



„Aukštųjų mokyklų tinklo optimizavimas ir studijų kokybės gerinimas Šiaulių universiteta prijungiant prie Vilniaus universiteto“ Nr. 09.3.1-ESFA-V-738-03-0001. Projektas finansuojamas iš Europos socialinio fondo.



Kurjame  
Lietuvos ateitį  
2014-2020 metų  
Europos Sąjungos  
fondų investicijų  
veiklos programa



# Acquisition du FLE et corpus linguistique : *diagnostic, usage, apprentissage*

dr Vitalija KAZLAUSKIENE  
Université de Vilnius, LITUANIE

*demander ou poser  
une question*

*produit vieilli  
ou périmé*

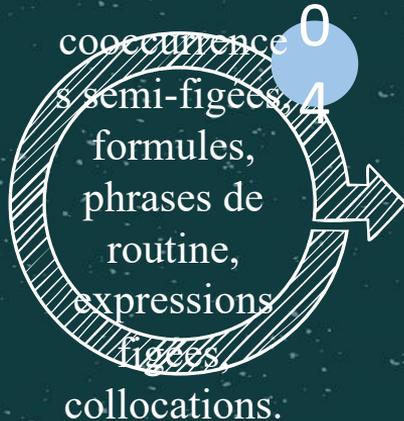
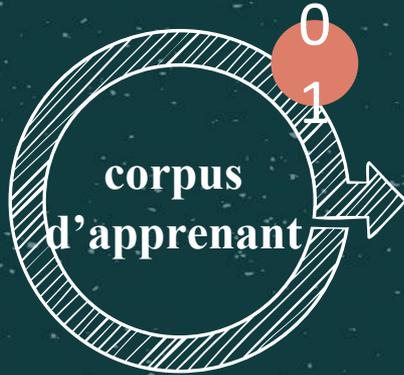
*faire ou prendre une photo*

*recevoir ou obtenir un diplôme*

*visa fini ou expiré*

- L'acquisition de la compétence lexicale est complexe, dans la mesure où il ne suffit pas pour les apprenants de distinguer la forme et la signification d'un élément lexical pour savoir comment l'intégrer correctement dans le contexte langagier. On ne communique pas par des mots isolés. Souvent le mot, dont la signification semble transparente, dans un contexte représente la signification inattendue et ce n'est que dans l'ensemble des mots, dans le contexte langagier, parfois même social qu'il est possible de le déchiffrer correctement.
- L'importance des usages stéréotypés de la langue est reconnue par la plupart des études didactiques sur le lexique (Granger and Paquot, 2008). Ce type de relation lexicale constitue un domaine du lexique difficile à maîtriser par les apprenants non natifs.
- Nous abordons des questions allocentrées de linguistique, autrement dit des questions que se pose un apprenant du français L2 et pour lesquelles il est difficile parfois de trouver des réponses dans les grammaires savantes ou manuels de FLE.
- Un diagnostic de la maîtrise de ces expressions par les apprenants aussi bien des difficultés qui se posent, dans la perspective d'élaborer des matériaux didactiques adaptés à leurs besoins.

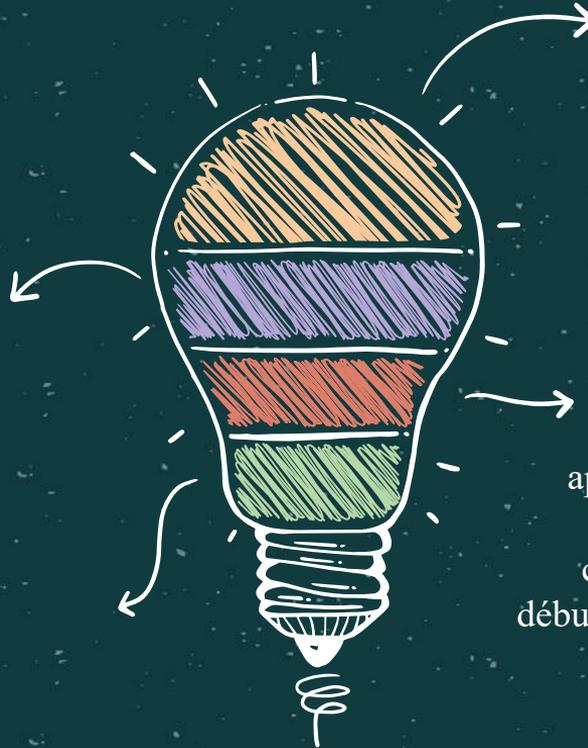
# CHALKBOARD BACKGROUND INFOGRAPHICS



# corpus d'apprenants

Les corpus collectifs des apprenants peuvent être révélateurs de la description de leur compétence transitoire.

