



9^e Colloque International de Linguistique de Corpus

9^o Congreso Internacional de Lingüística de Corpus

9th International Conference on Corpus Linguistics

MAY 31st to JUNE 2nd 2017

Corpus en classe de langue. Exemple avec les marqueurs d'exemplification et de reformulation

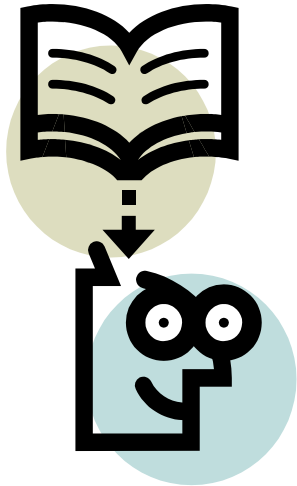
Thi Thu Hoai TRAN⁽¹⁾, Cristelle CAVALLA⁽²⁾

(1) Grammatica, Université d'Artois

(2) Diltec, Université Paris Sorbonne

Paris, 01 juin 2017

Plan



1. Cadre de travail : objectifs et hypothèse
2. Objet d'étude : marqueurs de reformulation et d'exemplification
3. Méthodologie de recherche
4. Expérimentation

Cadre de travail

- Travaux sur l'aide à l'écriture scientifique
- Objectifs didactiques :
 - Appropriation des spécificités des écrits scientifiques.
 - Utilisation du corpus en classe de langue
- Hypothèse :
 - Partir des représentations des étudiants sur la reformulation et l'exemplification et leur faire découvrir les corpus et les éléments linguistiques associés à ces deux valeurs

Cavalla, C., & Loiseau, M. (2013). Scientext comme corpus pour l'enseignement. In A. Tutin & F. Grossmann (Éd.), *L'écrit scientifique : du lexique au discours. Autour de scientext*. Rennes : Presses Universitaires de Rennes, 163-180.

Coxhead, A. (2000). « A New Academic Word List ». *TESOL Quarterly*, 34(2), 213-238.

Phal, A. (1971). *Vocabulaire général d'orientation scientifique*. Paris : Crédif.

9^e Colloque International de Linguistique de Corpus

Cadre de travail

Linguistique de
corpus

**Données authentiques
dans la classe de
langue (Boulton & Tyne,
2014; Chambers, 2010)**

Didactique du Français sur
Objectif Universitaire

(Mangiante & Parpette, 2011)

