



**HAL**  
open science

## Comment dirais-je ? Que veux-tu ? Comment ça va ?

Agnès Tutin

► **To cite this version:**

Agnès Tutin. Comment dirais-je ? Que veux-tu ? Comment ça va ? : Quelques observations sur les phrases interrogatives partielles préfabriquées dans les interactions orales et les dialogues romanesques. *Lingvisticae investigationes: International Journal of Linguistics and Language*, 2022, 45 (2), pp.172-196. 10.1075/li.00072.tut . hal-04171249

**HAL Id: hal-04171249**

**<https://hal.science/hal-04171249>**

Submitted on 28 Dec 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Tutin, A. (2022). Comment dirais-je? Que veux-tu? Comment ça va? Quelques observations sur les phrases interrogatives partielles préfabriquées dans les interactions orales et les dialogues romanesques. *Linguisticae Investigationes*, 45(2), 172-196.  
Version préliminaire

## ***Comment dirais-je ? Que veux-tu ? Comment ça va ? Quelques observations sur les phrases interrogatives partielles préfabriquées dans les interactions orales et les dialogues romanesques***<sup>12</sup>

Agnès Tutin

*Univ. Grenoble Alpes, LIDILEM*

### **Abstract**

This article focuses on prefabricated sentences specific to interactions, in which we observe a significant proportion of direct partial interrogative sentences such as *comment dirais-je ? (how shall I say?)* or *que veux-tu ? (what do you expect?)*. After a presentation of the main syntactic structures and functions of these sentences, we carry out a corpus study on a sample of 10 representative sentences from the ORFEO-CEFC and ESLO2 spoken corpora and the PhraseoRom novel corpus. The results show that these utterances, diverse in their semantic and pragmatic functions, are not characterized at first sight by syntactic fixity, but present a strong variation according to register and modality (even if variation within the same corpus is also noticeable). However, some syntactic features can be attributed to their prefabricated status in the spoken corpus (over-representation of clitic inversion, under-representation of *in situ* patterns).

**Keywords:** prefabricated sentences, conversational formulae, Qu- questions, spoken corpora

### **Résumé**

Cet article porte sur les phrases préfabriquées propres aux interactions, dans lesquelles on observe une proportion importante de phrases interrogatives directes partielles comme *comment dirais-je ?* ou *que veux-tu ?* Après une présentation des principales structures syntaxiques et des principales fonctions de ces énoncés, nous examinons un échantillon de 10 phrases préfabriquées représentatives à partir des corpus oraux ORFEO-CEFC et ESLO2 et du corpus de romans PhraseoRom. Les résultats montrent que ces énoncés, divers dans leurs fonctions sémantiques et pragmatiques, ne se caractérisent pas de prime abord par le figement syntaxique, mais présentent dans l'ensemble une forte variation selon le registre et la modalité écrite ou orale (même si une variation à l'intérieur du même corpus est aussi perceptible). Cependant, certaines caractéristiques syntaxiques peuvent clairement être attribuées au statut préfabriqué dans le corpus oral (surreprésentation de l'inversion clitique, sous-représentation de la structure *in situ*).

**Mots-clés :** phrases préfabriquées, formules de la conversation, interrogatives partielles, corpus oraux

### **Introduction**

Parmi les phrases préfabriquées fréquentes dans les interactions, la structure interrogative occupe une place primordiale. On rencontre en effet de nombreuses phrases toutes faites comme les suivantes dans les échanges conversationnels : *comment allez-vous ?*, *de quoi je me mêle ?*, *mais qu'est-ce que je raconte ?* et *à qui le dis-tu ?*<sup>3</sup>.

<sup>1</sup> Cette recherche a été réalisée dans le cadre du projet « Énoncés stéréotypés des conversations ordinaires en français hexagonal et en français québécois : construction d'une ressource numérique », PAI 2022-24, financé par la Région Auvergne Rhône Alpes.

<sup>2</sup> Je remercie tout particulièrement Gaétane Dostie, Pierre Larrivée, Francis Grossmann et Nathalie Rossi-Gensane pour leurs commentaires très avisés sur cet article.

<sup>3</sup> Par convention, dans cet article, nous emploierons systématiquement le point d'interrogation pour les expressions considérées. L'usage à l'écrit est toutefois flottant, comme nous le verrons à la section 3.

La présence de ces phrases « figées » dont certaines n’attendent pas de réponse a été déjà relevée (par exemple, Coveney, 2011 ; Lefeuve & Rossi-Gensane, 2015 ; Guryev, 2017), mais à notre connaissance, ce sujet n’a pas fait l’objet d’études systématiques.

Dans cet article, nous nous intéresserons spécifiquement aux phrases interrogatives partielles préfabriquées (désormais PIPP). À travers une étude portant sur un échantillon représentatif de PIPP exploitant un corpus oral contemporain (ORFEO-CEFC et ESLO2) et un corpus d’oral représenté dans des romans français (PhraseoRom), nous cherchons à mieux observer le fonctionnement pragmatique, sémantique et syntaxique de ces interrogatives préfabriquées. Nous souhaitons en particulier mieux définir cet objet phraséologique et observer comment se manifeste la variation – particulièrement marquée pour les interrogatives directes – entre oral authentique et oral romanesque. Retrouve-t-on, comme pour les interrogatives standard une opposition très nette entre oral et écrit (ici l’oral représenté) ? Ou la nature préfabriquée de certaines expressions neutralise-t-elle en quelque sorte cette opposition classique ?

Nous définirons dans un premier temps ce que nous entendons par phrases interrogatives partielles préfabriquées, puis nous présenterons quelques observations d’ensemble sur les grands types syntaxiques et fonctionnels de ces PIPP. Nous procéderons ensuite à une étude de corpus sur un échantillon de PIPP qui portera sur la productivité des énoncés et les structures syntaxiques afin de répondre aux questions précédentes, en comparant les deux genres discursifs examinés.

## 1. Qu’appelle-t-on phrases préfabriquées interrogatives partielles (PIPP) ?

Les linguistes s’intéressant aux structures interrogatives comme Coveney (2011) ou Lefeuve & Rossi-Gensane (2015) ont observé la présence de « locutions figées » ou « séquences figées » parmi des interrogatives fonctionnant comme des questions rhétoriques ou des marqueurs discursifs :

Quelques questions rhétoriques sont devenues tellement fréquentes qu’elles peuvent être considérées comme des locutions figées :

*Où allons-nous ?*

*Comment voulez-vous que je sache ?*

*Qu’est-ce que tu veux que je fasse ?*

(Coveney, 2011 : 139)

On peut compter des cas d’interrogation partielle avec inversion, dans des séquences figées, proches des marqueurs discursifs ; ce sont de fausses questions, et non des demandes d’information, ici dans le corpus d’extraits PFC [...] :

*Enfin, il y avait la..., oh la... oh comment dirais-je, la... (PFC, Roanne).*

(Lefeuve & Rossi-Gensane, 2015, sans pagination)

Par ailleurs, un premier inventaire de phrases préfabriquées (cf. section 2.2) effectué dans notre équipe a montré qu’on pouvait estimer à 10% environ la proportion de structures interrogatives parmi ces phrases. Il s’agit donc d’un phénomène qui est loin d’être marginal et qui mérite qu’on s’y arrête. Mais quelles sont exactement les caractéristiques de ces PIPP ?

À côté des collocations (ex. : *peur bleue*), des locutions figées (ex. : *coup d’œil*) ou des « phraséotermes » (ex. : *analyse syntaxique*)<sup>4</sup>, la langue regorge d’énoncés autonomes

<sup>4</sup> Les collocations peuvent être brièvement définies comme des cooccurrences privilégiées de deux unités lexicales, généralement compositionnelles, alors que les locutions renvoient à des expressions non compositionnelles. Les phraséotermes sont des expressions polylexicales propres à une langue de spécialité.

préconstruits, stockés dans la mémoire des locuteurs et souvent répertoriés dans les dictionnaires de langue.

Les phrases préfabriquées, fréquentes à l'oral, ont déjà fait l'objet d'un intérêt marqué, de la « phraséologie exclamative » de Bally (1951), aux « conversational routines » de Coulmas (1985), Corpas Pastor (1996) ou Lüger (2007), aux « speech formulae » de Cowie (2001), sans parler des « énoncés liés » de Fónagy (1997), des clichés de Mel'čuk (par exemple, 2015) ou des « formules de la conversation » de López-Simó (2016)<sup>5</sup>. Ce champ de recherche reste toutefois encore balbutiant et les descriptions assez peu nombreuses, du fait, entre autres, d'une disponibilité toute récente de corpus oraux exploitables pour des analyses systématiques. En outre, ces phénomènes phraséologiques apparaissent particulièrement complexes à décrire, car ils requièrent des analyses relevant de multiples champs de la linguistique, allant de la syntaxe du français parlé à la pragmatique en passant par la sémantique.