Le diagnostic des difficultés propres à l'interlangue permet de créer les activités plus efficaces.



Le diagnostic des difficultés propres à l'interlangue permet d'évaluer la progression d'une langue cible, de la décrire, de relever la variété hégémonique

Le diagnostic des écrits des apprenants à son tour démontre la nécessité nette d'introduire les cooccurrences semi-figées dès le début de l'apprentissage, de travailler maximum le bagage lexical, une variété de formules, phrases de routine, expressions figées, collocations.

- CLASSIFICATION DES COLLOCATIONS ?

???

## LA NOTION DE COLLOCATION

- D'après Tutin et Grossmann (2002), « on caractérise souvent les collocations par le fait que le sens est transparent en réception (il se « devine »), alors que, pour un locuteur non-natif, il est difficile de produire les lexèmes adéquats ».
- La rencontre fréquente de deux lexies qui se suivent ou sont séparés par d'autres mots. Haussmann soutient qu'« il ne suffit pas de voir la collocation comme une combinaison sous contrainte, il faut encore comprendre qu'elle est une combinaison orientée. En tenant compte de cette orientation on appelle l'un des éléments de la collocation – base et l'autre – collocatif.
- *célibataire endurci* En d'autres termes, la première unité lexicale (la base ou mot-clé) est choisie librement. La seconde unité lexicale (le collocatif) est choisie en fonction de la base (Heid et Raab 1989 ; Mel'cuk 1998 ; Mel'cuk et Polguère 2007).
- Les collocations forment une unité de sens, et elles sont soumises à certaines restrictions syntaxiques, sémantiques et pragmatiques. Ces unités linguistiques se caractérisent par un certain degré de figement et de récurrence.

# CLASSIFICATION DES COLLOCATIONS

- La *collocation grammaticale* est une combinaison courante, composée d'un mot lexical (du type verbe, nom ou adjectif) et d'un mot grammatical, le plus souvent d'une préposition. En revanche, la *collocation lexicale*, est composée de deux éléments lexicaux égaux.
- La typologie proposée par Cowie (1981 : 230) est : « open collocation » (collocation transparente) et « restricted collocation » (collocation restreinte). La première catégorie comprend des combinaisons de mots qui, sur le plan sémantique, sont tellement généraux qu'ils peuvent se combiner de façon assez libre. A propos des collocations restreintes, Cowie estime que le sens du mot-clé dans un groupe de mots limite les possibilités pour les autres éléments. Ces derniers sont beaucoup plus intéressants sur le plan lexicographique.
- La catégorisation de Mel'čuk (1993 ; 1995 ; 2008) est plutôt sémantique. Il souligne la nécessité de répertorier dans la description d'une langue les unités qu'il appelle phrasèmes et qui reposent sur quatre types : phrasèmes pragmatiques (pragmatèmes) ; phrasèmes complets ou expression idiomatique (complètement figée) ; demi-phrasèmes ou collocation ; quasi-phrasème ou expression quasi-idiomatique (où le sens inclut le sens de tous ces constituants, mais aussi un sens additionnel, imprévisible).

- collocations lexicales jouent un rôle primordial dans l'acquisition du vocabulaire du français langue étrangère et soulèvent souvent des difficultés pour les apprenants.
- Zgusta (1971 : 296) utilise le terme « anisomorphisme » pour désigner le phénomène de l'absence d'équivalent total entre les deux langues.
- les difficultés concernant l'acquisition des collocations pourraient s'expliquer par les points suivants : la liste non-exhaustive des collocations ; l'influence de la langue maternelle ; la faible présence des collocations dans les dictionnaires ou manuels. Ainsi, elle suggère que « la collocation doit être apprise comme un tout, une unité. »
- Pour s'exprimer correctement et sans efforts, un apprenant étranger doit donc disposer d'un large répertoire de collocations, comme celui que possède un locuteur natif dans son lexique mental.

