Marqueurs discursifs construits à partir de *shell nouns* en français et en roumain[1]

Discourse markers based on shell nouns in French and in Romanian

Cristina Petraş

Universitatea *Alexandru Ioan Cuza* din Iaşi (Iaşi, Roumanie)
cristina.petras@uaic.ro
https://orcid.org/0000-0002-0093-8002

Reçu le 6/11/2025, accepté le 28/4/2026, publié le 20/11/2026

Creative Commons Attribution 4.0 International (*CC BY 4.0*)
© 2026 Cristina Petraş

Pour citer cet article
Petraş, Cristina 2026. Marqueurs discursifs construits à partir de *shell nouns* en français et en roumain. *Studia linguistica romanica* 2026.17, 129-155. https://doi.org/10.25364/19.2026.17.6.

1 Cet article a été réalisé dans le cadre d'un projet de recherche financé par Ministerul Educației şi Cercetării, CNCS/CCCDI – UEFISCDI, PNCDI III, projet no. PN-III-P4-ID-PCE-2020-1505.

Résumé
Tout en servant à la focalisation de la séquence qu'elles introduisent, les constructions du type *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* etc., viennent marquer différentes relations discursives avec ce qui précède (d'explication, causale, d'énumération, contrastive) et par conséquent elles contribuent à la structuration discursive. Dans cet article nous nous proposons d'interroger dans une perspective comparative, à partir de corpus comme *French Web 2023*, *Romanian Web 2021* et *Frantext,* les types de fonctionnement sémantico-pragmatique de ces expressions, leur interaction avec le co(n)texte (définir, préciser, expliciter le rapport entre le contexte de gauche et le contexte de droite). Dans une perspective typologique, on s'intéressera aux propriétés des noms susceptibles de fonctionner dans des constructions qui deviennent marqueurs discursifs et aux mécanismes de changement.

Mot-clés
Nom capsule, focalisation, constructionnalisation.

Abstract
Constructions like *le fait est que* / *faptul este că* 'the fact is that', *le problème est que* / *problema este că* 'the problem is that', *l'idée est que* / *ideea este că* 'the idea is that', *la vérité est que* / *adevărul este că* 'the truth is that' etc. focalize the sequence they introduce and mark different discourse relations (explanatory, causal, enumerative, contrastive) with the preceding sequence. Thereby they contribute to discourse structuring. In this article, we examine the semantic-pragmatic functions of these expressions and their interaction with the co(n)text (defining, specifying, and explicating the relationship between the left and right contexts). From a typological perspective, we focus on the properties of nouns likely to function in constructions that become discourse markers and on the mechanisms of change. The study's approach is comparative, using corpora such as *French Web 2023*, *Romanian Web 2021* and *Frantext*.

Keywords
Shell noun, focalization, constructionalization.

DOI : https://doi.org/10.25364/19.2026.17.6 ISSN : 2663-9815

Sommaire

1 Introduction....132
2 Noms sous-spécifiés / noms coquilles / noms capsules....133
3 Des structures spécificationnelles aux marqueurs discursifs....136
3.1 *Le fait est que* / *fapt(ul) e(ste) că*....138
3.2 *Le problème est que* / *problema e(ste) că*....144
3.3 *L'idée est que* / *ideea este că*....146
3.4 *La vérité est que* / *adevărul este că*....148
3.5 *Le point est que* / *punctul este că*....149
3.6 *La chose est que* / *lucrul este că*....149
3.7 *Le truc est que* / *chestia este că* / *faza este că*....150
4 Conclusions....151
Abréviations et références bibliographiques....154

DOI : https://doi.org/10.25364/19.2026.17.6 ISSN : 2663-9815

1 Introduction

[1] Des expressions comme *le fait est que* / *fapt(ul) este că*, *le problème est que* / *problema este că*, *l'idée est que* / *ideea este că*, *la vérité est que* / *adevărul este că*, etc. fonctionnent respectivement en français et en roumain comme des structures qui, malgré l'aspect formel de matrice, ne portent pas l'information principale. C'est tout au contraire la séquence apparemment subordonnée du point de vue formel qui présente le sens assertif, alors que les expressions mentionnées, tout en attirant l'attention sur la séquence introduite, viennent marquer des rapports divers avec ce qui précède et par conséquent elles contribuent à la structuration discursive. Ainsi, dans l'exemple (1) ci-dessous, *le fait est que* introduit une justification du dire antérieur : la conclusion conformément à laquelle la maîtrise de l'arabe constitue un important atout sur le marché du travail est justifiée par le recours à l'argument selon lequel l'arabe est une langue internationale importante.

(1) instits.org[2]
Si en plus de vos compétences professionnelles, vous avez une parfaite maîtrise de l'arabe, alors vous disposez d'un important atout. *Le fait est que* l'arabe occupe une place de choix dans les affaires étrangères et le commerce international. Il s'agit d'une qualité très recherchée qui privilégie votre candidature pour des postes de grande envergure dans le commerce international, le tourisme et bien d'autres secteurs juteux.

Considérées comme marqueurs pragmatiques, marqueurs discursifs ayant une fonction de focalisation ou bien formules, ces expressions ont fait l'objet d'une série importante d'études, portant notamment sur l'anglais ou le français (pour ne citer ici que Schmid 2000, 2018 ; Legallois 2006 ; Aijmer 2007 ; Hundt 2022).

[2] L'emploi retenu est lié aux fonctions de la classe des noms sous-spécifiés, pour utiliser le terme de Legallois (2006), que celle-ci associe à l'indexicalité[3], à la modalité, à la connexité, ces propriétés se manifestant à des degrés différents parmi les différents éléments de la classe. La valeur de connexité s'exprime notamment dans le type de structure auquel nous avons déjà fait référence, qui est considéré comme une construction spécificationnelle. Ce type de construction vient marquer différentes relations discursives : d'explication (causale), d'énumération, contrastive. Rappelons que dans la perspective des grammaires de construction,

2 Sauf indication autre, les exemples proviennent des corpus *French Web 2023* et *Romanian Web 2021*, consultés sur la plateforme *Sketch engine*.

3 Toujours selon Legallois (2006 [78]), ces noms se caractérisent par un « comportement mi-pronominal, mi-nominal » : ils partagent certaines propriétés des noms, mais présentent aussi un fonctionnement qui les rend dépendants du contexte (indexical, proche du fonctionnement des pronoms). De la même manière, Schmid (2000 : 18), s'interrogeant sur le type de concepts auxquels renvoient les *shell nouns*, affirme que « the concepts created by shell nouns consist of a stable symbolic and a variable indexical part ».

DOI : https://doi.org/10.25364/19.2026.17.6 ISSN : 2663-9815

> [e]st considérée comme construction une unité qui associe 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 : 11)

Les constructions sont caractérisées par des degrés différents de schématicité et d'abstraction. Dans ce type d'approche théorique qui propose une autre manière de rendre compte du rapport entre lexique et syntaxe – qui ne sont pas envisagés comme clairement séparés –, les constructions sont considérées carrément comme des *unités de la langue* (Legallois & Patard 2017). Dans la suite de notre article, le terme *construction* sera utilisé avec ce sens spécialisé, pour désigner un certain type de structure spécificationnelle [N *être que* + P conjonctive], dont la signification n'est pas entièrement compositionnelle. C'est cette construction en particulier que nous retenons dans la perspective de l'émergence des marqueurs discursifs.

[3] Dans cet article nous nous proposons d'interroger dans une perspective comparative, à partir de corpus comme *French Web 2023*, *Romanian Web 2021* et *Frantext*, les types de fonctionnement sémantico-pragmatique des expressions retenues, leur interaction avec le co(n)texte (définir, préciser, expliciter le rapport entre le contexte gauche et le contexte droit). Dans une perspective typologique, seront scrutées les propriétés des unités lexicales susceptibles de fonctionner dans des structures qui deviennent marqueurs discursifs et aux mécanismes de changement (constructionnalisation, grammaticalisation/pragmaticalisation).

[4] Dans § 2 nous passerons en revue les particularités des noms sous-spécifiés qui leur permettent d'entrer dans la constitution des énoncés spécificationnels. § 3 sera consacrée à l'analyse comparative des emplois d'une série d'expressions autour de noms sous-spécifiés, qui, dans le patron d'une construction spécificationnelle, fonctionnent comme marqueurs discursifs. La section de conclusions nous permettra de relever les points communs et les divergences entre les deux langues (français et roumain) et, à l'intérieur d'une même langue, entre des expressions construites à partir de noms différents. Des renvois ponctuels seront faits à l'anglais pour élargir la portée des conclusions notamment du point de vue typologique.