Pour les définir brièvement, la classe des phrases préfabriquées présente les propriétés suivantes (cf. aussi Tutin, 2019 ; Pausé *et al.*, 2022) :

- Ce sont des énoncés complets ayant une fonction illocutoire. Certains de ces énoncés peuvent ne pas être autonomes (par exemple, les énoncés parenthétiques comme *comment dirais-je* dans *c'est un comment dirais-je une sorte de drapeau...*) ou comporter des dépendants facultatifs (*tu parles d'un imbécile*).
- Leur interprétation se fait dans le contexte énonciatif, et plus particulièrement dans une situation d'interaction. Ces phrases contiennent d'ailleurs pour nombre d'entre elles des référents actualisés dans la situation d'énonciation (ex. : *tu* comme allocutaire, dans *à qui le dis-tu ?*)
- La structure syntaxique est parfois contrainte, avec des éléments non déplaçables et des variantes morphologiques limitées, voire impossibles : *à qui le dis-tu ?* vs *\*tu le dis à qui ?* ou *\*à qui le disais-tu ?* comme on le verra dans le corpus examiné pour certaines phrases interrogatives préfabriquées.
- L'acte de langage réalisé n'est pas nécessairement conforme à ceux qui sont les plus habituels dans la structure syntaxique de la phrase préfabriquée considérée : *qu'est-ce que ça peut te faire ?* ne formule pas une demande (et n'appelle pas de réponse), mais a plutôt une fonction expressive.
- L'interprétation est souvent non littérale (avec une non-compositionnalité dans nombre de cas), comme le montre d'ailleurs le test de la traduction (fr : *Qu'est-ce que ça peut te faire ?* <-> eng : *What is it to you/what do you care?*)<sup>6</sup>
- Du fait de leur préfabrication, ces phrases toutes faites, en particulier les moins transparentes, sont souvent répertoriées dans les dictionnaires, comme le *Trésor de la Langue Française*. La figure 1 présente ainsi deux exemples de traitement pour les expressions *à quoi ça rime ?* ou *c'est pas bientôt fini ?*

s.v. RIMER

– Familier

• **À quoi ça rime?** Qu'est-ce que cela veut dire? Qu'est-ce que cela signifie? *À quoi ça rime tout ça, hein? Je te le demande?* (CÉLINE, *Voyage*, 1932, p. 561).

s.v. FINIR

– Loc., usuel, fam. **Ce (n'est pas) bientôt fini? Vous (n')avez (pas) bientôt fini?** Marquant l'impatience et la désapprobation du locuteur.

<sup>5</sup> Mentionnons également les travaux de Martins-Baltar (1995), Kauffer (2019), Dostie (2019), Krzyżanowska *et al.* (2021).

<sup>6</sup> Traductions les plus fréquentes observées dans les corpus parallèles français-anglais sur le site tradooit (<http://www.tradooit.com/>).

Fig. 1 : Exemple de traitement de phrases préfabriquées dans le *Trésor de la Langue Française*

Bien entendu, comme pour la plupart des expressions polylexicales, ces énoncés présentent de nombreuses ambiguïtés : d'une part, l'interprétation formulaire peut être en concurrence avec une interprétation littérale (ex. : *où va-t-on ?*) ; d'autre part, plusieurs interprétations peuvent co-exister pour une même phrase préfabriquée (ex. : *ça me dit rien*, examiné par Grossmann et Krzyzanowska dans ce numéro).

## 2. Typologies des phrases interrogatives partielles préfabriquées

Avant de procéder à l'étude de corpus, qui porte sur un échantillon de 10 PIPP, nous proposons d'effectuer un premier examen réalisé à partir d'une liste d'une centaine de PIPP collectées au LIDILEM<sup>7</sup> afin de repérer les structures syntaxiques et les principales fonctions pragmatiques et sémantiques de ces phrases.

### 2.1 Structures syntaxiques des interrogatives partielles

Dans cet article, nous nous limiterons aux phrases ayant une structure syntaxique correspondant classiquement aux « interrogatives directes partielles », définies dans *La Grande Grammaire du Français* comme « interrogeant sur un participant ou un élément de la situation, le mot interrogatif correspondant à l'information manquante » (Abeillé & Godard, 2021 : 1404). Le choix d'écarter dans un premier temps les interrogatives totales de notre étude est principalement motivé par des considérations pratiques : les interrogatives totales ne sont pas marquées pour la plupart morpho-syntaxiquement à l'oral, la modalité interrogative s'effectuant essentiellement par la prosodie<sup>8</sup>. Au contraire, les phrases interrogatives partielles sont facilement repérables en corpus puisqu'elles comportent toutes un mot interrogatif, qu'il s'agisse d'un pronom interrogatif (*à qui le dis-tu ?*), d'un adverbe interrogatif (*comment dirais-je ?* ou *où va-t-on ?*) ou d'un déterminant interrogatif (*quelle mouche te pique ?*).

Sur le plan formel, les phrases interrogatives partielles se présentent sous des structures diverses, propres au riche système des interrogatives en français, largement étudié pour sa diversité formelle (cf. par exemple, Coveney, 2011 ; Lefeuvre & Rossi-Gensane, 2015 ; Guryev, 2017 ; Larrivée & Guryev, 2021). Nous ne retiendrons ici que les interrogatives partielles à temps fini, mais il faut signaler la productivité des interrogatives partielles préfabriquées à l'infinitif (*que faire ? comment dire ? à qui se fier ?*) ou averbale (*à qui la faute ? pourquoi pas ?*), qui seraient à inclure dans un inventaire plus large.

Les PIPP qui nous intéressent peuvent apparaître sous des structures variées, comme le montre le tableau 1, reprenant la typologie de Coveney (2011). Comme nous le verrons dans la section 4 avec l'étude de corpus, des contraintes peuvent être observées selon l'expression et le registre.

<sup>7</sup> Cette liste de PIPP a été extraite d'un premier inventaire comportant à peu près 2500 phrases préfabriquées. Elle a été élaborée à partir de différentes ressources (lexique du LADL, *Dictionnaire électronique des mots* de Dubois et Dubois-Charlier, TlfPhraseo, dictionnaires d'expressions, observations sur corpus). Nous nous sommes en partie basée sur une ressource élaborée par D. Le Pesant (2019), que nous remercions pour la mise à disposition de ses données.

<sup>8</sup> La recherche des phrases préfabriquées interrogatives totales est rendue difficile par le fait que, d'une part la prosodie n'est souvent pas marquée dans les corpus oraux transcrits, et d'autre part, nombre de phrases préfabriquées peuvent apparaître à travers plusieurs modalités énonciatives comme le très banal *ça va*. De plus, il est parfois difficile de distinguer la prosodie interrogative de la prosodie exclamative, comme signalé par Kerbrat-Orecchioni (2001).

Structure	Phrase préfabriquée
Inversion du clitique (QV-CL)	<i>Où va-t-on<sup>9</sup> ?</i>
Antéposition (QSV)	<i>Où on va ?</i>
In situ (SVQ)	<i>On va où ?</i>
<i>Est-ce que</i> (QESV)	<i>Où est-ce qu'on va ?</i>
« que » complémenteur (QkSV)	<i>Comment que tu vas ?</i>
« c'est que » (QsekSV)	<i>Comment c'est que tu vas ?</i>
Inversion complexe (Q GN V-CL)	<i>Où le mal est-il ?</i>
Sujet (Q=S V)	<i>Qui va là ?</i>

Tableau 1 : Principales structures syntaxiques des PIPP  
(en reprenant la typologie de Coveney, 2011)

On peut ainsi identifier des structures variées pour les PPI. Nous reviendrons sur ce point en 4.2.

## 2.2 Types fonctionnels des phrases préfabriquées interrogatives

Comme cela a été signalé plus haut, il a souvent été remarqué que les phrases interrogatives « figées » – partielles ou non – ne correspondent pas à des demandes d'information, qu'elles se rapprochent de phrases expressives ou de marqueurs de discours à fonction métaénonciative. Si cela est probablement le cas, cette tendance (qui reste à confirmer sur une étude de corpus plus large) n'a cependant rien de systématique.

Certaines phrases interrogatives préfabriquées ont en effet bien une fonction « questionnante ». On pourra ainsi relever dans ce groupe des phrases que l'on peut associer à la notion de « pragmatème » (Blanco & Mejri, 2018) ou de « phrase sociale » (Dostie, 2019), comme les suivantes : *à qui ai-je l'honneur ?* et *c'est de la part de qui ?* pour lesquelles une réponse ou réplique est en principe attendue, même si certaines phrases rituelles comme les salutations complémentaires (ex. : *ça va ?*) ont une évidente fonction phatique et ne requièrent pas de réponse déclarative, mais éventuellement un retour rituel du type – *ça va ? – ça va ?* (Kerbrat-Orecchioni, 2016 : 118-120).

