

Les fonctions discursives du Lexique Scientifique Transdisciplinaire pour aider à la compréhension et à la production d'écrits académiques

The discursive functions of the Transdisciplinary Scientific Lexicon to aid in the understanding and production of academic writing

Jingrao Li

Beijing International Studies University

jingraoli2@163.com

<https://orcid.org/0000-0001-8220-2068>

Cristelle Cavalla

Université Sorbonne Nouvelle, DILTEC

cristelle.cavalla@sorbonne-nouvelle.fr

<https://orcid.org/0000-0002-8599-8521>

Abstract

The Transdisciplinary Scientific Lexicon (TSL) is a lexicon that transcends scientific disciplines (e.g., *hypothesis, advance an argument, the objective of this study is to determine...*). Mastering these elements is crucial because they fulfill essential discursive functions (DF) in the scientific genre, serving reasoning, structuring, and argumentation purposes. In this article, we propose using DFs as a bridge that links understanding and production of the TSL. To achieve this, a scientific writing training program was offered to learners, utilizing a corpus-based teaching approach and a lexical tool for scientific writing developed by our team. We will highlight the benefits of mastering DFs in terms of helping learners to understand and produce scientific texts, as well as the challenges learners face in using DFs to formulate ideas in writing.

Keywords: Transdisciplinary Scientific Lexicon, discursive functions, scientific writing, Chinese-speaking learners

1. INTRODUCTION

Dans les écrits scientifiques, il existe un ensemble d'éléments lexicaux qui ne sont pas spécifiques à un domaine scientifique particulier, mais sont partagés par plusieurs disciplines scientifiques (*analyse, à ce jour, l'objectif de cette étude est de déterminer...*). Ces éléments lexicaux participent à la construction du sens scientifique, à la structuration et à l'argumentation, et sont désignés par le « Lexique Scientifique Transdisciplinaire (LST) » (Jacques & Tutin, 2018 ; Li & Tutin, 2023). Ce LST, très présent dans les écrits scientifiques, joue un rôle important dans la compréhension du fonctionnement discursif du genre scientifique (Tran, 2014 ; Cavalla, 2018), et remplit des fonctions discursives (FD) essentielles dans ce type d'écrits. Par exemple, *faire une hypothèse* sert à présenter une hypothèse de l'étude, comme dans la phrase *on peut faire l'hypothèse que*.

L'objectif de notre recherche est l'aide à la rédaction scientifique auprès d'étudiants allophones du français. Cet article s'oriente vers des liens entre le LST, l'utilisation des corpus, la compréhension et la production des écrits académiques par des étudiants étrangers en Français Langue Étrangère (FLE). Il s'agit de l'analyse de l'utilisation des FD par les apprenants lors d'une séquence pédagogique pour les aider dans leur production écrite scientifique¹.

Nous utilisons une base de données existante : le LST, développé au sein du laboratoire LIDILEM. Cette ressource lexicale a été constituée à partir d'un corpus d'articles scientifiques dans des domaines des sciences humaines et sociales (SHS) – le TermITH-Transdisciplinaire, grâce à un travail d'équipe (Jacques & Tutin, 2018 ; Tran, 2014 ; Hatier, 2016 ; Hatier *et al.*, 2016 ; Yan, 2017 ; Ji, 2022 ; Wang, 2022 ; Li, 2023). À partir de ce corpus et pour atteindre l'objectif énoncé précédemment, nos questions de recherche sont les suivantes : Comment enseigner les FD du LST pour aider à la compréhension d'un écrit scientifique ? Dans quelle mesure la maîtrise des FD peut-elle influencer la rédaction des écrits académiques ? Afin de répondre à ces questions, nous présentons alors, dans un premier temps, les types d'unités lexicales du LST, la base de données que nous utilisons, ainsi que les FD que ces éléments remplissent dans le genre scientifique. Dans un second temps, nous exposons la formation offerte aux apprenants pour rédiger leur mémoire, en recourant aux FD comme entrée dans les caractéristiques du genre scientifique. Dans un troisième temps, nous montrons l'influence de la compréhension des FD lors de la production écrite des apprenants.

¹ Pour davantage d'informations sur le cadre théorique de ce travail, nous invitons les lecteurs à se reporter aux différents articles de l'équipe cités ici et à la thèse de Li (à paraître).

