

Université de Montréal

**Description du lexique spécialisé chinois et constitution d'une ressource didactique adaptée
pour locuteurs non sinophones**

Par

Zhiwei Han

Département de linguistique et de traduction

Faculté des arts et des sciences

Thèse présentée en vue de l'obtention du grade de Philosophiæ Doctor (Ph. D.)

en traduction, option terminologie

Octobre 2022

© Zhiwei Han, 2022

Université de Montréal

Cette thèse intitulée

**Description du lexique spécialisé chinois et constitution d'une ressource didactique adaptée
pour locuteurs non sinophones**

Présenté par

Zhiwei Han

a été évaluée par un jury composé des personnes suivantes

François Lareau

Président-rapporteur

Marie-Claude L'Homme

Directrice de recherche

Chantal Gagnon

Membre du jury

Pascale Elbaz

Examineur externe

Résumé

L'enseignement-apprentissage du lexique spécialisé chinois est un chemin semé d'obstacles. Pour les apprenants non natifs, les combinaisons lexicales spécialisées (CLS) (L'Homme, 2000) soulèvent des difficultés syntaxico-sémantiques et représentent ainsi un défi majeur dans l'acquisition de compétences lexicales. On recense, toutefois, peu de propositions méthodologiques pour résoudre ces difficultés dans la littérature consacrée à la didactique du chinois sur objectifs spécifiques (COS) (Q. Li, 2011).

Dans cette recherche, nous nous attachons à explorer de quelle manière une méthode de description lexicale basée sur une représentation sémantique et syntaxique assiste les apprenants non natifs dans la résolution des problèmes lexicaux soulevés par les CLS.

Notre thèse vise à concevoir une méthode de description des CLS en vue de la résolution de difficultés lexicales par les locuteurs non sinophones. La méthode mise au point est appliquée à l'élaboration du dictionnaire CHINOINFO, une ressource lexicale chinois-français portant sur le domaine de l'informatique. Cette ressource s'adresse aux apprenants francophones du chinois. L'objectif secondaire de notre thèse consiste à évaluer l'efficacité du CHINOINFO auprès des apprenants francophones qui reçoivent une formation de chinois dans un établissement universitaire au Québec ou en Chine.

Notre recherche fait appel à des notions empruntées à trois cadres théoriques. Premièrement, la Lexicologie explicative et combinatoire (Mel'čuk et al., 1995) nous sert d'appui théorique pour fonder la description des CLS sur la représentation sémantique du lexique spécialisé. Deuxièmement, notre démarche de collecte et d'analyse des CLS est guidée par l'approche lexico-sémantique à la terminologie (L'Homme, 2020a). Enfin, nous nous appuyons sur l'approche cognitive en didactique des langues secondes (Chastain, 1990) pour envisager une présentation structurée des connaissances lexicales.

Notre démarche méthodologique s'est déroulée en trois phases. Nous avons d'abord assemblé un corpus spécialisé chinois pour en extraire un échantillon de CLS et les renseignements permettant de les décrire. L'analyse des données collectées à partir du corpus nous a amenée à anticiper trois types de difficultés syntaxico-sémantiques soulevées par les CLS : 1) distinguer les acceptions

d'un polysème dans différentes CLS; 2) différencier les sens distincts de CLS de forme identique; 3) choisir les cooccurrents appropriés d'un terme.

À la deuxième phase, nous avons mobilisé différentes stratégies pour décrire les propriétés syntaxico-sémantiques des CLS. Une méthode descriptive intégrant les solutions proposées a ensuite été appliquée à la création du CHINOINFO. Cette ressource en ligne répertorie 91 termes fondamentaux du domaine de l'informatique, pour lesquels nous avons encodé au total 282 termes reliés et 644 CLS. La structuration des données au sein des articles s'est largement inspirée de l'adaptation du DiCoInfo (Observatoire de linguistique Sens-Texte, 2022) à un dictionnaire d'apprentissage (Alipour, 2014). Différents moyens techniques ont été mis en œuvre pour assurer la convivialité de la ressource.

La dernière phase de notre recherche consiste en une expérimentation comparative visant à évaluer l'efficacité pédagogique du CHINOINFO. Nous avons fait passer un test lexical à deux groupes d'apprenants francophones, soit le groupe contrôle (GC) et le groupe expérimental (GE), en leur proposant un nombre d'outils de référence. Seul le GE a eu accès à CHINOINFO. Nous avons aussi collecté, au moyen de questionnaires de sondage, le profil des participants ainsi que leur appréciation sur le test et les outils de référence proposés.

De manière générale, l'analyse comparative des résultats du test lexical montre que le GE a mieux réussi à résoudre les trois types de difficultés soulevées par les CLS. Les participants étaient plutôt satisfaits de l'organisation du test. Le GE a eu moins de difficultés à réaliser le test puisqu'il se sentait mieux outillé pour trouver des éléments de réponses aux questions du test par rapport GC. Le GE s'est exprimé favorablement quant à l'utilité du CHINOINFO pour résoudre les problèmes lexicaux dans le cadre de notre expérimentation.

Pour conclure, les résultats de notre analyse fournissent des indices sur l'apport du CHINOINFO en tant qu'une ressource d'apprentissage des CLS, ce qui laisse entrevoir l'intérêt de la méthode de description lexicale que nous avons proposée dans un contexte pédagogique.

Mots-clés : lexique spécialisé, combinaison lexicale spécialisée, polysémie, Lexicologie explicative et combinatoire, approche lexico-sémantique, approche cognitive en didactique des langues secondes, chinois sur objectifs spécifiques, ressource lexicale, approche basée sur le corpus.

Abstract

The teaching and learning of Chinese specialized lexicon is a path strewn with obstacles. For non-native learners, specialized lexical combinations (SLCs) (L'Homme, 2000) raise syntactic and semantic difficulties and thus represent a major challenge in the acquisition of lexical skills. However, there are few methodological proposals to solve these difficulties in the literature devoted to the teaching practice and applied research of Chinese for specific purposes (Q. Li, 2011).

In this research, we explore how a lexical description method based on semantic and syntactic representation assists non-native learners in solving lexical problems raised by SLCs.

This thesis aims at designing a method for describing SLCs to help non-Chinese speakers solve lexical difficulties. The proposed method is applied to develop CHINOINFO, a Chinese-French dictionary of computer science and information technology terms. This lexical resource is designed for French-speaking learners of Chinese and can also be used as a writing tool for language professionals (translators, technical writers, and proofreaders), as well as professionals in this field. The secondary objective of this thesis is to evaluate the pedagogical effectiveness of the developed resource among French-speaking university students.

This research draws on concepts derived from three theoretical frameworks. Firstly, the Explanatory and Combinatorial Lexicology (Mel'čuk et al., 1995) provides theoretical support for founding the lexical description on the representation of semantic features of the specialized lexicon. Secondly, the collection and analysis of SLCs are guided by the lexical-semantic approach to terminology (L'Homme, 2020a). Finally, we draw on the cognitive approach to second language didactics (Chastain, 1990) to explore the effective ways to organize and present the descriptive information of the specialized lexicon.

Our methodological approach was carried out in three stages. We started by assembling a specialized Chinese corpus to extract a sample of SLCs and their descriptive information. The analysis of the data collected from the corpus led us to anticipate three types of syntaxico-semantic difficulties raised by SLCs: 1) distinguishing polysemes in different SLCs; 2)

identifying, in a given context, the meaning of a lexical combination that is syntactically ambiguous; and 3) selecting appropriate co-occurrences for a term.

In the second stage, we deployed different strategies to describe the syntaxico-semantic features of SLCs. Subsequently, a descriptive method that incorporates the proposed solutions has been applied to the creation of CHINOINFO. This online lexical resource contains 91 basic terms related to computer science and information technology. For these terms, we encoded a total of 282 related terms and 644 SLCs. The organization of content in the entries has been largely inspired by the conversion of DiCoInfo (Observatoire de linguistique Sens-Texte, 2022) into a learner's dictionary (Alipour, 2014). We used various techniques to make the resource user-friendly.

The final stage of our research consists of a comparative experiment to evaluate the pedagogical effectiveness of CHINOINFO. We had two groups of French-speaking learners, the control group (CG) and the experimental group (EG), take a lexical test by providing them with several reference materials. Only the EG had access to CHINOINFO during the test. We also collected the information about the participants' learning profile and their appreciation of the test and the proposed reference materials.

Overall, the comparative analysis of the test results shows that the GE succeeded better in solving the three types of difficulties raised by the SLCs. The participants were quite satisfied with the organization of the lexical test. The EG encountered less difficulty in answering questions during the test since they felt better equipped to find elements of answers in the reference materials than the GC. The EG commented favorably on the utility of CHINOINFO in solving lexical problems.

To conclude, the results of our experiment provide clues about the pedagogical interest of CHINOINFO as a SLC learning resource, which suggests the relevance of the lexical description method we proposed in a pedagogical context.

Keywords: specialized lexicon, specialized lexical combination, polysemy, Explanatory and Combinatorial Lexicology, lexical-semantic approach, Cognitive-code learning theory, Chinese for Specific Purposes, lexical resource, corpus-based approach.

Table des matières

| | |
|---|----|
| Résumé | 5 |
| Abstract | 7 |
| Table des matières | 9 |
| Liste des tableaux | 15 |
| Liste des figures | 17 |
| Liste des sigles et abréviations | 21 |
| Conventions d'écriture | 23 |
| Remerciements | 27 |
| Chapitre 1 Introduction | 29 |
| 1.1 Contexte et problématique | 29 |
| 1.1.1 Enseignement du lexique spécialisé chinois : défis à relever..... | 30 |
| 1.1.2 Méthodes didactiques du lexique | 34 |
| 1.1.3 Problématique..... | 37 |
| 1.2 Objectifs | 38 |
| 1.3 Hypothèses | 39 |
| 1.4 Originalité de la thèse..... | 39 |
| 1.5 Structure de la thèse | 40 |
| Chapitre 2 État de la question | 42 |
| 2.1 Enseignement du lexique dans une perspective sémantique..... | 42 |
| 2.1.1 Lexique en langue générale..... | 42 |
| 2.1.2 Lexique spécialisé | 45 |
| 2.2 Étude de la combinatoire lexicale dans une perspective didactique | 47 |
| 2.2.1 Définition de la combinatoire lexicale en vue de son enseignement | 48 |
| 2.2.2 Enseignement de la combinatoire lexicale spécialisée..... | 50 |
| 2.3 Enseignement du lexique spécialisé chinois | 54 |
| 2.3.1 Enseignement du chinois sur objectif spécifique (COS)..... | 54 |
| 2.3.1.1 Qu'est-ce que l'on entend par « objectifs spécifiques » ? | 54 |
| 2.3.1.2 Profil du public ciblé | 57 |

| | |
|--|-----|
| 2.3.1.3 Objectifs d'enseignement..... | 58 |
| 2.3.2 Approches didactiques à l'œuvre..... | 60 |
| 2.3.3 Ressources d'apprentissage mises au point..... | 67 |
| 2.4 Synthèse..... | 73 |
| Chapitre 3 Cadre théorique..... | 74 |
| 3.1 Lexicologie explicative et combinatoire (LEC)..... | 74 |
| 3.1.1 Notions et principes fondamentaux de la LEC..... | 74 |
| 3.1.2 Applications lexicographiques de la LEC..... | 90 |
| 3.1.3 Application de la LEC à la didactique des langues..... | 91 |
| 3.1.4 Pertinence de la LEC comme cadre de référence..... | 92 |
| 3.2 Approche lexico-sémantique de la terminologie..... | 93 |
| 3.2.1 Principes fondamentaux de l'approche lexico-sémantique..... | 93 |
| 3.2.1.1 Notions de base de l'approche lexico-sémantique..... | 93 |
| 3.2.1.2 Cadres conceptuels de l'approche lexico-sémantique..... | 94 |
| 3.2.1.3 Propositions théoriques et méthodologiques de l'approche lexico-sémantique..... | 95 |
| 3.2.2 Application de l'approche lexico-sémantique à la confection de dictionnaires spécialisés..... | 97 |
| 3.2.3 Application de l'approche lexico-sémantique à l'enseignement du lexique spécialisé..... | 100 |
| 3.2.4 Intérêt de l'approche lexico-sémantique comme cadre de référence..... | 101 |
| 3.3 Approche cognitive en didactique des langues secondes..... | 102 |
| 3.3.1 Principes fondamentaux de l'approche cognitive..... | 103 |
| 3.3.1.1 Notions associées aux méthodes d'enseignement des langues..... | 103 |
| 3.3.1.2 Principes théoriques de l'approche cognitive..... | 104 |
| 3.3.2 Implication méthodologique dans l'enseignement du lexique..... | 106 |
| 3.3.3 Pertinence de l'approche cognitive comme cadre de référence..... | 108 |
| Chapitre 4 Méthodologie..... | 112 |
| 4.1 Extraire et analyser des CLS à partir d'un corpus..... | 113 |
| 4.1.1 Compilation du corpus..... | 113 |

| | |
|---|-----|
| 4.1.1.1 Critères de sélection | 114 |
| 4.1.1.2 Nettoyage et segmentation du corpus..... | 125 |
| 4.1.2 Identification des termes | 126 |
| 4.1.2.1 Extraction des candidats-termes (CT)..... | 126 |
| 4.1.2.2 Identification des unités terminologiques (UT) | 129 |
| 4.1.2.3 Distinction sémantique des UT | 131 |
| 4.1.2.4 Repérage des termes à partir des UT polysémiques..... | 137 |
| 4.1.3 Collecte et analyse des CLS | 143 |
| 4.1.3.1 Extraction des CLS..... | 143 |
| 4.1.3.2 Extraction des contextes d'utilisation des CLS..... | 150 |
| 4.1.3.3 Synthèse des renseignements recueillis..... | 153 |
| 4.1.3.4 Anticipation des difficultés syntaxico-sémantiques soulevées par les CLS | 158 |
| 4.1.4 Synthèse | 163 |
| 4.2 Concevoir une méthode descriptive des CLS en vue de l'élaboration d'une ressource lexicale | 163 |
| 4.2.1 Mise au point de la méthode de description..... | 163 |
| 4.2.1.1 Difficulté 1 : distinguer les acceptions d'un polysème dans différentes CLS | 163 |
| 4.2.1.2 Difficulté 2 : différencier les sens distincts de CLS de forme identique.. | 173 |
| 4.2.1.3 Difficulté 3 : choisir les cooccurrents appropriés d'un terme | 175 |
| 4.2.1.4 Synthèse des solutions proposées..... | 183 |
| 4.2.2 Création de la ressource lexicale | 184 |
| 4.2.2.1 Compilation des données terminologiques | 184 |
| 4.2.2.2 Encodage des données terminologiques..... | 187 |
| 4.2.2.3 Mise en ligne de la ressource | 190 |
| 4.2.3 Synthèse | 191 |
| 4.3 Évaluer l'efficacité de la ressource lexicale mise au point | 191 |
| 4.3.1 Conception d'une expérimentation comparative..... | 192 |
| 4.3.1.1 Objectifs de l'expérimentation | 192 |

| | |
|--|-----|
| 4.3.1.2 Profil des participants recherchés..... | 193 |
| 4.3.1.3 Outils de cueillette des données | 194 |
| 4.3.1.4 Recrutement des participants | 202 |
| 4.3.2 Protocole d'expérimentation | 204 |
| 4.3.2.1 Planification des séances expérimentales..... | 204 |
| 4.3.2.2 Collecte de données..... | 204 |
| 4.3.2.3 Analyse des données | 207 |
| 4.3.3 Synthèse | 208 |
| Chapitre 5 Présentation du CHINOINFO | 209 |
| 5.1 Nomenclature du CHINOINFO | 209 |
| 5.2 Structure interne des articles | 211 |
| 5.2.1 Renseignements généraux | 212 |
| 5.2.2 Définition et structure actancielle | 212 |
| 5.2.3 Contextes | 215 |
| 5.2.4 Termes reliés et combinaisons lexicales | 216 |
| 5.3 Fonctionnalités du CHINOINFO | 218 |
| 5.3.1 Recherche de termes..... | 219 |
| 5.3.2 Guide d'utilisation du CHINOINFO..... | 221 |
| 5.3.3 Renvois..... | 221 |
| 5.3.4 Bilan des liens lexicaux encodés | 223 |
| 5.4 Synthèse | 224 |
| Chapitre 6 Traitement des résultats expérimentaux | 226 |
| 6.1 Analyse des réponses au questionnaire de profil | 226 |
| 6.1.1 Méthode d'analyse des données collectées | 226 |
| 6.1.2 Résultats d'analyse..... | 227 |
| 6.2 Analyse des réponses au test lexical | 233 |
| 6.2.1 Correction des réponses aux questions lexicales..... | 233 |
| 6.2.2.1 Analyse intragroupe | 234 |
| 6.2.2.2 Analyse intergroupe | 238 |
| 6.2.3 Analyse des erreurs | 240 |

| | | |
|------------|---|-----|
| 6.2.3.1 | Méthode d'analyse | 240 |
| 6.2.3.2 | Résultats d'analyse | 241 |
| 6.2.4 | Synthèse des réponses à la question libre..... | 244 |
| 6.2.5 | Synthèse | 246 |
| 6.3 | Analyse des réponses aux questionnaires d'appréciation | 246 |
| 6.3.1 | Questionnaire d'appréciation et de commentaire..... | 246 |
| 6.3.1.1. | Questions à échelle d'évaluation..... | 247 |
| 6.3.1.2. | Questions à choix de réponses | 255 |
| 6.3.1.3. | Questions à réponse courte..... | 256 |
| 6.3.2 | Synthèse | 261 |
| 6.4 | Interprétation des résultats et discussion..... | 262 |
| 6.4.1 | Efficacité pédagogique du CHINOINFO..... | 263 |
| 6.4.2 | Autres facteurs sur les résultats expérimentaux | 264 |
| 6.4.3 | Limites de l'expérimentation | 266 |
| Chapitre 7 | Conclusion..... | 268 |
| 7.1 | Synthèse des résultats de la thèse | 268 |
| 7.2 | Apports de la thèse | 272 |
| 7.3 | Perspectives de recherche..... | 273 |
| Références | bibliographiques | 275 |
| Annexe 1 | – Article <i>admiration</i> du LAF | 307 |
| Annexe 2 | – Article <i>programme</i> du DiCoInfo | 308 |
| Annexe 3 | – Liste de références de notre corpus..... | 309 |
| Annexe 4 | – Script Python pour nettoyer un texte de notre corpus..... | 320 |
| Annexe 5 | – Liste de CT extraits de notre corpus | 321 |
| Annexe 6 | – Liste de termes regroupés en vocables | 330 |
| Annexe 7 | – Exemples de paires de CLS contenant les polysèmes relevés..... | 334 |
| Annexe 8 | – Modèle d'encodage du CHINOINFO (exemple de l'article <small>ruan jian</small> 软件 ₁ « logiciel ₁ »)..... | 335 |
| Annexe 9 | – Fiche <small>ruan jian</small> 软件 « logiciel » en XML | 339 |
| Annexe 10 | – Exemple d'article du CHINOINFO..... | 340 |
| Annexe 11 | – Test lexical portant sur le lexique de l'informatique | 341 |

| | |
|--|-----|
| Annexe 12 – Questionnaire 1 : Profil des participants..... | 344 |
| Annexe 13 – Questionnaire 2a : Appréciations et commentaires (Groupe contrôle)..... | 346 |
| Annexe 14 – Questionnaire 2b : Appréciations et commentaires (Groupe expérimental)..... | 348 |
| Annexe 15 – Attestation d’approbation d’éthique | 351 |
| Annexe 16 – Formulaire d’information et de consentement..... | 352 |
| Annexe 17 – Renseignements sur le test lexical (présentation narrée destinée au GC)..... | 355 |
| Annexe 18 – Liste d’outils de référence proposés au GC | 359 |
| Annexe 19 – Renseignements sur le test lexical (présentation narrée destinée au GE)..... | 361 |
| Annexe 20 – Liste d’outils de référence proposés au GE | 367 |
| Annexe 21 – Exemple de vocable du CHINOINFO | 369 |
| Annexe 22 – Synthèse des termes reliés répertoriés dans le CHINOINFO | 370 |
| Annexe 23 – Synthèse des CLS répertoriées dans le CHINOINFO | 372 |
| Annexe 24 – Grille de correction des questions 1 à 9 du test lexical..... | 375 |
| Annexe 25 – Analyse des erreurs repérées dans le test lexical | 377 |

Liste des tableaux

| | |
|--|-----|
| Tableau 1. – Deux unités lexicales associées à la forme <small>bǎo cún</small> 保存 | 32 |
| Tableau 2. – La construction lexicale <small>ān zhuāng chéng xù</small> 安装程序 (‘installer -programme’) associée à deux structures syntaxiques différentes | 33 |
| Tableau 3. – Quatre combinaisons lexicales chinoises et leur équivalent français comportant le collocatif verbal <i>enregistrer</i> | 34 |
| Tableau 4. – Trois combinaisons lexicales contenant <i>transférer</i> et leur équivalent..... | 36 |
| Tableau 5. – Nuances sémantiques exprimées par la combinaison lexicale <i>transférer un fichier</i> et les combinaisons équivalentes chinoises..... | 36 |
| Tableau 6. – Les domaines abordés en COS d’après Li (2011, p. 111-112) | 55 |
| Tableau 7. – Les premiers manuels de COS portant sur le domaine du commerce et des affaires..... | 55 |
| Tableau 8. – Typologie des structures morphosyntaxiques en chinois..... | 65 |
| Tableau 9. – Structure actancielle de <small>bèi fèn</small> 备份 (« sauvegarde ») | 80 |
| Tableau 10. –Description de FL non standards observées dans des CLS..... | 86 |
| Tableau 11. –Équivalence entre le verbe français <i>envoyer</i> et deux verbes chinois accompagnés de compléments d’objet direct | 110 |
| Tableau 12. –Thèmes et sous-thèmes des textes composant le corpus..... | 114 |
| Tableau 13. –Variation géographique des termes chinois en informatique..... | 116 |
| Tableau 14. –Résumé des critères de sélection des textes composant notre corpus..... | 120 |
| Tableau 15. –Types de mots multicatégoriels relevés à partir des UT | 138 |
| Tableau 16. –Délimitation de quatre acceptions de l’UT <small>yùn xíng</small> 运行 | 140 |
| Tableau 17. –Analyse d’exemples d’usage de <small>fǎng wèn zī yuán</small> 访问资源 (« accéder à une ressource ») extraits du corpus | 152 |
| Tableau 18. –Regroupement des CLS renfermant un polysème dont la structure syntaxique est la même mais ayant des formes différentes | 154 |
| Tableau 19. –Regroupement des CLS renfermant un polysème qui ont deux formes distinctes et deux structures syntaxiques différentes..... | 155 |
| Tableau 20. –Regroupement des CLS renfermant un polysème qui ont une même forme et une même structure syntaxique..... | 156 |

| | |
|--|-----|
| Tableau 21. –Regroupement des CLS renfermant un polysème qui ont une même forme et deux structures syntaxiques distinctes | 157 |
| Tableau 22. –Comparaison des CLS de type « prédicat-COD » contenant <u>传输</u> ^{chuán shū} ₁ « transférer » et <u>迁移</u> ^{qiān yí} ₁ « faire migrer »..... | 161 |
| Tableau 23. –Représentation de la structure actancielle de <u>传输</u> ^{chuán shū} ₁ (« transférer ») et <u>转发</u> ^{zhuǎn fā} (« retransmettre »)..... | 162 |
| Tableau 24. –Représentation des liens sémantiques entre les trois acceptions du vocable <u>设置</u> ^{shè zhì} | 165 |
| Tableau 25. –Représentation des liens sémantiques au sein des CLS verbales contenant <u>更新</u> ^{gēng xīn} ₂ (« mise à jour ₁ ») et <u>更新</u> ^{gēng xīn} ₃ (« mise à jour ₂ »), | 167 |
| Tableau 26. –Regroupement des FL en classes intermédiaires et génériques (Jousse et al., 2011, p. 138)..... | 169 |
| Tableau 27. –Liste de classes génériques du CHINOINFO | 170 |
| Tableau 28. –Regroupement des CLS contenant <u>同步</u> ^{tóng bù} ₂ (« synchronisation ») en classes génériques et intermédiaires | 171 |
| Tableau 29. –Mise en parallèle de la description des CLS <u>安装</u> ^{ān zhuāng} ₁ <u>程序</u> ^{chéng xù} (« installer un programme ») et <u>安装</u> ^{ān zhuāng} ₂ <u>程序</u> ^{chéng xù} (« programme d’installation »). | 174 |
| Tableau 30. –Mise en parallèle de la description des CLS contenant <u>检测</u> ^{jiǎn cè} ₁ (« détecter ») et <u>检测</u> ^{jiǎn cè} ₂ (« mettre à l’essai »)..... | 177 |
| Tableau 31. –Description des termes reliés à <u>迁移</u> ^{qiān yí} ₁ (« faire migrer ») et à <u>传输</u> ^{chuán shū} ₁ (« transférer ») | 178 |
| Tableau 32. –Comparaison des liens sémantiques de trois termes ayant le sens d’ <i>en ligne</i> | 182 |
| Tableau 33. –Types de vocables dans le CHINOINFO..... | 210 |
| Tableau 34. –Catégorisation des renseignements recueillis au moyen du questionnaire 1 | 227 |
| Tableau 35. –Grille de correction des sections I et II du test lexical..... | 233 |
| Tableau 36. –Tableau d’analyse des erreurs repérées dans le test lexical | 241 |
| Tableau 37. –Catégorisation des données collectées au moyen des questions à échelle d’évaluation dans les questionnaires 2a et 2b | 247 |
| Tableau 38. –Catégorisation des données collectées au moyen des questions à choix | 255 |
| Tableau 39. –Catégorisation des données collectées au moyen des questions à réponse courte.. | 257 |

Liste des figures

| | |
|---|-----|
| Figure 1. – Unités linguistiques structurant la langue chinoise | 31 |
| Figure 2. – Carte mentale de l'évolution de la colère (début, milieu, fin) (Cavalla et al., 2014) .. | 45 |
| Figure 3. – Résultats de recherche portant sur le collocatif <i>a raudales</i> dans le dictionnaire DiCE | 52 |
| Figure 4. – Interface de la base ARTES montrant un accès aux collocations par fonctions discursives (<i>ARTES : Aide à la Rédaction de Textes Scientifiques [CLILLAC-ARP]</i> , s. d.) | 53 |
| Figure 5. – Extrait d'un texte spécialisé assorti d'une liste de mots nouveaux dans le manuel ^{zhōng} <u>医汉语综合教程</u> [<i>Chinois de la médecine traditionnelle chinoise : cours intégré</i>] (J. Xu, 2013, p. 3) | 59 |
| Figure 6. – Réseau lexical construit à partir du morphème ^{wǎng} « réseau » | 63 |
| Figure 7. – Description lexicale dans le manuel ^{kē jì hàn yǔ} <u>科技汉语</u> : ^{zhōng jí yuè dú jiào chéng} <u>中级阅读教程</u> [<i>Chinois scientifique et technique : manuel de lecture intermédiaire</i>] (An et Shan, 2005, p. 17) | 68 |
| Figure 8. – Description lexicale dans le manuel ^{fǎ lǜ hàn yǔ} <u>法律汉语</u> : ^{shāng shì piān} <u>商事篇</u> [<i>Chinois juridique : droit des affaires</i>] (T. Zhang, 2007, p. 106) | 68 |
| Figure 9. – Activité de mise en réseau des termes dans le manuel ^{zhōng yī hàn yǔ tīng shuō jiào chéng} <u>中医汉语听说教程</u> [<i>Chinois de la médecine traditionnelle chinoise : compréhension et expression orales</i>] (Jin et Guo, 2017, p. 2) | 69 |
| Figure 10. – Exercice portant sur l'antonymie dans le manuel ^{zhōng yī hàn yǔ zōng hé jiào chéng} <u>中医汉语综合教程</u> [<i>Chinois de la médecine traditionnelle chinoise : cours intégré</i>] (J. Xu, 2013, p. 4) | 70 |
| Figure 11. – Module d'apprentissage des lexies de forme identique dans le manuel ^{zhōng yī hàn yǔ zōng hé} ^{jiào chéng} <u>教程</u> [<i>Chinois de la médecine traditionnelle chinoise : cours intégré</i>] (J. Xu, 2013, p. 6) | 71 |
| Figure 12. – Explication grammaticale portant sur la formation des lexies composées dans le manuel ^{jī chǔ kē jì hàn yǔ jiào chéng} <u>基础科技汉语教程</u> : ^{yuè dú kè běn xià} <u>阅读课本 (下)</u> [<i>Chinois scientifique et technique élémentaire : compréhension écrite (volume 2)</i>] (Du, 2012b, p. 4-5) | 72 |
| Figure 13. – Répartition de nombre de textes par thème abordé dans notre corpus | 123 |
| Figure 14. – Répartition de volume de textes par thème abordé dans notre corpus | 123 |
| Figure 15. – Répartition de nombre de textes par genre textuel dans notre corpus | 124 |

| | |
|---|-----|
| Figure 16. –Répartition de volume de textes par genre textuel dans notre corpus | 124 |
| Figure 17. –Interface de recherche illustrant les paramètres de notre extraction terminologique | 128 |
| Figure 18. –Les 50 premiers CT extraits par l’outil <i>Keywords</i> (triés par score de spécificité) | 129 |
| Figure 19. –Extrait de la liste de concordances de l’UT ^{gōng xīn} 更新 générée par l’outil Concordance..... | 132 |
| Figure 20. –Extrait de la liste de cooccurrents de ^{gōng jī} 攻击 _{nom} (« attaque informatique ») générée par l’outil <i>Word Sketch</i> | 146 |
| Figure 21. –Formulation d’une requête à l’aide du CQL dans l’interface de <i>Concordance</i> | 148 |
| Figure 22. –Formulation d’une expression CQL dans l’interface de <i>CQL Builder</i> | 149 |
| Figure 23. –Formulation d’une expression CQL complexe dans l’interface de <i>Concordance</i> | 149 |
| Figure 24. –Extrait des résultats d’une recherche avancée dans l’interface de <i>Concordance</i> | 150 |
| Figure 25. –Requête formulée dans l’interface de <i>CQL Builder</i> pour extraire des contextes contenant la combinaison ^{fǎng wèn zī yuán} 访问资源 (« accéder à une ressource ») | 151 |
| Figure 26. –Extrait des contextes d’utilisation de la CLS ^{fǎng wèn zī yuán} 访问资源 (« accéder à une ressource ») | 151 |
| Figure 27. –Lien sémantique entre ^{cè shì} 测试 ₁ (« tester ») et les deux premières acceptions du vocable ^{jiǎn cè} 检测 (« détecter » et « mettre à l’essai ») | 160 |
| Figure 28. –Structuration hiérarchique des CLS rattachées à ^{tóng bù} 同步 ₂ (« synchronisation ») dans le CHINOINFO..... | 172 |
| Figure 29. –Liens sémantiques entre les éléments composants des CLS ^{zì dòng xià zài} 自动 ₁ ^{xià zài} 下载 ₂ (« téléchargement automatique ») et ^{zì dòng xià zài} 自动 ₂ ^{xià zài} 下载 ₁ (« télécharger automatiquement ») | 175 |
| Figure 30. –Nomenclature du dictionnaire CHINOINFO | 180 |
| Figure 31. –Affichage plié du fichier XML consacré au vocable ^{ruǎn jiàn} 软件 | 188 |
| Figure 32. –Extrait du fichier XML de l’article ^{ruǎn jiàn} 软件 ₁ (« logiciel ₁ ») | 189 |
| Figure 33. –Structure arborescente de la partie « combinaisons lexicales » du fichier XML de l’article ^{ruǎn jiàn} 软件 ₁ (« logiciel ₁ ») | 190 |
| Figure 34. –Affichage des renseignements généraux du terme ^{wǎng luò} 网络 ₁ (« réseau ₁ ») | 212 |
| Figure 35. –Ruban affichant « Définition » (l’onglet actif) et « Structure actancielle » (onglet inactif) de ^{bèi fèn} 备份 ₁ (« sauvegarder ») | 213 |

| | |
|---|-----|
| Figure 36. –Ruban affichant « Définition » (l’onglet inactif) et « Structure actancielle » (onglet actif) de ^{bèi fèn} 备份 ₁ (« sauvegarder ») | 213 |
| Figure 37. –Ruban affichant « Définition » (l’onglet inactif) et « Structure actancielle » (onglet actif) de ^{rù qīn} 入侵 ₁ (« s’introduire ») | 214 |
| Figure 38. –Ruban contenant « Contextes » de ^{jiā mì} 加密 ₂ (« cryptage ») | 215 |
| Figure 39. –Rubrique « Termes reliés » de l’article ^{wǎng luò} 网络 ₁ (« réseau 1 ») | 216 |
| Figure 40. –Rubrique « Combinaisons lexicales » de l’article ^{wǎng luò} 网络 ₁ (« réseau 1 ») | 218 |
| Figure 41. –Interface de recherche du CHINOINFO | 220 |
| Figure 42. –Renvois à deux articles du DiCoInfo dans l’article ^{rù qīn} 入侵 ₂ (« intrusion ») | 222 |
| Figure 43. –Renvois à deux articles du CHINOINFO dans l’article ^{qiān yí} 迁移 ₂ (« migration ») | 223 |
| Figure 44. –Répartition des participants par affiliation universitaire | 228 |
| Figure 45. –Répartition des participants par niveau et programme d’études | 229 |
| Figure 46. –Répartition des participants par nombre d’heures d’apprentissage du chinois | 230 |
| Figure 47. –Résultats du sondage sur les ressources utilisées par les participants pour résoudre les difficultés lexicales (Q8) | 231 |
| Figure 48. –Résultats du sondage sur les besoins d’apprentissage non comblés par les ressources lexicales (Q9) | 232 |
| Figure 49. –Notes des participants pour le test lexical (Q1 à Q9) | 235 |
| Figure 50. –Moyennes des notes obtenues par le GC et le GE (sections I à III du test lexical) | 236 |
| Figure 51. –Taux d’erreur du test lexical chez le GC et le GE | 237 |
| Figure 52. –Moyennes des notes obtenues par le GC et le GE (Q1 à Q9 du test lexical) | 239 |
| Figure 53. –Appréciation de l’organisation du test par le GC (Q1.1-1.7 du questionnaire 2a) | 248 |
| Figure 54. –Appréciation de l’organisation du test par le GE (Q1.1-1.7 du questionnaire 2b) | 249 |
| Figure 55. –Appréciation du niveau de difficulté du test par le GC (Q2.1-2.4 du questionnaire 2a). | 250 |
| Figure 56. –Appréciation du niveau de difficulté du test par le GE (Q2.1-2.4 du questionnaire 2b). | 250 |
| Figure 57. –Opinions du GC sur l’utilité des outils de référence consultés [Q5 du questionnaire 2a] | 251 |

| | |
|--|-----|
| Figure 58. –Opinions du GE sur l'utilité des outils de référence consultés [Q3 et Q10 du questionnaire 2b] | 253 |
| Figure 59. –Niveau de satisfaction du GC envers les outils de référence utilisés (Q7 du questionnaire 2a) | 254 |
| Figure 60. –Niveau de satisfaction du GE envers les outils de référence utilisés (Q4 du questionnaire 2b) | 254 |
| Figure 61. –Motifs de recours aux ressources lexicales autres que le CHINOINFO | 256 |
| Figure 62. –Problème d'affichage de l'interface du CHINOINFO et solution proposée | 259 |

Liste des sigles et abréviations

AE : analyse des erreurs
ARTES : Aide à la Rédaction de TExtes Scientifiques
CLIL : *Content and Language Integrated Learning*
CLS : combinaison lexicale spécialisée
COD : complément d'objet direct
COI : complément d'objet indirect
COS : chinois sur objectifs spécifiques
CQL : *Corpus query langage*
CT : candidat-terme
DAFA : Dictionnaire d'apprentissage du français des affaires
DEC : Dictionnaire explicatif et combinatoire
DiCo : Dictionnaire de Combinatoire
DiCoEnviro : Dictionnaire fondamental de l'environnement
DiCoInfo : Dictionnaire fondamental de l'informatique et d'Internet
DiCouèbe : Dictionnaire en ligne de combinatoire du français
ESP : *English for Specific Purposes*
EMILE : Enseignement d'une matière par l'intégration d'une langue étrangère
FL : fonction lexicale
FLE : français langue étrangère
FOS : français sur objectifs spécifique
FOU : français sur objectifs universitaires
GC : groupe contrôle
GE : groupe expérimental
GRELEP : Groupe de recherche en lexicographie pédagogique
L1 : langue première
L2 : langue seconde
LAF : Lexique actif du français : l'apprentissage du vocabulaire fondé sur 20 000 dérivations sémantiques et collocations du français
LEC : Lexicologie explicative et combinatoire
LIDILEM : Laboratoire de Linguistique et didactique des langues étrangères et maternelles
LSG : langue scientifique générale
LSP : langue de spécialité
LST : lexique scientifique transdisciplinaire
MAO : méthode audio-orale
MOOC : *Massive Open Online Course*
MTC : médecine traditionnelle chinoise
OLST : Observatoire de linguistique Sens-Texte
RL-fr : Réseau Lexical du Français
SG : sujet grammatical
TGT : théorie générale de la terminologie
TST : Théorie Sens-Texte
UP : unité phraséologique
UT : unité terminologique
XML : eXtensible Markup Language

Conventions d'écriture

Caractère chinois (*sinogramme*) : 网

Caractère chinois assorti de transcription phonétique (*pinyin*) : 网^{wǎng}

Sens d'un caractère ou d'une lexie : 'réseau'

Équivalent français d'un caractère : « réseau informatique, en ligne »

Combinaison de caractères assortie de transcription phonétique : 网站^{wǎngzhàn}

Sens littéral d'une combinaison de caractères : 'réseau-site'

Équivalent français d'une combinaison de caractères : « site Web »

Expression inacceptable : *无线和社交网络^{wú xiàn hé shè jiāo wǎng luò}

Distinction des sens associés à une même forme lexicale : 网络^{wǎng luò}₁, 网络^{wǎng luò}₂

Unité lexicale assortie d'une étiquette indiquant la partie du discours : 在线^{zài xiàn}_{1/adj}, 在线^{zài xiàn}_{2/adv}

À la mémoire de mon grand-père, Zhiqiang YU, qui m'a tant apporté

Remerciements

À tous ceux et celles qui m'ont aidée et encouragée de près ou de loin dans la réalisation de ce travail, j'exprime ma profonde reconnaissance.

Je tiens d'abord à témoigner toute ma gratitude à ma directrice de recherche, professeure Marie-Claude L'Homme, qui a guidé mes pas tout au long de ce chemin semé d'embûches.

Merci Marie-Claude, pour tout le soin que tu as pris à me diriger. Merci pour ta grande disponibilité, ta rigueur, ton souci de précision, ton encadrement, ton ouverture d'esprit et tes remarques constructives. Chaque fois je sortais de ton bureau après notre rencontre, je me sentais rassurée et prête à me remettre à l'ouvrage. Merci d'avoir partagé ta vaste expérience en recherche et en enseignement. Cela était une véritable mine d'or dont j'ai pleinement profité.

Merci d'avoir été un modèle de l'humanité pour moi. J'apprécie énormément ton écoute attentive, ta compréhension, tes sages conseils pour la vie, ton encouragement et ta bonne humeur contagieuse. Tu es l'incarnation parfaite de l'esprit de ces vers célèbres de la poésie chinoise :

已识乾坤大，犹怜草木青。 « Après avoir découvert la magnificence du monde, on contemple toujours la tendre verdure des jeunes pousses. »

J'adresse aussi mes remerciements aux membres du jury de mon examen de synthèse, professeur Patrick Drouin et professeure Chantal Gagnon. Leurs remarques et commentaires me sont précieux.

Je suis reconnaissante aux professeurs du Département de linguistique et de traduction. Les cours et les conférences auxquels j'ai eu l'occasion d'assister m'ont permis d'approfondir mes connaissances, et le plus important, de garder un esprit ouvert pour pousser plus loin ma réflexion.

Je suis extrêmement redevable envers Benoit Robichaud, membre de l'Observatoire de linguistique Sens-Texte (OLST) qui a apporté son expertise à différentes étapes de la réalisation de ce projet. Je le remercie chaleureusement pour sa contribution généreuse à l'élaboration du dictionnaire CHINOINFO et pour sa grande disponibilité à répondre à mes questions.

J'en profite aussi pour remercier les étudiants qui ont participé à l'expérimentation de mon projet. Leur collaboration m'a permis de recueillir des données empiriques précieuses pour réaliser cette thèse visant à bonifier concrètement la pédagogie du lexique spécialisé.

J'exprime aussi ma gratitude à Pascale Bellemare, notre bibliothécaire, et à Clémence Bourget, notre technicienne en gestion des études, qui ont grandement facilité mon parcours universitaire.

Mes remerciements vont aussi à mes collègues de l'Observatoire de Linguistique Sens-Texte que j'ai eu l'occasion de côtoyer. Bien que leur quotidien fût très différent du mien, nous avons eu de petites causeries agréables et inspirantes lors de mon passage au labo.

Je suis reconnaissante au Département de linguistique et de traduction et à la Faculté des études supérieures de l'Université de Montréal pour les bourses d'études qu'ils m'ont octroyées durant ma scolarité. Ces bourses m'ont permis de me consacrer pleinement à ma recherche.

Je remercie ma famille pour le soutien inconditionnel qu'elle m'a apporté.

Merci, papa et maman, pour votre dévouement sans bornes. Vous m'avez encouragée à aller au bout de mes rêves. Vous m'avez appris à être généreuse, authentique et respectueuse envers les autres et moi-même. Grâce à vous, je ne perds jamais de vue que ces valeurs sont bien plus importantes que les succès professionnels.

Merci Yue, mon mari bien-aimé, pour ta compréhension des hauts et des bas dans la réalisation de ce projet de thèse. Merci d'avoir apporté tant de doses de bonne humeur qui m'ont permis de refaire le plein d'énergie et de retourner au travail le cœur léger.

Merci Julien, mon cher fils. Tu n'avais que 16 mois quand j'ai entamé cette aventure extraordinaire. Tu m'as inspirée à me dépasser et à rester exigeante envers moi-même tout le temps. Ton regard pétillant, tes beaux sourires, tes bisous et tes câlins me rendent courageuse et imbattable.

Enfin, je dédie ce modeste travail à mon regretté grand-père, Zhiqiang Yu, qui m'a transmis sa passion pour la langue et la culture chinoises dès mon jeune âge, qui m'a cultivé le goût de l'excellence et qui a été pour moi un modèle exceptionnel de persévérance et de détermination. Cher grand-papa, veille sur moi.

Chapitre 1 Introduction

L'enseignement-apprentissage du lexique spécialisé chinois est un chemin semé d'obstacles pour les apprenants non natifs. La particularité structurelle de la langue chinoise engendre des problèmes lexicaux dont la résolution soulève des défis importants pour les apprenants non natifs. Les difficultés auxquelles ils se heurtent sont situées respectivement au niveau des morphèmes, des lexies et des combinaisons lexicales. Dans le cadre de cette thèse, nous nous focalisons sur les combinaisons lexicales qui engendrent des difficultés sur le plan tant sémantique que syntaxique.

1.1 Contexte et problématique

Notre recherche s'inscrit dans le contexte de l'enseignement du lexique en chinois sur objectifs spécifiques (désormais COS).

Il convient de préciser, d'entrée de jeu, que le chinois est une macrolangue composée d'une dizaine de dialectes, dont le cantonais, le shanghaien, le hakka et le teochow. Ces dialectes sont parlés dans différentes régions de la Chine. La langue parlée par la majorité de la population chinoise est le chinois standard (普通话), imposé par le gouvernement de la République populaire de Chine dans les années 1950. Quant à l'écriture, la langue chinoise dispose de deux systèmes d'écriture : le chinois simplifié (简体中文) et le chinois traditionnel (繁体中文). Le chinois simplifié est utilisé en Chine continentale. Le chinois traditionnel est en usage à Hong Kong, à Macao et à Taïwan. En COS, on enseigne le chinois standard qui s'écrit en chinois simplifié.

Le COS est par nature une langue de spécialité (LSP). Il a fait son apparition en Chine dans les années 1980 marquées par une explosion des échanges internationaux du pays dans divers secteurs d'activités, d'où la nécessité d'adapter la formation du chinois aux besoins de communication dans des domaines de spécialité. De nos jours, l'enseignement des LSP constitue un sujet d'actualité qui est au cœur des préoccupations didactiques, comme le montre un numéro récent de la Revue de linguistique et de didactique des langues (Carras et Álvarez Martínez, 2022). On recense des recherches abordant cette question du point de vue de la constitution des

ressources didactiques, de l'analyse des propriétés linguistiques des LSP et de la transposition des LSP en objet d'enseignement.

Notre thèse est axée sur l'enseignement du lexique en COS auprès d'apprenants non natifs. Dans ce qui suit, nous donnons d'abord un aperçu des obstacles majeurs à l'acquisition du lexique spécialisé chinois. Nous exposons ensuite les lacunes en matière de description lexicale dans le matériel didactique en usage, pour formuler la problématique de notre recherche.

1.1.1 Enseignement du lexique spécialisé chinois : défis à relever

Notre expérience d'enseignement¹ nous a permis de repérer trois types de difficultés d'apprentissage auxquelles font face les locuteurs non sinophones : 1) la distinction des lexies de forme identique; 2) la distinction des combinaisons lexicales de forme identique et 3) le choix de cooccurrents pour une lexie.

Ces difficultés sont liées à différents éléments qui structurent la langue chinoise. Celle-ci se distingue des langues indo-européennes par l'unité minimale qui la structure, soit le *caractère*, nommé aussi *sinogramme*. Chaque caractère est associé à une forme graphique et correspond à au moins un son monosyllabique. Les caractères forment les *morphèmes*. Le morphème est l'unité minimale de sens en chinois, composée d'un seul caractère porteur de sens (par exemple, 电 ^{diàn} 'électricité' « électricité », 网 ^{wǎng} 'réseau' « réseau informatique/en ligne ») ou de plusieurs caractères (par exemple, 玻璃 ^{bō li} 'verre' « verre », 赫兹 ^{hè zī} 'hertz' « hertz », 麦克风 ^{mài kè fēng} 'microphone' « microphone »²). À partir des morphèmes sont construites les *unités lexicales* ou *lexies*. Une lexie est l'unité minimale qui véhicule un sens lexical et fonctionne de façon autonome en

¹ Nous enseignons le chinois à l'Université Laval depuis 2015. Nous prenons en charge les cours élémentaire et intermédiaire, destinés aux apprenants débutants qui n'ont aucune ou peu de connaissances du chinois. La grande majorité d'entre eux sont francophones. Leur niveau d'apprentissage est validé par un test de classement. À titre de chargée de cours, nos tâches consistent à planifier des activités pédagogiques, à élaborer des modules d'apprentissage multimédia, à organiser des séances d'animation culturelle, à corriger les travaux d'étudiants et à offrir des tutorats individuels ou en groupe.

² Tous les caractères chinois ne sont pas porteurs de sens. Dans les morphèmes qui correspondent à une combinaison de caractères, chaque caractère constitutif n'a de sens que dans cette combinaison. Nous avons donné ici trois exemples pour illustrer ce phénomène : le premier est un morphème qui résulte de la combinaison de deux caractères, les deux autres exemples sont les cas de transcriptions phonétiques de termes étrangers.

chinois. Elle peut être composée d'un seul morphème (开 'ouvrir' « ouvrir, allumer », 关 'fermer' « fermer, éteindre ») ou d'une combinaison de morphèmes (par exemple, 硬盘 'dur-disque' « disque dur », 电脑 'électricité-cerveau' « ordinateur »). Précisons que les lexies composées forment la plus grande partie du lexique du chinois. Le groupement des lexies peut donner lieu à des combinaisons lexicales (par exemple, 浏览网页 'survoler-voir-réseau-page' « visiter une page Web »). La figure 1 illustre la spécificité structurelle de la langue chinoise.

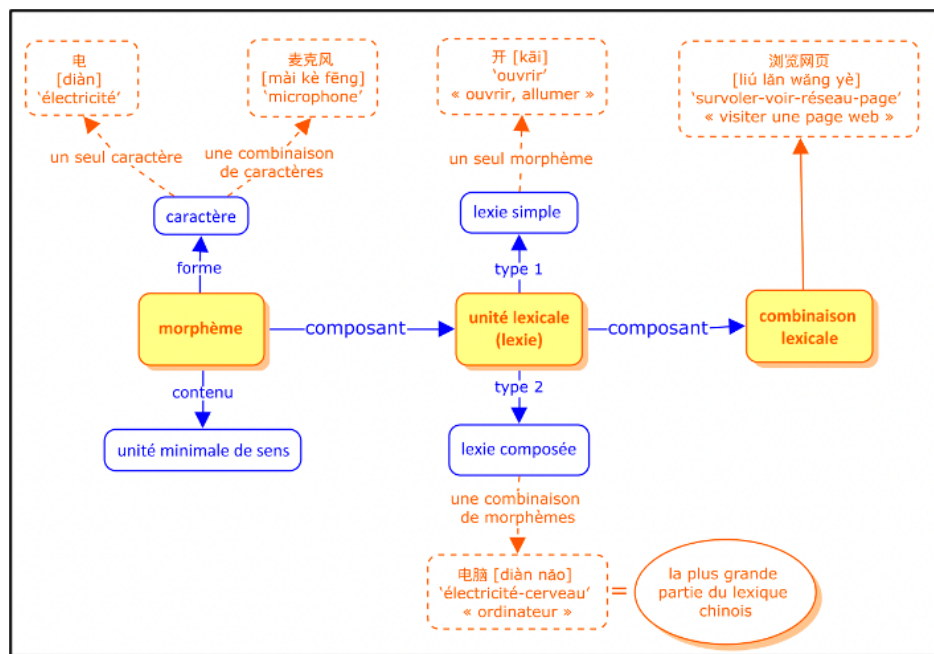


Figure 1. – Unités linguistiques structurant la langue chinoise

Dans cette recherche, nous nous focalisons sur les difficultés engendrées par les lexies et les combinaisons lexicales. Ces unités linguistiques sont dotées d'une autonomie sémantique et syntaxique en chinois. Nous allons explorer leur fonctionnement dans les textes spécialisés afin de proposer des solutions aux difficultés qu'elles soulèvent.

En premier lieu, la distinction sémantique des lexies de forme identique constitue une difficulté d'apprentissage importante. En chinois, certaines lexies ayant la même forme relèvent de

différentes parties du discours et entretiennent entre elles un lien de *dérivation sémantique*³. Le tableau 1 en montre un exemple.

| | Unité lexicale | Transcription phonétique | Sens littéral | Partie du discours | Équivalent français |
|------|----------------|--------------------------|---------------------|--------------------|---------------------|
| (1a) | 保存 | [bǎo cún] | ‘préserver-déposer’ | verbe | « enregistrer » |
| (1b) | 保存 | [bǎo cún] | ‘préserver-déposer’ | nom | « enregistrement » |

Tableau 1. – Deux unités lexicales associées à la forme ^{bǎo cún} 保存

Les deux lexies correspondent à la même forme, soit ^{bǎo cún} 保存 qui signifie littéralement ‘préserver-déposer’. La lexie (1a) est de nature verbale ; (1b) est un nom désignant l’action exprimée par (1a). Dans les outils de référence et les ressources d’apprentissage, la description de ^{bǎo cún} 保存 (‘préserver-déposer’) est souvent réduite à une énumération des sens qui y sont associés.

De plus, les lexies de forme identique peuvent se combiner avec une même lexie dans différentes combinaisons lexicales. Par exemple, les lexies (1a) et (1b) forment deux combinaisons lexicales avec ^{wén jiàn} 文件 (« fichier »), soit ^{bǎo cún} 保存_{1a}^{wén jiàn} 文件 (‘enregistrer-fichier’ « enregistrer un fichier ») et ^{wén jiàn} 文件^{bǎo cún} 保存_{1b} (‘fichier-enregistrer’ « enregistrement d’un fichier »), ce qui rend la tâche de distinction sémantique encore plus ardue.

En deuxième lieu, les apprenants sont confrontés à l’ambiguïté syntaxique des combinaisons lexicales. Cette difficulté découle du fait qu’un groupement lexical peut être associé à plus d’une structure syntaxique et véhiculer ainsi des sens différents. Examinons les deux combinaisons lexicales présentées dans le tableau 2.

³ La notion de *dérivation sémantique*³ est proposée par Mel’čuk et Polguère (2007, p. 18). Nous l’abordons dans la sous-section 3.1.1.

| | | |
|--------------------------------|---|------------------------------------|
| (2a) | ān zhuāngchéng xù 安 装 程 序 | |
| Unités lexicales constitutives | ān zhuāng 安 装 | chéng xù 程 序 |
| Sens littéral et signification | ‘Installer-équiper’ « Installer » | ‘Procédure-ordre’ « Programme » |
| Fonction syntaxique | Prédicat syntaxique | Complément d’objet direct |
| Partie du discours | Verbe | Nom |
| Équivalent français | « Installer un programme » | |
| (2b) | ān zhuāngchéng xù 安 装 程 序 | |
| Unités lexicales constitutives | ān zhuāng 安 装 | chéng xù 程 序 |
| Sens littéral et signification | ‘Installer-équiper’ « Installation » | ‘Procédure-ordre’ « Programme » |
| Fonction syntaxique | Modificateur | Modifié |
| Partie du discours | Nom | Nom |
| Équivalent français | « Programme d’installation » | |

Tableau 2. – La construction lexicale 安 装 程 序 (‘installer-programme’) associée à deux structures syntaxiques différentes

Les groupements lexicaux (2a) et (2b) sont constitués de deux composants de formes identiques. Le (2a) signifie « installer un programme ». 程 序 (« programme ») remplit la fonction syntaxique de complément d’objet direct (COD) du verbe 安 装 (« installer »). Le (2b) signifie « programme installeur ». 程 序 (« programme ») est modifié syntaxiquement par le nom 安 装 (« installation »), celui-ci étant le dérivé nominal du verbe 安 装 (« installer ») ayant la même forme. Dans ce cas, on parle d’un programme qui sert à installer un logiciel sur le disque dur d’un ordinateur.

Le choix de collocations apporte aussi son lot de défis. Une combinaison lexicale spécialisée (CLS) (L’Homme et Gemme, 1997; L’Homme et Meynard, 1998; Meynard, 1998; L’Homme et Bertrand, 2000; L’Homme, 2000, 2003, 2017) relie deux éléments constitutifs, nommés la *base* et le *collocatif* (Hausmann, 1979). La base, soit l’élément central de la combinaison, est sémantiquement autonome et détermine le choix de son collocatif. Cette association s’explique non seulement par une affinité syntaxico-sémantique, mais aussi par des usages préférentiels dans un domaine spécialisé.

Les apprenants non natifs ont souvent de la difficulté à trouver un collocatif approprié pour une lexie donnée. Sous l’influence de leur langue première (L1), ils ont tendance à associer une lexie

chinoise à un collocatif qui semble être ‘emprunté’ de leur L1. Considérons l’exemple reproduit dans le tableau 3.

| Combinaisons lexicales chinoises | | Équivalent français | |
|---|---|---------------------|------------------------|
| Collocatif | Base | Collocatif | Base |
| bǎo cún 保存 ‘préservier-déposer’ « enregistrer » | wén jiàn 文件 ‘texte-pièce’ « fichier » | Enregistrer | un fichier |
| zhù cè 注册 ‘inscrire-registre’ « inscrire » | yù míng 域名 ‘domaine-nom’ « nom de domaine » | | un nom de domaine |
| lù zhì 录制 ‘filmer-produire’ « filmer » | shì pín huì yì 视频会议 ‘voir-fréquence-rencontre-discussion’ « vidéoconférence » | | une vidéoconférence |
| jì lù 记录 ‘noter-inscrire’ « consigner » | jiàn rù 键入 ‘taper-entrer’ « frappes au clavier » | | les frappes au clavier |

Tableau 3. – Quatre combinaisons lexicales chinoises et leur équivalent français comportant le collocatif verbal *enregistrer*

Les quatre CLS chinoises contiennent chacune un collocatif verbal distinct. Le choix de ces collocatifs est déterminé par la lexie de base (文件 « fichier », 域名 « nom de domaine », 视频会议 « vidéoconférence », 键入 « frappes au clavier »). Les apprenants francophones sont susceptibles de confondre ces collocatifs qui sont tous traduits par *enregistrer* dans les CLS équivalentes françaises. Il s’agit d’une source d’erreur courante que nous avons observée dans notre pratique d’enseignement.