2 Noms sous-spécifiés / noms coquilles / noms capsules

[5] Les noms à la base des expressions qui font l'objet de ce travail ont été considérés comme des *noms sous-spécifiés* (Legallois 2006, 2008), *noms porteurs* (Huyghe 2018), *noms capsules* (Huyghe 2021), *containers* (Vendler 1968, cf. Huyghe 2021) ou *shell nouns* (Schmid 2000). Dans les études portant sur les noms sous-spécifiés, on se pose souvent la question de leur rapport avec les noms généraux. Des auteurs comme Cappeau & Schnedecker (2021) proposent de distinguer entre les deux catégories, à partir d'une série de critères identifiés dans des études portant sur le sujet, alors que Vajnovszki (2025) identifie une classe à part de noms sous-spécifiés et des classes proches, dont celle des noms généraux.

DOI : https://doi.org/10.25364/19.2026.17.6 ISSN : 2663-9815

D'autres auteurs ne distinguent pas entre les deux classes. Ainsi, Adler (2017) traite indifféremment des noms généraux ou sous-spécifiés. Troisièmement, les noms sous-spécifiés sont envisagés comme une sous-classe des noms généraux (Adler & Legallois 2018).

[6] À partir de l'approche de Cappeau & Schnedecker (2021), on pourra énumérer comme traits partagés par les deux catégories : un type de sémantisme général ou non spécifié qui dépend du contexte, la haute fréquence, le fort potentiel cohésif (l'emploi anaphorique). Prenant comme point de départ la définition et la classification des noms généraux de Halliday & Hasan (1976), où l'on retrouve des noms considérés ailleurs comme sous-spécifiés (*question*, *idea*)[4], Vajnovszki (2025 : 63-64) note que, si des ressemblances peuvent être identifiées entre les deux classes (par exemple, leur fonctionnement anaphorique déjà cité ci-dessus), ils diffèrent sur d'autres points – si les noms généraux sont des hypéronymes, les noms sous-spécifiés « sont avant tout des coquilles vides », ce qui rend compte de leur fonctionnement anaphorique. D'autres éléments différenciateurs sont : le caractère abstrait des noms sous-spécifiés, leur signification sous-spécifiée de nature dénotationnelle, leur rapport important avec la textualité, certains des noms sous-spécifiés dénommant des séquences d'enchaînement (Adler & Legallois 2018) (v. [8]). Si la généralité sémantique peut être invoquée aussi pour les noms sous-spécifiés, Vajnovszki (2025 : 57) l'envisage comme un continuum, se manifestant de façon plus ou moins marquée selon le nom. La caractéristique syntaxique qui fait de ces noms des noms sous-spécifiés est de figurer dans des structures spécificationnelles, qui apportent la spécification nécessaire à leur fonctionnement.

[7] La dépendance du contexte et la complémentation sémantique contextuelle est qualifiée en termes d'*encapsulation propositionnelle* par Huyghe (2021)[5]. Selon Legallois (2008), plus que de sous-spécification sémantique, il s'agit de sous-spécification informationnelle. La relation d'équivalence qui s'établit entre le nom et la séquence équivalente est contingente, résultant d'une expérience propre, donc non codée, comme le souligne Adler (2017 : 54) : « La relation entre référent et nom n'est pas codée, fixée ou nécessaire ; elle n'est pas inhérente ou obligatoire ; elle n'est ni scientifique ni lexicalisée, non plus ; elle est plutôt expérimentale ou ad hoc ».

[8] Du fait de leur contenu sémantique qui dépend du contexte, certains des noms sous-spécifiés (*problème*, *question*, *situation*, *réponse*, *solution*, *objectif*, *désir*, etc.) présentent des propriétés qui leur permettent d'avoir une fonction de connexion : comme mentionné dans [6], ils signalent les structures d'enchaînement dans un texte, contribuant à la structuration et à la cohésion textuelle (Legallois 2006 ; Stosic & Rebeyrolle 2021). Ceux-ci sont appelés par Legallois (2006)

4 Un nom comme *vérité* fonctionne, selon Vajnovszki (2025 : 68), tantôt comme nom général, tantôt comme nom sous-spécifié.

5 L'idée d'encapsulation est reprise, parmi d'autres, chez Francis (1986) (v. Schmid 2000 : 12).

DOI : https://doi.org/10.25364/19.2026.17.6 ISSN : 2663-9815

noms sous-spécifiés marqueurs. Ainsi, on peut résumer, avec Adler (2017), que les fonctions de ces noms dans les textes sont la structuration, la référence, l'encapsulation, la cohésion.

[9] Huyghe (2021) observe que la terminologie utilisée pour désigner la classe de ces noms dépend de la composante qui est mise en avant – leur propriété sémantique ou leur rôle dans la cohésion discursive : *noms sous-spécifiés*, *noms coquilles*, *containers*, *carrier nouns* / *noms porteurs* pour la première ; *labels*, *signalling nouns* / *noms signalants* pour le second. Par ailleurs, cet auteur privilégie la dénomination de *noms capsules*, à la place de celle de *noms sous-spécifiés* de Legallois (2006). Schmid (2000) rend compte de la dénomination de *shell nouns* (qu'il est le premier à avoir utilisé) par le fait que les noms en question sont utilisés pour désigner des coquilles conceptuelles pour les éléments d'information qui se manifestent dans les phrases ou les séquences de phrases de l'environnement discursif. Schmid (2000) définit la relation entre ces noms et la phrase postnominale (*postnominal clause*) comme une *experiential identity*. Si dans le titre de l'article nous avons retenu le terme anglais *shell nouns*, dans la suite du texte nous utiliserons les termes noms coquilles et noms sous-spécifiés.

[10] Parmi les noms coquilles, certains participent de structures spécificationnelles qui connaissent une grammaticalisation très avancée (v., par exemple, *le fait est que*, *le problème est que*, etc.), où le sémantisme de la structure s'avère non compositionnel. La contribution des constructions spécificationnelles à la cohérence discursive est favorisée par la fonction méta-(discursive ou langagière) que Cappeau & Schnedecker (2021) prêtent à des noms comme *problème*, *raison*, *idée*, etc.

[11] Le choix des éléments à analyser dans ce texte s'est fait en fonction de leur typicalité en tant que représentants de la classe des *shell nouns*. Il s'agit de ce que Schmid (2000 : 85) appelle *prime shell nouns*[6] : parmi d'autres, *fact*, *problem*,

6 Schmid (2000) identifie six classes sémantiques de *shell nouns* : *factual* (*fact*, *thing*, *point*, *problem*, *reason*, *difference*, *upshot*), *linguistic* (*news*, *message*, *rumour*, *report*, *order*, *proposal*, *question*), *mental* (*idea*, *notion*, *belief*, *assumption*, *aim*, *plan*, *decision*), *modal* (*possibility*, *truth*, *permission*, *obligation*, *need*, *ability*), *eventive* (*act*, *move*, *measure*, *reaction*, *attempt*, *tradition*, *trick*), *circumstantial* (*situation*, *context*, *place*, *area*, *time*, *way*, *approach*). Comme le montre ce même auteur, ces noms présentent les caractéristiques qui en font des *shell nouns* à des degrés différents, certains plus que d'autres. Certains sont des prototypes de la classe, d'autres se retrouvent à la périphérie de la classe. À part le classement de Schmid (2000), d'autres ont été proposés. Considérant les noms sous-spécifiés comme « des expressions fondamentalement modales », Adler & Legallois (2018 : 30-31) proposent un classement en six classes modales : aléthique (*fait*, *évidence*, *chose*, *logique*, *vérité*, *réalité*, *preuve*, *résultat*, *effet*, *conséquence*, *condition*, *façon*, *méthode*, *moyen*, *raison*, *motif*) ; épistémique (*sentiment*, *certitude*, *crainte*, *décision*, *conception*, *conviction*, *politique*, *question*, *réponse*, *argument*, *conclusion*, *propos*, *enseignement*, *mesure*) ; appréciatif (*comble*, *essentiel*, *exigence*, *important*, *inconvénient*, *intérêt*, *mieux*, *pire*, *progrès*, *urgence*, *erreur*, *scandale*, *drame*, *paradoxe*, *écueil*, *problème*, *risque*, *solution*, *différence*, *nouveauté*) ; axiologique (*courage*, *honneur*, *mérite*, *sagesse*, *vertu*, *habileté*, *faute*) ; boulique (*alternative*, *ambi-*

DOI : https://doi.org/10.25364/19.2026.17.6 ISSN : 2663-9815

thing, *idea*, *reason*, *notion*, *concept*, *issue*, *principle*, *rumour*, *legend*, *point*, *case*. Pour ce qui est des classes sémantiques des noms retenus, il s'agit notamment des noms factuels, mais aussi des noms mentaux et des noms linguistiques. Ce sont des noms d'un sémantisme très général, « nouns representing third-order entities which are the only means of conceptualizing the particular type of experience they represent » (Schmid 2000 : 85).

3 Des structures spécificationnelles aux marqueurs discursifs

[12] Les différentes constructions dans la constitution desquelles entrent les noms sous-spécifiés (Adler & Legallois 2018 en comptent onze) ont été étudiées pour le français et le roumain par Vajnovszki (2025) dans sa thèse, à travers une analyse outillée, qui prend en compte des éléments comme l'impact du genre textuel, les préférences constructionnelles des noms sous-spécifiés, leurs valeurs sémantiques.