2. LE LEXIQUE SCIENTIFIQUE TRANSDISCIPLINAIRE ET LES FONCTIONS DISCURSIVES

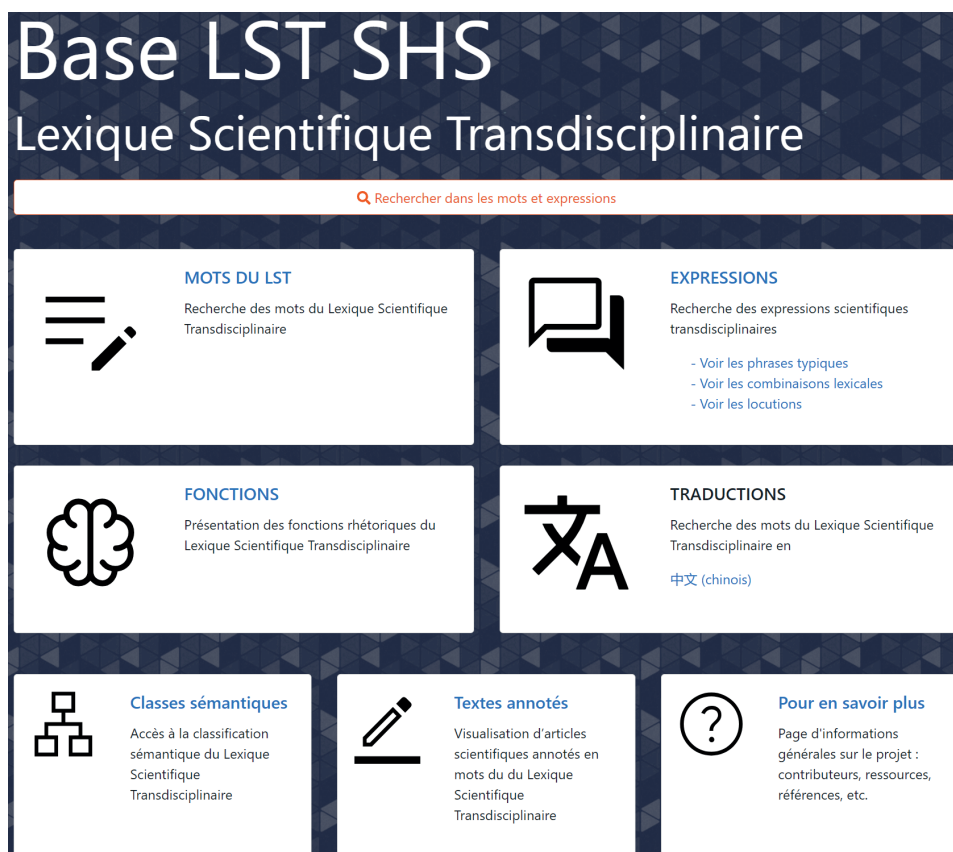
2.1. LA BASE DE DONNÉES LST-SHS

Le LST comprend des mots simples (*analyser, constituer*), ainsi que des éléments phraséologiques, comme les locutions/expressions figées (*à ce sujet, mettre au jour*), les collocations/combinaisons préférées des mots (*aborder un sujet, adopter une approche*) et les routines sémantico-rhétoriques/phrases typiques préconstruites (*Le but de cet article est de...*) (Luodonpää-Manni *et al.*, 2022).

La base de données du LST, que nous utilisons, comporte 1871 acceptions de mots simples et locutions (noms, verbes, adjectifs et adverbes), environ 1500 collocations, reliées aux mots simples et locutions, et un ensemble de routines sémantico-rhétoriques à fonction de guidage du lecteur (*Dans ce qui suit, nous allons tenter de...*), et de routines d'attitude (*On s'étonne que...*) (Ji, 2022 ; Wang, 2022). Les mots simples et les locutions sont accompagnés des informations lexicales suivantes : glose, catégorie grammaticale, construction syntaxique (pour les verbes), exemples du corpus, expressions/collocations, emploi par discipline, champ sémantique, les statistiques (fréquence, spécificité...) et la traduction (pour l'instant, seulement la traduction en chinois est disponible). Les collocations et les routines sont caractérisées par le patron syntaxique et les FD qu'elles remplissent dans les écrits scientifiques.

Pour faciliter l'exploitation de la base de données du LST par les linguistes, les techniciens de traitement automatique des langues, les didacticiens et les apprenants, une plateforme d'exploitation² a été conçue grâce à un travail d'équipe, ce que nous appelons « l'outil lexical LST-SHS ». Cet outil fournit plusieurs accès sémasiologiques (de la forme au sens), et accès onomasiologiques (du sens à la forme) pour interroger les données, y compris les accès par mots simples, par expressions, par fonctions discursives, par la traduction en chinois et par une barre de recherche centrale qui permet d'exploiter toute la base de données (Figure 1) :

² La plateforme est accessible sur : <http://lst.univ-grenoble-alpes.fr/>

Figure 1. Base de données LST-SHS¹

Dans cet article, nous nous concentrons sur les FD du LST.

2.2. FD DU LST ET LA LISTE DE FONCTIONS DANS LA BASE

Les éléments du LST ont une fonction discursive et donc sémantique³. Nous nous intéressons ici prioritairement aux éléments préconstruits phraséologiques présentés précédemment. Ces éléments préconstruits portent des fonctions discursives spécifiques qui sont à connaître ou reconnaître (par les apprenants) pour ensuite être capable de les insérer dans les discours scientifiques (surtout écrits pour cette recherche).

Dans les écrits scientifiques, la FD concerne tous les éléments langagiers nécessaires pour construire et transmettre les connaissances scientifiques de manière claire,

³ Pour davantage d'informations quant aux choix linguistiques pour cette étude, cf. la thèse et les articles de Jingrao Li ainsi que les articles de l'équipe du LIDILEM : Agnès Tutin et Sylvain Hatier notamment.

précise, cohérente, efficace et impartiale (Swales, 1990). Les éléments remplissant une ou des FD particulières peuvent apparaître dans toutes les parties du discours scientifique, allant de la revue de littérature à l'annonce des résultats, aux conclusions, en passant par la présentation des méthodes utilisées, l'analyse et l'interprétation des données.

À notre connaissance, peu d'études ont explicité la définition du concept de « fonction discursive ». Il n'en existe ni d'ensemble fixe ni de liste exhaustive. Les propositions de classification varient à plusieurs égards selon le corpus utilisé et des objectifs visés (Alonso-Ramos & García Salido, 2019). Notre étude s'appuie sur une liste de FD élaborée dans le cadre de la thèse de Li (à paraître), dans une perspective didactique des éléments du LST. Cette liste a été incorporée dans la base de données LST-SHS⁴ (exemple de 4 fonctions dans la Figure 2).

| | |
|--|--|
| Pour définir le thème essentiel de la recherche (sujet d'étude, objectif, hypothèse, problématique) ▼ | <i>L'objet/sujet d'étude concerne qc</i> |
| <p>Annoncer l'objet/sujet ? L'objet/sujet d'étude concerne qc</p> <p>Annoncer l'objectif ? Le premier objectif est de + Vinf</p> <p>Présenter les hypothèses ? Notre première hypothèse considère que...</p> <p>Présenter la problématique ? Qc (l'article/l'analyse...) est consacré à l'étude/ la question de qc</p> | |
| Présenter les méthodes, analyses, résultats, interprétations, difficultés et limitations ▼ | <i>Qn adopte une approche de...</i> |
| Pour insister sur un sujet, mettre en évidence un élément, ou attirer l'attention sur un aspect saillant ▼ | <i>On peut souligner...</i> |
| Pour évaluer un objet (évaluation/jugement, importance, originalité, singularité) ▼ | <i>Qn + V+ à juste titre+ l'objet</i> |

Figure 2. FD dans la base LST-SHS²

La base contient 12 FD avec 47 sous-fonctions. Il existe : 1) des fonctions orientées vers la recherche scientifique (*Définir le thème essentiel de la recherche, Présenter les méthodes*) ; 2) des fonctions orientées vers le texte (*Indiquer la structure du texte*) ; et 3) des fonctions rhétoriques orientées vers le lecteur visant à se positionner ou à persuader (*Insister sur un sujet*).

