



HAL
open science

Le corpus oral ESLO comme ressource didactique pour la formation universitaire en FLE et sciences du langage : l'exemple d'un module sur les mots du discours

Marie Skrovec, Britta Thörle, Chloé Sarah Tahar

► To cite this version:

Marie Skrovec, Britta Thörle, Chloé Sarah Tahar. Le corpus oral ESLO comme ressource didactique pour la formation universitaire en FLE et sciences du langage : l'exemple d'un module sur les mots du discours. 11es Journées internationales de Linguistique de Corpus, Jul 2023, Grenoble, France. pp.325-333. halshs-04478325

HAL Id: halshs-04478325

<https://shs.hal.science/halshs-04478325v1>

Submitted on 26 Feb 2024

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.

Le corpus oral ESLO comme ressource didactique pour la formation universitaire en FLE et sciences du langage : l'exemple d'un module sur les *mots du discours*

Marie Skrovec ¹, Britta Thörle ² et Chloé Tahar ¹

¹ Laboratoire Ligérien de Linguistique, Université d'Orléans

² Département de Romanistik, Université de Siegen

prenom1.nom1@univ-UNIV.fr, prenom2.nom2@univ-UNIV1.fr, prenom3.nom3@univ3-UNIV2.fr

Introduction

Depuis une trentaine d'années, des chercheurs à l'interface entre la linguistique et la didactique des langues se penchent sur la question de savoir comment les corpus de langues peuvent être utilisés d'un point de vue didactique. Si les initiatives dans ce domaine se sont d'abord concentrées sur l'espace anglophone, on constate depuis une dizaine d'années une augmentation des projets concernant les corpus francophones. La présente contribution est issue du projet ESLO-FLEU (ESLO pour le FLE et la Linguistique dans l'Enseignement Universitaire), qui a pour objectif de développer des ressources d'enseignement-apprentissage basées sur des corpus pour le français langue étrangère dans le contexte universitaire. Issus du corpus ESLO (Enquêtes SocioLinguistiques à Orléans), des modules d'enseignement-apprentissage sont élaborés sur différents thèmes à l'interface entre didactique des langues et sciences du langage et visent à développer chez les apprenants une meilleure connaissance du français tel qu'il se parle. Nous présentons la conception d'un module d'enseignement-apprentissage sur les mots du discours (MD)². Si les MD ont été reconnus comme un outil central de la compétence communicative et interactionnelle dans la didactique des langues étrangères, les concepts didactiques et le matériel pour leur enseignement font cependant défaut. Le travail avec des corpus de conversations naturelles, intégrées dans des situations d'interaction authentiques, représente ici une ressource particulièrement précieuse. Dans ce qui suit, le projet sera présenté à l'exemple du module « Le parler jeune » qui traite, entre autres, les MD en français parlé.

État de l'art : corpus oraux et enseignement des mots du discours

Les MD sont un élément central de la compétence discursive telle qu'elle est définie dans le CECR (cf. ch. 5.2.3.1). Ils sont des moyens linguistiques indispensables lorsqu'il s'agit de commencer une conversation, de la maintenir ou de la terminer, de prendre ou de céder le tour

¹ <http://eslo.huma-num.fr/>

² Ce module est construit pour familiariser les apprenants avec un ensemble d'objets actuellement décrits en linguistique sous l'appellation « marqueurs discursifs ». Devant certaines difficultés de catégorisation liées à la nature-même de ces marqueurs (formes polysémiques, ensemble non homogène d'un point de vue syntaxique, etc.), et pour répondre à une nécessaire adaptation terminologique pour un public non expert, le terme « mot du discours », renvoyant aux *Gesprächswörter* de la tradition germanophone (cf. Koch/Oesterreicher 1990), a semblé pertinent pour désigner une diversité de phénomènes incluant entre autres interjections, connecteurs et marqueurs de l'interaction.