[13] Comme déjà précisé, dans ce travail, nous nous intéressons pratiquement à une seule construction spécificationnelle, à savoir [N + *être que* complétive], sous une forme particulière, celle qui donne lieu à des fonctionnements comme marqueurs discursifs. Le tableau ci-dessous présente les expressions retenues en français et en roumain. Nous y avons inclus les expressions correspondantes en anglais.

	Français	**Roumain**	**Anglais**
fait / fapt / fact	*le fait est que*	*faptul / fapt este că*	*(the) fact is (that)*
problème / problemă / problem	*le (gros) problème est que*	*problema (marea problemă) este că*	*the (big) problem is (that)*
idée / ideea / idea	*l'idée est que*	*ideea este că*	*the idea is (that)*
? point / Ø / point	*? le point est que*	*Ø*	*the point is (that)*
vérité / adevăr / truth	*la vérité est que*	*adevărul este că*	*the truth is (that)*
chose / ? lucru / thing	*la chose est que*	*Ø*	*(the) thing is (that)*
truc / chestie / față / thing	*le truc est que*	*(toată) chestia este că / faza este că*	*(the) thing is (that)*

Tableau 1 : Constructions spécificationnelles étudiées

Pour la plupart de ces expressions, on remarquera des parallélismes entre les trois langues. Il existe aussi des cas particuliers à chacune ou des absences dans l'une ou autre. Par exemple, le français et le roumain ne présentent pas d'expression équivalente à *the point is that*, l'expression enregistrée en français semblant être un calque sur l'anglais. En roumain, il ne s'est pas à proprement parler développé d'équivalent pour l'anglais *the thing is that*. Par contre, *lucru* se retrouve dans d'autres constructions spécificationnelles, comme le montre Vajnovszki (2025).

tion, *but*, *choix*, *intention*, *objectif*, *objet*, *préoccupation*, *désir*, *volonté*, *priorité*, *projet*, *esprit*, *idée*, *proposition*) ; déontique (*devoir*, *responsabilité*, *mission*, *rôle*, *principe*).

DOI : https://doi.org/10.25364/19.2026.17.6 ISSN : 2663-9815

[14] Dans les expressions qui nous intéressent ici, la construction spécificationnelle [N *être que* + P conjonctive] se fige sous la forme *article défini* + *nom* + *est* (présent de l'indicatif) + *conjonction*. Les différentes structures connaissent des degrés différents de figement, des plus figées (*le fait est que*) aux moins figées. L'utilisation d'un autre déterminant, comme *notre* dans l'exemple suivant, est associé à une construction libre :

(2) senat.fr
Sur ces dernières semaines, nous avons installé sept personnes de la police française en Roumanie, pour six mois (...). J'en ai prépositionné quatre en Bulgarie sur l'affaire des réseaux de trafiquants de prostitution (...). En Colombie, nous allons aussi installer un certain nombre de fonctionnaires. *Notre idée est que* plus nous les installons en amont, plus nous sommes efficients et moins nous avons besoin de fonctionnaires.

Dans une seule situation (*fapt*), le fonctionnement sans déterminant est possible (v. § 3.1)[7]. La dislocation et la reprise anaphorique indiquent aussi un degré moindre de figement. Le phénomène est possible avec des noms comme *problème*, *idée*, *truc* (*le problème, c'est que* ; *l'idée, c'est que* ; *le truc, c'est que*), mais non avec *fait* (**le fait, c'est que*).

[15] Nous suivons ainsi Hundt (2022) qui considère qu'il s'agit de constructions (avec l'acception de la grammaire de construction) semi-figées, en rajoutant à son argument concernant la compositionnalité du sens, un autre : si c'est le patron syntaxique article défini + nom + verbe *être* (+que) qui privilégie l'émergence du marqueur discursif, on remarque, pour le français *fait* la variation pour ce qui est de la forme verbale (*est* / *était*).

[16] La position à gauche semble être prototypique. Des exemples de rejet à la périphérie droite comme (3) sont rares dans le corpus exploité. Hundt (2022) signale la même situation pour l'anglais.

(3) april.org
Par exemple si quelqu'un s'exprime d'une manière un peu véhémente contre un ministre, se pose la question : « Peut-il être passible de poursuites pour diffamation » ? Il est arrivé que le gouvernement ou des tiers demandent à un hébergeur ou au prestataire, un site internet, de retirer ses contenus, *le fait est*.

Le fait est apparaît aussi, dans le corpus analysé, en position médiane :

7 Il ne s'agit pas pour autant d'une nouvelle construction. La construction spécificationnelle étant [N *être que* + P conjonctive], les différentes réalisations du déterminant (article défini, possessif) ou son absence ne devraient pas être associées, selon nous, à de nouvelles constructions, mais elles devraient être ramenées à une même structure d'un degré plus élevé d'abstraction.

DOI : https://doi.org/10.25364/19.2026.17.6 ISSN : 2663-9815

(4) carjager.com
Le moteur central combiné à un empattement relativement long (2,73 m) contribue à définir un équilibre particulièrement exigeant si l'on s'avise de pousser la voiture dans ses retranchements – d'autant qu'en dehors de l'ABS, les seules béquilles électroniques (entièrement déconnectables) se résument à un antipatinage et à un différentiel à commande automatique. […] La Carrera GT, *le fait est*, n'est pas l'une de ces machines bardées de fonctionnalités intrusives vouées à pallier l'incompétence potentielle de leurs conducteurs.

La place dans la périphérie gauche de l'expression (*pre-front field*) est associée par Aijmer (2007) avec le fonctionnement comme marqueur pragmatique. Comme telle, l'expression a une fonction de cadrage ou de signalement – il s'agit d'indiquer le type de relation entre les deux séquences, impliquant des fonctions interpersonnelles (modale et évaluative) et rhétoriques.

[17] Là où le contenu sémantique du nom est plus spécifique, on remarque la possibilité d'intégrer un modificateur : *le gros problème est que*, *marea problemă este că*, *the big problem is that*, *toată chestia este că*.

[18] En termes d'organisation de l'information, dans le type de construction spécificationnelle étudié, on voit à l'œuvre une double opération de thématisation/topicalisation : un *problème*, un *fait*, une *solution* apparaît comme thème (v. l'utilisation de l'article défini, lié au caractère thématique du nom, renvoyant à une réalité présupposée comme connue) ; la séquence introduite par la structure [Le N *est que*], qui correspond au rhème (à l'apport informationnel proprement dit), se trouve focalisée (Legallois 2006 ; Huyghe 2018). À son tour, Aijmer (2007) montre, pour l'anglais, que, malgré l'aspect formel qui indiquerait une subordination introduite par la conjonction *that*, ce n'est pas *the fact is* qui contient l'information principale, mais bien la séquence introduite par *that*. De la même manière, Delahunty (2012 : 60-61) soutient que la valeur de vérité de la phrase entière implique la valeur de vérité de la complétive et vice versa. Comme la matrice ne participe pas aux conditions de vérité, c'est la complétive qui porte le contenu informationnel principal. En citant Huyghe (2018), on pourra retenir deux valeurs des structures de type [Le N *est que* complétive] : focalisation de la complétive et recatégorisation selon la signification du nom porteur, pour utiliser la terminologie de l'auteur.

3.1 *Le fait est que* / *fapt(ul) e(ste) că*

[19] Le nom *fait* (*fapt* en roumain, *fact* en anglais) semble être un prototype de la classe envisagée et par conséquent il a fait, par excellence, l'objet d'études portant sur la catégorie des noms en question et sur les constructions spécificationnelles auxquelles ils donnent lieu. *Fait* se caractérise, selon Huyghe (2021, [60]), par un « haut degré de généralité sémantique », même dans les emplois qui

DOI : https://doi.org/10.25364/19.2026.17.6 ISSN : 2663-9815

ne dépendent pas du contexte. Son fonctionnement est décrit dans les termes suivants :

> Le rôle de *fait* est essentiellement de permettre à un contenu propositionnel d'accéder à la référence nominale et ce faisant, de réifier un contenu abstrait. Ne décrivant pas la forme de l'encapsulation, *fait* s'emploie souvent comme simple introducteur de propositions, favorisant par son statut nominal la topicalisation. Ainsi, l'emploi d'une proposition conjonctive en position de sujet (i.e. canoniquement comme thème de l'énoncé) est facilité lorsque celle-ci est introduite par *le fait*. (Huyghe 2021, [61])

Relevant, selon Schmid (2000), de la catégorie des noms capsules (*shell nouns*) factuels, *fact* présente en anglais des emplois neutres. Dans la construction spécificationnelle qui fait l'objet de ce travail, il s'agit d'une fonction de focalisation. Selon le même auteur, la structure *the fact is that* met en évidence l'affirmation selon laquelle le contenu de la phrase introduite est vrai (v. aussi le classement modal d'Adler & Legallois (2018), selon lequel *fait* relève de la catégorie aléthique). La nécessité d'une telle stratégie devient compréhensible puisqu'en fait le besoin de mettre en évidence que la phrase introduite est vraie laisse entendre qu'on s'attendrait à autre chose, d'où la signification contrastive. Eu égard au sens factuel du nom, la phrase introduite sera notamment construite avec des formes verbales de présent et de passé.