C'est en élaborant cette liste que nos questions de recherche ont émergées et ont été renforcées par les deux hypothèses suivantes : la première est que la sensibilisation à l'existence de ces éléments préconstruits, et à leurs FD, aident les apprenants à comprendre le fonctionnement discursif de ces écrits. La deuxième est que les FD servent de pont entre l'exposition à ces éléments préconstruits pendant la lecture et leur production dans la phase d'écriture. Pour valider nos hypothèses, nous avons dispensé des cours à l'aide des corpus numériques.

⁴ La liste complète des FD est accessible sur : <http://lst.univ-grenoble-alpes.fr/fonctions>.

3. FORMATION OFFERTE

Cette formation avait un double objectif : 1) familiariser les apprenants au genre scientifique, y compris la structure, les FD et le LST (présenté dans cet article) ; 2) tester l'utilité de l'outil lexical LST-SHS afin de l'améliorer. Le processus d'enseignement et d'apprentissage choisi peut être schématisé comme suit (Figure 3) :

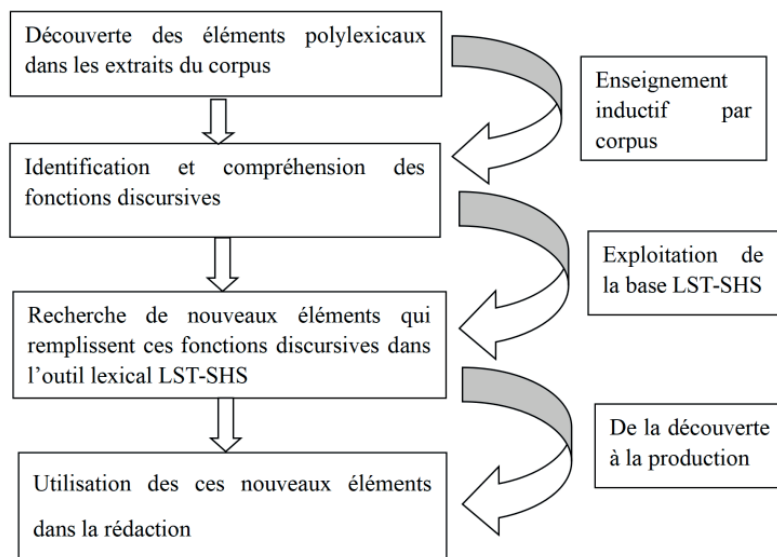


Figure 3. Processus d'enseignement et d'apprentissage de la formation

Avant de poursuivre, voici la présentation du contexte d'enseignement et d'apprentissage.

3.1. PARTICIPANTS

Les participants à l'étude ont été contactés par courrier électronique afin de leur proposer une assistance dans la rédaction de leur mémoire. Le public prioritaire était les apprenants sinophones en M2, étant donné que la base LST-SHS intègre un accès par la traduction en chinois. Ces apprenants, contraints par le temps de soumission du mémoire, devaient accepter de suivre la formation non créditée. Nous avons admis quelques apprenants rédigeant un rapport de stage, en supposant qu'ils puissent également acquérir des expressions utiles pour ces rapports.

Nous avons accueilli neuf apprenants sinophones en M1 et M2 de linguistique, industrie de la langue et communication à l'université Grenoble Alpes. Leur niveau de

français est B2 et supérieur selon l'échelle du CECRL (2001). Trois participants ont abandonné la formation pour des raisons personnelles. Pour garantir l'anonymat dans les analyses, nous avons attribué à ceux qui ont suivi le protocole entier des numéros allant de 1 à 6.

3.2. ACTIVITÉS PROPOSÉES

La formation, d'une durée totale de 12 heures, était répartie en six séances de deux heures chacune. La première séance a été consacrée à une introduction générale à l'écrit scientifique et aux LST. Les cinq suivantes ont porté sur les éléments remplissant des fonctions discursives propres à chaque partie du mémoire : introduction, cadrage théorique, méthodologie, résultats et analyses, conclusion. La figure 4 présente les fonctions que nous avons abordées dans les cinq parties du mémoire. Les noms de ces fonctions diffèrent de la base car au cours nous avons choisi de les nommer de façon longue afin de bien les expliquer.