En résumé, différents obstacles se dressent devant les apprenants non natifs sur leur chemin d’apprentissage de la langue chinoise, qu’il s’agisse du chinois langue générale ou du chinois langue de spécialité. Il importe de les prendre en compte dans une perspective didactique, afin d’assister les apprenants dans la résolution de problèmes lexicaux auxquels ils se heurtent.

1.1.2 Méthodes didactiques du lexique

Malgré les nombreux défis à relever en enseignement-apprentissage du lexique spécialisé chinois, le matériel didactique en usage⁴ offre rarement des solutions concrètes aux difficultés décrites dans la section précédente.

⁴Dans le cadre de notre thèse, nous nous intéresserons aux ressources didactiques élaborées depuis les deux dernières décennies, dont Li et Ding (2002), An et Shan (2005, 2008), Zhang (2007), Li (2008a, 2008b, 2008c, 2008d), Bai (2014), Du (2011a, 2011a, 2011b, 2012b, 2012a), Han (2012), Xu et al. (2012), Xu (2013), Jin et Guo (2017), Lu et al. (2017), et Zhang (2014). Ces manuels

Premièrement, le matériel didactique accorde peu d'attention au fonctionnement linguistique des unités lexicales. La description du vocabulaire prend la forme d'une liste de lexies associées à un ou plusieurs équivalents dans une autre langue. Rares sont les tentatives de représentation des propriétés sémantique et combinatoire des lexies. Par exemple, la description de 网站^{wǎngzhàn} ('réseau-site' « site Web »), dans un bon nombre de ressources d'apprentissage consacrées au lexique, se limite à présenter son équivalent anglais, sans prendre en compte ni ses combinaisons lexicales (访问网站^{fǎng wèn wǎngzhàn} « accéder à un site Web », 运营网站^{yùn yíng wǎngzhàn} « gérer un site Web », 恶意网站^{è yì wǎngzhàn} « site Web malveillant »), ni les lexies sémantiquement proches (网址^{wǎng zhǐ} « adresse de site Web », 网页^{wǎng yè} « page Web », 网民^{wǎng mǐn} « internaute »). Pour les apprenants non natifs, notamment les débutants, une telle méthode descriptive ne permet qu'un décodage rapide.

À cela s'ajoutent des lacunes observées dans les explications données dans une autre langue que le chinois. La présentation du vocabulaire, réduite à un simple répertoire bilingue, associe souvent plusieurs équivalents partiels à une forme lexicale donnée, ce qui constitue une source de confusion pour les apprenants. Nous identifions deux points faibles à cet égard dans le matériel didactique élaboré.

Premièrement, nous constatons des lacunes quant à la distinction sémantique de lexies de forme identique. Par exemple, la forme 应用^{yìngyòng} ('répondre-utiliser') correspond à deux lexies nominales. La première lexie désigne un type de logiciel qui sert à accomplir une tâche ou une activité particulière. Elle donne lieu à des combinaisons telles que 应用软件^{yìngyòng ruǎn jiàn} (« logiciel d'application ») et 应用商店^{yìngyòng shāng diàn} (« magasin d'applications »). La deuxième lexie renvoie à l'utilisation de moyens informatiques pour répondre à un besoin spécifique. Elle figure dans les combinaisons telles que 应用场景^{yìngyòng chǎng jǐng} (« scénario d'application ») et 应用模式^{yìngyòng mó shì} (« mode d'application »). Dans le matériel didactique en usage, la forme lexicale 应用^{yìngyòng} ('répondre-utiliser') ne fait l'objet que d'une seule entrée. Elle est associée, soit à un seul équivalent partiel, soit à deux équivalents qui

constituent un échantillon représentatif du matériel didactique existant en COS. Du point de vue de la matière enseignée, ces manuels portent sur les domaines de spécialité les plus abordés en COS. En outre, la conception de ces manuels est documentée. On recense des recherches traitant des méthodes utilisées pour élaborer ces ressources d'apprentissage didactiques (Lu, 2006; Shan, 2008; Shan et An, 2009; Zhang, 2014). Nous présentons les ouvrages recensés dans les sous-sections 2.3.1 et 2.3.3.

correspondent chacun à l'un des deux sens de 应用^{yīngyòng}. Aucune autre précision n'est donnée sur le fonctionnement de ces lexies.

Deuxièmement, l'absence de traitement de la combinatoire lexicale confronte les apprenants aux difficultés liées au choix de collocatif, comme nous l'avons expliqué dans la sous-section 1.1.1 (voir le tableau 3). Le tableau 4 donne un autre exemple de ce type de difficulté.

| Combinaison française | | Combinaison équivalente chinoise | |
|-----------------------|--------------------------|---|--|
| Transférer | des données | ^{chuán shū} 传输 'faire passer-transporter' | ^{shù jù} 数据 'chiffres-preuve' « données » |
| | un courrier électronique | ^{zhuǎn fā} 转发 'transmettre-envoyer' | ^{yóu jiàn} 邮件 'poste-pièce' « courrier électronique » |
| | un appel | ^{zhuǎn yí} 转移 'transmettre-déplacer' | ^{hū jiào} 呼叫 'interpeller-appeler' « appel » |

Tableau 4. – Trois combinaisons lexicales contenant *transférer* et leur équivalent

Dans le domaine de l'informatique, le verbe transférer constitue l'équivalent de la lexie 传输^{chuán shū} lorsqu'il se combine avec données. Ce verbe français peut véhiculer d'autres sens et correspondre ainsi à d'autres verbes chinois dans différentes combinaisons lexicales, telles que transférer un courrier électronique et transférer un appel. Le choix du collocatif verbal est déterminé par le terme de base (courrier électronique et appel), ce qui donne du fil à retordre pour les apprenants non natifs.

Le contexte d'utilisation n'est non plus pris en considération dans la description lexicale. Reprenons le verbe transférer comme exemple. La combinaison transférer un fichier exprime des nuances sémantiques en fonction de son environnement linguistique (tableau 5). Les apprenants peuvent avoir de la difficulté à trouver l'expression la plus appropriée pour exprimer un sens spécifique.

| Combinaison française | Nuances sémantiques | Combinaison équivalente chinoise | |
|-----------------------|---|--|--|
| Transférer un fichier | d'un support électronique à un autre | ^{chuán shū} 传输 'faire passer-transporter' | ^{wén jiàn} 文件 'texte-pièce' « fichier » |
| | à un destinataire | ^{zhuǎn fā} 转发 'transmettre-envoyer' | |
| | pour le sauvegarder dans un autre support | ^{zhuǎn cún} 转存 'transmettre-sauvegarder' | |

Tableau 5. – Nuances sémantiques exprimées par la combinaison lexicale *transférer un fichier* et les combinaisons équivalentes chinoises

En résumé, l'analyse d'un nombre de manuels de COS nous permet de constater que le matériel didactique présente certaines lacunes à combler en ce qui concerne la description des propriétés sémantique et combinatoire des lexies. Nous examinerons plus attentivement les manuels recensés à la sous-section 2.3.3.

1.1.3 Problématique

Dans notre recherche, nous prenons comme objet d'étude les combinaisons lexicales spécialisées (CLS). Cette notion est mise de l'avant dans les travaux terminologiques précédents (L'Homme et Gemme, 1997; L'Homme et Meynard, 1998; Meynard, 1998; L'Homme et Bertrand, 2000; L'Homme, 2000, 2003, 2017). Une CLS est l'association d'un terme, défini comme la base du groupement, et d'un autre terme ou unité lexicale de nature générale, défini(e) comme le collocatif. Ce groupement lexical est conventionnel, c'est-à-dire il correspond à un usage spécifique dans un domaine de spécialité (L'Homme, 2017, p. 220), comme nous l'avons mentionné dans la sous-section 1.1.1.

Dans le cadre de notre thèse, la notion de CLS renvoie à deux catégories de combinaisons lexicales : 1) les collocations⁵, à savoir les cooccurrences lexicales restreintes composées d'un terme et de son collocatif, par exemple, la CLS ^{kāi yuán ruǎn jiàn} 开源软件 ('source ouverte-logiciel' « logiciel libre ») reliant le terme ^{ruǎn jiàn} 软件 (« logiciel ») à son collocatif ^{kāi yuán} 开源 (« source libre »); 2) les liens observés dans les structures actanciennes de termes prédicatifs. Le sens d'une lexie prédicative est exprimé par un prédicat sémantique et les actants sémantiques (Mel'čuk et al., 1995, p. 76). Un prédicat sémantique dénote un fait qui implique nécessairement des participants, tels un événement, un processus, un état, une propriété, une relation, etc. Les actants sémantiques d'une lexie prédicative sont les sens participant au fait dénoté par le prédicat sémantique. La structure actancielle désigne la « micro-structure sémantique associée au prédicat (sémantique) dans le lexique » (Mel'čuk et Polguère, 2008 paragraphe 4). Par exemple, la CLS ^{rù qīn jīn cè} 入侵检测 ('intrusion-détection' « détection d'intrusion ») relie le terme prédicatif ^{jīn cè} 检测 (« détection ») à la lexie ^{rù qīn} 入侵 (« intrusion ») qui réalise son deuxième actant sémantique.

⁵ Nous aborderons la notion de collocation dans la sous-section 2.2.1.

Du point de vue didactique, l'exploitation de ces deux types de CLS est intéressante à plus d'un titre. D'une part, la description des collocations restreintes met en évidence le lien d'affinité entre les composantes de ces groupements lexicaux sur le plan sémantique et syntaxique (Grossmann et Tutin, 2003 ; Tutin, 2013 ; Tutin et al., 2006). Elle sensibilise aussi les apprenants au fait que ces combinaisons lexicales sont soumises aux contraintes imposées par les usages préférentiels du domaine abordé (L'Homme, 2017). D'autre part, l'étude de la deuxième catégorie de CLS vise à rendre explicite le lien entre un terme prédicatif et ses actants sémantiques, à savoir les éléments impliqués dans le sens du terme.

La prise en compte de la combinatoire spécialisée exige un examen plus approfondi du comportement linguistique des termes, comme le souligne l'Homme (2017, p. 220), et plus précisément, des propriétés sémantique et combinatoire des termes. Pour pallier les lacunes observées dans le matériel didactique de COS à cet égard, nous nous attachons à explorer de quelle manière une méthode de description lexicale basée sur une représentation sémantique et syntaxique assiste les apprenants dans la résolution des problèmes lexicaux soulevés par les CLS.

1.2 Objectifs

Notre recherche a pour objectif général de concevoir une méthode de description de la combinatoire spécialisée en vue de la résolution de difficultés lexicales par les locuteurs non sinophones. Dans le cadre de cette thèse, nous nous intéressons aux apprenants provenant d'autres pays que ceux de la sinosphère (dont le Japon, la Corée du Sud, la Corée du Nord, le Singapour et la Malaisie), plus spécifiquement, aux apprenants francophones.

Notre recherche vise aussi à appliquer la méthode à l'élaboration du dictionnaire CHINOINFO, une ressource lexicale chinois-français portant sur le domaine de l'informatique. Cette ressource s'adresse aux apprenants francophones du chinois.

L'objectif secondaire de notre thèse consiste à évaluer l'efficacité de la ressource élaborée. Notre analyse s'appuie sur les données collectées au moyen d'une expérimentation comparative auprès

des apprenants francophones débutants⁶ qui suivent des cours de chinois dans une université au Québec ou en Chine.

1.3 Hypothèses

Notre recherche s’articule autour de trois hypothèses portant respectivement sur les difficultés d’apprentissage des CLS, les solutions aux difficultés identifiées et les facteurs favorisant l’acquisition des connaissances lexicales. Partant de ces hypothèses, nous nous attachons à élaborer une méthode de description lexicale qui aide les apprenants à résoudre les difficultés soulevées par les CLS au cours de leur apprentissage du lexique spécialisé chinois, tel est l’objectif général de notre thèse.

1. La résolution des difficultés liées à la distinction sémantique des CLS et à l’interprétation syntaxique des celles-ci représente un défi de taille pour les apprenants non natifs.
2. La représentation sémantique des CLS et l’explication de leur structure syntaxique assistent les apprenants non natifs dans la résolution de difficultés lexicales.
3. La présentation structurée et intuitive des connaissances lexicales favorise leur assimilation par les apprenants.

1.4 Originalité de la thèse

Notre recherche se distingue des travaux précédents par le contexte de recherche, le problème ciblé, la prise en compte des propriétés syntaxico-sémantiques du lexique chinois dans sa description, ainsi que le choix méthodologique pour collecter des données empiriques.

D’abord, notre travail aborde l’enseignement du lexique spécialisé du domaine de l’informatique dans le cadre du COS. La méthode de description lexicale que nous élaborons vise à répondre aux besoins spécifiques des apprenants francophones en milieu universitaire.

⁶ Dans le cadre de ce travail, les apprenants débutants désignent les apprenants qui font preuve d’un niveau élémentaire de connaissance du chinois selon la norme d’évaluation de l’État chinois (Le ministère de l’Éducation de la République populaire de Chine et Commission d’État pour la langue et l’écriture chinoises, 2021). Nous aborderons les niveaux de maîtrise de la langue chinoise dans la sous-section 4.3.1.2.

Dans ce contexte bien circonscrit, notre recherche cible la combinatoire spécialisée. Même si les collocations en chinois ont commencé à faire l'objet de travaux descriptifs (X. Zhao, 2018; Zhiwei Han et L'Homme, 2022), cela reste un phénomène marginal. En COS, l'enseignement de la combinatoire spécialisée a occupé jusqu'ici une place périphérique, contrairement à ce qui arrive en français où on recense un bon nombre de travaux sur ce sujet. Par ailleurs, on aborde rarement les CLS dans les ressources didactiques. Pour combler cette lacune, nous nous attachons à proposer une méthode descriptive permettant de résoudre les difficultés syntaxico-sémantiques soulevées par les CLS. En cela, notre travail est susceptible d'apporter une contribution inédite à l'enseignement du lexique en COS.

De plus, la méthode descriptive que nous proposons s'adapte aux particularités syntaxico-sémantiques du lexique spécialisé chinois. Outre la représentation des propriétés sémantiques des termes, soit les lexies véhiculant un sens relevant d'un domaine de spécialité, nous décrivons également la structure syntaxique des CLS. Cette méthode vise à sensibiliser les apprenants non natifs à la structure linguistique du chinois qui régit le fonctionnement de la combinatoire spécialisée.

Enfin, nous avons recours à une méthode expérimentale pour collecter des données auprès d'apprenants francophones du chinois. Ce choix méthodologique, inspiré du travail d'Alipour (2018), vise à tracer un portrait de l'intérêt que présente notre méthode de description dans un contexte d'enseignement.

Pour conclure, notre recherche, comparée à des travaux antérieurs sur la didactique du lexique spécialisé, s'inscrit dans un contexte particulier, soit l'enseignement du lexique spécialisé chinois visant les apprenants non natifs. Nous explorons un aspect qui demeure jusqu'ici peu exploré en didactique du COS, soit la combinatoire spécialisée. Notre recherche se caractérise enfin par le recours aux données empiriques recueillies au moyen d'expérimentation.

1.5 Structure de la thèse

Outre cette introduction, la suite de notre thèse s'organise de la façon suivante.

Le chapitre 2 dresse un état de l'art des travaux sur l'enseignement du lexique spécialisé. Nous recensons les recherches axées sur trois thèmes qui sont en rapport avec notre problématique (2.1), notre objet d'étude (2.2) et le contexte de recherche dans lequel s'inscrit notre travail (2.3).

Le chapitre 3 présente notre cadre théorique qui comprend trois volets, axés respectivement sur la lexicologie (3.1), la terminologie (3.2) et la didactique des langues secondes (3.3). Les deux premiers volets fournissent des repères théoriques et méthodologiques de la description du lexique spécialisé. Le volet didactique aborde la transmission des connaissances lexicales.

Le chapitre 4 expose en détail notre démarche méthodologique qui se divise en trois phases. Nous abordons, dans un premier temps, la compilation d'un corpus spécialisé ainsi que la collecte et l'analyse des CLS à partir du corpus (4.1). Nous expliquons par la suite la conception de la méthode de description lexicale et la création de la ressource CHINOINFO (4.2). À la fin du chapitre, nous décrivons la conception de notre expérimentation comparative et la collecte de données auprès d'apprenants francophones (4.3).

Le chapitre 5 est consacré à la description du CHINOINFO. Nous donnons d'abord un aperçu du contenu du dictionnaire, y compris la nomenclature (5.1) et la structure des renseignements descriptifs au sein des articles (5.2). Nous expliquons ensuite les fonctionnalités de la ressource (5.3).

Le chapitre 6 aborde le traitement des données expérimentales ainsi que les résultats que nous avons obtenus. Nous décrivons notre démarche pour analyser les renseignements sur le profil des participants (6.1), leurs réponses à un test de compétences lexicales (6.2) ainsi que leurs commentaires sur le test et les ressources lexicales auxquelles ils ont recours au cours du test (6.3). À la lumière des résultats de notre analyse, nous discutons de l'apport pédagogique du CHINOINFO ainsi que des limites de notre expérimentation (6.4).

Le chapitre 7 présente une conclusion générale sur notre thèse. Nous récapitulons les grandes lignes des chapitres précédents et formulons des remarques conclusives sur les résultats obtenus en regard de nos objectifs de recherche et nos hypothèses initiales, ce qui nous permet de répondre à notre problématique. Nous abordons ensuite les contributions de notre recherche et proposons enfin des pistes de réflexion qui pourraient être approfondies dans le futur.

Chapitre 2 État de la question

Ce chapitre dresse un état de l'art de la didactique du lexique spécialisé. Nous passons en revue les travaux antérieurs portant sur les thèmes suivants : l'enseignement du lexique basé sur une approche sémantique (section 2.1) ; l'étude de la combinatoire spécialisée dans une perspective didactique (section 2.2) et l'enseignement du lexique en chinois sur objectifs spécifiques (COS) (section 2.3). Nous faisons enfin une synthèse de la revue de littérature pour situer notre recherche par rapport aux travaux recensés (section 2.4).

2.1 Enseignement du lexique dans une perspective sémantique

Cette section présente les travaux en didactique du lexique adoptant une approche sémantique. Nous examinons le traitement sémantique du lexique dans différents cadres applicatifs, notamment le développement de dispositifs d'enseignement et la création de ressources didactiques. Les deux sous-sections abordent respectivement l'enseignement du lexique en langue générale (2.1.1) et en langues de spécialité (2.1.2).

2.1.1 Lexique en langue générale

De façon générale, on identifie, parmi les principales orientations en didactique du lexique, une approche centrée sur le sens (Sardier, 2015, p. 9 -16). La prise en compte des propriétés sémantiques des unités lexicales, selon cette approche, se trouve au cœur de la modélisation du lexique. On distingue deux modèles descriptifs du lexique en linguistique : un modèle componentiel et un modèle relationnel (Jousse, 2010, p. 10 -13). Le premier modèle définit le sens des unités lexicales en termes d'unités sémantiques plus petites, nommées *sèmes* (Pottier, 1992) ou *primitifs sémantiques* (Wierzbicka, 1996). Le modèle relationnel envisage le lexique comme un ensemble d'unités lexicales reliées par des relations sémantiques variées. Ce modèle étudie le sens des unités lexicales non seulement de façon isolée, mais aussi dans un réseau⁷

⁷ En didactique du lexique, la notion de réseau a été utilisée de différentes façons, comme l'explique Grossmann (2011, p. 175). Dans cette section, le *réseau* se réfère à un moyen de structurer la réflexion linguistique, comme c'est le cas de *réseaux sémantiques* de Picoche (1992, 1993a), « constitués à partir de mots de haute fréquence, des caractéristiques sémantiques de leurs actants, des associés moins fréquents » (Grossmann, 2011, p. 175), et le cas de *réseau sémantique* constitué des liens

sémantique. Comme notre recherche vise la description de la combinatoire, nous nous intéressons d'évidence à la représentation des relations sémantiques.

Parmi les travaux didactiques s'inscrivant dans le modèle relationnel, citons notamment ceux qui sont inspirés de la LEC. En tant que branche lexicologique de la Théorie Sens-Texte (TST), la LEC fonde la description des phénomènes lexicaux sur la modélisation sémantique des lexies. La LEC modélise deux types de liens sémantiques, soit les liens *paradigmatiques* (synonymes, antonymes et dérivés syntaxiques) et les liens *syntagmatiques* (collocations restreintes) (Polguère, 1998, p. 11). Ce modèle descriptif mène naturellement à la prise en compte de la combinatoire lexicale. Nous l'aborderons dans la section 2.2.

La LEC a été appliquée en didactique du lexique général par un certain nombre de chercheurs. On recense des travaux portant sur la conception de dispositifs d'enseignement en classe (Cauchon, 2003 ; Tremblay, 2003, 2004, 2009 ; Anctil, 2005, 2010) et la création de ressources d'apprentissage (Mel'čuk et Polguère, 2007 ; Milićević et Hamel, 2007 ; Milićević, 2008). Ces travaux traitent de l'enseignement du lexique à différents niveaux scolaires et ciblent les apprenants de langue première (L1) et de langue seconde (L2). Les recherches sur l'enseignement du français L1 au Québec ciblent le niveau primaire (Cauchon, 2003 ; Tremblay, 2003, 2004, 2009), secondaire (Anctil, 2010) et universitaire (Anctil, 2005). Anctil fait appel à la LEC pour traiter les erreurs lexicales (Anctil, 2010) et assister les étudiants universitaires dans la résolution des problèmes lexicaux (Anctil, 2005). Cauchon (2003) s'intéresse à l'enseignement de la structure prédicative du verbe au primaire. Tremblay s'inspire de la LEC pour proposer un système de notions, afin de soutenir l'enseignement-apprentissage du lexique au primaire, aussi bien dans la conception d'activités en classe (Tremblay, 2003, 2004) que dans la formation d'enseignants (Tremblay, 2009). Milićević et Hamel abordent l'enseignement du français L2. Leurs recherches (Milićević et Hamel, 2007 ; Milićević, 2008) visent l'élaboration d'un dictionnaire de reformulation destiné aux étudiants de niveaux intermédiaire et avancé.

syntagmatiques et paradigmatiques, qui représente formellement un sens lexical, tel que proposé par la LEC (Mel'čuk et al., 1995, p. 73, 235),

Picoche propose aussi une approche didactique fondée sur l'analyse sémantique. Elle s'intéresse notamment à la polysémie de mots de haute fréquence (Picoche, 1993a, 1993b, 2016). Son approche vise à clarifier les structures mentales que recouvrent les polysèmes. Picoche propose de faire éclater les polysèmes en homonymes et regrouper, au moyen des champs actanciels, les polysèmes ayant la même structure de sens. Ce premier travail mène à l'établissement de réseaux sémantiques « comportant les lexies associées par des liens sémantiques et/ou morphologiques, ainsi qu'à travers la structure actancielle » (Picoche et Souhaité, 2004, p. 221). En cela, cette démarche pédagogique permet l'enrichissement progressif du vocabulaire des apprenants. Elle a débouché sur l'élaboration du *Dictionnaire du français usuel* (Picoche et Rolland, 2002), un dictionnaire à orientation pédagogique. De plus, l'approche actancielle de Picoche intègre l'enseignement syntaxique à la didactique du lexique. Notons que les réflexions pédagogiques de Picoche se limitent au lexique en langue générale. L'auteure laisse de côté le lexique spécialisé, en prétendant que « (les) mots techniques, les termes propres aux différentes spécialités [...] s'apprennent tout naturellement quand on pratique ladite spécialité » (Picoche, 1999, p. 430).

L'intérêt des réseaux sémantiques pour l'acquisition du lexique a été démontré par d'autres didacticiens. Dans les travaux de Cavalla portant sur la didactique du français langue étrangère (FLE), les réseaux sémantiques sont mis à contribution pour enseigner le lexique des émotions auprès des étudiants non natifs (Cavalla, 2013, 2015, 2016 ; Cavalla et al., 2014 ; Tutin et al., 2006). Les noms d'affect, à titre d'exemple, sont classés en réseau en fonction de leurs propriétés sémantique et combinatoire (Tutin et al., 2006). Les réseaux élaborés sont représentés au moyen de cartes mentales. La figure 2 illustre la carte mentale rassemblant des verbes permettant de désigner le stade évolutif de la colère (Cavalla et al., 2014). Les recherches de Cavalla, visant la conception des séquences pédagogiques en classe du FLE, se distinguent par deux traits : l'intégration des réseaux collocationnels (Williams, 1999, 2001) à l'enseignement du lexique, ainsi que la démarche inductive, autrement dit, les étudiants sont amenés à construire des réseaux sémantiques par eux-mêmes.

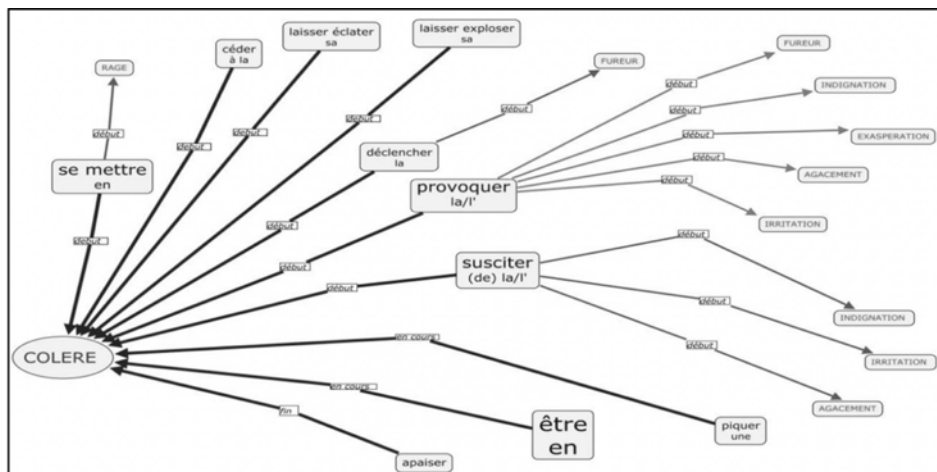


Figure 2. – Carte mentale de l'évolution de la colère (début, milieu, fin) (Cavalla et al., 2014)

Lavoie et Cellier proposent aussi la mise en réseau des unités lexicales. Cellier (2008, 2011) s'intéresse à l'enseignement du français L1 dans les classes du primaire en France. Lavoie (2015 ; Lavoie et Hoa, 2014) traite de l'enseignement du français L2 au primaire en contexte québécois.

En résumé, les travaux recensés dans cette sous-section fondent la didactique du lexique sur la description systématique du sens des unités lexicales. Le modèle relationnel est mis à contribution pour favoriser l'enrichissement du vocabulaire chez les apprenants.

2.1.2 Lexique spécialisé

En didactique du lexique spécialisé, la description sémantique des termes, soit les unités lexicales dotées d'un sens spécialisé, vise à aider les apprenants à résoudre des difficultés rencontrées tant dans la compréhension que dans la production écrite.

Du point de vue des problèmes ciblés, on recense d'importants travaux portant sur la représentation sémantique des termes. Binon et ses collègues (Binon et al., 1998) abordent le traitement de la synonymie dans un dictionnaire d'apprentissage visant principalement les apprenants non natifs. Alipour (2018) propose des méthodes d'enseignement de la polysémie et de la (quasi-) synonymie. Dechamps (2004, 2013a, 2013b), pour sa part, aborde la polysémie et la synonymie en français juridique. Selon ces auteurs, le sens des termes doit être appréhendé dans le contexte linguistique de ceux-ci, notamment la combinatoire lexicale.

Du point de vue du domaine abordé et du public ciblé, Dechamps (2004, 2013a, 2013b) aborde l'enseignement du français juridique à de futurs traducteurs lusophones. Alipour (2018)

s'intéresse aussi à la formation de traducteurs spécialisés, pour les initier au vocabulaire lié à la pollution de l'eau. Les travaux menés par le Groupe de recherche en lexicographie pédagogique (GRELEP) (Binon et al., 1998, 2000, 2005 ; Binon, Selva, et al., 2004 ; Binon et Verlinde, 2004 ; Verlinde et al., 2003, 2005) ont débouché sur l'élaboration du *Dictionnaire d'apprentissage du français des affaires* (DAFA), destiné aux apprenants de niveau avancé de français L2⁸.

Du point de vue de l'objectif de recherche, on recense des travaux sur l'enseignement du lexique spécialisé auprès d'étudiants en traduction. Alipour (2018) s'inspire de l'approche lexico-sémantique de la terminologie pour concevoir des activités pédagogiques en classe, afin de développer l'habileté des apprenants à résoudre les problèmes lexicaux dans la traduction spécialisée à l'aide de contextes. L'efficacité de ces activités a été évaluée au moyen d'une expérimentation dans un cours de terminologie du premier cycle. Les résultats ont confirmé l'utilité pédagogique des activités mises au point. Les recherches de Dechamps (2004, 2013a, 2013b) visent à élaborer un outil pédagogique, soit une plateforme d'exercices, afin de favoriser l'acquisition du lexique juridique par de futurs traducteurs.

Outre la conception d'activités en classe, on recense quelques travaux consacrés à l'élaboration de ressources d'apprentissage. En plus du projet de recherche du GRELEP sur le DAFA, citons Fuertes Olivera (2010), un recueil d'articles de recherche de divers aspects de la lexicographie pédagogique en langue de spécialité. La première partie de l'ouvrage traite de la contribution de la lexicographie fonctionnelle (Tarp, 2005, 2008) à l'élaboration de dictionnaires spécialisés pour les apprenants. La deuxième partie présente quelques pistes de recherche inspirées de la lexicographie pédagogique en langue générale, y compris la prise en compte des aspects culturels et du sens figuré des termes, la description des propriétés linguistiques des termes et le recours au corpus. La dernière partie de l'ouvrage aborde de nouvelles voies de recherche en matière de création de dictionnaires spécialisés à vocation pédagogique. Publié quinze ans après la parution de *Manual of Specialised Lexicography* (Bergenholtz et Tarp, 1995) qui a longtemps été un ouvrage de référence important en matière de lexicographie en langue de spécialité, Fuertes Olivera (2010) donne un aperçu des progrès réalisés dans ce domaine de recherche et de

⁸ Le GRELEP a aussi mis au point le *Dictionnaire d'apprentissage du français langue étrangère ou seconde* (DAFLES), un outil d'apprentissage du lexique de la langue générale.

l'évolution de la lexicographie pédagogique depuis 1995. Cet ouvrage a inspiré des travaux qui débouchent sur des résultats tangibles, dont Alipour (2014). L'auteure s'inspire en partie de la lexicographie fonctionnelle (Tarp, 2005, 2008) pour explorer la conversion de dictionnaires spécialisés en dictionnaires d'apprentissage. Cette recherche a le mérite d'élaborer un modèle de dictionnaire nommé *AppDiCo*⁹ à partir de données terminologiques du DiCoInfo. La conception de ce modèle rend compte des besoins des apprenants du lexique spécialisé, plus précisément, les traducteurs et les professionnels d'un domaine spécialisé.

Comme les apprenants du lexique spécialisé sont majoritairement des adultes ayant des objectifs spécifiques, il importe ainsi de leur offrir des outils d'apprentissage permettant une acquisition efficace et autonome des connaissances lexicales. En cela, Binon et al. (1998, p. 77) insistent sur l'importance de « la présentation claire et ordonnée des relations sémantiques des termes » lors de l'élaboration d'un dictionnaire d'apprentissage, étant donné que la maîtrise des relations sémantiques entre les termes se trouve au cœur des processus de compréhension et de production. Ce principe a aussi guidé la conception du *Dictionnaire bilingue (anglais-français) de la distribution*, un outil destiné à des traducteurs ou à des professionnels du domaine commercial. Dans ce dictionnaire, les traits sémantiques des termes ainsi que les relations sémantiques que ceux-ci entretiennent ont été explicités, mettant ainsi en évidence l'organisation conceptuelle du domaine (Dancette et Réthoré, 1997, p. 229).

Ce survol des travaux précédents montre que le traitement sémantique joue un rôle essentiel dans l'enseignement du lexique, tant en langue générale qu'en langues de spécialité. La description des relations sémantiques mène naturellement à la prise en compte de la combinatoire lexicale.

2.2 Étude de la combinatoire lexicale dans une perspective didactique

Dans cette section, nous recensons les recherches sur la combinatoire lexicale à des fins didactiques. Nous donnons d'abord un aperçu de différentes définitions de la combinatoire lexicale, pour aborder ensuite son enseignement. Comme notre projet vise les apprenants non

⁹ L'*AppDiCo* est repérable à l'adresse suivante : <http://olst.ling.umontreal.ca/cgi-bin/appdico/search.cgi?>

natifs de langues de spécialité, nous nous focalisons sur les travaux didactiques s'inscrivant dans ce contexte.

2.2.1 Définition de la combinatoire lexicale en vue de son enseignement

En langue générale, le terme *collocation* a été introduit par Firth (1957) pour désigner l'association habituelle et récurrente des mots sémantiquement apparentés. Selon les travaux du contextualisme britannique qui sont dans la même lignée, la collocation se définit par la fréquence d'occurrence des groupements lexicaux (Halliday, 1966 ; Sinclair, 1987, 1991). Certains linguistes adoptent une vision sémantique et définissent la collocation comme une cooccurrence lexicale restreinte (Hausmann, 1979, 1989 ; Benson, 1989 ; Kromann et al., 1991 ; Mel'čuk et al., 1984, 1995 ; Mel'čuk, 1996).

La collocation est également définie en termes de contraintes qu'elle exerce sur l'utilisation du lexique. Selon Tutin et Grossmann, qui se réfèrent à Hausmann (1979, 1989), une collocation résulte d'une cooccurrence privilégiée de deux constituants linguistiques entretenant une relation sémantique et syntaxique (Grossmann et Tutin, 2003 ; Tutin, 2013 ; Tutin et al., 2006).

En langues de spécialité, les combinaisons lexicales sont désignées par diverses appellations. De nombreux auteurs retiennent le terme *collocation* et fondent leurs réflexions sur l'enseignement du lexique général, dont les chercheurs du GRELEP (Binon et Verlinde, 2003, 2004 ; Verlinde et al., 2003), Dechamps (2004, 2012, 2013b, 2015, 2017), Cavalla (2008a, 2008b, 2015a, 2015b), Tutin (2010b, 2010a, 2013 ; Tutin et Grossmann, 2002 ; Grossmann et Tutin, 2003) et Pecman (2012 ; Pecman et al., 2009). Certains terminologues préfèrent le terme *cooccurrent* pour insister sur la récurrence de certaines combinaisons préférentielles dans un contexte d'usage spécifique (Cohen, 1986 ; Thoiron et Béjoint, 1989 ; Caignon, 2001). D'autres terminologues parlent de *phraséologismes*¹⁰ (Pavel, 1993 a, 1993b, 1994) ou de *phrasèmes*¹¹ (Meyer et Mackintosh, 1994),

¹⁰ Pavel (1994, p. 4) définit le *phraséologisme* en langue de spécialité comme « une combinaison préférentielle ou une solidarité lexicale d'un terme, appelé noyau ou base, et de mots qui cooccurrent dans la même phrase ». Ce terme, comme le fait remarquer L'Homme (2017, p. 219), recouvre une acception différente en linguistique, à savoir « les expressions multilexémiques non compositionnelles comme les expressions figées ».

soulignant ainsi les propriétés syntaxico-sémantiques des combinaisons spécialisées¹². Notons que la phraséologie, dans les travaux sur le lexique transdisciplinaire de la langue scientifique générale (LSG) (Pecman, 2005; Tutin, 2014), recouvre un sens plus large. Elle se réfère aux « séquences polylexicales récurrentes spécifiques au genre des écrits scientifiques, qu’elles soient ou non figées, et, quel que soit le statut syntaxique des éléments constituant l’expression » (Tutin, 2007, 2014). Pecman (2005, p. 114-115) souligne que les unités phraséologiques¹³ comprennent aussi bien les « collocations restrictives »¹⁴ que les combinaisons libres. En ce sens, la phraséologie établit le pont entre le lexique et le discours spécialisé, situant ainsi l’étude de la combinatoire lexicale dans une perspective discursive (Pecman, 2012).

Rappelons que notre thèse aborde les combinaisons lexicales spécialisées (CLS). Celles-ci se définissent comme les collocations contrôlées par une lexie à sens spécialisé dans les travaux de recherche en terminologie (L’Homme et Gemme, 1997 ; L’Homme et Meynard, 1998 ; L’Homme et Bertrand, 2000 ; L’Homme, 2000, 2003, 2017). Dans le cadre de notre travail, les CLS renvoient aussi aux liens observés dans les structures actanciennes qui comportent une lexie à sens prédicatif et ses actants sémantiques (voir aussi la section 1.3).

La maîtrise de la combinatoire lexicale est essentielle pour acquérir une langue, comme le soulignent plusieurs chercheurs (Lerat, 1995; Binon et Verlinde, 2003, 2004; Dechamps, 2004), et comme nous l’avons illustré avec des exemples chinois à la section 1.1. Il est essentiel de

¹¹ La notion de phrasème est définie autrement en Lexicologie explicative et combinatoire (LEC). Il s’agit des unités lexicales multi-lexémiques qui peuvent être divisées en trois sous-catégories selon leur degré de figement et leur compositionnalité, le tout constituant un continuum allant des locutions idiomatiques jusqu’aux collocations partiellement compositionnelles. Les expressions libres sont donc exclues (Mel’čuk et al., 1995, p. 46).

¹² Pavel (1994, p. 4) cite Rousseau (1993, p. 9) qui fait remarquer que « (la plupart des auteurs attribuent au phraséologisme un caractère syntaxique et un caractère notionnel particulier qui va au-delà de la notion désignée par le terme (élément de base du phraséologisme spécialisé) ».

¹³ Pecman (2007) s’aligne sur la définition d’*unité phraséologique* proposée par Zinglé (Zinglé et Brobeck-Zinglé, 2003, p. iii). Ce terme désigne « indifféremment les collocations, les colligations, les locutions et les syntagmes divers ». Nous référons au lecteur à Pecman (2004, p. 18-28) pour une analyse des différents types d’unités polylexicales.

¹⁴ Les collocations restrictives sont désignées sous le terme simple *collocations* (Grossmann et Tutin, 2003 ; Tutin, 2013 ; Tutin et al., 2006).

réserver une place à ce phénomène en didactique du lexique. Dans ce qui suit, nous nous focalisons sur les travaux abordant l'enseignement de la combinatoire spécialisée auprès des apprenants non natifs.

2.2.2 Enseignement de la combinatoire lexicale spécialisée

En didactique du lexique spécialisé, les CLS soulèvent d'abord des difficultés sur le plan réceptif. Binon et Verlinde (2003, p. 17) font remarquer que le sens d'une CLS n'est pas prédictible à partir de ses composantes ni à partir de la langue maternelle des apprenants. Dechamps (2004, 2013a, 2013b, 2017), qui œuvre dans la formation de traducteurs spécialisés, s'intéresse à la distinction sémantique des CLS verbales en langue juridique. Le choix de collocations verbales, selon Dechamps, permet de saisir les nuances sémantiques des CLS, ce qui influence la recherche d'équivalents dans la langue cible (Dechamps, 2013b, p. 12). Pour exemplifier son propos, Dechamps analyse deux CLS associées au terme *ordonnance* qui sont des sens proches : 1) *annuler une ordonnance* et 2) *abroger une ordonnance*. La première CLS relève du pouvoir exécutif et exige un sujet désignant une institution gouvernementale, comme *conseil d'État*; la deuxième CLS relève du pouvoir législatif et se combine ainsi avec des sujets comme *Parlement* ou *loi*. Face aux difficultés relevées, ces didacticiens proposent le recours au contexte linguistique pour appréhender, soit le sens des CLS, soit le sens des éléments constitutifs de celles-ci (Binon et al., 1998 ; Dechamps, 2013a, 2013b).

Pour les apprenants non natifs, la combinatoire spécialisée présente un obstacle davantage sur le plan de la production. Si les apprenants peuvent recourir au contexte pour inférer le sens des CLS, ils auront à s'approprier des CLS pour formuler leurs idées, tout en tenant compte des contraintes pragmatiques (registre de langue, genre discursif, etc.). La difficulté que présente le maniement de la combinatoire a été soulignée par plusieurs didacticiens en FLE (Binon et Verlinde, 2004, p. 274 ; Grossmann et Tutin, 2003, p. 9 ; Pecman, 2005, p. 116). Ces constats poussent les chercheurs à se pencher sur l'enseignement de la combinatoire dans une perspective d'aide à la production écrite. Certains travaux explorent les méthodes de description des CLS, citons les

études sur le lexique scientifique transdisciplinaire (LST)¹⁵ réalisées par le Laboratoire de Linguistique et didactique des langues étrangères et maternelles (LIDILEM) (Cavalla, 2008, 2015, 2018a, 2018b ; Tutin, 2014, 2018 ; Tutin et Grossmann, 2014 ; Tutin et Kraif, 2016). D'autres recherches ont débouché sur l'élaboration de ressources pédagogiques. Pour n'en citer que quelques-unes, le DAFA mis au point par le GRELEP ; la base ARTES (Aide à la Rédaction de Textes Scientifiques) élaborée par l'Université Paris Diderot (Pecman, 2005, 2012 ; Pecman et al., 2009 ; Pecman et Kübler, 2011), ainsi que le DiCE, un dictionnaire de collocations de l'espagnol académique, élaboré par des chercheurs de l'Université de La Corogne (Alonso Ramos et al., 2008, 2010 ; Garcia et al., 2019 ; Wanner et al., 2013).

Les travaux recensés mettent l'accent sur le traitement sémantique de la combinatoire. Selon Cavalla (2016, p. 66), une CLS « est simultanément une association lexicale, une combinatoire syntaxique spécifique et un sens particulier ». Tous ces points d'accès doivent être pris en compte dans l'enseignement de la combinatoire, et le sens doit occuper une place prioritaire. Pour ce faire, l'auteure fait appel à la sémantique lexicale pour développer une description sémantique des CLS du français scientifique transversal (Cavalla, 2008a).

En espagnol académique, l'élaboration du dictionnaire DiCE a été guidée par le même principe. Les CLS sont encodées au moyen de fonctions lexicales (FL). Des gloses en langue naturelle sont aussi fournies pour expliciter les liens sémantiques exprimés par les FL (Alonso Ramos, 2005 ; Alonso Ramos, et al., 2010 ; Vincze et al., 2011). La figure 3 illustre les résultats obtenus pour une recherche portant sur *a raudales* (« en abondance ») dans le dictionnaire en ligne DiCE. On y trouve les termes qui sont reliés au collocatif *a raudales* par la FL **Magn** exprimant le sens de 'très', 'intense', 'à un degré élevé' (Mel'čuk et al., 1995, p. 136). On explique le lien sémantique entre chaque terme de base et le collocatif *a raudales* au moyen d'une glose (*intenso, intensa et grande*). Des exemples d'usage des CLS sont aussi fournis.

¹⁵ Dans les travaux de recherche recensés, le lexique scientifique transdisciplinaire (LST) relève d'une langue de spécialité nommée français académique, français sur objectifs universitaires (FOU) ou langue scientifique générale (LSG) (Grossmann et Tutin, 2003, p. 9; Pecman, 2005, p. 116).

Encontradas 4 colocaciones, listadas del 1 al 4 (página 1 de 1)
 << página anterior | | página siguiente >>

Magn (4 valores en total)

afecto 2a (*Sentimiento*) [ver ejemplos]
 Glosa
 intenso
 Ejemplos
 1. el Dr. Inchausti y Pepe les dispensan su mayor admiración y afecto a raudales.

alegría 1a (*Sentimiento*) [ver ejemplos]
 Glosa
 intensa
 Ejemplos
 1. Alegría a raudales, que diría un cursi.
 2. Y entonces nos pusimos a firmar papeles y papeles, mientras Matías, Paula y Gonso derrochaban alegría a raudales. (web)

simpatía 2 (*Cualidad*) [ver ejemplos]
 Glosa
 grande
 Ejemplos
 1. Agassi desprende simpatía a raudales

Figure 3. – Résultats de recherche portant sur le collocatif *a raudales* dans le dictionnaire DiCE

La conception du DAFA est orientée vers les besoins du public visé, soit les apprenants allophones du français des affaires (Binon et Verlinde, 2003). Les CLS sont organisées selon des critères à la fois morphosyntaxiques, sémantiques et paradigmatiques. La description sémantique des CLS s’inspire aussi des travaux de Mel’čuk sur l’élaboration du *Dictionnaire explicatif et combinatoire* (DEC) (Mel’čuk, 1992 ; Mel’čuk et al., 1984, 1999 ; Mel’čuk et Clas, 1988, 1992). D’abord, les FL servent de repères pour organiser la combinatoire des termes. Ensuite, pour les combinaisons verbales, le recours aux schémas actanciels permet d’explicitier de façon systématique tous les actants et circonstants des verbes (Binon et al., 2005, p. 221-222 ; Binon et Verlinde, 2004, p. 280-281). Les travaux de Dechamps (2013b, 2017) s’inscrivent dans la même lignée. L’auteure étudie le traitement des CLS dans les dictionnaires destinés aux apprentis traducteurs juridiques. Elle envisage l’organisation de la combinatoire verbale au moyen de la structure actancielle, inspirée des recherches en terminologie (L’Homme, 1997 ; Pimentel, 2012 ; Pimentel et al., 2011) et en lexicographie pédagogique (Verlinde et al., 2003). Plus précisément, les différents actants d’un verbe sont explicités et organisés en classes d’objets. Les propriétés sémantiques et les contraintes d’usage de ces actants sont décrites à l’intérieur des classes établies. Cette démarche facilite la distinction des verbes synonymes et la recherche de leur équivalent dans la langue cible.

Dans le cadre de l’élaboration de la base ARTES, la combinatoire spécialisée a été prise en compte dans sa dimension discursive. Cette base de données est destinée aux locuteurs de langues de spécialités, notamment les traducteurs et les scientifiques qui rédigent dans leur L2

(Pecman, 2012). On y trouve des informations sur la phraséologie des écrits scientifiques. Les unités phraséologiques (UP) transdisciplinaires sont organisées en paradigmes combinatoires selon leur fonction de désignation d'une notion dans les discours scientifiques. Cette démarche, selon Pecman (2007, p. 93), met en évidence «les divergences dans les paradigmes combinatoires des différentes langues dans le cadre d'un type de discours particulier». La base ARTES offre un double accès aux UP, soit à partir de leur forme et de leur contenu sémantique (Pecman, 2005, p. 109, 2012, p. 116). Les fonctions discursives des collocations sont prises en compte dans le traitement des collocations transdisciplinaires. La figure 4 illustre l'interface de recherche de la base ARTES donnant un accès aux collocations transdisciplinaires par fonctions discursives. On y trouve une liste de collocations qui servent à présenter ses objectifs de recherche. En fin de compte, les ressources ainsi élaborées aident les locuteurs non natifs à manier la combinatoire dans leurs productions écrites.

The screenshot shows the 'DICTIONNAIRE ARTES' interface. At the top, there are three tabs: 'Terminologie en contexte', 'Phraséologie discursive', and 'Recherche multicritères'. The 'Phraséologie discursive' tab is active, showing a search for 'Fonctions discursives'. A yellow box highlights 'Point d'accès aux collocations spécifiques aux genres'. Below this, there is a search bar and a table of discursive functions with examples. A red circle highlights the function 'Présenter ses objectifs de recherche'. To the right, a section titled 'Présenter ses objectifs de recherche' shows a search for 'Collocations associées'. Below this, a table lists collocations with their constructions and associated discourses. A yellow box highlights 'Collocations répertoriées selon la fonction discursive à laquelle elles sont associées'.

| Fonctions discursives | Exemple |
|--|---------------------|
| Annoncer la ou les conclusions | ex. our conclusio |
| Evoquer le sujet de son étude | ex. in the present |
| Présenter ses objectifs de recherche | ex. one of the pri |
| Faire un renvoi à une partie dans le discours en cours | ex. some concret |
| Evoquer son positionnement ou le contexte théorique dans lequel s'in | ex. this work is ai |
| Présenter ses méthodes, outils, ses approches, ses techniques | ex. to pioneer a n |
| Présenter ses hypothèses ou ses prémisses de travail | ex. to put forward |
| Discuter des difficultés, problèmes ou limitations rencontrés | ex. the problem v |
| Evoquer les autres points d'intérêt pour l'étude | ex. this issue can |
| Faire des observations empiriques | ex. a considerabl |
| Parler des expérimentations, des tests et des vérifications | ex. our descriptio |
| Faire des illustrations ou donner des exemples | ex. an interesting |
| Décrire, interpréter et analyser les données ou phénomènes observés | ex. to describe st |
| Expliquer les conditions ou les circonstances dans lesquelles se dérou | ex. to occur unde |
| Faire des estimations, des calculs et des interprétations | ex. estimates hav |
| Evoquer des caractéristiques, propriétés, spécificités | ex. understandin |
| Parler des lois, règles, formalismes, algorithmes, fonctions | ex. these laws pe |

| Collocation | Construction | Discours |
|---------------------------------------|------------------|--------------------|
| the primary objective of sth is to | schéma phrastiqu | discours socio-pc |
| the aim of the present study is to vb | GN vb. prép. | discours scientifi |
| the purpose of this study | nom prép. nom | discours universit |
| for the purpose of sth/vb ing | prép. N prép. | discours universit |
| with a view to vb | prép. N prép. | discours multireg |
| to study sth | vb. nom | discours scientifi |
| the aim of this work is to | GN vb. prép. | discours scientifi |
| the study highlights | nom vb. | discours scientifi |
| the present chapter will attempt to | introduceur d'ém | discours multireg |
| the aim of this study | GN prép. nom | discours scientifi |
| the prime focus of | adj. N prép. N | discours socio-éc |
| to achieve this | vb. nom | discours multireg |

Figure 4. – Interface de la base ARTES montrant un accès aux collocations par fonctions discursives (ARTES : Aide à la Rédaction de Textes Scientifiques [CLILLAC-ARP], s. d.)

En guise de conclusion, la maîtrise de la combinatoire lexicale constitue « la clef de voûte » de l'enseignement-apprentissage d'une L2 (Binon et Verlinde, 2003, p. 16). Les recherches dans ce domaine s'avèrent fructueuses. De nombreuses pistes de solution ont été proposées, notamment en ce qui concerne le développement de méthodes d'enseignement et la création de ressources

d'apprentissage. Ces travaux, bien que menés dans diverses perspectives, attachent tous une importance au traitement sémantique des CLS. Nous nous en inspirons pour envisager notre propre méthode de description de la combinatoire spécialisée en chinois.

2.3 Enseignement du lexique spécialisé chinois

Cette section dresse un état des lieux de l'enseignement lexical en COS. Nous présentons d'abord le contexte général (2.3.1) et les approches didactiques du lexique spécialisé (2.3.2). Nous examinons enfin la description lexicale dans les manuels de COS (2.3.3).

2.3.1 Enseignement du chinois sur objectif spécifique (COS)

Cette sous-section propose un descriptif du COS. Nous présentons les domaines abordés (2.3.1.1), le profil des apprenants (2.3.1.2) et les objectifs d'enseignement qui orientent la conception de dispositifs pédagogiques (2.3.1.3).

2.3.1.1 Qu'est-ce que l'on entend par « objectifs spécifiques » ?

Le COS a émergé au début des années 1980, où la Chine a intensifié ses échanges internationaux dans divers secteurs d'activités. À l'époque, les universités chinoises ont commencé à accueillir un nombre grandissant d'étudiants internationaux, d'où la nécessité d'offrir des formations langagières adaptées aux besoins spécifiques de ces apprenants.

La conception de la formation en COS s'inspire des travaux de recherche sur l'anglais sur objectifs spécifiques (*English for Specific Purposes*, ou ESP)¹⁶ (Zhai, 2009; Q. Li, 2011; L. Zhang et al., 2016). Le COS vise à préparer les apprenants à leur intégration dans un milieu universitaire ou professionnel. Ainsi, l'*objectif spécifique* se définit d'une part, par les disciplines universitaires (comme la chimie, la physique et la science politique), et, d'autre part, par la visée professionnelle des apprenants (comme le commerce, le tourisme et l'hôtellerie). Li (2011, p. 111-112) fait état des domaines abordés en COS (tableau 6).

¹⁶ L'enseignement de l'anglais sur objectifs spécifiques (ESP) a donné lieu à une littérature de recherche abondante. Faute d'espace, nous nous limitons à citer quelques travaux qui brossent un portrait de ce domaine de recherche (Hyland, 2007, 2022) ou inspirent l'enseignement du lexique en COS (Coxhead, 2017; Hyland et Tse, 2007).

| | |
|--|---|
| COS défini en fonction de disciplines universitaires | le chinois du STIM (science, technologie, ingénierie et mathématiques) |
| | le chinois médical (la médecine traditionnelle chinoise et la médecine occidentale) |
| | le chinois de la littérature, de l'histoire et de la philosophie |
| | le chinois de la science économique, de la science politique et du droit |
| COS défini en fonction de visées professionnelles | le chinois de la diplomatie |
| | le chinois des affaires |
| | le chinois du commerce |
| | le chinois des médias |
| | le chinois militaire |
| | le chinois du tourisme |
| | le chinois de l'ingénierie |
| | le chinois de l'entreprise |
| | le chinois de l'aéronautique |
| | le chinois de l'hôtellerie |
| le chinois du bureau | |

Tableau 6. – Les domaines abordés en COS d'après Li (2011, p. 111-112)

Les premiers cours de COS ont eu pour thème le chinois du commerce, le chinois scientifique et technique, le chinois médical et le chinois de l'ingénierie (B. Lü, 1990, p. 69-72). Quant à la compilation de matériel didactique, plus de 150 manuels de COS ont été publiés depuis les années 1980s (Li, 2011 ; Lu, 2006). Les premiers manuels de COS, axés principalement sur le commerce et les affaires, ont accordé la primauté à la communication orale et proposé ainsi un enseignement contextualisé du lexique (tableau 7).

| Titre | Auteur(s) | Maison d'édition | Date de publication |
|--|--|--|---------------------|
| wài mào qiān wǔ bǎi jù hàn yīng duì zhào 外贸洽谈五百句 (汉英对照) [500 Sentences for Foreign Trade Negotiations (Chinese-English)] | Beijing Language Institute et Beijing Institute of Foreign Trade | huá yǔ jiāo xué chū bǎn shè 华语教学出版社 [Sinolingua Press] | 1982 |
| shāng yè hàn yǔ 商业汉语 [Business Chinese]. | Liyang Lianni et Li Gengxin | huá yǔ jiāo xué chū bǎn shè 华语教学出版社 [Sinolingua Press] | 1990 |
| jīng mào chū jí hàn yǔ kǒu yǔ 经贸初级汉语口语 [Elementary Spoken Chinese in Economics and Trade] | Huang Weizhi | duì wài mào yì jiāo yù chū bǎn shè 对外贸易教育出版社 [Foreign Trade Education Press] | 1993 |
| shāng yòng hàn yǔ huì huà 商用汉语会话 [Business Chinese Conversation] | Guo Li | běi jīng dà xué chū bǎn shè 北京大学出版社 [Peking University Press] | 1993 |

Tableau 7. – Les premiers manuels de COS portant sur le domaine du commerce et des affaires

Au fur et à mesure de l'élargissement de l'ouverture de la Chine, l'enseignement du COS est en pleine évolution depuis les vingt dernières années. Les didacticiens et les linguistes chinois sont

amenés à repenser la pédagogie du COS en tant qu'une partie intégrante du chinois langue seconde¹⁷.