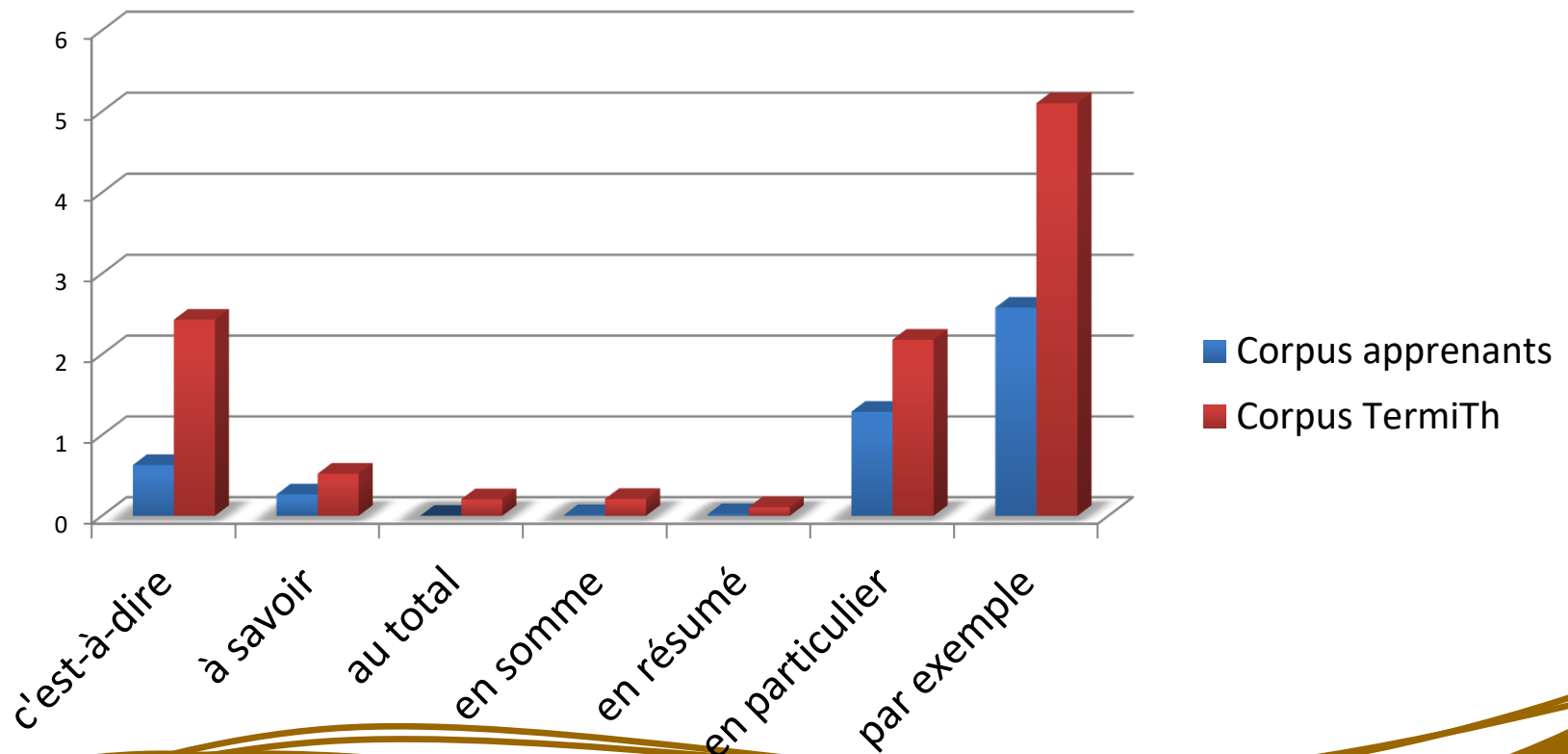
Analyse du discours
scientifique

Boulton, A., Tyne, H. 2014. Des documents authentiques aux corpus. Démarches pour l'apprentissage des langues. Paris : Didier.
Chambers, A. 2010. « L'apprentissage de l'écriture en langue seconde à l'aide d'un corpus spécialisé ». Revue française de linguistique appliquée, noXV, p. 9-20.

Mangiante, J.-M. & Parpette, C. (2011). Le français sur objectif universitaire. Grenoble : Presses universitaires de Grenoble.

Objet d'études : les marqueurs discursifs

Comparaison de fréquence des marqueurs de reformulation et d'exemplification dans le corpus d'apprenants et dans le corpus TermiTh



Objet d'études : marqueurs de reformulation

- La reformulation fait partie des trois macro-compétences transversales de la production écrite (Mangiante & Parpette, 2011)
- Comparaison avec le corpus d'apprenants :
 - Absence d'organiseurs de texte → organisation confuse d'idées ou fragmentation d'idées (Le, 2012)
 - Absence des marqueurs de reformulation (Tran, 2014)

Le, N. B. 2013. L'utilisation des connecteurs argumentatifs dans les écrits universitaires des étudiants vietnamiens : le cas du discours géographique. Pour une perspective didactique. Thèse de doctorat en Sciences du langage Spécialité Français Langue Etrangère, Université Grenoble .