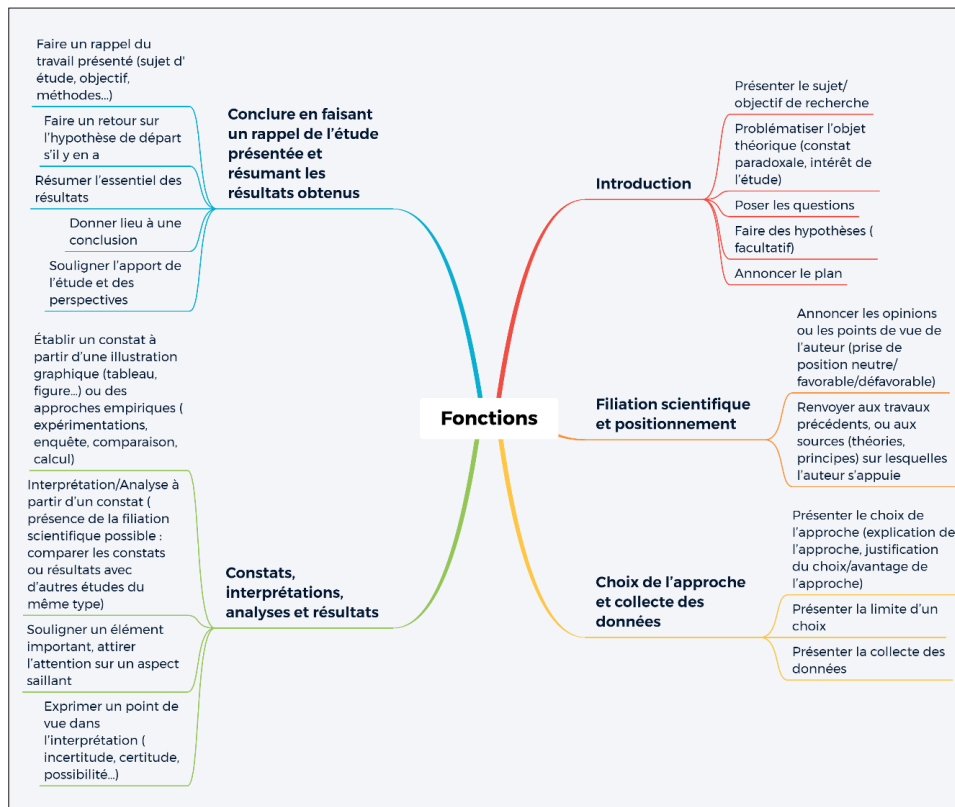


Figure 4. Les FD étudiées pendant la formation

Les activités proposées en classe consistaient en trois étapes : 1) la découverte du LST et des FD à partir des extraits du corpus, 2) l'approfondissement des connaissances à travers des exercices, et 3) le réemploi des connaissances dans une production libre. L'organisation de la formation a été inspirée des séquences didactiques de Yan (2017). Ces séquences ont été adaptées pour inclure des exercices qui impliquent l'exploitation de la base LST-SHS pour aider à la production écrite. Afin de familiariser les apprenants aux FD, quatre types d'exercices ont été proposés : association des expressions aux fonctions correspondantes, repérage, reformulation et regroupement. Voici la présentation de chaque type d'exercices avec leur justification pédagogique.

3.2.1. EXERCICE D'ASSOCIATION AUX FONCTIONS DISCURSIVES

Cet exercice a été présenté aux apprenants lors de la première séance, avant l'enseignement des FD. Les étiquettes correspondaient exactement à celles présentées dans la base LST-SHS (donc plus courtes que celles présentées ensuite). L'objectif d'un tel exercice est de voir si ces FD sont intuitives chez les apprenants (Figure 5).

1.2 Reliez les expressions soulignées et les fonctions suivantes

1. Ces deux analyses ainsi que les résultats obtenus font l'objet de cet article.
2. dans l'exemple (16) cité ci-dessus,
3. Nous nous sommes demandé si les projets professionnels...
4. Trois recueils de données ont été menés.
5. En d'autres termes,
6. On peut effectivement remarquer comme cet auteur que...
7. Elles ont, en effet, globalement des projets professionnels moins ambitieux qu'eux

- A. Présenter la problématique
- B. Présenter la méthode
- C. Reformuler avec d'autres mots
- D. Annoncer l'objet de l'étude
- E. Présenter un constat
- F. Nuancer un propos, énoncer une opinion qui semble paradoxale
- G. Renvoyer à une autre partie du texte

Figure 5. Exercice : relier les expressions aux FD

3.2.2. EXERCICE DE REPÉRAGE

L'objectif de ce type d'exercice est de vérifier si les apprenants ont saisi les explications sur les objets linguistiques visés et sont alors capables de les repérer dans des extraits. Les apprenants devaient repérer des expressions à FD spécifiques dans plusieurs extraits. La figure 6 présente un exemple tiré de la séance dédiée à la méthodologie. Les étudiants ont travaillé en binôme et confronté leurs réponses avec celles d'un autre groupe. Ensuite, l'enseignant a mis en évidence les expressions pertinentes et a souligné leur rôle dans le fonctionnement global du texte.

II. Analyse des phénomènes linguistiques liés à la méthodologie (30m)

La partie méthodologique peut être divisée en trois champs sémantiques : le choix de l'approche (explication de l'approche, justification du choix/avantage de l'approche), la limite d'un choix et la collecte des données.

Consigne : Relevez les expressions qui portent sur le choix de la méthodologie.

Extrait 1 :

Pour mener à bien cet objectif, nous nous sommes appuyés sur un dispositif unique, l'élément "dessine avec ton doigt", présenté dans l'exposition des 3/5 ans de la Cité des enfants, à la Cité des sciences et de l'industrie. Par ailleurs, nous avons privilégié une approche qualitative basée sur une démarche d'analyse approfondie, portant par conséquent sur un échantillon relativement restreint de participants.

[sed-the-69-body] – L'impact cognitif des dispositifs médiatiques sur les enfants d'âge préscolaire en situation d'apprentissage avec un adulte (BERNARD)

Consigne : Relevez les expressions qui portent sur la limite d'un choix.

Extrait 2 :

Même si cette étude présente des limites dues aux contingences de terrain et nécessite d'être répliquée avec d'autres échantillons, elle est néanmoins la première à avoir eu recours à l'analyse en cluster dans le but d'identifier les profils motivationnels des individus dans le domaine du travail. Cette recherche nous a donné des informations importantes sur les différents profils motivationnels des salariés d'une entreprise privée française ainsi que sur les conséquences qui y sont associées. (Texte_psych_10)

Consigne : Relevez les expressions qui portent sur la collecte et la présentation des données.

Extrait 3 :

Afin de recueillir les requêtes [3] formulées par les élèves, nous avons procédé à l'enregistrement des adresses web (URL[4]) des moteurs de recherche dans les historiques des navigateurs (exclusivement Internet Explorer) utilisés pendant la séance durant laquelle les élèves avaient fourni un travail autonome de recherche sur Internet. Nous avons ainsi recueilli les historiques de 9 sessions, c'est-à-dire le travail de 18 élèves, étant donné qu'ils travaillaient en binômes. À partir de ces historiques nous avons retenu 95 requêtes pour notre analyse [5].

[sed-the-356-body] - Travail collectif des enseignants et pratiques d'enseignement : le cas de la prise en charge des élèves dits en difficulté au sein de l'école primaire (Lefevre)

Figure 6. Exemple d'exercice de repérage

3.2.3. EXERCICE DE REFORMULATION

Dans cet exercice, les apprenants reformulaient la phrase avec les expressions ayant une même FD, à l'aide de l'outil LST-SHS (Figure 7). Ils devaient d'abord identifier la FD à laquelle correspond l'expression soulignée et après chercher d'autres expressions qui remplissent cette fonction dans l'outil, pour reformuler la partie soulignée. L'objectif ici est double : vérifier et renforcer leur capacité à repérer les FD et, surtout, avancer vers leur autonomie à utiliser l'outil et à faire des choix lexicaux quand cela est possible.