Des recherches s'intensifient sur les théories et la pratique d'enseignement du COS. Zhai (2009) s'inspire des travaux de recherche en didactique de l'ESP pour redéfinir les domaines de spécialité abordés en COS. Li (2011) brosse un portrait de la prestation de formation en COS en Chine. Il aborde les aspects fondamentaux des recherches en enseignement du COS, dont la définition du COS, le cadre théorique et les principes de base de la didactique du COS, l'analyse de besoins d'apprentissage, l'élaboration de curriculum de formation, la conception de ressources didactiques et la formation des enseignants.

La première monographie qui présente de manière systématique les théories et les méthodes de l'enseignement du COS, à notre connaissance, a été parue en 2016 (L. Zhang et al., 2016). Il s'agit d'un manuel destiné aux étudiants inscrits au programme de maîtrise en enseignement du chinois langue seconde. Cette monographie est composée de sept chapitres s'articulant autour de trois thèmes : l'analyse de besoins d'apprentissage, l'analyse du COS à des fins didactiques et la conception de séquences pédagogiques du COS.

Tao et Chen (2019) dresse un état de l'art des recherches en didactique du COS. Dans cet ouvrage collectif, on rassemble 17 articles rédigés en anglais. Ces travaux sont axés sur les aspects théoriques et appliqués de l'enseignement du chinois académique, du chinois des affaires, du chinois de la médecine et la formation de chinois offerte dans d'autres domaines (diplomatie, tourisme, dégustation de vins, etc.).

Des ressources didactiques ont été élaborés pour soutenir l'enseignement du COS dans différents domaines de spécialités. Nous avons recensé des manuels du chinois scientifique et technique (An et Shan, 2005, 2008 ; Bai, 2014 ; Du, 2011a, 2011b, 2012a, 2012b ; Han, 2012 ; Lu et al., 2017 ; Xu et al., 2012), du chinois du commerce (L. Li et Ding, 2002; X. Li, 2008a, 2008b, 2008d, 2008c), du chinois de la médecine traditionnelle chinoise (MTC) (Xu, 2013 ; Jin et Guo,

¹⁷ En didactique du chinois, l'appellation du *chinois langue seconde* est souvent confondue avec le *chinois langue étrangère*. Dans notre travail, le *chinois langue seconde* fait référence à l'enseignement-apprentissage du chinois visant tous les apprenants non natifs.

2017), du chinois de l'informatique (S. Zhang, 2014) et du chinois du droit des affaires (T. Zhang, 2007). Ces manuels abordent les matières les plus enseignées en COS et enrichissent les recherches sur la conception des ressources didactiques (Zhiying Lu, 2006; Shan, 2008; Shan et An, 2009; Y. Zhang, 2014). Nous les examinons plus attentivement à la section 2.3.3.

En résumé, le COS vise les apprenants qui se destinent à un domaine d'activité spécifique. L'enjeu de cette forme pédagogique consiste à adapter l'enseignement aux besoins des apprenants.

2.3.1.2 Profil du public ciblé

Le COS s'adresse à un public assez diversifié. Nous pouvons caractériser ces apprenants selon leur motif d'apprentissage, leur compétence linguistique et leurs connaissances du domaine abordé.

En premier lieu, nous identifions trois types d'apprenants souhaitant acquérir une formation de type COS. D'abord, le COS s'adresse aux étudiants internationaux qui suivent un cursus universitaire en Chine. Ils doivent suivre une formation préparatoire du chinois avant d'être intégrés au programme d'études de leur choix. Ensuite, le COS vise les professionnels d'un milieu spécialisé. L'objectif de ces apprenants consiste à acquérir rapidement des compétences linguistiques afin de pouvoir communiquer en milieu professionnel où l'on utilise le chinois. Ils optent souvent pour une formation « sur mesure » offerte par un établissement privé. D'autres apprenants sont des autodidactes. Ceux-ci souhaitent soit se donner un défi personnel, soit se préparer à une vie professionnelle dans un domaine particulier tel que le tourisme ou la restauration. Ces apprenants s'instruisent à l'aide de toutes sortes de ressources didactiques, notamment les manuels et les cours en ligne.

En second lieu, le niveau de compétences linguistiques des apprenants n'est pas homogène. Les débutants forment la majorité du public visé du COS, sauf certains qui ont déjà vécu une expérience de séjour en Chine. De plus, le bagage linguistique des apprenants affecte leur apprentissage. Par exemple, les apprenants provenant des pays où on utilise le chinois sont susceptibles de démontrer une habileté à reconnaître les caractères chinois, ce qui facilite grandement leur acquisition du lexique, comparé aux apprenants du reste du monde.

En dernier lieu, le bagage cognitif des apprenants est variable. Leur niveau de connaissances spécialisées varie en fonction de leurs expériences académique et professionnelle. Parmi les étudiants internationaux qui séjournent en Chine, ceux qui comptent entreprendre un programme de premier cycle ont généralement peu de connaissances du domaine spécialisé de leur choix. Les étudiants des cycles supérieurs ont déjà acquis de solides connaissances dans leur domaine.

Pour résumer, le COS s'adresse à des étudiants ayant des profils divers. Les besoins des apprenants varient selon leur motif d'apprentissage, leur compétence linguistique et leurs connaissances du domaine abordé. Ces apprenants démontrent une forte motivation. Leur apprentissage est soumis à des contraintes de temps imposées par leur projet académique ou professionnel. Leurs besoins prioritaires sont définis par la nécessité d'être rapidement « opérationnels » dans un milieu spécialisé.

Dans ce qui suit, nous nous focalisons sur l'enseignement du COS en milieu universitaire, vu que la conception pédagogique dans ce contexte rend compte des traits communs du public cible du COS.

2.3.1.3 Objectifs d'enseignement

L'enseignement du COS se fixe deux objectifs : le développement d'habiletés linguistiques et la remise à niveau des connaissances spécialisées. Ces objectifs président à la conception du contenu d'enseignement. Prenons comme exemple le manuel zhōng yī hàn yǔ zōng hé jiàochéng 中医汉语综合教程 [*Chinois de la médecine traditionnelle chinoise : cours intégré*] (J. Xu, 2013). Les modules d'apprentissage de ce manuel s'articulent autour de six thèmes en médecine traditionnelle chinoise (désormais MTC). Cette structuration thématique amène les apprenants du COS à construire progressivement leur bagage cognitif dans différents sous-domaines de la MTC, dont les notions fondamentales de la MTC (le yin et le yang, les cinq éléments, le *qi*, etc.), le diagnostic, l'acupuncture, la pharmacopée chinoise, les connaissances générales sur le bien-être humain, etc. Chaque module comprend un texte spécialisé et une liste de mots nécessaires à la compréhension du texte. L'apprentissage du lexique est ainsi intégré à des séquences de lecture. La figure 5 illustre l'extrait d'un module d'apprentissage portant sur les fondements philosophiques de la MTC dans le manuel *Chinois de la médecine traditionnelle chinoise : cours intégré* (J. Xu, 2013, p. 3). On y trouve un texte de compréhension écrite portant sur le soleil, la lune, le yin et le yang. Le texte est

assorti d'une liste de mots nouveaux, parmi lesquels figurent 哲学 (« philosophie »), 阴 (« yin »), 阳 (« yang ») et 寒 (« froid »). Chaque mot est assorti de sa transcription phonétique, sa partie du discours et son équivalent anglais.

课文 **Texte spécialisé sur les connaissances de base de la médecine traditionnelle chinoise**

TEXT

日月与阴阳 Le soleil, la lune, le yin et le yang

中国古代**哲学**认为,世界上任何事物都有**阴**、**阳**两个方面。



太阳为阳,月亮为阴;山的南面向着太阳,为阳,山的北面背着太阳,为阴;**白昼**属阳,黑夜属阴;男人属阳,女人属阴;天在上,**属**阳,地在下,属阴;火温暖而向上,属阳,水寒冷而向下,属阴,等等。

中医学也用阴阳来**解释**人体,比如:头在上,为阳,脚在下,为阴;体表为阳,体内为阴;背部为阳,腹部为阴;气为阳,血为阴。色泽鲜明、发热、口渴喜冷、兴奋等**症状**属阳,色泽晦暗、畏寒、口不渴或口渴喜热、**抑制**等**症状**属阴。

阴阳**平衡**才能保证身体健康,任何一方面**太过**或者**不足**,都会**引发**疾病。

生词 **NEW WORDS**

Liste de mots nouveaux figurant dans le texte

1. 哲学 *zhéxué* n. philosophy
2. 任何 *rènyǐ* pron. any
3. 阴 *yīn* n. *yīn*, the feminine or negative principle in nature
4. 阳 *yáng* n. *yáng*, the masculine or positive principle in nature
5. 方面 *fāngmiàn* n. aspect, side
6. 白昼 *báizhòu* n. daytime
7. 属 *shǔ* v. belong to
8. 解释 *jiěshì* v. commentate, explain
9. 人体 *rēntǐ* n. human body
10. 腹 *fù* n. abdomen
11. 喜 *xǐ* v. be fond of
12. 兴奋 *xīngfèn* adj. excited
13. 症状 *zhèngzhuàng* n. symptom
14. 色泽 *sèzé* n. colour
15. 晦暗 *huì'àn* adj. dark
16. 畏 *wèi* v. fear
17. 寒 *hán* adj. cold

Figure 5. – Extrait d'un texte spécialisé assorti d'une liste de mots nouveaux dans le manuel *中医汉语综合教程 [Chinois de la médecine traditionnelle chinoise : cours intégré]* (J. Xu, 2013, p. 3)

Comme on peut le constater, l'approche pédagogique de COS combine l'enseignement du chinois L2 et l'enseignement des connaissances d'un domaine de spécialité, ce qui est comparable à l'enseignement d'une matière par l'intégration d'une langue étrangère (EMILE) ou *Content and Language Integrated Learning (CLIL)* en anglais. L'approche EMILE est conçue par des experts en éducation et en acquisition de langues dans le cadre éducatif européen dans les années 1990 (Coyle, 2007, 2008; Coyle et al., 2010; Dalton-Puffer, 2007; Dalton-Puffer et al., 2010; Pérez-Cañado, 2012), où le multilinguisme et la diversité linguistique sont devenus une question cruciale dans le monde entier (Gabillon, 2020a, 2020b). Cette approche intègre

l'apprentissage des langues étrangères dans l'acquisition des contenus disciplinaires. Elle vise à créer un environnement d'apprentissage qui favorise l'engagement cognitif des apprenants et le développement de la capacité des élèves à résoudre des problèmes.

L'approche EMILE se distingue de l'approche pédagogique du COS par ses contextes d'application. Comparativement à l'enseignement du COS qui s'adresse majoritairement aux apprenants adultes, l'EMILE s'applique à une variété de contextes d'enseignement pour répondre aux besoins changeants des apprenants de différents groupes d'âge (Gabillon, 2020b). Par ailleurs, l'EMILE est mis en pratique davantage dans la conception des activités en classe (Corino et Onesti, 2019; Coyle, 2005; Urmeneta et Sola, 2009), comparativement à l'approche pédagogique du COS qui sous-tend aussi la conception des ressources didactiques.

Dans cette sous-section (2.3.1), nous avons donné un aperçu de l'enseignement du COS. Cette forme pédagogique est centrée sur les objectifs des apprenants. Malgré la diversité de leur profil, les apprenants poursuivent un objectif commun : développer des habiletés langagières propres à un domaine spécialisé auquel ils se destinent. Venons-en maintenant à l'enseignement du lexique en COS.

2.3.2 Approches didactiques à l'œuvre

En COS, on distingue trois approches pédagogiques selon leur point d'accès à l'enseignement du lexique : approches dites 词本位 (cǐ běn wèi « entrée par le 'mot'¹⁸ »), 字本位 (zì běn wèi « entrée par le caractère ») et 语素本位 (yǔ sù běn wèi « entrée par le morphème »).

En didactique du chinois L2, un débat s'est articulé depuis les années 1990 autour de 本位 (běn wèi « unité de base », soit l'unité primitive d'enseignement, comme le fait remarquer Bellassen (2010, p. 30; Bellassen et al., 2012, p. 25-42). Selon les partisans de l'approche 词本位 (cǐ běn wèi « entrée par le 'mot' »), l'enseignement du lexique doit être ancré dans les activités de communication. Les unités lexicales, considérées comme la brique de base du discours, sont enseignées comme entités indécomposables. Cette approche est devenue prédominante depuis son émergence dans les

¹⁸ Rappelons que le « mot » fait référence à l'unité lexicale dans ce paragraphe (voir aussi la note de bas de page 8).

années 1960-1970 (voir la sous-section 1.1.2). Elle a été massivement mise en œuvre dans les classes de COS et présidé à la conception de matériel didactique. Dans les ressources d'apprentissage du lexique que nous avons recensées¹⁹, la description d'un mot composé n'aborde pas ses morphèmes constitutifs. Par exemple, 硬盘^{yìng pán} (« disque dur ») n'est abordé que comme une entité unique. Les deux morphèmes qui le composent, à savoir 硬^{yìng} (« dur » « physique ») et 盘^{pán} (« disque » « disque »), ne sont pas expliqués : on ne décrit pas leur sens ni les autres possibilités combinatoires dans lesquelles on peut les retrouver (telles que 硬件^{yìngjiàn} [yìng jiàn] « dur-pièce » « matériel informatique », 光盘^{guāng pán} (« lumière-disque » « disque optique »)). Une telle méthode de description, selon Bellassen (2010, p. 39), réduit la « transparence sémantique »²⁰ d'une grande partie du lexique chinois à néant. Pour les apprenants non sinophones, cela alourdit grandement la tâche de mémorisation et ne permet pas, de ce fait, une acquisition efficace du lexique.

D'autres didacticiens insistent sur la relation entre la langue chinoise et son système d'écriture. Ils soutiennent l'approche 字本位^{zì běn wèi} (« entrée par le caractère »). Selon cette approche, le caractère constitue l'unité de base en enseignement du lexique et doit être abordé sous ses facettes graphique, sémantique et combinatoire (Bellassen et Zhang, 1990, 1991 ; Zhang, 1992 ; Xu, 1994a, 1994b, 2005 ; Bellassen et Wong, 1995 ; Lu, 2011). L'avantage de cette approche est qu'elle assiste les apprenants, notamment ceux qui proviennent des pays n'appartenant pas à la sinosphère, dans l'acquisition de caractères. Cette tâche constitue un défi majeur pour ces apprenants, vu qu'un caractère résulte d'une association arbitraire entre un ou plusieurs sons, un ou plusieurs sens et une forme graphique. Bien qu'ardue, cette tâche constitue une étape incontournable dans l'acquisition du lexique chinois.

¹⁹ Nous avons analysé des ressources didactiques élaborées depuis les deux dernières décennies, dont Li et Ding (2002), An et Shan (2005, 2008), Zhang (2007), Li (2008a, 2008b, 2008c, 2008d), Bai (2014), Du (2011a, 2011b, 2012b, 2012a), Han (2012), Xu et al. (2012), Xu (2013), Jin et Guo (2017), Lu et al. (2017), et Zhang (2014).

²⁰ La « transparence sémantique » se réfère ici à la compositionnalité d'une unité lexicale. Par exemple, 无线网^{wú xiàn wǎng} ([wú xiàn wǎng] « sans-fil-réseau » « réseau sans fil ») est sémantiquement transparente vu que son sens est constitué de la combinaison des sens de chacune de ses composants. Notons que le concept de transparence sémantique s'applique aussi aux combinaisons lexicales. Par exemple, l'expression 访问网站^{fǎngwèn wǎngzhàn} (« visiter-site web » « visiter un site web ») est sémantiquement transparente.

Par contre, l'approche 字本位 (^{zì běn wèi} «entrée par le caractère») a ses limites quant au traitement sémantique du lexique. En fait, tous les caractères ne sont pas porteurs de sens. Comme nous l'avons mentionné dans la sous-section 1.1.1, les unités minimales du sens en chinois sont les morphèmes, qui peuvent correspondre à un seul caractère (par exemple, 电 ^{diàn} 'électricité' «électricité») ou à une combinaison de caractères (par exemple, 垃圾 ^{lā jī} 'déchet' «déchet»). Les unités lexicales sont composées d'un seul ou de plusieurs morphèmes. De ce fait, la description sémantique du lexique doit se fonder sur les morphèmes au lieu des caractères.

Partant de ce constat, certains didacticiens proposent une approche pédagogique alternative basée sur le morphème (Sheng, 1990 ; Lü, 1999). Cette approche considère le morphème comme point d'accès à l'enseignement du lexique et accorde ainsi une grande importance au développement de la conscience morphologique²¹. Cette initiative pédagogique a été davantage développée au cours des deux dernières décennies. De nombreuses recherches empiriques ont montré les effets bénéfiques de la conscience morphologique sur l'acquisition des compétences lexicales chez les apprenants non sinophones (Feng, 2002, 2003, 2009; Feng et Song, 2004; Gan, 2009; Xing, 2003, 2006; J. Wang et Xing, 2010; W. Zhao, 2017). Plus précisément, l'acquisition des connaissances morphologiques assiste les apprenants dans la réalisation de différentes tâches d'apprentissage du lexique : l'interprétation du sens d'une lexie composée inconnue, la mémorisation des lexies comportant un même morphème et la distinction sémantique des lexies comportant des morphèmes homonymes. D'autres études ont abordé la mise en application de la pédagogie morphologique (Xiao, 2002 ; Li, 2002 ; Xie, 2004 ; Yang, 2011 ; Poizat, 2011, 2017 ; Jiang, 2016).

La pédagogie 语素本位 (^{yǔ sù běn wèi} «entrée par le morphème») s'appuie en grande partie sur l'approche cognitive en didactique des L2 (Carroll, 1964, 1966 ; Chastain, 1970, 1971), celle-ci faisant partie de notre cadre théorique (voir la section 3.3). La pédagogie centrée sur le morphème vise un

²¹ Le terme *conscience morphologique* (*morphological awareness*) (Carlisle, 1995; Kirby, 2004) est issu de la psycholinguistique. En enseignement du chinois langue seconde, ce terme désigne la capacité des apprenants à reconnaître les morphèmes (à savoir les éléments minimaux porteurs de sens), à identifier la structure interne des unités lexicales et à comprendre les règles de formation des unités lexicales.

enseignement structuré du lexique. Elle fait appel à deux stratégies d’enseignement : la mise en réseau des unités lexicales et la sensibilisation au fonctionnement du système lexical.

Sur le plan sémantique, la construction de réseaux lexicaux à partir de morphèmes permet une présentation structurée et systématique du lexique (J. Liu et Wei, 2013; Z. Wang et Qing, 2004; Yang, 2011). La figure 6 illustre l’application de cette stratégie à la description de lexies d’informatique.

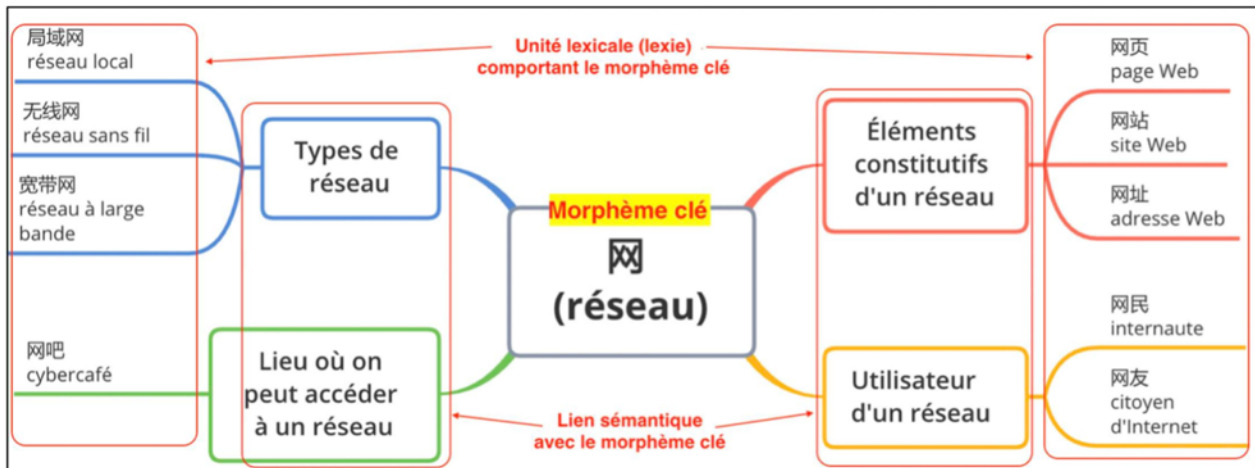


Figure 6. – Réseau lexical construit à partir du morphème ^{wǎng}网 « réseau »

Le réseau lexical s’organise autour du morphème ^{wǎng}网 (« réseau informatique ») et réunit un ensemble de lexies du domaine informatique. Celles-ci sont regroupées en cinq sous-catégories selon la relation sémantique qu’elles entretiennent avec le morphème clé. Certaines composantes de ce réseau, par exemple, ^{wú xiàn wǎng}无线网 (« réseau sans fil »), peuvent contenir un morphème (^{wú xiàn}无线 « sans fil ») qui donne lieu à un autre réseau, dans lequel on trouve ^{wú xiàn shǔ biāo}无线鼠标 (« souris sans fil ») et ^{wú xiàn ěr jī}无线耳机 (« écouteur sans fil »). Cette manière de structurer les connaissances lexicales facilite l’encodage ²², considéré par l’approche cognitive comme élément central de l’apprentissage. La mise en réseau permet ainsi aux apprenants de repérer de façon systématique les lexies acquises. Le tout vise à améliorer l’efficacité de l’acquisition du lexique.

²² Nous utilisons le terme *encodage*, tel que défini par Gagné (1965), pour désigner la structuration de nouvelles informations en vue de leur archivage dans la mémoire des apprenants. Cette notion sera abordée dans la section 3.3.

Sur le plan structurel, l'enseignement explicite des règles morphosyntaxiques vise à sensibiliser les apprenants aux particularités structurelles de la langue chinoise. D'une part, la plus grande partie du lexique chinois se forme par composition. D'autre part, les unités lexicales, les syntagmes (à savoir les combinaisons lexicales) et les phrases sont construits suivant les mêmes règles morphosyntaxiques, comme le montre le tableau 8²³. L'acquisition consciente de ces règles permet de comprendre non seulement le fonctionnement du système lexical, mais aussi la formation des énoncés et leur fonctionnement dans les discours.

Le tableau 8 illustre les cinq types de structures morphosyntaxiques en chinois, à partir desquelles sont formées les différentes unités linguistiques (Zhiwei Lu, 1964; Chao, 1968; Y. Liu et al., 2004, p. 5-15). Les exemples présentés sont issus du domaine informatique. Nous pouvons en dégager certaines caractéristiques structurelles du chinois. Par exemple, une unité lexicale composée (计算 'compter-calculer' « calculer ») peut jouer le rôle d'un morphème libre et entrer dans la constitution d'une autre unité lexicale (计算机 'calculer-machine' « ordinateur »). De plus, une combinaison lexicale peut être associée à différentes structures internes et revêt ainsi des sens différents : 备份数据 ('sauvegarder-données') peut signifier « sauvegarder des données » ou « données de sauvegarde » en fonction de la relation syntaxique entre ses deux composantes. La distinction sémantique est d'autant plus difficile qu'il s'agit d'un syntagme complexe. Par exemple, 下载备份数据 ('télécharger-sauvegarder-données') signifie « télécharger et sauvegarder des données » ou « télécharger des données de sauvegarde »

²³ Nous nous limitons à présenter la formation des unités lexicales et des syntagmes. La construction phrastique, qui ne fait pas l'objet de notre étude, est ainsi exclue du tableau.

| Type de structure morphosyntaxique | Type d'unité linguistique | Forme écrite | Sens littéral | Équivalent français |
|--|---|---|---|---|
| Construction coordonnée | Unité lexicale | cúnchǔ 存储 | 'conserver-stocker' | « stocker » |
| | | jì suàn 计算 | 'compter-calculer' | « calculer » |
| | Combinaison biléxicale | shū rù shūchū [输入][输出] | ['transporter-entrer'] ['transférer-sortir'] | « entrée-sortie » |
| | Syntagme complexe ²⁴ | yǒuxiàn wúxiàn wǎnglù [有线][无线][网络] | ['avoir-fil'] ['sans-fil'] ['réseau-filet'] | « réseaux filaire et sans fil » |
| Modificateur-Modifié | Unité lexicale | yìngpán 硬盘 | 'dur-disque' | « disque dur » |
| | | jì suàn jī [计算][机] | ['compter-calculer'] ['machine'] | « ordinateur » |
| | Combinaison biléxicale | yí dòng yìngpán [移动][硬盘] | ['déplacer-bouger'] ['dur-disque'] | « disque dur amovible » |
| | | shā dù ruǎnjiàn [杀毒][软件] | ['tuer-virus'] ['souple-pièce'] | « logiciel antivirus » |
| Syntagme complexe | miǎnfèi fān yì ruǎnjiàn [免费][翻译][软件] | ['exempter-frais'] ['tourner-traduire'] ['souple-pièce'] | « logiciel de traduction gratuit » | |
| Verbe-COD | Unité lexicale | bèifèn 备份 | 'préparer-exemplaire' | « sauvegarder » |
| | Combinaison biléxicale | xiàzài gēngxīn [下载][更新] | ['décendre-charger'] ['modification-nouveau'] | « télécharger une mise à jour » |
| | | chuán shū shù jù [传输][数据] | ['faire passer-transporter'] ['chiffre-preuve'] | « transférer des données » |
| | Syntagme complexe | xiàzài bèifèn shù jù [下载][备份][数据] | ['décendre-charger'] ['préparer-exemplaire'] ['chiffre-preuve'] | « télécharger et sauvegarder des données » « télécharger des données de sauvegarde » |
| Verbe-Complément circonstanciel | Unité lexicale | qīnglíng 清零 | 'éliminer-zéro' | « remettre à zéro » |
| | | jiàn rù 键入 | 'enfoncer une touche-entrer' | « entrer quelque chose au clavier » |
| | Combinaison biléxicale | huī fù chū chǎng [恢复][出厂] | ['rétablir-restaurer'] ['sortir-usine'] | « restaurer quelque chose à son état d'usine » |
| Syntagme complexe | tóng bù zhì yún pán [同步][至][云盘] | ['uniformiser-marche'] ['à'] ['nuage-disque'] | « synchroniser sur le nuage informatique » | |
| Sujet-Prédicat²⁵ | Unité lexicale | hóngfàn 洪泛 | 'inondation-déborder' | « inondation ²⁶ » |
| | Combinaison biléxicale | xì tǒng bēngkuì [系统][崩溃] | ['système-ensemble'] ['s'effondrer-briser'] | « un système informatique plante » |
| | | huì huà guò qī [会话][过期] | ['réunion-dialogue'] ['dépasser-délai'] | « une session expire » |
| | Syntagme complexe | cí pán kōngjiān bù zú [磁盘][空间][不足] | ['magnétique-disque'] ['vide-espace'] ['pas-suffisant'] | « espace disque insuffisant » |

Tableau 8. – Typologie des structures morphosyntaxiques en chinois

²⁴ Nous entendons par « syntagme complexe » un groupement lexical composé d'au moins trois lexies.

²⁵ En chinois, le prédicat syntaxique peut être un verbe (崩溃 ^{bēngkuì} « planter ») ou un adjectif (不足 ^{bù zú} « insuffisant »).

²⁶ Le terme *inondation* est défini comme « (t)echnique de piratage qui consiste à inonder un serveur de milliers de requêtes simultanées, dans le but de le saturer et d'entraîner sa défaillance » (Source : *Le Grand dictionnaire terminologique*. Repéré à < http://gdt.oqlf.gouv.qc.ca/ficheOqlf.aspx?Id_Fiche=8363693 >, consulté le 17 août 2020). Ce terme français et son équivalent chinois (洪泛 ^{hóngfàn}) véhiculent un sens métaphorique dans le domaine de l'informatique. La forme lexicale de ces termes est associé à un autre sens non spécialisé en français et en chinois.

Pour sensibiliser les apprenants non natifs à ces particularités morphosyntaxiques, la méthode pédagogique que nous présentons ici part de l'analyse sémantique des morphèmes, pour procéder à l'identification de la structure interne des lexies, ce qui mène enfin à l'acquisition des règles de formation des énoncés. Au cours de ce processus d'apprentissage, l'enseignant peut adopter une approche déductive (expliquer une règle et proposer ensuite des exercices) ou inductive (amener les apprenants à découvrir une règle au moyen d'activités d'observation).

La pédagogie axée sur le morphème est mise en œuvre davantage dans des activités en classe que dans la conception de matériel didactique (Ju, 2005; J. Wang et Xing, 2010; Jiang, 2016). Dans les classes de chinois L2, une séquence d'apprentissage lexicale basée sur le morphème consiste généralement en trois phases. D'abord, l'enseignant propose des activités de mise en situation, afin d'activer les connaissances antérieures des apprenants. Ceux-ci sont invités à faire appel à leur bagage linguistique pour, par exemple, identifier le sens d'un morphème ou regrouper des unités lexicales composées d'un même morphème en fonction de leur structure interne. Ensuite, l'enseignant procède à la présentation du nouveau matériel, qui se déroule en deux étapes : l'analyse morphosyntaxique d'une unité lexicale suivie de la construction d'un réseau. Par exemple, pour enseigner le mot 网课 (wǎng kè « cours en ligne »), l'enseignant le décompose d'abord en deux morphèmes (网 wǎng « en ligne » et 课 kè « cours »), il explique ensuite que cette unité lexicale est une construction endocentrique (modificateur-nom modifié), le premier morphème 网 wǎng étant le modificateur du deuxième morphème 课 kè. Par la suite, l'enseignant incite les apprenants à rappeler d'autres unités lexicales comportant le morphème 网 wǎng, telles que 网站 (wǎngzhàn « site Web »), et à deviner le sens de certaines unités lexicales inconnues, telles que 网店 (wǎngdiàn « boutique en ligne ») et 网购 (wǎnggòu « achat en ligne »). Ces activités amènent les apprenants à découvrir le mode de composition des unités lexicales et à distinguer les deux sens exprimés par le morphème 网 wǎng (« réseau » et « en ligne ») dans différentes lexies contenant ce morphème. Enfin, l'enseignant propose des exercices contextualisés pour entraîner les apprenants à appliquer le contenu enseigné dans différentes situations. Un de ces exercices, à titre d'exemple, consiste à choisir le mot approprié parmi une liste des mots composés d'un même morphème pour compléter une phrase ou un dialogue. Pour résumer, une telle séquence d'apprentissage vise à entraîner les apprenants à décoder des mots inconnus à partir d'un morphème acquis, ce qui favorise l'expansion de leur vocabulaire et améliore ainsi leur performance en lecture.

Quant à la conception de matériel didactique, les connaissances morphologiques ne sont pas intégrées à la description du lexique. Les propriétés sémantique et combinatoire des morphèmes ne sont abordées que de façon ponctuelle dans certains manuels de COS (An et Shan, 2005 ; Du, 2012a, 2012b ; Jin et Guo, 2017). À notre connaissance, aucune méthode de description lexicale en usage ne rend compte des morphèmes de façon systématique.

En guise de conclusion, les trois approches présentées abordent l'enseignement-apprentissage du lexique via différents points d'entrée. Dans le cadre de notre thèse, nous nous intéressons en particulier au principe de l'enseignement structuré du lexique. Nous appliquerons ce principe, ainsi que les méthodes didactiques qui en découlent, à la description de la combinatoire spécialisée.

2.3.3 Ressources d'apprentissage mises au point

De façon générale, les manuels de COS sont conçus pour soutenir un « apprentissage incident » du lexique. Grossmann (2011, p. 164) a dressé un constat similaire en didactique du français. Dans les classes de COS, l'enseignement du lexique est souvent intégré à des activités de compréhension écrite. Ainsi, les unités lexicales sont décrites en vue d'un décodage rapide. Dans un bon nombre de ressources didactiques que nous avons recensées, y compris des manuels du chinois scientifique et technique (An et Shan, 2005, 2008 ; Bai, 2014 ; Du, 2011a, 2011b, 2012a, 2012b ; Han, 2012 ; Lu et al., 2017 ; Xu et al., 2012), du chinois de l'informatique (S. Zhang, 2014), du chinois de la médecine traditionnelle chinoise (MTC) (Xu, 2013 ; Jin et Guo, 2017), du chinois du droit des affaires (T. Zhang, 2007) et du chinois du commerce (L. Li et Ding, 2002 ; X. Li, 2008a, 2008b, 2008d, 2008c), la description lexicale prend la forme d'un glossaire bilingue, comme l'illustrent les figures 7 et 8. Les entrées du glossaire sont sélectionnées selon leur valeur communicative, c'est-à-dire leur utilité pour la compréhension d'un texte donné.

| | Information grammaticale | Transcription phonétique (pinyin) | Équivalent anglais |
|---------|--------------------------|-----------------------------------|------------------------------|
| 1. 温室效应 | (名) | wēnshì xiàoyīng | greenhouse effects |
| 2. 阻止 | (动) | zǔzhǐ | to hinder; to prevent |
| 3. 反射 | (动) | fǎnshè | to reflect |
| 4. 辐射 | (动) | fúshè | to radiate |
| 5. 宇宙 | (名) | yǔzhòu | universe |
| 6. 频繁 | (形) | pínpán | frequentl; repeated |
| 7. 加剧 | (动) | jiājù | to exacerbate; to accelerate |
| 8. 生态环境 | (名) | shēngtài huánjīng | ecological environment |
| 9. 不良 | (形) | bùliáng | harmful; not good |

Figure 7. – Description lexicale dans le manuel ^{kē jì hàn yǔ} 科技汉语: ^{zhōng jí yuè dú jiào chéng} 中级阅读教程 [Chinois scientifique et technique : manuel de lecture intermédiaire] (An et Shan, 2005, p. 17)

| | | | |
|----|--|-----------------------------------|--------------------|
| 1 | 独资化 (dúzīhuà): privatize | | |
| 2 | 比重 (bǐzhòng): proportion; ratio | Transcription phonétique (pinyin) | |
| 3 | 受让 (shòuràng): assign; authorize | | |
| 4 | 股权 (gǔquán): stockownership | | |
| 5 | 增资扩股 (zēng zī kuò gǔ): increase capital and stock shares | | |
| 6 | 外方控股 (wàifāng kònggǔ): a foreign company has controlling interest | | Équivalent anglais |
| 7 | 股东转让 (gǔdōng zhuǎnràng): a share holder transfers (his/her) shares | | |
| 8 | 审批 (shēnpī): approve; approval | | |
| 9 | 股权交割 (gǔquán jiāogē): deliver stocks | | |
| 10 | 份额 (fèn'ér): share; portion | | |
| 11 | 合资变局 (hézī biànjú): joint venture transformation | | |
| 12 | 战略管理咨询公司 (zhànlüè guǎnlǐ zīxún gōngsī): strategic management and consulting firm (company) | | |
| 13 | 回避冲突 (huìbì chōngtū): avoid conflict | | |
| 14 | 赢利 (yínglì): profit; bonus | | |

Figure 8. – Description lexicale dans le manuel ^{fǎ lǜ hàn yǔ} 法律汉语: ^{shāng shì piān} 商事篇 [Chinois juridique : droit des affaires] (T. Zhang, 2007, p. 106)

Les réseaux lexicaux sont utilisés dans l'élaboration d'activités pour les séances de lecture. Par exemple, en chinois de la médecine traditionnelle chinoise (MTC), certains manuels (Xu, 2013 ; Jin et Guo, 2017) amènent les apprenants à classer les termes en réseaux conceptuels portant, par exemple, sur les phénomènes naturels et le système fonctionnel du corps humain. La figure 9 illustre une activité d'apprentissage proposée dans la première leçon du manuel *Chinois de la médecine traditionnelle chinoise : compréhension et expression orales* (Jin et Guo, 2017, p. 2). Cette leçon a trois objectifs d'apprentissage : en ce qui concerne les compétences linguistiques, les apprenants ont à acquérir des termes en rapport avec les concepts 阴 (« yīn ») et 阳 (« yáng »); en ce qui concerne les connaissances spécialisées, les apprenants ont à comprendre les influences de ces concepts sur la MTC et à distinguer les syndromes relevant de *yīn* et de *yáng*.

L'activité proposée consiste à classer un nombre de termes en deux réseaux conceptuels qui s'articulent autour de *yin* et de *yang*. On y trouve deux types de termes : 1) les termes désignant des éléments de la nature, ainsi que des états ou des faits naturels, dont l'eau, le feu, les quatre saisons, le jour et la nuit, le mouvement et le calme; 2) les termes désignant les éléments du système fonctionnel du corps humain, tels que des parties du corps et les méridiens humains.

第一课
Lesson 1

阴 阳
Yin and Yang

学习目标 Learning Objectives

1. 能听懂并理解“阴”“阳”及相关词汇，并能举出具体的例子。
To be able to understand *yin, yang*, and other relevant terms in the audio material and to be able to provide specific examples.
2. 能说出“阴”“阳”这两个概念对中医的影响。
To be able to tell the influence of *yin* and *yang* on traditional Chinese medicine.
3. 能说出并区分“阳证”和“阴证”的症状。
To be able to tell and differentiate the *yang*-syndromes and *yin*-syndromes.

Objectifs d'apprentissage :
Acquisition des termes en rapport avec les concepts 阴 « *yin* » et 阳 « *yang* »

主题热身 Topic Lead-in

分组讨论，将下列词语归类。
Discuss in groups and categorize the following words.

自然界：秋冬、白昼、下、夜晚、上、火、轻、水、春夏、上浮、重、动、静、下降

Termes désignant des phénomènes naturels

人体：四肢外侧、三焦、人体上部、心、小肠、肝、背、腹、脾、人体下部、体表、体内、四肢内侧、肺、肾、大肠、膀胱、胆、胃

Termes désignant le corps humain

阴

阳

Classer les termes en deux réseaux conceptuels :
1. le réseau conceptuel de 阴 « *yin* » ;
2. le réseau conceptuel de 阳 « *yang* »

zhōng yī hàn yǔ tīngshuō jiàochéng

Figure 9. – Activité de mise en réseau des termes dans le manuel *中医汉语听说教程 [Chinois de la médecine traditionnelle chinoise : compréhension et expression orales]* (Jin et Guo, 2017, p. 2)

Au sein d'un réseau élaboré, les apprenants sont amenés à réaliser des tâches précises, par exemple, mettre les termes en ordre (tels que les termes désignant les méridiens humains) ou identifier les antonymes (J. Xu, 2013, p. 4), comme l'illustre la figure 10.

(三) 根据课文, 写出下列词语的反义词
Write the antonyms of the following words based on the text

| | | | | | | |
|----------|--------------|------------|-----------|---------------|--|------------------|
| yin 阴 | | nuit 黑夜 | | nord 北 | | excitation 兴奋 |
| | soleil 太阳 | | ciel 天 | chaleur 温暖 | | tête 头 |
| | | | | | | dos 背部 |

进一步理解“阴阳”的含义 **Comprendre davantage les concepts 阴 « yin » et 阳 « yang »**

Figure 10. – Exercice portant sur l’antonymie dans le manuel *中医汉语综合教程 [Chinois de la médecine traditionnelle chinoise : cours intégré]* (J. Xu, 2013, p. 4)

Le recours aux réseaux lexicaux a un double objectif. Cette méthode permet, d’une part, d’expliciter les relations sémantiques entre les termes mis en réseau et, d’une part, de remettre à niveau des connaissances spécialisées des apprenants. C’est notamment le cas de la MTC, puisque certaines lexies, dont 阴 (« yin »), 阳 (« yang ») et 和 (« harmonie »), reflètent la base philosophique et symbolique de la MTC. D’autres lexies correspondent à des concepts qui sont propres à la MTC, par exemple, 寒 (« froid »), 热 (« chaleur ») et 湿 (« humidité »). Cette méthode d’enseignement du lexique spécialisé permet ainsi aux apprenants d’acquérir des connaissances du domaine abordé. Par contre, les enseignants n’utilisent pas la stratégie de mise en réseau pour décrire le lexique de façon systématique.

Quant aux liens sémantiques, certains manuels se limitent à aborder la synonymie et l’antonymie au moyen d’exercices contextualisés (An et Shan, 2005 ; Du, 2012a, 2012b ; Xu, 2013). Par exemple, on propose des exercices de substitution en contexte (remplacer un terme par son synonyme dans une phrase donnée), des exercices d’appariement (associer un terme avec ses synonymes ou antonymes) ou des textes à trous (identifier les liens de synonymie ou d’antonymie selon le contexte). D’autres liens sémantiques, qu’ils soient paradigmatiques ou syntagmatiques, ne sont pas pris en compte. En dernière analyse, les relations sémantiques des unités lexicales ne sont pas décrites de façon explicite dans les manuels recensés. Par conséquent, la combinatoire spécialisée est passée sous silence.

Rares sont les tentatives de traitement systématique de la polysémie dans les manuels didactiques, hormis Xu (2013). Dans ce manuel du chinois de la MTC, chaque leçon contient des modules de description lexicale, en complément à une explication de texte. Dans un module, on regroupe plusieurs lexies ayant une même forme lexicale. Par exemple, le module illustré dans la figure 11 s’articule autour du caractère chinois 畏 qui a deux sens et correspond ainsi à deux

verbes. Le premier verbe désigne les syndromes des patients et signifie ‘craindre’ ou ‘avoir la phobie de’, le deuxième verbe est un terme pharmacologique qui désigne le fait qu’un médicament réduit ou élimine les réactions toxiques ou les effets secondaires d’un autre médicament. Chaque lexie au sein d’un module est décrite de façon systématique. Les renseignements linguistiques consignés comprennent la partie du discours de la lexie vedette, sa définition et des exemples illustrant son emploi dans un contexte particulier. Même si ces informations sont utiles, nous constatons dans ce manuel un manque d’explication systématique des propriétés syntaxico-sémantique des combinaisons lexicales.

Des exercices contextualisés sont élaborés pour faire la distinction entre les lexies d’un même module. Par exemple, on demande aux apprenants d’expliquer le sens du caractère 畏 dans des phrases données (figure 11). La pédagogie basée sur le morphème est aussi appliquée. Chaque lexie vedette associée au caractère 畏 peut agir comme un morphème libre. On demande aux apprenants d’identifier les unités lexicales qui renferment ce morphème (figure 11). Cet exercice permet aux apprenants d’associer le nouveau contenu enseigné à leurs connaissances antérieures.

Forme lexicale polysémique

(三) 畏

| Partie du discours | Définition | Exemples d'usage |
|--------------------|--|---------------------|
| 词性 | 释义 | 举例 (译释) |
| Sens 1 动词 | 害怕、不喜欢 Craindre, avoir la phobie de | 畏光 (害怕光亮) 畏难 畏寒 |
| Sens 2 动词 | 相畏: 中药学术语, 即一种药物的毒性反应或副作用, 能被另一种药物减轻或消除。 Un médicament réduit ou élimine les réactions toxiques ou les effets secondaires d'un autre médicament | 半夏畏生姜 (生姜可以减轻半夏的毒性) |

1. 写出你知道的含有“畏”字的词语。 ← Identifier les unités lexicales qui renferment le caractère 畏

2. 试着解释下列句子中“畏”字的意思。 ← Expliquer le sens associé au caractère 畏 dans les phrases suivantes

① 学习汉语不能有畏难情绪。 _____

② 她的身体很虚弱, 总是畏寒, 手脚冰凉。 _____

③ 常山畏陈皮, 意思是陈皮可以缓和常山引起恶心、呕吐等反应。 _____

Figure 11. – Module d’apprentissage des lexies de forme identique dans le manuel ^{zhōng yī hàn yǔ zōng hé jiào} 中医汉语综合教程 [Chinois de la médecine traditionnelle chinoise : cours intégré] (J. Xu, 2013, p. 6)

D'autres travaux sensibilisent les apprenants au fonctionnement du système lexical du chinois. Plusieurs manuels sur le chinois scientifique et technique (An et Shan, 2005 ; Du, 2012a, 2012b) proposent d'expliquer les modes de formation des lexies composées. Ces explications renferment des exemples représentatifs du domaine abordé. Par exemple, des morphèmes à haute fréquence tels que 体^{tǐ} (« état de substance ») sont cités comme exemples pour expliquer la formation des termes 液体^{yè tǐ} (« substance liquide ») ou 绝缘体^{jué yuán tǐ} (« substance diélectrique »), comme l'illustre la figure 12. Notons que ces renseignements sont donnés dans les modules consacrés à la grammaire, au lieu d'être intégrés à la description lexicale.

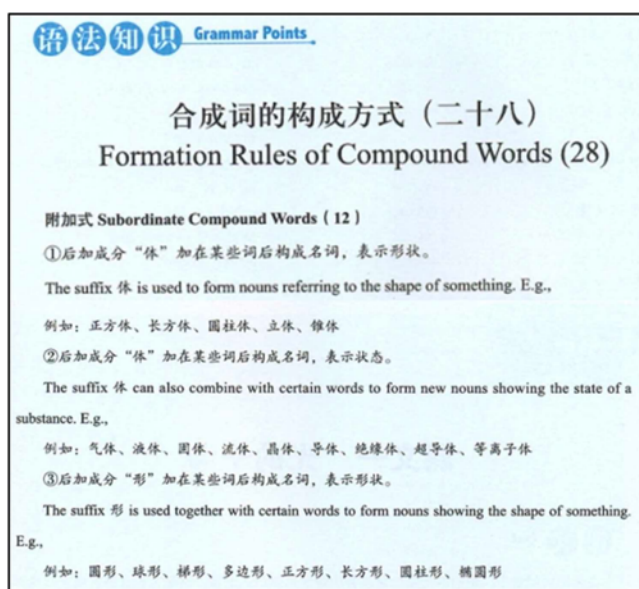


Figure 12. – Explication grammaticale portant sur la formation des lexies composées dans le manuel *基础科技汉语教程：阅读课本（下）* [Chinois scientifique et technique élémentaire : compréhension écrite (volume 2)] (Du, 2012b, p. 4-5)

Pour conclure, les ressources didactiques recensées intègrent l'apprentissage du lexique dans les tâches de lecture et d'explication de textes. Dans ces ouvrages, la description lexicale, visant plutôt un décodage rapide qu'une bonne maîtrise du lexique, gagnerait à être enrichie de différentes manières : la distinction fine des lexies ayant la même forme, l'enseignement explicite et systématique des relations sémantiques entre les lexies et la mise en réseau des lexies ayant une affinité sémantique.

2.4 Synthèse

Dans ce chapitre, nous avons présenté des méthodes descriptives du lexique en vue de son enseignement. Nous avons examiné de plus près l'enseignement du lexique spécialisé auprès des locuteurs non natifs ainsi que les méthodes utilisées pour l'enseignement du COS, ce qui s'accorde avec notre contexte de recherche. Des travaux antérieurs (Binon et Verlinde, 2003, 2004; Dechamps, 2004; Pecman, 2012; Cavalla, 2018b) semblent montrer qu'une description fine et systématique du lexique spécialisé favorise son acquisition par les apprenants non natifs. Dans cette recherche, nous nous attachons à combler certaines lacunes observées en COS. Nous nous inspirons des travaux recensés tout en apportant notre propre contribution à la didactique de la combinatoire spécialisée chinoise.

Chapitre 3 Cadre théorique

Notre cadre théorique se décline en trois volets, axés respectivement sur la lexicologie, la terminologie et la didactique des langues secondes. Nous présentons d'abord la Lexicologie explicative et combinatoire (LEC) (Mel'čuk et al., 1995) qui nous sert d'assise pour définir les principes de base de la description du lexique chinois (section 3.1). Nous exposons l'approche lexico-sémantique de la terminologie (section 3.2). Cette approche guide notre démarche méthodologique du traitement du lexique spécialisé. La section 3.3 est consacrée à l'approche cognitive en didactiques des langues secondes. Nous nous appuyons sur cette approche pour envisager une présentation à la fois structurée et intuitive des connaissances lexicales.

3.1 Lexicologie explicative et combinatoire (LEC)

La LEC constitue le volet lexicologique de notre cadre théorique. Dans ce qui suit, nous exposons d'abord les propositions théoriques de la LEC (3.1.1), pour aborder ensuite l'application de la LEC à la confection de ressources lexicales (3.1.2) et à la didactique des langues (3.1.3). Nous justifions enfin la pertinence de la LEC comme cadre théorique de notre thèse (3.1.4).

3.1.1 Notions et principes fondamentaux de la LEC

La LEC est une théorie lexicologique qui s'inscrit dans le cadre de la Théorie Sens-Texte (TST) (Žolkovskij et Mel'čuk, 1965)²⁷, une théorie linguistique qui présuppose la correspondance entre le sens et l'ensemble des énoncés paraphrastiques qui expriment ce sens en langue naturelle, à savoir le texte. La TST vise « la description de la correspondance Sens↔Texte [...] au moyen de la construction de modèles formels » (Polguère, 1998, p. 4). La modélisation des phénomènes linguistiques en TST « part d'une représentation sémantique pour construire des arbres

²⁷ Faute d'espace, nous ne donnons pas une liste exhaustive de la littérature portant sur la Théorie Sens-Texte. Nous nous limitons donc à des travaux fondamentaux de la TST, dont Mel'čuk (1981) Mel'čuk (1992), Mel'čuk (1993), Mel'čuk et al. (1995), Mel'čuk (1997), Polguère (1998), Mel'čuk (2003) et Milićević (2006). D'autres repères bibliographiques de la littérature Sens-Texte ont été proposées par Mel'čuk et al. (1992, p. 59-66), Polguère (1998) et Milićević (2006).

syntaxiques À L'AIDE DU LEXIQUE » (Mel'čuk et al., 1995, p. 17), d'où la place centrale de la lexicologie dans les recherches sémantiques Sens-Texte, comme le signale Mel'čuk (1997, p. 32).

La LEC vise la description formelle des propriétés linguistiques du lexique. Elle constitue le cadre théorique pour l'élaboration du *Dictionnaire explicatif et combinatoire* (DEC) (Mel'čuk, 1992 ; Mel'čuk et al., 1984, 1999 ; Mel'čuk et Clas, 1988, 1992). Les principes sous-jacents à la compilation du DEC sont axés non seulement sur la description des lexies particulières, mais aussi sur la modélisation du lexique dans son ensemble. La LEC met en avant un modèle descriptif sur lequel nous nous appuyons pour envisager une méthode de description adaptée à la spécificité du lexique spécialisé.

Il convient d'explicitier les notions de base sur lesquelles repose notre recherche. Nous précisons de quelle manière ces notions s'intègrent au cadre théorique qui guide notre démarche méthodologique.

Lexie (unité lexicale)

La *lexie*, nommée aussi *unité lexicale*, désigne l'objet d'étude de la lexicologie. Une lexie est un mot (*lexème*) ou une locution (*phrasème*) pris(e) dans une seule acception bien circonscrite (Mel'čuk et al., 1995, p. 15-16). La lexie constitue l'unité de base du DEC. Plus précisément, chaque article du DEC décrit une lexie distincte, munie de tous les renseignements qui spécifient totalement son comportement dans un texte (Mel'čuk et al., 1995, p. 56-57). La lexie, comme le souligne Polguère (2016, p. 292), est « l'unité de structuration du lexique; sa description est donc l'unité de description du lexique ».

En chinois, une lexie est l'unité minimale qui véhicule un sens lexical et fonctionne de façon autonome. Elle peut contenir un seul morphème ou une combinaison de morphèmes, comme nous l'avons expliqué dans la sous-section 1.1.1. Rappelons que les lexies composées d'une combinaison de morphèmes forment la plus grande partie du lexique du chinois.

Dans notre travail, nous nous alignons sur la définition de la lexie proposée par la LEC (Mel'čuk et al., 1995, p. 15-16) pour envisager l'unité de base du lexique spécialisé, soit le *terme*. Un terme est une unité lexicale dotée d'un sens lié à un domaine de spécialité, tel que proposé par l'approche lexico-sémantique de la terminologie (L'Homme, 2020a, p. 36). Nous reviendrons sur

la définition de terme dans la section 3.2. Quant à la description du lexique spécialisé, chaque terme fait l'objet d'un article distinct dans la ressource lexicale que nous envisageons de créer.

Fonction lexicale

La LEC modélise les relations entre lexies au moyen des *fonctions lexicales* (FL), un outil formel conçu sur le modèle des fonctions mathématiques (Mel'čuk et Polguère, 2021). Une FL f « décrit une relation entre une lexie L — appelée l'*argument* de f — et un ensemble de lexies ou de syntagmes appelé la *valeur* de l'*application* de f à la lexie L » (Polguère, 2016, p. 196).

Les FL encodent deux types de relations lexicales récurrentes en langue : les relations *paradigmatiques* qui associent une lexie à ses dérivés sémantiques ainsi que les relations *syntagmatiques* qui décrivent les collocations contrôlées par une lexie (Mel'čuk et al., 1995, p. 129; Mel'čuk et Polguère, 2007, p. 17-22; Polguère, 1998, p. 11). Polguère (2003, p. 125-126) souligne que les FL doivent être considérées comme des « métalexies » étant donné qu'elles sont caractérisées par des propriétés de sens et de combinatoire. Elles existent au même titre que les notions de partie du discours ou de genre grammatical.

Du point de vue formel, une FL est représentée par la formule suivante qui est inspirée du formalisme mathématique $f(x) = y$:

Lien lexical (lexie) = expressions possibles du lien lexical (Mel'čuk et Polguère, 2007, p. 23)

Dans cette formule, la fonction exprime un lien lexical. La lexie entre parenthèses est l'argument ou le mot-clé de la fonction. L'application de la fonction à l'argument retourne comme valeur un ensemble d'expressions linguistiques qui suit le signe d'égalité (Mel'čuk et al., 1995, p. 126; Mel'čuk et Polguère, 2007, p. 23).

Les FL sont indépendantes des langues, comme le souligne Mel'čuk (1992, p. 33-35). Dans ce qui suit, nous illustrons, à l'aide d'exemples concrets, des relations lexicales modélisées au moyen des FL. Les exemples sont tirés du corpus chinois portant sur l'informatique et les technologies de l'information que nous avons élaboré aux fins de cette recherche (voir aussi la sous-section 4.1.1).

FL paradigmatique

Les FL paradigmatiques modélisent les liens de *dérivation sémantique* entre lexies. Il s'agit d'une relation fondée sur une parenté de sens. Une dérivation sémantique s'établit entre deux lexies L₁ et L₂ si les trois conditions suivantes sont remplies, comme l'expliquent Mel'čuk et Polguère (2007, p. 18) :

- 1) L₁ et L₂ possèdent des composantes de sens communes. Par exemple, le sens de la lexie ^{ruǎn}软件 (« logiciel ») [= L₂] fait appel à ^{chéng xù}程序 (« programme ») [= L₁], car un 'logiciel' signifie 'un ensemble de programmes constituant une unité destinée à effectuer un traitement particulier sur un ordinateur' (Office québécois de la langue française, 2020b).
- 2) La relation sémantique entre L₁ et L₂ est récurrente dans la langue. Le lien sémantique établi entre ^{ruǎn jiàn}软件 (« logiciel ») [= L₂] et ^{chéng xù}程序 (« programme ») [= L₁] est également observé entre ^{wǎng luò}网络 (« réseau ») [= L₂] et ^{yìng jiàn}硬件 (« matériel ») [= L₁] étant donné que L₂ est défini au moyen de L₁, qui désigne un élément constitutif de L₂. Plus précisément, 'matériel' se définit comme l'« ensemble des éléments informatiques physiques qu'un utilisateur utilise pour effectuer une tâche sur un ordinateur » (Observatoire de linguistique Sens-Texte, 2015 b), tandis que 'réseau' se définit comme « l'ensemble des moyens matériels et logiciels qui permettent l'échange de données, de services ou de ressources entre plusieurs appareils informatiques, y compris ces appareils eux-mêmes » (Office québécois de la langue française, 2021).
- 3) Il arrive que la relation s'exprime de façon morphologique dans la langue. Notons que cette condition ne s'applique pas à toutes les langues. En chinois, comme les lexies ne subissent pas de dérivation morphologique, les lexies reliées par certaines relations de dérivation sémantique ont souvent la même forme lexicale, donnant ainsi lieu à la polysémie. Nous reviendrons sur la spécificité morphosyntaxique du chinois dans la sous-section 4.1.2.4.