## INTERLANGUE

- **La notion d'interlangue** (Selinker, 1972) décrit **la variété de langue qui se forme chez l'apprenant** et qui ne coïncide pas totalement avec la langue d'apprentissage ; cette variété de langue est **provisoire** et évolue en principe au fur et à mesure de l'apprentissage – pour le professeur est essentiel **d'éviter la fossilisation de l'interlangue.**

Le transfert négatif de la langue maternelle pourrait causer des erreurs interlinguistiques  
Les particularités syntaxiques et sémantiques des collocations françaises posent des difficultés

*\*demander une question (poser), \* recevoir (obtenir) un diplôme, \*faire (prendre) une photo, \*visa fini (expiré), \*produit vieilli (périmé), etc.*

A la condition que le phénomène collocatif apparaisse l'une des sources de nombreuses erreurs chez les apprenants étrangers, l'enseignement des collocations devient alors comme une nécessité.

La stratégie de la « collocation compétence (compétence collocationnelle) » est proposée par Brown (1994 : 25). D'après lui, c'est une stratégie active pour les apprenants de haut niveau. La compétence collocationnelle est un critère important pour évaluer la profondeur des connaissances du vocabulaire chez des apprenants, elle occupe une place importante dans l'enseignement d'une langue étrangère.

Ces routines sont définies comme « des énoncés récurrents construits autour d'un verbe et d'arguments remplissant différents rôles sémantiques dont une partie est actualisée dans le discours. Chaque routine peut être associée à une fonction discursive/ rhétorique spécifique » (Tutin & Kraif, 2016, p. 121).

**les typologies d'erreurs** aux trois niveaux : inspiration contrastive, syntaxique et sémantique.

### **Typologie d'erreurs d'inspiration contrastive**

en adoptant à la fois le principe de l'analyse d'erreurs et celui de l'analyse contrastive entre langue cible et langue source (Tréville et Duquette 1996 ; Millischer 2000), on voit que les erreurs sont réparties en quatre catégories : erreur interlinguistique, erreur intralinguistique, erreur extralinguistique (méthodologique), erreur de stratégie communicative.

### **Typologie d'erreurs en fonction syntaxique**

En fonction de la grammaire française, les erreurs de collocations commises par des apprenants se présentent sous forme de : omission, ajouts, déplacement, substitution, pléonasme, interférence, etc.

Nous remarquons que le type « verbe + nom » est le plus fréquent, ce qui représente presque la moitié des erreurs. Il est suivi par d'autres types d'erreurs tels que « adjectif + nom », « nom + adjectif », « nom + nom », « adverbe + nom », « nom + verbe ». Un type d'erreur n'a pas été repéré dans le corpus, c'est le type « adverbe + adjectif ». Cela pourrait s'expliquer par la taille du corpus et la différence des niveaux de français des apprenants.

### **Typologie d'erreurs en fonction sémantique**

En fonction de la lexicologie explicative et combinatoire, branche lexicographique de la Théorie Sens-Texte (Mel'čuk, 1998), les erreurs lexicales liées aux propriétés de combinatoire lexicale, à savoir les problèmes de collocation, sont regroupés en quatre groupes : *choix d'un mauvais collocatif (collocation inexistante), choix d'un collocatif valable (dont le sens véhiculé n'est pas le bon), non-respect de contraintes d'emploi du collocatif, et choix d'un collocatif qui appartient à une autre lexie du même vocable* (Anctil, 2010 : lii).

# FACTEURS INTERVENANT DANS LE PROCESSUS D'ACQUISITION DU VOCABULAIRE

## Transfert négatif de la langue maternelle

- Le rôle de la langue maternelle et son influence dans l'acquisition d'une langue étrangère a donné lieu à de nombreuses hypothèses. (Klein, 1989 : 42)
- D'un point de vue pédagogique, il est bon de connaître les possibilités de transfert positif pour les favoriser.
- d'une traduction littérale de la langue maternelle : \**gagner un salaire*

## Nuance des synonymes français

- Les nuances des synonymes en termes de collocations ne sont pas évidentes aux yeux des apprenants. Ex :
- (1) \**la génération dernière*
- (2) \**un mûr consommateur*