2.2.3 : Reformulez les parties soulignées avec formules ayant une même fonction à travers l'accès par fonction de l'outil (variations syntaxiques permises)

Ex : Le terme de " variations " désigne un ensemble de phénomènes de nature hétérogène, repérables dans la forme des énoncés et qui ne se confondent pas avec les valeurs du mot.

Fonction correspondante : Pour définir les termes - Donner une définition

Reformulation : Le terme de " variations " renvoie à un ensemble de phénomènes...

- 1) Prenons l'exemple des étudiants arrivant d'Europe de l'Est qui sont assez mobiles et qui, pourtant, rencontrent de grandes difficultés d'argent : ils sont plus nombreux à citer les difficultés financières que les Africains (52 %).

Fonction correspondante :

Reformulation :

- 2) Certains auteurs défendent la thèse selon laquelle la préposition est désémanisée lorsqu'elle introduit un argument du verbe : elle ne serait que le support fonctionnel de l'introduction d'un argument du verbe.

Fonction correspondante :

Reformulation :

Figure 7. Exemple d'exercice de reformulation

3.2.4. EXERCICE DE CLASSEMENT

Cet exercice vise à approfondir les connaissances des apprenants en leur proposant de classer des expressions selon les FD abordées dans la partie concernée (Figure 8).

L'objectif principal de cet exercice était de vérifier leur compréhension à la fin de chaque séance. De plus, ces activités permettaient aux apprenants de découvrir un ensemble d'expressions tout en facilitant leur mémorisation.

IV. Classement des constructions étudiées (10m)

Consigne : Vous venez d'apprendre des expressions servant à présenter un choix de l'approche, la limite d'un choix et la collecte des données. Classez les expressions ci-dessous dans les catégories proposées dans le tableau.

1. *Nous nous limitons à étudier .../à l'étude*
2. *Rassembler des articles/des textes*
3. *Ayant pour objectif de..., nous avons adopté l'approche ...*
4. *L'étude est basée sur l'approche/le corpus...*
5. *Pour mener à bien cet objectif, nous nous sommes appuyés sur...*
6. *Afin de..., nous avons choisi de....*
7. *Nous adoptons/retenons/privilegions l'approche*
8. *Regrouper les données*
9. *On se restreint à l'analyse...*
10. *Nous optons pour la méthode*
11. *Les données sont recueillies*

| Présenter un choix favorable | Présenter la limite d'un choix | Présenter la collecte des données |
|------------------------------|--------------------------------|-----------------------------------|
| | | |

Figure 8. Exemple d'exercice de classement

3.2.5. EXERCICE DE RÉDACTION LIBRE

Cet exercice permet le réemploi des FD en utilisant notre base LST-SHS. À l'issue de chaque séance les étudiants ont travaillé directement sur leur mémoire pendant 30 min. Les étudiants ont rédigé quelques paragraphes de la partie abordée en classe à l'aide de la base. Des questions (Figure 9) les aidaient à organiser leurs idées. La figure 9 montre un exemple de rédaction libre avec des questions d'incitation pour la partie méthodologique.

V. Production écrite d'un paragraphe sur la méthodologie de recherche (30m)

Consigne : Dans le cadre de votre mémoire, vous allez présenter en quelques lignes l'approche adoptée et la collecte des données en essayant de répondre aux questions suivantes :

1. Avec quel objectif de recherche, souhaitez-vous constituer un corpus/collecter des données ?
2. Quel est le type de corpus/de données recueillies (articles, productions écrites/orales, etc.) ?
3. De quelle façon les données ont-elles été recueillies (enregistrement, entretien, questionnaire, etc.) ? Auprès de quel public ?
4. Quel type de méthode d'analyse utilisez-vous (empirique/statistique/quantitatif/qualitatif) ?
5. Pourquoi choisir cette méthode de collecte ou d'analyse ? Est-elle appuyée sur des théories ou inspirée des études précédentes ?
6. Concernant la représentativité du corpus/des données/d'analyse, y a-t-il une limite de recherche ?

Figure 9. Exemple d'exercice de rédaction libre

En somme, ces exercices portent à la fois sur la compréhension des FD et sur leur production. Ils ont été fournis aux apprenants au format papier et leurs réponses ont été scannées avant d'être partagées pour la correction collective. Le nombre d'exercices collectés pour chaque partie variait en raison des absences occasionnelles des participants. Hormis les réponses aux exercices, d'autres éléments ont été recueillis pour cette recherche ; en voici un aperçu.

4. DONNÉES COLLECTÉES

À l'issue de la formation, nous avons collecté trois types de données : les réponses des apprenants aux exercices, les captures vidéo d'écran (CVE) et les réponses au questionnaire.

4.1. CAPTURE VIDÉO D'ÉCRAN (CVE)

À l'instar des travaux d'Hamel (2012) et de Vincze et Alonso Ramos (2013), la technologie de CVE est utilisée pour observer les manipulations de l'outil par les apprenants pendant les exercices de reformulation et de rédaction libre. Par rapport à d'autres méthodes pour connaître le processus d'écriture des apprenants, comme les protocoles de réflexion à haute voix (Janssen, Van Waes & Van den Bergh, 1996), le logiciel de CVE est souvent exécuté de manière invisible sur un ordinateur. Cela permet de recueillir discrètement les données sans interrompre la réflexion des apprenants, ce qui contribue ainsi à réduire toutes perturbations dans l'écriture. Par ailleurs, la CVE nous a permis de recueillir des travaux d'étudiants faits chez eux, quand cela était impossible en cours.

Ainsi, dans les CVE, nous voyons les manipulations de l'outil, comme les endroits cliqués, les pages consultées, les accès utilisés, etc., et l'élaboration du texte. Les exercices présentés précédemment ont été fournis aux apprenants au format papier, nous avons alors combiné les CVE et les réponses écrites sur les copies. Le tableau 1 résume les vidéos que nous avons collectées pour les exercices de reformulation et pour la rédaction libre. Les vidéos pour la reformulation durent au total trois heures et demie et la durée moyenne d'une vidéo est d'environ 12 min : temps moyen consacré à un exercice de reformulation (avec 5 questions) par les apprenants. Les vidéos pour la rédaction libre durent de 10 à 50 min.