de parole, d'introduire ou de conclure un sujet ou de s'assurer de la compréhension mutuelle. Alors que la pertinence des moyens linguistiques – et particulièrement des connecteurs – pour la cohérence et de la cohésion du texte est mise en évidence de manière relativement concrète dans le CECR, le rôle des MD pour les autres aspects de la compétence discursive reste vague (cf. CECR 5.2.3.1). Les études sur l'acquisition des MD en langue étrangère montrent toutefois que les apprenants disposent d'un répertoire extrêmement limité de MD, qui ne se différencie formellement et fonctionnellement qu'à des niveaux très avancés (Borreguero Zuloaga/Thörle 2016). Les apprenants sont ainsi privés d'un prérequis important pour une participation compétente à la conversation. D'où la nécessité d'une prise en compte plus systématique des corpus de conversations naturelles en interaction authentique dans la didactique des langues, qui permettent l'observation des fonctions pragmatiques et sociales des MD et contribuent – contrairement aux purs inventaires de connecteurs – à une prise de conscience de leur fonctionnement (cf. aussi Beeching 2014).

Corpus et méthodologie

De précédentes études sur le corpus ESLO ont observé une fréquence importante de l'usage des MD dans les enregistrements impliquant des locuteurs jeunes, ceux-ci privilégiant par ailleurs leurs emplois les plus pragmatiques (voir Skrovec et al. 2022 sur *après*, Abouda 2022 sur *du coup*). C'est ce constat qui a guidé le choix de constituer un module de formation sur les MD à partir d'une sélection d'extraits sélectionnés selon l'âge des locuteurs.

Le module « Le parler jeune », composé de 20 extraits issus de situations informelles, implique majoritairement des locuteurs de la tranche d'âge 15/25 ans. Les interactions se caractérisent par une interactivité très élevée et une absence de préparation. Ce sous-corpus, de 10 000 mots environ et annoté sous TXM, a tout d'abord fait l'objet d'un repérage manuel, non exhaustif mais large, de formes susceptibles de recevoir une analyse linguistique comme marqueurs discursifs. On y dénombre 2125 occurrences pour 48 types de MD. Une annotation plus précise a ensuite été menée sur 23 MD pour spécifier ces formes d'un point de vue sémantico-pragmatique, selon 4 grands ensembles fonctionnels :

- le domaine *Interaction* regroupe les marqueurs relatifs à la structure de l'interaction et à la co-construction des tours (*alors* en ouverture de tour, *hm* de réception), la gestion intersubjective (*mais* de désaccord) et à la gestion des connaissances partagées (*hein* de sollicitation de l'approbation, *voilà* de ratification) ;
- le domaine *Formulation* comprend les MD associés au travail de formulation et de gestion de la progression discursive, au sens de Güllich (1993), comme les procédures de correction (*enfin*, *quoi*), d'hésitation (*euuh*), d'approximation (*genre*), d'atténuation (*quand même*), de proposition de formulation (*voilà*), de précision (*mais*), etc.
- le domaine *Connexion* recense les emplois comme connecteurs logico-argumentatifs, pouvant porter sur du contenu propositionnel autant que sur des éléments contextuels explicites ou implicites, inférables de la situation ou des savoirs partagés, comme les emplois consécutifs de *donc* ou *du coup*, les emplois contrastifs de *mais* ou *quand même*, etc.
- le domaine *Information* identifie les emplois liés à la gestion des topics conversationnels, comme les procédés d'activation de nouveau topic (*alors*), de développement topical (*du coup*), de parenthèse (*bon*), ou de clôture de topics (*voilà*, *bref*).

Ces annotations sont disponibles au format TXM pour les utilisateurs experts ou au format html pour une utilisation en classe, avec fonction afficher/masquer. Ainsi, il est possible de prévoir une phase de découverte du document sonore transcrit sans les annotations, puis d'afficher

l'ensemble des MD repérés et leurs annotations. Cela doit permettre de guider les apprenants dans l'observation des formes et la formulation d'hypothèses sur leurs différentes fonctions pragmatiques. Un guide d'annotation, à destination des utilisateurs experts ou semi-experts, est en cours d'élaboration.