[20] La question du degré de figement de l'expression *le fait est que* est sujet à débat. Ainsi, selon Van de Velde (2021), *le fait est que* est une expression figée. Témoignent du figement, selon cette auteure, l'impossibilité de remplacement du verbe *être* par *consister* (comme c'est le cas avec une complétive comme attribut) (**Le fait consiste qu'il a réussi* vs. *Le fait important consiste en ce qu'il a réussi*) ; la liaison obligatoire entre *fait* et *est*, mais non quand *fait* est accompagné d'un modificateur (*Le principal fait est qu'on a trouvé une autre possibilité de trouver une planète habitable*)[8]. Pour Lenepveu (2009), la locution *le fait est que* reste relativement figée au niveau syntaxique (v., par exemple, la possibilité d'intégrer à l'intérieur de l'expression un connecteur comme *cependant* ou *donc* ou un adjectif), la signification étant partiellement non transparente. Avec l'utilisation du verbe *être* à l'imparfait comme dans (5) (cas de figure que Lenepveu (2009) relève dans *Frantext* dans un texte du 19e siècle), il s'agit de transposer dans le registre du récit la forme qui relie les deux séquences :

(5) forumsactifs.com
Pour l'heure, Tiky se contentait de mettre sa main en visière afin de ne pas se faire aveugler par les rayons oranges du soleil qui apparaissaient à l'autre bout de la rue, qu'il traversait pour se rendre au lieu où il devait retrouver Asch. *Le fait était que* les jumeaux avaient eu un moment de libre dans la

8 Dans ce dernier cas, on peut utiliser soit le subjonctif, soit l'indicatif.

DOI : https://doi.org/10.25364/19.2026.17.6 ISSN : 2663-9815

journée et n'avaient rien trouvé de mieux que de se rendre en ville pour se dégourdir les jambes.

Force est de noter que c'est toujours la forme d'imparfait qui est utilisée dans ce genre de contextes, jamais le passé composé.

[21] Du point de vue de la signification, le caractère figé de l'expression peut être vérifié par la possibilité de la remplacer par des unités comme *certes*, *assurément* (v. Lenepveu 2009). Pour ce qui est du fonctionnement sémantico-pragmatique, nous observerons que *TLFi* (s. v. *fait*) enregistre l'expression *le fait est que* comme « ce qui est arrivé, ce qui existe » et, plus particulièrement, « ce qui existe réellement, la chose réelle (…), tout ce qui peut être constaté de manière certaine ». L'expression *le fait est que* y est paraphrasée par « il est vrai, indéniable que, il faut admettre que » et on y propose comme expressions équivalentes *il est de fait que*, *c'est un fait que*.

[22] Les emplois que nous avons identifiés correspondent aux paraphrases mentionnées ci-dessus. Ainsi, *le fait est que* apparaît pour introduire une séquence d'une même orientation argumentative : celle-ci se présente comme une justification de l'énoncé antérieur, même comme un argument supplémentaire en faveur de l'idée soutenue (v. déjà (1)). Dans (6), par exemple, il s'agit d'étayer l'idée conformément à laquelle les écoles ne s'intéressent pas à l'orientation professionnelle des futurs étudiants :

(6) newsnet.fr
J'ai fait des études de Droit parce que j'aimais la discussion, et tout le monde me disait que je ferais un bon avocat. Je n'ai pas bénéficié de la moindre orientation professionnelle. *Le fait est que* les écoles d'élite lançaient à la rue des vagues de jeunes bacheliers totalement dépourvus de connaissances politiques élémentaires.

Cette observation rejoint la conclusion de Van de Velde (2021, [73]), conformément à laquelle « il existe une expression *le fait est*, non modifiable, qui est une sorte de renforcement de la force assertive de la phrase ».

[23] Dans un exemple comme (7) la séquence introduite par *le fait est que* s'avère nécessaire pour rendre compte de l'utilisation d'un mot qui peut être perçu comme fort pour qualifier le vote (*psychodrame*). Dans un premier temps l'emploi de ce mot est justifié par la séquence introduite par *car*. Tout ce développement (v. aussi la suite : *Raison pour laquelle ...*) va dans le sens de l'idée qu'un projet de référendum doit être bien préparé avant que le peuple ne soit appelé pour s'y exprimer :

(7) ge.ch
Ce vote a été le psychodrame de 1996, car le projet n'était pas défini, et ce n'est qu'après qu'on a eu ce vote – pont, tunnel, petit tracé, grand tracé –

DOI : https://doi.org/10.25364/19.2026.17.6 ISSN : 2663-9815

qui sortait de l'initiative de 1988. *Le fait est que* cela a perturbé la décision du peuple, qui aurait dû pouvoir se prononcer sur un objet précis. Raison pour laquelle nous souhaitons que le peuple puisse, cette fois, voter sur un projet particulier, qui a effectivement fait l'objet d'études.

Si, comme le soutient Van de Velde (2021), *fait* semble inutile, car formuler une assertion constitue du même coup un fait (v. la valeur aléthique), il ne l'est pas du point de vue de l'outil énonciatif qu'il est :

> Unique, donc, même à l'intérieur de sa propre sous-classe, le nom *fait* l'est encore en ceci qu'il est, d'une certaine façon, tout à fait inutile (sauf, bien entendu, comme outil énonciatif - et en tant que tel il est irremplaçable) : en effet, il n'est besoin, pour « dire » qu'une situation ou un événement constitue un fait, que de former une phrase assertive complète. Le nom *fait* explicite, et fait porter l'emphase sur, une propriété de la dénotation des phrases qui est tellement fondamentale qu'elle est déjà fondée dans leur structure syntaxique. C'est ce que fait déjà de manière exemplaire l'expression *le fait est que* placée devant une phrase assertive, qui n'a pas besoin d'elle pour être comprise comme énonçant un fait. (Van de Velde 2021, [96])

L'emphase évoquée par la même auteure comme valeur de l'emploi du nom *fait*, dans les conditions rappelées, doit être reliée à la nuance oppositive que peut revêtir *le fait est que*. Cette signification est bien illustrée par l'exemple suivant :

(8) senat.fr
S'agissant de l'amiante, la branche n'a absolument pas attendu l'interdiction de 1996-1997 pour formuler un certain nombre de recommandations. Plusieurs centaines de communications ont été effectuées au sujet de l'amiante. La branche avait assez largement évoqué tous les aspects du problème. *Le fait est que* cette information, qui paraissait abondante, n'a peut-être pas été suffisamment diffusée.

À partir de la séquence précédant *le fait est que*, on peut envisager la conclusion selon laquelle la problématique est largement connue. La phrase suivante vient restreindre la portée de cette conclusion. Dans le même sens, Aijmer (2007 : 40) montre que *the fact is* indique une prise en charge, orientée vers une possible opposition ou contre-argument. Les valeurs associées à l'expression sont, selon la même auteure, rejet, élaboration ou justification, renforcement. Lenepveu (2009) rapproche les différents emplois de l'expression *le fait est que* de ceux de deux autres locutions autour du nom *fait* : avec la valeur de confirmation, *le fait est que* est proche de l'expression *de fait*, alors que dans les contextes oppositifs ou de restriction ou spécification, *le fait est que* peut être remplacé par *en fait*.

[24] Pour ce qui est de la dimension historique, l'expression est présente dans *Frantext*, à partir de Diderot et Voltaire (deuxième moitié du 18e siècle), où elle fonctionne déjà comme dans le discours contemporain :

DOI : https://doi.org/10.25364/19.2026.17.6 ISSN : 2663-9815

(9) Voltaire, *Dictionnaire philosophique*, 1764, *Frantext*
Ils le prouvent encore par des passages d'Isaïe et d'Ézéchiel ; mais les hébreux auxquels parlait Moïse ne pouvaient avoir lu ni Ézéchiel ni Isaïe, qui ne vinrent que plusieurs siècles après. Il est très inutile de disputer sur les sentiments secrets de Moïse. *Le fait est que* dans les lois publiques, il n'a jamais parlé d'une vie à venir, qu'il borne tous les châtiments et toutes les récompenses au temps présent.

Si les premières attestations datent de la deuxième moitié du 18e siècle, les statistiques fournies par *Frantext* montrent que la construction est très bien représentée dans les textes des 19e, 20e et 21e siècles.