On peut également mentionner des PIPP à fonction métaénonciative, hétéroadressées, comme celles qui adressent à l'allocutaire une demande de précision/reformulation comme *que veux-tu dire ?* ou *où voulez-vous en venir ?*.

Enfin, on relève des PIPP qui correspondent à des phrases situationnelles (par exemple, Anscombe, 2000), c'est-à-dire à des phrases non expressives, interprétables en contexte, mais qui ne sont pas exclusives d'une situation d'interaction : *qu'est-ce qui se passe ? qu'est-ce qu'il y a de neuf ?*.

Cependant, nombre de PIPP n'ont pas une fonction de demande. Si une réaction chez l'allocutaire peut parfois suivre la phrase interrogative, elle ne portera pas sur les éléments correspondant au segment interrogatif (ex. : *à qui* ou *qu'est-ce que*), mais sera en lien avec le sens global de l'énoncé interrogatif, comme on peut l'observer dans (1).

<sup>9</sup> Ici au sens phraséologique, bien entendu.

- (1) Le gardien de prison s'assied sur une roche en disant :
- Qu'est-ce que tu me chantes là ?
  - Tu comprendrais si tu passais tes journées dans le centre de la ville. Tu es dans la prison. Loin de la crasse. Loin du vice. De la puanteur, de la corruption... Loin de l'avarice.
- (Corpus PhraseoRom, Bernard Clavel, *La révolte à deux sous*, 1992)

On évoque souvent ici pour cet emploi la figure de style de la « question rhétorique » (Borillo, 1981) ou « interrogation oratoire » (Riegel, Pellat & Rioul, 2018) mais cette notion ne permet pas, nous semble-t-il, d'embrasser la totalité des cas qui nous intéressent<sup>10</sup>.

Un premier groupe de phrases a une fonction expressive marquée, qui peut être centrée sur le locuteur et exprime des sentiments comme le désarroi (*où va-t-on ?*) ou l'impuissance (*que voulez-vous ?*). Plus nombreuses semblent être les PIPP expressives hétéroadressées, qui expriment un sentiment négatif ou une orientation polémique (*dans quel monde vis-tu ?*, *qu'est-ce que ça peut te faire ?*, *de quoi je me mêle ?*, *qu'est-ce tu me chantes ?*).

Plusieurs types de phrases à fonction métaénonciative n'appellent pas de réponse non plus. Cela inclut les PIPP autoadressées, qui visent à corriger la formulation (*que dis-je ?*), rechercher l'expression adéquate (*comment dirais-je ?*) ou retrouver le fil de conversation (*que voulais-je dire ?*). On relève également dans ce groupe une PIPP hétéroadressée, qui a souvent une fonction d'amorce discursive (*tu sais quoi ?*).

Enfin, on observe un petit groupe de PIPP à fonction plus directement interactionnelle, exprimant avant tout l'accord (*à qui le dis-tu ?*) ou le désaccord teinté d'émotion négative (*qu'est-ce vous croyez ?*).

Les PIPP se révèlent donc extrêmement diverses dans leur fonctionnement pragmatique et on ne peut pas les réduire à un ou deux stéréotypes. L'étude de corpus permettra d'approfondir la compréhension de ces objets linguistiques.

### 3. Méthode d'analyse de corpus : une comparaison oral authentique vs oral représenté dans les romans

Pour mieux comprendre le fonctionnement des PIPP, nous avons procédé à un sondage en corpus à partir d'un échantillon de quelques PIPP représentatives.

L'étude de corpus que nous avons effectuée est avant tout qualitative. Elle exploite deux corpus :

- un corpus d'oral authentique, représenté par la partie orale du corpus ORFEO-CEFC<sup>11</sup> (Benzitoun & Debaisieux, 2020) et le corpus ESLO2<sup>12</sup> (Abouda & Baude, 2006) totalisant à eux deux 5,4 millions de tokens<sup>13</sup>. Le corpus contient des conversations informelles, des entretiens, des réunions diverses et des discours publics et il correspond à des interactions récentes (postérieures aux années 2000).
- un corpus de romans contemporains en français (1950-2020), développé dans le cadre du projet PhrasoRom (Novakova & Siepmann, 2019), de 171 millions de tokens, comprenant de nombreux dialogues.

Les PIPP examinées se limitent à 10 phrases : ce choix s'explique par le fait qu'il fallait pouvoir disposer d'un nombre suffisant d'attestations afin d'examiner les différentes variantes.

<sup>10</sup> La question rhétorique est généralement définie comme une question qui correspond en fait à une assertion indiquant le contraire du contenu propositionnel. Si cette définition fonctionne assez bien pour des PIPP comme *où va-t-on ?* (que l'on peut grossièrement paraphraser par « je ne sais pas où l'on va »), elle est plus difficilement applicable pour des énoncés comme *que voulez-vous/que veux-tu ?* ou *à qui le dis-tu ?*

<sup>11</sup> <https://orfeo.ortolang.fr>

<sup>12</sup> <http://eslo.huma-num.fr/>

<sup>13</sup> Un token est un mot, un chiffre ou un signe de ponctuation.

Les corpus oraux examinés restent de taille modeste et comportent par ailleurs peu de PIPP familières et polémiques, propres à la conversation intime. Paradoxalement, en effet les expressions conflictuelles comme *qu'est-ce que ça peut te foutre ?* ou *à quoi tu joues ?* sont moins représentées dans ces corpus oraux que dans les écrits romanesques contemporains qui mettent davantage en scène la conflictualité du quotidien. En outre, le petit nombre de PIPP examinées est aussi lié à l'important temps de traitement manuel des PIPP dans les corpus, assez fastidieux, du fait de l'ambiguïté des PIPP et de l'absence de ponctuation dans les transcriptions des corpus oraux.

Les PIPP retenues sont des énoncés relativement fréquents, qui apparaissent dans les deux corpus. Elles sont de plusieurs sortes, selon la typologie fonctionnelle présentée à la section précédente (voir Tableau 2). Les variantes syntaxiques seront présentées à la section 4.

PIPP	Question	Type
<i>Qu'est-ce ça peut faire<sup>14</sup> ?</i>	Non	Fonction expressive (centrée sur le locuteur)
<i>Que veux-tu/voulez-vous ?</i>	Non	Fonction expressive (centrée sur le locuteur)
<i>Comment dirais-je ?</i>	Non	Fonction métaénonciative (auto-adressée)
<i>Que dis-je ?</i>	Non	Fonction métaénonciative (auto-adressée)
<i>Que voulais-je dire ?</i>	Non	Fonction métaénonciative (auto-adressée)
<i>Tu sais/vous savez quoi ?</i>	Non	Fonction métaénonciative (hétéro-adressée)
<i>Comment ça se fait ?</i>	Oui	Phrase situationnelle
<i>Que se passe-t-il ?</i>	Oui	Phrase situationnelle
<i>Comment allez-vous /vas-tu ?</i>	Oui	Pragmatème (phrase rituelle de salutation)
<i>Comment ça va ?</i>	Oui	Pragmatème (phrase rituelle de salutation)

Tableau 2 : PIPP sélectionnées et types fonctionnels

L'outil utilisé pour interroger le corpus est *le Lexicoscope*<sup>15</sup> (Kraif, 2019), qui intègre, outre le corpus PhraseoRom, les deux corpus oraux (Pausé *et al.*, 2022), avec l'alignement texte-son et une analyse syntaxique en dépendances. Ce système permet d'effectuer des recherches ciblées, intégrant des paramètres syntaxiques et indiquant la place dans le tour de parole. Le logiciel permet aussi de repérer les spécificités de chaque sous-corpus. La figure 2 présente ainsi la surreprésentation nette de l'énoncé *comment dirais-je ?* dans les sous-corpus d'entretiens dont le CFPP (*Corpus de Français Parlé Parisien*, Branca-Rosoff *et al.*, 2012), alors que l'expression *tu sais quoi/vous savez quoi ?* n'y apparaît pas une seule fois, ce qui montre bien aussi la nécessité de prendre en compte les genres spécifiques de l'oral.

<sup>14</sup> Nous n'avons pas inclus ici les occurrences de *qu'est-ce que ça peut te/vous faire* dont le fonctionnement interactionnel est assez différent, puisqu'il est nettement hétéroadressé.