Application en cours de linguistique française à l'étranger

Un prototype du module "Parler jeune" a été utilisé à l'université de Siegen dans un séminaire de linguistique française pour les étudiants de licence avancés (niveau Bachelor en Allemagne). Dans ce cours, les participants travaillaient avec les extraits sonores transcrits, mais pas encore annotés, étant donné que l'annotation était encore en cours à l'époque. Le groupe cible du séminaire était en grande partie constitué d'étudiants en formation de futurs professeurs de FLE. Sur la base des résultats de travaux de recherche antérieurs sur l'acquisition du français par des apprenants avancés (Bartning/Kirchmeyer 2003, Koch/Thörle 2021), on peut supposer que ce groupe d'étudiants n'a acquis l'usage des MD que partiellement et ne dispose pas de connaissances métalinguistiques systématiques sur leur rôle en français parlé. Les objectifs du module étaient donc : 1. de sensibiliser les apprenants au phénomène des MD, 2. de leur faire connaître les MD en tant que classe fonctionnelle d'expressions afin qu'ils puissent identifier les MD dans le contexte de la conversation et de les distinguer d'autres classes de mots, 3. de leur faire reconnaître dans le corpus les fonctions décrites dans la littérature de recherche et, le cas échéant, de formuler eux-mêmes des hypothèses sur les fonctions pragmatiques dans un contexte d'interaction concret. Pour une unité d'enseignement de deux séances de séminaire, les MD *bon*, *enfin*, *quoi* ont été choisis en raison de leur fréquence relativement élevée. De plus, on peut supposer que les expressions sont familières aux apprenants, mais pas forcément connues en tant que MD. Les activités pédagogiques sont résumées dans la fiche pédagogique ci-dessous (Figure 1).

Résultats provisoires et perspectives

Si ce travail d'annotation n'a pas été mené en vue d'une analyse quantitative, il est néanmoins intéressant de se pencher sur le relevé du nombre des occurrences de chaque marqueur, et les domaines fonctionnels qu'ils occupent (Table 1).

MD	Occurrences	MD	Occurrences
euh	448	du coup	14
mais	124	et tout	14
donc	84	genre	11
quoi	63	tout ça	9
enfin	62	attends	4
en fait	32	en même temps	3
voilà	31	alors là	2
bon	24	au final	2
après	17	bref	2
quand même	16	d'ailleurs	2
alors	14	en effet	2

Table 1 : Occurrences des MD annotés dans le Module "Parler jeune"

Ces résultats peuvent permettre d'orienter les choix didactiques des enseignants : ainsi, il semble pertinent d'attirer l'attention des apprenants sur des marqueurs fréquents et particulièrement polyfonctionnels comme *mais*, *donc*, *quoi*, *enfin*, *en fait*, *voilà*, conformément à l'expérimentation menée à Siegen qui s'est focalisée sur *enfin*, *quoi* et *bon*.

L'annotation sémantico-pragmatique permet d'orienter l'observation en pointant les MD en contexte pour suggérer une catégorisation de ces derniers et mettre en évidence leur caractère polyfonctionnel, comme dans le cas de *quoi* illustré ci-dessous (table 2).

Information	Clôture_topic		
Connexion			
Interaction	Emphase		
Formulation	Approximation	Précision	Proposition_formulation

Table 2 : Domaines fonctionnels et jeu d'étiquettes pour le MD *quoi*

En ce qui concerne l'application en cours de linguistique, le travail sur corpus a été généralement apprécié par les étudiants. L'application du Module ESLO dans deux séminaires à l'université de Siegen qui a été évaluée à l'aide d'un questionnaire. Les réponses des dix participants à l'enquête montrent que l'authenticité des données, le support sonore et la contextualisation situative et sociale ont été évalués positivement, comme en témoigne les remarques suivantes :

« Ce qui m'a particulièrement plu était qu'on a pu travailler avec des enregistrements « authentiques ». C'était intéressant de faire des analyses par rapport aux marqueurs du discours qui peuvent avoir plusieurs sens. »

« Ce qui m'a particulièrement plu, c'est le langage authentique et sa représentation en écrit (les transcriptions). Les transcriptions ESLO sont très faciles à lire et comprendre. »

En guise de conclusion, nous proposons l'aperçu d'un "livrable" de notre expérience avec le prototype des activités pédagogiques applicables dans des séances de cours universitaires en linguistique française avec un public de FLE (Figure 1). À la différence du cours décrit plus haut et pour réduire la complexité, les activités proposées dans la fiche sont centrées sur les MD *enfin* et *quoi*.