Les dérivations sémantiques représentées par les FL paradigmatiques peuvent être regroupées en trois catégories (Mel'čuk et Polguère, 2007, p. 19) :

- 1) Le premier type de dérivation sémantique s'établit entre deux lexies qui ont (approximativement) le même sens. Ce type regroupe les cas de figure suivants :

synonymie exacte (检测 « mettre à l'essai » et 测试 « tester »), synonymie approximative (攻击 « attaque informatique » et 入侵 « intrusion »), conversion (发送 « envoyer » et 接收 « recevoir » : X 向 Z 发送 Y à Z « X envoie Y à Z » est équivalent à Z 接收 来自 X 的 Y « Z reçoit Y provenant de X »), changement de partie du discours (安装 /nom « installation » et le nom d'action 安装 /v « installer ») et hyperonymie (软件 « logiciel » est le terme générique de 应用 « application »).

- 2) Le deuxième type de dérivation sémantique relie deux lexies ayant des sens opposés, c'est le cas de l'antonymie exacte (加密 « crypter » versus 解密 « décrypter ») et approximative (删除 « supprimer » versus 恢复 « restaurer »). Certains cas de conversion peuvent aussi donner lieu à des sens opposés, comme 下载 (« télécharger ») et 上传 (« téléverser »).
- 3) Le troisième type regroupe les cas où l'une des deux lexies désigne un élément de la situation dénotée par l'autre. Le dérivé sémantique d'une lexie prédicative²⁸ peut correspondre à l'un de ses actants (入侵 « intrusion » est relié à 黑客 « pirate informatique » qui réalise le premier actant d'une (intrusion), 存储 « stockage » est relié à 存储器 « mémoire » qui est la destination d'un (stockage)) ou une caractéristique d'un actant (加密 /v « crypter » → 加密 /adj « crypté » qui décrit l'état d'une entité qui subit l'action de (crypter)).

Les FL paradigmatiques, tout comme les FL syntagmatiques, expriment des liens lexicaux orientés (Mel'čuk et Polguère, 2007, p. 22). Une FL relie son mot-clé à un ensemble de dérivés sémantiques ou à un ensemble de collocatifs. Reprenons quelques exemples présentés plus haut pour illustrer la modélisation sémantique au moyen des FL paradigmatiques :

La relation de synonymie entre 检测 (« mettre à l'essai ») et 测试 « tester » peut être représentée au moyen de la FL [Syn] comme suit :

Syn (检测 « mettre à l'essai ») = 测试 (« tester »).

²⁸ Rappelons que le sens d'une lexie prédicative est exprimé par un *prédicat sémantique* ainsi que les *actants sémantiques* qui participent au fait dénoté (Mel'čuk et al., 1995, p. 76), comme nous l'avons évoqué dans la sous-section 1.1.3.

La FL [S₀] relie son mot-clé à un dérivé sémantique nominal qui a nécessairement la même composante sémantique que le mot-clé (Mel'čuk et al., 1995, p. 133), par exemple, la relation entre le verbe 安装^{ān zhuāng}_{/v} (« installer ») et le nom d'action équivalant à ce verbe est représentée comme suit :

S₀ (安装^{ān zhuāng}_{/v} « installer ») = 安装^{ān zhuāng}_{/nom} (« installation »).

Les FL [S_i] associent une lexie prédicative au nom qui réalise son « i-ème » (1^{er}, 2^e, 3^e) actant (Mel'čuk et al., 1995, p. 135; Polguère, 2016, p. 199). La FL [S₁], par exemple, met en relation le mot-clé (tel que 入侵^{rù qīn} « intrusion ») au nom dénotant son premier actant, soit 黑客^{hēi kè} « pirate informatique » :

S₁ (入侵^{rù qīn} « intrusion ») = 黑客^{hēi kè} (« pirate informatique »).

Il en va de même pour le dérivé sémantique adjectival (Mel'čuk et al., 1995, p. 135). La FL [A₂] modélise le lien entre le mot-clé et un modificateur adjectival d'une lexie qui réalise le deuxième actant du mot-clé.

A₂ (加密^{jiā mì}_{/v} « crypter ») = 加密^{jiā mì}_{/adj} (« crypté »).

Les FL paradigmatiques couvrent aussi les dérivés sémantiques qui se rapportent aux circonstants du mot-clé, tels que le nom d'instrument [S_{instr}]. On modélise ainsi le lien entre 发送^{fā sòng} (« envoyer ») et l'instrument utilisé pour réaliser l'action d'envoyer :

S_{instr} (发送^{fā sòng} « envoyer ») = 电子邮箱^{diàn zǐ yóu xiāng} (« courrier électronique »).

FL syntagmatique

Les FL syntagmatiques modélisent les relations sémantiques au sein des collocations contrôlées par une lexie. Une collocation résulte d'une cooccurrence privilégiée de deux lexies entretenant une relation sémantique et syntaxique (Grossmann et Tutin, 2003 ; Tutin, 2013 ; Tutin et al., 2006). Il s'agit d'une expression semi-idiomatique composée d'une base et d'un collocatif. La base est une lexie sémantiquement autonome. Elle retient son sens dans la collocation et détermine le choix du collocatif (Jousse et Polguère, 2005, p. 7) (voir aussi la section 2.2).

Les FL syntagmatiques décrivent les phénomènes collocationnels sur deux axes : sur le plan sémantique, elles rendent compte du sens exprimé par le collocatif; sur le plan syntaxique, elles

rendent compte du rôle syntaxique que le collocatif joue auprès de la base (Polguère, 2016, p. 201). Polguère (2018) fait remarquer qu'un collocatif est « actanciel » dans le cas où il participe à l'expression d'un des actants de la base. L'Homme et Leroyer (2009, p. 271-272), pour leur part, soulignent aussi que dans de nombreuses collocations, le collocatif fait appel à un actant sémantique de la base. Dans notre corpus, nous avons relevé ce type de collocations, telles que 全量备份 ('totalité-sauvegarde' « sauvegarde complète ») et 增量备份 ('incrément-sauvegarde' « sauvegarde incrémentale »). Ces collocations désignent deux sortes de sauvegarde : la première se fait sur la totalité des données, la deuxième se fait sur les éléments des données non présents lors de la dernière sauvegarde. Les collocatifs adjectivaux 全量 et 增量 sont actanciels puisqu'ils caractérisent le deuxième actant de la base 备份 (« sauvegarde »), soit les données. La structure actancielle de 备份 (« sauvegarde ») est représentée comme suit (tableau 9).

| Premier actant sémantique | | Troisième actant sémantique | | Deuxième actant sémantique | Prédicat sémantique |
|--|----------|-----------------------------|----------|----------------------------|---------------------|
| yòng hù 用户 | zài 在 | cún chǔ jiè zhì 存储介质 | duì 对 | shù jù 数据 | bèi fèn 备份 |
| Utilisateur | Sur | Support de stockage | Envers | Données | Sauvegarde |
| Sauvegarde de <u>données</u> [2 ^e actant] sur <u>support de stockage</u> [3 ^e actant] par <u>utilisateur</u> [1 ^{er} actant] | | | | | |

Tableau 9. – Structure actancielle de 备份 (« sauvegarde »)

Nous présentons ici trois types FL syntagmatiques que nous avons observées dans notre corpus portant sur l'informatique et les technologies de l'information (voir aussi la sous-section 4.1.1). Ces FL décrivent trois types de collocatifs : les modificateurs syntaxiques, les verbes supports et les verbes de mode de fonctionnement (Mel'čuk et Polguère, 2021).

1) Collocatifs modificateurs syntaxiques. Dans une collocation, le modificateur syntaxique fonctionne auprès de la base comme un adjectif si celle-ci est nominale, ou comme un adverbe, si la base est un verbe ou un adjectif. Parmi les FL syntagmatiques qui modélisent ce type de collocations, citons la FL **Epit** qui associe le mot-clé à son épithète pléonastique à valeur stylistique, à savoir une lexie adjectivale ou adverbiale utilisée comme modificateur du mot-clé pour des fins stylistiques et sans réelle motivation informative forte. Par exemple, l'application de la FL **Epit** au nom 入侵 (« intrusion ») retourne comme valeur 非法 (« illicite ») :

Epit (入侵 « intrusion ») = 非法 (« illicite »).

Ce collocatif adjectival signifie ‘qui est fait sans autorisation’. Il s’agit d’une épithète pléonastique de 入侵^{rù qīn} (« intrusion ») qui désigne « accès non autorisé à un système informatique ou à un réseau, obtenu en contournant ou en désamorçant les dispositifs de sécurité en place » (Office québécois de la langue française, 2020a).

2) Collocatif de type verbes supports. Certaines FL modélisent la relation entre le mot-clé et un verbe support. Il s’agit d’un « collocatif verbal sémantiquement vide dans le contexte de la collocation », comme l’explique Polguère (2016, p. 205). Un verbe support sert à « verbaliser » sa base nominale, pour que celle-ci fonctionne dans la phrase comme si elle était un verbe.

Les FL **Oper_i**, **Func_i** et **Labor_i**, qui associent leur mot-clé à un verbe support, rendent compte non seulement de la relation sémantique dans une collocation, mais aussi du rôle syntaxique du mot-clé par rapport au collocatif (Mel’čuk et al., 1995, p. 138-140). Dans une collocation modélisée par ces FL, le mot-clé est respectivement le complément d’objet direct (COD), le sujet grammatical (SG) ou le complément d’objet indirect (COI) de son collocatif. L’indice *i* qui se rattache à chaque FL est un indice actanciel qui est déterminé par le rôle des actants syntaxiques profonds (ASyntP)²⁹ du mot-clé (Mel’čuk et al., 1995, p. 138).

La FL **Oper₁**, par exemple, modélise la relation entre le nom d’action 更新^{gēng xīn} (« mise à jour ») et son collocatif 进行^{jìn xíng} « effectuer ». L’indice 1 signifie que le SG du verbe support (进行^{jìn xíng} « effectuer ») est une lexie (telle que 用户^{yòng hù} « utilisateur ») réalisant le premier actant sémantique (ASyntP I) du mot-clé. La relation entre la base et son collocatif est exprimée comme suit : **Oper₁** (更新^{gēng xīn} « mise à jour ») = 进行^{jìn xíng} (« effectuer »). En d’autres termes, la FL **Oper₁** relie un verbe support qui a comme SG le premier actant sémantique du mot-clé au mot-clé qui est le COD de ce verbe.

3) Collocatif de type verbes de mode de fonctionnement.

Les FL **Obstr**, **Stop** et **Excess** s’appliquent à un mot-clé nominal et retournent comme valeur les verbes qui prennent le mot-clé comme ASyntP I (sujet syntaxique) et dénotent trois modes de

²⁹ Selon la LEC, un *actant syntaxique profond* (ASyntP) de la lexie L se définit comme « un syntagme qui dépend de L syntaxiquement et en exprime un actant sémantique » (Mel’čuk et al., 1995, p. 117).

fonctionnement anormaux du mot-clé : fonctionnement entravé, fonctionnement interrompu et fonctionnement excessif (Mel'čuk et Polguère, 2021).

La FL **Stop** relie à son mot-clé un verbe collocatif signifiant 's'interrompre'. Cette FL s'applique souvent un terme dénotant un processus en cours d'exécution, tel que 连接 (lián jiē « connexion ») et 传输 (chuán shū « transfert »), et retourne comme valeur un verbe exprimant le sens de 's'arrêter de façon anormale'³⁰ :

Stop (连接 lián jiē « connexion ») = 中断 (zhōng duàn « s'interrompre »).

Stop (传输 chuán shū « transfert ») = 中断 (zhōng duàn « s'interrompre »).

La FL **Obstr** associe son mot-clé à un verbe collocatif qui signifie 'ne pas fonctionner comme il faut'. Cette FL s'applique souvent à un terme dénotant une tâche (comme 下载 xià zài « téléchargement ») :

Obstr (下载 xià zài « téléchargement ») = 失败 (shī bài « échoue »).

Il convient de souligner que les collocatifs de type verbes de mode de fonctionnement sont fréquemment observés dans notre corpus.

Distinction entre les relations paradigmatiques et syntagmatiques

Polguère (2016, p. 204) et Jousse (2010, p. 32) soulignent que la distinction entre les relations paradigmatiques et syntagmatiques n'est pas étanche. Il arrive qu'une FL syntagmatique instancie une relation lexicale paradigmatique. Plus précisément, l'application d'une FL syntagmatique peut retourner comme valeur fusionnée³¹ une lexie qui exprime simultanément le sens du mot-clé et la FL (Polguère, 2016, p. 204). Cette valeur fusionnée peut être reliée au mot-clé par une relation paradigmatique comme l'hyponymie.

³⁰La FL **Stop** est équivalente à l'une des deux FL complexes, soit **AntiVerFinFunc0** et **AntiVerFinFact0**, selon le sens de son mot-clé (Mel'čuk et Polguère, 2021). Dans les cas de 连接 (lián jiē « connexion ») et 传输 (chuán shū « transfert »), on peut utiliser la FL **AntiVerFinFunc0** pour exprimer le lien sémantique qui les relie au verbe 中断 (zhōng duàn « s'interrompre »). Nous reviendrons sur les FL complexe plus loin dans cette sous-section.

³¹ Une valeur fusionnée est indiquée par le symbole « // » dans la LEC (Polguère, 2016, p. 204)

La relation entre (^{ruǎnjiàn}软件 « logiciel ») et (^{fáng huǒ qiáng}防火墙 (« coupe-feu »)), par exemple, peut être envisagée sous deux angles. D'une part, les deux lexies entretiennent une relation d'hyponymie qui peut être modélisée par la FL **Hypo** comme suit : **Hypo** (^{ruǎnjiàn}软件 « logiciel ») = (^{fáng huǒ qiáng}防火墙 (« coupe-feu »)). D'autre part, elles sont reliées par une relation syntagmatique puisqu'un 'coupe-feu' est « un logiciel dont la fonction est de prévenir les intrusions non souhaitées par l'utilisateur » (Observatoire de linguistique Sens-Texte, 2013). La modélisation de cette relation syntagmatique fait appel à un appareil descriptif autre que les FL standards simples (Mel'čuk et al., 1995, p. 127-129). Nous y reviendrons un peu plus loin dans cette sous-section.

FL standards simples

Les FL « s'organisent en système autour d'un noyau d'une soixantaine de FL dites standards », comme le soulignent Mel'čuk et Polguère (2021). Une FL est standard si elle satisfait aux conditions suivantes : 1) avoir un nombre élevé d'arguments, soit des mots-clés (Polguère, 2016, p. 196), 2) retourner comme valeur un nombre important de lexies, 3) donner lieu à des valeurs diversifiées, 4) être indépendante des langues (Mel'čuk et al., 1995, p. 127-128; Jousse, 2010, p. 73-75).

Une FL standard est dite simple si elle ne peut être décomposée en d'autres fonctions lexicales (Polguère, 2016, p. 201). La LEC propose au total 66 FL standards simples qui décrivent les relations paradigmatiques et syntagmatiques entre les lexies (Mel'čuk et Polguère, 2021).

Le système des FL est très riche : outre les FL simples, on fait appel aux FL dites « irrégulières » pour décrire les relations entre lexies, c'est notamment le cas de la description des phénomènes collocatifs. Il arrive que la relation entre la base et le collocatif d'une collocation ne puisse être modélisée au moyen d'une FL simple. Dans ce cas, il est nécessaire de recourir à d'autres variétés de FL. Mel'čuk et al. (1995, p. 148-151) énumère trois types de FL irrégulières : FL complexes, configurations de FL et FL non standards. Nous présentons ci-dessous ces types de FL à l'aide d'exemples tirés de notre corpus.

FL « irrégulières » 1 : FL complexes

Une FL complexe se définit comme « un enchaînement de FL simples syntaxiquement liées [...] qui exprime, de façon indécomposable, le sens de l'enchaînement entier » (Mel'čuk et al., 1995, p. 148). Considérons un cas précis d'application de FL complexe. Dans la collocation (^{ān zhuāng yìng yòng}安装应用

(« installer une application »), la relation lexicale entre la base (应用 « application ») et le collocatif verbal (安装 « installer ») est représentée comme suit :

Prepar₁Fact₀^I (应用 « application ») = 安装 (« installer »)

La FL **Prepar_i** modélise la relation syntagmatique entre le mot-clé et un collocatif verbal ayant le sens de ‘préparer pour’. L’indice actanciel *i* signifie que l’action exprimée par le verbe de préparation est réalisée par le « i-ème » actant du mot-clé. Cette FL se combine souvent avec le verbe de réalisation **Fact₀** (Mel’čuk et al., 1995, p. 146).

La FL **Fact_i** associe le mot-clé à un verbe de réalisation qui exprime le sens de ‘réaliser les objectifs inhérents de l’entité désignée par le mot-clé’. L’indice actanciel *i* renvoie à l’ASyntP_{*i*} du mot-clé. Le chiffre romain en exposant indique le degré de réalisation du verbe. Le chiffre I signifie que le verbe exprime la première phase de la réalisation des objectifs du mot-clé (Mel’čuk et al., 1995, p. 142).

Les deux FL standards simples se combinent pour former une FL complexe qui s’applique à la lexie 应用 « application ». La valeur retournée est le verbe 安装 (« installer ») qui exprime le sens de ‘préparer une application pour permettre à l’application de fonctionner’, il s’agit de la première des deux phases de la réalisation des objectifs inhérents du mot-clé. Le verbe de réalisation renvoyant à la seconde phase de 应用 (« application ») est 设置 (« configurer »). Ce verbe exprime le sens de ‘définir les paramètres de fonctionnement d’une application afin qu’il fonctionne en conformité avec du matériel ou un logiciel spécifique’, ce qui implique que l’application est déjà installée dans un système informatique composé du matériel et du logiciel.

FL « irrégulières » 2 : Configuration de FL

La configuration de FL est « une suite de FL simples qui ne sont pas syntaxiquement liées entre elles, mais qui ont le même mot-clé, cette suite ayant une valeur globale cumulative qui exprime de façon indécomposable le sens de la suite entière » (Mel’čuk et al., 1995, p. 149). Nous donnons un exemple concret d’une configuration de FL. Dans la collocation 恶意软件 (« logiciel malveillant »), la relation lexicale entre la base (软件 « logiciel ») et le collocatif (恶意 « malveillant ») est représentée comme suit :

AntiBon + Real₁ (软件 « logiciel ») = 恶意 (« malveillant ») :

Cette configuration de FL comporte une FL complexe **AntiBon** et une FL standard simple **Real_i**. La FL **AntiBon** résulte de la combinaison de deux FL standards simples : la FL **Bon** (pour *Laudatif*) qui associe le mot-clé à son modificateur adjectival ou adverbial exprimant le sens de ‘bon’ en tant qu’approbation subjective (Mel’čuk et al., 1995, p. 136) et la FL **Anti** (pour *Antonyme*). Cette FL complexe, qui signifie ‘le contraire de laudatif’, exprime la qualification négative du mot-clé.

La FL **Real_i** associe le mot-clé à un verbe de réalisation. L’indice 1 signifie que l’action exprimée par le verbe est réalisée par l’ASyntP I du mot-clé. (Mel’čuk et al., 1995, p. 142)

La configuration de FL ainsi formée permet de décrire la relation sémantique au sein de la collocation étudiée : le modificateur 恶意 (à yì « malveillant ») du mot-clé 软件 (ruǎnjiàn « logiciel ») désigne un ‘logiciel qui est conçu dans l’objectif de [Real_i] nuire au [AntiBon] fonctionnement de composants informatiques’. Par ailleurs, il convient de mentionner que la collocation 恶意软件 (à yì ruǎnjiàn « logiciel malveillant ») entretient une relation d’hyponymie avec 软件 (ruǎnjiàn « logiciel »).

Pour conclure sur la présentation des FL, elles modélisent la structure relationnelle du lexique. Cet outil descriptif identifie deux aspects d’une relation lexicale dans le cas des FL syntagmatiques, à savoir son sens et son patron syntaxique, comme le souligne Jousse (2010, p. 71). La LEC, qui applique les FL à la compilation de ressources lexicales, assure une description homogène et systématique des dérivations sémantiques et des combinaisons lexicales restreintes. Cette approche descriptive fournit au locuteur la totalité des choix lexicaux nécessaires pour exprimer sa pensée de façon précise et appropriée (Mel’čuk et al., 1995, p. 129).

- **FL « irrégulières » 3 : FL non standards**

Les FL non standards s’appliquent notamment à la description des collocations très marquées culturellement ou techniquement, dont le sens est trop spécifique et non généralisable (Mel’čuk et al., 1995, p. 150). En d’autres termes, les FL non standards se caractérisent par leur nature non systématique. Elles s’appliquent aux liens observés dans les domaines spécialisés liés aux sciences et aux technologies. Dans notre corpus portant sur l’informatique, nous avons observé un bon nombre de liens qui peuvent être représentés par les FL non standards. C’est notamment le cas des liens reliant un modificateur à un terme dénotant une entité informatique (telle que 软

件 « logiciel ») ou une action impliquant des entités informatiques et des intervenants du domaine (telle que 攻击 « attaque informatique »).

Par exemple, le terme 软件 (« logiciel ») se combine avec des modificateurs tels que 开源 (« source libre ») et 间谍 (« espion ») dans des CLS dénotant un type spécifique de logiciel. Le terme 攻击 (« attaque informatique ») est souvent associé à des modificateurs pour désigner différents types d'attaques informatiques. Ces modificateurs peuvent être une unité lexicale (telle que 中间人 « intermédiaire ») ou un syntagme tel que (拒绝服务 « déni de service »). Les liens observés dans ces CLS ne peuvent être exprimés qu'à l'aide des FL non standards étant donné leur nature technique et non systématique. Quant à la description de ces liens à des fins didactiques, nous avons expliqué ces liens sémantiques à l'aide de définitions lors de l'élaboration du CHINOINFO, une ressource lexicale du lexique chinois destinée aux apprenants non natifs, comme le montre le tableau 10. La définition terminologique des CLS décrites permet aux apprenants de comprendre les connaissances spécialisées véhiculées par les CLS.

| CLS | Équivalent | Définition de la CLS dans laquelle on observe un lien sémantique exprimé par la FL non standard |
|---|-----------------------------|---|
| jiàn dié ruǎn jiàn 间谍软件 'espion'-'logiciel' | logiciel espion | Logiciel généralement indétectable qui s'infiltré dans un système informatique lors d'un téléchargement ou d'une installation dans le but d'employer la connexion Internet d'un utilisateur pour recueillir et transmettre ses renseignements personnels à son insu (Office québécois de la langue française, 2020c). |
| kāi yuán ruǎn jiàn 开源软件 'source ouverte'-'logiciel' | logiciel libre | Logiciel fourni avec son code source de manière qu'il puisse être étudié, copié, modifié et redistribué |
| zhōng jiān rén gōng jī 中间人攻击 'intermédiaire'-'attaque' | Attaque par interception | Cyberattaque au cours de laquelle une personne intercepte les communications entre deux parties, notamment à l'aide d'une clé cryptographique obtenue frauduleusement, ce qui lui permet de se faire passer pour l'une des parties, ou les deux, ou encore de récupérer de l'information sensible (Office québécois de la langue française, 2021b). |
| jù jué fú wù gōng jī 拒绝服务攻击 'refuser-service'- 'attaque' | Attaque par déni de service | Cyberattaque qui consiste à submerger de requêtes le système informatique d'une organisation afin de le rendre inopérant et d'en bloquer l'accès aux utilisateurs légitimes (Office québécois de la langue française, 2021a). |

Tableau 10. – Description de FL non standards observées dans des CLS

Encodage de FL

Encoder une FL consiste à paraphraser la relation lexicale qu'elle modélise au moyen d'un langage formel ou d'une langue contrôlée (Polguère, 2003). Dans le DEC, les FL sont encodées

au moyen d'une formule inspirée de la fonction mathématique, comme nous l'avons illustré plus haut. Il a été montré que ce formalisme d'encodage traditionnel peut être amélioré (Kahane et Polguère, 2001; Polguère, 2003). Kahane et Polguère (2001) soulignent l'importance d'adapter l'encodage de FL à différents contextes d'application, ils proposent un encodage algébrique adapté au traitement automatique et un encodage explicite ou vulgarisé qui s'applique à la confection de dictionnaire destiné au grand public (Polguère, 2000 a, 2000b).

Un encodage vulgarisé de FL consiste à expliciter les liens sémantiques qu'expriment les FL au moyen de paraphrases linguistiques. L'élaboration des paraphrases doit faire appel à une langue contrôlée, pour assurer la rigueur et la systématisme de la description des liens lexicaux. Ce type d'encodage a été mis en œuvre dans des ressources lexicales dont l'élaboration est guidée par les principes théoriques de la LEC, citons entre autres le dictionnaire vulgarisé *Lexique actif du français : l'apprentissage du vocabulaire fondé sur 20 000 dérivations sémantiques et collocations du français* (LAF) (Polguère, 2000a, 2000b; Mel'čuk et Polguère, 2007) ainsi que les dictionnaires spécialisés DiCoInfo³² (L'Homme et Alipour, 2009; Observatoire de linguistique Sens-Texte, 2022b) et DiCoEnviro³³ (L'Homme et Laneville, 2009; Observatoire de linguistique Sens-Texte, 2022a).

Dans le LAF, les signes et formules décrivant les liens lexicaux sont explicités au moyen d'un métalangage compréhensible par tous. Il s'agit d'une sorte de français standardisé (Mel'čuk et Polguère, 2007, p. 24). Examinons un exemple concret : dans l'article ADMIRATION (Mel'čuk et Polguère, 2007, p. 96) (voir l'annexe 1), la description des collocations contrôlées par la lexie vedette est enrichie de gloses explicitant la relation lexicale entre la base et le collocatif. La glose **Intense** explicite la relation lexicale d'intensification, modélisée quant à celle-ci par la FL **Magn**. Cette glose est suivie d'un ensemble de collocatifs modificateurs qui exprime le sens de 'intense' (Mel'čuk et al., 1995, p. 136), tels que *grande, vive, profonde, sans borne*, etc. La glose **A. devient plus intense** sert à expliciter la FL complexe **IncepPredPlus** qui décrit un collocatif verbal exprimant le sens de 'devenir plus **Magn**' (Mel'čuk et al., 1995, p. 136), tel que *grandir*.

³² Dictionnaire fondamental de l'informatique et d'Internet

³³ Dictionnaire fondamental de l'environnement

Dans notre recherche, nous nous sommes inspirée des travaux portant sur la vulgarisation des FL et nous avons eu recours à un système de gloses adaptées aux termes que nous avons retenus dans le CHINOINFO. Par exemple, le terme ^{chuán shū} 传输 (« transfert ») est associé au verbe ^{zhōng duàn} 中断 (« s'interrompre ») par un lien sémantique représenté la FL **Stop** qui signifie 's'arrêter de façon anormale', comme nous l'avons évoqué plus haut de cette sous-section. Nous avons expliqué ce lien sémantique à l'aide de la glose « un transfert cesse de fonctionner de façon anormale » dans le CHINOINFO. Nous avons aussi utilisé des définitions pour expliquer les liens sémantiques exprimés par les FL non standard. Nous reviendrons sur l'élaboration de gloses et de définitions dans la sous-section 4.2.1.1.

Vocable

Dans notre thèse, la définition de certaines notions (*vocable* et *polysème*) diffère de celles données dans la LEC. Cette démarche adaptative nous est nécessaire pour appliquer ces notions à la description du lexique spécialisé chinois. Nous apportons ici des précisions sur les écarts définitoires.

Comme nous l'avons vu plus haut, la lexie est l'unité de base de la description lexicale. La LEC regroupe les lexies ayant un signifiant identique et des signifiés apparentés en *vocables*. Chaque lexie d'un vocable, nommée *acception*, correspond à un sens lexical et fait ainsi l'objet d'un *article* dans le DEC. Un vocable est décrit par un *superarticle* du DEC (Mel'čuk et al., 1995, p. 155-156; Polguère, 2016, p. 291-292). Le superarticle regroupe donc des d'articles décrivant des sens apparentés.

Le regroupement des lexies en vocable repose sur le *pont sémantique*, soit une composante commune caractéristique dans la définition de deux lexies (Mel'čuk et al., 1995, p. 157). Autrement dit, les lexies regroupées dans un même vocable doivent satisfaire les deux conditions suivantes : 1) avoir un identifiant identique; 2) les signifiés de ces lexies sont apparentés.

Considérons les lexies ^{yí dòng} 移动₁ (« mobile ») et ^{yí dòng} 移动₂ (« amovible »). La première lexie se dit d'un appareil informatique, une application ou un réseau sans fil et autonome qu'on peut utiliser en se déplaçant (Office québécois de la langue française, 2016). La seconde lexie se dit d'un périphérique qui peut être retiré de l'unité centrale et remis à volonté (Cajolet-Laganière et al., 2022; Observatoire de linguistique Sens-Texte, 2015b). Ces deux lexies appartiennent à un même

vocabulaire puisque l'on dégage deux composantes sémantiques communes dans leur définition : 'détachable' et 'transportable'. Si des lexies ayant la même forme lexicale ne possèdent pas de pont sémantique, elles sont considérées comme homographes et consignées dans des articles consacrés à des vocables différents.

Selon la LEC, deux parties du discours distinctes sont associées à des vocables différents. Dans le *Dictionnaire en ligne de combinatoire du français* (DiCouèbe) (Mel'čuk et al., 2013), par exemple, le nom *criminel* et l'adjectif *criminel* relèvent de deux vocables distincts.

Dans notre travail, un vocable peut contenir les termes appartenant à des parties du discours distinctes. Nous nous éloignons de la terminologie de la LEC pour plusieurs raisons. Premièrement, dans les dictionnaires du chinois, les lexies de forme identique sont souvent regroupées dans un même article, quelle que soit la partie du discours à laquelle ces lexies appartiennent.

Deuxièmement, nous adaptons la définition de vocable pour rendre compte de la spécificité morphosyntaxique de la langue chinoise. Comme nous l'avons évoqué plus haut dans cette sous-section, les lexies chinoises ne subissent pas de modification morphologique. Ainsi, le lien de dérivation syntaxique entre lexies est souvent exprimé par la polysémie en chinois. En d'autres termes, les lexies qui relèvent de différentes parties du discours et entretiennent entre elles un lien de dérivation sémantique (Mel'čuk et Polguère, 2007, p. 18) peuvent partager la même forme lexicale, tel est le phénomène de *mot multicatégoriel* (兼类词) qui a été étudié dans des travaux de linguistique chinoise (Arcodia et Basciano, 2021, p. 276-280; Lu, 1994, 2003; Packard, 2000, p. 34-36; Zhu, 1982). Zhu (1982) et Lu (1994, 2003a) font remarquer le consensus des linguistes chinois sur le fait que le mot multicatégoriel est un cas particulier de polysémie lexicale. Nous y reviendrons dans la sous-section 4.1.2.4.

Troisièmement, les lexies de forme identique qui relèvent de différentes parties du discours ont souvent des sens très proches dans les domaines de spécialité. En informatique, par exemple, le terme 下载_v (« télécharger ») et son dérivé syntaxique nominal 下载_{nom} (« téléchargement ») ont une forme lexicale identique. Ce dernier désigne l'action de 'télécharger'. Dans notre travail, ces deux lexies sont regroupées dans un même vocable étant donné leur parenté sémantique.

D'un point de vue didactique, la distinction de ces termes sémantiquement proches exige une analyse ancrée dans leur contexte d'utilisation, ce qui représente une difficulté de taille pour les apprenants selon notre expérience d'enseignement auprès d'étudiants francophones. Nous proposons ainsi de regrouper ces lexies dans un même vocable pour faciliter l'accès à ces lexies et faire ressortir la distinction entre elles.

Polysème

Dans le même ordre d'idée, la notion de *polysème* couvre un sens plus général dans notre thèse. Un vocable du DEC correspond à la description d'un mot polysème qui est l'unité de base des dictionnaires courants. On peut en déduire qu'un *polysème*, selon la LEC, est un ensemble d'unités lexicales de forme identique qui appartiennent en général à la même partie du discours (Mel'čuk et al., 1995, p. 59, 156). Dans notre travail, un *polysème* fait référence à tous les sens associés à une même forme lexicale qui peuvent relever de parties de discours distinctes. Nous nous écartons de la définition proposée par la LEC afin de rendre compte de la polysémie engendrée par la dérivation sémantique en chinois.

Pour conclure, la LEC se caractérise par l'importance qu'elle accorde à la représentation sémantique du lexique. Elle rend compte non seulement de la description formelle des lexies, mais aussi de la structuration du lexique comme ensemble de lexies.

3.1.2 Applications lexicographiques de la LEC

La LEC a été mise en œuvre dans la compilation de ressources lexicales en langue générale. Outre le DEC, citons le *Lexique actif du français : l'apprentissage du vocabulaire fondé sur 20 000 dérivations sémantiques et collocations du français* (Mel'čuk et Polguère, 2007), le *Dictionnaire de Combinatoire* (DiCo), une base de données lexicale du français et sa version en ligne, le *Dictionnaire en ligne de combinatoire du français* (DiCouèbe) (Mel'čuk et al., 2013)³⁴, ainsi que le *Réseau Lexical du Français* (RL-fr), une base de données en cours de construction au laboratoire ATILF du CNRS (Lux-Pogodalla et Polguère, 2011; Polguère, 2014).

³⁴ Pour une documentation détaillée de l'élaboration de ces ressources, nous référons le lecteur aux travaux portant sur l'approche relationnelle de la modélisation lexicale (Jousse et al., 2008; Jousse et Polguère, 2005; Polguère, 2003).

La LEC, ainsi que son application dans le DiCo et le DiCouèbe, a aussi servi d'appui théorique à l'élaboration de dictionnaires spécialisés, dont le DiCoInfo (L'Homme et Alipour, 2009; Observatoire de linguistique Sens-Texte, 2022b) et le DiCoEnviro (L'Homme et Laneville, 2009; Observatoire de linguistique Sens-Texte, 2022a), ressources terminologiques multilingues mises au point par l'Observatoire de linguistique Sens-Texte (OLST), ainsi que le DiCE, un dictionnaire de collocations de l'espagnol académique, élaboré par une équipe de recherche de l'Université de La Corogne (Alonso Ramos, 2000, 2005b; Alonso Ramos et al., 2010). Calvi (2022) a présenté la description des collocations dans le cadre de la création de la base de données DIACOM-fr, une ressource terminologique portant sur le commerce international qui est mise au point par l'Université de Vérone.

D'ailleurs, la LEC a été mise à contribution pour modéliser des phénomènes linguistiques en chinois. Les premières tentatives d'application de la LEC au chinois remontent aux années 1990 Cheng et Polguère (1993) explorent l'application de l'approche DiCo de la lexicographie computationnelle au chinois. Nguyen (2006) aborde la morphologie dans le but de proposer un modèle de dictionnaire d'encodage à l'intention d'apprenants francophones. Zhao (2018) présente un modèle de dictionnaire des collocations portant sur le champ sémantique des émotions. Han et L'Homme (2022) présente une méthodologie et un modèle descriptif conçu pour encoder les collocations chinoises dans le DiCoEnviro (L'Homme et Laneville, 2009; Observatoire de linguistique Sens-Texte, 2022a), une base de données terminologique multilingue portant sur le domaine de l'environnement.

3.1.3 Application de la LEC à la didactique des langues

L'intérêt pédagogique de la LEC a été mis en évidence dans des travaux variés. En didactique du lexique général, la LEC a été appliquée à la conception de dispositifs d'enseignement en classe (Cauchon, 2003 ; Tremblay, 2003, 2004, 2009 ; Anctil, 2005, 2010) et à l'élaboration de ressources d'apprentissage (Mel'čuk et Polguère, 2007 ; Milićević et Hamel, 2007 ; Milićević, 2008), comme nous l'avons évoqué dans la sous-section 2.1.1.

En didactique du lexique spécialisé, il a été montré que les ressources lexicales élaborées selon les principes de la LEC peuvent être adaptées aux besoins des apprenants des langues de spécialité. Alipour (2014) propose une méthodologie de conversion de dictionnaires spécialisés

en dictionnaires d'apprentissage et l'applique à l'adaptation du DiCoInfo (L'Homme et Alipour, 2009; Observatoire de linguistique Sens-Texte, 2022b), un dictionnaire spécialisé portant sur le domaine de l'informatique.

3.1.4 Pertinence de la LEC comme cadre de référence

Le volet lexicologique constitue la composante fondamentale de notre cadre théorique. Nous avons choisi la LEC comme cadre de référence pour trois raisons principales.

En premier lieu, la LEC place la représentation sémantique au cœur de la description du lexique, ce qui constitue le principe de base de notre travail visant la conception d'une méthode descriptive du lexique spécialisé. D'une part, la mise en correspondance entre le sens d'une lexie et son comportement linguistique se fait sous l'angle de la synthèse, c'est-à-dire à partir du sens vers l'ensemble des expressions linguistiques correspondantes. Cette démarche se prête notamment à l'élaboration d'ouvrages de référence, comme le signale Polguère (1998). Appliquée à notre thèse, la description synthétique du lexique spécialisé permet de fournir au locuteur un vaste choix lexical pour exprimer un sens donné. D'autre part, la LEC propose un modèle descriptif relationnel. À la différence des modèles basés sur l'analyse componentielle du sens, limités aux mots isolés, la LEC vise la mise en relation des lexies, sur le plan tant paradigmatique et syntagmatique. En cela, la LEC s'avère pertinente pour modéliser la combinatoire lexicale.

En deuxième lieu, la mise en application de la LEC a débouché sur des résultats concrets. Cette approche a été mise en œuvre dans la compilation de ressources lexicales en langue générale et en langue de spécialité, comme nous l'avons expliqué dans la sous-section 3.1.2. D'ailleurs, la LEC a été appliquée à la description du lexique chinois (sous-section 3.1.2). Tout compte fait, la LEC est pertinente pour encadrer notre thèse qui débouche sur l'élaboration d'une ressource de référence du lexique spécialisé.

En dernier lieu, le potentiel didactique de la LEC a été reconnu par différents chercheurs (sous-section 3.1.3). La LEC présente un intérêt à l'enseignement du lexique en L2, notamment à l'élaboration de ressources lexicales, ce qui s'accorde avec notre objectif de créer une ressource d'apprentissage de la combinatoire spécialisée visant les locuteurs non natifs.

Pour conclure, la pertinence de la LEC comme cadre de référence de notre thèse se justifie à plusieurs titres, notamment son orientation sémantique, son application lexicologique et son intérêt pédagogique.

3.2 Approche lexico-sémantique de la terminologie

Cette section porte sur l'approche lexico-sémantique de la terminologie, qui constitue un cadre descriptif du lexique spécialisé. Nous décrivons, dans un premier temps, les principes fondamentaux de l'approche lexico-sémantique (3.2.1). Nous présentons ensuite l'application de celle-ci à la confection de dictionnaires spécialisés (3.2.2) et à l'enseignement du lexique spécialisé (3.2.3). À la fin de cette section, nous expliquons l'intérêt de l'approche lexico-sémantique comme cadre de référence pour le traitement du lexique spécialisé (3.2.4).

3.2.1 Principes fondamentaux de l'approche lexico-sémantique

Nous apportons, d'entrée de jeu, des précisions sur les notions fondamentales de l'approche lexico-sémantique, pour expliquer ensuite les cadres conceptuels dans lesquels celle-ci s'inscrit. Nous présentons enfin les propositions théoriques et méthodologiques de cette approche.

3.2.1.1 Notions de base de l'approche lexico-sémantique

La première notion que nous évoquons souvent dans ce travail est la *terminologie*. Elle se réfère à l'étude des *termes*, à savoir des expressions linguistiques désignant des éléments de connaissance dans un domaine de spécialité (L'Homme, 2020b, p. 5). La terminologie est ancrée dans la pratique, étant donné que son objectif principal consiste à proposer des modèles théoriques et méthodologiques pour les pratiques visant la description des termes (L'Homme, 2020a, p. 17).

L'approche lexico-sémantique définit le *terme* comme un type d'unité lexicale (L'Homme, 2020 c, p. 67-71). Plus précisément, un terme résulte de l'association d'une forme linguistique et d'un sens lexical, telle est la définition de l'unité lexicale en sémantique lexicale (Cruse, 1986; Polguère, 2016). L'Homme (2020a, p. 36) précise que les termes constituent un sous-ensemble du lexique d'une langue. Il s'agit « des unités lexicales dont le sens est envisagé par rapport à un domaine de spécialité, c'est-à-dire un domaine de la connaissance humaine, souvent associé à une activité socioprofessionnelle » (L'Homme, 2020a, p. 24).

3.2.1.2 Cadres conceptuels de l'approche lexico-sémantique

L'approche lexico-sémantique propose un modèle de description terminologique basé sur la sémantique lexicale (Cruse, 1986; Polguère, 2016) et la LEC (Mel'čuk et al., 1995)³⁵, d'où l'importance primordiale accordée au sens véhiculé par le terme en tant qu'unité lexicale.

En premier lieu, l'approche lexico-sémantique s'inscrit dans la mouvance des approches linguistiques à la terminologie, dont la socioterminologie (Gaudin, 1993, 2003) et la terminologie textuelle (Bourigault et Slodzian, 1999). Elle envisage les termes dans une optique linguistique. Ainsi, la description des termes rend compte des propriétés linguistiques des termes en tant qu'un sous-ensemble du lexique, ce qui a des conséquences méthodologiques sur différentes étapes de travail terminographique³⁶. Nous y reviendrons dans la sous-section 3.2.1.3.

En deuxième lieu, l'approche lexico-sémantique s'inspire de la LEC pour décrire les propriétés sémantiques des termes. Appliquée à la terminologie, la LEC est notamment utile pour décrire la structure actancielle des termes prédicatifs et modéliser les relations sémantiques, à savoir les relations paradigmatiques et syntagmatiques entre les termes (L'Homme, 2020 c, p. 39-43).

Plus récemment, l'approche lexico-sémantique fait aussi appel à la Sémantique des cadres (Fillmore, 1982; Fillmore et Baker, 2015) pour rendre compte de la structuration des connaissances véhiculées par le lexique d'un domaine de spécialité. Plus précisément, la Sémantique des cadres présuppose que le lexique est construit sur la base de « connaissances d'arrière-plan » et propose de représenter ces connaissances au moyen de « cadres » (L'Homme, 2016, 2020c). Un cadre se définit comme la modélisation schématique d'un scénario conceptuel dans lequel interviennent des participants. On trouve dans un cadre les termes qui « évoquent » ce scénario. L'approche lexico-sémantique s'inspire de la Sémantique des cadres et de son application dans la ressource FrameNet (Fillmore et al., 2003; Ruppenhofer et al., 2010) pour relier la description des propriétés lexico-sémantiques des termes (leur structure actancielle et les

³⁵ La LEC constitue le volet lexicologique du cadre théorique de notre thèse. Nous reportons le lecteur à la section 3.1 pour la présentation de la LEC.

³⁶ Nous reportons le lecteur à L'Homme (2020c) pour en savoir davantage sur l'apport de la sémantique lexicale à la terminologie.

relations sémantiques qui les lient dans le réseau lexical d'une langue) à une représentation conceptuelle (L'Homme, 2016).

Dans ce qui suit, nous développons les propositions de l'approche pour décrire les propriétés lexico-sémantiques des termes. Ce sont ces aspects que nous retenons pour notre travail. Les propositions inspirées de la Sémantique de cadre ne seront pas discutées dans le cadre de notre thèse.

3.2.1.3 Propositions théoriques et méthodologiques de l'approche lexico-sémantique

Dans le champ disciplinaire de la terminologie, on distingue différents modèles théoriques qui envisagent les termes sous des angles différents. La terminologie classique³⁷ adopte une optique conceptuelle. La TGT préconise une démarche onomasiologique. Autrement dit, elle postule l'existence du concept clairement délimité *a priori*. Le terme est le résultat d'un procédé de dénomination du concept. La description terminologique consiste donc à découvrir les formes linguistiques qui dénomment un concept et à mettre en évidence les relations entre les concepts dénotés par les termes. La TGT préconise la biunivocité entre le concept et le terme qui désigne celui-ci. Elle cherche à prévenir la synonymie et la polysémie en terminologie. Au sein d'un domaine de spécialité, les termes sont structurés de façon hiérarchique pour représenter les relations conceptuelles (L'Homme, 2005, p. 1114-1115). Ces principes ont fait l'objet de critiques depuis le milieu des années 1990 (Gaudin, 1993, 2003 ; Bourigault et Slodzian, 1999 ; Temmerman, 2000 ; Cabré, 2003).

Les remises en question de la TGT ont entraîné « une réhabilitation de la linguistique dans les analyses terminologiques » (L'Homme et Vandaele, 2007, p. 4). À la différence de la TGT qui place le concept au cœur de la description terminologique, l'approche lexico-sémantique prend le terme comme point de départ. Elle se caractérise par sa démarche sémasiologique pour appréhender le sens des termes à partir de leur comportement dans les discours spécialisés. La prise en compte des propriétés sémantiques des termes en tant qu'unités lexicales a des conséquences méthodologiques, comme le souligne L'Homme (2005, p. 1125-1129).

³⁷ Le terme *terminologie classique* se réfère à la *théorie générale de la terminologie* (TGT) (Wüster, 1974, 1974, 1976; Felber, 1980),

D'abord, l'identification des termes repose sur quatre critères sémantiques (L'Homme, 2020b, p. 71-74). Les unités lexicales relevant de différentes parties du discours sont reconnues comme termes si elles véhiculent un sens spécialisé. Par exemple, les termes que nous avons relevés à partir de notre corpus (voir aussi la sous-section 4.1.2.2) peuvent appartenir aux parties du discours du nom (网络 «réseau»), du verbe (更新 «mettre à jour»), de l'adjectif (移动 «mobile») et de l'adverbe (实时 «en temps réel»). Nous abordons dans la sous-section 4.1.2.2 l'application de certains de ces critères au repérage des termes à partir de notre corpus.

Ensuite, la synonymie et la polysémie sont reconnues, vu la correspondance multivoque entre la forme et le sens du terme. Précisons que l'approche lexico-sémantique applique des modèles linguistiques au traitement de la polysémie. D'une part, chaque sens associé à une forme lexicale doit être décrit de façon distincte (L'Homme, 2020b, p. 103-110). Ce principe guide notre démarche d'analyse des unités terminologiques³⁸ extraites de notre corpus. Par exemple, nous avons identifié deux sens associés à la forme lexicale 移动 : 移动₁ «mobile» et 移动₂ «amovible». Chaque sens fait l'objet d'un article dans la ressource lexicale que nous avons créée. D'autre part, la distinction sémantique fait appel aux tests lexico-sémantiques (L'Homme, 2020a, p. 76-80), inspirés par les tests mis au point en sémantique lexicale (Cruse, 1986, p. 55-57; Mel'čuk et al., 1995, p. 64-68). Dans notre travail, nous avons mis en œuvre trois tests de distinction sémantique pour analyser les unités terminologiques polysémiques relevées de notre corpus (voir la sous-section 4.1.2.3). En outre, L'Homme (2020d) présente différents cas de polysémie spécialisée et aborde des méthodes de description lexicale permettant d'en rendre compte dans des dictionnaires spécialisés, par exemple, la représentation des relations lexicales de chaque terme associé à une même forme lexicale, l'explicitation de la structure actancielle et l'identification des cadres sémantiques évoqués par les termes. Dans notre thèse, nous nous sommes inspirée des deux premières solutions, lors de l'élaboration de notre ressource

³⁸ Les unités terminologiques (UT) sont les unités linguistiques ayant un ou plusieurs sens envisagé(s) par rapport à un domaine de spécialité. Elles peuvent être monosémiques ou polysémiques (Alipour, 2018, p. 60, 99-101).

spécialisée, dans l'objectif de rendre explicite la distinction sémantique entre les unités polysémiques.

Enfin, la structuration terminologique met en évidence les relations lexicales que les termes entretiennent (à savoir les liens syntagmatiques et paradigmatiques), plutôt que les relations conceptuelles envisagées par la TGT. Nous illustrons, dans la sous-section 3.2.2, l'organisation des données dans le DiCoInfo, un dictionnaire spécialisé dont l'élaboration s'appuie sur l'analyse terminologique à optique lexico-sémantique.

En résumé, l'approche lexico-sémantique propose un cadre descriptif terminologique à orientation linguistique. Elle se caractérise notamment par son ancrage dans les discours spécialisés et la primauté accordée aux propriétés sémantiques des termes.

3.2.2 Application de l'approche lexico-sémantique à la confection de dictionnaires spécialisés

La terminologie vise à fournir des repères fondamentaux aux différentes applications relevant de la terminographie, comme le souligne L'Homme (2020a, p. 17). Celle-ci « regroupe un ensemble d'activités dont l'objectif principal est de décrire des termes dans les dictionnaires spécialisés ou les banques de terminologie » (L'Homme, 2020a, p. 23).

L'approche lexico-sémantique est appliquée à la confection de dictionnaires spécialisés, dont le DiCoInfo (Observatoire de linguistique Sens-Texte, 2022b) et le DiCoEnviro (Observatoire de linguistique Sens-Texte, 2022a). Dans ce qui suit, nous présentons l'élaboration du DiCoInfo ainsi que le contenu du dictionnaire, afin de montrer de quelle manière les principes méthodologiques de l'approche lexico-sémantique appuient le travail de description des termes à différentes étapes de travail.

Le DiCoInfo (Observatoire de linguistique Sens-Texte, 2022b) est un dictionnaire spécialisé multilingue portant sur les termes fondamentaux dans le domaine de l'informatique et d'Internet. Ce dictionnaire se distingue de la majorité des ressources terminologiques (dictionnaires spécialisés et banques de terminologie) en ce sens qu'il fournit une description des propriétés linguistiques, notamment des propriétés lexico-sémantiques des termes répertoriés.

Le travail terminographique menant à l'élaboration du DiCoInfo se décline en deux phases (L'Homme, 2008). La première phase consiste à collecter des données et à établir la nomenclature du dictionnaire. Cette phase comporte quatre étapes : 1) l'élaboration d'un corpus; 2) le repérage de termes; 3) la sélection de contextes et 4) l'établissement de distinctions sémantiques. La deuxième phase regroupe trois tâches reliées à la rédaction des articles : 5) la définition de la structure actancielle; 6) l'établissement de la liste de liens lexicaux et 7) l'encodage des liens lexicaux.

La mise en forme d'un corpus spécialisé constitue le point de départ du travail terminographique. La collecte et l'analyse des données terminologiques reposent sur les observations faites par les terminologues dans le corpus. Un corpus est élaboré pour chacune des langues abordées dans le DiCoInfo (L'Homme, 2008).

La deuxième étape du travail terminographique, à savoir le repérage de terme, fait appel à des critères sémantiques, comme nous l'avons expliqué dans la sous-section 3.2.1.3 (L'Homme, 2020b, p. 71-74). Le corpus est d'abord soumis à un dépouillement automatique à l'aide de TermoStat (Drouin, 2003), un extracteur terminologique. Les résultats générés par l'extraction automatique, nommés candidats-termes (CT), sont ensuite analysés à l'aide des quatre critères visant à valider le statut terminologique des CT : 1) la dénotation d'une entité liée au domaine abordé; 2) la présence d'actants sémantiques de nature spécialisée; 3) la parenté morphologique accompagnée d'une parenté sémantique et 4) tout autre lien paradigmatique (L'Homme, 2008).

La troisième étape consiste à extraire les contextes d'utilisation des termes repérés à partir du corpus. Pour chaque terme, on extrait entre 15 et 20 contextes. Ceux-ci illustrent l'emploi des termes dans les textes spécialisés.

Enfin, l'établissement de la nomenclature nécessite des distinctions sémantiques de certaines unités terminologiques ayant plus d'un sens spécialisé. Au cours de cette étape de travail, on fait appel aux tests lexico-sémantiques (L'Homme, 2020a, p. 76-80) pour délimiter les acceptions de ces unités terminologiques polysémiques : 1) la substitution par un synonyme; 2) l'opposition différentielle; 3) la dérivation morphologique différentielle, 4) la présence de liens paradigmatiques différentiels, 5) la cooccurrence compatible et 6) la cooccurrence différentielle (Mel'čuk et al., 1995, p. 64-68).

Quant à la rédaction des articles, le DiCoInfo s'inspire de la LEC et de son application à l'élaboration du *Dictionnaire de Combinatoire* (DiCo) et du *Dictionnaire en ligne de combinatoire du français* (DiCouèbe) (Mel'čuk et al., 2013)³⁹ pour décrire les propriétés lexico-sémantiques des termes. Les acceptions associées à la même forme lexicale sont regroupées en vocables. Chaque acception fait l'objet d'un article distinct. Dans un article du DiCoInfo, on trouve des renseignements sur le sens et le fonctionnement linguistique des termes, tels que la structure actancielle, la combinatoire et les relations sémantiques (L'Homme et Alipour, 2009).

Premièrement, la structure actancielle fournit des renseignements sur le sens du terme vedette : on décrit les actants sémantiques de celui-ci, le rôle sémantique de chaque actant par rapport au terme vedette ainsi que les réalisations linguistiques des actants.

Deuxièmement, pour chaque terme décrit, le DiCoInfo propose une liste de liens lexicaux dans laquelle sont présentés les liens paradigmatiques et syntagmatiques du terme. Examinons la liste des liens lexicaux dans l'article *programme* du DiCoInfo (Observatoire de linguistique Sens-Texte, 2015d) (annexe 2).

Dans la version française du DiCoInfo, pour laquelle un travail a été fait afin de la rendre plus conviviale et adaptée aux besoins des apprenants du lexique spécialisé (Alipour, 2014), on trouve également la description des liens sémantiques.

Les liens paradigmatiques qui relient le terme vedette à d'autres termes du domaine sont présentés dans deux rubans de l'article. Le deuxième ruban contient les synonymes, les contraires et les voisins du terme décrit. Dans l'article *programme* (annexe 2) (Observatoire de linguistique Sens-Texte, 2015d), on présente dans l'onglet « Synonyme(s) » le terme *programme informatique* qui est le synonyme du terme vedette. Sur le quatrième ruban figurent les liens paradigmatiques autres que les synonymes, les voisins et les contraires. On y trouve trois onglets intitulés « Famille de mots », « Sortes de » et « Autres liens ». Dans l'onglet « Famille de mots »,

³⁹ Pour une documentation détaillée de l'élaboration de ces ressources, nous référons le lecteur aux travaux portant sur l'approche relationnelle de la modélisation lexicale (Jousse et al., 2008; Jousse et Polguère, 2005; Polguère, 2003).

par exemple, on répertorie les termes reliés au terme décrit (*programme*) par un lien morphologique, tels que *programmer*, *programmation* et *sous-programme*.