<sup>15</sup> [http://phraseotext.univ-grenoble-alpes.fr/lexicoscope\\_2.0](http://phraseotext.univ-grenoble-alpes.fr/lexicoscope_2.0)

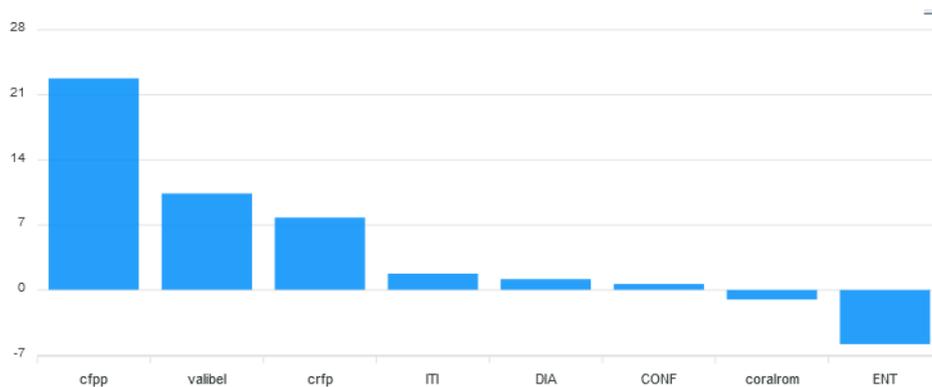


Figure 2 : Spécificités des PIPP *comment dirais-je* dans les corpus ORFEO-CEFC et ESLO2

Le repérage des énoncés génère un bruit important et une désambiguïisation manuelle est nécessaire. Elle apparaît particulièrement complexe dans les corpus oraux du fait de l'absence de ponctuation dans les transcriptions et des problèmes de segmentation rencontrés. Dans les deux types de corpus, il faut en outre distinguer l'emploi préfabriqué des emplois standard qui co-existent, comme dans le cas de *que voulez-vous/veux-tu* qui correspond la plupart du temps à une simple demande (2) plutôt qu'à une phrase expressive (3), comme l'illustrent les exemples suivants :

- (2) Un homme d'une soixantaine d'années au visage marqué apparut à la fenêtre du rez-de-chaussée.  
 – Oui, que voulez-vous ?  
 – Je suis l'inspecteur Cadin, de Toulouse. Voici le brigadier Lardenne, mon adjoint. Je désire vous parler, en privé.  
 (Corpus PhraseoRom, D. Daeninckx, *Meurtres pour mémoire*, 1964).
- (3) [ilcBM1] nos grands-parents comme nous / un jour sont amenés disparaître hein  
 [ilcBM1] que voulez -vous oui  
 [ilcMF0] mm  
 (Corpus ORFEO-CEFC, Valibel)

Enfin, il faut mentionner dans l'étude de corpus quelques spécificités liées à la ponctuation dans le corpus Phraseorom. Les requêtes que nous avons effectuées sur ce corpus ne se limitent pas à des expressions suivies d'un point d'interrogation. En effet, on observe pour certaines PIPP comme *que voulez-vous/que veux-tu ? (qu'est-ce que tu veux/vous voulez. ?)* et *qu'est-ce que ça peut faire ?* une ponctuation variable. À côté du point d'interrogation, qui reste le plus fréquent, on relève aussi des points d'exclamation, comme dans (4), voire une absence complète de marque comme dans (5)<sup>16</sup>.

- (4) Je suis serviable, que veux-tu ! Quand je vais de Tabelbala à Béni Abbès bien sûr. Parce que quand je reviens de Béni Abbès à Tabelbala, alors là, jamais personne à prendre sur la piste. Tabelbala, ça se vide, ça ne se remplit pas !  
 (Corpus PhraseoRom, Michel Tournier, *La goutte d'or*, 1985).
- (5) – Mais, Mme Meunier, dans cet article il n'y a rien d'autre qui vous choque ?  
 – Quoi donc, qu'il parle encore de Commodore ? Qu'est-ce que vous voulez, ça fait vendre cette histoire...

<sup>16</sup> Comme l'a fait remarquer N. Rossi-Gensane (que je remercie), l'absence de ponctuation forte est probablement liée en (5) à la position initiale de la PIPP.

(Corpus PhraseoRom, Serge Joncour, *L'écrivain national*, 2014)

Cette hésitation des scripteurs est intéressante en ce qu'elle indique le statut ambivalent de ces expressions. Sur le plan syntaxique, il s'agit bien de phrases interrogatives directes partielles, avec un mot interrogatif (*qu'est-ce que* ou *que*) ayant un statut de complément d'objet direct<sup>17</sup>. Sur le plan pragmatique cependant, la phrase n'est pas véritablement une demande, comme nous l'avons souligné plus haut, ce qui fait dire d'ailleurs à Bally (1951)<sup>18</sup> :

Il est [...] préférable de supprimer cette distinction usuelle entre ton exclamatif et ton interrogatif qui en réalité ne sont séparés par rien d'essentiel : disons que l'interrogation qui n'interroge pas appartient au langage dit exclamatif, quitte à donner à ce terme un sens tout à fait général. (Bally, 1951 : 59)

Toutefois, face à la difficulté de définir ce qu'est la modalité exclamative en s'appuyant sur des critères sémantiques et pragmatiques, nous préférons ici nous baser sur des critères formels pour délimiter les PIPP, même si cette classe intègre ce qu'Abeillé et Godard (2021 : 1430) appellent des « interrogatives exclamantes ».

#### 4. Analyse des PIPP dans le corpus oral et le corpus PhraseoRom

Pour les phrases sélectionnées, nous avons repéré la productivité des PIPP sous leurs différentes variantes syntaxiques dans les deux corpus examinés, telles que schématisées par Coveney (2011) et présentées dans le tableau 1.

Pour des commodités de lecture, nous avons reproduit dans le Tableau 3 la forme des différentes variantes. La productivité des structures a été observée en indiquant la fréquence relative (pour 1 million de tokens) et, entre parenthèses, la fréquence absolue<sup>19</sup>. Bien entendu, certaines structures ne sont syntaxiquement pas possibles (par exemple, la forme à antéposition avec le pronom interrogatif faible *que* : *\*que tu veux ?*) et, de ce fait, n'ont pas été examinées (elles correspondent aux cases vides). En outre, les variantes avec « complémenteur » (*comment que tu vas ?*) et avec variante en *c'est* (*comment c'est que tu vas ?*) sont très marginales dans les deux corpus et n'ont pas été retenues ici<sup>20</sup>. Enfin, certaines constructions contraignent à des variations morphologiques mineures. Par exemple, l'inversion complexe (*comment ça va ?* vs *comment cela va-t-il ?*) oblige à une transformation du *ça* en *cela*, pour rester conforme au registre formel. Les résultats sont consignés dans le tableau 3 et sont commentés plus en détail dans les sections qui suivent.

<sup>17</sup> Contrairement aux *qu'est-ce que* examinés par Dekhissi & Coveney (2021), par exemple dans l'extrait cité par les auteurs *Qu'est-ce qu'on vient me poser ces questions-là ?! C'est pas les questions qui préoccupent les Français*. (Sarkozy, *TF1*, 2012). On pourrait d'ailleurs aussi analyser ces phrases comme des exclamatives.

<sup>18</sup> Cité par Borillo (1991).

<sup>19</sup> Du fait du très grand nombre de réponses pour le corpus PhraseoRom, la fréquence a parfois été estimée à partir d'un échantillon des 50 premières occurrences.

<sup>20</sup> Cela pose d'ailleurs des questions sur la nature des corpus oraux examinés qui comportent peu d'oral très familier.