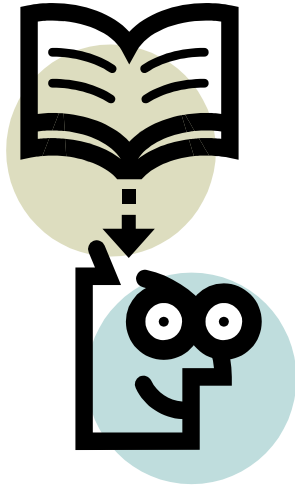
Tran, T.-T.-H. 2014. Développement d'une aide à l'écrit scientifique. Description de la phraséologie scientifique et réflexion didactique pour l'enseignement à des étudiants non natifs. Thèse de doctorat en Sciences du langage Spécialité Français Langue Etrangère, Université Grenoble Alpes.

Exemples de reformulation dans le corpus d'apprenants

1. Parmi là, trois sur cinq d'activités ont lieu le samedi et le dimanche chaque semaine, **c'est-à-dire** on n'a pas beaucoup de choix lors de la visite aux jours quotidiens. [geo-06]
2. Enfin, pour le WWF (2001), l'écotourisme « simple » se mue en « écotourisme communautaire », **à savoir**, selon laquelle la communauté locale contrôle de façon significative son développement et sa gestion tout en y étant impliquée, et dans laquelle une proportion importante des bénéfices reste au sein des communautés. » [geo-09]

4 occurrences *c'est-à-dire que*, 12 occurrences en tête de phrase

Plan



1. Cadre de travail : objectifs et hypothèse
2. Objet d'étude : marqueurs de reformulation et d'exemplification
3. Méthodologie de recherche
4. Expérimentation

Méthodologie

- Corpus de travail :
 - Corpus TermiTh-SHS, environ 5M de mots, composé des articles en 10 disciplines en sciences humaines et sociales (Hatier et al., 2016; Tran, 2014)
 - ✓ Adresse : <http://scientext.msh-alpes.fr/scientext-site/spip.php?article9>
 - ✓ Corpus étiqueté morpho-syntaxiquement (Syntec) et annoté semi-automatiquement.
- Corpus d'apprenants :
 - Mémoires de recherche d'étudiants chinois et vietnamiens (~ 1 000 000 mots)
 - Disciplines en sciences humaines et sociales : économie, géographie, littérature, traduction, linguistique et psychologie
 - Outil d'extractions : Antconc

Méthodologie

ScienQuest – Corpus «Écrits scientifiques en français»

Scientext

Corpus | Textes | Recherche | Critères | Liste

Actuellement votre sélection comporte 5 063 315 mots (205 textes), sur les 5 063 315 mots (205 textes) du corpus.
Cliquez pour voir les détails de la sélection...

Discipline

- Sciences humaines**
 - Linguistique (1 633 579 mots, 67 textes)
 - Psychologie (528 554 mots, 17 textes)
 - Sciences de l'éducation (1 280 515 mots, 60 textes)
 - TAL (753 025 mots, 17 textes)
- Sciences expérimentales**
 - Biologie (744 534 mots, 28 textes)
 - Médecine (97 139 mots, 8 textes)
- Sciences appliquées**
 - Électronique (419 234 mots, 5 textes)
 - Mécanique (139 912 mots, 3 textes)

Tout

Rien

Genre

- Article (348 168 mots, 45 textes)
- Communication (532 333 mots, 112 textes)
- Thèse (4 198 966 mots, 41 textes)
- HDR (517 025 mots, 7 textes)

Tout

Rien

Parties textuelles

- Parties principales**
 - Développement (4 310 279 mots)
 - Introduction (154 732 mots)
 - Conclusion (202 371 mots)
- Autres parties**
 - Résumé (42 330 mots)
 - Notes de bas de page (249 156 mots)
 - Remerciements (22 090 mots)
 - Annexes (71 878 mots)
 - Avant-propos (4 145 mots)
 - Mots-clés (6 334 mots)

Tout

Rien

Méthodologie

Mot 1

Forme

Lemme

Catégorie

Mot 2

Forme

Lemme

Catégorie

+

Ordre des mots: Aucune relation syntaxique n'est définie. Les mots seront recherchés exactement dans cet ordre.