Les liens syntagmatiques qui relient le terme décrit à ses collocatifs dans des collocations sont décrits dans le dernier ruban de l'article. Ces collocations sont répertoriées dans différentes classes selon les liens sémantiques observés. Dans l'article *programme* (annexe 2), on trouve dans l'onglet « Utiliser/Ne pas utiliser » deux collocations dans lesquelles le terme vedette *programme* se combine avec deux verbes exprimant le sens de 'préparer un programme afin de le faire fonctionner' (*appeler un programme* et *charger un programme*), de même que deux collocations qui contiennent le dérivé syntaxique nominal de ces verbes (*appel d'un programme* et *chargement d'un programme*)⁴⁰.

Comme on peut le constater, la méthodologie menant à l'élaboration du DiCoInfo comprend un ensemble de tâches terminographiques guidées par les principes de l'approche lexico-sémantique que nous avons présentés dans la sous-section 3.2.1.3.

3.2.3 Application de l'approche lexico-sémantique à l'enseignement du lexique spécialisé

L'approche lexico-sémantique est aussi appliquée à la didactique du lexique spécialisé. Elle a été mise à contribution pour élaborer des méthodes d'enseignement en classe. Alipour (2018) aborde l'enseignement du lexique spécialisé auprès d'étudiants en traduction. L'auteure s'attache à concevoir des activités pédagogiques visant à aider les futurs traducteurs à résoudre les difficultés liées à la polysémie et à la (quasi-) synonymie.

Dans sa recherche, Alipour (2018) se situe dans une perspective lexico-sémantique pour envisager la notion de terme, ainsi que les relations sémantiques entre les termes, à savoir la polysémie et la (quasi-) synonymie. L'auteure adopte une approche basée sur le corpus. Plus précisément, elle s'appuie sur la démarche sémasiologique, telle que proposée par l'approche lexico-sémantique, pour amener les apprenants à appréhender le sens des termes dans les discours

⁴⁰ Nous reportons le lecteur à L'Homme (2008) (sous-section 3.2.3) et à L'Homme (2009, p. 20-26) pour la présentation des liens lexicaux dans le DiCoInfo.

spécialisés. Lors de la conception des modèles de tâches et d'activités, l'auteure fait appel à la méthodologie lexico-sémantique pour sélectionner et analyser les termes à partir du corpus. D'une part, on valide le statut terminologique des unités lexicales extraites du corpus à l'aide des critères d'identification des termes (L'Homme, 2020b, p. 71-74), d'autre part, la distinction sémantique des unités terminologiques polysémiques repose sur les tests lexico-sémantiques (L'Homme, 2020a, p. 76-80), comme nous l'avons expliqué dans la sous-section 3.2.1.3. Quant à la recherche de solutions pédagogiques, l'auteure propose des stratégies d'analyse de contexte pour aider les apprenants à faire face aux difficultés soulevées par les (quasi-)synonymes. La pertinence des activités pédagogiques et l'efficacité des stratégies d'analyse des contextes sont confirmées au moyen d'une expérimentation comparative.

Une autre application de l'approche lexico-sémantique à l'enseignement du lexique spécialisé est la création de ressources didactiques. Alipour (2014) propose une méthodologie pour convertir le DiCoInfo (Observatoire de linguistique Sens-Texte, 2022b) en un dictionnaire d'apprentissage destiné aux apprenants de la langue de l'informatique. L'élaboration du DiCoInfo, comme nous l'avons expliqué dans la sous-section 3.2.2, repose sur les principes méthodologiques de l'approche lexico-sémantique. Le DiCoInfo fournit une description détaillée des propriétés lexico-sémantiques des termes, c'est-à-dire leur sens et leur fonctionnement linguistique dans les textes spécialisés. Pour adapter ce dictionnaire aux besoins des utilisateurs, à savoir les traducteurs, réviseurs et rédacteurs techniques, l'auteur se penche sur la présentation des données terminologiques. Plus précisément, elle apporte des modifications à la mise en page des articles, à l'organisation des données, à la navigation entre les articles et à l'utilisation des outils audio et visuels. Ces stratégies visent à fournir une présentation simplifiée des renseignements descriptifs des termes.

3.2.4 Intérêt de l'approche lexico-sémantique comme cadre de référence

L'approche lexico-sémantique de la terminologie s'attarde sur la description du lexique spécialisé. Sa pertinence comme cadre de référence se manifeste sur le plan tant théorique que pratique.

D'un point de vue théorique, l'approche lexico-sémantique définit les principes méthodologiques du traitement du lexique spécialisé. Nous nous intéressons en particulier à la démarche

sémasiologique et au traitement de la polysémie. Dans notre travail, la sélection de termes et l'analyse terminologique s'appuient sur les observations dans un corpus spécialisé. Nous avons construit un corpus chinois pour en extraire une liste des termes, un échantillon de CLS contenant ces termes ainsi que les données descriptives des CLS repérées. Les renseignements recueillis nous ont servi d'assise pour envisager une méthode de description visant à aider les apprenants à résoudre les difficultés soulevées par la combinatoire lexicale spécialisée (voir aussi la section 4.1).

Quant au traitement de la polysémie, nous avons abordé effectué des analyses des contextes d'utilisation des termes relevés, afin d'identifier les différentes parties du discours associées à une unité terminologique polysémique. Nous avons ensuite appliqué les tests lexico-sémantiques (L'Homme, 2020a, p. 76-80) pour délimiter les acceptions des unités terminologiques polysémiques relevant de la même partie du discours (voir aussi la sous-section 4.1.2.3).

D'un point de vue appliqué, l'approche lexico-sémantique a été mise en œuvre dans la création de dictionnaires spécialisés, dont le DiCoInfo (Observatoire de linguistique Sens-Texte, 2022b) et le DiCoEnviro (Observatoire de linguistique Sens-Texte, 2022a). Elle a fourni des repères méthodologiques pour adapter le DiCoInfo à un dictionnaire d'apprentissage (Alipour, 2014) et adapter l'encodage des combinaisons lexicales dans le DiCoEnviro au chinois.

Ces ressources terminologiques nous servent de modèles pour envisager une méthode de description systématique à des fins de confection de ressources lexicales spécialisées, tel est l'objectif primaire de notre recherche.

Pour terminer, l'approche lexico-sémantique de la terminologie guide notre démarche du traitement du lexique spécialisé chinois. Les deux premiers volets présentés plus haut, axés respectivement sur la lexicologie (section 3.1) et la terminologie (section 3.2), encadrent la description du lexique. Dans la section 3.3, notre objet de réflexion passe à la transmission des connaissances lexicales, ce qui nous amène au volet didactique de notre cadre théorique.

3.3 Approche cognitive en didactique des langues secondes

Les théories en didactique de L2 rendent compte des fondements linguistiques, psychologiques ou socioculturels de la pratique enseignante. Notre travail s'appuie sur l'approche cognitive

(*Cognitive-code learning theory*) (Hinkel, 2012, p. 625-626) qui soutient un enseignement structuré de la combinatoire spécialisée.

Nous dressons d'abord un portrait de l'approche cognitive (3.3.1), pour donner ensuite un aperçu des méthodes d'enseignement qui lui sont associées (3.3.2). Nous expliquons enfin, à l'aide d'exemples, quelques pistes dégagées de cette approche pour enseigner les CLS de façon organisée.

3.3.1 Principes fondamentaux de l'approche cognitive

Nous définissons en premier lieu les notions de base en didactique des langues, notamment la notion d'*approche*. Partant de cette définition, nous décrivons ensuite les principes théoriques de l'approche cognitive.

3.3.1.1 Notions associées aux méthodes d'enseignement des langues

En didactique des langues, on recense des travaux qui proposent des modèles pour décrire, analyser et comparer les méthodes d'enseignement. Le linguiste appliqué états-unien Edward Anthony (1963, p. 63-67) élabore un schéma à trois niveaux (*approach-method-technique*) pour modéliser les méthodes d'enseignement des langues. Ce modèle est révisé par Richards et Rodgers (2001, p. 18-34) qui soutiennent que toute *méthode* peut être envisagée en termes de l'approche (*approach*), de la conception (*design*) et de la procédure (*procedure*).

L'*approche* se définit comme un ensemble de théories sur la nature de la langue et celle de l'apprentissage d'une langue. Une approche sert ainsi d'appuis théoriques à l'enseignement des langues (Richards et Rodgers, 2001, p. 20). La *conception* rend compte d'un nombre d'éléments menant à l'élaboration d'une méthode, à savoir les objectifs d'une méthode, la sélection et l'organisation du contenu d'enseignement, les types de tâches d'apprentissage, les types d'activités pédagogiques, les rôles des apprenants, ceux des enseignants ainsi que les fonctions du matériel didactique (Richards et Rodgers, 2001, p. 24). La *procédure* se concentre sur la manière dont une méthode est mise en œuvre en classe, plus précisément, on analyse de quelle manière la conception d'une méthode d'enseignement se concrétise dans la pratique (Richards et Rodgers, 2001, p. 31).

Dans le cadre de notre travail, nous nous alignons sur la proposition de Richards et Rodgers (2001, p. 18-34). On entend par *approche* un ensemble de principes théoriques qui sous-tendent le choix d'une *méthode* d'enseignement⁴¹. Une approche se caractérise par deux aspects essentiels : un aspect linguistique abordant la nature de la langue, ainsi qu'un aspect psycholinguistique axé sur la nature de l'apprentissage d'une langue.

Notons que plusieurs didacticiens et linguistes chinois (Liu, 2000; Xu et Wu, 2005) décrivent les méthodes d'enseignement des langues en termes de trois disciplines de référence, à savoir la linguistique, la psychologie générale et la psychologie de l'éducation. Les recherches en didactique des langues qui sont menées dans ces perspectives se penchent respectivement sur trois aspects : 1) l'objet d'enseignement, soit la langue; 2) le sujet qui apprend une langue et 3) le rapport entre l'objet d'enseignement et l'apprenant. À notre avis, le premier aspect correspond au volet linguistique d'une *approche* didactique, tel qu'elle est définie par Richards et Rodgers (2001, p. 20), les deux autres aspects correspondent au volet psycholinguistique de celle-ci.

Dans ce qui suit, nous allons examiner l'approche cognitive en didactique des langues dans les perspectives linguistique et psycholinguistique.

3.3.1.2 Principes théoriques de l'approche cognitive

L'approche cognitive a été proposée aux États-Unis dans les années 1960 (Carroll, 1964, 1966 ; Chastain, 1970, 1971), sous l'impulsion des avancées de la linguistique appliquée et des théories de l'apprentissage. De façon générale, cette approche met en évidence le rôle actif des apprenants dans l'acquisition d'une L2, plus spécifiquement, la manière dont les apprenants traitent les nouvelles connaissances et emmagasinent celles-ci dans leur mémoire.

D'un point de vue linguistique, l'approche cognitive s'appuie sur la théorie de la grammaire générative et transformationnelle (Chomsky, 1957) selon laquelle la langue est un système régi par un ensemble de règles qui président notamment à la syntaxe. Selon l'approche cognitive,

⁴¹ Notons toutefois que notre thèse ne vise pas l'élaboration d'une *méthode d'enseignement*. Notre objectif de recherche consiste à mettre au point une méthode de description lexicale pour favoriser l'acquisition des CLS. Les résultats de notre recherche peuvent être mis à contribution pour concevoir des séquences d'apprentissage du lexique, ou pour élaborer des ressources didactiques. Ces moyens et ressources déployés font partie intégrante d'une méthode d'enseignement à proprement parler.

l'apprentissage d'une L2 est défini comme un processus d'acquisition consciente des règles grammaticales, lexicales et phonétiques (Carroll, 1964; Chastain, 1971). Cette approche remet en cause le point de vue béhavioriste (Skinner, 1957), selon lequel les connaissances d'une langue peuvent être décomposées en connexions observables entre un stimulus et une réponse (Tremblay, 2009, p. 127). Autrement dit, l'acquisition d'une L2 est une formation d'habitudes à travers l'association "stimuli-réponses", telle est la critique formulée par Chomsky (1959) à la théorie béhavioriste de Skinner (1957).

L'approche cognitive s'appuie sur la psychologie cognitive (Bruner, 1960 ; Ausubel, 1968) pour rendre compte de la façon dont les apprenants acquièrent une langue. Selon les premiers partisans de cette approche, dont Carroll (1966) et Chastain (1970), l'apprentissage est avant tout un processus mental actif de traitement de l'information, dans lequel les apprenants perçoivent, organisent et stockent les nouvelles connaissances dans leur mémoire à long terme pour les extraire plus tard. Il s'agit d'un processus d'interaction entre l'objet d'enseignement et l'apprenant qui est contrôlé par ce dernier. En second lieu, l'apprentissage est constructif et repose sur l'association des nouvelles informations reçues aux connaissances antérieures des apprenants, organisées quant à celles-ci sous forme de *structure cognitive* (Ausubel, 1963 ; Bruner, 1960). Il en résulte que l'organisation des connaissances joue un rôle fondamental dans l'apprentissage. Gagné (1965), qui propose un modèle de traitement de l'information pour décrire les différentes phases de l'apprentissage, souligne aussi l'importance de l'*encodage*, à savoir la structuration de nouvelles informations en vue de leur archivage dans la mémoire des apprenants. Un encodage de qualité, comme le signale Tardif (1992), facilite également le repérage des connaissances acquises ultérieurement.

Partant de cette conception de la langue et de l'apprentissage, l'approche cognitive analyse aussi les facteurs censés favoriser l'acquisition d'une L2. Premièrement, Carroll (1966, p. 102) préconise l'enseignement explicite des règles grammaticales et lexicales. Il souligne l'importance de la pratique efficace en enseignement-apprentissage d'une L2. Selon l'auteur, les apprenants doivent prendre conscience du fonctionnement de la L2, afin d'appliquer les règles enseignées dans une nouvelle situation d'utilisation de la langue. Deuxièmement, les connaissances doivent être présentées de façon structurée (Ausubel, 1968), que ce soit en classe ou dans le matériel didactique, afin de faciliter l'encodage des connaissances par les apprenants. Troisièmement, le

bagage cognitif des apprenants doit être pris en compte, afin de faciliter le *transfert d'apprentissage* (Tardif, 1999), à savoir l'association du contenu enseigné aux connaissances antérieures des apprenants. En didactique des L2, une attention particulière est accordée au *transfert négatif* (Doca, 1981 ; Tardif, 1992, 1999), défini comme l'influence négative des connaissances acquises sur l'apprentissage d'une nouvelle langue. Tardif (1999) précise qu'un transfert négatif peut être causé par plusieurs facteurs, notamment l'interférence de la L1 des apprenants et la surgénéralisation des connaissances acquises en L2.

Pour conclure, l'approche cognitive met en lumière les aspects cognitifs en jeu dans l'apprentissage d'une L2. Elle souligne non seulement le rôle actif des apprenants dans le traitement de contenu enseigné, mais aussi le rôle des enseignants et des concepteurs de ressources didactiques dans la présentation de nouveaux contenus (Chastain, 1990, p. 25). Cette prise en compte du processus cognitif stimule la réflexion des enseignants et soutient constamment le développement de stratégies pédagogiques en enseignement des L2. Nous abordons cette question en détail dans la sous-section 3.3.2.

3.3.2 Implication méthodologique dans l'enseignement du lexique

L'approche cognitive est une théorie proposée en réaction contre la méthode audio-orale (MAO) qui a dominé l'enseignement des L2 aux États-Unis pendant les années 1940-1960. Malgré la critique formulée aux fondements behavioristes de la MAO, l'approche cognitive n'a pas débouché sur des retombées concrètes (Byram, 2002, p. 174 ; Richards et Rodgers, 2001, p. 66). Selon les auteurs, aucune nouvelle méthode didactique n'a été mise de l'avant comme solution de remplacement à la MAO. Par contre, l'approche cognitive a pu entraîner un changement d'orientation subtil en didactique des L2, comme le signale Chastain (1990, p. 23-26), étant donné que ses principes fondamentaux se sont intégrés aux approches pédagogiques à l'œuvre, depuis son émergence jusqu'aux années 1990. Tardif et al. (1994) font remarquer que la conception cognitiviste prévalait en psychopédagogie dans les années 1990. Au Québec, Tardif (1992) présente pour la première fois une vision générale et élaborée d'une conception cognitiviste de l'enseignement (M. Tardif et al., 1994). L'approche cognitive contribue aussi au développement d'une méthode d'analyse des erreurs (AE) (Corder, 1967, 1981). Cette méthode, basée sur les productions linguistiques des apprenants, propose une analyse systématique des erreurs en fonction de différents paramètres (nature des erreurs, leur fréquence et leur probabilité

d'apparition). Cette méthode permet ainsi aux enseignants d'intervenir de façon proactive pour mettre les apprenants en garde contre des erreurs authentiques.

En enseignement du chinois L2, l'approche cognitive sert d'appui théorique à la pédagogie basée sur le morphème (voir la sous-section 2.3.2), mise de l'avant dans les années 1990 où la morphologie a suscité un intérêt grandissant chez les linguistes et les didacticiens (Sheng, 1990; Lü, 1999). Cette méthode se caractérise par deux traits fondamentaux. D'abord, elle vise à développer la conscience morphologique⁴² des apprenants. Ensuite, elle préconise la présentation structurée des connaissances lexicales au moyen de réseaux sémantiques. Dans la pratique, comme nous l'avons mentionné dans la sous-section 2.3.2, cette pédagogie s'applique notamment à la conception de séquences d'apprentissage du lexique en classe.

De plus, l'approche cognitive continue d'alimenter la réflexion sur la conception de dispositifs pédagogiques. On recense des travaux plus récents ayant recours à cette approche pour améliorer l'efficacité de l'acquisition d'une L2 dans différents contextes. Redmond et Emirkanian (2018) font appel à l'approche cognitive pour envisager un enseignement structuré de la polysémie verbale en français L2. Elles proposent plus précisément de décrire les différentes acceptions du verbe *prendre* au moyen d'une analyse sémantique systématique basée sur l'approche cognitive. Une étude empirique, menée par la suite auprès de 191 apprenants du français L2, a démontré l'efficacité de l'analyse sémantique proposée. Cette étude a aussi mis en évidence l'impact de transfert sur l'acquisition du lexique en L2. Plusieurs partisans de l'approche cognitive, dont Doca (1981) et Tardif (1992, 1999), ont pour leur part souligné l'influence du transfert des connaissances au cours de l'apprentissage d'une L2. Ying (2013) applique l'approche cognitive pour analyser les erreurs écrites des apprenants chinois du FLE. Cette analyse vise à déceler le processus qui conduit à l'erreur, afin de proposer des pistes pour bonifier l'enseignement du FLE. Un corpus a été construit pour en extraire 1020 erreurs d'ordre lexical, grammatical et sociopragmatique. Cette recherche a aussi confirmé l'utilité de l'approche cognitive pour aider

⁴² Le terme *conscience morphologique* (*morphological awareness*) (Carlisle, 1995; Kirby, 2004) est issu de la psycholinguistique. En enseignement du chinois langue seconde, ce terme désigne la capacité des apprenants à reconnaître les morphèmes (à savoir les éléments minimaux porteurs de sens), à identifier la structure interne des unités lexicales et à comprendre les règles de formation des unités lexicales.

les apprenants à contourner les erreurs principales identifiées. Huong (2010) se penche sur la didactique du français L2 au Vietnam. L'auteur aborde plus spécifiquement la formation des enseignants des sciences dans le cadre d'un programme d'enseignement intensif du français. L'approche cognitive est mise en œuvre pour élaborer des stratégies pédagogiques visant à favoriser l'acquisition des connaissances lexicales et grammaticales. Parmi les stratégies déployées, citons la construction de réseaux sémantiques, le recours aux exercices contextualisés et l'enseignement systématique des règles grammaticales.

En résumé, l'approche cognitive a apporté du nouveau aux fondements théoriques de la didactique des L2, comme le signale Chastain (1990, p. 26). Cette approche est mise en application dans différents contextes d'enseignement. En didactique du chinois L2, qui nous intéresse plus particulièrement, cette approche soutient la pédagogie axée sur le morphème.

3.3.3 Pertinence de l'approche cognitive comme cadre de référence

L'approche cognitive constitue le volet didactique de notre cadre théorique. Elle fournit un cadre de référence pertinent pour envisager un enseignement structuré de la combinatoire spécialisée du chinois.

Premièrement, l'approche cognitive préconise un enseignement explicite et systématique des règles grammaticales. Dans notre travail, nous appliquons ce principe à la présentation des CLS. Plus précisément, nous proposons de mettre en évidence la structure syntaxique des CLS afin d'aider les apprenants à résoudre l'ambiguïté syntaxique des CLS, c'est-à-dire distinguer les différents sens associés aux CLS ayant la même forme mais des structures syntaxiques différentes (voir aussi la sous-section 4.2.1.2).

Deuxièmement, l'approche cognitive envisage l'apprentissage comme un processus constructif de traitement de l'information (voir la sous-section 3.3.1.2). Elle insiste sur la présentation organisée des connaissances au cours de l'apprentissage. Le rôle de l'enseignant et du concepteur de matériel didactique, comme le soutient Chastain (1990, p. 25), consiste en grande partie à structurer le contenu enseigné afin qu'il soit compréhensible et compatible avec le niveau d'apprentissage des apprenants. Ce principe nous amène à explorer différents moyens pour organiser les CLS en différentes classes hiérarchiques (voir les sous-sections 4.2.1.1 et 4.2.1.2), afin d'amener les apprenants à encoder, à savoir emmagasiner dans leur mémoire les CLS selon

la structure syntaxique de celles-ci et le lien sémantique observé entre les deux éléments composants d'une CLS.

Par ailleurs, l'approche cognitive présente un intérêt particulier pour l'enseignement des CLS du point de vue des types de connaissance. Tardif (1992) fait remarquer que la psychologie cognitive distingue trois types de connaissance : les connaissances déclaratives, procédurales et conditionnelles. Les connaissances déclaratives concernent les faits, les règles, les lois ou les principes théoriques (Tremblay, 2009, p. 131), par exemple, la connaissance des différents types de structures syntaxiques en chinois. Les connaissances procédurales sont les savoir-faire, il s'agit d'un ensemble d'étapes à suivre pour réaliser une tâche ou résoudre un problème, par exemple, chercher le sens d'un terme dans un dictionnaire spécialisé. Les connaissances conditionnelles concernent le contexte dans lequel on applique les connaissances déclaratives et procédurales, par exemple, déduire le sens d'un mot à partir du contexte dans lequel il apparaît (Tremblay, 2009, p. 132). Ces types de connaissances sont étroitement liés à la notion de *compétence*, comme le souligne Tremblay (2009, p. 132). Pour favoriser l'acquisition des compétences lexicales des apprenants, il faut les aider à construire ces trois types de connaissances. Appliquée à l'enseignement des CLS, l'approche cognitive guide notre démarche afin de :

- 1) sensibiliser les apprenants aux structures syntaxiques des CLS et au sens que celles-ci véhiculent (connaissances déclaratives);
- 2) inciter les apprenants à utiliser un dictionnaire spécialisé pour trouver, de manière autonome, des éléments de réponse à des difficultés soulevées des CLS et à évaluer la qualité de l'information fournie par une ressource spécialisée (connaissances procédurales);
- 2) inciter les apprenants à évaluer la qualité de l'information fournie par les ressources lexicales spécialisées afin de trouver, de manière autonome, des éléments de réponse à des difficultés soulevées des CLS (connaissances procédurales);
- 3) développer les habiletés des apprenants à résoudre trois types de problèmes lexicaux authentiques qu'ils rencontrent dans les textes spécialisés (connaissances conditionnelles), à savoir identifier le sens d'une forme lexicale dans une CLS donnée, distinguer les sens associés à

des CLS de forme identique dans différents contextes et choisir les cooccurrents appropriés d'un terme selon le contexte donné (voir aussi la sous-section 4.1.3.4).

Enfin, l'approche cognitive accorde une attention particulière au transfert de l'apprentissage, à savoir l'influence des connaissances antérieures des apprenants sur l'acquisition d'une L2. Nous nous intéressons plus spécifiquement au transfert négatif (Doca, 1981 ; Tardif, 1992, 1999) causé par l'interférence de la L1. Le tableau 11 illustre un cas de l'interférence du français sur l'apprentissage des combinaisons lexicales en chinois.

| Unité lexicale | Équivalent chinois et combinaisons lexicales | |
|----------------|--|--------------------------------|
| Envoyer | 寄 « faire parvenir quelque chose par la poste » | 邮件 « courrier » |
| | | 信 « lettre » |
| | | 邮包 « colis » |
| | 发送 « faire parvenir quelque chose par voie électronique » | 电子邮件 « courrier électronique » |
| | | 短信 « message texte » |
| | | 语音信息 « message vocal » |

Tableau 11. – Équivalence entre le verbe français *envoyer* et deux verbes chinois accompagnés de compléments d'objet direct

Il s'agit d'un cas de passage délicat du français au chinois. En français, le verbe envoyer véhicule le sens de « faire parvenir quelque chose à quelqu'un » (Académie française, 2019 ; Robert et al., 2020). En chinois, par contre, on fait appel à deux verbes distincts pour exprimer le sens de 'faire parvenir'. Chacun de ces verbes correspond à un intermédiaire spécifique : 寄 s'applique à l'envoi par la poste et 发送, à l'envoi électronique. Sous l'influence de leur langue maternelle, les apprenants francophones ont tendance à associer le verbe 寄 à un envoi électronique, par exemple, *寄短信 (« envoyer un message texte »)⁴³ et *寄电子邮件 (« envoyer un courrier électronique »). Ces deux constructions figurent parmi les erreurs les plus fréquentes tant à l'oral qu'à l'écrit selon mon expérience d'enseignement auprès d'étudiants francophones, vu que les COD, à savoir 短信 « message texte » et 电子邮件 « courrier électronique », contiennent

⁴³ Les constructions lexicales précédées d'un astérisque sont des exemples d'erreurs observées en situation d'apprentissage du chinois langue seconde.

respectivement le morphème 信^{xìn} (« lettre ») et 邮件^{yóu jiàn} (« courrier »). Ceux-ci évoquent le contexte d'envoi postal et constituent ainsi une source de confusion supplémentaire.

Comme on peut le constater, l'interférence de la L1 doit être prise en compte lors de la conception de matériel didactique, afin de mettre les apprenants en garde contre des erreurs en ce qui concerne le choix d'un cooccurrent approprié dans une CLS.

En conclusion, l'approche cognitive rend compte des aspects cognitifs en situation d'apprentissage d'une L2. Elle redéfinit ce que sont des moyens efficaces pour transmettre des connaissances aux apprenants non natifs. Nous nous appuyons sur cette approche pour explorer des moyens permettant une description structurée de la combinatoire spécialisée, notamment l'enseignement explicite des règles morphosyntaxiques et la présentation structurée des CLS.

Chapitre 4 Méthodologie

Notre démarche méthodologique, qui se déploie en trois phases, a été définie en fonction de nos objectifs de recherche, à savoir élaborer une méthode de description de la combinatoire spécialisée en vue de l'apprentissage du chinois par des locuteurs non sinophones et évaluer l'intérêt pédagogique d'une ressource lexicale faisant appel à cette méthode. Nous avons assemblé, dans un premier temps, un corpus spécialisé pour en extraire un échantillon de CLS ainsi que les données descriptives des CLS repérées. Ces renseignements nous ont servi de matériel de base pour concevoir, au cours de la deuxième phase, une méthode de description des CLS. Au cours de cette seconde étape, nous avons anticipé des difficultés faisant obstacle à l'acquisition de la combinatoire spécialisée et proposé différentes stratégies pour résoudre ces difficultés. Une méthode descriptive intégrant les solutions proposées a été appliquée à la création d'une ressource lexicale chinois-français intitulée CHINOINFO. La dernière phase consiste en une expérimentation comparative visant à évaluer l'efficacité de la ressource mise au point et l'intérêt qu'elle suscite auprès d'apprenants francophones. Nous avons mis au point un test lexical afin de faire passer le test à deux groupes de participants en mettant à leur disposition un nombre d'outils de référence. Seul un groupe a eu accès à CHINOINFO. Nous avons aussi élaboré des questionnaires afin de recueillir les renseignements sur le profil d'apprentissage des participants et leur recours à des ressources lexicales au cours du test.

À chaque phase de la réalisation du projet, nos choix méthodologiques ont été guidés par les trois volets de notre cadre théorique, soit le traitement des CLS, la représentation des propriétés sémantiques des CLS ainsi que la structuration des connaissances lexicales visant à améliorer l'efficacité pédagogique.

Ce chapitre est composé de trois sections. La section 4.1 aborde la collecte et le traitement des CLS extraites à partir d'un corpus. La section 4.2 décrit en détail la conception de la méthode de description, ainsi que l'application de cette méthode à la création du CHINOINFO. La dernière section est consacrée à la mise au point d'une expérimentation comparative visant à évaluer l'apport pédagogique du CHINOINFO et par son intermédiaire, l'intérêt pédagogique d'une partie de la méthode descriptive. Nous présentons la conception de l'expérimentation, le

déroulement de la collecte de données et la méthode d'analyse des données recueillies. Les résultats obtenus seront présentés dans les deux chapitres suivants.

4.1 Extraire et analyser des CLS à partir d'un corpus

Nous expliquons d'abord les critères pris en compte pour compiler un corpus aux fins de cette recherche (4.1.1). Nous décrivons ensuite la démarche adoptée pour repérer une liste de termes à partir de notre corpus (4.1.2) et extraire des CLS contenant les termes relevés pour analyser les propriétés syntaxico-sémantiques des CLS repérées (4.1.3).

4.1.1 Compilation du corpus

Nous avons choisi de collecter des CLS à partir d'un corpus pour deux raisons. Premièrement, notre démarche méthodologique du traitement des CLS est guidée par le principe sémasiologique de l'approche lexico-sémantique. Ce choix de fonder l'analyse des CLS sur leurs usages réels dans les discours spécialisés permet de résoudre les difficultés syntaxico-sémantiques soulevées par les CLS. Citons entre autres la polysémie des CLS dont la résolution nécessite une analyse ancrée dans le contexte linguistique des CLS. Nous expliquerons ces difficultés dans les sections 4.1.3.3 et 4.1.3.4.

Deuxièmement, le recours à un corpus spécialisé permet de mieux adapter notre recherche aux besoins des apprenants du COS. Comme nous l'avons expliqué dans la sous-section 2.3.1 portant sur l'enseignement du COS, l'objectif des apprenants est d'acquérir des compétences linguistiques qui leur permettent d'être « opérationnels », c'est-à-dire être en mesure de s'intégrer à un milieu académique ou professionnel sinophone. Notre choix d'assembler un corpus aux fins de cette recherche vise à réunir les textes adaptés au niveau d'apprentissage de notre public visé. Nous avons sélectionné les textes qui sont accessibles aux apprenants du COS du point de vue des thèmes abordés, du niveau de spécialité et du genre textuel. De cette manière, la collecte des CLS basée sur ce corpus permet de rendre compte des problèmes lexicaux authentiques auxquels se heurtent les apprenants.

Les textes composant notre corpus ont été choisis selon les critères que nous présentons ci-dessous.

4.1.1.1 Critères de sélection

Nous avons défini les critères suivants pour construire notre corpus : le domaine de spécialité, la langue de rédaction, le genre textuel, les destinataires, le niveau de spécialisation et la date de parution (Bowker et Pearson, 2002, p. 45-54; L'Homme, 2020a, p. 140-142).

Domaine de spécialité

Nous abordons, dans le cadre de cette thèse, les CLS en informatique et technologies de l'information. Les textes choisis portent plus précisément sur l'un des cinq thèmes suivants : 1) système informatique; 2) sécurité informatique; 3) génie logiciel; 4) traitement de données; 5) réseaux informatiques. Ces thèmes sont divisés en différents sous-thèmes, comme l'illustre le tableau 12.

| | Thèmes abordés | Sous-thèmes |
|---------|-----------------------|--|
| Thème 1 | Système informatique | Composantes et architectures de systèmes d'exploitation Installation et mise en service de systèmes d'exploitation (Microsoft et Linux) |
| Thème 2 | Sécurité informatique | Cybersécurité Sécurité dans les réseaux informatiques |
| Thème 3 | Génie logiciel | Conception de logiciels Développement d'applications mobiles Utilisation de logiciels |
| Thème 4 | Traitement de données | Gestion de bases de données Infonuagique |
| Thème 5 | Réseaux informatiques | Réseaux mobiles Internet des objets |

Tableau 12. – Thèmes et sous-thèmes des textes composant le corpus

Nous avons arrêté notre choix sur ces thèmes selon deux critères définis en fonction des objectifs pédagogiques de notre thèse. Premièrement, nous avons ciblé les thèmes traitant des applications concrètes de l'informatique, étant donné la visée de l'enseignement du lexique en COS (voir la sous-section 2.3.1). Deuxièmement, les données terminologiques recueillies à partir du corpus servent à élaborer une ressource lexicale qui intéresse les apprenants débutants du COS. Parmi les cas de figure, on trouve ceux qui suivent un cursus universitaire en Chine ou préparent leur entrée dans un programme d'études en informatique offert par une université chinoise. Compte tenu du profil des utilisateurs potentiels de la ressource, nous avons ciblé les thèmes accessibles aux étudiants inscrits à la première année d'un baccalauréat en informatique. Pour cerner les thèmes

faisant l'objet de notre corpus, nous avons consulté, à titre de références, les thèmes de recherche du Département d'informatique et de recherche opérationnelle (Département d'informatique et de recherche opérationnelle, 2020) et le répertoire des cours du premier cycle en informatique (Université de Montréal, 2020). Nous avons aussi visité les sites de quelques bibliothèques universitaires qui se classent parmi les plus grandes bibliothèques de la Chine, dont celle de l'Université de Pékin⁴⁴ et celle de l'Université Tsinghua⁴⁵, pour repérer les manuels destinés aux étudiants de baccalauréat en informatique de ces universités.

Langue de rédaction

Nous avons assemblé un corpus monolingue chinois. Comme la langue enseignée en COS s'écrit en chinois simplifié, nous avons veillé à choisir les textes conformes aux normes graphiques de ce système d'écriture. Ce choix se justifie en outre par la prise en compte de la variation géographique (ou topolectale) de la langue chinoise. Il existe en fait des différences importantes entre la langue chinoise utilisée en Chine continentale et celle qui est utilisée à Hongkong, à Macao et à Taïwan. En plus du système d'écriture des caractères (le chinois simplifié en Chine continentale et le chinois traditionnel à Hongkong, à Macao et à Taïwan), la variation linguistique touche aussi le lexique utilisé dans ces régions. Comme le montre le tableau 13, un même concept (par exemple, « souris ») peut être désigné de façon différente (le terme se traduit par ^{shǔ}鼠 ^{biāo}标 en Chine continentale et ^{huá shǔ}滑鼠 dans les trois autres régions). Une même forme lexicale (par exemple, ^{chéng xù}程序) peut véhiculer deux sens différents en fonction des territoires où elle est en usage (« programme » en Chine continental et « processus » dans les trois autres régions).

⁴⁴ Le site de la bibliothèque de l'Université de Pékin est repérable à l'adresse suivante : <https://www.lib.pku.edu.cn/portal/>.

⁴⁵ Le site de la bibliothèque de l'Université Tsinghua est repérable à l'adresse suivante : <https://lib.tsinghua.edu.cn/>.

| Variation géographique de termes en informatique | | | |
|--|--|---|---|
| Colonne 1 Région : Chine continentale Script : Chinois simplifié | Colonne 2 Régions : Hongkong/Macao/Taiwan Script : Chinois simplifié | Colonne 3 Régions : Hongkong/Macao/Taiwan Script : Chinois traditionnel | Équivalent français du terme (colonne 1) et de ses variants (colonnes 2 et 3) |
| 鼠标 | 滑鼠 | 滑鼠 | souris |
| 信息 | 资讯 | 資訊 | information |
| 网络 | 网路 | 網路 | réseau |
| 程序 | 程式 | 程式 | programme |
| 卸载 | 解除安装 | 解除安裝 | désinstaller |
| 应用程序/应用 | 應用程式/應用 | 應用程式/應用 | application |
| 进程 | 程序 | 程序 | processus |
| 硬盘 | 硬碟 | 硬碟 | disque dur |
| 内存 | 记忆体 | 記憶體 | mémoire interne |
| 硬件 | 硬体 | 硬體 | matériel |
| 软件 | 软体 | 軟體 | logiciel |
| 操作系统 | 作业系统 | 作業系統 | système d'exploitation |
| 移动设备 | 流动装置 | 流動裝置 | appareil mobile |
| 调试 | 侦错/除错 | 偵錯/除錯 | déboguer |
| 台式机 | 桌上型电脑 | 桌上型電腦 | ordinateur de bureau |
| 局域网 | 区域网路 | 區域網路 | réseau local |
| 文件 | 档案 | 檔案 | fichier |
| 文件夹 | 档案夹 | 檔案夾 | dossier |

Tableau 13. – Variation géographique des termes chinois en informatique

Pour éviter toute interférence causée par la variation terminologique topolectale, nous avons exclu les documents rédigés en chinois traditionnel et ceux qui proviennent des régions où le chinois traditionnel est en usage.

Genre textuel

L'étude de la combinatoire lexicale spécialisée, à savoir le point focal de notre recherche, doit s'appuyer sur un corpus de textes représentatifs de la spécificité discursive du milieu considéré. D'une part, la combinatoire s'explique par des usages préférentiels dans un milieu spécialisé (L'Homme, 2017, p. 219). D'autre part, l'analyse discursive des collocations en langue de spécialité (Pecman, 2012) met en évidence leur rôle spécifique au niveau du genre et du type de discours.

Pour élaborer une ressource lexicale des CLS qui reflète davantage leur fonctionnement dans le discours spécialisé, nous avons accordé une attention particulière au genre textuel lors de la construction du corpus. En premier lieu, le genre textuel permet de cerner les particularités expressives des textes spécialisés. L'Homme (2020a, p. 141) fait remarquer la différence entre les *types de documents* qui définissent la forme de la publication et le *genre textuel* qui distingue les textes en fonction de leur contenu, notamment leurs caractéristiques linguistiques telles que le vocabulaire utilisé et l'organisation du discours. Condamines (2003, p. 40) souligne, pour sa part,

la corrélation entre le genre et les régularités linguistiques des textes spécialisés. En deuxième lieu, le genre textuel caractérise la situation de production des textes (Condamines, 2003, p. 70-71) et en délimite ainsi le niveau de spécialisation.

Nous avons adopté la proposition de Bowker et Pearson (2002, p. 28) qui identifient trois types de situations de communications en langues de spécialité : 1) la communication entre les experts d'un même domaine; 2) les experts s'adressent aux semi-experts (étudiants ou experts d'un domaine connexe); 3) les experts s'adressent aux non-experts (grand public sans formation préalable dans le domaine abordé). Nous avons ciblé les deux premiers types de communication, étant donné que les lectorats visés sont comparables aux destinataires de la ressource que nous souhaitons créer dans le cadre de cette thèse. Les textes de vulgarisation n'ont pas été retenus. Ces textes, adaptés pour rejoindre un lectorat non expert, contiennent des explications en langue générale et évitent, dans la mesure du possible, le recours à des expressions hautement techniques. Par conséquent, la densité des CLS dans ces textes est relativement moins élevée par rapport aux deux autres types de textes spécialisés.

Les textes composant notre corpus sont divisés en cinq groupes selon leur genre.

1) Textes didactiques : ce genre textuel regroupe des manuels pédagogiques, du matériel de cours en ligne ouverts et massifs (nommés aussi MOOC⁴⁶) ainsi que des blogues et des chroniques tenus par des spécialistes reconnus dans leur domaine d'expertise. Ces textes sont destinés aux professionnels en devenir, à savoir les lecteurs en formation qui comprennent les étudiants allant jusqu'au premier cycle universitaire, selon la définition de L'Homme (2019). Les textes didactiques présentent les connaissances de base liées à un sujet spécialisé. Ils sont de nature explicative et riches en renseignements définitoires des concepts fondamentaux. Ces textes se caractérisent en outre par la densité élevée des termes et le recours aux expressions propres au domaine abordé.

2) Textes académiques : ce genre textuel correspond à des articles de périodiques scientifiques. Il s'agit d'un moyen de communication au sein d'une communauté d'experts qui possèdent des connaissances profondes dans un domaine. Outre les chercheurs universitaires et les étudiants aux

⁴⁶ MOOC est l'acronyme de *Massive Open Online Course*.

cycles supérieurs, le lectorat comprend aussi les professionnels qui ont reçu une formation spécialisée et exercent une profession dans le même domaine (L'Homme, 2019). Les textes de ce genre abordent des sujets de recherche très pointus et sont évalués par les pairs. On trouve peu de renseignements explicatifs ou définitoires dans ces textes hautement spécialisés.

3) Textes d'actualités professionnelles : on regroupe dans cette catégorie des articles de périodiques professionnels, des rapports officiels, ainsi que des contributions à des forums de discussion destinés aux professionnels en informatique. Ces textes abordent l'actualité dans le domaine et présentent les nouveautés ayant un impact sur le développement de l'industrie. Des rapports préparés par un organisme international ou gouvernemental se focalisent sur l'évolution dans un secteur d'activité spécifique, par exemple, la conception des applications mobiles. Les lecteurs visés comprennent aussi les professionnels en devenir qui s'orientent vers une carrière en informatique. Nous avons évalué la fiabilité des forums de discussion à l'aide des critères suivants : l'autorité de la source, la date de la dernière mise à jour, la fréquence des mises à jour et les liens vers des sites extérieurs pertinents et crédibles.

4) Textes publicitaires : nous avons réuni des publicités de produits et services visant les professionnels et les professionnels en devenir. Ces textes portent sur des outils informatiques (logiciel de programmation, logiciel de gestion de base de données, logiciel de courriel sécurisé, etc.) et des services informatiques (test logiciel, migration de base de données, etc.). Le public cible est censé avoir acquis des connaissances et des savoir-faire en informatique pour pouvoir exercer une profession dans ce domaine. Les textes retenus proviennent des sites Web des principaux fournisseurs de produits et de services informatiques reconnus en Chine, comme Huawei, Tencent et Baidu.

5) Tutoriels et guides d'utilisation : nous avons aussi recueilli des tutoriels et des guides d'utilisation, dont la plupart proviennent des mêmes sites Web où nous avons recueilli les publicités. Les textes sélectionnés visent à donner des instructions précises à suivre pour utiliser un produit informatique (logiciel de gestion de base de données) ou un service (solution de sécurité informatique et de réseau pour entreprise). Le contenu est organisé selon une structure énumérative. Les textes de ce genre, tout comme les publicités, sont destinés aux professionnels et aux professionnels en devenir et reflètent les usages préférentiels des CLS au sein de cette communauté.

Destinataires

Les destinataires des textes choisis se divisent en deux groupes selon leur niveau d'expertise dans le domaine abordé : 1) les experts ayant reçu une formation spécialisée en informatique, on regroupe dans cette catégorie les chercheurs universitaires, les étudiants aux cycles supérieurs et les professionnels qui exerce un métier dans l'un des sous-domaines présentés plus haut; 2) les semi-experts, ou experts en devenir, désignent les lecteurs en formation qui s'orientent vers une carrière professionnelle en informatique, à savoir les étudiants au premier cycle universitaire. Nous avons ciblé ces deux types de lectorat pour collecter les textes spécialisés qui sont à la portée des apprenants débutants du COS.

Niveau de spécialisation

Le niveau de spécialisation des textes est fonction du niveau d'expertise des lecteurs visés, à savoir les connaissances et les savoir-faire que ces derniers ont acquis au cours de leur cursus universitaire ou leur expérience professionnelle. Nous avons ciblé les textes présentant un niveau de spécialisation élevé ou intermédiaire. Premièrement, nous avons choisi des articles de périodiques scientifiques. Ces textes sont hautement spécialisés et s'adressent aux experts du domaine abordé. Deuxièmement, nous avons opté pour les textes d'un niveau de spécialisation intermédiaire : en plus des ressources didactiques, nous avons également repéré les textes portant sur les actualités professionnelles ou l'état actuel du développement du domaine, ainsi que les publicités et les tutoriels de produits et de services en informatique. Ces textes sont accessibles aux experts en devenir, à savoir les étudiants au premier cycle universitaire.

Le tableau 14 résume les différents critères pris en compte pour sélectionner les textes de notre corpus. Ce tableau nous permet de constater un certain regroupement entre le genre textuel, les types de documents, le niveau de spécialisation et le lectorat visé. Chacun de ces critères, comme le fait remarquer L'Homme (2020a, p. 141), met l'accent sur un aspect spécifique des textes choisis.

| Genre textuel | Types de documents | Situation de communication | Niveau de spécialisation | Lectorat visé |
|--|--|--|--------------------------|--|
| G1 Textes didactiques | 1) Manuel pédagogique; 2) Matériel de cours en ligne ouverts et massifs (MOOC); 3) Blogues et chroniques tenus par des spécialistes. | Experts ⇒ Semi-experts | Intermédiaire | Professionnels en devenir (étudiants allant jusqu'au premier cycle universitaire) |
| G2 Textes académiques | Articles de périodiques scientifiques | Experts ⇒ Experts | Élevé | 1) Chercheurs universitaires; 2) Étudiants aux cycles supérieurs; 3) Professionnels. |
| G3 Textes d'actualités professionnelles | 1) Articles de périodiques professionnels; 2) Rapports officiels; 3) Contributions à des forums de discussion destinés aux professionnels | Experts ⇒ Experts; Experts ⇒ Semi-experts | Élevé | 1) Professionnels; 2) Professionnels en devenir. |
| G4 Textes publicitaires | Publicités de produits et services en informatique | Experts ⇒ Experts; Experts ⇒ Semi-experts | Intermédiaire | 1) Professionnels; 2) Professionnels en devenir. |
| G5 Tutoriels et guides d'utilisation | 1) Tutoriels d'un produit informatique; 2) Guide d'utilisation et de soutien technique pour les utilisateurs d'un produit ou d'un service en informatique | Experts ⇒ Experts; Experts ⇒ Semi-experts | Intermédiaire | 1) Professionnels; 2) Professionnels en devenir. |

Tableau 14. – Résumé des critères de sélection des textes composant notre corpus

Date de parution

Nous avons délimité la période de parution des textes en fonction de leur genre, afin de nous assurer que le contenu des textes retenus soit actuel et pertinent. Pour les textes d'actualités professionnelles (G3), les textes publicitaires (G4) et les tutoriels et guides d'utilisation (G5), nous avons ciblé les documents parus entre 2018 et 2022. Comme ces genres de textes rendent compte des nouveautés dans le milieu professionnel, il est essentiel de réunir une documentation très récente compte tenu de l'évolution constante et rapide des technologies de l'information. Les textes scientifiques (G2) retenus ont été publiés au cours des dix dernières années, étant donné que ce genre de textes se focalise sur les avancées scientifiques ayant une influence à moyen terme ou à long terme sur le développement du domaine. Quant aux textes didactiques (G1), nous

avons ciblé les publications datant des vingt dernières années. Cette extension temporelle nous semble raisonnable puisque les concepts de base et les fondements théoriques dans ce domaine sont relativement stables. En outre, le rythme de publication des manuels est plus lent que certains autres types de textes, notamment les textes d'actualités professionnelles (G3) et les textes publicitaires (G4).

Équilibre du corpus

Nous avons veillé à l'équilibre du corpus quant au thème abordé et au genre textuel, afin d'assurer la représentativité de la documentation assemblée.

Pour trouver des textes respectant nos critères de sélection, nous avons effectué des recherches documentaires de trois façons différentes. Dans un premier temps, nous avons visité les sites de plusieurs bibliothèques publiques ou universitaires qui se classent parmi les plus grandes bibliothèques de la Chine, y compris la Bibliothèque nationale de Chine et la bibliothèque de l'Université de Pékin. Nous avons repéré des manuels pédagogiques en format électronique au moyen de recherches par mot-clé et par discipline. À l'aide des outils de recherche de ces bibliothèques, nous avons également accédé à diverses ressources documentaires en ligne, notamment des bases de données de publications scientifiques. Nous avons exploré deux bases de données qui sont largement utilisées par les chercheurs universitaires et les professionnels du domaine abordé : *CNKI.net* (China National Knowledge Infrastructure) et *Wanfang data*. Ces deux plateformes de diffusion des connaissances dans plusieurs domaines différents rassemblent des documents spécialisés numériques de nature variée, y compris les thèses et mémoires des universités chinoises, les périodiques scientifiques, les actes de colloque, les revues professionnelles, les brevets et les normes, etc. Les interfaces de ces bases de données permettent aux utilisateurs de rechercher les documents selon différents critères (mot-clé, discipline, auteur, année de publication, etc.) et d'y accéder. Outre les catalogues de bibliothèques, nous avons ensuite utilisé des moteurs de recherche pour retrouver d'autres documents dans le Web. Nous avons effectué des recherches par mot-clé dans trois moteurs de recherche : *Google Scholar*, qui indexe un large éventail de publications scientifiques et permet de chercher des mots-clés dans le texte intégral des documents; *Google* et *Baidu*, qui permettent de repérer les rapports officiels, les forums de discussion destinés aux professionnels ainsi que les blogues et chroniques tenus par des spécialistes. Enfin, nous avons effectué des recherches par mot-clé dans les sites Web de trois

principaux fournisseurs de produits et de services informatiques en Chine, à savoir Huawei, Tencent et Baidu, pour repérer les catalogues de produits et de services, les tutoriels de produits ainsi que les guides de soutien technique.

Pour ne retenir que les documents rédigés en chinois simplifié, nous avons affiné les paramètres de notre recherche documentaire dans les bases de données bibliographiques et les moteurs de recherche. Outre la langue de rédaction, nous avons aussi vérifié l'affiliation des auteurs afin d'exclure les documents en provenance des régions où le chinois traditionnel est en usage.

Notre corpus réunit 113 textes comportant au total 637,591 caractères. Nous avons eu recours à Microsoft Word pour calculer le nombre de caractères, compte tenu de la facilité d'utilisation et l'efficacité de ce logiciel de traitement textuel. Le nombre total de mots formés à partir de ces caractères s'élève à 341,676. Comme nous l'avons indiqué dans la section 1.1, le mot est l'unité minimale qui fonctionne de façon indépendante en chinois (Lu, 2003, p. 19). Un mot correspond à une forme lexicale qui comporte un seul caractère ou une combinaison de caractères chinois et qui peut être associée à une ou plusieurs unités lexicales. Le compte de mots a été réalisé à l'aide de Sketch Engine (Kilgarriff et al., 2014), une plateforme de gestion et d'analyse de corpus qui offre des fonctionnalités d'aide à la compilation de corpus. Tous les textes faisant partie de notre corpus sont numérotés. Les renseignements détaillés de ces textes sont présentés dans l'annexe 3, y compris la référence bibliographique, le thème abordé, le genre textuel, le nombre de caractères et le nombre de mots.

La répartition de nombre de textes par thème (figure 13) dans notre corpus est relativement équilibrée. Comme on peut le constater, les textes relevant du thème 3 (T3), soit le génie logiciel, représentent la proportion la plus importante (29 %) du corpus, vu que le T3 regroupe plus de sous-thèmes que les autres thèmes (voir aussi le tableau 14). La figure 14 illustre la répartition de volume de textes par thème abordé. Ces deux bilans permettent de constater une corrélation entre le nombre de textes et le volume de la documentation réunie pour les thèmes T2, T4 et T5. La documentation portant sur le système informatique (T1) est anormalement volumineuse puisqu'elle se compose majoritairement de manuels didactiques et de guides de soutien technique étoffés. En revanche, la documentation du T3, composée d'un bon nombre de textes courts, est de taille réduite malgré le nombre élevé des textes retenus.

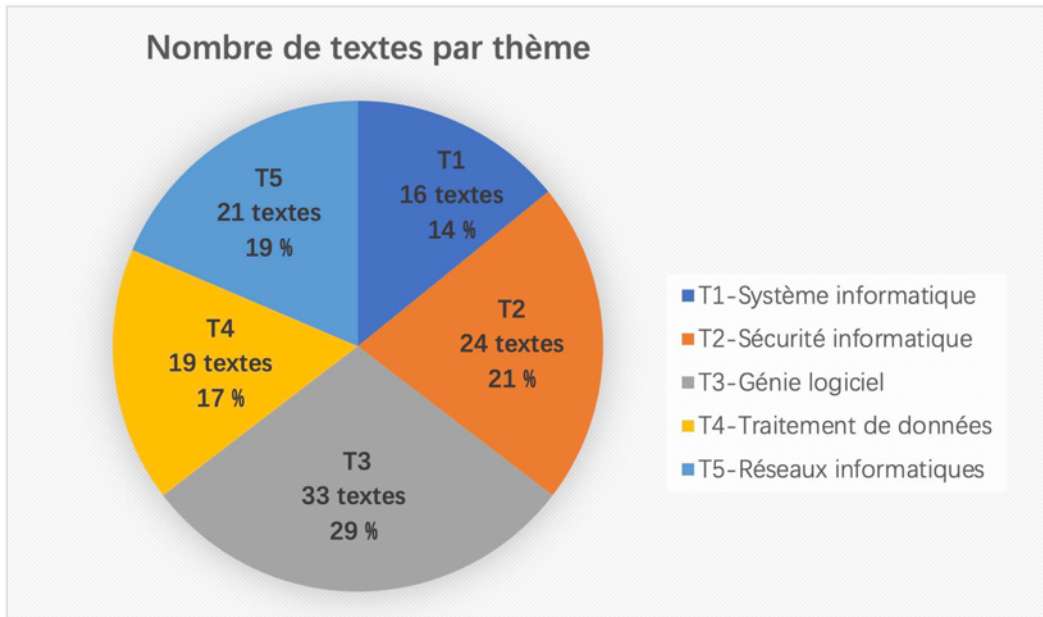


Figure 13. – Répartition de nombre de textes par thème abordé dans notre corpus

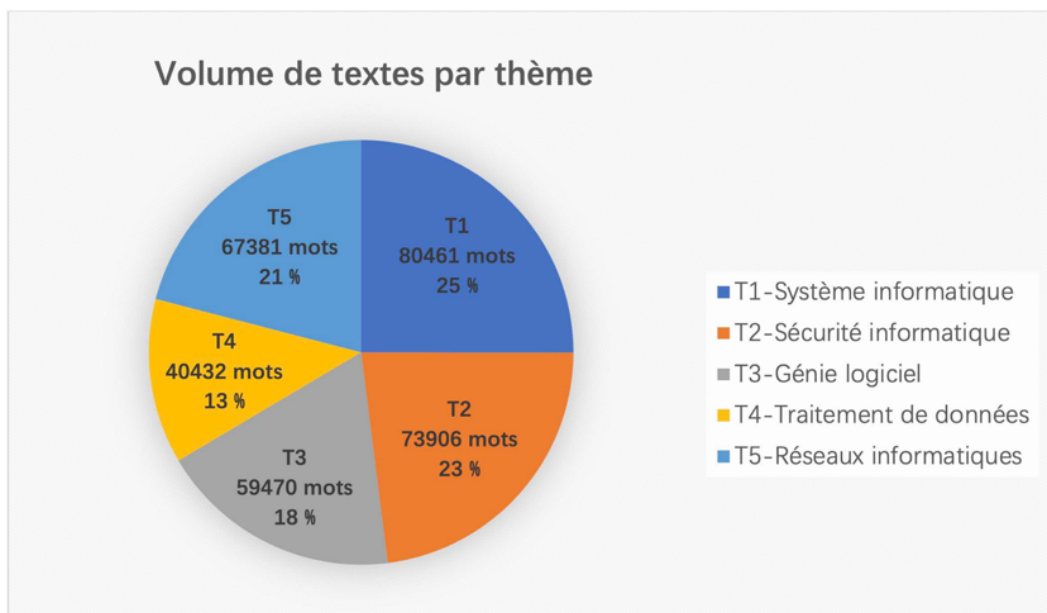


Figure 14. – Répartition de volume de textes par thème abordé dans notre corpus

Quant à la répartition de nombre de textes par genre (figure 15), les textes d'actualités professionnelles (G3) et les textes didactiques (G1) sont les plus nombreux vu que ces deux genres regroupent plus de types de documents que les autres genres. Du point de vue du nombre de mots, le G1 représente la proportion la plus importante (38 %) du corpus, compte tenu de la taille volumineuse de matériel didactique. Les genres G3, G4 et G5 se composent principalement

des textes concis (publicités, contributions à des forums, guide d'utilisation, etc.), ce qui explique la taille réduite de la documentation assemblée (figure 16).