Inversion du clitique	Antéposition QSV	In Situ SVQ	Est-ce que	Inversion complexe	Fréquence totale (toutes structures)
		<i>Ca peut faire quoi ?</i> Or. : 0, Ec. : 0	<i>Qu'est-ce que ça peut faire ?</i> Or. : 5 Ec. : 195	<i>Que cela peut-il faire ?</i> Or. : 0 Ec. : 0	Or. : 0,92 (5) Ec. : 1,14 (195)
<i>Que veux-tu/voulez-vous ?</i> Or. : 0,55 (3), Ec. : 0,73 (125)		<i>Vous voulez/tu veux quoi ?</i> Or. : 0, Ec. : 0	<i>Qu'est-ce que tu veux/vous voulez ?</i> Or. : 2,96 (16), Ec. : 0,23 (40)		Or. : 3,51 (19) Ec. : 0,96 (165)
<i>Comment dirais-je ?</i> Or. : 9,6 (52) Ec. : 0,26 (45)	<i>Comment je dirais ?</i> Or. : 5,4 (9) Ec. : 0	<i>Je dirais comment ?</i> Or. : 0,18 (1) Ec. : 0	<i>Comment est-ce que je dirais ?</i> Or. : 0 Ec. : 0		Or. : 11,4 (62) Ec. : 0,26 (45)
<i>Que dis-je ?</i> Or. : 0,36 (2), Ec. : 1,19 (204)		<i>Je dis quoi ?</i> Or. : 0,36 (2), Ec. : 0	<i>Qu'est-ce que je dis ?</i> Or. : 1,3 (7) Ec. : 0,04 (7)		Or. : 2,03 (11) Ec. : 1,23 (211)
<i>Que voulais-je dire ?</i> Or. : 0 Ec. : 0 (1)		<i>Je voulais dire quoi ?</i> Or. : 0 Ec. : 0	<i>Qu'est-ce que je voulais dire ?</i> Or. : 3,15 (17) Ec. : 0,02 (4)		Or. : 3,14 (17) Ec. : 0,02 (5)
<i>Que sais-tu/savez-vous ?</i> Or. : 0 Ec. : 0		<i>Tu sais/vous savez quoi ?</i> Or. : 3,15 (17) Ec. : 2,76 (473)	<i>Qu'est-ce que tu sais ?</i> Or. : 0 Ec. : 0		Or. : 3,14 (17) Ec. : 2,76 (473)
	<i>Comment ça se fait ?</i> Or. : 10,2 (55) Ec. : 0,45 (77)	<i>Ca se fait comment ?</i> Or. : 0, Ec. : 0	<i>Comment est-ce que ça se fait ?</i> Or. : 0, Ec. : 0	<i>Comment cela se fait-il ?</i> Or. : 0 Ec. : 0,19 (33)	Or. : 10,1 (55) Ec. : 0,64 (110)
<i>Que se passe-t-il ?</i> Or. : 0,72 (4) Ec. : 12 (2054)		<i>Il se passe quoi ?</i> Or. : 0,36 (2) Ec. : 0,1 (18)	<i>Qu'est-ce qui se passe ?</i> Or. : 0,92 (5) Ec. : 10,1 (1732)		Or. : 2,03 (11) Ec. : 22,4 (3804)
	<i>Comment ça va ?</i> Or. : 1,48 (8), Ec. : 1,81 (310)	<i>Ça va comment ?</i> Or. : 0,18 (1), Ec. : 0,03 (5)	<i>Comment est-ce que ça va ?</i> Or. : 0,36 (2), Ec. : 0	<i>Comment cela/ça va-t-il ?</i> Or. : 0, Ec. : 0,08 (15)	Or. : 2,03 (11) Ec. : 1,92 (330)
<i>Comment allez-vous /vas-tu ?</i> Or. : 0,18 (1), Ec. : 3,8 (649)	<i>Comment tu vas/vous allez ?</i> Or. : 0,36 (2), Ec. : 0,33 (57)	<i>Vous allez/ tu vas comment ?</i> Or. : 0, Ec. : 0,01 (2)	<i>Comment est-ce que tu vas/vous allez ?</i> Or. : 0, Ec. : 0		Or. : 0,55 (3) Ec. : 4,14 (708)

Tableau 3 : Corpus des PIPP examinées avec les variantes syntaxiques et les fréquences (par million de tokens).

[Les variantes les plus productives dans chaque corpus apparaissent dans une case grisée]

#### 4.1 Productivité des PIPP dans les deux corpus

Les PIPP qui nous intéressent apparaissent à la fois dans le corpus oral et dans le corpus romanesque PhraseoRom, et plus spécifiquement pour ce dernier, dans les passages dialogués, qui constituent une proportion non négligeable du texte (un peu moins d'un tiers du texte d'après une première estimation<sup>21</sup>). Il peut alors être intéressant de comparer l'oral authentique et l'oral représenté, évidemment stylisé dans le roman, qui ne cherche d'ailleurs pas nécessairement à reproduire le style oral. Dans une étude linguistique sur le dialogue dans le roman, Durrer (2005) a ainsi relevé que certains éléments propres aux interactions orales comme les éléments phatiques (*bonjour*), les régulateurs (*ah, hum, d'accord*) étaient de manière générale sous-représentés dans les dialogues romanesques. Dans une étude antérieure (Tutin & Gharbi, 2020), nous avons de notre côté relevé une surreprésentation dans les romans contemporains par rapport au corpus ORFEO-CEFC pour des formules expressives familières comme *c'est plus fort que moi* ou *bon sang* par rapport à des formules plus quotidiennes comme *c'est pas grave* ou *c'est dommage* nettement moins représentées, probablement pour renforcer l'effet d'expressivité dans le roman.

Suite à ces observations, nous faisons l'hypothèse que les expressions de type métaénonciatif, surtout celles qui sont liées à la régulation du discours (auto-correction, reformulation), et qui « parasitent » en quelque sorte le dialogue, seront sous-représentées dans le corpus romanesque, de façon à permettre une meilleure fluidité. Il serait aussi intéressant d'observer comment se comportent les formules expressives dans les deux corpus. Pour ce faire et pour présenter des résultats plus lisibles que dans le tableau 3, nous avons observé la productivité des expressions dans un histogramme en prenant en compte la fréquence relative (cf. figure 3).

Toutes les variantes syntaxiques sont ici fusionnées, de façon à mettre l'accent sur les fonctions des PIPP. Par exemple, *que se passe-t-il, il se passe quoi* et *qu'est-ce qui se passe ?* sont comptabilisés comme une seule expression.

---

<sup>21</sup> Une annotation des passages dialogués dans les romans est en cours au LIDILEM (mais n'a pas encore été réalisée à ce stade). Il n'est donc pas possible d'indiquer précisément la proportion de texte consacrée aux dialogues dans le corpus PhraseoRom. Néanmoins, une annotation systématique sur un échantillon de 11 romans récents (Tutin & Gharbi, 2020) a montré que 29,2% des mots du texte apparaissaient dans les passages dialogués.

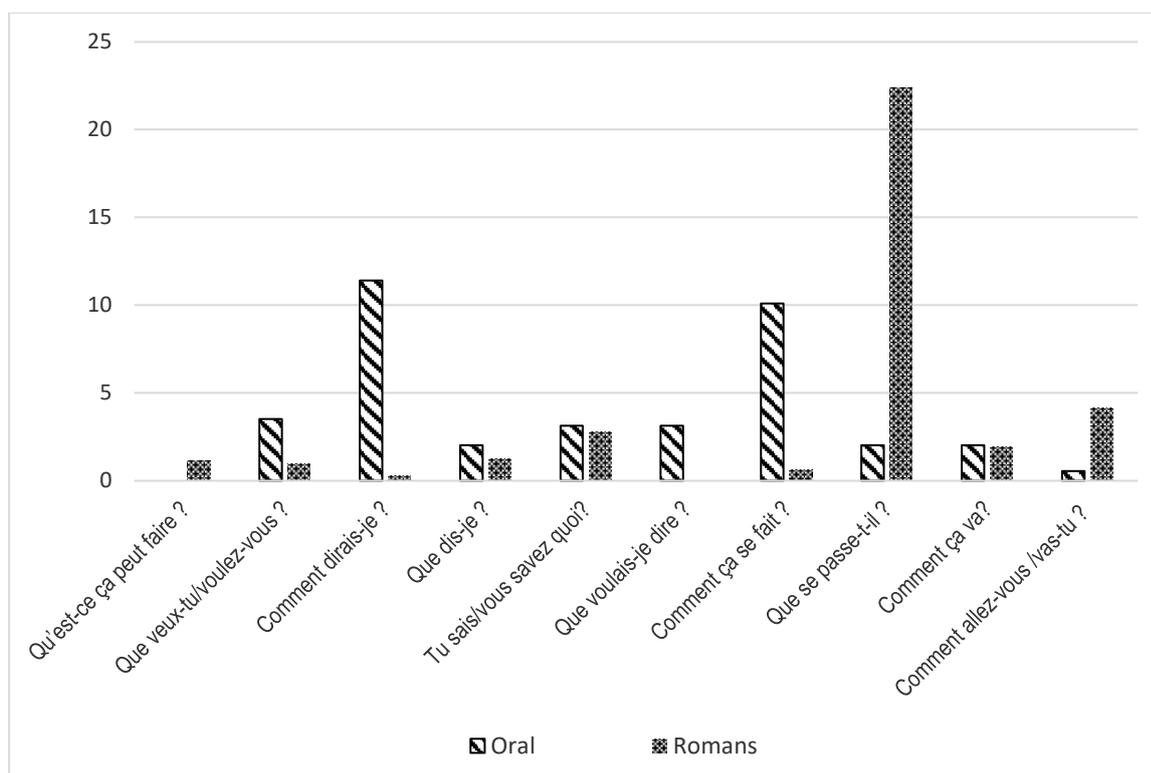


Figure 3 : Distribution des PIPP (toutes variantes syntaxiques) selon les 2 corpus (en fréquence relative, pour 1 million de tokens)

Comme attendu, les résultats apparaissent assez contrastés entre les deux corpus. De manière générale, les PIPP sont assez présentes dans les romans, ce qui est tout à fait remarquable, car les passages dialogués où elles sont les plus courantes ne représentent, comme on l'a vu, qu'une partie du texte<sup>22</sup>. Les romanciers semblent donc bien mettre en scène l'oralité à travers des phrases préfabriquées fréquentes dans les interactions orales.