## LES DIFFICULTÉS DANS L'ACQUISITION DU FRANÇAIS LANGUE SECONDE

1. <...> les technologies #modernes\$ *moderne* #ce n'est pas\$ *snapas* #une\$ #bonne\$ **bien** #solution\$ *solutions*. (2011E301898)
2. *Bonjour, mon* #bon\$ **bien ami**. *Ça va?* (2013L165432)
3. *Peut être il* #commençait\$ *commençait beaucoup* #de\$ *des* #médicaments\$ *médicaments pour* #une\$ *un* #longue\$ **large vie** ? (2011E301931)
4. *Mais il y a* #un\$ **une** #sérieux\$ **sérieuse** #problème\$ **probleme**. (2015L2643D35)
5. *J'ai aussi peur que cette chose soit contemporaine et* **le jour** #suivant\$ **prochain** *tout va changer*. (2011E301892)
6. *Néanmoins, quelquefois les jeunes sont trop sous une pression* **des personnes mûres**. (2015E2643D13)
7. #le\$ *Les test de français était* #très\$ **très** #difficile\$ **dificile** *et je* #avais\$ *peur que je ne* #finirerais\$ *finiserais* #mon\$ *ma école* #aux\$ *avec* #résultats\$ **resultats** #mauvais résultats\$ **mal**. (2015L2643D04)

## COOCCURRENCES SEMI-FIGÉES LE BAGAGE LEXICAL, UNE VARIÉTÉ DE FORMULES, PHRASES DE ROUTINE, EXPRESSIONS FIGÉES, COLLOCATIONS ET L'APPRENTISSAGE

- La question des aides et des moyens possibles pour atteindre ce but se pose. Les dictionnaires peuvent certainement proposer des solutions, dans une certaine mesure, mais ils vont avoir tendance à s'orienter plutôt vers des mots isolés. Les cooccurrences lexicales, les collocations « on les retrouve la plupart du temps dans la section réservée aux exemples, mais toutes les collocations possiblement contrôlées par un mot n'y figurent pas forcément » (Tremblay, 2014, p. 75). Avec le développement de la linguistique de corpus, les dictionnaires collocationnels deviennent de plus en plus répandus. Le *Dictionnaire des combinaisons de mots* (Le Fur, 2007) en est un exemple.
- Quoi qu'il en soit, Martinez affirme que l'apprentissage des expressions semi-figées fait partie des acquisitions langagières inaccomplies des programmes, à tous les niveaux d'enseignement et d'apprentissage du FLE (Martinez, 2018, p. 36). Cependant, si on essaie de comparer les méthodes d'avant l'approche communicative, la perspective actionnelle et celles d'aujourd'hui, il est manifeste que ce type des relations lexicales apparaissent au fur et à mesure de l'enseignement du français dès le niveau faux débutant. Néanmoins, les listes de mots et d'expressions figées sont encore fréquentes, ainsi que les tableaux de verbes conjugués qui sont nombreux, eux aussi.

# Collocations dans les manuels du FLE

*Faire du shopping*

*Etre passionné par la lecture*

*J'ai une grande nouvelle*

*Elle va bien*

*Vous avez l'heure, svp ?*

*Les horaires d'ouverture*

*Etc.*

## Réflexion sur l'enseignement des collocations dans le FLE

L'observation et l'analyse par rapport aux erreurs rencontrées en production sont nécessaires mais pas suffisantes. Il faut aller plus loin dans cette démarche et trouver un remède tendant vers l'autonomie de l'apprenant. Face à ces erreurs, l'enseignant doit être capable d'apporter des outils à l'apprenant afin qu'il n'y ait pas de fossilisation ni de stagnation de l'apprentissage. Ainsi, l'enseignant et l'apprenant se servent d'outils variés en fonction des différents types d'erreurs rencontrées.

# CORPUS POUR L'APPRENTISSAGE

- Le corpus d'une langue authentique constitue une ressource lexicale intégrant les patrons verbaux, les routines verbales permettent la réalisation d'activités d'apprentissage dans la perspective de l'aide à la rédaction scientifique.
- Inscrite dans la linguistique de corpus, l'approche adoptée est fondée sur le modèle Corpus Pattern Analysis (CPA, analyse des patrons basée sur les corpus) (Hanks, 2013). Ce modèle visant à analyser les usages réguliers des verbes en contexte, permet d'accéder aux sens du verbe à travers l'analyse des constructions syntaxiques et des co-occurrences qui lui sont associées, soit l'analyse des emplois usuels du verbe dans le domaine de l'écrit scientifique avec une visée didactique. La méthode d'analyse a combiné des techniques de traitement automatique du langage (TAL) et l'analyse linguistique manuelle.
- Il serait donc faux de croire qu'il faut commencer par présenter des mots isolés pour passer ensuite aux combinaisons de mots. Certaines collocations et expressions très courantes méritent au contraire d'être introduites dès le début, de même que certaines expressions ou formules de communication, certains bouts de phrases.
- L'analyse des corpus de la fréquence de ces unités polylexicales dans la langue démontre d'ailleurs que leur fréquence dépasse souvent celle de mots tout à fait courants. On pensera ici, par exemple, à des unités polylexicales comme *faire l'objet de*, *avoir besoin de*, *jouer un rôle*, *mettre en scène*, qui présentent une fréquence d'apparition qui les situerait dans la tranche des 2000 mots les plus fréquents du français (S. Verlinde et T. Selva, 2001).
- Travailler le corpus de Google; YouTube, les corpus de la langue et les corpus des apprenants pour l'autocorrection