[25] Le roumain enregistre, à côté de l'expression *faptul e(ste) că*, où le nom est accompagné de l'article défini, une variante sans déterminant, *fapt e(ste) că*. En fait, c'est la forme sans déterminant qui est la plus usuelle. De source livresque, sous cette forme, *faptul e(ste) că* se place en général en tête de phrase. L'expression est rarement introduite par un connecteur (par exemple, *dar* 'mais', *cu toate acestea* 'pourtant, malgré tout'). *Fapt e(ste) că* est paraphrasé dans *Dexonline* par « nu e mai puțin adevărat că » 'il n'est pas moins vrai que', « adevărul e că » 'la vérité est que', « în orice caz » 'en tout cas' ; dans *DLR*, *fapt e că* « ceea ce este adevărat, ce nu se poate contesta, e că… » 'ce qui est vrai, ce qui ne peut pas être contesté, c'est que…' (avec la précision qu'il ressemble à l'italien *è un fatto* ...), est enregistré sous la signification 'réalité' (avec le commentaire que la signification suit le français moderne).

[26] Dans son approche de l'emploi du défini du nom *fait* (emploi particulier puisqu'il n'y a pas de fait posé, pour qu'on s'y rapporte ensuite par *le fait*), Van de Velde (2021, note 13) rappelle l'italien *fatto sta che* ou le français *il est de fait que* (« particules de renforcement de l'assertion qu'elles introduisent »). Le roumain *fapt este că* rentre dans cette même catégorie. Vajnovszki (2025 : 177) exclut des constructions qu'elle retient celles où le nom n'a pas de déterminant, dont *fapt e că* et *dovadă e că*, en reconnaissant à ces constructions le caractère figé et une signification découlant du changement sémantique (« en voie de grammaticalisation »).

[27] Comme dans le cas du français *le fait est que*, en roumain *fapt este că* se manifeste comme une expression de la justification. Dans (10) ci-dessous, la séquence introduite par *fapt este că* ('en effet', 'ce qui est vrai, c'est que' ou 'quoi qu'il en soit') vient restreindre la portée d'une conclusion qui pourrait être tirée à partir de l'énonciation antérieure : en fait, le monde lit, mais il faut que la littérature soit de qualité.

(10) compania.ro
Pînă şi ficţiunea naţională – domeniu-far al oricărei producţii naţionale – pune probleme uriaşe. Sigur că vorba «lumea nu mai citeşte» e o prostie.

DOI : https://doi.org/10.25364/19.2026.17.6 ISSN : 2663-9815

Fapt e că e nevoie de calitate, de mult mai multă bună credinţă şi mai mult entuziasm.
'Même la production fictionnelle – domaine phare de toute production nationale – soulève de gros problèmes. Certes, la phrase « le monde ne lit plus » est une bêtise. *Le fait est qu*'on a besoin de qualité, de beaucoup plus de sincérité et plus d'enthousiasme.'

Il en est de même de *faptul este că* dans (11). L'énumération d'un nombre d'exemples d'espèces animales disparues (*doar câteva*) conduit à l'idée qu'il y en a beaucoup d'autres, ce qui est d'ailleurs explicité dans la séquence introduite par *faptul este că* :

(11) buru-news.com
Acestea sunt doar câteva exemple de animale care au dispărut. *Faptul este că* există multe specii aflate în proces de dispariţie,
'Ce ne sont que quelques exemples d'animaux disparus. *Le fait est qu*'il existe beaucoup d'espèces en voie de disparition.'

Pour continuer la comparaison avec le français, on notera que la valeur oppositive n'est pas évidente pour le roumain *faptul este că*.

[28] Pour ce qui est du type de figement, la forme avec le verbe *être* à l'imparfait est rare. Il s'agit, comme pour le français, de transposer dans le passé ce qui constitue une forme de commentaire sur la relation entre les deux séquences reliées :

(12) confluente.org
Evenimentul îl preocupase într-atât, încât toată noaptea dinainte numai într-un coşmar o petrecuse. Se făcea, între altele, că scăpase calul în vânzoleala circulaţiei din piaţa Gării de Nord, cum nimerise cu el tocmai acolo nu-şi putea explica. *Fapt era că* patrupedul o luase razna printre troleibuze, tramvaie şi taxiuri, alerga înnebunit, […]
'L'événement l'avait préoccupé à tel point qu'il avait eu des cauchemars toute la nuit de la veille. Parmi d'autres, il se voyait lâcher prise le cheval dans l'agitation de la circulation de la place de la Gare du Nord, il ne pouvait pas s'expliquer comment il en était arrivé là. *Le fait était que* le quadrupède s'était enfui parmi les trolleybus, les trams et les taxis, il courait affolé, […]'

L'interrogation de la partie roumaine du corpus de textes littéraires *ELTeC* fait voir la présence des deux expressions au 19e siècle, où elles fonctionnent de la même manière que dans l'usage récent :

DOI : https://doi.org/10.25364/19.2026.17.6 ISSN : 2663-9815

(13) Dimitrie Bolintineanu, *Doritorii nebuni*, 1864, *ELTeC*
– Ba înțeleg, răspunse Edem, însă înțeleg asemenea că Rusia nu este încă dispusă să te scoață din domnie, eşti încă tare. Ea te va lovi în ziua cînd va înțelege că te-ai făcut o stavilă la realizarea planurilor sele în Principate. Insă atunci poți compta pe partida liberală şi iată pentru ce îți zic să îi menagezi încă, o să-ți fie trebuincioşi. – Aşa, zise domnitorul, dar *faptul este că* nu sînt nici cu Rusia, nici cu partida națională!
'– Si, je comprends bien, répondit Edem, mais je comprends aussi que la Russie n'est pas encore disposée à t'éloigner du pouvoir, tu es encore fort. Elle te frappera au moment où tu seras devenu un obstacle dans la réalisation de ses projets dans les Principautés. Mais à ce moment-là tu peux compter sur le parti libéral et voilà pourquoi je te conseille de continuer à les ménager, tu en auras besoin. – D'accord, dit le prince, mais *le fait est que* je ne suis ni le partisan de la Russie, ni celui du parti national !'

(14) Dumitru Moruzi, 1850-1914, *ELTeC*
Dar ce să vă mai povestesc lucruri pe care le cunoaşteți cu toții, uneia pentru că le-ați văzut, alția pentru că le-ați auzit de la părinții sau bunicii voştri?! *Fapt este că* din ziua proclamării independenței noastre, nu au mai fost nici conservatori, nici liberali, nici singuri fii la mame văduve, ci numai români vrednici, care şi-au făcut cu toții datoria cu prisosință!
'Mais pourquoi vous raconter à nouveau des choses que vous connaissez tous, certains parce que vous les avez vus, d'autres parce que vous les avez entendus chez vos parents ou vos grands-parents ?! *Le fait est qu'*à partir du jour de la proclamation de notre indépendance, il n'y a eu ni conservateurs, ni libéraux, ni fils uniques de mères veuves, mais uniquement de braves Roumains, qui ont tous accompli leur devoir pleinement !'

3.2 *Le problème est que* / *problema e(ste) că*

[29] Selon Van de Velde (2021), par rapport à *fait*, d'autres noms comme *problème*, *scandale* ou *obstacle* pourront être considérés comme relevant de la catégorie du *fait qualifié*. *Problème* fait partie de la catégorie appréciative, selon Adler & Legallois (2018). Ainsi, le contenu introduit par l'expression *le problème est que* est présenté comme problématique.

[30] Dans l'expression *le problème est que* le sens est compositionnel dans une plus grande mesure que pour d'autres expressions, le rôle de capsule du nom et sa contribution à la focalisation du contenu de la phrase ainsi introduite étant plus évidents, comme il ressort des exemples (15) et (16) ci-dessous :

(15) 7id.fr
Des traces d'humidité apparaissent sur les murs et les moisissures envahissent le bois, les plafonds... À la longue, les pièces touchées deviennent

DOI : https://doi.org/10.25364/19.2026.17.6 ISSN : 2663-9815

invivables et sont source de maladie, particulièrement pour les enfants. *Le problème est qu*'il est quelquefois difficile de déceler l'origine d'un écoulement d'eau, comme une microfissure sur une canalisation invisible.

(16) cadtm.org
En ce moment, les créanciers du Nord subissent des contradictions internes dues au sauvetage de leurs systèmes financiers nationaux respectifs et du système financier international. Une position radicale des pays du Sud pourrait déboucher sur des solutions favorables à leurs intérêts. *Le problème est que*, comme il n'existe pas de sentiment d'urgence par rapport à la crise, les gouvernements du Sud pensent qu'ils peuvent continuer à refinancer leur dette et contracter de nouvelles dettes sans grand problème.

Le rôle de l'expression est d'introduire un fait qui semble contredire la conclusion qu'on pourrait tirer de la séquence précédente. Ainsi, l'exemple (15) pourrait être paraphrasé de la manière suivante : *s'il y a des problèmes d'humidité qui rendent les pièces invivables, alors il faut procéder à des réparations, mais ceci est difficile, vu qu'on ne peut pas identifier la source de l'écoulement* ; dans (16), la paraphrase pourrait être la suivante : *la situation financière devrait conduire les pays du Sud à prendre des mesures dans le sens de la réduction de la dette, mais ils considèrent qu'ils peuvent refinancer leur dette*.