Conformément à notre hypothèse, les PIPP à fonction métaénonciative sont nettement sous-représentées dans le corpus romanesque, en particulier celles qui gèrent le flux du discours (*comment dirais-je ? que voulais-je dire ?*). On relève toutefois que l'expression *tu sais quoi*, très orale dans sa forme avec le pronom interrogatif *in situ*, est assez fréquente dans les dialogues romanesques, comme dans (6) (nous y reviendrons à la section suivante) :

- (6) - Pas de message, dit Marthe.  
 - Pas grave, dit Marc, à nouveau résolu à garder le silence.  
 - Mais tu sais quoi ? reprit-il. Louis veut m'enrôler comme homme qui court.  
 (Corpus PhraseoRom, Fred Vargas, *Un peu plus loin sur la droite*, 1996).

On ne dégage pas de tendance claire pour les deux PIPP expressives. Pour les formules de salutation, on observe une nette surreprésentation de *comment allez-vous/vas-tu ?* dans les romans, peut-être liée à la multiplicité des interactions dans ce genre textuel, plusieurs scènes avec différents personnages pouvant se succéder. Une autre explication peut être le caractère un peu plus formel de *comment allez-vous/vas-tu* par rapport aux PIPP plus neutres *comment ça va ?* et *ça va ?*. Enfin, la fréquence élevée de *que se passe-t-il ?* dans le corpus romanesque semble lié bien entendu aux aspects narratifs et à la présence d'événements successifs à relater, moins fréquents dans les interactions orales.

<sup>22</sup> Il serait intéressant dans une recherche ultérieure de comparer plus précisément la productivité entre oral authentique et passages dialogués.

En bref, la différence de distribution entre les expressions (toutes variantes syntaxiques confondues) dans les deux sous-corpus semble liée à différents facteurs, comme l'importance informationnelle de la PIPP (faible dans les expressions autoadressées comme *comment dirais-je ?*) ou la dynamique narrative. L'examen réalisé ici reste toutefois limité à quelques expressions et rend les généralisations difficiles.

## 4.2 Structures syntaxiques

On sait que la modalité écrite vs orale et le registre influent fortement sur les structures des interrogatives directes, l'inversion clitique (ex. : *où vas-tu ?*) étant moins présente dans l'oral informel qui préfère de manière générale les structures *in situ* (ex. : *tu vas où ?*) (Coveney 2011 ; Lefeuvre & Rossi-Gensane 2015 ; Guryev 2017). Toutefois, il a été observé que les interrogatives figées avaient parfois un comportement syntaxique atypique, avec des caractéristiques propres à un état de langue plus ancien<sup>23</sup>, qui privilégiait des emplois comme l'inversion clitique. De plus, certaines PIPP comme *à qui le dis-tu ?* présentent une distribution syntaxique lacunaire<sup>24</sup>, n'acceptant qu'une seule structure. Cette pluralité de paramètres rend particulièrement intéressante – mais aussi complexe – l'étude des variations syntaxiques pour les expressions qui nous concernent.

### 4.2.1 De rares structures syntaxiques figées

Tout d'abord, on observe que la grande majorité des PIPP présentent des variations syntaxiques : seules 2 expressions sur 10 (*tu sais/vous savez quoi ?* et *qu'est-ce que ça peut faire ?*) apparaissent sans aucune variante, ce qui n'est nullement lié à la structure syntaxique même des énoncés<sup>25</sup>, mais à ces expressions spécifiques. On ne peut donc pas affirmer que la syntaxe des PIPP est de manière générale figée.

La PIPP *tu sais quoi ?*, correspondant d'après Quillard (2001) à une « question introductive stéréotypée » ou à une « préannonce » (Coveney 2011)<sup>26</sup>, a une fonction d'anticipation par rapport à la suite du propos, servant « à des fins de projection d'une action subséquente » (Guryev 2021 : 98), comme on l'observe dans l'exemple (7). Cette fonction s'observe aussi dans les PPI interrogatives totales *tu connais la meilleure ?* ou *tu sais la nouvelle ?*.

- (7) [L1] hum et ton kyste là que tu avais maintenant c'est bon / tu as plus mal  
 [L2] ouais non c'est bon / ça fait euh ça fait déjà un bout de temps c'est pas revenu /  
 normalement tous les s~ normalement les s~ les pr~ les  
 [L1] que tu marchais comme un cowboy  
 [L2] c'est tous les deux mois / ah ben au début c'était  
 [L2], mais tu sais quoi moi j'ai un pote qui a eu le même truc que moi tu vois  
 (Corpus ORFEO-CEFC, Tcof, Masc\_dom\_sd)

<sup>23</sup> Comme l'absence de déterminant en position objet (*faire flèche de tout bois*).

<sup>24</sup> On pourra ainsi comparer *à qui le dis-tu ?* et *à qui le donnes-tu ?* :

clitique : *à qui le dis-tu ?* vs *à qui le donnes-tu ?* ; antéposition : *\*à qui tu le dis ?* vs *à qui tu le donnes ?* ; in situ : *\*tu le dis à qui ?* vs *tu le donnes à qui ?* ; est-ce que : *\*à qui est-ce que tu le dis ?* vs *à qui est-ce que tu le donnes ?* De la même façon, on peut comparer *qu'est-ce que ça peut faire* à *qu'est-ce que ça peut produire ?*. La seconde phrase accepte tout à fait la structure *in situ* : *ça peut produire quoi ?*.

<sup>25</sup> Par exemple *tu aimes quoi ?* alterne bien avec *qu'aimes-tu ?* *qu'est-ce tu aimes ?* et *qu'est-ce ça peut donner ?* avec *ça peut donner quoi ?*.

<sup>26</sup> Coveney (2011) considère d'ailleurs qu'il s'agit d'interrogatives totales. Cette analyse peut être confirmée par la quasi-synonymie avec les interrogatives totales préfabriquées *tu connais la meilleure ?* ou *tu sais la nouvelle ?*.

Sur le plan communicatif, l'accent est ici mis sur *quoi* et, de ce fait, la forme faible *que* apparaît impossible. L'hypothèse de Coveney (1997) (citée par Lefevre & Rossi-Gensane 2015) selon laquelle la structure *in situ* donnerait au pronom une forte valeur rhématique semble ici validée. Nous pensons aussi que le fonctionnement quasi cataphorique du *quoi* post-verbal peut expliquer sa position. En outre, il est intéressant de signaler que l'expression est presque aussi fréquente dans l'oral romanesque (cf. ex. 6) que l'oral authentique, ce qui peut être lié à la dynamique de suspense qu'elle crée dans la narration.

En ce qui concerne l'autre expression, *qu'est-ce que ça peut faire ?*, l'explication est en partie syntaxique. On relève un certain nombre de phrases en *qu'est-ce que* qui ont de façon privilégiée une fonction expressive et polémique (*qu'est-ce tu me chantes ?*, *qu'est-ce que tu veux ?*), mais ces phrases peuvent aussi apparaître dans la structure avec clitique (*que me chantes-tu ?*, *que veux-tu ?*). La limitation de structure pour *qu'est-ce que ça peut faire ?* nous semble principalement liée à la présence du pronom *ça*, qui ne permet pas l'inversion clitique.

#### 4.2.2 Une importante variation entre corpus oral et corpus romanesque, mais également au sein de chaque corpus

La plupart des PIPP analysées acceptent une variation syntaxique, entre les deux corpus bien sûr, mais aussi à l'intérieur du même corpus, ce qui montre bien que l'on a affaire à des tendances plutôt qu'à des distributions tranchées et complémentaires. Rappelons aussi que le corpus oral est assez hétérogène, ce qui peut expliquer cette variation (voir distribution de *comment dirais-je ?* dans l'oral authentique figure 2).

L'examen de la structure SVQ dans le tableau montre qu'en dehors de *tu sais quoi ?*, seule structure possible pour cet énoncé, les positions *in situ* dans le corpus oral sont loin d'être les structures majoritaires, alors que d'après les observations de différents linguistes compilées par Guryev (2017 : 117), ce seraient les variantes les plus courantes pour l'oral informel (même si elles sont un peu moins fréquentes avec les interrogatives partielles qu'avec les interrogatives totales). Les PIPP semblent donc se différencier ici des structures standard. Notons toutefois qu'avec le mot interrogatif *comment* (présent dans 4 PIPP sur 10) les structures *in situ* semblent de manière générale moins fréquentes que les structures QSV (Lefevre & Rossi-Gensane 2015 ; Guryev 2017), ce que nous observons également dans notre corpus où les structures QSV (*comment ça va ?*, *comment ça se fait ?*) sont nettement plus fréquentes que les structures SVQ.