Elle a été appliquée au corpus Transdisciplinaire-TermITH qui se compose de 500 articles représentant 10 disciplines des sciences humaines et sociales (Hatier et coll., 2016). L'extraction et la modélisation des routines verbales ont été appliquées à une liste des verbes établie par Hatier (2016) dans le cadre du projet TermITH1, selon les étapes suivantes :

Extraction des cadres de sous-catégorisation

Repérage des acceptions

Classification des verbes et traitement des routines

Effectivement, nous pouvons imaginer une leçon au cours de laquelle l'enseignant veut introduire le lexique du domaine écologique, celle-ci peut être réalisée, sans détailler la terminologie de la collocation, en abordant les différences réelles entre *ville verte*, *espace verte*, etc. ou *poser et demander une question?*

Une accessibilité et une adaptation langagière s'imposent pour attirer l'attention des jeunes et faciliter la compréhension à l'aide de supports ludiques comme les vidéos animées visualisables et téléchargeables sur YouTube

La méthode de classe inversée marche parfaitement dans le travail avec des corpus, l'analyse individuelle et le travail collectif ensuite.

Des petits projets linguistiques, liste des collocations syntaxiques, pour les douées les collocations lexicales ou même pragmatèmes. Toujours le texte et le contexte, des documents authentiques.

# SKETCHENGINE

The screenshot displays the SketchEngine web interface for the word "poser". The browser address bar shows "app.sketchengine.eu". The page title is "WORD SKETCH" and the source is "EUROPARL7 sample, French". The search term is "poser as verb 3,106x".

The interface is divided into several columns, each representing a different grammatical or semantic category for the word "poser":

- objects of "poser"**: Includes "question", "problème", "jalon", "fondation", "base", "fondement", "défi", "risque", "choix", "candidature", "pierre", and "diagnostic".
- subjects of "poser"**: Includes "situation", "rapport", and "question".
- modifiers of "poser"**: Includes "alors", "aujourd'hui", "actuellement", "directement", "certainement", "maintenant", "effectivement", "toutefois", "cependant", "ensuite", "ici", and "comme".
- pronominal objects of "poser"**: Includes "se", "vous", "me", and "nous".
- pronominal subjects of "poser"**: Includes "vous", "elle", "on", "je", "nous", and "il".
- prepositional phrases**: Includes "poser" du, "poser" de, "poser" au, "poser" en, "poser" quant, and "poser" sans.

A vertical sidebar on the left contains navigation icons for home, search, and other functions. The top right corner includes a login status ("Not logged in") and various utility icons.

