc/lgge_0458-726x_1981_num_15_63_1875.

Krifka 2015 = Manfred Krifka 2015. Bias in commitment space semantics: Declarative questions, negated questions, and question tags. Sarah D'Antonio, Mary Moroney, Carol Rose Little (éds.). *Proceedings of the 25th Semantics and Linguistic Theory Conference. Stanford University, May 15-17, 2015*. Linguistic Society of America, Cornell Linguistics Circle, 328-345. https://doi.org/10.3765/salt.v25i0.3078.

Lefeuvre 2012 = Florence Lefeuvre 2012. *Eh bien* comme évaluateur de discours. *Travaux de Linguistique* 65, 123-143. https://shs.cairn.info/revue-travaux-de-linguistique-2012-2-page-123?lang=fr.

Legallois 2006 = Dominique Legallois 2006. Quand le texte signale sa structure : la fonction textuelle des noms sous-spécifiés. *Corela. Cognition, représentation, langage* HS-5, 1-24. https://doi.org/10.4000/corela.1465.

Lopes & Carrilho 2020 = Ana C. Macário Lopes, Ernestina Carrilho 2020. Discurso e marcadores discursivos. Eduardo Paiva Raposo, Maria F. Nascimento, Maria A. Mota, Luísa Segura, Amália Mendes (éds.). *Gramática do português. Vol. 3*. 2e édition. Fundação Calouste Gulbenkian, 2667-2698.

Marques & Duarte 2015 = M. Aldina Marques, Isabel M. Duarte 2015. *Cá* e *lá*: Atenuação, reforço e outros valores modais em PE. *Acta semiótica et lingvistica* 20.2, 115-128. https://periodicos.ufpb.br/index.php/actas/article/view/28020.

Marques & Duarte 2017 = M. Aldina Marques, Isabel M. Duarte 2017. *Lá*, atenuador em interações informais do Português Europeu. *Studia Universitatis Babeş-Bolyai Philologia*. LXII 4, 17-34.

Marques & Duarte 2019 = M. Aldina Marques, Isabel M. Duarte 2019. *Lá* em sequências narrativas orais. M. Aldina Marques, Xosé M. Sánchez Rei (éds.). *Estudos atuais de linguística galego-portuguesa*. Laiovento, 211-238.

MD = Marqueur discursif.

Meisnitzer 2012 = Benjamin Meisnitzer 2012. Modality in the Romance languages: Modal verbs and modal particles. Werner Abraham, Elisabeth Leiss (éds.). *Modality and theory of mind elements across languages*. De Gruyter Mouton, 335-360.

Meisnitzer & Müller 2024 = Benjamin Meisnitzer, Lukas Müller 2024. Partículas modais e partículas modalizadoras no português europeu numa perspetiva sincrónica e diacrónica. María de F. Gil, Rute Soares, Rogério Madeira (éds.). *Olhares cruzados: espaço de língua portuguesa – espaço de língua alemã*. CITCEM, 223-238. https://ler.letras.up.pt/uploads/ficheiros/20229.pdf.

Molinier & Levrier 2000 = Christian Molinier, Françoise Levrier 2000. *Grammaire des adverbes. Description des formes en* -ment. Droz.

Müller 2024 = Lukas Müller 2024. European Portuguese *lá*: Use-conditional meaning and pragmaticalization. *Languages* 9.6.189, 1-21. https://doi.org/10.3390/languages9060189.

Müller, Mamede & Baptista 2023 = Izabela M. G. Müller, Nuno Mamede, Jorge M. Baptista 2023. Advérbios compostos do português do Brasil. *Revista da Associação portuguesa de linguística* 10, 230-250. https://doi.org/10.26334/2183-9077/rapln10ano2023a13.

Müller, Mamede & Baptista 2024 = Izabela M. G. Müller, Nuno Mamede, Jorge M. Baptista 2024. Diferenciação dos advérbios compostos do português brasileiro e do português europeu. *Revista da Associação portuguesa de linguística* 11, 136-163. https://doi.org/10.26334/2183-9077/rapln11ano2024a7.

ParlE = Parlement Européen (éd.) 1994-. *Séances plénières. Débats et vidéos*. https://www.europarl.europa.eu/plenary/fr/debates-video.html.

Pons Bordería 1998 = Salvador Pons Bordería 1998. *Conexión y conectores: Estudio de su relación*

DOI : https://doi.org/10.25364/19.2026.17.1 ISSN : 2663-9815

en el registro informal de la lengua. Universitat de Valencia. https://www.uv.es/ponss/PDFs/Conexion.pdf.

Roda Viva = Isaac Souza de Miranda Jr., Gabriela Wick-Pedro, Cláudia Dias de Barros, Oto Araújo Vale 2024. *Corpus Roda Viva*. https://github.com/LeGOS-UFSCar/Roda-Viva.

Romanian Web 2021 = Lexical computing 2021. *Romanian Web 2021*. https://www.sketchengine.eu/rotenten-romanian-corpus/.

Schiffrin 1987 = Deborah Schiffrin 1987. *Discourse markers*. Cambridge University Press.

Schmid 2000 = Hans-Jörg Schmid 2000. *English abstract nouns as conceptual shells. From corpus to cognition*. Mouton de Gruyter.

Silvano, Gómez González & Cordeiro 2025 = Maria da Purificação Moura Silvano, Maria de los Ángeles Gómez González, João Cordeiro 2025. Tag questions in English and Portuguese monologues: types, features and functions. *Linguistics Vanguard* 11.s2, 139-158.

Zufferey & Degand 2024 = Sandrine Zufferey, Liesbeth Degand 2024. Defining connectives and discourse relations. Sandrine Zufferey, Liesbeth Degand (éds.). *Connectives and discourse relations*. Cambridge University Press, 1-22.

DOI : https://doi.org/10.25364/19.2026.17.1 ISSN : 2663-9815