Tableau 1. CVE collectées pendant la formation

| Type d'exercice | Exercices de reformulation | Rédaction libre | | | |
|----------------------------|----------------------------|-----------------|---------|--------------|-----------------------|
| Partie | | Introduction | Théorie | Méthodologie | Analyses et résultats |
| Nombre de vidéos collectés | 15 | 6 | 3 | 4 | 5 |

4.2. QUESTIONNAIRE

Un questionnaire a été distribué aux apprenants à la fin de la formation afin d'évaluer l'outil lexical et la méthode d'enseignement par corpus. Le questionnaire comptait cinq entrées : 1) la convivialité de l'outil ; 2) la qualité du contenu ; 3) l'organisation et la présentation des données ; 4) l'utilité de la ressource pour résoudre les problèmes lexicaux dans la rédaction et 5) des propositions d'amélioration de l'outil. En ce qui concerne la formation, nous souhaitons connaître les connaissances acquises par les participants, la compréhension du contenu de la formation et les difficultés pendant l'apprentissage pour mesurer l'efficacité pédagogique. Nous avons recolté 6 réponses qui constituent un petit échantillon de données dont nous présentons ci-dessous les analyses.

5. ANALYSES ET RÉSULTATS

Pour avancer vers notre objectif, les analyses portent sur trois aspects : 1) la compréhension des FD à l'aide de l'enseignement et sa contribution à la compréhension de la structure générale du genre scientifique, 2) l'identification des FD à partir des expressions données – comme une sensibilisation à l'utilité des FD – et 3) le réemploi des FD dans la rédaction, à travers l'utilisation des expressions sous les étiquettes fonctionnelles dans la base LST-SHS.

5.1. COMPRÉHENSION DES FONCTIONS DISCURSIVES DANS LES EXERCICES

Parmi les cinq types d'exercices présentés précédemment, trois d'entre eux portent sur la compréhension des FD (relier, repérer, classer). Le premier exercice pose davantage de problèmes aux apprenants que les deux derniers, car réalisé sans formation et présenté avec les étiquettes plus courtes. En combinant l'observation des erreurs et les réponses issues du questionnaire, nous avons remarqué que les apprenants avaient des difficultés à comprendre certains termes, comme « problématique » et « renvoyer ». Si une partie de ces difficultés sont attribuables à des lacunes lexicales chez les apprenants, l'autre partie peut résider dans la méconnaissance de certaines notions scientifiques. Pour les aider, nous pouvons simplifier davantage les étiquettes fonctionnelles dans la base LST-SHS par des mots moins spécialisés. Par exemple, nous pouvons remplacer « renvoyer » par « citer », « reformuler » par « redire », avec des descriptions simplifiées de ces fonctions. En revanche, « problématique » ou « constat » sont difficiles à reformuler. Il faudrait des enseignements spécifiques pour ces notions.

Les apprenants avaient une meilleure performance dans les exercices de repérage et de classement. D'après le questionnaire, les apprenants ont beaucoup apprécié la méthode d'enseignement avec le corpus. En outre, ils ont affirmé que ces fonctions les ont aidés à mieux comprendre les éléments à présenter dans chaque partie du mémoire.

5.2. IDENTIFICATION DES FONCTIONS DISCURSIVES DANS LES EXERCICES DE REFORMULATION

Deux exercices de reformulation ont été proposés : le premier dans la première séance, sans formation préalable pour tester l'intuition des apprenants, où les fonctions pouvaient concerner toutes les parties d'un écrit scientifique. Le deuxième était dans la partie du cadre théorique, qui a porté exclusivement sur les FD enseignées dans la partie théorique.

Les apprenants ont eu des difficultés à identifier les FD à partir des expressions. C'est ce qui ressort des exercices de ce type. D'après les CVE, les apprenants ont

passé du temps à parcourir la liste de fonctions dans la base LST-SHS pour faire le premier exercice de reformulation. Si les FD comme « donner des exemples » ont été bien repérées, les FD « exprimer un point de vue » étaient sujettes à des erreurs. Le point positif est que les apprenants ont bien reformulé la phrase avec les bonnes expressions, une fois qu'ils avaient identifié les fonctions appropriées.

Les apprenants ont mieux réussi le deuxième exercice de reformulation, en commettant peu d'erreurs et en moins de temps (la question du temps ne sera pas abordée dans cet article). Leur réussite peut s'expliquer par le fait que l'exercice a porté exclusivement sur les FD étudiées dans la partie théorique, à savoir « renvoyer aux autres auteurs » et « signaler la théorie, le principe, ou le domaine d'où on part ». Les exercices d'entraînement et les explications de l'enseignant ont aidé les apprenants à mieux comprendre l'utilisation de ces deux FD.

5.3. UTILISATION DE L'ACCÈS PAR LES FONCTIONS DISCURSIVES DANS LA RÉDACTION LIBRE

Les FD servent de pont entre 1) l'exposition à des éléments préconstruits dans les écrits scientifiques et 2) leur production dans la rédaction par les apprenants. En revanche, en observant les CVE, nous avons remarqué qu'il n'était pas évident pour les apprenants d'utiliser l'accès onomasiologique, c'est-à-dire d'entrer par les FD pour formuler leurs idées, et ceci même après les entraînements.

Lors de la rédaction libre, nous pouvons voir que les apprenants n'utilisaient pas souvent l'accès par les FD. La figure 10 présente le pourcentage d'utilisation de l'accès par les FD par rapport à d'autres accès dans la base LST-SHS, en fonction des apprenants (1 à 6). On peut voir que l'accès par FD est moins utilisé que les autres.