[31] La possibilité d'intégration d'un adjectif épithète (*le gros problème*, qui tend à se figer sous cette forme), ainsi que la variation du verbe *être* (*le problème était que*) indiquent, à côté de la compositionnalité sémantique, une expression partiellement figée.

(17) menora.info
La Grande-Bretagne obtint la Palestine, y compris ce qui est aujourd'hui la Jordanie, tandis que la France obtenait le Liban et la Syrie. *Le problème était que* les Arabes palestiniens se considéraient comme Syriens et étaient considérés comme tels par d'autres Syriens.

C'est une structure qui semble de formation récente ou du moins ayant son origine dans le discours des médias. Ceci peut expliquer pourquoi, dans *Frantext*, les attestations (au nombre de 11) remontent à la fin du 20e siècle.

[32] Une variante orale signalée par Legallois (2006) est celle avec le pronom de reprise : *le problème, c'est que …* Ce type de construction a été rapproché par différents auteurs (pour ne citer que Legallois 2006 ; Vajnovszki 2025) des constructions pseudo-clivées, avec lesquelles l'expression partage le mécanisme d'organisation de l'information.

[33] Comme dans les cas de l'expression française, avec *problema este că* en roumain est introduit un autre moment dans l'argumentation, la paraphrase pouvant être *or* ou *mais*. Plus le contenu du nom est riche, plus son emploi s'avère

DOI : https://doi.org/10.25364/19.2026.17.6 ISSN : 2663-9815

nécessaire pour désambiguïser la relation entre le contexte de gauche et le contexte de droite :

(18) tvr.ro
Deputații europeni au început să dezbată la Strasbourg cum va arăta viitoarea relație dintre Uniunea Europeană şi Marea Britanie. În general, îşi doresc un acord cât mai strâns între cele două maluri ale Mânecii. *Problema este că* Boris Johnson intenționează să promoveze o separare mult mai importantă de Uniunea Europeană.
'Les députés européens ont commencé à Strasbourg le débat sur la manière dont se présentera la future relation entre l'Union Européenne et la Grande Bretagne. Généralement, ils souhaitent l'accord le plus étroit possible entre les deux rives de la Manche. *Le problème est que* Boris Johnson a l'intention de promouvoir une séparation beaucoup plus importante de l'Union Européenne.'

(19) mymed.ro
Într-un interviu pentru Adevărul, cardiologul Radu Ciudin, specialist în electrofiziologie, declara că « puține morți sunt chiar subite, mulți au avut semne înainte ». *Problema este că* mulți pacienți neglijează semnele premonitorii ale morții subite, punându-le pe seama oboselii sau vârstei.
'Dans un entretien pour Adevărul, le cardiologue Radu Ciudin, spécialiste en electrophysiologie, déclarait que « peu de morts sont réellement subites, beaucoup ont eu des signes avant ». *Le problème est* que beaucoup de patients négligent les signes avant-coureurs de la mort subite, les associant à la fatigue et à l'âge.'

3.3 *L'idée est que* / *ideea este că*

[34] L'expression française n'est pas présente dans *Frantext*, mais elle est bien représentée dans *French Web 2023*. Il en est de même pour la construction roumaine dans *Romanian Web 2021*. Un premier emploi des deux expressions est en rapport direct avec le sens des noms *idée* / *idee*. Paraphrasables par *ce qui est vrai* ou *c'est que*, *l'idée est que* / *ideea este că* introduisent une séquence qui représente une justification d'un dire antérieur, et en cela sont proches de *le fait est que* / *fapt(ul) este că* :

(20) weka.ch
Comme pour toute chose, faites vous aussi vos expériences et tirez vos propres conclusions. Si cette méthode ne vous convainc pas, ne l'appliquez pas. Le mimétisme sert à inspirer rapidement confiance et à se faire accepter par l'autre. *L'idée est que* l'on apprécie souvent les personnes qui nous ressemblent.

DOI : https://doi.org/10.25364/19.2026.17.6 ISSN : 2663-9815

(21) teodormincu.ro
Între zilele de schiat am dat o fugă în capitala principatului, Andorra la Vella. Impresia pe care mi-a făcut-o cel puțin zona comercială, dacă nu tot orașul, este cea a unui Duty Free imens. *Ideea este că* în Andorra nu există TVA și deci prețurile sunt mult mai reduse decât în țările vecine.
'Entre les jours de ski je suis allé voir la capitale de la principauté, Andorra la Vella. L'impression que m'a laissée du moins la zone commerciale, sinon la ville entière, est celle d'un Duty Free immense. *L'idée est qu*'il n'y a pas de TVA en Androrre et donc les prix sont beaucoup plus réduits que dans les pays voisins.'

(22) femeia.ro
Notele de la școală reflectă inteligență. Greșit! Cum ar putea niște teste standardizate să reflecte întregul potențial al copiilor? A primi note proaste la lucrările de control nu înseamnă că nu ești inteligent. *Ideea este că* părintele deștept o să ia în seamă notele copilului numai după o onestă și sinceră evaluarea a copilului.
'Les résultats obtenus à l'école sont un reflet de l'intelligence. Faux ! Comment imaginer que des tests standardisés puissent refléter tout le potentiel des enfants ? Avoir de mauvaises notes aux épreuves écrites ne signifie pas qu'on n'est pas intelligent. *L'idée est que* le parent intelligent ne prendra en compte les notes de l'enfant qu'après une évaluation honnête et sincère de l'enfant.'

L'effet d'explicitation du dire antérieur, adjacent dans les exemples (20), (21), (22), est plus évident dans le suivant :

(23) univ-tlse2.fr
Patricia Kuhl a mené toute une série d'études en perception qui l'ont conduite à développer le principe du prototype puis de « l'effet magnet ». *L'idée est qu*'un individu stocke en mémoire une sorte de résumé abstrait se rapportant à des éléments formant une catégorie.

La présence du subjonctif dans la séquence introduite par *l'idée est que* indique une véritable subordination, une dépendance de celle-ci par rapport à la matrice, comme dans (24) :

(24) pipcke.fr
C'est un appartement que j'achète qui doit être rénové entièrement. Peux-être prendre en considération la possibilité d'un lit escamotable. Le but est d'avoir un studio entièrement réaménagé. L'appartement sera un investissement locatif. *L'idée est qu*'il soit neutre et premium.

DOI : https://doi.org/10.25364/19.2026.17.6 ISSN : 2663-9815

C'est dans ce genre de contexte que le nom peut être caractérisé de boulique dans la terminologie d'Adler & Legallois (2018). Dans les emplois comme marqueurs discursifs que nous avons analysés, il s'agit plutôt d'une valeur aléthique.

3.4 *La vérité est que* / *adevărul este că*

[35] L'un des emplois de ces expressions est d'introduire un argument en faveur de l'idée exprimée antérieurement ou une justification du dire antérieur, comme s'il s'agissait de faire découvrir une vérité (paraphrasable par *ce qui ne saurait étonner* ou *vu que*), ce qui est en rapport avec la valeur aléthique du nom :

(25) lejdd.fr
On pourrait se féliciter que le gouvernement veuille améliorer les conditions d'accueil et d'exercice des 10 000 médecins étrangers exerçant en France avec un diplôme hors Union européenne. *La vérité est que* ces médecins sont aujourd'hui indispensables au fonctionnement de nos hôpitaux, victimes d'une politique prolongée d'austérité.

(26) unica.ro
Cei mai frumoşi pantofi Louboutin purtați de vedete. Nu e niciun secret că divele adoră pantofii Louboutin. *Adevărul este că* vorbim despre opere de artă, nu de simple încălțări.
'Les plus belles chaussures Louboutin portées par les stars. Il n'y a aucun secret que les stars adorent les chaussures Louboutin. *La vérité est qu*'il s'agit là d'œuvres d'art, non de simples chaussures.'

Dans un autre type d'emploi, l'expression vient introduire un contenu inédit, auquel on ne s'attendait pas, mais dont la dimension négative sera atténuée par l'utilisation de cette expression qui fait semblant de révéler la vérité.

(27) kiiky.com
Avec autant de styles d'apprentissage et d'objectifs d'apprentissage, vous vous demandez peut-être par où commencer. *La vérité est qu*'il n'y a pas d'approche unique pour enseigner à votre enfant comment apprendre.

(28) cosmopolitan.ro
Cu siguranță cunoşti şi tu foarte multe persoane trecute de 30 de ani care se confruntă cu probleme mai mult sau mai puțin grave, care se îndoapă cu medicamente, dar care refuză să facă sport pentru că «nu au timp». *Adevărul este că* oricare dintre noi are timp să facă şi sport, doar că cei mai mulți preferă să iasă în oraş cu prietenii la o bere, să se uite la seriale sau să piardă timpul pe Internet. Cu alte cuvinte, nu îşi fac timp.
'Tu connais sans doute toi aussi énormément de personnes de plus de 30 ans qui se confrontent à des problèmes plus ou moins graves, qui prennent plein de médicaments, mais qui refusent de faire du sport parce qu'ils « n'ont pas

DOI : https://doi.org/10.25364/19.2026.17.6 ISSN : 2663-9815

le temps ». *La vérité est que* chacun de nous a le temps de faire du sport aussi, mais la plupart préfèrent sortir en ville avec les amis prendre un verre, regarder des séries télévisées ou perdre le temps sur internet. En d'autres mots, ils ne trouvent pas le temps pour faire du sport.'