Fiche pédagogique ESLO-FLEU :
***enfin* et *quoi* comme mots du discours dans le module « Parler jeune »**

Public cible :	étudiant.e.s d'un cours de linguistique française en Allemagne (troisième année d'études)
Corpus/Module/Extrait :	module ESLO-FLEU « Parler jeune » ; extraits « Guitare », « Les autres »
Références :	Bertrand, R./Chanet, C. (2005): Fonctions pragmatiques et prosodie de <i>enfin</i> en français spontané. <i>Revue de sémantique et pragmatique</i> 17, 41-68. Dostie, G./Pusch, C.D. (2007): Introduction. Les marqueurs discursifs. Sens et variation ». <i>Langue française</i> 154/2, 3-12. Hölker, K. (2010): Frz. 'quoi' als Diskursmarker". <i>Linguistik online</i> 44. < https://bop.unibe.ch/linguis.k-online/ar.cle/view/405 >

Objectifs :

Les étudiant.e.s

- sont sensibilisé.e.s au phénomène des mots du discours (MD)
- connaissent les MD en tant que classe fonctionnelle d'expressions. Il.elle.s savent identifier les MD dans le contexte de la conversation et les distinguer d'autres classes de mots
- se familiarisent particulièrement avec les MD *enfin* et *quoi*
- appliquent de manière critique des travaux de recherche exemplaires sur *enfin* et *quoi* au corpus
- formulent eux-mêmes des hypothèses sur les fonctions pragmatiques de ces MD dans un contexte d'interaction concret.

Conditions préliminaires :

- Les étudiant.e.s ont été familiarisé.e.s avec la notion des « mots du discours ». Il.elle.s connaissent quelques particularités généralement attribuées aux MD (Dostie/Pusch 2007).
- Il.elle.s ont à leur disposition les articles de Bertrand/Chanet (2005) et Hölker (2010) qui ont été discutés en classe.

Tâches :

Compréhension de l'extrait et identification des MD :

- Écoutez les extraits « Guitare » et « Les autres » et lisez les transcriptions. Décrivez brièvement la situation et le thème des conversations.
- Soulignez tous les *enfin* et *quoi* utilisés comme MD dans l'extrait. Justifiez pourquoi il s'agit de MD.
- Donnez un exemple où *quoi* n'est pas un MD.

Observations concernant la forme et la position :

- Écoutez de nouveau l'extrait et faites attention à la prononciation de *enfin*. Comment ce MD est-il réalisé ?
- Essayez de déterminer la portée des MD dans deux occurrences de *enfin* et deux occurrences de *quoi* de votre choix. Qu'observez-vous quant à la position de *enfin* et *quoi* ?

Attribution de fonctions pragmatiques et combinaisons avec d'autres MD :

- Parmi les occurrences de *enfin* et *quoi*, lesquelles correspondent aux descriptions de Bertrand/Chanet et Hölker ?
- Avec quels autres MD *enfin* et *quoi* apparaissent-ils en combinaison ?
- Regardez d'autres extraits du module « Parler jeune ». Y a-t-il des occurrences de *enfin* et *quoi* qui ne se laissent pas/se laissent difficilement décrire par les théories proposées par ces auteurs ? Faites des hypothèses sur leurs fonctions.

Résultats possibles :