# CONCORDANCE

EUROPARL7 sample, French

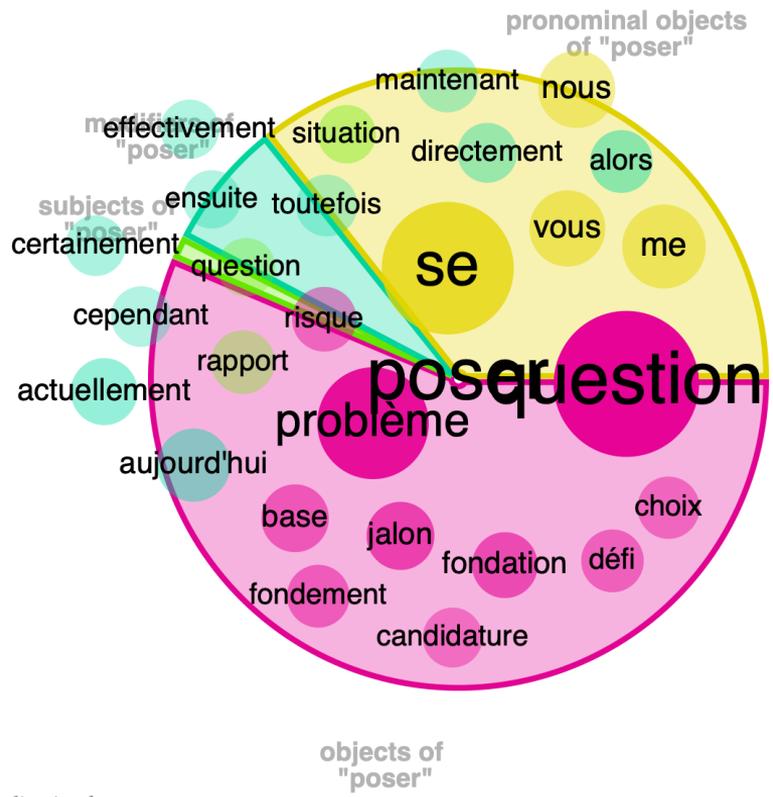
Not logged in  
[Log in](#)

CQL **poser + question** • 791  
42.35 per million tokens • 0.0042%

[🔍](#) [📄](#) [🔍](#) [☰](#) [👁](#) [✂](#) [≡](#) [≡](#) [GD EX](#) [📄](#) [↔](#) [📄](#) **KWIC** [🔍](#)

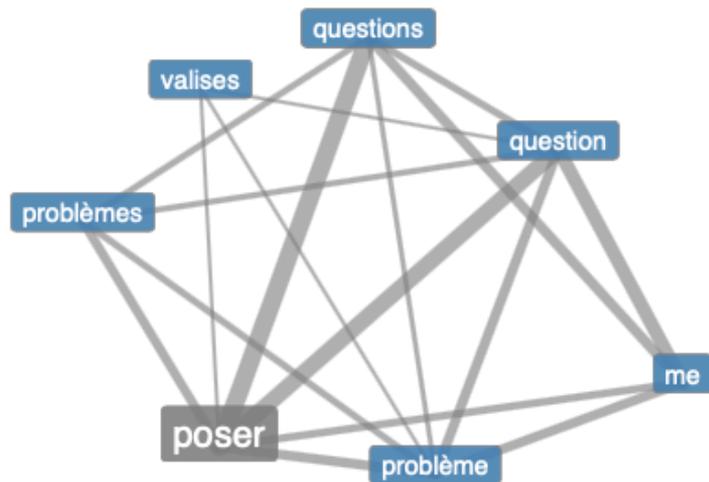
Details  Left context  KWIC  Right context