Relations syntaxiques

+

Chercher 487 occurrences

	#	Contexte gauche	Occurrences	Contexte droit	Texte
<input type="checkbox"/>	1	le milieu et comment sont -ils reliés institutionnellement ou personnellement	au savoir	mathématique ? Comment les élèves sont -ils susceptibles de les	#1 - HDR - Sciences de l'éducation - Introduction
<input type="checkbox"/>	2	celles dont la formulation et la validation permettent de donner accès	au savoir	visé . Il s' agit de transformer des " de	#1 - HDR - Sciences de l'éducation - Développement
<input type="checkbox"/>	3	est à la fois épistémologique et cognitive : Épistémologique car elle réfère	au savoir	mathématique ; Cognitive car elle traite de la relation entre action	#1 - HDR - Sciences de l'éducation - Développement
<input type="checkbox"/>	4	envisagée permet -elle effectivement de faire apparaître des connaissances associées	aux savoirs	visés ? " Ces confusions ont rendu nécessaire la reprise	#1 - HDR - Sciences de l'éducation - Développement
<input checked="" type="checkbox"/>	5	Tn tende vers la longueur de la courbe limite ,	à savoir	la base du triangle de départ , il suffit que	#1 - HDR - Sciences de l'éducation - Développement
<input checked="" type="checkbox"/>	6	de cette suite , celui qui dépend de n ,	à savoir	(4/9) n ; ceci implique que les élèves	#1 - HDR - Sciences de l'éducation - Développement
<input checked="" type="checkbox"/>	7	en particulier la difficulté à identifier n comme variable , et	à savoir	faire le lien entre une expression qui est variable et	#1 - HDR - Sciences de l'éducation - Développement
<input checked="" type="checkbox"/>	8	au départ pour trancher entre les deux conjectures pour Pn ,	à savoir	Pn n' a pas de maximum ou Pn a un	#1 - HDR - Sciences de l'éducation - Développement

Analyses linguistiques

Contexte gauche

Contexte droit

p **Marqueur discursif** **q**

Paramètres	Valeurs		
Position syntaxique	initiale	médiane	finale
Portée syntaxique	globale	locale	
Valeurs sémantiques	[] _p MD V.SEMANTIQUE [] _q		

Paillard, D., & Vu, T.-N. (2012). *Inventaire raisonné des marqueurs discursifs du français. Description. Comparaison. Didactique*. Paris : AUF.

Analyses linguistiques

p en somme q

Relier les idées concordantes

p en résumé q

Relier les idées concordantes
ou différentes.

Delphine restera quatre ans et demi dans cette entreprise, d'abord comme ingénieur de recherche puis comme « biochemist » au sein d'une antenne nouvellement créée aux États-Unis, avant de devenir « chargée de recherche prospective » auprès d'une marque de luxe concurrente. [...] **En somme**, cette courte séquence biographique de déambulation professionnelle a semble-t-il eu pour fonction de faire accepter à Delphine un destin auquel elle n'adhérait qu'à moitié et de lui faire goûter [...]. [soc-art-118-body]

Analyses linguistiques

p en somme q

Relier les idées concordantes

p en résumé q

Relier les idées semblables ou différentes.

**Position
syntaxique**

[]_p en somme []_q

[]_p en résumé []_q

**Portée
syntaxique**

$\left(\begin{array}{l} \text{Proposition} \\ \text{Paragraphe} \end{array} \right)_p$ En somme $\left(\begin{array}{l} \text{Proposition} \end{array} \right)_q$

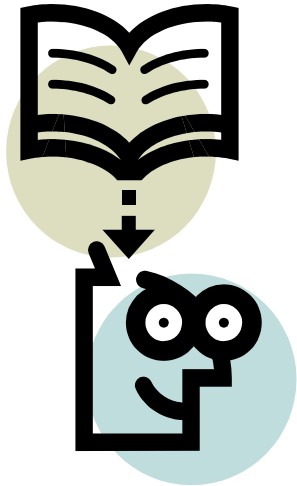
$\left(\begin{array}{l} \text{Proposition} \\ \text{paragraphe} \end{array} \right)_p$ En résumé $\left(\begin{array}{l} \text{Proposition} \\ \text{paragraphe} \end{array} \right)_q$