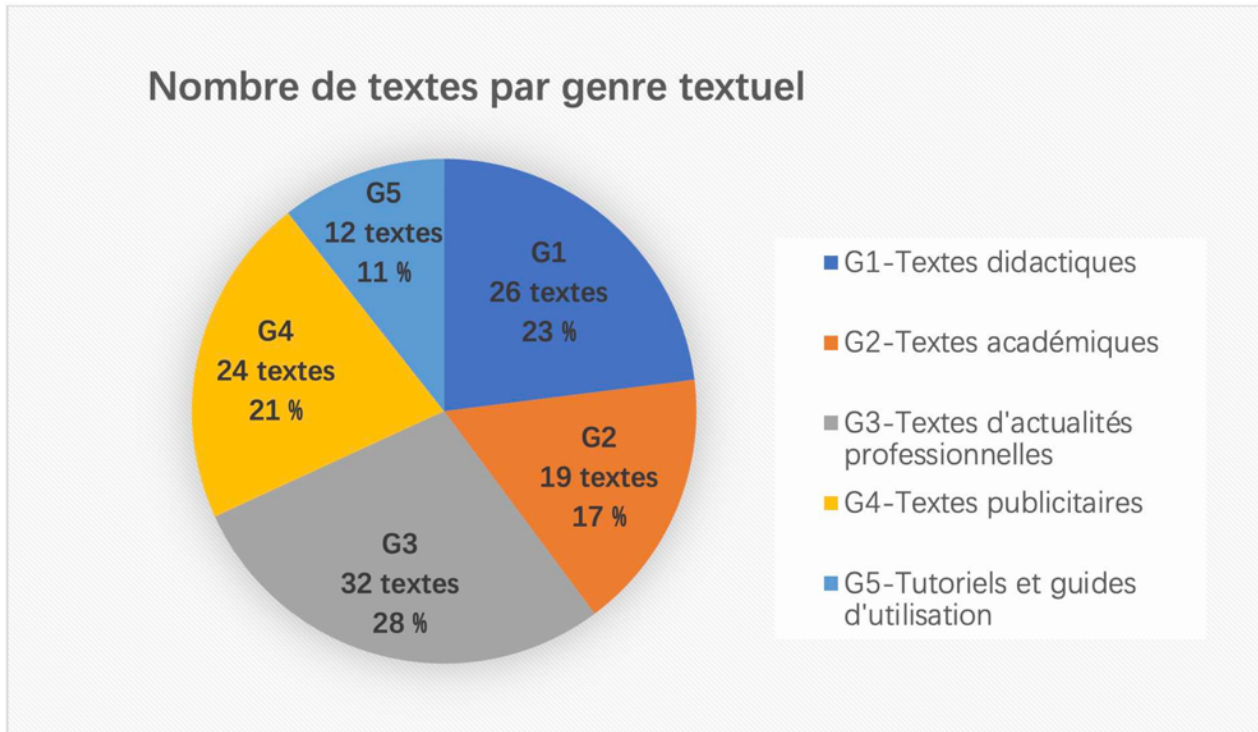


Figure 15. – Répartition de nombre de textes par genre textuel dans notre corpus

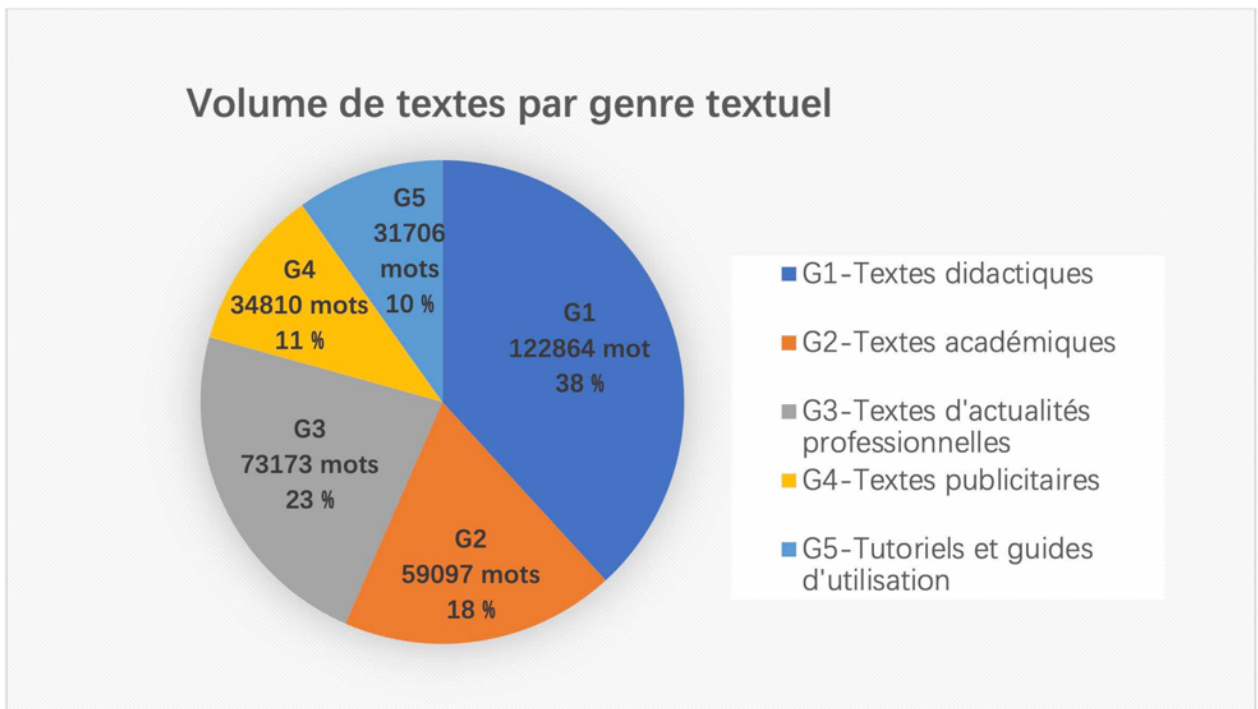


Figure 16. – Répartition de volume de textes par genre textuel dans notre corpus

4.1.1.2 Nettoyage et segmentation du corpus

Les documents recueillis ont fait l'objet d'un nettoyage et d'un traitement semi-automatique avant d'être soumis à un dépouillement terminologique. La première étape du nettoyage consiste à retirer les éléments textuels jugés inutiles pour l'extraction des CLS, à savoir les références bibliographiques, les mentions légales, les tableaux, les illustrations et les chapitres de manuel qui ne portent pas sur les thèmes de notre corpus. Les documents élagués contiennent encore des signes de ponctuation et des caractères spéciaux. Comme ces signes ne sont pas des éléments constitutifs des mots, ils doivent être exclus lors du compte de caractères composant les mots. Pour ce faire, nous avons utilisé le langage de programmation Python pour concevoir un script qui élimine de manière automatique les signes de ponctuation, des caractères spéciaux et des espaces entre les caractères. L'annexe 4 illustre le script Python qui nous a permis de supprimer un ensemble de signes prédéfinis dans un texte faisant partie du corpus. Le script a été exécuté de façon répétitive dans chacun des textes composant notre corpus. Le compte de caractères dans Word et le compte de mots dans Sketch Engine ont été réalisés après ce nettoyage.

Les documents nettoyés ont ensuite été soumis à Sketch Engine. Cet outil permet de créer un corpus à partir des fichiers téléversés par un utilisateur. Au cours de la compilation de notre corpus, Sketch Engine réalise de façon automatique la segmentation et l'étiquetage des textes soumis. L'étiquetage consiste à indiquer la partie du discours (nom, adjectif, verbe, pronom, etc.) des unités lexicales apparaissant dans les textes et à expliciter la fonction grammaticale (déterminant, modificateur de nom) de certaines unités lexicales. Certaines étiquettes s'avèrent particulièrement utiles pour expliquer les spécificités grammaticales de la langue chinoise. Par exemple, l'étiquette *BA* (*把*^{bǎ} in *ba-construction*) précise que l'unité lexicale *把*^{bǎ} introduit une construction verbale formée d'un verbe d'action et d'un complément de résultat. En chinois, le temps n'est pas marqué dans la forme du verbe. On utilise les aspects pour exprimer la façon dont une action est envisagée dans son déroulement ou sa réalisation (Bellassen et Arslangul, 2014, p. 122). L'étiquette *AS* (*Aspect marker*) attaché au suffixe verbal *着*^{zhe} précise que ce dernier marque un aspect de l'état prolongé qui exprime la continuité d'une action ou d'un état.

Le corpus compilé a été ensuite dépouillé à l'aide de plusieurs outils offerts par Sketch Engine, dont *Keywords*, *Word Sketch* et *Concordance*, afin d'en extraire un échantillon de CLS et les

contextes d'utilisation de celles-ci. L'utilisation que nous avons faite de ces outils est décrite dans les sections suivantes.

4.1.2 Identification des termes

Une fois le corpus mis en forme, nous avons procédé à son exploitation à l'aide de Sketch Engine. Notre première tâche consiste à repérer un nombre de termes qui soulèvent des difficultés syntaxico-sémantiques pour les apprenants lorsque ces termes forment des CLS. Cette tâche a été réalisée en trois étapes : 1) extraire une liste des candidats-termes (CT), soit « des mots ou des suites de mots qui sont susceptibles de correspondre à des unités terminologiques » (L'Homme, 2020a, p. 186); 2) identifier les unités terminologiques (UT)⁴⁷ à partir des CT repérés; 3) délimiter les acceptions des unités polysémiques identifiées et établir enfin la liste des termes dont les CLS seront extraites et analysées par la suite.

4.1.2.1 Extraction des candidats-termes (CT)

D'abord, nous avons eu recours à l'outil *Keywords* pour élaborer une liste de CT. Cet outil permet d'extraire automatiquement des mots dont la fréquence d'emploi est proportionnellement plus élevée dans notre corpus spécialisé que dans un corpus de référence. Ce dernier est choisi par l'utilisateur parmi une liste de corpus de référence disponibles dans Sketch Engine. Le corpus de référence par défaut pour la langue chinoise est *Chinese Web corpus 2017 Simplified (zhTenTen17)* et c'est celui auquel nous avons eu recours. Ce corpus est composé de textes recueillis dans le Web au moyen de *SpiderLing*, un outil de recherche automatisée de documents sur Internet. Il s'agit du plus grand corpus chinois dans les bases de données de Sketch Engine, dont le nombre total de mots s'établit à 13,5 milliards.

L'extraction génère deux types de résultats : *Keywords* et *Terms*. *Keywords* regroupe les CT composés d'un seul *mot*, soit l'unité minimale qui s'utilise de façon indépendante en chinois. *Terms* regroupe les CT composés d'au moins deux mots. Il est possible de définir, avant de lancer l'extraction terminologique, le type de CT que l'on souhaite repérer. Dans le cadre de cette

⁴⁷ Dans le cadre de notre recherche, les UT se définissent comme les unités linguistiques ayant un ou plusieurs sens envisagé(s) par rapport à un domaine de spécialité. Elles comprennent les unités monosémiques, soit les termes, ainsi que les unités polysémiques dont chaque sens peut correspondre à un terme (Alipour, 2018, p. 60, 99-101).

recherche, nous avons défini les paramètres d'extraction en fonction de nos critères de sélection de termes. En ce qui concerne la forme des termes à extraire, nous avons ciblé les termes simples qui correspondent à un seul mot. Les termes complexes ont été écartés puisqu'ils risquent d'être confondus avec les CLS.

La figure 17 illustre l'interface de recherche de l'outil *Keywords*. Nous avons coché l'option *Identify keywords* pour repérer exclusivement les CT simples. L'option *Focus on* permet d'identifier les CT dont le score de spécificité (*Keyness Score*) se situe dans un intervalle donné. Le score de spécificité est calculé au moyen de *Simple maths*, une méthode visant à établir un rapport quantitatif entre la fréquence d'emploi d'un CT dans un corpus spécialisé et la fréquence d'emploi du CT dans un corpus de référence. Dans l'option *Focus on*, le paramètre par défaut permet de localiser les CT dont le score de spécificité s'articule autour de 1. L'utilisateur peut modifier le paramètre de deux façons : il peut taper une valeur dans l'espace en dessous de *Focus On* ou positionner le curseur sur l'axe qui se trouve dans la même section. Les deux extrémités de l'axe sont respectivement nommées *rare* et *common*. Plus le curseur se déplace vers *rare*, plus les CT extraits sont spécifiques au corpus spécialisé. De cette manière, on se concentre sur les CT ayant une fréquence d'emploi comparativement faible dans le corpus de référence. En revanche, plus le curseur approche de l'extrémité *common*, plus la fréquence d'emploi d'un candidat-terme est élevée dans le corpus de référence.

Pour notre extraction terminologique, nous avons fixé le score de spécificité à 50, une valeur qui se positionne légèrement à gauche du centre de l'axe (figure 17), pour repérer les CT ayant un niveau de spécificité moyen. Cette démarche vise à relever les termes qui sont accessibles aux apprenants du chinois ayant une connaissance de base de l'informatique, par exemple, les étudiants au premier cycle universitaire.

Nous avons opté pour ce paramètre après avoir fait quelques observations préliminaires des résultats d'extraction à l'aide du paramètre par défaut. Ceux-ci présentent un taux de bruit élevé : on y trouve notamment les sigles en anglais (HTTP, ISP, SNMP, TELNET) et les noms de produits informatiques (FineReader, TiDB) qui ne sont pas pertinents pour la suite de notre travail.

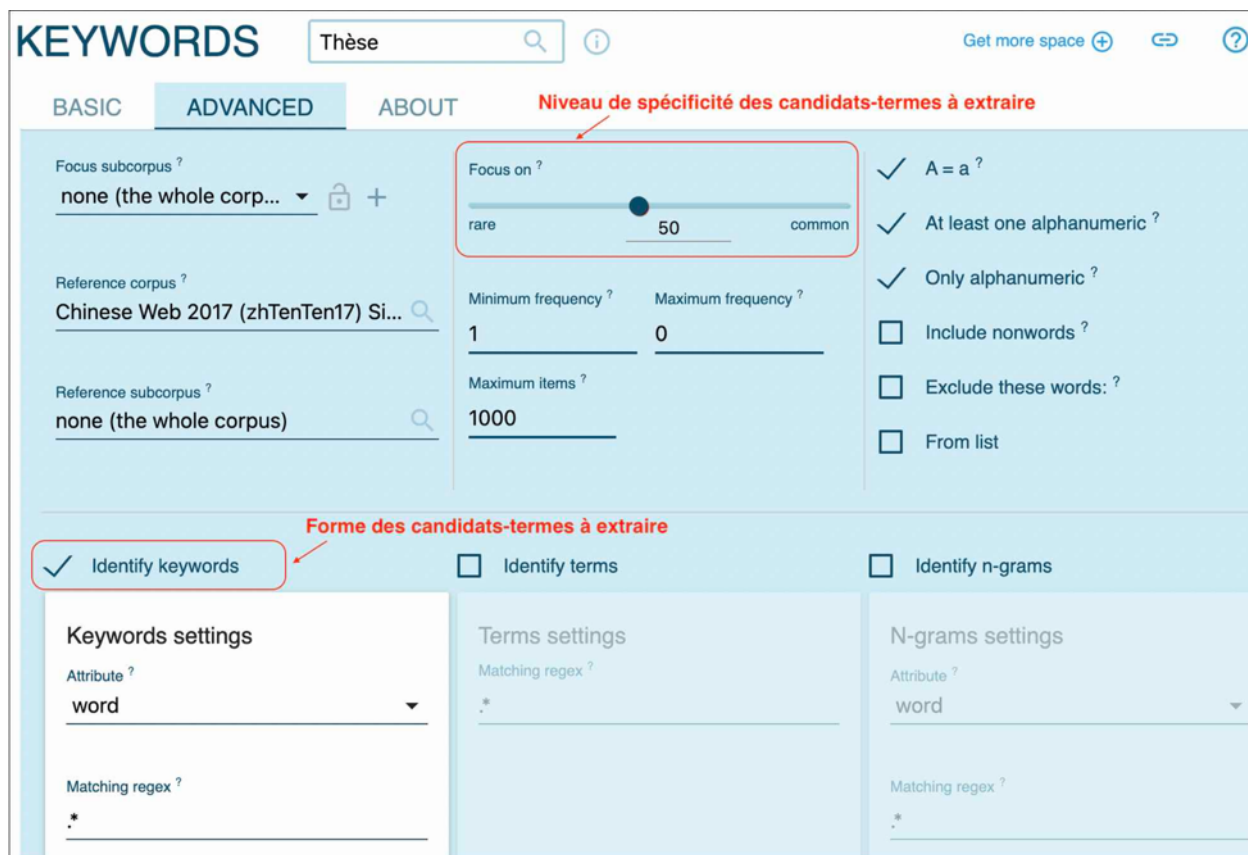


Figure 17. – Interface de recherche illustrant les paramètres de notre extraction terminologique

À la suite de l'extraction automatique, les résultats ont été triés par score de spécificité décroissant (figure 18). Ce mode d'organisation permet de repérer rapidement les termes les plus saillants dans notre corpus. Pour établir une liste de CT à analyser à la prochaine étape, nous avons filtré les résultats proposés selon trois critères : outre le score spécificité des CT, nous avons aussi tenu compte de leur fréquence d'emploi et de leur distribution dans le corpus. Le seuil de filtrage correspond à la valeur médiane des 1000 premiers résultats générés. Les CT retenus doivent respecter les critères suivants : 1) avoir un score de spécificité supérieur à 2,772; 2) avoir un nombre d'occurrences supérieur à 69, soit la médiane arrondie des 1000 premiers CT extraits; 3) apparaître dans plus de 19 documents dans notre corpus spécialisé, soit la médiane arrondie des 1000 premiers résultats. Cette démarche nous a permis de repérer 308 CT jugés significatifs d'un point de vue quantitatif et de leur spécificité.

KEYWORDS Thèse

Get more space

Paramètres d'extraction des CT

RESULT INFORMATION

Keywords attribute word × Reference corpus preloaded/zhTencent17_simplified_stf2 Simple N 50 ×

Only alphanumeric "yes" • Ignore case "yes" • Keywords regex .* • max_items 1000 • Minimum frequency 1 • N-grams attribute word • N-grams regex .* • At least one alphanumeric "yes" • Reference subcorpus none • Terms regex .* • Subcorpus none

SINGLE-WORDS ✓

reference corpus: Chinese Web 2017 (zhTenTen17) Simplified (items: 15,207)

Nombre d'occurrences Nombre de documents Score de spécificité (critère de tri des résultats)

| Word | Frequency | Document frequency | Score ? | Word | Frequency | Document frequency | Score ? | Word | Frequency | Document frequency | Score ? |
|------------|-----------|--------------------|----------|--------|-----------|--------------------|----------|----------|-----------|--------------------|----------|
| 1 程序 | 2,174 | 90 | 33.4 ... | 18 文件 | 1,307 | 83 | 17.9 ... | 35 云计算 | 257 | 25 | 13.1 ... |
| 2 加密 | 704 | 45 | 33.0 ... | 19 口令 | 305 | 16 | 17.5 ... | 36 Linux | 229 | 24 | 12.8 ... |
| 3 漏洞 | 694 | 38 | 29.9 ... | 20 数据库 | 516 | 54 | 17.4 ... | 37 缓存 | 226 | 24 | 12.6 ... |
| 4 恶意 | 611 | 44 | 27.7 ... | 21 密钥 | 301 | 20 | 17.3 ... | 38 文档 | 314 | 42 | 12.4 ... |
| 5 计算机 | 1,254 | 69 | 27.6 ... | 22 权限 | 385 | 44 | 17.2 ... | 39 安全 | 2,743 | 88 | 12.2 ... |
| 6 存储 | 924 | 74 | 27.5 ... | 23 数据 | 2,980 | 99 | 17.0 ... | 40 安全性 | 305 | 53 | 12.2 ... |
| 7 访问 | 977 | 79 | 27.3 ... | 24 密码 | 489 | 45 | 16.8 ... | 41 通信 | 457 | 55 | 11.9 ... |
| 8 软件 | 2,584 | 93 | 25.9 ... | 25 终端 | 543 | 43 | 16.4 ... | 42 协议 | 607 | 47 | 11.6 ... |
| 9 攻击 | 985 | 48 | 24.5 ... | 26 备份 | 330 | 33 | 16.0 ... | 43 进程 | 342 | 34 | 11.5 ... |
| 10 Windows | 502 | 40 | 24.4 ... | 27 网络 | 2,653 | 91 | 15.7 ... | 44 算法 | 288 | 32 | 11.5 ... |
| 11 Web | 425 | 34 | 22.8 ... | 28 操作 | 1,526 | 96 | 15.2 ... | 45 运行 | 772 | 84 | 11.3 ... |
| 12 服务器 | 736 | 68 | 22.3 ... | 29 防火墙 | 272 | 19 | 14.8 ... | 46 云 | 567 | 43 | 11.3 ... |
| 13 邮件 | 669 | 40 | 21.8 ... | 30 病毒 | 484 | 43 | 14.1 ... | 47 连接 | 467 | 67 | 11.0 ... |
| 14 移动 | 1,390 | 66 | 19.0 ... | 31 代码 | 386 | 58 | 14.0 ... | 48 主机 | 255 | 33 | 10.9 ... |
| 15 应用 | 2,467 | 103 | 18.9 ... | 32 系统 | 3,752 | 106 | 14.0 ... | 49 检测 | 649 | 46 | 10.9 ... |
| 16 用户 | 2,710 | 102 | 18.9 ... | 33 攻击者 | 232 | 28 | 13.8 ... | 50 远程 | 295 | 39 | 10.9 ... |
| 17 传输 | 520 | 59 | 18.6 ... | 34 单击 | 254 | 15 | 13.8 ... | | | | |

Rows per page: 50 1-50 of 1,000 < 1 /20 >

Figure 18. – Les 50 premiers CT extraits par l’outil *Keywords* (triés par score de spécificité)

4.1.2.2 Identification des unités terminologiques (UT)

Une fois les CT relevés, nous avons procédé à la validation de leur statut terminologique, afin d’identifier les unités terminologiques (UT). Comme nous l’avons évoqué en début de la section 4.1.2, les UT se divisent en deux types : 1) les *unités monosémiques* ou *termes*, à savoir les unités lexicales dont le sens est envisagé par rapport au domaine abordé; 2) les *unités polysémiques* ou *polysèmes*, qui résultent de l’association d’une forme lexicale à plusieurs sens spécialisés. Il s’agit, autrement dit, d’un groupe de termes ayant une forme lexicale identique (Alipour, 2018, p. 60).

En premier lieu, nous avons éliminé les CT qui peuvent être facilement identifiés comme faisant partie du bruit, à savoir : 1) les CT dont le sens n’est pas lié au domaine abordé (L’Homme, 2020a, p. 214-215), par exemple, les conjonctions (如 « par exemple », 即 « à savoir »,

bìng « et »), les pronoms (它 « il », 它们 « ils », 您 « vous »); 2) les mots qui ne sont pas segmentés de façon appropriée, tel que 数据流 (« flux de données »). Ce terme a été segmenté en deux éléments : 数据 (« données ») et 流 (« flux ») qui ont été proposés comme deux CT. Toutefois, le second élément 流 (« flux ») est un morphème lié, il s'agit d'une unité de sens qui ne peut fonctionner de façon indépendante en chinois. Il se combine avec d'autres morphèmes pour former une unité lexicale. Les chaînes de caractères en anglais, dont les abréviations (IP, App) et les marques déposées (Windows, Android), ont aussi été exclues, étant donné que notre projet vise la description des termes chinois.

Nous avons ensuite appliqué trois critères lexico-sémantiques (L'Homme, 2020a, p. 72-74) pour repérer les UT à partir des autres CT. Selon le critère a), un terme véhicule un sens lié au domaine de spécialité abordé (L'Homme, 2020a, p. 72). Nous avons ainsi identifié les UT comme 软件 (« logiciel »), 程序 (« programme »), 网络 (« réseau ») et 计算机 (« ordinateur ») qui sont directement liés à l'informatique et aux technologies de l'information, plus précisément, à l'un des cinq thèmes que nous avons définis dans la section 4.1.1.1.

Selon le critère b), un terme précatif peut être identifié selon la nature de ses actants sémantiques (L'Homme, 2020a, p. 72-73). Par exemple, le verbe 安装 (« installer ») fait appel à trois actants sémantiques : l'*agent* qui entreprend l'action d'installer, le *patient* qui subit l'action et la *destination* qui renvoie au lieu visé par l'action. Le premier actant peut être réalisé par des termes comme 用户 (« utilisateur ») et 管理员 (« administrateur »). Le deuxième actant peut être réalisé par des termes tels que 程序 (« programme »), 软件 (« logiciel ») et 应用 (« application »). Le troisième actant est réalisé le plus souvent par le terme 计算机 (« ordinateur ») ou son synonyme 电脑 (« ordinateur »). Comme les termes réalisant les actants sémantiques de 安装 (« installer ») véhiculent un sens spécialisé en informatique, ce verbe est reconnu comme UT.

Le critère c) propose d'identifier un terme selon la parenté morphologique (L'Homme, 2020a, p. 73). Si l'on valide le statut terminologique d'une unité lexicale en vertu des critères a) et b), les dérivés de ce terme sont forcément spécialisés. Ce critère n'est pas applicable à la langue chinoise puisque les unités lexicales ne peuvent pas donner lieu à des dérivés morphologiques.

Selon le critère d), toute autre relation paradigmatique partagée par une unité lexicale avec un terme déjà admis relève d'un sens spécialisé (L'Homme, 2020a, p. 74). Par exemple, 网络

(« réseau »), qui désigne l'ensemble des moyens matériels et logiciels qui permettent l'échange de données, de services ou de ressources entre plusieurs appareils informatiques, est admis comme UT selon le critère a). Ainsi, toute unité lexicale dénotant un élément constitutif du réseau, par exemple, 服务器 (^{fú wù qì} « serveur »), 主机 (^{zhǔ jī} « ordinateur hôte »), 防火墙 (^{fáng huǒ qiáng} « pare-feu »), doit être identifiée comme UT puisqu'elle entretient une relation partie-tout avec le « réseau ».

À l'issue de cette étape, nous avons repéré 228 UT parmi les 308 CT, ce qui équivaut à un taux de bruit de 25,9 %. Rappelons que ces UT incluent des *termes* (ou *unités monosémiques*) tels que 程序 (^{chéng xù} « programme »), ainsi que des *polysèmes*⁴⁸ (ou *unités polysémiques*) tels que 加密 qui correspond à trois termes : 加密_{1/v} (^{jiā mì} « crypter »), 加密_{2/nom} (^{jiā mì} « cryptage ») et 加密_{3/adj} (^{jiā mì} « crypté »). L'annexe 5 présente la liste des CT extraits de notre corpus, y compris les UT retenues et les CT non terminologiques.

4.1.2.3 Distinction sémantique des UT

Nous avons ensuite procédé à l'analyse sémantique des UT sélectionnées dans le but d'identifier les polysèmes et de distinguer les acceptions de ces polysèmes. L'UT 加密 (^{jiā mì}) que nous avons présentée dans la sous-section 4.1.2.2, par exemple, est un polysème qui regroupe une acception verbale (« crypter »), une acception nominale (« cryptage ») et une acception adjectivale (« crypté »). Chaque acception d'un polysème correspond à un terme et fait l'objet de l'extraction des CLS par la suite. L'analyse des CLS contenant des polysèmes nous permet d'anticiper, d'une façon plus générale, les difficultés syntaxico-sémantiques faisant obstacle à l'acquisition des CLS (voir aussi la section 1.1 consacrée à la problématique).

Comme notre analyse s'appuie sur les contextes d'utilisation des UT, nous avons eu recours à l'outil *Concordance* dans Sketch Engine qui permet de localiser une UT dans les énoncés où elle apparaît. La figure 19 illustre la liste de concordances de l'UT 更新 (^{gēng xīn} « mettre à jour ») générée par Sketch Engine. Les résultats sont présentés sous forme d'index KWIC (*Key Word In Context*). Ce

⁴⁸ Dans notre travail, le terme *polysème* fait référence à une forme lexicale prise dans toutes ses acceptions. Selon la LEC, un polysème est un ensemble d'unités lexicales de forme identique qui appartiennent en général à la même partie du discours (Mel'čuk et al., 1995, p. 59, 156). À la différence de la LEC, nous examinons également des polysème dont les acceptions sont reliées à des parties du discours distinctes.

mode d'affichage permet de présenter les occurrences de l'UT recherchée assorties de ses contextes immédiats. Il est possible de cliquer sur l'une des concordances pour avoir accès à un contexte plus étendu.

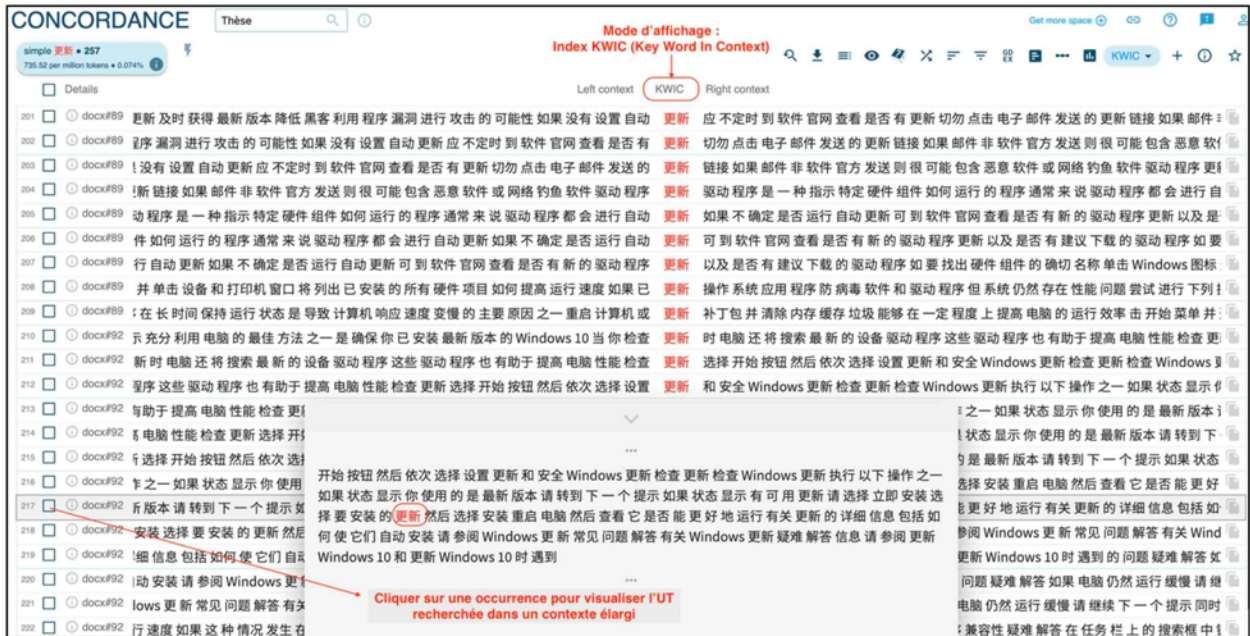


Figure 19. – Extrait de la liste de concordances de l'UT ^{gēng xīn}更新 générée par l'outil *Concordance*

Notre démarche de distinction sémantique s'est déroulée en deux temps : 1) identifier les parties du discours associées à une UT; 2) distinguer les sens d'une même partie du discours. En premier lieu, nous avons d'abord distingué les parties du discours d'une UT. Bien que Sketch Engine étiquette notre corpus et indique les parties du discours de certaines UT, on constate souvent des erreurs d'étiquetage. Le dérivé sémantique nominal ou adjectival d'un verbe, tel que 加密_{2/nom} (« cryptage ») et 加密_{3/adj} (« crypté »), est parfois étiqueté comme un verbe (加密_{1/v} « crypter »). De ce fait, nous avons tenté d'identifier les parties du discours en tenant compte de la compatibilité de l'UT avec d'autres éléments dans une phrase où elle apparaît. Considérons les trois phrases contenant l'UT ^{gēng xīn}更新 que nous avons extraites de notre corpus :

1) 及时^{jí shí}/adv 更新^{gēng xīn}/v 系统^{xì tǒng}/nom 能够^{néng gòu} 更好^{gèng hǎo} 地^{de}⁴⁹ 避免^{bì miǎn} 病毒^{bìng dú} 和^{hé} 其他^{qí tā}
 À temps mettre à jour système pouvoir mieux DE₂ éviter virus et autre
 恶意^{è yì} 软件^{ruǎnjiàn} 的^{de} 入侵^{rù qīn}。
 malveillant logiciel DE₁⁵⁰ intrusion.

« Le fait de **mettre à jour** le système à temps permet de mieux le protéger contre les virus et autres logiciels malveillants. » [Notre traduction]

2) 用户^{yòng hù} 应^{yīng} 及时^{jí shí} 对^{duì} 应用程序^{yìngyòngchéng xù} 和^{hé} 操作系统^{cāo zuò xì tǒng} 进行^{jìn xíng}/v 更新^{gēng xīn}/nom^o
 Utilisateur devoir à temps envers applications et système effectuer mise à jour.
 d'exploitation

« L'utilisateur doit effectuer à temps la **mise à jour** des applications et la **mise à jour** du système d'exploitation. » [Notre traduction]

3) Windows Update 可^{kě} 自动^{zì dòng} 安装^{ān zhuāng}/v [Windows] 更新^{gēng xīn}/nom 以及^{yǐ jí}
 Windows Update pouvoir automatiquement installer [Windows] mise à jour et
 一些^{yì xiē} [应用程序^{yìngyòngchéng xù}] 的^{de} 更新^{gēng xīn}/nom^o
 un nombre de applications DE₂ mise à jour

« Windows Update peut installer automatiquement les **mises à jour** de [Windows] et les mises à jour d'[applications] » [Notre traduction]

Nous y avons identifié deux parties du discours associées à l'UT étudiée : 更新^{gēng xīn}/v (« mettre à jour ») et 更新^{gēng xīn}/nom (« mise à jour »). Dans la première phrase, l'UT est de nature verbale (更新^{gēng xīn}/v « mettre à jour ») puisqu'elle est modifiée par l'adverbe 及时^{jí shí} (« à temps ») et suivie d'un COD 系统^{xì tǒng} (« système »). Dans les deux autres phrases, l'UT est de nature nominale (更新^{gēng xīn}/nom « mise à

⁴⁹ 地 (« DE₁ ») est une particule structurale qui marque la modification du verbe. Cette particule sert à relier un ou plusieurs modificateur(s) de nature adverbiale au verbe modifié (Bellassen et Arslangul, 2014, p. 106).

⁵⁰ 的 (« DE₂ ») est une particule structurale qui marque la modification du nom. Cette particule sert de lien entre un ou plusieurs modificateur(s) adjectival(aux) ou nominal(aux) et le nom modifié (Bellassen et Arslangul, 2014, p. 73).

jour ») pour deux raisons : premièrement, elle remplit la fonction syntaxique de COD des verbes 进行 (jìn xíng « effectuer ») et 安装 (ān zhuāng « installer »); deuxièmement, elle accepte un déterminant qui peut être le nom d'un système d'exploitation (Windows) ou un terme désignant un type de logiciel (应用程序 yìng yòng chéng xù « application »), introduit dans la plupart des cas par la particule structurale 的 (de)⁵¹.

En deuxième lieu, nous avons procédé à la distinction des sens à l'intérieur d'une même partie du discours. Pour ce faire, nous avons appliqué quelques tests lexico-sémantiques (L'Homme, 2020a, p. 76-80), inspirés par ceux de la sémantique lexicale (Cruse, 1986, p. 55-57; Mel'čuk et al., 1995, p. 64-68), afin de valider le fait que nous ayons affaire à des sens distincts.

Nous présentons ci-dessous les trois critères de distinction sémantique mis en œuvre.

1) La substitution par un synonyme. Une UT ne peut être remplacée par un même synonyme si elle véhicule plus d'un sens dans différents contextes. Ce critère nous permet de confirmer que l'UT 移动 (yí dòng) revêt deux sens différents (移动₁ « mobile » et 移动₂ « amovible ») dans les CLS 移动₁互联网 (yí dòng hù lián wǎng « Internet mobile ») et 移动₂硬盘 (yí dòng yìng pán « disque dur amovible ») relevées dans notre corpus. Dans le premier contexte, il est possible de remplacer 移动₂ (« amovible ») par 便携式 (biàn xié shì « portatif »), tandis que le premier contexte n'admet pas la substitution par la même unité lexicale.

2) La présence de liens paradigmatiques différentiels. Ce critère vérifie la possibilité de dégager des ensembles de liens paradigmatiques différentiels correspondant à des sens distincts véhiculés par une même UT. Nous avons appliqué ce test pour identifier certains polysèmes verbaux. Par exemple, nous avons dégagé deux acceptions de l'UT 访问 (fǎng wèn) (访问₁ « accéder 1 ») et 访问₂ (« accéder 2 ») dans les CLS 访问₁磁盘 (fǎng wèn cí pán « accéder à un disque magnétique ») et 访问₂服务器 (fǎng wèn fú wù qì « accéder à un serveur »). Ces deux sens correspondent à deux termes qui apparaissent dans des paradigmes distincts. 访问₁ (« accéder 1 ») signifie « un lecteur, un processeur ou un utilisateur commence à intervenir sur des données placées dans une mémoire ou sur un support de stockage après avoir localisé ces données » (Observatoire de linguistique Sens-Texte, 2008). Ce

⁵¹ En chinois, la particule structurale 的 marque la modification du nom. Elle sert de lien entre le(s) déterminant et le nom déterminé dans un syntagme nominal (Bellassen et Arslangul, 2014, p. 73).

terme entre dans le même paradigme que les verbes 寻址 (xún zhǐ « adresser ») et 读取 (dú qǔ « lire »). Le terme 访问₂ (fǎng wèn signifie « un utilisateur s’est connecté et commence à utiliser un réseau ou de l’information se trouvant dans un réseau » (Observatoire de linguistique Sens-Texte, 2015). Il entretient une relation paradigmatique avec les verbes 连接 (lián jiē « se connecter ») et 获取 (huò qǔ « avoir accès à »).

3) La cooccurrence compatible et la cooccurrence différentielle. Ce test tente de combiner une UT avec des cooccurrents différents. Si les cooccurrents auxquels une UT s’associe ne sont pas compatibles, l’UT est associée à plus d’un sens (Mel’čuk et al., 1995, p. 64-68). Reprenons l’exemple de 访问 : comme cette UT exprime deux sens différents dans les CLS 访问₁ 磁盘 (fǎng wèn cí pán « accéder à un disque magnétique ») et 访问₂ 网站 (fǎng wèn wǎngzhàn « accéder à un site Web »), il est impossible de l’associer en même temps à ces deux cooccurrents, soit 磁盘 (cí pán « disque magnétique ») et 网站 (wǎngzhàn « site Web »). L’expression ainsi obtenue est inacceptable : *访问磁盘和网站 (fǎng wèn cí pán hé wǎngzhàn « accéder à un disque magnétique et à un site Web »). Ce test s’est avéré particulièrement utile pour distinguer les sens des UT verbales. Nous avons confirmé, par exemple, deux acceptions de 检测 (jiǎn cè « détecter ») et 检测₂ « mettre à l’essai » dans les CLS 检测₁ 病毒 (jiǎn cè bìng dú « détecter un virus ») et 检测₂ 软件 (jiǎn cè ruǎnjiàn « mettre un logiciel à l’essai »), compte tenu de l’incompatibilité des cooccurrents (*检测₂ 病毒和软件 (jiǎn cè bìng dú hé ruǎnjiàn « détecter un virus et un logiciel »). Dans la première CLS, l’UT exprime le sens de « déceler la présence d’un virus ou d’une erreur » (Observatoire de linguistique Sens-Texte, 2007), alors que dans la seconde CLS, on dégage une autre acception de l’UT, soit « tester, à l’aide d’un programme approprié, l’état d’un logiciel, d’un réseau ou d’un système informatique, afin d’en évaluer le niveau de sécurité et d’en déterminer les vulnérabilités ». L’application de ce test nous permet également de faire la distinction sémantique subtile des polysèmes nominaux. Nous avons observé que l’UT 网络 (wǎng luò) exprime deux sens (网络₁ « réseau₁ » et 网络₂ « réseau₂ ») dans les CLS 无线网络 (wú xiàn wǎng luò « réseau sans fil ») et 社交网络 (shè jiāo wǎng luò « réseau social ») : la formulation erronée *无线和社交网络 (wú xiàn hé shè jiāo wǎng luò « réseaux sans fil et social) révèle que 无线 (wú xiàn « sans fil ») et 社交 (shè jiāo « social ») se combinent avec deux termes différents.

Ce test nous a permis de confirmer que l’UT 更新 (gēng xīn) véhicule deux sens nominaux différents, en plus d’un sens verbal (更新₁ (gēng xīn « mettre à jour ») que nous avons identifié à l’étape précédente. Reprenons les trois phrases suivantes extraites de notre corpus :

1) 及时^{jí shí}/adv 更新^{gēng xīn}/v 系统^{xì tǒng}/nom 能够^{néng gòu} 更好^{gèng hǎo} 地^{de} 52 避免^{bì miǎn} 病毒^{bìng dú} 和^{hé} 其他^{qí tā}
 À temps mettre à jour système pouvoir mieux DE₂ éviter virus et autre
 的^{de} 恶意^{è yì} 软件^{ruǎnjiàn} 的^{de} 入侵^{rù qīn}。
 malveillant logiciel DE₁⁵³ intrusion.

« Le fait de **mettre à jour** le système à temps permet de mieux le protéger contre les virus et autres logiciels malveillants. » [Notre traduction]

2) 用户^{yòng hù} 应^{yīng} 及时^{jí shí} 对^{duì} 应用程序^{yìngyòngchéng xù} 和^{hé} 操作系统^{cāozuò xì tǒng} 进行^{jìn xíng}/v 更新^{gēng xīn}/nom^o
 Utilisateur devoir à temps envers applications et système effectuer mise à jour.
 d'exploitation

« L'utilisateur doit effectuer à temps la **mise à jour** des applications et la **mise à jour** du système d'exploitation. » [Notre traduction]

3) Windows Update 可^{kě} 自动^{zì dòng} 安装^{ān zhuāng}/v [Windows] 更新^{gēng xīn}/nom 以及^{yǐ jí}
 Windows Update pouvoir automatiquement installer [Windows] mise à jour et
 一些^{yì xiē} [应用程序^{yìngyòngchéng xù}] 的^{de} 更新^{gēng xīn}/nom^o
 un nombre de applications DE₂ mise à jour

« Windows Update peut installer automatiquement les **mises à jour** de [Windows] et les mises à jour d'[applications] » [Notre traduction]

Dans les phrases 2) et 3), l'UT 更新^{gēng xīn}/nom (« mise à jour ») se combine respectivement avec 进行^{jìn xíng} (« effectuer ») et 安装^{ān zhuāng} (« installer »). L'incompatibilité de ces deux cooccurrents verbaux (*进行和安装更新^{xíng hé ān zhuāng gēng xīn} (« effectuer et installer une mise à jour ») révèle que l'on a affaire à deux sens

⁵² 地 (« DE₁ ») est une particule structurale qui marque la modification du verbe. Cette particule sert à relier un ou plusieurs modificateur(s) de nature adverbiale au verbe modifié (Bellassen et Arslangul, 2014, p. 106).

⁵³ 的 (« DE₂ ») est une particule structurale qui marque la modification du nom. Cette particule sert de lien entre un ou plusieurs modificateur(s) adjectival(aux) ou nominal(aux) et le nom modifié (Bellassen et Arslangul, 2014, p. 73).

différents. 更新₂ (« mise à jour₁ ») est un nom d'action tandis que 更新₃ (« mise à jour₂ ») constitue un nom de résultat.

4.1.2.4 Repérage des termes à partir des UT polysémiques

À la suite de l'analyse sémantique des UT relevées, nous avons identifié 33 polysèmes correspondant à 91 termes. Ces polysèmes se divisent en deux types : 1) les *mots multicatégoriels* (兼类词) (Zhu, 1982; Lu, 1994, 2003; Packard, 2000, p. 34-36; Arcodia et Basciano, 2021, p. 276-280) qui regroupent les termes relevant de différentes parties du discours et 2) les polysèmes réunissant les termes d'une même partie du discours. Nous présentons, dans cette sous-section, la délimitation des termes à partir de ces polysèmes.

Le premier type de polysèmes est lié à la spécificité morphosyntaxique du chinois. À la différence des langues indo-européennes, le chinois se caractérise par l'absence de dérivation morphologique. En chinois, les unités lexicales d'une même famille, à savoir celles qui relèvent de différentes parties du discours et entretiennent entre elles un lien de dérivation sémantique (Mel'čuk et Polguère, 2007, p. 18), peuvent partager la même forme lexicale et forment ainsi un ensemble nommé *mot multicatégoriel*. Nous avons repéré, à titre d'exemple, le mot multicatégoriel 下载 qui est un polysème à deux acceptions : il regroupe le verbe 下载₁ (« télécharger ») et son dérivé sémantique nominal 下载₂ (« téléchargement ») qui désigne l'action de 'télécharger'. Il s'agit d'un cas particulier de polysémie lexicale, comme le soulignent Zhu (1982) et Lu (1994, 2003a). Ces deux termes se combinent respectivement avec 软件 (« logiciel ») pour former deux CLS de structure syntaxique différente : 下载₁ 软件 (« télécharger un logiciel ») et 软件下载₂ (« téléchargement de logiciel »). La distinction des acceptions d'un polysème est essentielle pour appréhender le sens des CLS qui le contiennent. Cette tâche est d'autant plus difficile que celles-ci sont associées à une même forme. Considérons les exemples suivants : 1) 备份₁ 数据 (« sauvegarder des données ») et 2) 备份₂ 数据 (« données sauvegardées »). La première CLS est une combinaison verbale, dans laquelle le verbe 备份₁ (« sauvegarder ») remplit la fonction syntaxique de prédicat verbal. Dans la seconde CLS de nature nominale, 备份₂ (« sauvegardé ») agit comme le modificateur du nom 数据 (« données »). La distinction de ces CLS qui sont syntaxiquement ambiguës repose sur l'analyse des contextes de celles-ci. La démarche que nous avons adoptée pour discerner le sens des CLS sera présentée dans la section suivante.

Les mots multicatégoriels mettent en jeu différents ensembles de parties du discours. Ce phénomène touche notamment les verbes, les noms et les adjectifs (An, 2005; Guo, 2002; Hu, 1996a, 1996b). Une étude quantitative (Hu, 1996b) montre que le type verbe/nom⁵⁴ représente de 12,91 % à 19,33 % du nombre total des mots multicatégoriels en chinois. An (2005) et Yu et al. (2005) font aussi remarquer que le type verbe/nom s’observe souvent dans le lexique chinois. Nous avons repéré, à partir des UT identifiés à l’étape précédente, quatre types de mots multicatégoriels, comme le montre le tableau 15. Ces résultats confirment des constats faits par d’autres chercheurs avant nous.

| Type de mots multicatégoriels | Unité lexicale | Transcription phonétique (pinyin) | Partie du discours | Équivalent français | Nombre de mots multicatégoriels relevés |
|-------------------------------|--|-----------------------------------|--------------------|-----------------------|---|
| Verbe/nom | 传输 ₁ <small>chuán shū</small> | [chuán shū] | Verbe | transférer | 18 |
| | 传输 ₂ <small>chuán shū</small> | | Nom | transfert | |
| Verbe/nom/adj | 加密 ₁ <small>jiā mì</small> | [jiā mì] | Verbe | crypter | 3 |
| | 加密 ₂ <small>jiā mì</small> | | Nom | cryptage | |
| | 加密 ₃ <small>jiā mì</small> | | Adjectif | crypté,e | |
| Nom/adjectif | 安全 ₁ <small>ān quán</small> | [ān quán] | Nom | sécurité informatique | 3 |
| | 安全 ₂ <small>ān quán</small> | | Adjectif | sécurisé | |
| Adjectif/adverbe | 远程 ₁ <small>yuǎn chéng</small> | [yuǎn chéng] | Adjectif | distant | 6 |
| | 远程 ₂ <small>yuǎn chéng</small> | | Adverbe | à distance | |

Tableau 15. – Types de mots multicatégoriels relevés à partir des UT

Outre des mots multicatégoriels, nous avons aussi identifié des polysèmes qui comprennent des termes ayant des sens apparentés et relevant d’une même partie du discours. Ces polysèmes sont divisées en quatre sous-types suivants selon la partie du discours des termes regroupés et le lien sémantique entre eux-ci :

⁵⁴ Le mot multicatégoriel de type verbe/nom regroupe un verbe et ses dérivés sémantiques nominaux. Ces derniers peuvent dénoter un même fait, tel un événement, une action, un état, une relation, etc. (Mel’čuk et al., 1995, p. 76), ou désigner le résultat de l’action exprimée par le verbe.

1) Regroupement de verbes et de leur dérivé sémantique nominal [verbe 1 + verbe 2 + nom d'action 1 + nom d'action 2]. Nous avons identifié des polysèmes réunissant des termes verbaux qui font appel à différents actants sémantiques⁵⁵ et leur dérivé sémantique nominal, soit les noms exprimant le même sens que ces verbes. Il s'agit en fait d'un ensemble de mots multicatégoriels reliés par un rapport d'affinité sémantique. Par exemple, l'UT 运行^{yùn xíng} recouvre quatre acceptions, comme le montre le tableau 16. Les verbes 运行^{yùn xíng}₁ (« s'exécuter ») et 运行^{yùn xíng}₂ (« exécuter ») ont des structures actanciennes différentes. 运行^{yùn xíng}₁ (« s'exécuter ») désigne « un programme fonctionne sur un ordinateur ou sous un système d'exploitation » (Observatoire de linguistique Sens-Texte, 2013 b). Ce terme fait appel à deux actants : 1) le *patient* renvoie à l'entité qui subit l'action exprimée par le terme, cet actant est réalisé par des termes typiques comme 程序^{chéng xù} (« programme »); 2) l'*environnement* renvoie à l'ensemble des éléments nécessaires à la réalisation de l'action exprimée, cet actant est réalisé par des termes typiques tels que 计算机^{jì suàn jī} (« ordinateur ») ou 操作系统^{cāo zuò xì tǒng} (« système d'exploitation »). 运行^{yùn xíng}₂ (« exécuter ») signifie « un utilisateur fait en sorte qu'un programme fonctionne sur un ordinateur ou sous un système d'exploitation » (Observatoire de linguistique Sens-Texte, 2013b). Il fait appel à trois actants, à la différence de 运行^{yùn xíng}₁ (« s'exécuter ») : outre le *patient* et l'*environnement*, 运行^{yùn xíng}₂ (« exécuter ») implique aussi un *agent* qui entreprend l'action exprimée (tableau 16).

Nous avons ensuite décrit les relations lexicales⁵⁶ entre les différentes acceptions au moyen d'un système de gloses et de définitions adaptées aux termes retenus dans le CHINOINFO (sous-section 3.1.1). Nous reviendrons sur l'encodage des liens sémantiques dans la sous-section 4.2.1.1.

⁵⁵ Rappelons que le sens d'un terme prédicatif est exprimé par un *prédicat sémantique* qui dénote un fait, tel un événement, une action, un état, une relation, etc., ainsi que les *actants sémantiques* qui participent au fait dénoté (L'Homme, 2020a, p. 70-71; Mel'čuk et al., 1995, p. 76).

⁵⁶ Le lien sémantique entre les différentes acceptions d'un polysème peut être modélisé au moyen des FL. Par exemple, la FL **Caus** qui exprime la causation d'un état ou d'un événement (Mel'čuk et al., 1995, p. 143) permet de représenter la relation entre 运行^{yùn xíng}₁ (« s'exécuter ») et 运行^{yùn xíng}₂ (« exécuter ») de la façon suivante : **Caus** (运行^{yùn xíng}₁) = 运行^{yùn xíng}₂ (faire en sorte qu'un programme s'exécute). Le lien de dérivation syntaxique entre ces verbes et leur nominalisation (运行^{yùn xíng}₃ « exécution_{1a} » et 运行^{yùn xíng}₄ « exécution_{1b} ») peut être représenté par la FL **nominalisation [S₀]** comme suit : S₀ (运行^{yùn xíng}₁) = 运行^{yùn xíng}₃ ; S₀ (运行^{yùn xíng}₂) = 运行^{yùn xíng}₄. Rappelons que, dans notre recherche, les FL nous servent d'appui dans notre démarche d'analyse des liens sémantiques.

| Terme | Équivalent en français | Définition | Structure actancielle | | | | | |
|-------|------------------------|--|----------------------------|----------------|--------------------------|----------------------------------|----------------------------------|------------------|
| | | | Nature de participant | Premier actant | Deuxième actant | Prédicat | | |
| 运行1 | s'exécuter | Un programme fonctionne sur un ordinateur ou sous un système d'exploitation. [Source : Dicolnfo] | Nature de participant | Premier actant | Deuxième actant | Prédicat | | |
| | | | Terme typique & traduction | 程序 programme | (在) 计算机 (sur) ordinateur | 运行1 s'exécuter | | |
| | | | Rôle sémantique | Patient | Environnement | Prédicat | | |
| 运行2 | exécuter | Un utilisateur fait en sorte qu'un programme fonctionne sur un ordinateur ou sous un système d'exploitation. [Source : Dicolnfo] | Nature de participant | Premier actant | Deuxième actant | Prédicat | Troisième actant | |
| | | | Terme typique & traduction | 用户 utilisateur | (在) 计算机 (sur) ordinateur | 运行2 exécuter | 程序 programme | |
| | | | Rôle sémantique | Agent | Environnement | Prédicat | Patient | |
| 运行3 | exécution 1a | Processus pendant lequel un programme fonctionne sur un ordinateur ou sous un système d'exploitation. [Source : Dicolnfo] | Nature de participant | Premier actant | Deuxième actant | | Prédicat | |
| | | | Terme typique & traduction | 程序 programme | (在) 计算机 (sur) ordinateur | (的) <i>particule structurale</i> | 运行3 exécution 1a | |
| | | | Rôle sémantique | Patient | Environnement | | Prédicat | |
| 运行4 | exécution 1b | Action réalisée par un utilisateur qui fait en sorte qu'un programme fonctionne sur un ordinateur ou sous un système d'exploitation. [Source : Dicolnfo] | Nature de participant | Premier actant | Deuxième actant | Troisième actant | Prédicat | |
| | | | Terme typique & traduction | 用户 utilisateur | (在) 计算机 (sur) ordinateur | (对) 程序 (envers) programme | (的) <i>particule structurale</i> | 运行4 exécution 1b |
| | | | Rôle sémantique | Agent | Environnement | Patient | | Prédicat |

Tableau 16. – Délimitation de quatre acceptions de l'UT ^{yùn xíng} 运行

2) Regroupement de deux noms à sens prédicatif [nom d'action + nom de résultat]. Nous avons aussi identifié des polysèmes constitués de deux termes nominaux à sens prédicatif, ceux-ci désignent respectivement une action et le résultat de l'action. Prenons les exemples de ^{gēng xīn}更新₂ (« mise à jour₁ ») et ^{gēng xīn}更新₃ (« mise à jour₂ ») qui ont été présentés dans la sous-section 4.1.2.3. Il est possible de différencier ces deux termes en prenant compte de la relation de dérivation sémantique qu'ils entretiennent avec le verbe ^{gēng xīn}更新₁ (« mettre à jour ») regroupé dans un même mot multicatégoriel : ^{gēng xīn}更新₂ (« mise à jour₁ ») est le nom d'action de ^{gēng xīn}更新₁ (« mettre à jour »), tandis que ^{gēng xīn}更新₃ est le nom de résultat du verbe. Ces liens de dérivation sémantique peuvent être modélisés au moyen des FL [S_{res}] (nom de résultat) et [S₀] (nom de même sens) de la manière suivante :

$$S_{res} (\text{更新}_1) = \text{更新}_3$$

$$S_0 (\text{更新}_1) = \text{更新}_2$$

3) Regroupement de deux noms reliés par un lien d'holonymie [holonyme-méronyme]. Ce type de polysèmes recouvrent deux acceptions reliées par une relation sémantique de tout-partie. Considérons l'UT ^{ruǎn jiàn}软件 qui semble, à la première vue, dotée d'un seul sens, soit « logiciel ». Pour confirmer notre intuition, nous avons appliqué le test de la présence des liens paradigmatiques

(voir la sous-section 4.1.2.3) pour vérifier si 软件^{ruǎnjiàn} est polysémique. Ce test vérifie la possibilité de dégager des ensembles de liens paradigmatiques correspondant à des sens distincts véhiculés par une même UT. Nous avons identifié, à l'aide de ce critère, deux sens associés à 软件^{ruǎnjiàn} dans les deux phrases suivantes extraites de notre corpus :

1) 如果^{rú guǒ} 您^{nín} 选择^{xuǎn zé} 使用^{shǐ yòng} 第三方^{dì sān fāng} 软件^{ruǎnjiàn}₁, 请^{qǐng} 务必^{wù bì}
 Si vous choisir de utiliser tiers logiciel, être prié de sans faute
 从^{cóng} 可信^{kě xìn} 来源^{lái yuán} 下载^{xià zài} 软件^{ruǎnjiàn}₁。
 en provenance de fiable source télécharger logiciel.