De façon attendue, la variante avec inversion complexe (*comment cela se fait-il ?*) ne se rencontre que dans le corpus romanesque, où elle apparaît très rare. En revanche, l'inversion clitique, en principe exceptionnelle en français informel, est pratiquement présente dans toutes les expressions pour lesquelles elle est syntaxiquement possible à l'exception de *tu sais quoi ?* La figure 4 compare les structures qui acceptent les variantes avec inversion clitique et avec *est-ce que*, en observant la productivité de chaque structure.

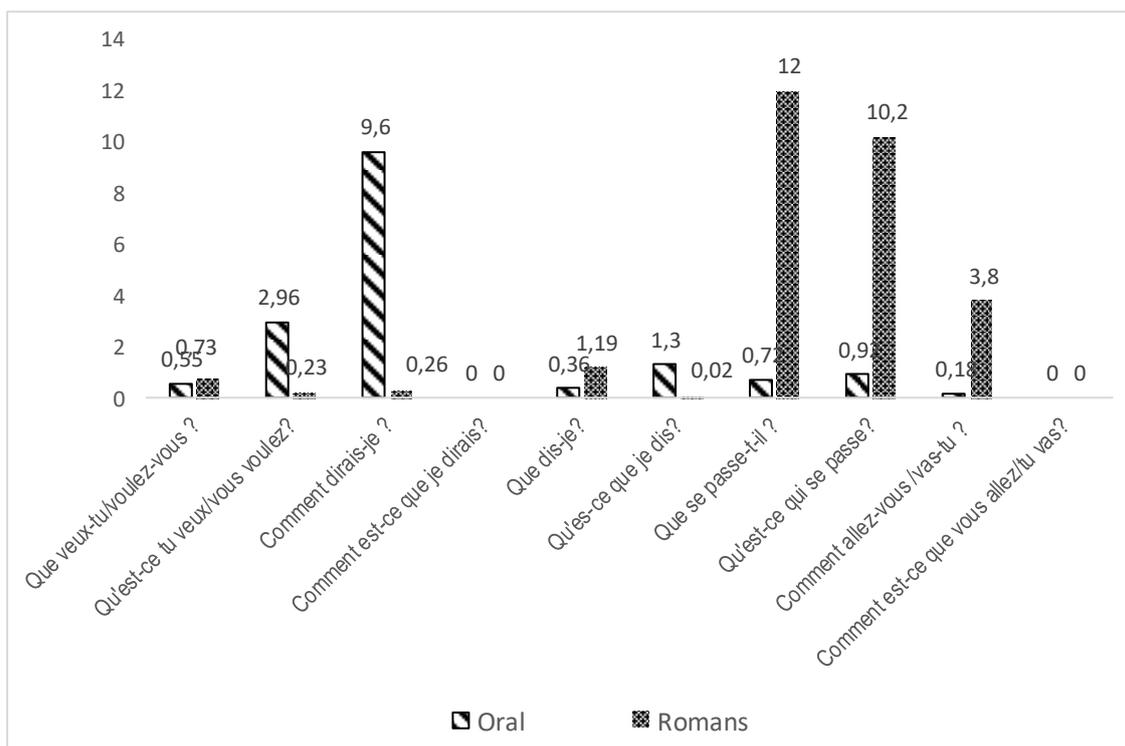


Figure 4 : Comparaison des structures avec inversion clitique et *est-ce que* pour un échantillon de PIPP

Même si l'inversion clitique apparaît moins fréquente dans l'oral authentique (en dehors de l'expression *comment dirais-je ?* où elle est majoritaire) que dans les romans, elle est attestée dans le corpus oral pour toutes les PIPP qui acceptent cette structure<sup>27</sup>. Cela confirme les analyses de Stark et Binder (2021) qui observent une proportion importante de tournures figées (12,2%) dans les inversions clitiques de l'oral (sur le corpus ORFEO-CEFC). De façon intéressante, on observe par ailleurs que des tournures avec clitique comme *comment dirais-je*, qui sont souvent considérées comme des marqueurs discursifs (cf. citation de Lefevre et Rossi-Gensane, 2015, plus haut), connaissent tout de même des alternances syntaxiques<sup>28</sup>, en particulier avec la structure QSV, qui est assez fréquente dans le corpus oral :

- (8) [spk2] oh je leur ai fait la totale hein / c' était bien / hm hm / ah ouais c' était un b- un / un bon souvenir / je me souv- je m' en souviens très bien parce que je suis parti là - bas j' ai- j' avais lu depuis huit jours je lisais que les poèmes contre-révolutionnaires / de Julian Beck / qui est un grand auteur de théâtre euh new yorkais je crois / qui a écrit beaucoup de choses très très engagées un peu à la Michael Moore, mais  
[spk2], mais plus comment je dirais euh plus euh  
[spk2] hm années soixante quoi  
(Corpus ESLO2\_ENT\_1029\_C)

L'observation de la structure avec *est-ce que* confirme plusieurs analyses antérieures (en particulier Quillard, 2001 ; Guryev, 2017) : la séquence interrogative *comment est-ce que* est très peu utilisée avec les PIPP de notre corpus : seules 2 occurrences apparaissent dans le corpus

<sup>27</sup> Rappelons ici que les sujets avec *ça* ne permettent pas cette structure.

<sup>28</sup> C'est un argument pour nous en faveur de l'intégration de ces marqueurs discursifs à structure phrastique dans la classe des phrases préfabriquées.

oral (mais aucune dans le corpus PhraseoRom). La rareté de cette structure peut s'expliquer d'une part par le caractère très familier de la forme interrogative (par comparaison au tout à fait standard *qu'est-ce que*), et peut-être d'autre part, par le caractère redondant de la particule *est-ce que*, inutile dans la phrase interrogative avec *comment* en tête de phrase sans inversion du sujet (*Comment ça va ?* vs *Comment est-ce que ça va ?*). En revanche, *qu'est-ce que* est tout à fait productif dans les deux corpus, mais avec une nette surreprésentation dans le corpus oral. En bref, le corpus oral préfère – sans exclure l'inversion clitique – la structure *qu'est-ce* alors que le corpus romanesque tend à favoriser l'inversion clitique, se rapprochant ici du registre formel.

Au terme de ce sondage, l'examen des structures syntaxiques sur un échantillon de PIPP confirme bien, comme mentionné par Larrivée et Guryev (2021 : 16), que « la sélection des variantes de l'interrogation se trouve conditionnée par une multitude de facteurs. La concurrence entre elles ne se manifeste pas sur un seul plan, mais implique une combinaison de contraintes : morphosyntaxiques, sémantico-pragmatiques et sociolinguistiques ».

On retiendra de ces données parfois difficiles à démêler les principaux points suivants :

- Le comportement syntaxique des expressions examinées est assez hétérogène, et incite à élargir les corpus de phrases examinées ainsi que les sources textuelles pour y voir plus clair.
- Le figement syntaxique existe bien avec certaines PIPP, mais est loin d'être majoritaire. Comme les interrogatives standard, la plupart des PIPP sont susceptibles de variations syntaxiques, y compris à l'intérieur du même corpus.
- On observe un net effet du registre et de la modalité écrite/orale sur les réalisations syntaxiques des PIPP : l'oral représenté, tout en reproduisant des dialogues parfois familiers, comporte bien les caractéristiques syntaxiques du registre formel, avec une surreprésentation des inversions clitiques et une sous-représentation des structures SVQ.
- La préfabrication semble bien avoir un effet réel sur la structure syntaxique des PIPP. On relève ainsi dans le corpus oral (par rapport aux données de la littérature sur les interrogatives standard) une surreprésentation des inversions clitiques et une sous-représentation des structures SVQ. Une comparaison approfondie avec des interrogatives standard de même structure (avec les mots interrogatifs *comment* et *que/quoi* par exemple) dans les corpus ORFEO-CEFC et ESLO2 permettrait de confirmer ce résultat.

## 5. Pour conclure

Cette étude exploratoire sur les PIPP a permis d'en circonscrire certaines caractéristiques intéressantes.

Le premier élément à souligner est la productivité de cette classe de phrases qui dépasse la notion de « question rhétorique », souvent abordée dans la littérature. Un ensemble de PIPP se caractérise bien par une fonction non questionnante (expressivité, commentaire métaénonciatif, expression de l'accord ou désaccord). On relève toutefois aussi un ensemble de PIPP qui correspondent classiquement à des demandes, parmi les phrases rituelles, les demandes métaénonciatives ou les phrases situationnelles. Il apparaît donc essentiel de s'intéresser à ces énoncés préfabriqués dans toute leur diversité.