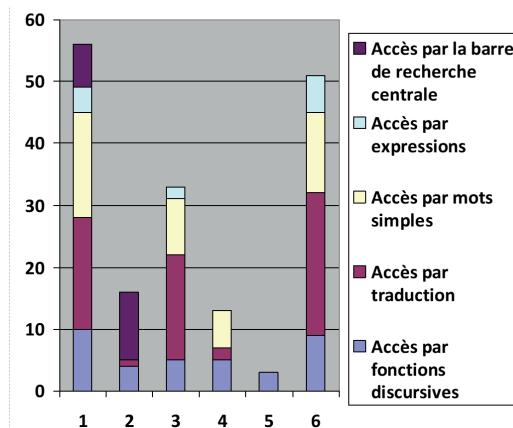


Figure 10. Nombre d'utilisation de l'accès par les FD par rapport à d'autres accès dans la base LST-SHS en fonction des apprenants (1 à 6)

Pour mesurer l'effet de l'enseignement, nous avons observé que les fonctions abordées dans la formation sont plus utilisées que celles qui n'ont pas été étudiées. Si nous analysons plus précisément les FD que les apprenants ont utilisées, nous constatons que la plupart servent à annoncer les grandes lignes d'une étude (l'objectif, la problématique, les hypothèses...), comme « Annoncer l'objet/le sujet », « Présenter la problématique/les hypothèses ». Il est intéressant de voir que les apprenants ont tendance à consulter plusieurs fois une même fonction, ce qui reflète leur processus de réflexion. De plus, les fonctions pour présenter les interprétations et les résultats ont souvent été exploitées de manière successive. Nous supposons que les apprenants avaient l'intention de comparer les expressions dans ces deux fonctions, ce qui pourrait également refléter une certaine difficulté à bien les différencier.

Les deuxièmes FD les plus utilisées par les apprenants portaient sur : la structuration : « Annoncer le plan », « Pour conclure » et la référence aux travaux ou théories précédents : « Signaler la théorie d'où on part », « Renvoyer aux autres auteurs ».

Certaines FD, comme « Exprimer le but », « Reformuler en d'autres termes », ont été consultées une fois par différents apprenants, tandis que les autres n'ont jamais été consultées. Ces fonctions négligées servent à mettre en évidence un élément, à évaluer un objet, à définir les termes. Ce constat rejoint celui de Yan (2017) à propos des difficultés des apprenants sinophones à insister sur un point et exprimer leur point de vue. Il est surprenant de voir que la définition des termes dans la partie théorique est absente dans les écrits des apprenants. Nous supposons que ces fonctions peu utilisées peuvent refléter en partie le manque de connaissance de ces notions exigées par le genre scientifique, comme le positionnement et la définition des termes. Elles demandent donc des sensibilisations particulières.

Par ailleurs, l'exploitation des FD ne signifie pas leur compréhension par les apprenants. Nous pouvons évaluer la compréhension grâce à l'adéquation des résultats de recherche dans les vidéos. Autrement dit, nous avons examiné si les FD choisies par les apprenants correspondaient à leur intention de recherche et si elles leur permettaient d'obtenir des résultats pertinents. Nous pouvons distinguer les cas suivants pour mesurer la maîtrise des FD utilisées par les apprenants :

- 1) FD bien choisie, expressions bien utilisées : réussite de recherche ;
- 2) FD bien choisie, l'apprenant a négligé les expressions qui auraient pu l'aider, ou a utilisé des expressions erronées : échec de recherche ;
- 3) FD qui ne correspondait pas à l'intention de recherche de l'apprenant : échec de recherche ;
- 4) Pas de reflet dans la rédaction après la recherche des FD, de sorte que nous ne pouvons pas identifier l'intention de recherche de l'apprenant.

Les réussites de recherche se concentrent sur les fonctions les plus fréquemment utilisées par les apprenants, comme celles pour annoncer les grandes lignes de la recherche et celles sur la structuration. À l'aide de l'outil, les apprenants avaient accès à des expressions variées pour aider à la rédaction.

D'après les CVE, il existe plusieurs cas où les fonctions ont été bien choisies, mais les apprenants ont négligé des expressions qui auraient pu les aider. Par ailleurs, les fonctions moins utilisées sont souvent sources d'erreurs. Par exemple, un apprenant a choisi la fonction « Préciser un sujet » et obtenu l'expression *plus précisément*, mais il l'a utilisée dans la phrase *Cette étude aborde le sujet d'étudier l'exactitude des citations. Plus précisément, d'après une étude, il y a une moyenne taux d'erreur de 25.4 % pour les citations.* Nous faisons l'hypothèse que l'apprenant n'a pas bien compris que cette FD sert à préciser une idée mentionnée précédemment.

Une confusion entre les FD « Pour conclure » et « Pour énumérer des éléments » a été observée. Cette confusion pourrait être expliquée par le fait qu'en chinois, le mot 最后 (*zui hòu*) correspond à la fois au mot *au final* pour donner une conclusion et à *dernièrement* pour indiquer le dernier élément d'une liste. Il se peut que l'apprenant pense au mot 最后 (*zui hòu*) en chinois, mais qu'il ne comprenne pas s'il s'agit de la conclusion ou du dernier élément. Cela rejoint les réponses des apprenants dans le questionnaire : ils avaient parfois du mal à saisir la différence entre certaines FD.

Par ailleurs, dans la rédaction libre, de nombreuses recherches effectuées dans la base n'ont pas abouti à des résultats dans la rédaction finale. Dans ce cas, nous ne pouvons pas identifier exactement les raisons pour lesquelles les recherches des FD n'ont pas eu d'écho dans la rédaction, mais une interprétation possible est que les FD choisies ne correspondaient pas à l'intention de communication des apprenants, et ils n'ont pas réussi à en trouver une appropriée. Cela peut s'expliquer par trois éléments : 1) la non-exhaustivité de la liste en termes de FD, 2) la maîtrise insuffisante des FD par les apprenants, qui sont passés à côté des fonctions qui auraient pu les aider, ou 3) la trop grande spécialisation terminologique de l'outil.

Pour résumer, nous voyons que si les fonctions simples, comme l'annonce des grandes lignes d'une étude (objectif, problématique...) et la structuration sont à peu près maîtrisées par les apprenants, les FD plus complexes sur le positionnement ou la topicalisation nécessitent un enseignement et une sensibilisation. Le fait que les apprenants soient plus à l'aise avec les FD présentées pendant la formation, comparées à celles non abordées, confirme l'impact positif de cette dernière. Nous n'avons pas abordé les FD de mise en relief et de reformulation parce qu'elles ne sont pas spécifiques à une partie de l'écrit scientifique, mais apparaissent tout au long de l'écrit. Cependant, ces FD représentent des difficultés récurrentes pour les apprenants et mériteraient d'être intégrées dans une formation future.