« Si vous choisissez d'utiliser un logiciel tiers, assurez-vous de le télécharger à partir d'une source fiable. » [Notre traduction]

2) 在^{zài} 计算机^{jì suàn jī} 系统^{xì tǒng} 中^{zhōng}, 所有^{suǒ yǒu} 硬件^{yìng jiàn} 部件^{bù jiàn} (如^{rú} CPU、
 À ordinateur système à l'intérieur de tous matériel composant par exemple unité centrale
 存储器^{cún chǔ qì}、 输入^{shū rù} 输出^{shū chū} 设备^{shè bèi} 等^{děng}) 称^{chēng} 做^{zuò} 硬件^{yìng jiàn} 资源^{zī yuán};
 mémoire entrée sortie périphérique etc. appeler comme matériel ressource;
 而^{ér} 程序^{chéng xù} 和^{hé} 数据^{shù jù} 等^{děng} 信息^{xìn xī} 称^{chēng} 为^{wéi} 软件^{ruǎnjiàn}₂ 资源^{zī yuán}。
 tandis que programme et données etc. information appeler comme logiciel ressource.

« Dans un système informatique, tous les composants matériels (par exemple, l'unité centrale, la mémoire, les périphériques d'entrée et de sortie, etc.) sont appelés ressources matérielles, tandis que les informations telles que les programmes et les données sont appelées ressources logicielles. » [Notre traduction]

L'UT 软件^{ruǎnjiàn} correspond à deux termes apparaissant dans des paradigmes distincts. Dans la première phrase, 软件^{ruǎnjiàn}₁ (« logiciel₁ ») renvoie à l'« ensemble des programmes constituant une unité destinée à effectuer un traitement particulier sur un ordinateur » (Observatoire de linguistique Sens-Texte, 2013d). Ce terme est l'hyperonyme 应用^{yìng yòng} (« logiciel d'application ») et entre dans le même paradigme que 版本^{bǎn běn} (« version ») qui désigne l'état d'un logiciel à un moment donné de son développement. 软件^{ruǎnjiàn}₂ (« logiciel₂ ») fait référence à l'« ensemble des ressources logiques utilisées par un utilisateur pour accomplir des tâches sur un ordinateur » (Observatoire de linguistique Sens-Texte, 2013d). Ce terme est le co-méronyme (L'Homme,

2020a, p. 110) de 硬件^{yǐng jiàn} (« matériel ») qui désigne l'« ensemble des éléments informatiques physiques qu'un utilisateur utilise pour effectuer une tâche sur un ordinateur » (Observatoire de linguistique Sens-Texte, 2015 b). 软件^{ruǎn jiàn}₂ (« logiciel₂ ») et 硬件^{yǐng jiàn} (« matériel ») ont un holonyme commun qui est 计算机系统^{jì suàn jī xì tǒng} (« système informatique ») puisqu'ils renvoient aux deux composants essentiels d'un système informatique.

Du point de vue syntaxique, 软件^{ruǎn jiàn}₁ (« logiciel₁ ») se combine avec le verbe 下载^{xià zài} (« télécharger ») pour former une CLS régie par la structure syntaxique « prédicat verbal-complément d'objet ». Quant à la CLS 软件^{ruǎn jiàn}₂资源^{zī yuán} (« ressources logicielles »), il s'agit d'une construction appositive dans laquelle 软件^{ruǎn jiàn}₂ (« logiciel₂ ») et l'apposition de 资源^{zī yuán} (« ressources »). Du point de vue sémantique, cette CLS renvoie à la même entité que 软件^{ruǎn jiàn}₂ (« logiciel₂ »). Elle met l'accent sur le fait que « logiciel₂ » est envisagé en tant qu'un ensemble des ressources.

Cette analyse nous a permis d'identifier la relation sémantique entre les deux acceptions du polysème 软件^{ruǎn jiàn} : 软件^{ruǎn jiàn}₁ (« logiciel₁ ») est le méronyme de 软件^{ruǎn jiàn}₂ (« logiciel₂ ») et désigne un élément constitutif de ce dernier. En revanche, 软件^{ruǎn jiàn}₂ (« logiciel₂ ») est l'holonyme de 软件^{ruǎn jiàn}₁ (« logiciel₁ ») et désigne l'ensemble auquel ce dernier appartient.

4) Regroupement de deux adjectifs ayant une composante sémantique commune [adjectif 1

+ adjectif 2]. Un polysème peut réunir deux termes adjectivaux ayant des sens apparentés. Nous avons délimité, par exemple, deux acceptions de l'UT 移动^{yí dòng} : 移动^{yí dòng}₁ « mobile » et 移动^{yí dòng}₂ « amovible » (voir aussi la sous-section 4.1.2.3). Les deux adjectifs expriment le sens de « qui n'est pas fixe, qui peut être déplacé ». Pour distinguer ces deux termes dans les CLS 移动^{yí dòng}₁互联^{hù lián}网^{wǎng} (« Internet mobile ») et 移动^{yí dòng}₂硬盘^{yìng pán} (« disque dur amovible »), nous avons appliqué deux tests lexico-sémantiques : la substitution par un synonyme et la présence des liens paradigmatiques. D'abord, seul 移动^{yí dòng}₂ (« amovible ») peut être remplacé par 便携式^{biàn xié shì} (« portatif ») ou 可移动^{kě yí dòng} (« détachable »). Ensuite, 移动^{yí dòng}₂ (« amovible ») entretient un lien d'antonymie avec 固定^{gù dìng} (« fixe »), tandis que 移动^{yí dòng}₁ « mobile » entre dans le même paradigme que 无线^{wú xiàn} (« sans fil »).

À l'issue de l'extraction terminologique, nous avons établi une liste de 91 termes, regroupés en 33 vocables⁵⁷. Chaque terme fait l'objet d'un article dans le dictionnaire CHINOINFO que nous présentons dans le chapitre 5. Chaque terme est accompagné d'un numéro afin de distinguer les différentes acceptions d'un vocable. Les acceptions au sein d'un vocable sont ordonnées selon leur partie du discours et le lien de dérivation sémantique entre elles. Le numéro d'acception est attribué dans l'ordre suivant : 1) verbe intransitif; 2) verbe transitif; 3) nom d'action, 4) nom de résultat; 5) adjectif; 6) adverbe. L'annexe 6 présente la liste complète des termes sélectionnés.

Les contextes extraits du corpus à cette étape sont consignés dans un fichier de travail Excel pour dégager, par la suite, les relations sémantiques qu'un terme sélectionné entretient avec d'autres termes du domaine. Nous aborderons la description des liens sémantiques dans la section 4.2.1.

4.1.3 Collecte et analyse des CLS

À la suite du repérage des termes, nous avons procédé à la collecte des CLS et à l'analyse de leurs propriétés syntaxico-sémantiques. Notre démarche comprend quatre étapes : 1) extraire un échantillon de CLS renfermant des termes identifiés (4.1.3.1); 2) relever un nombre de contextes d'utilisation des CLS extraites (4.1.3.2); 3) analyser les données recueillies (4.1.3.3) pour anticiper les difficultés syntaxico-sémantiques soulevées par les CLS (4.1.3.4). Les données descriptives des CLS ont ensuite été compilées aux fins d'élaboration du dictionnaire CHINOINFO.

4.1.3.1 Extraction des CLS

Notre première tâche consiste à élaborer une liste de CLS pour chacun des termes relevés. Pour ce faire, nous avons adopté une méthode d'extraction semi-automatique en faisant appel aux outils *Word Sketch* et *Concordance* de Sketch Engine.

⁵⁷ Rappelons que nous nous éloignons de la terminologie utilisée par la LEC dans le cas de *vocable*. Selon la LEC, deux parties du discours distinctes sont associées à des vocables différents. Dans notre travail, le terme *vocable* fait référence à un ensemble de termes de forme identique qui sont associées à une même partie du discours ou à des parties du discours distinctes (voir aussi la sous-section 3.1.1).

En premier lieu, nous avons eu recours à la fonction *Word Sketch* pour obtenir un résumé rapide du comportement combinatoire d'un terme. Cette fonction génère automatiquement une liste de cooccurrents pour un terme et classe les résultats selon la structure syntaxique des CLS formées. Si le terme recherché est une acception d'un mot multicatégoriel, *Word Sketch* extrait des cooccurrents de façon distincte pour chacune des parties du discours associées à ce mot. Par exemple, l'outil propose deux listes de résultats pour la forme lexicale 攻击^{gōng jī} qui correspond à deux termes : 攻击^{gōng jī}/_v (« attaquer ») et 攻击^{gōng jī}/_{nom} (« attaque informatique »). Les cooccurrents du terme 攻击^{gōng jī}/_{nom} (« attaque informatique ») sont regroupés en différentes catégories (figure 20) selon la fonction syntaxique du terme par rapport à ces cooccurrents, y compris le modificateur, le nom modifié, le complément d'objet, le sujet, etc. Dans la deuxième catégorie « N_Modifier », nous avons relevé un nombre de modificateurs nominaux qui se combinent avec 攻击^{gōng jī}/_{nom} (« attaque informatique ») dans des CLS telles que 黑客攻击^{hēi kè gōng jī}/_{nom} (« pirate-attaque » « piratage informatique »), 钓鱼攻击^{diào yú gōng jī}/_{nom} (« hameçonnage-attaque » « attaque par hameçonnage ») et 中间人攻击^{zhōng jiān rén gōng jī}/_{nom} (« intermédiaire-attaque » « attaque par interception »). La troisième catégorie répertorie un nombre de verbes qui forment des CLS de type « prédicat verbal-complément d'objet » avec 攻击^{gōng jī}/_{nom} (« attaque informatique »), dont 实施攻击^{shí shī gōng jī}/_{nom} (« effectuer une attaque informatique »), 发起攻击^{fā qǐ gōng jī}/_{nom} (« lancer une attaque informatique ») et 拦截攻击^{lán jié gōng jī}/_{nom} (« intercepter une attaque informatique »).

Par défaut, les résultats sont triés par score de typicalité (*typicality score*) calculé au moyen de *logDice*, un algorithme visant à mesurer le degré d'association des éléments d'une combinaison lexicale (Rychlý, 2008). Il convient d'expliquer de façon simplifiée le fonctionnement de *logDice*. Pour mesurer le degré d'association entre le terme 程序^{chéng xù} (« programme ») et ses cooccurrents verbaux, l'algorithme repère d'abord toutes les occurrences de combinaisons contenant un verbe et 程序^{chéng xù} (« programme ») dans le corpus. Ensuite, Sketch Engine recherche, pour chaque verbe identifié (comme 下载^{xià zài}₁ « télécharger »), toutes les combinaisons contenant ce verbe et un nom dans le corpus. Chaque fois que le verbe 下载^{xià zài}₁ (« télécharger ») se trouve en combinaison avec 程序^{chéng xù} (« programme »), ce verbe obtient un point positif. En revanche, chaque fois que le verbe 下载^{xià zài}₁ (« télécharger ») se trouve en combinaison avec un autre nom, il obtient un point négatif. En conséquence, les verbes qui ont tendance à se combiner avec un grand nombre de noms obtiendront un score relativement faible puisque leur usage combinatoire est

relativement flexible. Les verbes qui ne se combinent qu'avec 程序 (chéng xù « programme ») ou un nombre restreint d'autres noms obtiendront un score relativement élevé puisque leur usage combinatoire est plus restreint et spécifique. Autrement dit, plus le score de typicité est élevé, plus l'association entre le terme et son cooccurrent est forte (Rychlý, 2008; « logDice », 2016; « Most frequent or most typical collocations? », 2017).

Le tri des résultats par score permet d'identifier les combinaisons lexicales les plus typiques d'un terme. Reprenons l'exemple de 攻击 (gōng jī /nom « attaque informatique »), le cooccurrent 黑客 (hēi kè « pirate informatique ») a un score de 10,9 et apparaît en cooccurrence avec 攻击 (gōng jī /nom) dans 13 contextes. Comparé à 网络 (wǎng luò « réseau »), un cooccurrent qui s'utilise plus fréquemment (37 occurrences dans notre corpus) mais qui a un score moins élevé (9,2), 黑客 (hēi kè « pirate informatique ») est un résultat moins prévisible mais plus approprié pour décrire le comportement combinatoire du terme 攻击 (gōng jī /nom « attaque informatique »).

Les options d'affichage de *Word Sketch* permettent de trier les résultats par fréquence. Il est aussi possible d'afficher un exemple d'usage pour chaque cooccurrent extrait, comme l'illustre la figure 20. Ce mode d'affichage nous permet de constater que le cooccurrent verbal 拦截 (lán jié « intercepter ») s'utilise de façon disjointe avec 攻击 (gōng jī /nom « attaque informatique ») dans des contextes comme 拦截钓鱼攻击 (lán jié diào yú gōng jī « intercepter-hameçonnage-attaque » « intercepter une attaque par hameçonnage »). Chaque résultat est d'ailleurs assorti de liens vers d'autres outils tels que *Concordance* et *Multiword sketch*, ce qui permet à l'utilisateur d'obtenir davantage de renseignements sur le fonctionnement combinatoire d'un terme (figure 20). La fonction *Concordance* nous est d'une grande utilité puisqu'elle permet de localiser des CLS non contiguës, soit des contextes dans lesquels un terme (comme 攻击 (gōng jī /nom « attaque informatique ») et son cooccurrent (拦截 (lán jié « intercepter ») sont séparés par d'autres unités linguistiques (telles que 钓鱼 (diào yú « hameçonnage »).

WORD SKETCH Thèse

攻击 as noun 776x

Fonction syntaxique du terme par rapport aux cooccurrents

| Modifies | N_Modifier | Object_of | Subject_of | Possession |
|----------------------------|------------------------|------------------------------|-----------------------------|--------------------|
| 行为 的攻击行为 10.3 | APT APT 攻击中 10.9 | 受到 容易受到攻击 10.3 | 指 攻击是指攻击者 9.3 | 方法 进行攻击的方法 10.1 |
| 目的 攻击目的路由器 10.2 | 黑客 黑客攻击 10.5 | 溢出 缓冲区溢出攻击的 9.9 | 导致 攻击也可导致 9.2 | 方式 攻击的方式 9.1 |
| 手段 攻击手段 10.1 | DDoS 的 DDoS 攻击 10.5 | 发动 发动攻击的 9.6 | 攻击 攻击无论是攻击 9.2 | |
| 事件 攻击事件 10.1 | 钓鱼 钓鱼攻击等 10.2 | 注入 SQL 注入攻击的 9.6 | 发生 攻击发生在 9.1 | |
| 路径 攻击路径链表 9.9 | 盗号 天遭遇疑似盗号攻击事件约 9.3 | 受 受攻击的 9.4 | 可能 攻击可能 8.8 | |
| 目标 攻击目标的 9.7 | 中间人 攻击中间人攻击 9.3 | 实施 实施攻击 9.4 | 入侵 的攻击入侵检测的 8.6 | |
| 数量 的攻击数量 9.6 | 网络 网络攻击 9.2 | 发起 发起攻击的 9.3 | 来自 攻击来自 8.6 | |
| 方式 的攻击方式 9.6 | 入侵 入侵攻击事件 9.1 | 检测 检测到的攻击 9.3 | 影响 攻击严重影响 8.6 | |
| 攻击者 的攻击攻击者 9.3 | 撞库 撞库或撞库攻击关于带毒 9.0 | 遭受 遭受攻击时会 8.9 | 带来 攻击带来的 8.6 | |
| 图 基于攻击图 9.1 | 邮箱 的一种邮箱攻击 9.0 | 遭到 遭到攻击 8.9 | 涉及 攻击都涉及到 8.5 | |
| 手法 的黑客攻击手法 9.0 | 鱼叉 使用的是鱼叉攻击 8.9 | 拦截 拦截钓鱼攻击 8.6 | 会 攻击也不会 8.4 | |
| 方法 的攻击方法 8.9 | 漏洞 漏洞攻击 IP 的地域 8.9 | 防范 防范未经由防火墙的攻击 8.6 | 利用 攻击主要是利用 8.4 | |
| 源 攻击源路由器及端口 8.9 | 欺骗 欺骗攻击 8.9 | 进行 进行攻击 8.6 | 使用 中 79.2% 的攻击使用的是鱼叉 7.6 | |
| 场景 评估具体网络攻击场景的行为安全 8.8 | 服务 拒绝服务攻击 8.7 | 看 看 2016 年全球针对工业系统的攻击 8.4 | 包括 攻击包括 7.2 | |
| 溯源 检测与 DDoS 攻击溯源这两种 8.7 | 物理 物理攻击 8.7 | 挂 挂马攻击 8.3 | 到 攻击都涉及到 7.0 | |
| 代码 攻击代码 8.7 | 内部 内部攻击和外部 8.6 | 攻击 攻击该攻击 8.3 | 可以 这种攻击可以 7.0 | |

Typicality score

Text

Exemple d'usage

攻击 + 发生

攻击 + 发生

发生

Liens vers d'autres outils tels que « Concordance » et « Multiword Sketch »

Figure 20. – Extrait de la liste de cooccurrents de 攻击/nom (« attaque informatique ») générée par l'outil *Word Sketch*

L'outil *Word Sketch* propose des pistes très utiles afin de repérer des CLS valides, mais les résultats proposés doivent être validés. Nous avons dû résoudre plusieurs problèmes au cours de l'extraction des CLS. Premièrement, *Word Sketch* ne distingue pas les différentes acceptions d'un polysème. Pour un terme cherché, la liste des résultats générée peut inclure des cooccurrents d'un autre terme associé à la même forme lexicale. Ce problème est causé par des erreurs d'étiquetage du corpus. Par exemple, on trouve l'UT 主动 qui correspond à 主动/adj (« active ») et 主动/adv (« activement ») dans la liste des cooccurrents du terme 攻击/v (« attaquer »). L'UT 主动 peut former deux CLS avec l'UT 攻击 : 主动/adv 攻击/v (« attaquer de façon active ») et 主动/adj 攻击/nom (« attaque active »). Pour vérifier si le résultat généré par *Word Sketch* est le modificateur adverbial du verbe 攻击/v, nous avons dû analyser la fonction syntaxique du cooccurrent proposé dans ses contextes d'utilisation. Considérons la phrase suivante extraite de notre corpus :

fonctionnel que pour les noms et les verbes. Il est impossible d'extraire des cooccurents d'un terme étiqueté comme adjectif ou adverbe, tel que 恶意^{è y1}/_{adj} (« malveillant ») ou 恶意^{è y1}/_{adv} (« de manière malveillante »).

Pour pallier ces problèmes, nous avons eu recours à l'outil *Concordance* pour relever les cooccurents d'un terme à partir des contextes d'utilisation de celui-ci. Cet outil permet d'effectuer des recherches avancées à l'aide du *Corpus Query Language* (CQL), un langage de requête utilisé dans Sketch Engine. Le CQL sert à formuler des interrogations paramétrées, afin de repérer les concordances répondant à des critères de recherche complexes qui ne peuvent être définis dans l'interface utilisateur standard (Lexical Computing, 2020).

Cette fonction de recherche a été d'une grande utilité. D'abord, une requête formulée en CQL permet de localiser un terme relevant d'une partie du discours spécifique. La figure 21 illustre l'interface de recherche avancée de l'outil *Concordance*. Nous avons formulé une expression CQL pour rechercher la forme lexicale 加密^{jiā mì} étiquetée comme adjectif (« crypté ») dans notre corpus. Il est possible de saisir directement une requête dans le champ de recherche ou saisir les paramètres de recherche dans l'interface de *CQL Builder*. Ce dernier permet de générer automatiquement une expression CQL selon les paramètres définis par l'utilisateur (figure 21).

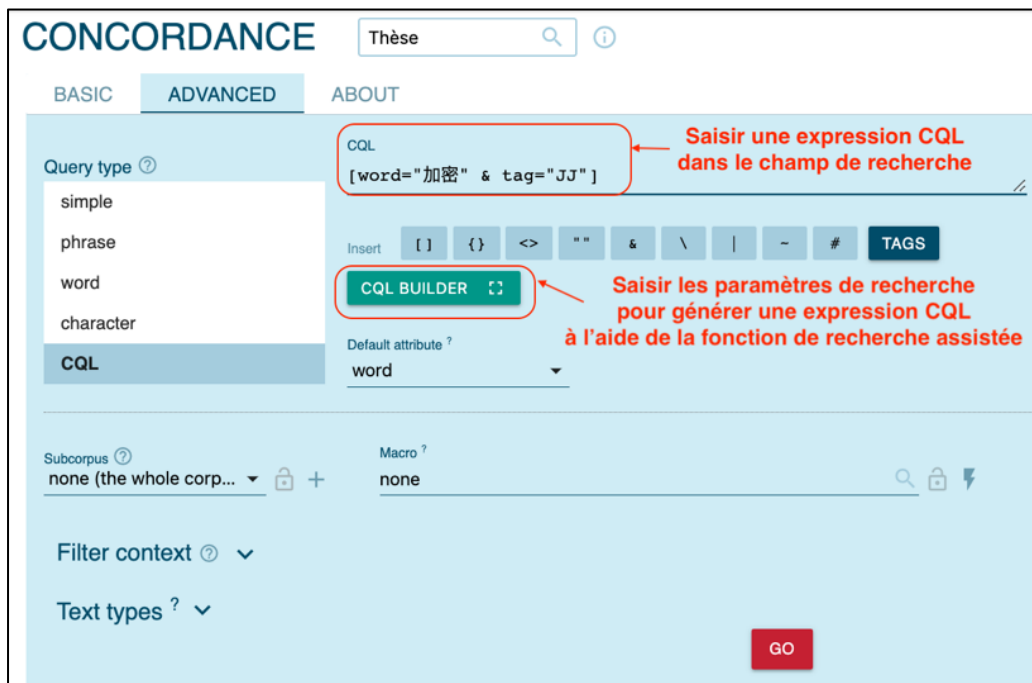


Figure 21. – Formulation d'une requête à l'aide du CQL dans l'interface de *Concordance*

Nous avons aussi formulé des expressions CQL complexes pour repérer les cooccurents d'un terme qui peuvent être insérés dans une configuration syntaxique spécifique. Comme l'illustre la figure 22, la requête saisie vise à localiser les contextes dans lesquels un cooccurent nominal de 加密^{jiǎo mì}/adj (« crypté ») apparaît à une distance maximum d'un mot du terme, ce qui nous permet de repérer les combinaisons lexicales nominales, contiguës ou disjointes, dans lesquelles 加密^{jiǎo mì}/adj (« crypté ») remplit la fonction syntaxique de modificateur adjectival, telles que 加密^{jiǎo mì}/adj 数据^{shù jù} (« données cryptées »). Nous avons autorisé une distance d'un mot entre 加密^{jiǎo mì}/adj (« crypté ») et son cooccurent puisque le modificateur et le nom modifié sont souvent reliés par la particule structurale 的^{de} en chinois (voir aussi la sous-section 4.1.2.3). La valeur de distance peut être modifiée selon la configuration syntaxique des CLS à extraire.



Figure 22. – Formulation d'une expression CQL dans l'interface de *CQL Builder*

La figure 23 illustre un extrait des résultats de la recherche ainsi formulée. Les concordances sont affichées en mode KWIC et triées par ordre alphabétique selon leur transcription phonétique en pinyin.

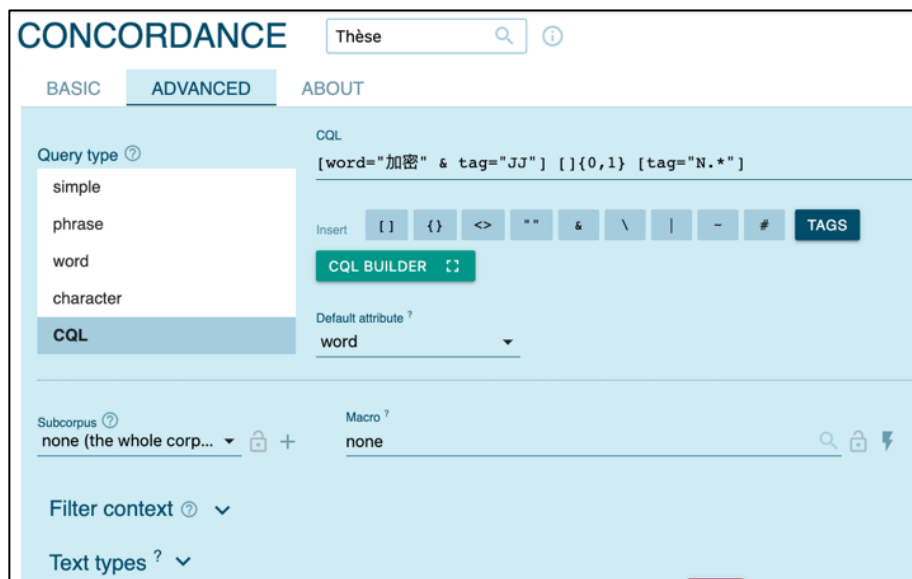


Figure 23. – Formulation d'une expression CQL complexe dans l'interface de *Concordance*

Il arrive que l’outil *Concordance*, tout comme *Word Sketch*, puisse générer des résultats non pertinents à cause d’erreurs d’étiquetage. On confond souvent le nom et l’adjectif puisque ces deux parties du discours peuvent modifier un nom dans une structure syntaxique « modificateur-nom modifié ». Reconsidérons la requête pour localiser les contextes contenant 加密 /adj (« crypté ») et le nom qu’il modifie. Nous avons remarqué que dans certaines combinaisons relevées, le modificateur 加密 étiqueté comme adjectif est en fait un terme nominal (« cryptage »), par exemple, 加密算法 (« algorithme de cryptage ») (figure 24). Les modificateurs nominaux (tels que 加密/nom « cryptage ») servent à expliquer la finalité du nom modifié (算法 « algorithme »), contrairement aux modificateurs adjectivaux (tels que 加密/adj « crypté ») qui s’emploient pour décrire les caractéristiques du nom modifié (数据 « données »).



Figure 24. – Extrait des résultats d’une recherche avancée dans l’interface de *Concordance*

À la fin de cette étape de collecte, nous avons relevé au total 644 CLS. Les résultats d’extraction ont été consignés dans un fichier Excel. Pour chacun des termes identifiés, les CLS collectées ont été classées selon leur structure syntaxique.

4.1.3.2 Extraction des contextes d’utilisation des CLS

Nous avons procédé à l’extraction des contextes d’utilisation des CLS à l’aide de l’outil *Concordance*. Cette démarche vise à sélectionner un nombre d’exemples d’usage représentatifs à partir desquels nous analysons les propriétés syntaxico-sémantiques des CLS repérées. Les

résultats d'analyse nous permettront d'anticiper les difficultés d'apprentissage soulevées par celles-ci.

Pour chaque CLS, nous avons effectué une recherche avancée au moyen du CQL. Il convient d'autoriser une certaine distance entre les deux éléments constitutifs d'une CLS afin de retrouver les CLS disjointes. Par exemple, nous avons formulé une requête pour repérer les exemples d'usage de la combinaison verbale ^{fǎng wèn zī yuán} 访问资源 (« accéder à une ressource »), comme l'illustre la figure 25.



Figure 25. – Requête formulée dans l'interface de *CQL Builder* pour extraire des contextes contenant la combinaison ^{fǎng wèn zī yuán} 访问资源 (« accéder à une ressource »)

Cette expression CQL permet de localiser les contextes dans lesquels le terme ^{zī yuán} 资源 « ressource » apparaît à une distance maximum de trois mots à droite du terme ^{fǎng wèn} 访问 « accéder à ». Les résultats de recherche (figure 26) montrent que la base (^{zī yuán} 资源 « ressource ») et son collocatif (^{fǎng wèn} 访问 « accéder à ») sont souvent séparés par d'autres éléments.

The screenshot shows a concordance tool interface. At the top, the search query is displayed: `CQL [word="访问"] [0..3] [word="资源"] + 36`. Below the query, there is a table with columns: **Details**, **Left context**, **KWIC**, and **Right context**. The table contains 22 rows of search results, each with a document ID (e.g., doc#2, doc#7) and corresponding text snippets. The KWIC column highlights the search terms in red. The right context column shows the surrounding text.

Figure 26. – Extrait des contextes d'utilisation de la CLS ^{fǎng wèn zī yuán} 访问资源 (« accéder à une ressource »)

Le terme 资源 (^{zī yuán} « ressource ») est parfois précédé d'un ou plusieurs modificateurs. Cette fonction syntaxique peut être remplie par un adjectif (par exemple, 本地^{běn dì} « local », 网络^{wǎng luò} « en ligne »), un nom (系统^{xì tǒng} « système »), un syntagme nominal (移动互联网^{yí dòng hù liánwǎng} « Internet mobile ») ou un syntagme adjectival qui est relié au terme au moyen de la particule structurale 的^{de} (par exemple, 有限的^{yǒu xiàn de} 'limité-DE particule structurale' « qui est limitée », 其他计算机上的^{qí tā jì suàn jī shàng de} 'autre-ordinateur-sur-DE particule structurale' « qui se trouve sur d'autres ordinateurs »). Le terme de base 资源 (^{zī yuán} « ressource ») se combine parfois avec plusieurs collocatifs verbaux coordonnés, par exemple, 访问或删除^{fǎng wèn huò shān chú} ('accéder à-ou-supprimer' « accéder à une ressource ou la supprimer »). Le tableau 17 présente quelques exemples d'usage de la CLS 访问资源 (^{fǎng wèn zī yuán} « accéder à une ressource ») extraits de notre corpus. Pour chaque CLS sélectionnée, nous avons extrait entre 5 et 10 contextes. Toutes les données terminologiques collectées à partir de notre corpus ont été analysées et compilées lors de l'élaboration de la ressource CHINOINFO.

| Exemple | Collocatif | | Modificateur | Base | Équivalent en français |
|-----------|-------------|---|---|-------------------|--|
| Exemple 1 | 访问 | | | 资源 | accéder à une ressource / accéder à des ressources |
| | 'accéder à' | | | 'ressource' | |
| Exemple 2 | 访问 | | Modificateur = adjectif | Base | Équivalent en français |
| | 'accéder à' | | [本地] 'local' | 资源 'ressource' | |
| Exemple 3 | 访问 | | Modificateur = adjectif | Base | Équivalent en français |
| | 'accéder à' | | [网络] 'en ligne' | 资源 'ressource' | |
| Exemple 4 | 访问 | | Modificateur = syntagme nominal | Base | Équivalent en français |
| | 'accéder à' | | [移动互联网] 'mobile-Internet' | 资源 'ressource' | |
| Exemple 5 | 访问 | | Modificateur = syntagme adjectival | Base | Équivalent en français |
| | 'accéder à' | | [其他计算机上][的] 'autre-ordinateur-sur"particule structurale' | 资源 'ressource' | |
| Exemple 6 | 访问 | Modificateur 1 = syntagme adjectival | Modificateur 2 = nom | Base | Équivalent en français |
| | 'accéder à' | [有限][的] 'limité'-'particule structurale' | [计算机] 'ordinateur' | 资源 'ressource' | |
| Exemple 7 | 访问 | Collocatif 2 = conjonction de coordination | Modificateur = nom | Base | Équivalent en français |
| | 'accéder à' | [或][删除] 'ou'-'supprimer' | [系统] 'système' | 资源 'ressource' | |

Tableau 17. – Analyse d'exemples d'usage de 访问资源 (^{fǎng wèn zī yuán} « accéder à une ressource ») extraits du corpus

4.1.3.3 Synthèse des renseignements recueillis

À la suite du dépouillement du corpus, nous avons analysé les CLS assorties de leurs contextes d'utilisation au sein de chacun des 33 vocables relevés. Cette démarche vise à examiner les propriétés syntaxico-sémantiques des CLS des termes appartenant à un même polysème. Ces CLS peuvent être regroupées en quatre types de « paires » selon leur forme, leur structure syntaxique et la fonction syntaxique du polysème. Nous les expliquerons, dans ce qui suit, à l'aide d'exemples. Nous proposons en outre une analyse détaillée de ces cas de figure dans l'annexe 7.

Il importe de rappeler que, dans notre thèse, les termes *polysème* et *vocable* sont définis façon différente que ce qui est proposé par la LEC, pour rendre compte du phénomène de mot multicatégoriel que l'on observe couramment en chinois. Le *polysème* fait référence à tous les sens associés à une même forme lexicale qui peuvent relever de parties de discours distinctes. Le *vocable* regroupe un ensemble de termes ayant une forme lexicale identique, ces derniers peuvent être associés à une seule partie du discours ou à plusieurs parties de discours (voir aussi les sous-sections 4.1.2.3 et 4.1.3.1).

1) Deux CLS ayant des formes distinctes et une même structure syntaxique. Les acceptions d'un même vocable peuvent former des CLS avec des unités lexicales distinctes. Ces CLS sont associées à la même structure syntaxique. Par exemple, le vocable ^{yí dòng}移动 a deux acceptions, soit ^{yí dòng}移动₁ (« mobile ») et ^{yí dòng}移动₂ (« amovible »). Ces acceptions peuvent se retrouver dans les CLS ^{yí dòng yīngyòng}移动₁应用₁ (« application mobile ») et ^{yí dòng yìngpán}移动₂硬盘 (« disque dur amovible »). Ces deux CLS ont la même structure syntaxique, dans lesquelles ^{yí dòng}移动₁ (« mobile ») et ^{yí dòng}移动₂ (« amovible ») remplissent la fonction syntaxique de modificateur, comme l'illustre le tableau 18. De la même manière, le vocable ^{jiǎn cè}检测 regroupe deux acceptions : ^{jiǎn cè}检测₁ (« détecter ») et ^{jiǎn cè}检测₂ (« mettre à l'essai »). Du point de vue syntaxique, ils exigent deux COD différents pour former deux CLS verbales de type « Prédicat-COD » (tableau 18). Du point de vue sémantique, ^{jiǎn cè}检测₁ (« détecter ») signifie « déceler la présence d'une attaque informatique visant à accéder à des données et, éventuellement, à nuire au fonctionnement du matériel ou à l'intégrité des données ». Ce terme fait appel à deux actants : l'*agent* qui effectue l'action et le *patient* qui subit l'action, réalisé par des termes comme ^{gōng jī}攻击₂ (« attaque informatique ») ou ^{bìng dú}病毒 (« virus »). ^{jiǎn cè}检测₂ (« mettre à l'essai ») a aussi deux actants sémantiques et exprime le sens de tester, à l'aide d'un programme approprié, l'état d'un logiciel,

d'un réseau ou d'un système informatique afin d'en évaluer le niveau de sécurité et d'en déterminer les vulnérabilités. Le patient de 检测₂ (« mettre à l'essai ») peut être réalisé par 软件₁ (« logiciel₁ »). Dans les CLS 检测₁攻击₂ (« détecter une attaque informatique ») et 检测₂软件₁ (« mettre un logiciel à l'essai »), le patient de 检测₁ (« détecter ») et celui de 检测₂ (« mettre à l'essai ») sont réalisés syntaxiquement comme COD.

| Regroupement des CLS contenant un polysème | Forme des CLS | Structure syntaxique des CLS | Exemples | Structure syntaxique des CLS | CLS | Équivalent de la CLS | Acceptation du polysème | Partie du discours | Équivalent de l'acceptation | Fonction syntaxique du polysème |
|---|---------------|------------------------------|-----------|------------------------------|--------|-----------------------------------|-------------------------|--------------------|---|---------------------------------|
| 2 CLS ⇒ 2 formes ⇒ 1 structure syntaxique | Différente | Identique | Exemple 1 | [Modificateur] + Nom modifié | 移动1应用1 | application mobile | 移动1 | adj | mobile | Modificateur |
| | | | | | 移动2硬盘 | disque dur amovible | 移动2 | adj | amovible | |
| | | | Exemple 2 | Modificateur + [Nom modifié] | 原生应用1 | application native | 应用1 | nom | application1 (= logiciel d'application) | Nom modifié |
| | | | | | 商业应用2 | application commerciale | 应用2 | nom | application2 (= application de l'informatique) | |
| | | | Exemple 3 | [Prédicat] + COD | 检测1攻击2 | détecter une attaque informatique | 检测1 | verbe | détecter | Prédicat |
| | | | | | 检测2软件1 | mettre un logiciel à l'essai | 检测2 | verbe | mettre à l'essai | |

Tableau 18. – Regroupement des CLS renfermant un polysème dont la structure syntaxique est la même mais ayant des formes différentes

2) Deux CLS ayant des formes distinctes et des structures syntaxiques distinctes. Nous avons observé, au sein des vocables identifiés, des paires de CLS qui ont deux formes différentes et deux structures syntaxiques distinctes (tableau 19). Le vocable 测试, à titre d'exemple, est un mot multicatégoriel (voir la sous-section 4.1.2.4) qui regroupe le verbe 测试₁ (« tester ») et son dérivé sémantique 测试₂ (« test »). Les deux termes se combinent avec une même unité lexicale, soit 软件₁ (« logiciel₁ »), pour former deux CLS associées à deux structures syntaxiques distinctes. 测试₁软件₁ (« tester un logiciel₁ ») est une CLS verbale, dans laquelle le cooccurrent 软件₁ (« logiciel₁ ») est le COD du terme 测试₁ (« tester »), tandis que dans la CLS 软件₁测试₂ (« test logiciel »), le même cooccurrent 软件₁ (« logiciel₁ ») sert à modifier le terme 测试₂ (« test »). Il s'agit d'un modificateur nominal qui équivaut à l'adjectif relationnel *logiciel* en français. Dans un vocable, un verbe (comme 传输₁ « transférer ») et son dérivé sémantique 传输₂ « transfert ») peuvent s'associer à des termes nominaux dans des CLS de structures syntaxiques différentes. Par exemple, 传输₁ est le prédicat verbal dans la CLS 传输₁数据 (« transférer des données »). Dans 传输₂介质 (« support de transmission de données »), 传输₂ (« transfert ») est un modificateur nominal du terme 介质 (« support »).

| Regroupement des CLS contenant un polysème | Forme des CLS | Structure syntaxique des CLS | Exemples | Structure syntaxique des CLS | CLS | Équivalent de la CLS | Acception du polysème | Partie du discours | Équivalent de l'acception | Fonction syntaxique du polysème |
|---|---------------|------------------------------|-----------|------------------------------|--------|--------------------------------------|-----------------------|--------------------|---------------------------|---------------------------------|
| 2 CLS ⇒ 2 formes ⇒ 2 structures syntaxiques | Différente | Différente | Exemple 1 | [Prédicat] + COD | 测试1软件1 | tester un logiciel | 测试1 | verbe | tester | Prédicat |
| | | | | Modificateur + [Nom modifié] | 软件1测试2 | test logiciel | 测试2 | nom | test | Nom modifié |
| | | | Exemple 2 | [Prédicat] + COD | 传输1数据 | transférer des données | 传输1 | verbe | transférer | Prédicat |
| | | | | [Modificateur] + Nom modifié | 传输2介质 | support de transmission (de données) | 传输2 | nom | transfert | Modificateur |

Tableau 19. – Regroupement des CLS renfermant un polysème qui ont deux formes distinctes et deux structures syntaxiques différentes

3) Deux CLS ayant une forme identique et une structure syntaxique identique. Des paires de CLS sont polysémiques dans le cas où elles partagent une même forme. Les acceptions d'un vocable peuvent se combiner avec une même unité lexicale dans une même configuration syntaxique. Examinons les deux exemples illustrés au tableau 20. 网络₁ (« réseau ») et 网络₃ (« en ligne ») modifient un même terme 攻击 (« attaque informatique ») dans deux CLS. 网络₁ 攻击 (« attaque envers un réseau ») revoie aux actions visant à endommager, à forcer ou à détourner un réseau pour commettre un acte préjudiciable. 网络₁ (« réseau ») est un modificateur nominal qui sert à indiquer la cible d'une attaque informatique. 网络₃ 攻击 (« cyberattaque ») désigne les actions malveillantes qui ciblent différents dispositifs informatiques par l'intermédiaire d'un réseau. L'adjectif 网络₃ (« en ligne ») remplit également la fonction syntaxique de modificateur et précise l'intermédiaire pour réaliser une attaque informatique. Il en va de même pour le vocable 设置 qui regroupe le verbe 设置₁ (« configurer ») et ses dérivés nominaux 设置₂ (« configuration-nom d'action ») et 设置₃ (« configuration-nom de résultat »). Les deux dérivés nominaux peuvent se combiner avec le même modificateur 系统 (« système ») : la CLS 系统设置₂ (« configuration d'un système ») ainsi formée exprime l'activité de définir les paramètres d'un système afin d'en assurer le bon fonctionnement, tandis que 系统设置₃ (« configuration d'un système ») fait référence à l'ensemble des paramètres associés à un système pour en assurer le bon fonctionnement.

| Regroupement des CLS contenant un polysème | Forme des CLS | Structure syntaxique des CLS | Exemples | Structure syntaxique des CLS | CLS | Équivalent de la CLS | Acception du polysème | Partie du discours | Équivalent de l'acception | Fonction syntaxique du polysème |
|--|---------------|------------------------------|-----------|------------------------------|-------|----------------------------|-----------------------|--------------------|---------------------------------|---------------------------------|
| 2 CLS ⇒ 1 forme ⇒ 1 structure syntaxique | Identique | Identique | Exemple 1 | [Modificateur] + Nom modifié | 网络1攻击 | attaque envers un réseau | 网络1 | nom | réseau | Modificateur |
| | | | | | 网络3攻击 | cyberattaque | 网络3 | adj | en ligne | |
| | | | Exemple 2 | Modificateur + [Nom modifié] | 系统设置2 | configuration d'un système | 设置2 | nom | configuration (nom d'action) | Nom modifié |
| | | | | | 系统设置3 | configuration d'un système | 设置3 | nom | configuration (nom de résultat) | |

Tableau 20. – Regroupement des CLS renfermant un polysème qui ont une même forme et une même structure syntaxique

4) Deux CLS ayant une même forme et deux structures syntaxiques distinctes. Les CLS ayant une forme identique et des sens différents contiennent souvent un mot multicatégoriel, à savoir un polysème qui regroupe les termes relevant de différentes parties du discours, comme l'illustre le tableau 21. Observons le vocable 加密 qui a trois acceptions : 加密_{1/v} (« crypter »), 加密_{2/nom} (« cryptage ») et 加密_{3/adj} (« crypté ») (voir aussi la sous-section 4.1.2.2). La première acception 加密₁ (« crypter ») et son dérivé sémantique adjectival 加密₃ (« crypté ») se combinent avec le même terme 数据 (« données ») et forment deux CLS de structure syntaxique différente : 加密₁数据 (« crypter des données ») est une CLS composée d'un prédicat (加密₁ « crypter ») associé à son COD 数据 (« données »), 加密₃数据 (« données cryptées ») est une CLS nominale dans laquelle 加密₃ (« crypté ») est le modificateur adjectival de 数据 (« données »). Du point de vue sémantique, 加密₁ (« crypter ») et 加密₃ (« crypté ») possèdent un actant sémantique en commun : 加密₁ (« crypter ») dénote le fait de transformer une information en un cryptogramme, tandis que 加密₃ (« crypté ») dénote l'état d'une entité qui subit l'action de 'crypter'. Ces deux acceptions font appel à un patient qui peut être réalisé par le terme 数据 (« données »).

| Regroupement des CLS contenant un polysème | Forme des CLS | Structure syntaxique des CLS | Exemples | Structure syntaxique des CLS | CLS | Équivalent de la CLS | Acception du polysème | Partie du discours | Équivalent de l'acception | Fonction syntaxique du polysème |
|--|---------------|------------------------------|-----------|-------------------------------------|--------|-----------------------------|-----------------------|--------------------|---------------------------|---------------------------------|
| 2 CLS ⇒ 1 forme ⇒ 2 structures syntaxiques | Identique | Différente | Exemple 1 | [Prédicat] + COD | 加密1数据 | crypter des données | 加密1 | verbe | crypter | Prédicat |
| | | | | [Modificateur] + Nom modifié | 加密3数据 | données cryptées | 加密3 | adj | crypté | Modificateur |
| | | | Exemple 2 | [Modificateur] + Nom modifié | 自动1下载2 | téléchargement automatique | 自动1 | adj | automatique | Modificateur |
| | | | | | | | 下载2 | nom | téléchargement | Nom modifié |
| | | | | [Modificateur adverbial] + Prédicat | 自动2下载1 | télécharger automatiquement | 自动2 | adv | automatiquement | Modificateur adverbial |
| | | | | | | | 下载1 | verbe | télécharger | Prédicat |

Tableau 21. – Regroupement des CLS renfermant un polysème qui ont une même forme et deux structures syntaxiques distinctes

Considérons le second exemple présenté au tableau 21. Il s'agit d'une paire de CLS ayant une forme identique et des sens différents. Chacune de ces CLS est composée de deux éléments polysémiques : le premier élément ^{zì dòng}自动 est un polysème qui regroupe l'adjectif ^{zì dòng}自动₁ (« automatique ») et son dérivé adverbial ^{zì dòng}自动₂ (« automatiquement »), le deuxième élément ^{xià zài}下载 regroupe le verbe ^{xià zài}下载₁ (« télécharger ») et son dérivé sémantique ^{xià zài}下载₂ (« téléchargement »). Ces éléments forment deux CLS associées à une forme identique et à deux structures syntaxiques distinctes, soit ^{zì dòng xià zài}自动₁下载₂ (« téléchargement automatique ») et ^{zì dòng xià zài}自动₂下载₁ (« télécharger automatiquement »). La distinction sémantique de ces CLS qui sont syntaxiquement ambiguës nécessite une analyse ancrée dans le contexte. Les deux phrases suivantes extraites de notre corpus, à titre d'exemples, nous ont permis d'appréhender le sens des CLS en question et identifier la structure syntaxique de celles-ci :

1) ^{zhǔ jī}主机 ^{zài}在 ^{liánwǎng}联网 ^{qíngkuàng}情况 ^{xià}下 ^{jiāng}[将] ^{zì dòng}自动₂ ^{xià zài}下载₁
Hôte à être connecté à réseau circonstance sous aller⁵⁸ automatiquement installer
^{bìng}[并] ^{ān zhuāng}安装₁ ^{xì tǒng}系统 ^{gēng xīn}更新₃。
et installer système mise à jour.

« Lorsque l'hôte est connecté à Internet, il téléchargera et installera automatiquement les mises à jour du système. » [Notre traduction]

⁵⁸ 将 (« aller ») est un verbe auxiliaire qui sert à exprimer une action à venir. r

| | | | | | |
|------------------------------------|--------------------------------------|------------------------|------------------------------------|------------------------------------|------------------------|
| 2) 开启 ^{kāi qǐ} | WLAN | 环境 ^{huánjìng} | 自动 ^{zì dòng} ₁ | 下载 ^{xià zài} ₂ | 和 ^{hé} |
| Activer | réseau local sans fil | environnement | automatique | téléchargement | et |
| 夜间 ^{yè jiān} | 安装 ^{ān zhuāng} ₂ | 开关 ^{kāi guān} | 平板 ^{píng bǎn} | 将在 ^{jiāng zài} | 凌晨 ^{língchén} |
| pendant la nuit | installation | bouton | tablette | aller à | début de matin |
| 自动 ^{zì dòng} ₂ | 安装 ^{ān zhuāng} ₁ | 升级 ^{shēng jí} | 包 ^{bāo} | 2 h | à 4 h |
| automatiquement | installer | mise à niveau | paquet | | |

« Activez le téléchargement automatique et l’installation pendant la nuit dans un environnement de réseau local sans fil, la tablette installera automatiquement le progiciel de mise à niveau entre 2 h et 4 h du matin. » [Notre traduction]

Dans la première phrase, nous avons repéré plusieurs indices qui permettent de confirmer que la CLS 自动下载^{zì dòng xià zài} est un syntagme verbal (« télécharger automatiquement »), dans lequel le verbe 下载^{xià zài}₁ (« télécharger ») est modifié par l’adverbe 自动^{zì dòng}₂ (« automatiquement »). D’abord, cette CLS est précédée de l’adverbe 将^{jiāng} qui exprime une action à réaliser dans le futur. Ensuite, nous avons identifié le sujet et le COD du verbe 下载^{xià zài}₁ (« télécharger »), soit le terme 主机^{zhǔ jī} (« hôte ») et la CLS nominale 系统更新^{xì tǒng gēng xīn} (« mise à jour du système »). Du point de vue sémantique, ils réalisent deux actants sémantiques de 下载^{xià zài}₁ (« télécharger ») : 主机^{zhǔ jī} (« hôte ») réalise l’agent qui effectue l’action de ‘télécharger’, tandis que 系统更新^{xì tǒng gēng xīn} (« mise à jour du système ») réalise le patient qui subit l’action dénotée. Enfin, nous avons identifié la conjonction de coordination 并 (« et ») qui sert à relier deux verbes, soit 下载^{xià zài}₁ (« télécharger ») et 安装^{ān zhuāng}₁ (« installer »). Dans la deuxième phrase, la CLS 自动下载^{zì dòng xià zài} est nominale (« téléchargement automatique ») puisqu’elle remplit la fonction syntaxique de COD du verbe 开启^{kāi qǐ} (« activer »). Dans cette CLS, le terme 下载^{xià zài}₂ (« téléchargement ») est modifié par l’adjectif 自动^{zì dòng}₁ (« automatique »).

En résumé, nous avons analysé, à cette étape, les renseignements recueillis pour chaque terme dans la nomenclature du CHINOINFO et les CLS dont le terme fait partie sous les angles formel, syntaxique et sémantique. Les constats dressés nous ont permis d’anticiper les difficultés d’apprentissage des CLS auxquelles font face les locuteurs non sinophones.

4.1.3.4 Anticipation des difficultés syntaxico-sémantiques soulevées par les CLS

L’analyse des propriétés syntaxico-sémantiques des CLS des termes appartenant à un même polysème nous a amenée à explorer davantage les difficultés qui sont susceptibles d’être soulevées par les CLS.

1) Distinguer les acceptions d'un polysème dans différentes CLS. Cette difficulté correspond aux trois premiers constats dressés à la suite de notre analyse des propriétés syntaxico-sémantiques des CLS (voir la sous-section 4.1.3.3). Une paire de CLS de formes différentes peut renfermer deux acceptions d'un polysème. Par exemple, on identifie deux sens différents associés à la forme lexicale 移动 dans les CLS 移动₁互联网 (« Internet mobile ») et 移动₂硬盘 (« disque dur amovible ») (voir aussi la sous-section 4.1.2.4). Les apprenants auront à distinguer ces acceptions pour pouvoir appréhender le sens des CLS qui les contiennent et identifier la structure syntaxique des CLS. La distinction sémantique est davantage difficile quand un polysème se combine avec les termes d'une même partie du discours pour former des CLS de différentes structures syntaxiques. Les CLS 传输₁数据 (« transférer des données ») et 传输₂介质 (« support de transmission de données »), par exemple, sont associées à différentes structures syntaxiques. La première CLS est une combinaison de type « prédicat verbal-COD » dans laquelle 传输₁ (« transférer ») remplit la fonction syntaxique de prédicat verbal. La deuxième CLS est un syntagme nominal dans lequel 传输₂ (« transfert ») modifie le terme nominal 介质 (« support »).

2) Différencier les sens distincts de CLS de forme identique dans un contexte donné. Cette difficulté correspond au dernier constat qui se dégage de notre analyse des propriétés syntaxico-sémantiques des CLS (voir la sous-section 4.1.3.3). Elle découle du fait que deux CLS de même forme sont associées à des structures syntaxiques distinctes et véhiculent ainsi des sens différents. Les CLS 备份₁数据 (« sauvegarder des données ») et 备份₂数据 (« données sauvegardées ») illustrent bien la confusion engendrée par la polysémie des CLS : la première CLS est un syntagme verbal composé d'un prédicat verbal (备份₁ « sauvegarder ») et du COD du verbe, la seconde CLS est un syntagme nominal, dans lequel 备份₂ (« sauvegardé ») remplit la fonction syntaxique de modificateur. La distinction sémantique est susceptible de se compliquer davantage pour les apprenants dans le cas où une paire des CLS renferment deux polysèmes. Par exemple, dans les CLS 自动₁下载₂ (« téléchargement automatique ») et 自动₂下载₁ (« télécharger automatiquement »), les deux éléments composants appartiennent respectivement à un polysème. Les apprenants auront à appréhender le sens de ces CLS dans le contexte d'utilisation de celles-ci, comme nous l'avons expliqué dans la sous-section 4.1.3.3.

3) Choisir les cooccurents appropriés d'un terme. Ce défi à relever est lié à plusieurs difficultés d'apprentissage que nous avons observées dans la pratique d'enseignement auprès d'étudiants francophones. En premier lieu, les apprenants confondent parfois les acceptions d'un polysème et les termes qui y sont associés par un lien sémantique. Par exemple, nous avons identifié, dans la sous-section 4.1.2.3, deux acceptions verbales de 检测 (检测₁ « détecter » et 检测₂ « mettre à l'essai ») qui forment les CLS 检测₁病毒 (« détecter un virus ») et 检测₂软件₁ (« mettre un logiciel à l'essai »). Seule la deuxième acception 检测₂ (« mettre à l'essai ») est reliée par une relation de synonymie au terme 测试₁ (« tester »). Autrement dit, 检测₂ (« mettre à l'essai ») et 测试₁ (« tester ») sont interchangeables dans les CLS qui les contiennent. Les apprenants pourraient confondre les deux acceptions verbales de 检测 et établir à tort une relation de synonymie entre 检测₁ (« détecter ») et 测试₁ (« tester »), ce qui les amènerait à formuler l'expression *测试₁病毒 (« tester un virus ») qui est inacceptable du point de vue sémantique, comme le montre la figure 27.