**Valeurs
sémantiques**

[]_p en somme|V.
RECAPITULATIVE| []_q

[]_p en résumé|V.
EPISTEMIQUE| []_q

Plan



1. Cadre de travail : objectifs et hypothèse
2. Objet d'étude : marqueurs de reformulation et d'exemplification
3. Méthodologie de recherche
4. Expérimentation

Expérimentation : public

- Etudiants allophones DUFLEPES, Université d'Artois, niveau A2-B1.
- Cours de méthodologie des écrits scientifiques, 36h, dont 16h autour des adverbes.
 - Structure d'un écrit scientifique
 - Adverbes de reformulation
 - Introduction du corpus

Expérimentation : approche

Aspect lexico-syntaxique

- Observer le contexte gauche et le contexte droit des MD
- Observer la position des MD

Aspect sémantique

- Faire différencier les différents sens des MD

Expérimentation : exemples

Consigne

1. Observez le contexte gauche et le contexte droit de **c'est-à-dire**. Quelle est leur nature syntaxique ?
2. Définissez la relation entre le contexte gauche et le contexte droit de **c'est-à-dire**.
3. D'après vous, quel est le rôle de **c'est-à-dire** dans ces phrases ? Quelle est la valeur qu'il véhicule ?

Tous les échanges, horizontaux et verticaux, se font de manière asynchrone, **c'est-à-dire que** les participants et le modérateur animateur se connectent à l'heure de leur choix, sans obligation de simultanéité à des moments particuliers. [infcom-art-455-body]

Expérimentation effectuée

Consigne

1. Est-ce que **c'est-à-dire** et **à savoir** sont interchangeables dans ces exemples ? Si non, expliquez quelle est la raison qui vous permet d'utiliser **c'est-à-dire** à la place de **à savoir**, et vice versa ? Vous pouvez vous appuyer sur les critères syntaxiques et sémantiques de chaque marqueur.
2. Quels sont donc les emplois de chaque marqueur dans ces exemples ?

1. Les mêmes musiques que celles utilisées dans l'étude princeps étaient reprises dans notre étude, _____ un extrait provenant de l'Allegro de Mozart (Petite Musique de nuit, Divertimento). [psy-art-246-body]
2. Il permettait également d'associer les deux noyaux historiques de la ville, _____ l'implantation portuaire orientale (parallèle à la Breedstraat) et la partie longtemps demeurée agricole à l'ouest. [his-art-363-body]

Expérimentation : résultats

- Remarques des étudiants :
 - Prise de conscience de l'emploi de certains MD, par exemple le rôle de « c'est-à-dire », même s'ils rencontrent encore des difficultés pour d'autres marqueurs « au total », « en fin de compte », etc.
 - Découverte d'un nouveau genre.
 - Stratégies de lecture

Expérimentation effectuée

- Difficultés et limites :
 - Voir l'autonomie et l'investissement des étudiants à cause de leur faible niveau en langue
 - ✓ Quelques difficultés pour introduire des exercices sur l'aspect sémantique.
 - Question du temps pour des activités de production écrite
 - Développement d'exercices plus dynamiques
 - ✓ Peut-être utiliser quelques outils TICE d'aide à la rédaction.
 - ✓ Difficulté d'évaluer les progrès des étudiants en utilisant le corpus.

Conclusion

- Rôle important de l'enseignant dans le filtrage des exemples
 - prise de conscience du rôle de la part des enseignants.
- Travailler en parallèle les ADV et les autres éléments de la phrase
 - aborder l'aspect rhétorique des écrits scientifiques.
- Nécessité de diversifier les activités et les modalités de travail.

Merci