1	<input type="checkbox"/>	<a href="#">🔍</a> #113523	unique européen s' élève environ à 1 milliard d' euros.</s></s>Paolo Costa	<b>pose</b>	une première <b>question</b> :	la Commission peut -elle se satisfaire d' une telle
2	<input type="checkbox"/>	<a href="#">🔍</a> #128028	a à l' Institut européen de technologie?</s></s>En même temps, j' aimerais	<b>poser</b>	la <b>question</b> suivante:	ces parrains privés imposeront -ils aussi leurs conditi
3	<input type="checkbox"/>	<a href="#">🔍</a> #174786	s initier un suivi des conclusions de la commission d' enquête sur la CIA et	<b>poser</b>	la <b>question</b> formelle suivante:	quelles mesures le Conseil et la Commissio
4	<input type="checkbox"/>	<a href="#">🔍</a> #175432	d' anti-américaine.</s></s>Cependant ce sont en réalité les Américains qui	<b>posent</b>	les <b>questions</b> critiques,	qui ont réalisé des évaluations globales ces dernie
5	<input type="checkbox"/>	<a href="#">🔍</a> #182342	a lettre et qui sont impliqués dans le terrorisme.</s></s>Nous devons nous	<b>poser</b>	la <b>question</b> de savoir pourquoi.</s></s>Madame la Présidente,	le terrorism
6	<input type="checkbox"/>	<a href="#">🔍</a> #183261	ste.</s></s>(EL) Madame la Présidente, permettez -moi de commencer par	<b>poser</b>	deux <b>questions</b> fondamentales au Conseil.</s></s>Premièrement, pourqu	
7	<input type="checkbox"/>	<a href="#">🔍</a> #217000	ie.</s></s>En ce qui concerne les points mentionnés ci-dessus, je voudrais	<b>poser</b>	une <b>question</b> à la Commission:	la Commission annoncera -t-elle, lors de la
8	<input type="checkbox"/>	<a href="#">🔍</a> #230449	unauté internationale, y compris l' UE, est impuissante.</s></s>Je souhaite	<b>poser</b>	une courte <b>question</b> supplémentaire au Conseil,	étant donné que la présid
9	<input type="checkbox"/>	<a href="#">🔍</a> #245060	ikhe.</s></s>En ce qui concerne la Chine et la Russie, nous devrions nous	<b>poser</b>	la <b>question</b> de savoir si les dialogues -	au cours desquels nous avons expi
10	<input type="checkbox"/>	<a href="#">🔍</a> #245923	'age et les sanctions spécifiques instaurées par l' UE.</s></s>Et nous nous	<b>posons</b>	la <b>question</b> de savoir pourquoi nos mesures sont inefficaces	et pourquoi le
11	<input type="checkbox"/>	<a href="#">🔍</a> #259345	èglement Rome I.</s></s>Nous ne pouvons poursuivre infiniment en nous	<b>posant</b>	la <b>question</b> de savoir quel droit sera applicable.</s></s>Nous avons besoin	
12	<input type="checkbox"/>	<a href="#">🔍</a> #276645	l défi pour tous les pays.</s></s>Ces questions sanitaires vont, à nouveau,	<b>poser</b>	la <b>question</b> de la qualité de la surveillance de nos frontières extérieures	ta
13	<input type="checkbox"/>	<a href="#">🔍</a> #315407	'écurité d' approvisionnement.</s></s>C' est pourquoi il est important de se	<b>poser</b>	les <b>questions</b> que plusieurs intervenants ont déjà soulevées	clairement au
14	<input type="checkbox"/>	<a href="#">🔍</a> #359475	u élevé de plomb dans les peintures utilisées pour ces jouets.</s></s>Cela	<b>pose</b>	la <b>question</b> très sérieuse du type de futur que nous souhaitons pour nos	er
15	<input type="checkbox"/>	<a href="#">🔍</a> #387800	née européenne contre la peine de mort, il conviendrait de demander, et je	<b>pose</b>	la <b>question</b> à la présidence portugaise,	dans quelles circonstances un pay
16	<input type="checkbox"/>	<a href="#">🔍</a> #405283	énir?</s></s>Il est trop tôt pour le dire, mais il est tout à fait important de se	<b>poser</b>	la <b>question</b> .</s></s>Et je remercie M. Savary d' avoir dit qu' en effet,	il y a
17	<input type="checkbox"/>	<a href="#">🔍</a> #420661	ponses apportées par le Commissaire.</s></s>Cela fait trois ans que nous	<b>posons</b>	nos <b>questions</b> et, à ce jour,	nous ne sommes pas encore parvenus à obter
18	<input type="checkbox"/>	<a href="#">🔍</a> #472909	avons besoin des meilleurs cerveaux du monde.</s></s>Les citoyens nous	<b>posent</b>	aussi une autre <b>question</b> :	" Prenez -vous nos préoccupations au sérieux è
19	<input type="checkbox"/>	<a href="#">🔍</a> #533812	l, je souhaiterais m' exprimer maintenant au nom de M. Navarro, et ensuite	<b>poser</b>	ma propre <b>question</b> ,	ou plutôt réitérer une question.</s></s>Ma question, :
20	<input type="checkbox"/>	<a href="#">🔍</a> #533967	'endant, je vous donnerai 15 secondes supplémentaires, si vous souhaitez	<b>poser</b>	votre <b>question</b> immédiatement car,	en vertu du règlement, je ne peux auto



visualization by

▲ Graphique



[https://corpora.uni-leipzig.de/fr/res?corpusId=fra\\_mixed\\_2012&word=posér](https://corpora.uni-leipzig.de/fr/res?corpusId=fra_mixed_2012&word=posér)

<https://learningapps.org/display?v=pijmj8stn22>

L'usage des corpus linguistiques est utile pour remédier, faire évaluer la langue des apprenants. L'enseignement du lexique peut s'avérer beaucoup plus efficace s'il est fait de manière structurée et systématique avec la prise en compte de la combinatoire lexicale et syntaxique des mots.