- Les étudiant.e.s écoutent deux entretiens avec deux jeunes femmes. Dans « Guitare » une jeune vendeuse de 25/35 ans parle de sa pratique de la musique. Dans « Les autres », une étudiante de 15/25 ans au diplôme d'éducateurs de jeunes enfants expose les questions de son futur concours.
- Les étudiant.e.s trouvent une dizaine d'occurrences de *enfin* et *quoi* dont la plupart sont des MD (à quelques exceptions près comme, par exemple, *je saurais pas quoi répondre comme ça*).
- Il.elle.s observent que *enfin* est souvent réalisé de manière tronquée (*fin*). Dans la majorité des occurrences, *enfin* indique une reformulation et introduit l'élément reformulateur. À l'aide de l'aperçu dressé par Bertrand/Chanet (2005), il.elle.s pourront lui attribuer une valeur « corrective » ou « de synthèse, conclusive ».
- Les étudiant.e.s constatent que *quoi*, à différence de *enfin*, se trouve toujours en position finale, mais remplit aussi une fonction reformulatrice dans les extraits, indiquant, entre autres choses, une conclusion inférable par les informations du co-texte, le résumé de ce qui a été dit auparavant ou la clôture d'une explication. Les occurrences se laissent décrire par la fonction de base « marqueur de la répétition référentielle » décrite par Hölker (2010).
- Les étudiant.e.s trouveront dans les deux extraits des combinaisons fréquentes de MD comme *enfin bon, enfin je veux dire, enfin ... quoi*.
- Si la plupart des exemples peuvent être décrits à l'aide des approches mentionnées, il y a quelques occurrences pour lesquelles les étudiant.e.s pourraient rencontrer des difficultés à attribuer une des fonctions établies et où ils doivent formuler leurs propres hypothèses sur la valeur pragmatique de ces MD.

Figure 1 : Fiche pédagogique : *enfin* et *quoi* comme mots du discours dans le module « Parler jeune »

Références

Abouda, Lotfi (2022). « L'émergence du marqueur méta-discursif *du coup* : de la conséquence à l'actualisation énonciative », *Langages*, vol. 226, no. 2, 99-116.

Abouda, Lotfi & Skrovec, Marie. (2018). "Micro-diachronie d'un marqueur discursif. *En même temps* : simultanéité, coexistence, adversativité". *DIA 3 : Le français innovant*, Mar 2018, Paris, France. 245-270. [_halshs-03023109](https://halshs-03023109)

Bartning, Inge, Kirchmeyer, Nathalie (2003). « Le développement de la compétence textuelle à travers les stades acquisitionnels en français L2 », *Acquisition et Interaction en Langue Étrangère* 19, 9–39. <https://doi.org/10.4000/aile.1112>

Beeching, Kate (2014). « Corpora in language teaching and learning », *Recherches en didactique des langues et des cultures*, 11-1.

Bertrand, Roxane, Chanet, Catherine (2005). « Fonctions pragmatiques et prosodie de *enfin* en français spontané », *Revue de sémantique et pragmatique* 17, 41-68.

Borreguero Zuloaga, Margarita, Thörle, Britta (2016), « Introduction. Discourse Markers in Second Language Acquisition: Studies on Italian and French as L2 », *Language, Interaction and Acquisition* 7:1, 1-16.

CECRL, Conseil de l'Europe (2001), *Cadre européen commun de référence. Apprendre, enseigner, évaluer*.

Dostie, Gaétane (2004). *Pragmaticalisation et marqueurs discursifs. Analyse sémantique et traitement lexicographique*. De Boeck Supérieur, 2004.

Dostie, Gaetane, Pusch, Claus D. (2007), « Introduction. Les marqueurs discursifs. Sens et variation ». *Langue française* 154/2, 3–12.

Gulich, Elisabeth (1993). « Procédés de formulation et travail conversationnel : éléments d'une théorie des processus de la production discursive ». In : Hilty G, ed. 2. *Actes du XXe Congrès International de Linguistique et Philologie Romanes. Section II - analyse de la conversation, Section III - la fragmentation linguistique de la Roumanie*. Actes du XXe Congrès International de Linguistique et Philologie Romanes. Vol 2. Tübingen: Niemeyer; 1993: 137-151.

Hölker, Klaus (2010). « Frz. 'quoi' als Diskursmarker », *Linguistik online* 44.

Koch, Christian, Thörle, Britta (2021). « Metadiscursive Activities in Oral Discourse Production in French L2: A Study on Learner Profiles », *Corpus Pragmatics* 5/1, 153-186.

Koch, Peter, Oesterreicher, Wulf (1990), *Gesprochene Sprache in der Roumanie : Französisch, Italienisch, Spanisch*, Tübingen, Niemeyer.

Skrovec, Marie, Loyal Kanaan-Caillol, et Hisae Akihiro (2022). « Le marqueur après à l'oral : une approche micro-diachronique, variationniste et interactionnelle », *Langages*, vol. 226, no. 2, 2022, 117-131.