Sur le plan syntaxique, les PIPP présentent des structures étonnamment variées et finalement assez peu figées, si on les compare aux interrogatives standard décrites dans la littérature. L'examen de 10 PIPP et de leurs variantes dans un corpus d'oral authentique et un corpus romanesque contemporain permet de montrer que, comme on s'y attendait, la variation est étroitement liée à la modalité écrite/orale et au registre : les structures syntaxiques de l'oral représenté relèvent en grande partie du registre formel avec une sur

représentation des inversions clitiques et une sous-représentation des structures *in situ*. Toutefois, on observe dans le corpus oral, en comparaison avec les données de littérature, des différences qui semblent attribuables à la nature préfabriquée des expressions : une surreprésentation des inversions clitiques (qui confirme l'étude de Stark & Binder, 2021) et une sous-représentation des structures *in situ*, deux caractéristiques relevant habituellement davantage du registre formel et/ou d'un état de langue plus ancien.

Bien entendu, d'autres paramètres restent à explorer pour mieux comprendre le fonctionnement des PIPP et leur variation : a) il serait essentiel de mieux prendre en compte les fonctions globales de ces PIPP ; on peut ainsi faire l'hypothèse que, sur le plan formel, les énoncés de type métaénonciatif fonctionnent différemment des énoncés à fonction expressive ; b) il serait aussi pertinent d'observer plus finement la distribution des expressions dans les interactions (début de tour de parole ou non ; présence ou non dans le monologue, ...) ; c) enfin, la variation intracorpus entrevue dans cette étude montre qu'il faut mieux prendre en compte les différents sous-types de l'oral et de l'écrit romanesque ; les PIPP de l'interview ne sont de toute évidence pas celles de la conversation informelle entre proches ou de la réunion de travail.

## Références

- Abeillé, A., & Godard, D. (2021). *La grande grammaire du français*. Arles : Éditions Actes Sud.
- Abouda, L., Baude, O. (2006). Constituer et exploiter un grand corpus oral : choix et enjeux théoriques. Le cas des ESLO. In Rastier, F., Ballabriga, M. (Eds.), *Corpus en lettres et sciences sociales : des documents numériques à l'interprétation* (pp. 143-50). Paris : Texto.
- Anscombre, J. C. (2000). Parole proverbiale et structures métriques. *Langages*, 6-26.
- Bally, Ch. (1951). *Traité de stylistique française*. Paris : Klincksieck.
- Benzitoun, C., Debaisieux, J. M. (eds)(2020). Orféo: un corpus et une plateforme pour l'étude du français contemporain. *Langages*, 219(3).
- Blanco, X., Mejri, S. (2018). *Les pragmatèmes*. Paris : Champion.
- Borillo, A. (1981). Quelques aspects de la question rhétorique en français. *DRLAV. Documentation et Recherche En Linguistique Allemande Vincennes*, 25(1), 1-33.
- Branca-Rosoff, S., Fleury, S., Lefeuvre, F., & Pires, M. (2012), Discours sur la ville. Présentation du Corpus de Français Parlé Parisien des années 2000 (CFPP2000). [consulté le 14-04-2022, <http://cfpp2000.univ-paris3.fr/CFPP2000.pdf>]
- Corpas-Pastor, G. (1996). *Manual de fraseología española*. Gredos.
- Coulmas, F. (1985). *Conversational routine: Explorations in standardized communication situations and prepatterned speech*. Berlin : Walter de Gruyter.
- Coveney, A. (2011). L'interrogation directe. *Travaux de linguistique*, 63(2), 112-145.
- Cowie, A. P. (2001). Speech formulae in English: problems of analysis and dictionary treatment. *GAGL: Groninger Arbeiten zur germanistischen Linguistik*, (44), 1-12.
- Dekhissi, L., & Coveney, A. (2021). Le contexte linguistique des questions rhétoriques conflictuelles et la variation entre *pourquoi* et *qu'est-ce que*. *Langue française*, 212(4), 123-137.
- Dostie, G. (2019). Paramètres pour définir et classer les phrases préfabriquées: *La vengeance est un plat qui se mange froid. Bon appétit!*. *Cahiers de lexicologie*, 2019(114), 27-61.
- Durrer, S. (2005). *Le dialogue dans le roman*. Paris : Armand Colin.
- Fónagy, I. (1997). Figement et changement sémantique. In M. Martins-Baltar (Ed.), *La locution entre langue et usages* (pp. 131-164). Fontenay-Saint-Cloud : ENS Éditions.

- Guryev, A. (2017). *La forme des interrogatives dans le Corpus suisse de SMS en français: étude multidimensionnelle* (Doctoral dissertation, Université de Neuchâtel).
- Guryev, A. (2021). *Car je suis un... tu sais quoi? Rappeur*. Étude pragma-syntaxique de la séquence *Tu sais quoi?* *Lexique*, 219, 97-115.
- Kauffer, M. (2019). Les “actes de langage stéréotypés”: essai de synthèse critique. *Cahiers de lexicologie*, 2019(114), 149-171.
- Kerbrat-Orrechioni, C. (2016). *Les Actes de langage dans le discours*. Paris : NathanArmand Colin.
- Kerbrat-Orrechioni, C. (ed.)(2001). *La question*. Lyon : PUL.
- Kraif, O. (2019). Explorer la combinatoire lexico-syntaxique des mots et expressions avec le LEXICOSCOPE. *Langue française*, 203(3), 67-82.
- Krzyżanowska, A., Grossmann, F., Kwapisz-Osadnik, K. (2021). *Les formules expressives de la conversation Analyse contrastive: français-polonais-italien*. Lublin : Episteme.
- Larrivée, P., & Guryev, A. (2021). Variantes formelles de l’interrogation. Présentation. *Langue française*, 212(4), 9-24.
- Le Pesant, D. (2019). Suggestions méthodologiques et outils de traitement de corpus pour l’étude des Phrases Préfabriquées des Interactions. *Cahiers de lexicologie*, 2019(114), 93-118.
- Lefevre, F., & Rossi-Gensane, N. (2015). Interrogation. *Projet Fracov*. <http://www.univ-paris3.fr/index-des-fiches-227311.kjsp>.
- López Simó, M. (2016). *Fórmulas de la conversación. Propuesta de definición y clasificación con vistas a su traducción español-francés, francés-español*. Thèse de doctorat. Université d’Alicante.
- Lüger, H. (2007). Pragmatische Phraseme: Routineforme. In Burger, H. (Eds). *Phraseologie: ein internationales Handbuch der zeitgenössischen Forschung* (pp. 444–459). Berlin : Mouton de Gruyter.
- Martins-Baltar, M. (1995). Énoncés de motif usuels: figures de phrase et procès en déraison. *Cahiers du français contemporain*, 2, 87-118.
- Mel’čuk, I. (2015). Clichés, an understudied subclass of phrasemes. *Yearbook of Phraseology*, 6(1), 55-86.
- Novakova, I., & Siepman, D. (Eds.). (2019). *Phraseology and Style in Subgenres of the Novel: A Synthesis of Corpus and Literary Perspectives*. Springer Nature.
- Pausé, M. S., Tutin, A., Kraif, O., & Coavoux, M. (2022). Extraction de Phrases Préfabriquées des Interactions à partir d’un corpus arboré du français parlé: une étude exploratoire. In *SHS Web of Conferences* (CMLF, Tours). EDP Sciences.
- Quillard, V. (2001). La diversité des formes interrogatives : comment l’interpréter ?, *Langage et société*, 95, 57-72.
- Riegel, M., Pellat, J. C., & Rioul, R. (2018). *Grammaire méthodique du français*. 7<sup>ème</sup> édition. Paris : Presses Universitaires de France.
- Stark, E., & Binder, L. (2021). L’inversion du sujet clitique en français oral: ultime apanage des interrogatives?. *Langue française*, 212(4), 25-40.
- Tutin, A. (2019). Phrases préfabriquées des interactions: quelques observations sur le corpus CLAPI. *Cahiers de lexicologie*, 2019(114), 63-91.
- Tutin, A. Gharbi N. (2020). Phrases préfabriquées à fonction expressive dans les dialogues de romans contemporains. In Fesenmeier, L., Novakova, I. (Eds). *Phraseologie et stylistique de la langue littéraire* (pp.17-29). Berlin : Peter Lang.

Author address:

Agnès Tutin  
Laboratoire Lidilem  
Université Grenoble Alpes  
Bâtiment Stendhal  
CS40700  
38058 Grenoble cedex 9  
France  
[agnes.tutin@univ-grenoble-alpes.fr](mailto:agnes.tutin@univ-grenoble-alpes.fr)