Les données recueillies dans *Frantext* montrent que la structure *la vérité est que* est attestée depuis le 14e siècle. À cette première étape, elle est précédée de marqueurs comme *mais*, *toutefois* ou *or*. En corroborant ces informations avec les conclusions de Hundt (2022) sur la constructionnalisation en anglais à partir de *the truth is that*, on pourrait considérer qu'en français aussi l'émergence de la construction retenue dans ce travail a comme point de départ *la vérité est que*. Plus tard, d'autres noms apparaîtront dans la même construction.

3.5 *Le point est que* / *punctul este că*

[36] Absente dans *Frantext*, la construction est présente dans *French Web 2023*. Les 246 occurrences semblent indiquer un usage émergent, probablement à partir d'un calque sur l'anglais. Avec cette construction il s'agit d'introduire un élément qui vient réorienter l'argumentation. La valeur y est celle d'un marqueur comme *or* :

(29) sylvainpoirier.fr
On pourrait tenter de mettre à l'épreuve ce raisonnement en considérant que l'acheteur serait intéressé de savoir que l'objet qu'il achète garde une valeur lorsqu'il ne fonctionne plus. *Le point est qu*'il ne le sait pas, et que ce n'est pas sa responsabilité.

(30) unblog.fr
Je constate que de plus en plus de citoyens sont en friction avec des concepts comme le peuple la nation ou la souveraineté. *Le point est que* ces trois éléments sont les piliers de ce qu'on appelle une démocratie. Il n'y a pas de démocratie sans peuple sans nation ou sans souveraineté.

Dans *Romanian Web 2021* les quelques occurrences de l'expression *punctul este că* apparaissent dans des contextes où, selon toute apparence, on a affaire à des calques sur l'anglais.

3.6 *La chose est que* / *lucrul este că*

[37] Le roumain *lucru* ne participe pas de ce type de construction (*Romanian Web 2021* enregistre un nombre limité d'occurrences de *lucrul este că*, selon toute apparence un calque selon l'anglais), alors que le français *chose* y est présent. Par contre, comme le montre Vajnovszki (2025), *lucru* fonctionne dans d'autres types de constructions spécificationnelles, mais toujours modifié par un adjectif épithète, un numéral, un adjectif paradigmatisant, une relative ou un supin.

DOI : https://doi.org/10.25364/19.2026.17.6 ISSN : 2663-9815

[38] Faisant partie de la même catégorie modale – aléthique – que fait (selon Adler & Legallois 2008, v. note 6) et de la même catégorie sémantique – factuel –, dans les termes de Schmid (2000), *chose* partage avec *fait* dans la construction spécificationnelle étudiée ici la valeur de confirmation (comme dans (31) et (32) ci-dessous). Il s'agit d'introduire une justification du dire antérieur, emploi dans lequel l'expression peut être paraphrasée par *car*.

(31) astucesos.com
Bicarbonate de soude et vinaigre. Cette combinaison peut être vraiment efficace pour le nettoyage, mais c'est quelque chose que vous devriez éviter à tout prix. *La chose est* que le bicarbonate de soude ou le bicarbonate de sodium réagit avec l'acide acétique, qui se trouve dans le vinaigre, et produit de l'acétate de sodium, de l'eau et du dioxyde de carbone en conséquence.

(32) orangesmile.com
Dans de nombreux hôtels au-dessus des lavabos il n'y a pas de robinets. *La chose est que* les habitants de la ville sont très attentifs à l'utilisation des ressources naturelles, y compris de l'eau potable. Ils remplissent le lavabo d'eau, et quand ils terminent de se laver, ils vidangent tout simplement l'eau sale.

3.7 *Le truc est que* / *chestia este că* / *faza este că*

[39] Dans *le truc est que*, *truc* n'est pas le synonyme dans le registre familier de *chose*. C'est la signification de 'astuce' qui est mise à profit dans cette expression. Paraphrasable par *or*, *ce qui est inédit* ou *c'est que*, *le truc est que* vient introduire un autre mouvement argumentatif, pouvant être qualifié dans les termes de Vajnovszki (2025 : 432-433) comme *marqueur de changement de topique*.

(33) asayamind.com
Allez dans n'importe quel café, vous aurez probablement le choix entre deux options différentes : une infusion froide ou un café glacé. *Le truc est que* les deux sont froids et contiennent tous les deux de la glace – alors qu'est-ce qui les différencie ? Voyons ce qui rend les deux si différents.

Caractérisant surtout l'oral, le nom *chestie* 'truc' ou 'machin' est enregistré par *Dexonline* avec la signification « fapt divers » 'fait divers', synonyme de *fapt* 'fait', *lucru* 'chose', *problemă* 'problème', *întâmplare* 'histoire'. Dans *Romanian Web 2021 chestia este că* fonctionne à côté de *toată chestia este că* :

(34) superclipuri.com
Articole cu 2 prețuri. Unele magazine japoneze afișează 2 prețuri pentru un articol. *Chestia este că* TVA sau impozitul pe vânzări sunt de obicei adăugate automat la bunuri.

DOI : https://doi.org/10.25364/19.2026.17.6 ISSN : 2663-9815

‘Certains magasins japonais affichent deux prix pour un article. *Le truc est que* la TVA ou l’impôt sur les ventes sont habituellement ajoutés automatiquement sur les biens.’

(35) link-academy.com
Einstein şi mulți alți savanți au prezis-o [călătoria în timp] cu ajutorul unor ecuații complicate, dar până acum totul a fost doar teorie. Soluția era foarte simplă şi la îndemâna oricui. *Toată chestia este că* oamenii de știință au darul de a se complica în găsirea unor soluții.
‘Einstein et beaucoup d’autres savants l’ont prédit à l’aide d’équations compliquées, mais jusqu’à présent ce n’était que théorie. La solution était très simple et à la portée de tous. *Le truc est que* les scientifiques ont tendance à compliquer les choses lorsqu’il s’agit de trouver des solutions.’

La valeur de l’expression (*toată*) *chestia este că* dans les exemples ci-dessus est liée à la signification enregistrée par *Dexonline* pour *chestie*, à savoir ‘fait divers’, pouvant être paraphrasée par *ce este interesant* ou *inedit este că* / *ce qui est intéressant* ou *inédit, c’est que* : dans (34) l’éventuel inédit de la situation de l’existence de deux prix sur les étiquettes trouve vite une explication dans la séquence introduite par le marqueur ; dans (35) la bizarrerie du fait qu’on n’avait pas encore trouvé la solution du voyage dans le temps, alors que cette solution était très simple, est mise en rapport avec l’idée que les scientifiques compliquent les choses inutilement.

[40] Un fonctionnement proche se manifeste avec l’expression familière *faza este că*, comme dans l’exemple suivant :

(36) goodreads.com
Autorul a avut un mod de-a povesti foarte plăcut, multă narațiune – ceea ce ador eu cel mai mult la o carte, în afara subiectului –, iar parcurgerea se putea face uşor. Stilul a fost unul antrenant. *Faza este că* eu mă aşteptam să găsesc cu totul altceva în carte, cel puțin coperta şi descrierea îmi transmiteau altceva, iar începutul încerca să-mi confirme bănuiala.
‘L’auteur a fait preuve d’une manière de raconter très agréable, beaucoup de narration – ce que j’adore le plus dans un livre, en dehors du sujet –, et il était facile de suivre. Le style était entraînant. *Le truc est que* je m’attendais à trouver quelque chose de complètement différent dans le livre, du moins la couverture et la présentation me transmettaient quelque chose d’autre et le début était sur le point de confirmer mon soupçon.’

4 Conclusions

[41] La brève analyse comparative que nous avons menée nous permet de tirer quelques conclusions sur les types de fonctionnement des expressions envisagées, sur leur rôle dans l’organisation du discours, sur leur rapport avec les carac-

DOI : https://doi.org/10.25364/19.2026.17.6 ISSN : 2663-9815

téristiques des noms autour desquelles elles se construisent. Du fait de leur valeur de focalisation de la séquence introduite, les expressions du type spécificationnel [N *être que* complétive] jouent un rôle dans la désambiguïsation du rapport entre contexte gauche et contexte droit, selon le type de sémantisme plus ou moins général du nom. Dans les termes de Huyghe (2018), la complétive se trouve recatégorisée selon la signification du nom qui entre dans la constitution de la structure spécificationnelle : comme *fait*, comme *chose*, comme fait problématique avec le nom *problème*, comme fait inédit avec les noms *truc* ou *chestie*.