6. CONCLUSION

Dans l'optique de faciliter la compréhension et la production de textes scientifiques, nous avons enseigné le LST à des apprenants de langue première chinoise. Cette ap-

proche pédagogique s'est appuyée sur l'analyse des FD par les apprenants et a mis en œuvre une méthodologie d'enseignement basée sur corpus.

Les FD spécifiques à chaque partie d'un écrit scientifique jouent un rôle essentiel dans la compréhension des éléments à présenter et dans le fonctionnement discursif propre au genre scientifique. L'enseignement des FD doit donc intégrer non seulement des aspects linguistiques, mais aussi la conceptualisation des notions scientifiques, car les difficultés des apprenants résident à la fois dans la maîtrise de la langue et dans la compréhension des concepts. Par exemple, la compréhension insuffisante des notions scientifiques, comme « problématique », « constat », « positionnement », entrave la maîtrise de certaines FD. De plus, l'exposition des apprenants aux éléments préconstruits qui portent les FD, n'induit pas obligatoirement la compréhension des éléments, même s'ils sont capables d'en reconnaître le signifié. En outre, le signifiant est primordial pour la production et nous avons vu que l'incompréhension empêche une utilisation à bon escient. Dans ce cas, les FD peuvent servir de lien entre la compréhension des éléments préconstruits en lecture et la production de nouveaux éléments dans la rédaction académique des apprenants. Ce processus, allant de la compréhension à la production, requiert non seulement une simplification des descriptions des FD, mais aussi la mise en place de formations spécifiques visant à aider les apprenants à développer des aptitudes à utiliser des accès onomasiologiques pour améliorer leurs compétences rédactionnelles.

Notre formation étant courte, nous n'avons pas pu aborder toutes les FD de la base LST-SHS, mais cela constitue une piste à développer à l'avenir, lors des formations à l'utilisation de cet outil.

BIBLIOGRAPHIE

- Alonso-Ramos, M. & García Salido, M. (2019). Testing the Use of a Collocation Retrieval Tool Without Prior Training by Learners of Spanish. *International Journal of Lexicography*, 32 (4), 480-497. <https://doi.org/10.1093/ijl/ecz016>.
- Cavalla, C. (2018). Lexique transdisciplinaire et enseignement aux étudiants allophones. In A. Tutin & M.-P. Jacques (éds.), *Lexique transversal et formules discursives des sciences humaines* (pp. 191-214). London : ISTE Editions.
- CECRL. (2001). *Cadre européen commun de référence pour les langues*. www.coe.int/t/dg4/linguistic/Source/Framework_fr.pdf.
- Hamel, M.-J. (2012). Testing aspects of the usability of an online learner dictionary prototype: A product- and process-oriented study. *Computer Assisted Language Learning*, 25 (4), 339-365. <https://doi.org/10.1080/09588221.2011.591805>.
- Hatier, S. (2016). *Identification et analyse linguistique du LST. Approche outillée sur un corpus d'articles de recherche en SHS* (doctoral dissertation). Université Grenoble Alpes.
- Hatier, S., Augustyn, M., Yan, R., Tran, T., Tutin, A. & Jacques, M.-P. (2016). French cross-disciplinary scientific lexicon: Extraction and linguistic analysis. In T. Margalitadze & G. Meladze (éds.), *Proceedings of EURALEX*, (pp. 355-366). London : ISTE Editions.
- Jacques, M.-P. & Tutin, A. (éds.) (2018). *Lexique transversal et formules discursives des sciences humaines*. London : ISTE Editions.
- Janssen, D., Van Waes, L. & Van den Bergh, H. (1996). Effects of thinking aloud on writing processes. In C.M. Levy & S. Ransdell (éds.), *The science of writing: Theories, methods, individual differences & applications* (pp. 233-250). New Jersey : Lawrence Erlbaum Associates.
- Ji, Y. (2022). *Les routines de guidage du lecteur dans les écrits scientifiques en français* (doctoral dissertation). Université Grenoble Alpes.
- Li, J. (2023). Accès en chinois à un outil lexical pour aider à la rédaction scientifique des apprenants sinophones. *Recherche et pratiques pédagogiques en langues. Cahiers de l'Aplut*, 42 (2). <https://journals.openedition.org/apliut/11021>.
- Li, J. (2024). *Réflexion et mise en oeuvre d'une adaptation didactique d'une ressource lexicale pour l'aide à la rédaction scientifique* (doctoral dissertation). Université Grenoble Alpes.
- Li, J. & Tutin, A. (2023). Quels traitements pour la phraséologie scientifique transdisciplinaire dans une perspective d'aide à la rédaction ? In P. Frassi (éd.), *Phraséologie et terminologie* (pp. 105-130). Berlin, Boston : De Gruyter.
- Luodonpää-Manni, M., Grossmann, F. & Tutin, A. (2022). *Les routines discursives dans le discours scientifique oral et écrit*. London : UGA Éditions.
- Swales, J.M. (1990). *Genre Analysis: English in Academic and Research Settings*. Cambridge : Cambridge University Press.
- Tran, T.T.H. (2014). *Description de la phraséologie transdisciplinaire des écrits scientifiques et réflexions didactiques pour l'enseignement à des étudiants non-natifs : Application aux marqueurs discursifs* (doctoral dissertation). Université Grenoble Alpes.
- Vincze, O. & Alonso Ramos, M. (2013). Testing an electronic collocation dictionary interface: Diccionario de Colocaciones del Español. In I. Kosem, J. Kallas, P. Gantar, S. Krek, M. Langemets, K. Tuulik (éds.), *Electronic lexicography in the 21st century: thinking outside the paper. Proceedings of the eLex 2013 conference, Tallinn, Estonia, 17-19 October 2013* (pp. 328-337). Ljubljana-Tallinn: Trojina, Institute for Applied Slovene Studies-Eesti Keele Instituut.
- Wang, Y. (2022). *Repérage et annotation des routines sémantico-rhétoriques transdisciplinaires dans les écrits scientifiques* (master thesis). Université Grenoble Alpes.

Yan, R. (2017). *Étude des constructions verbales scientifiques dans une perspective didactique : Utilisation des corpus dans le diagnostic des besoins langagiers en FLE à l'aide des techniques de TAL* (doctoral dissertation). Université Grenoble Alpes.