[42] Il existe des parallélismes entre les constructions dans les deux langues (français et roumain), que nous avons comparées aussi avec l'anglais. On remarque des fonctionnements très proches notamment pour les expressions autour de noms comme *fait* / *fapt* / *fact*, *problème* / *problemă* / *problem*, *idée* / *idee* / *idea*, et *vérité* / *adevăr* / *truth*. Certaines différences ont été enregistrées entre français et roumain. Ainsi, contrairement à l'expression *le fait est que*, *fapt este că* n'apparaît pas dans des contextes oppositifs.

[43] Dans d'autres situations, il existe des divergences entre les trois langues : ainsi, l'anglais *the point is that* fait figure à part, du fait de la signification particulière de *point* dans cette construction ('ce qu'on veut dire'). Bien que la présence en français d'une expression équivalente ait été interprétée comme l'effet d'un calque sur l'anglais, l'hypothèse de l'émergence d'une construction *le point est que* qui exploite les potentialités sémantiques du nom *point* n'est pas tout à fait insoutenable car, comme c'est précisé dans le *TLFi*, dans l'un de ses emplois *point* est synonyme de *matière*, *question* ou *sujet*. Toujours pour ce qui est des différences, un nom comme *lucru* en roumain ne figure pas dans la construction [N *être que* complétive], mais se retrouve dans d'autres structures spécificationnelles.

[44] Si l'on revenait à la perspective de la grammaire de construction et plus particulièrement à une grammaire de construction diachronique, les expressions étudiées illustrent une lexicalisation de la construction, telle que décrite, par exemple, par Trousdale (2012). Hundt (2022) intègre le changement aboutissant aux marqueurs discursifs formés autour des noms capsules dans la constructionnalisation. Elle considère les structures envisagées comme semi-figées, car une partie de leur signification est déductible à partir des éléments composants. C'est effectivement le cas des expressions que nous avons retenues. Comme nous l'avons montré, on pourrait rajouter aux éléments qui en font des structures semi-figées l'alternance verbale présent / imparfait et la possibilité d'intégrer un adjectif épithète. Dans le même sens, *the thing is* en anglais est considéré par Schmid (2000) comme une collocation, alors que Aijmer (2007) considère *the fact is that* comme une formule. Pourtant, si la grammaticalisation est envisagée dans sa deuxième acception – « grammaticalization as expansion » (Traugott & Trousdale 2013 : 32) –, qui rend compte de l'émergence des marqueurs discursifs, supposant l'élargissement de la portée de l'expression, alors on peut invoquer dans le cas qui nous intéresse une grammaticalisation de la construction. Cette évolution de la proposi-

DOI : https://doi.org/10.25364/19.2026.17.6 ISSN : 2663-9815

tion principale (matrice) au marqueur discursif présente des caractéristiques comme les changements sémantiques et syntaxiques, la décatégorisation, l'évolution sémantique, qui fait passer un sens concret vers un sens plus abstrait (mise en évidence de l'importance de la phrase introduite) (Wang 2016). Contrairement à l'anglais, où ces constructions connaissent des emplois parenthétiques, sans article et/ou sans conjonction (*(the) thing is (that)*, *(the) problem is*, *(the) fact is*), en français et en roumain on ne remarque pas (encore ?) la disparition de la conjonction de subordination. Les exemples qui illustrent l'emploi parenthétique de *le fait est* sont plutôt des exceptions.

DOI : https://doi.org/10.25364/19.2026.17.6 ISSN : 2663-9815

Abréviations et références bibliographiques

Adler 2017 = Silvia Adler 2017. Les noms généraux – « *shell nouns* » – participent-ils à une lecture taxinomique de type Hiérarchie-être ? *Syntaxe & Sémantique* 18, 45-66. https://shs.cairn.info/revue-syntaxe-et-semantique-2017-1-page-45?lang=fr.

Adler & Legallois 2018 = Silvia Adler, Dominique Legallois 2018. Les noms sous-spécifiés dans le débat parlementaire : Analyse fréquentielle et catégorisation modale. *Langue française* 198, 19-34.

Aijmer 2007 = Karin Aijmer 2007. The interface between discourse and grammar: *The fact is that*. Agnes Celle, Ruth Huart (éds.). *Connectives as discourse landmarks*. Benjamins, 31-46.

Cappeau & Schnedecker 2021 = Paul Cappeau, Catherine Schnedecker 2021. Les noms d'humains généraux : contribution à la différenciation noms sous spécifiés / noms généraux. *Corela. Cognition, représentation, langage* HS-34, 1-20. https://doi.org/10.4000/corela.13517.

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

Delahunty 2012 = Gerald P. Delahunty 2012. An analysis of *The thing is that S* sentences. *Pragmatics* 22.1, 41-78. https://doi.org/10.1075/prag.22.1.02del.

Dexonline = Dexonline (éd.) 2004-. *Dexonline. Dicționare ale limbii române*. https://dexonline.ro/.

DLR = Academia Română (éd.) 2021-2022. *Dicționarul limbii române (DA / DLR). Ediție anastatică digitală*. https://clre.solirom.ro/.

ELTeC = Lexical computing 2021. *European literary text collection (ELTeC)*. https://www.sketchengine.eu/eltec-european-literary-text-collection-corpora/.

Francis 1986 = Gill Francis 1986. *Anaphoric nouns*. English Language Research.

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/.

Halliday & Hasan 1976 = Michael A. K. Halliday, Ruqaiya Hasan 1976. *Cohesion in English*. Routledge.

Hundt 2022 = Marianne Hundt 2022. *N-is* focalizers as semi-fixed constructions: Modeling variation across world Englishes. *Journal of English Linguistics* 50.2, 115-141. https://doi.org/10.1177/00754242221081241.

Huyghe 2018 = Richard Huyghe 2018. Généralité sémantique et portage propositionnel : Le cas de *fait*. *Langue française* 198, 35-49.

Huyghe 2021 = Richard Huyghe 2021. Noms généraux et noms sous-spécifiés : des relations à préciser. *Corela. Cognition, représentation, langage* HS-34, 1-19. https://doi.org/10.4000/corela.13475.

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.

Legallois 2008 = Dominique Legallois 2008. Sur quelques caractéristiques des noms sous-spécifiés. *Scolia* 23, 109-127. https://www.persee.fr/doc/scoli_1253-9708_2008_act_23_1_1118.

Legallois & Patard 2017 = Dominique Legallois, Adeline Patard 2017. Les constructions comme unités de la langue : illustrations, évaluation, critique. *Langue française* 194, 5-14.

Lenepveu 2009 = Véronique Lenepveu 2009. À propos de l'expression nominale *le fait est que*. *Le français moderne* 77.2, 199-218.

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

Schmid 2000 = Hans-Jörg Schmid 2000. *English abstract nouns as conceptual shells. From cor-*

DOI : https://doi.org/10.25364/19.2026.17.6 ISSN : 2663-9815

pus to cognition. Mouton de Gruyter.
Schmid 2018 = Hans-Jörg Schmid 2018. Shell nouns in English – a personal roundup. *Caplettra. Revista internacional de filologia* 64, 109-128. https://doi.org/10.7203/caplletra.64.11368.
Sketch engine = Lexical computing 2003-. *Sketch engine. Corpus query and management system*. https://www.sketchengine.eu/.
Stosic & Rebeyrolle 2021 = Dejan Stosic, Josette Rebeyrolle 2021. Les noms sous-spécifiés en français : du lexique au discours. Présentation. *Corela. Cognition, représentation, langage* HS-34, 1-6. https://doi.org/10.4000/corela.13467.
TLFi = Analyse et traitement informatique de la langue française (ATILF) (éd.) 2004. *Trésor de la langue française informatisé*. http://atilf.atilf.fr/tlfi.
Traugott & Trousdale 2013 = Elizabeth C. Traugott, Graeme Trousdale 2013. *Constructionalization and constructional changes*. Oxford University Press.
Trousdale 2012 = Graeme Trousdale 2012. Grammaticalization, constructions and the grammaticalization of constructions. Kristin Davidse, Tine Breban, Lieselotte Brems, Tanja Mortelmans (éds). *Grammaticalization and language change. New reflections*. Benjamins, 167-198.
Vajnovszki 2025 = Anaïs Vajnovszki 2025. *Sémantique, syntaxe et fonctions textuelles des Nss (noms sous-spécifiés) en français et en roumain*. Thèse de doctorat, Universiteit Gent, Université Sorbonne Nouvelle.
Van de Velde 2021 = Danièle Van de Velde 2021. La place du nom *fait* parmi les noms sous-spécifiés. *Corela. Cognition, représentation, langage* HS-34, 1-16. https://doi.org/10.4000/corela.13899.
Vendler 1968 = Zeno Vendler 1968. *Adjectives and nominalizations*. Mouton.
Wang 2016 = Haixia Wang 2016. *The (X) thing is*: From a matrix clause to a discourse marker. *Poznań Studies in Contemporary Linguistics* 52.3, 555-577.

DOI : https://doi.org/10.25364/19.2026.17.6 ISSN : 2663-9815