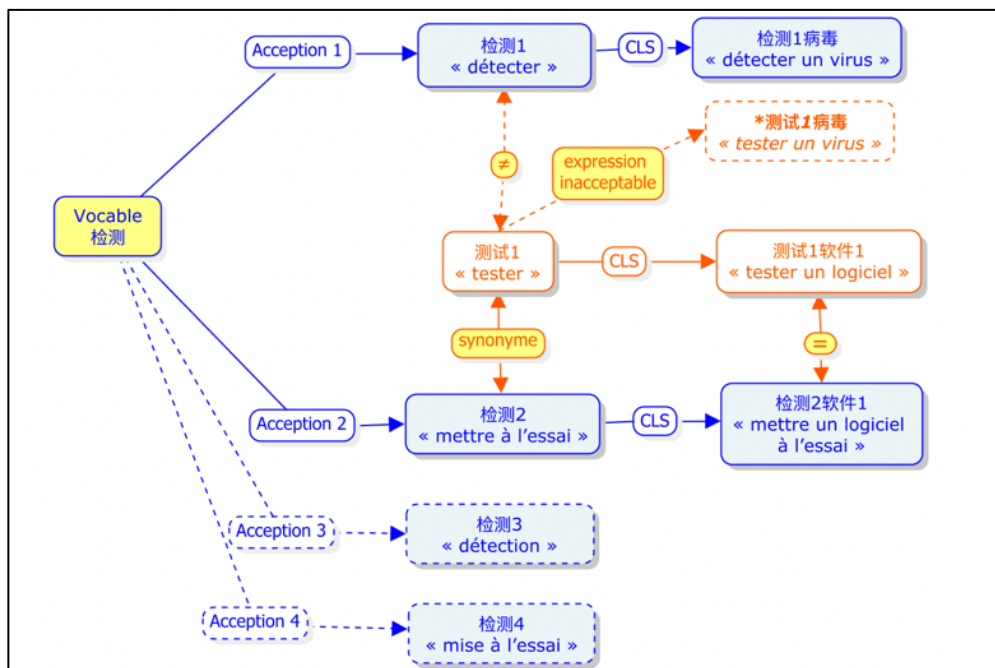


Figure 27. – Lien sémantique entre 测试₁ (« tester ») et les deux premières acceptions du vocable 检测 (« détecter » et « mettre à l'essai »)

En second lieu, on peut prévoir que les apprenants auront de la difficulté à distinguer les termes ayant une composante sémantique commune. Par exemple, 传输₁ (« transférer ») et 迁移₁

(« faire migrer ») sont sémantiquement proches : les deux termes expriment l'action de 'faire parvenir des données d'un emplacement de mémoire vers un autre'. Le deuxième actant sémantique de ces termes peut être réalisé par des termes dénotant une entité de données, tels que 数据 (« données ») et 文件 (« fichier »). Par contre, 迁移₁ (« faire migrer ») décrit l'action de 'déplacer un ensemble de données et d'effectuer les adaptations nécessaires pour assurer le fonctionnement adéquat des données relocalisées'. Certains termes typiques qui réalisent le patient de 迁移₁ (« faire migrer »), tels que 系统 (« système »), 数据库 (« base de données ») et 磁盘 (« disque magnétique »), ne peuvent pas se combiner avec 传输₁ (« transférer »), comme le montre le tableau 22.

| Terme | CLS | |
|-------------------------------------|---|---|
| | [Fonction syntaxique] = Prédicat [Rôle sémantique] = Prédicat sémantique | [Fonction syntaxique] = COD [Rôle sémantique] = Patient |
| 传输 ₁ « transférer » | 传输 « transférer » | 数据 « données » 文件 « fichier » * 数据库 « base de données » * 系统 « système » * 磁盘 « disque magnétique » |
| 迁移 ₁ « faire migrer » | 迁移 « faire migrer » | 数据 « données » 文件 « fichier » 数据库 « base de données » 系统 « système » 磁盘 « disque magnétique » |

Tableau 22. – Comparaison des CLS de type « prédicat-COD » contenant 传输₁ « transférer » et 迁移₁ « faire migrer »

Enfin, les apprenants sont souvent confrontés à la confusion engendrée par l'équivalence partielle interlinguistique. Autrement dit, il est probable qu'ils éprouvent des difficultés à distinguer les termes qui peuvent être traduits par un même mot dans leur langue première. Par exemple, les termes 传输₁ (« transférer ») et 转发 (« retransmettre ») sont souvent associés à un même équivalent français, soit « transférer », malgré leur nuance sémantique. 传输₁ (« transférer ») signifie « acheminer des données d'un emplacement de mémoire vers un autre » et, 转发 (« retransmettre ») exprime l'action de 'réacheminer un courrier électronique ou un message reçu

par voie électronique vers un autre destinataire’. Leur structure actancielle se présente comme suit (tableau 23).

| Actant 1 | Actant 3 | | Prédicat | Actant 2 | Actant 4 | |
|---------------|-----------|---------------------------|----------------------|--------------------------|--------------|---------------------------|
| Agent | Source | | | Patient | Destination | |
| yòng hù 用户 | cóng 从 | cún chǔ wèi zhì 存储位置 | chuán shū 传输 | shù jù 数据 | dào 到 | cún chǔ wèi zhì 存储位置 |
| utilisateur | de | emplacement de mémoire | transférer | données | à | emplacement de mémoire |
| Actant 1 | | | Prédicat | Actant 2 | Actant 3 | |
| Agent | | | | Patient | Destinataire | |
| yòng hù 用户 | | | zhuǎn fā 转发 | diàn zǐ yóu jiàn 电子邮件 | gěi 给 | shōu jiàn rén 收件人 |
| utilisateur | | | retransmettre | courrier électronique | à | destinataire |

Tableau 23. – Représentation de la structure actancielle de 传输₁ (« transférer ») et 转发 (« retransmettre »)

Comme on peut le constater, ces verbes font appel à un patient qui est réalisé par des termes distincts : 传输₁ (« transférer ») exige un patient qui renvoie à une entité de données, telle que 数据 (« données »), tandis que l’action de 转发 (« retransmettre ») s’applique à une entité reçue par voie électronique, telle que 电子邮件 (« courrier électronique »). Les apprenants confondraient les cooccurrents de ces termes et formuleraient ainsi des CLS inappropriées du point de vue sémantique, telles que * 传输₁ 电子邮件 (« transférer un courrier électronique ») au sens de « retransmettre un courrier électronique ».

Dans le cas où les deux termes se combinent avec un même terme, tel que 文件 (« fichier »), on observe les nuances sémantiques entre les CLS formées 传输₁ 文件 (« transférer un fichier ») fait référence au déplacement d’un fichier qui se trouve dans un emplacement de mémoire, tandis que 转发文件 (« retransmettre un fichier ») implique que le fichier en question est reçu par voie électronique et s’adresse à un destinataire spécifique.

En résumé, nous avons prévu trois types de difficulté d’apprentissage en nous appuyant sur l’analyse des propriétés syntaxico-sémantique des CLS extraits de notre corpus. Dans la section suivante, nous présenterons les méthodes adoptées pour aider les apprenants non natifs à surmonter ces obstacles.

4.1.4 Synthèse

La section 4.1 expose la méthode que nous avons adoptée pour extraire et analyser un échantillon de CLS à partir d'un corpus. La première étape de notre démarche consiste à compiler un corpus selon un ensemble de critères (4.1.1). Nous avons ensuite exploité le corpus mis en forme à l'aide de Sketch Engine, afin d'extraire un nombre de termes qui soulèvent des difficultés syntaxico-sémantiques pour les apprenants non natifs (4.1.2). À l'issue de cette étape, nous avons repéré 91 termes regroupés en 33 vocables. Nous avons ensuite extrait 644 CLS contenant ces termes et analysé les propriétés syntaxico-sémantiques des CLS relevées (4.1.3). Les données terminologiques collectées à partir de notre corpus ont été consignées dans un fichier de travail. Elles seront compilées en vue d'élaboration du dictionnaire CHINOINFO. Nous aborderons l'encodage des données terminologiques dans la section 4.2.2.

4.2 Concevoir une méthode descriptive des CLS en vue de l'élaboration d'une ressource lexicale

La deuxième phase de notre recherche comporte deux volets : dans le premier, nous esquissons quelques pistes de solutions aux difficultés d'apprentissage anticipées (4.2.1); dans le deuxième, une méthode intégrant les solutions proposées est mise en œuvre pour élaborer le dictionnaire CHINOINFO, une ressource didactique de la combinatoire spécialisée (4.2.2).

4.2.1 Mise au point de la méthode de description

Nous commençons par mobiliser différentes stratégies pour représenter les propriétés syntaxico-sémantiques des CLS. Ces pistes de solutions s'inspirent, comme on le verra, du DiCoInfo (L'Homme, 2008; L'Homme et Alipour, 2009; L'Homme, 2009; Jousse et al., 2011; L'Homme et al., 2012; L'Homme et Jia, 2015) et des travaux dont il s'est inspiré (Jousse et al., 2008; Jousse, 2010). Nous nous sommes appuyée sur ces travaux pour envisager une description systématique de la combinatoire spécialisée en vue de son enseignement.

4.2.1.1 Difficulté 1 : distinguer les acceptions d'un polysème dans différentes CLS

Inspirée de la LEC, nous avons regroupé en vocable les termes ayant une forme identique et des sens apparentés (voir aussi la sous-section 4.1.2.4) lors de notre analyse des UT polysémiques

(l'annexe 6). Il importe de rappeler l'écart entre la LEC et notre travail quant à l'utilisation du terme *vocable*. Selon la LEC, un vocable est associé à une seule partie du discours. Dans notre travail, un vocable peut contenir les termes relevant de parties du discours distinctes. Chaque terme correspond à une acception d'un vocable et fait l'objet d'un article distinct dans le CHINOINFO. Pour aider les apprenants à distinguer les acceptions d'un vocable dans les CLS qui les contiennent, nous avons recours à deux stratégies pour présenter les connaissances lexicales au sein d'un article : la description des relations sémantiques au moyen d'un système de gloses et de définitions, ainsi que la classification hiérarchique des CLS.

D'abord, nous nous sommes appuyée sur le système des FL pour analyser les relations sémantiques des termes décrits. En premier lieu, nous avons décrit les liens entre les acceptions d'un vocable. Dans le vocable ^{shè zhì} 设置 (voir aussi la sous-section 4.1.3.3), par exemple, nous avons observé que les trois acceptions entretiennent des liens sémantiques qui peuvent être exprimés par des FL paradigmatiques et syntagmatiques, comme l'illustre le tableau 24. D'une part, les liens de dérivation sémantique entre ces trois acceptions sont explicités à l'aide des gloses *famille de mots-X*, dans lesquelles on précise la partie du discours X auquel appartient le terme relié. Pour les dérivés sémantiques nominaux, nous avons distingué le nom d'action du nom de résultat. Les gloses utilisées pour encoder les liens sémantiques expliquent les trois FL paradigmatiques (verbalisation [V₀], nom d'action [S₀] et nom de résultat [S_{res}]) (Mel'čuk et al., 1995, p. 133) correspondantes. D'autre part, les liens qui relient le nom de résultat ^{shè zhì} 设置₃ aux deux autres acceptions sont modélisés au moyen des FL syntagmatiques **Caus₁Func₀** et **S₀Caus₁Func₀**⁵⁹. Dans le CHINOINFO, ces liens sémantiques sont glosés de la manière suivante : **Caus₁Func₀** = faire en sorte que *N* (le terme décrit) existe; **S₀Caus₁Func₀** = l'action de faire en sorte que *N* (le terme décrit) existe.

⁵⁹ Dans notre travail, les liens sémantiques sont encodés au moyen des gloses et non pas des FL. Ce choix méthodologique se justifie par deux raisons. Premièrement, le formalisme d'encodage traditionnel des FL peut être difficile d'accès pour les apprenants qui n'ont aucune connaissance de la LEC. Deuxièmement, tous les liens sémantiques ne sont pas encodés par les FL lorsqu'il est question de décrire les propriétés sémantiques des CLS. Dans notre thèse, les CLS renvoient à deux types de regroupements lexicaux : 1) les collocations restreintes contrôlées par un terme et 2) les liens observés dans les structures actanciennes. Les FL ne rendent compte que des relations sémantiques entre la base et le collocatif du premier type des CLS.

| FL | Lien sémantique | Encodage de lien sémantique |
|--|--|--|
| V ₀ | V ₀ (^{shèzhì} 设置 ₂ « configuration-nom d'action ») = ^{shèzhì} 设置 ₁ « configurer » | Famille de mots-verbe |
| S ₀ | S ₀ (^{shèzhì} 设置 ₁ « configurer ») = ^{shèzhì} 设置 ₂ « configuration-nom d'action » | Famille de mots-nom d'action |
| S _{res} | S _{res} (^{shèzhì} 设置 ₁ « configurer ») = ^{shèzhì} 设置 ₃ « configuration-nom de résultat » S _{res} (^{shèzhì} 设置 ₂ « configuration-nom d'action ») = ^{shèzhì} 设置 ₃ « configuration-nom de résultat » | Famille de mots-nom de résultat |
| Caus ₁ Func ₀ | Caus ₁ Func ₀ (^{shèzhì} 设置 ₃ « configuration-nom de résultat ») = ^{shèzhì} 设置 ₁ « configurer » | Faire en sorte qu'une configuration existe |
| S ₀ Caus ₁ Func ₀ | S ₀ Caus ₁ Func ₀ (^{shèzhì} 设置 ₃ « configuration-nom de résultat ») = ^{shèzhì} 设置 ₂ « configuration-nom d'action » | L'action de faire en sorte qu'une configuration existe |

Tableau 24. – Représentation des liens sémantiques entre les trois acceptions du vocable ^{shè zhì}设置

Cette piste de solution est utile pour la distinction des acceptions d'un mot multicatégoriel (sous-section 4.1.2.4) puisqu'elle met en évidence les liens sémantiques entre celles-ci. Appliquée à l'enseignement des CLS, elle vise à aider les apprenants à différencier les CLS dans lesquelles les acceptions d'un polysème se combinent avec une même unité lexicale, par exemple, la CLS ^{xì tǒng}系统设置₂ (« configuration d'un système ») renvoie à 'la définition des paramètres d'un système afin d'en assurer le bon fonctionnement', tandis que ^{xì tǒng shè zhì}系统设置₃ (« configuration d'un système ») fait référence à 'l'ensemble des paramètres qui résulte de l'action de configurer un système'.

La représentation des liens paradigmatiques permet également de distinguer les acceptions reliées par une relation de méronymie au sein d'un vocable. Appliquée à la description des deux acceptions du vocable ^{ruǎn jiàn}软件, cette solution aide les apprenants à distinguer si la forme lexicale désigne un seul logiciel (^{ruǎn jiàn}软件₁ « logiciel₁ ») ou un ensemble de ressources (^{ruǎn jiàn}软件₂ « logiciel₂ ») dans les CLS qui la contiennent, par exemple, ^{ruǎn jiàn xià zài}软件₁下载 (« téléchargement d'un logiciel ») et ^{ruǎn jiàn zī yuán}软件₂资源 (« ressource logicielle ») (sous-section 4.1.2.4).

En deuxième lieu, les liens paradigmatiques qui relient un terme décrit à d'autres termes du domaine ont été mis en évidence. Le vocable ^{yìng yòng}应用, par exemple, regroupe deux acceptions nominales : ^{yìng yòng}应用₁ (« application₁ ») est le synonyme de ^{yìng yòng ruǎn jiàn}应用₁软件₁ (« logiciel d'application »). ^{yìng yòng}应用₂ (« application₂ ») signifie 'l'utilisation de moyens informatiques pour répondre à un besoin déterminé et faciliter les processus de travail des utilisateurs' (Office québécois de la langue française, 2005). Il est probable que les apprenants aient de la difficulté à distinguer les sens

associés à 应用^{yìngyòng} dans les CLS 移动^{yí dòng}应用^{yìngyòng}₁ (« application mobile ») et 商业^{shāng yè}应用^{yìngyòng}₂ (« application commerciale »), d'autant plus que les deux CLS ont la même structure syntaxique. Pour résoudre cette difficulté, nous proposons de présenter les termes qui sont respectivement reliés à ces deux acceptions par un lien paradigmatique. 应用^{yìngyòng}₁ (« application₁ »), qui désigne une sorte spécifique de logiciel, entretient une relation d'hyponymie avec le terme 软件^{ruǎn jiàn}₁ (« logiciel₁ »). 应用^{yìngyòng}₂ (« application₂ ») est l'hyponyme de 使用^{shǐ yòng}₂ (« utilisation »). Cette solution vise à sensibiliser les apprenants aux réseaux lexicaux dans lesquels s'insèrent respectivement 应用^{yìngyòng}₁ (« application₁ ») et 应用^{yìngyòng}₂ (« application₂ »). Autrement dit, la comparaison des termes reliés à ces deux acceptions fournit des indices pour différencier le sens de celles-ci.

Quant à la description des CLS au sein d'un article, nous nous sommes abondamment inspirée de la méthodologie utilisée pour confectionner le DiCoInfo (L'Homme, 2002, 2008; L'Homme et Alipour, 2009; L'Homme, 2009; Jousse et al., 2011; L'Homme et al., 2012; L'Homme et Jia, 2015), notamment un ordonnancement des collocations s'inspirant de la LEC (Jousse et al., 2008; Jousse, 2010). Dans notre travail, nous avons adapté cette méthode à la description des CLS. Rappelons que dans notre travail, les CLS d'un terme incluent deux types de regroupements lexicaux : 1) les collocations contrôlées par le terme vedette, 2) la combinaison du terme vedette à une autre lexie réalisant son ASyntP (Mel'čuk et al., 1995, p. 117). Il s'agit d'une collocation contrôlée par l'ASyntP (voir aussi la sous-section 1.1.3).

Premièrement, nous avons encodé, pour chaque collocation contrôlée par un terme vedette, le lien syntagmatique entre la base et le collocatif au moyen d'une glose, comme nous le faisons pour un lien paradigmatique. Par exemple, dans la CLS 非法^{fēi fǎ}入侵^{rù qīn} (« 'illicite-intrusion' « intrusion illicite »), le lien syntagmatique qui associe le terme de base 入侵^{rù qīn}₂ (« intrusion ») à son collocatif 非法^{fēi fǎ} (« illicite ») est paraphrasé comme suit : « qui est faite sans autorisation ». Cette glose explique en fait la relation lexicale exprimée par la FL complexe **AntiBonOper**₁⁶⁰. Par ailleurs, dans la CLS 入侵^{rù qīn}₂防御^{fáng yù} (« 'intrusion-prévention' « prévention d'intrusion »), 入侵^{rù qīn}₂ (« intrusion ») est relié à son collocatif 防御^{fáng yù} « prévention » par la FL **S₀Caus@NonFunc**₂. La relation sémantique ainsi

⁶⁰ La FL **AntiBon** exprime le sens du 'contraire de laudatif'. Elle se combine avec le verbe support **Oper**₁ pour exprimer le sens d'être fait d'une manière non laudative (voir aussi l'explication des FL complexes dans la sous-section 3.1.1).

modélisée est expliquée par la glose suivante : « action réalisée par quelqu’un ou quelque chose qui fait en sorte qu’une intrusion ne se produise pas (sur un ordinateur) ».

La représentation des liens syntagmatiques offre l’avantage de permettre aux apprenants de comparer les CLS répertoriées dans différents articles au sein d’un vocable. Reconsidérons le vocable ^{gēng xīn}更新 qui comprend deux acceptions nominales, soit le nom d’action ^{gēng xīn}更新₂ (« mise à jour₁ ») et le nom de résultat ^{gēng xīn}更新₃ (« mise à jour₂ »), comme nous l’avons expliqué dans la sous-section 4.1.2.3. Les apprenants pourraient confondre ces deux acceptions puisqu’elles ont la même fonction syntaxique (COD) dans les CLS qu’elles forment avec un verbe. Du point de vue sémantique, ces termes sont reliés à leurs collocatifs verbaux par des liens lexicaux différents. Le lien entre ^{gēng xīn}更新₂ (« mise à jour₁-nom d’action ») et le verbe ^{jìn xíng}进行 (« effectuer »), comme nous l’avons mentionné plus haut, est exprimé par la FL simple **Oper₁**. La relation entre ^{gēng xīn}更新₃ (« mise à jour₂- nom de résultat ») et ses collocatifs verbaux, tels que ^{ān zhuāng}安装₁ (« installer ») et ^{qǐ yòng}启用 (« activer ») est modélisée à l’aide de FL complexes⁶¹. Par exemple, la FL complexe **Caus₁Able₁Real₁** qui relie ^{gēng xīn}更新₃ (« mise à jour₂- nom de résultat ») à ^{qǐ yòng}启用 (« activer ») exprime le sens de ‘faire en sorte que (le mot-clé) puisse être utilisée’.

La mise en parallèle du comportement combinatoire des acceptions d’un même vocable, comme l’illustre le tableau 25, fournit des indices aux apprenants pour les distinguer.

| Terme en entrée | Collocatif verbal | Représentation de lien sémantique au moyen de FL | Glose | CLS |
|--|---|--|--|---|
| ^{gēng xīn} 更新 ₂ « mise à jour ₁ » nom d’action | ^{jìn xíng} 进行 « effectuer » | Oper ₁ (^{gēng xīn} 更新 ₂) = 进行 | L'utilisateur fait une mise à jour | ^{jìn xíng} 进行更新 ₂ « effectuer une mise à jour » |
| ^{gēng xīn} 更新 ₃ « mise à jour ₂ » nom de résultat | ^{ān zhuāng} 安装 ₁ « installer » | Prepar ₁ Fact ₀ [phase I](^{gēng xīn} 更新 ₃) = 安装 ₁ | L'utilisateur prépare une mise à jour pour permettre à la mise à jour de fonctionner | ^{ān zhuāng} 安装 ₁ 更新 ₃ « installer une mise à jour » |
| | ^{qǐ yòng} 启用 « activer » | Caus ₁ Able ₁ Real ₁ (^{gēng xīn} 更新 ₃) = 启用 | L'utilisateur fait en sorte qu'une mise à jour puisse être utilisée | ^{qǐ yòng} 启用更新 ₃ « activer une mise à jour » |

Tableau 25. – Représentation des liens sémantiques au sein des CLS verbales contenant ^{gēng xīn}更新₂ (« mise à jour₁ ») et ^{gēng xīn}更新₃ (« mise à jour₂ »),

⁶¹ voir aussi l’explication des FL complexes dans la sous-section 3.1.1

La deuxième stratégie que nous mobilisons rend compte des CLS qui résultent de l'association d'un terme à l'un de ses ASyntP. Plus précisément, nous envisageons une classification hiérarchique des CLS au sein d'un article du CHINOINFO, afin d'explicitier la structure syntaxique des CLS et de mettre en évidence les ASyntP du terme vedette (voir la définition de dans la sous-section 3.1.1).

Cette solution est inspirée de la structuration des données terminologiques dans le DiCoInfo, qui s'inspire quant à elle de la méthode proposée par Jousse (2010) pour organiser les relations lexicales. Cette méthode est appliquée à l'élaboration de la base de données du DiCo, où les FL sont organisées en fonction de leurs propriétés combinatoires (Jousse, 2010, p. 192). Autrement dit, la description des collocations rend compte non seulement de leurs propriétés sémantiques (c'est-à-dire la relation syntagmatique entre la base et le collocatif représentée par une FL), mais aussi de leur configuration syntaxique. Ce mode d'organisation des collocations permet de localiser le collocatif qui conviendrait le mieux pour être inséré dans une configuration syntaxique, comme le souligne Jousse (2010, p. 192). Appliquée à la création des ressources lexicales pédagogiques, cette méthode présente un intérêt pédagogique dans une perspective d'aide à la rédaction.

Dans le DiCoInfo, les FL syntagmatiques sont utilisées non seulement pour modéliser les relations lexicales, mais aussi pour organiser les collocations en différentes classes hiérarchiques (Jousse et al., 2011). Par exemple, toutes les FL syntagmatiques qui expriment l'action d'utiliser sont regroupées dans différentes *classes intermédiaires*. Considérons un exemple concret : dans l'article *antivirus* du DiCoInfo (Observatoire de linguistique Sens-Texte, 2013a), la FL **IncepReal₁** relie le mot-clé à trois collocatifs verbaux, à savoir *charger*, *appeler* et *lancer* qui expriment le sens de 'commencer à utiliser' un antivirus. Cette FL, tout comme d'autres FL syntagmatiques qui expriment un sens similaire, telles que **IncepReal₂** reliant *système informatique* à son collocatif *accéder (à)*, est regroupée dans la classe intermédiaire « Commencer à utiliser/Apparaître ». Celle-ci, ainsi que d'autres classes intermédiaires telles que « Permettre l'utilisation », « Préparer l'utilisation » ou « Cesser d'utiliser/de faire fonctionner » sont regroupées dans la *classe générique* « UTILISER/NE PAS UTILISER », comme le montre le tableau 26 (Jousse et al., 2011, p. 138). La définition des classes repose sur les relations sémantiques observées dans le domaine de l'informatique, comme le font remarquer

Jousse et al. (2011, p. 137-139). Ce classement hiérarchique des collocations vise à offrir un accès facile et rapide aux collocations exprimant un sens spécifique, dans le but d'aider l'utilisateur dans la production ou la traduction de textes spécialisés (L'Homme, 2009; L'Homme et Leroyer, 2009).

| Lexical function | Intermediate class | Generic class |
|---|---|-----------------------------------|
| Caus1Able1Fact0, Caus1Able1Real1, Caus1Able1Real3, PermFact0, Perm1Fact0 | Permettre l'utilisation / Activer | UTILISER / NE PAS UTILISER |
| Prepar@, Prepar1, Prepar1Fact0, Prepar1Real1, Prepar2Real3, Prepar@Fact0, De_nouveauPrepar1, De_nouveauPrepar1Fact0 | Préparer l'utilisation / le fonctionnement | |
| IncepLabreal12, IncepReal@, IncepReal1, IncepReal2, IncepReal3 | Commencer à utiliser / Apparaître | |
| Caus1Fact0, Caus@Fact0, Labreal@2, Labreal12, Labreal123, QLabreal12, Real@, Real1, Real12, Real123 | Utiliser / Faire fonctionner | |
| FinLabreal12, FinReal1, FinReal2, Liqu1Fact0, Liqu@Fact0 | Cesser d'utiliser / de faire fonctionner | |

Tableau 26. – Regroupement des FL en classes intermédiaires et génériques (Jousse et al., 2011, p. 138)

Dans le CHINOINFO, nous nous alignons sur les travaux précédents pour envisager une organisation hiérarchique des CLS. À la différence du DiCoInfo qui utilise un modèle ascendant (du particulier au général) pour structurer des collocations en classes sémantiques, nous avons opté pour un modèle descendant lors de l'élaboration des classes des CLS. Plus précisément, toutes les CLS sont d'abord classées en 13 classes génériques en fonction de leur structure syntaxique, comme l'illustre le tableau 27. Le nom des classes génériques indique la structure syntaxique des CLS répertoriées. La dénomination de certaines classes génériques rend compte des propriétés sémantiques des CLS, c'est le cas des classes de CLS composées d'un prédicat et d'un complément de direction, de manière ou de résultat. La fonction syntaxique du terme vedette est mise entre crochets.

| Classes génériques | Exemple | Équivalent français |
|--|--|--|
| [Modificateur] – Nom modifié | yí dòng wǎngluò [移动 ₁] 网络 ₁ | Réseau [mobile] |
| [Modificateur] – Nom modifié apposé | yìngyòng ruǎnjiàn [应用 ₁] 软件 | Logiciel d'[application] = application |
| Modificateur – [Nom modifié] | zì dòng gǎngxīn 自动 ₁ [更新 ₂] | [Mise à jour] automatique |
| Modificateur adverbial – [Prédicat] | zì dòng gǎngxīn 自动 ₂ [更新 ₁] | [Mettre à jour] automatiquement |
| [Modificateur adverbial] – Prédicat | yuǎnchéng fǎngwèn [远 程 ₂] 访问 ₁ | Accéder (à...) [à distance] |
| Prédicat – [Complément d'objet] | cè shì ruǎnjiàn [测试 ₁] 软件 | [Tester] un logiciel |
| [Prédicat] – Complément d'objet | yùnxíng chéngxù [运行 ₂] 程序 | [Faire exécuter] un programme |
| [Prédicat] – Complément de direction | xiàzài dào [下载 ₁] 到... | [Télécharger] vers... |
| [Prédicat] – Complément de manière | yùnxíng huǎnmàn [运行 ₁] 缓慢 | S'exécuter lentement |
| [Prédicat] – Complément de résultat | jiǎn cè dào [检测 ₁] 到... | Réussir à [détecter]... |
| [Sujet] – Prédicat | chuán shū zhōngduàn [传输 ₂] 中断 | [(Un) transfert] s'interrompt |
| Structure coordonnée = [Prédicat1] – Prédicat2 | fù zhì zhāntiē [复制 ₂] 粘 贴 ₁ | [Copier]-coller ₁ (verbe) |
| Structure coordonnée = [Nom1] – Nom2 | fù zhì zhāntiē [复制 ₄] 粘 贴 ₂ | [Copier]-coller ₂ (nom) |

Tableau 27. – Liste de classes génériques du CHINOINFO

Les CLS répertoriées dans une même classe générique ont ensuite été analysées de façon plus fine pour être organisées en classes intermédiaires. L'élaboration de classes intermédiaires rend compte, d'une part, des liens syntagmatiques entre la base et le collocatif d'une collocation, comme ce qui a été proposé dans le DiCoInfo, et, d'autre part, de la structure actancielle du terme décrit, dans le cas où le terme se combine avec une lexie réalisant son ASyntP.

Observons la description des CLS renfermant le terme 同步₂ (« synchronisation ») (tableau 28). Les CLS sont structurées dans trois classes génériques. Dans la première classe générique « Modificateur – [Nom modifié] » se trouve la classe intermédiaire « Sortes de ». Celle-ci regroupe trois CLS à base nominale qui contiennent 同步₂ (« synchronisation ») et un modificateur. Les deux premières CLS (数据同步 « synchronisation de données » et 文件同步 « synchronisation de fichier ») résultent en fait de l'association du terme décrit à son ASyntP : les modificateurs nominaux 数据 (« données ») et 文件 (« fichier ») désignent une entité qui subit une 'synchronisation' et réalisent ainsi le deuxième actant de 同步₂ (« synchronisation »). Dans la

troisième CLS, le terme décrit est modifié par l'adjectif 实时^{shí shí}₁ (« en temps réel ») qui exprime le sens de 'qui se fait en réduisant le temps de réponse au minimum'. La classe intermédiaire « Instrument » qui est organisée dans la classe générique « [Modificateur]-Nom modifié », répertorie une CLS dans laquelle 同步^{tóng bù}₂ (« synchronisation ») modifie 软件^{ruǎn jiàn}₁ (« logiciel ») qui désigne un instrument que l'on utilise pour réaliser une 'synchronisation'.

| Terme en entrée : 同步 ^{tóng bù} ₂ (« synchronisation ») | | | | |
|--|----------------------|---|---------------------------------|-------------------------------|
| Classe générique | Classe intermédiaire | Glose | CLS | Équivalent français |
| Modificateur- <u>nom modifié</u> | Sortes de | Qui porte sur un ensemble de données | 数据同步 ₂ | synchronisation de données |
| | | Qui porte sur un fichier [= entité structurée contenant des données ayant un format spécifique qu'un utilisateur crée en lui attribuant ce format et place sur un support de stockage] [DICOFICH] | 文件同步 ₂ | synchronisation de fichier |
| | | Qui se fait en réduisant le temps de réponse au minimum | 实时 ₁ 同步 ₂ | synchronisation en temps réel |
| <u>Modificateur</u> -nom modifié | Instrument | Logiciel servant à effectuer une synchronisation | 同步 ₂ 软件 ₁ | logiciel de synchronisation |
| <u>Prédictat</u> - <u>Complément</u> <u>d'objet</u> | Faire/Avoir lieu | L'utilisateur fait une synchronisation | 进行同步 ₂ | effectuer une synchronisation |
| | Arrêter de faire | L'utilisateur met fin à une synchronisation provisoirement | 暂停同步 ₂ | suspendre une synchronisation |

Tableau 28. – Regroupement des CLS contenant 同步^{tóng bù}₂ (« synchronisation ») en classes génériques et intermédiaires

Comme notre méthode de description lexicale est appliquée au domaine de l'informatique, nous avons emprunté un nombre de classes intermédiaires du DiCoInfo, étant donné que celles-ci sont définies en fonction des relations sémantiques observées dans le même domaine de spécialité (Jousse et al., 2011, p. 137-139), y compris « Utiliser/Faire fonctionner », « Commencer à utiliser », « Préparer l'utilisation/le fonctionnement », « Permettre l'utilisation/Activer », « Empêcher l'utilisation », « Cesser d'utiliser/de fonctionner », « Créer/Supprimer », « Mettre quelque part » et « Transformer ». Le nom de ces classes indique le sens des collocatifs associés au terme décrit. D'autres classes intermédiaires sont nommées pour mettre en évidence l'ASyntP du terme décrit, c'est le cas de « installer un objet spécifique » dans l'article 安装^{ān zhuāng}₁ (« installer »). Nous présentons dans l'annexe 23 la liste des classes élaborées pour encoder les CLS dans le CHINOINFO.

Chaque CLS est assortie d'une glose qui sert à expliciter le lien sémantique entre le terme décrit et son cooccurrent ou d'une définition de la CLS. Certaines gloses se rapprochent d'une ébauche

de définition lexicologique du cooccurrent du terme décrit, comme le fait remarquer Polguère (2003). Par exemple, pour la CLS 文件同步^{wén jiàntóng bù} (« synchronisation de fichier »), nous expliquons, dans la glose attachée à cette CLS, qu'il s'agit d'une synchronisation « (qui) porte sur un fichier ». De plus, nous proposons, dans la même glose, une définition de *fichier*, pour expliquer l'entité dénotée par ce terme (tableau 28), ce qui permet à l'utilisateur d'accéder aux connaissances spécialisées véhiculées par une CLS. De cette manière, nous adaptons la description des CLS aux besoins du public visé par notre recherche, soit les apprenants débutants du COS qui ne disposent que de connaissances élémentaires du domaine de l'informatique.

La classification hiérarchique des CLS, comme on peut le constater dans la figure 28, rend compte de trois types de propriétés syntaxico-sémantiques de CLS, à savoir leur structure syntaxique, le lien syntagmatique entre la base et le collocatif dans une collocation et la structure actancielle du terme vedette.

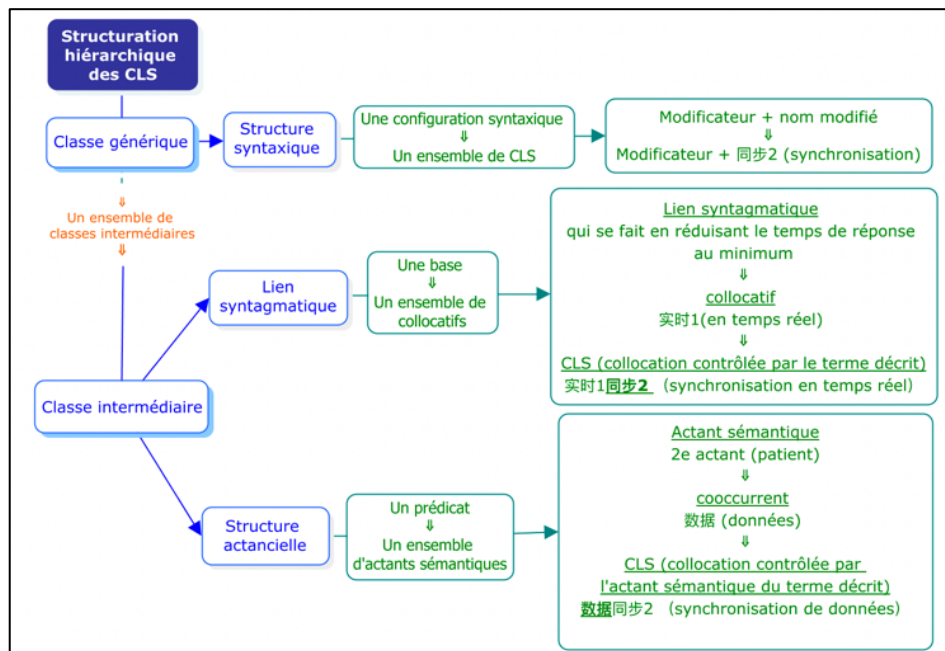


Figure 28. – Structuration hiérarchique des CLS rattachées à 同步₂ (« synchronisation ») dans le CHINOINFO

Cette solution est susceptible de sensibiliser les apprenants à la distinction des CLS contenant un mot multicatégoriel. Comme nous l'avons expliqué dans la sous-section 4.1.3.4, les acceptions d'un mot multicatégoriel forment souvent des CLS avec les termes d'une même partie du discours. 传输₁^{chuán shū} (« transférer ») et 传输₂^{chuán shū} (« transfert »), par exemple, se combinent respectivement

avec un nom dans les CLS ^{chuán shū shù jù} 传输₁数据 (« transférer des données ») et ^{chuán shū jiè zhì} 传输₂介质 (« support de transmission de données »). Pour les apprenants qui confondent les deux acceptions du polysème, ces CLS pourraient sembler associées à la même structure syntaxique.

Pour dissiper les confusions éventuelles qui pourraient entraver l'apprentissage des apprenants non natifs, ces deux CLS sont répertoriés dans différentes classes génériques et intermédiaires dans les articles. ^{chuán shū shù jù} 传输₁数据 (« transférer des données ») est répertoriée dans la classe générique « [Prédicat]-Complément d'objet ». Comme ^{shù jù} 数据 (« données ») réalise le deuxième actant du verbe ^{chuán shū} 传输₁ (« transférer »), nous avons répertorié cette CLS dans la classe intermédiaire « Transférer un objet spécifique ». Celle-ci regroupe deux autres CLS dans lesquelles le complément d'objet de ^{chuán shū} 传输₁ (« transférer ») désigne une entité qui subit l'action de 'transférer' : ^{chuán shū wén jiàn} 传输₁文件 (« transférer un fichier ») et ^{chuán shū xìn xī} 传输₁信息 (« transférer de l'information »). Le nom de cette classe intermédiaire indique le sens général des CLS répertoriées. Quant à ^{chuán shū jiè zhì} 传输₂介质 (« support de transmission de données »), cette CLS est d'abord organisée dans la classe générique « [Modificateur] – Nom modifié ». Les CLS regroupées dans cette classe sont subdivisées en deux classes intermédiaires « Moyen » et « Fonctionnement » en fonction de leur sens. La CLS ^{chuán shū jiè zhì} 传输₂介质 (« support de transmission de données ») est répertoriée dans la classe intermédiaire « Moyen » puisque le terme ^{jiè zhì} 介质 (« support ») désignant le moyen d'un 'transfert'. Cette manière de structurer les CLS vise à sensibiliser les apprenants, d'une part, aux configurations syntaxiques dans lesquelles peuvent s'insérer ^{chuán shū} 传输₁ (« transférer ») et ^{chuán shū} 传输₂ (« transfert »), d'autre part, au rôle sémantique des cooccurents par rapport au terme décrit.

4.2.1.2 Difficulté 2 : différencier les sens distincts de CLS de forme identique

Dans un deuxième temps, nous avons exploré les moyens de description qui pourraient aider les apprenants à distinguer les sens distincts de CLS qui ont la même forme. Comme nous l'avons expliqué dans la sous-section 4.1.3.4, cette difficulté découle du fait qu'une combinaison lexicale renferme au moins un polysème et est associée à des structures syntaxiques différentes. Pour résoudre cette difficulté, nous optons pour deux stratégies descriptives : la mise en évidence de la structure syntaxique des CLS et l'accès aux CLS de forme identique via chacun des polysèmes qu'elles renferment.

En premier lieu, nous proposons d'expliciter la structure syntaxique des CLS au moyen des classes génériques (voir aussi la sous-section 4.2.1.1). Cette piste de solution, sous-tendue par l'approche cognitive en didactique des L2 (voir aussi la section 3.3), vise à présenter de façon systématique les propriétés syntaxiques des CLS. Examinons les deux acceptions du vocable ^{ān}安^{zhuāng}装 : le verbe ^{ān}安^{zhuāng}装₁ (« installer ») et son dérivé sémantique nominal ^{ān}安^{zhuāng}装₂ (« installation ») se combinent tous les deux avec ^{chéng xù}程序 (« programme »), donnant ainsi lieu à deux CLS de sens différents, soit la CLS verbale ^{ān}安^{zhuāng}装₁ ^{chéng xù}程序 (« installer un programme ») et la CLS de nature nominale ^{ān}安^{zhuāng}装₂ ^{chéng xù}程序 (« programme d'installation »). Comme on peut le constater au tableau 29, la comparaison des renseignements descriptifs de ces deux CLS fait ressortir leurs différences sémantique et structurelle : ^{ān}安^{zhuāng}装₁ ^{chéng xù}程序 (« installer un programme ») est un syntagme verbal, dans lequel le terme en entrée ^{ān}安^{zhuāng}装₁ (« installer ») réalise le prédicat verbal de son COD ^{chéng xù}程序 (« programme »). ^{ān}安^{zhuāng}装₂ ^{chéng xù}程序 (« programme d'installation ») est une CLS à base nominale, dans laquelle ^{ān}安^{zhuāng}装₂ (« installation ») est un modificateur qui explique la fonctionnalité d'un logiciel.

| Terme | Classe générique | Classe intermédiaire | Glose | CLS | Équivalent français |
|--|---|----------------------------------|--|---|--|
| 安 ^{zhuāng} 装 ₁ installer | Prédicat- Complément d'objet | Installer un objet spécifique | Programme [= suite d'instructions écrites sous une forme que l'ordinateur peut comprendre pour traiter un problème ou pour effectuer une tâche] [GDTPROG] | 安 ^{zhuāng} 装 ₁ ^{chéng xù} 程序 | installer un programme |
| 安 ^{zhuāng} 装 ₂ installation | Modificateur -nom modifié | Instrument | Programme conçu pour installer correctement un logiciel sur le disque dur d'un ordinateur. [ADINST] | 安 ^{zhuāng} 装 ₂ ^{chéng xù} 程序 | programme d'installation = installeur |

Tableau 29. – Mise en parallèle de la description des CLS ^{ān}安^{zhuāng}装₁ ^{chéng xù}程序 (« installer un programme ») et ^{ān}安^{zhuāng}装₂ ^{chéng xù}程序 (« programme d'installation »).

La distinction du sens des CLS risque de constituer un défi de taille pour les apprenants dans le cas où deux CLS de même forme ont des sens différents. Par exemple, les CLS ^{zì dòng}自动₁ ^{xià zài}下载₂ (« téléchargement automatique ») et ^{zì dòng}自动₂ ^{xià zài}下载₁ (« télécharger automatiquement ») renferment respectivement deux composants polysémiques. Nous proposons de rendre ces CLS accessibles via chaque élément polysémique lorsque le moment est venu de créer la ressource CHINOINFO. La première CLS est présentée dans les articles consacrés à ^{zì dòng}自动₁ (« automatique ») et à ^{xià zài}下载₂ (« téléchargement »). La deuxième CLS est accessible via les articles consacrés à ^{zì dòng}自动₂ (« automatiquement ») et ^{xià zài}下载₁ (« télécharger »). Dans chacun de ces quatre articles, nous explicitons les liens sémantiques qui associent les acceptions de ces deux polysèmes, comme

l'illustre la figure 29. Lorsque les apprenants consultent l'une de ces CLS dans l'article consacré à la base de cette CLS, un lien les dirige vers le collocatif qui relève d'un autre vocable. Cette piste de solution offre un accès rapide à une paire de CLS ayant des sens différents et une forme identique. Nous reviendrons sur la création des renvois lorsque nous aborderons l'encodage des données terminologiques dans la sous-section 4.2.2.

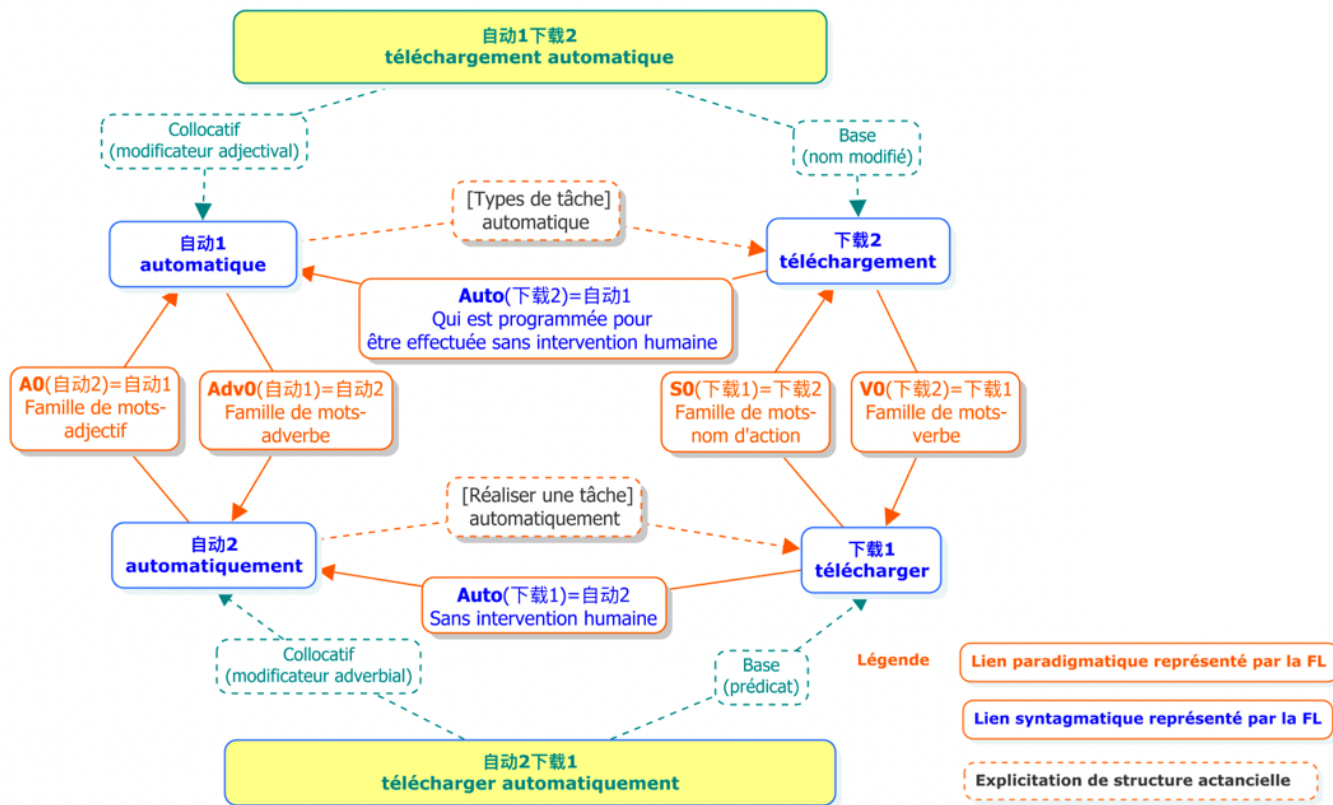


Figure 29. – Liens sémantiques entre les éléments composants des CLS ^{zì dòng xià zǎi}自动₁下载₂ (« téléchargement automatique ») et ^{zì dòng xià zǎi}自动₂下载₁ (« télécharger automatiquement »)

4.2.1.3 Difficulté 3 : choisir les cooccurrents appropriés d'un terme

La troisième difficulté relève du maniement des CLS en vue de productions écrites. Le choix de cooccurrents représente un défi important pour les apprenants non natifs. Nous avons abordé, entre autres, trois difficultés d'apprentissage liées à cette tâche : 1) distinguer les liens lexicaux de différentes acceptions d'un polysème; 2) distinguer les termes ayant une composante sémantique commune et 3) contourner l'interférence engendrée par l'équivalence interlinguistique (voir aussi la sous-section 4.1.3.4).

Pour aider les apprenants à surmonter ces obstacles, nous mobilisons trois stratégies de description lexicale. Outre le classement hiérarchique des CLS (voir aussi la sous-section 4.2.1.1), nous présentons les termes sémantiquement proches comme étant liés par une relation paradigmatique. De plus, nous proposons un double accès aux CLS, c'est-à-dire via leur forme et leur équivalent français, pour rendre compte de l'impact de la L1 des apprenants sur l'acquisition des CLS.

La première solution consiste à regrouper les CLS en classes génériques et intermédiaires, comme nous l'avons abordée dans la sous-section 4.2.1.1. Le tableau 30 illustre la description des CLS contenant les deux acceptions verbales du vocable 检测, à savoir 检测₁ (« détecter ») et 检测₂ (« mettre à l'essai »). Cette manière de présenter les CLS permet aux apprenants d'accéder à un ensemble de choix lexicaux pour exprimer un sens spécifique. Par exemple, dans la classe intermédiaire « Détecter un objet spécifique », on répertorie six termes nominaux qui dénotent une entité nuisible au fonctionnement d'un système informatique, tels que 病毒 (« virus ») et 攻击 (« attaque informatique »). Il est aussi possible de comparer les CLS ayant la même structure syntaxique qui contiennent différentes acceptions d'un vocable. Par exemple, le rapprochement entre les CLS de type « prédicat-COD » contenant 检测₁ (« détecter ») et 检测₂ (« mettre à l'essai ») met en évidence les termes typiques qui remplissent la fonction de COD de ces verbes. Du point de vue sémantique, ces termes typiques réalisent le rôle sémantique de patient auprès des verbes en entrée. En plus, cette manière de présenter les CLS laisse paraître que seule la première acception du vocable (检测₁ « détecter ») admet un complément de résultat. Ce dernier est introduit par la particule structurale 出 (« parvenir à ») ou 到 (« arriver à ») et contient un terme qui réalise le COD du verbe. Une CLS renfermant 检测₁ (« détecter ») et un complément de résultat exprime le sens de 'réussir à détecter une entité nuisible', par exemple 检测₁出病毒 (« réussir à détecter un virus »). Quant à 检测₂ (« mettre à l'essai »), ce verbe n'admet pas de complément de résultat puisqu'il dénote l'action de 'soumettre un logiciel, un réseau ou un système informatique à un test'.

| Terme | Classe générique | Classe intermédiaire | Glose | CLS | Équivalent français |
|-------------------------|----------------------------------|--|---|--------------------|--|
| 检测1 détecter | Prédictat-COD | Détecter un objet spécifique | Logiciel malveillant le plus souvent transmis par un réseau ou un support de stockage externe, ayant pour but d'infecter un fichier qui, lorsqu'il est exécuté, lui permettra de se propager et de produire les effets pour lesquels il a été conçu [GDTVIRU] | 检测1病毒 | détecter un virus |
| | | | Propriété du matériel ou d'un logiciel qui peut subir une attaque ou d'être endommagé par un virus. [DIVULN] | 检测1漏洞 | détecter une vulnérabilité |
| | | | Cessation de l'aptitude d'un système complet ou d'un élément matériel ou logiciel à exécuter une fonction requise. [GDTDEFAI] | 检测1故障 | détecter une défaillance |
| | | | Action réalisée par un pirate ou un virus qui entre dans un réseau ou un système informatique illégalement dans le but d'accéder à des données et, éventuellement, de nuire au fonctionnement du matériel ou à l'intégrité des données [DIINTR] | 检测1攻击2 | détecter une attaque |
| | | | Divergence entre une valeur ou une condition calculée, observée ou mesurée, et la valeur ou la condition vraie, spécifique ou théoriquement correcte correspondante [GDTREUR] | 检测1错误 | détecter une erreur |
| | | | Événement potentiel et appréhendé, susceptible de porter atteinte à un système informatique. [GDTMENACE] | 检测1威胁 | détecter une menace |
| | Prédictat-Complément de résultat | Détecter pour aboutir à un résultat spécifique | De manière à identifier un objet spécifique (par exemple, un virus ou une erreur) | 检测1到... 检测1出... | réussir à détecter ... réussir à détecter ... |
| 检测2 mettre à l'essai | Prédictat-COD | Mettre à l'essai un objet spécifique | Logiciel [= ensemble de programmes utilisé par un utilisateur pour exécuter une ou plusieurs applications spécifiques sur un ordinateur] [DILOGI] | 检测2软件1 | mettre un logiciel à l'essai |

Tableau 30. – Mise en parallèle de la description des CLS contenant 检测₁ (« détecter ») et 检测₂ (« mettre à l'essai »)

La deuxième solution consiste à regrouper les termes qui sont reliés entre eux par des liens paradigmatiques. Ce classement offre un accès rapide à un ensemble de termes sémantiquement proches. Considérons les termes 传输₁ (« transférer ») et 迁移₁ (« faire migrer ») : les deux verbes désignent le passage des données d'un emplacement de mémoire vers un autre (voir aussi la sous-section 4.1.3.4). Dans l'article consacré à 迁移₁ (« faire migrer »), les termes reliés au terme en entrée par une relation paradigmatique sont présentés dans deux rubriques. La rubrique « Sens proche » répertorie le terme 传输₁ (« transférer ») qui a une composante sémantique commune avec 迁移₁ (« faire migrer »). L'autre rubrique intitulée « Famille de mot-résultat » regroupe la deuxième acception qui relève du même vocable que 迁移₁ (« faire migrer »), soit 迁移₂ (« migration ») qui est un dérivé sémantique du terme en entrée. La même stratégie descriptive s'applique à la description des termes reliés au terme 传输₁ (« transférer »), comme le montre le tableau 31.

| Terme en entrée | Lien sémantique | Terme relié | Équivalent |
|--|------------------------------|------------------------------|--------------|
| qiān yí 迁移 ₁ « faire migrer » | Sens proche | chuán shū 传输 ₁ | transférer |
| | Famille de mots-résultat | qiān yí 迁移 ₂ | migration |
| Terme en entrée | Lien sémantique | Terme relié | Équivalent |
| chuán shū 传输 ₁ « transférer » | Sens proche | qiān yí 迁移 ₁ | faire migrer |
| | Famille de mots-nom d'action | chuán shū 传输 ₂ | transfert |

Tableau 31. – Description des termes reliés à 迁移₁ (« faire migrer ») et à 传输₁ (« transférer »)

Au lieu d'explicitier la nuance sémantique entre les termes répertoriés dans la rubrique « Sens proche » et le terme en entrée, nous incitons les apprenants à distinguer les sens de ces termes à l'aide des renseignements descriptifs qui leur sont fournis. Cette démarche vise à soutenir un apprentissage autonome et actif. Outre les termes reliés et les CLS, nous présentons, pour chaque terme en entrée, les renseignements grammaticaux, la définition, la structure actancielle et les contextes d'utilisation. Nous reviendrons sur la compilation des données terminologiques dans la section suivante. Quand il est question d'élaborer la ressource CHINOINFO, nous proposons de créer, dans chaque article, des hyperliens vers les autres entrées qui entretiennent une relation paradigmatique avec le terme décrit, afin de faciliter la consultation des apprenants.

La troisième solution rend compte de l'influence de la L1 des apprenants sur l'acquisition des CLS. Nous avons observé, dans la pratique d'enseignement auprès d'étudiants francophones, le phénomène de *greffe collocationnelle interlinguistique* définie par Polguère (2007, p. 262). Il s'agit d'une collocation erronée constituée d'au moins deux éléments lexicaux A₁ + B₂. On suppose que cette expression a été produite par interférence intralinguistique mettant en jeu une collocation valide A₁ + B₁ (appartenant à une langue L1), initialement visée par le locuteur, et une collocation également valide A₂ + B₂ (appartenant à autre langue L2) ayant parasité la production de A₁+ B₁. En enseignement des L2, les apprenants non natifs ont tendance à associer une lexie à un collocatif inapproprié qui semble 'emprunté' de leur langue maternelle, il s'agit

d'un cas particulier de *greffe collocationnelle interlinguistique*, dont la source⁶² est la langue maternelle des apprenants.

En effet, on prévoit que ces apprenants auront de la difficulté à distinguer les termes dont l'équivalent est associé à une même forme lexicale dans leur L1. Dans ce cas, ils sont susceptibles de faire des associations erronées entre ces termes et les collocatifs respectifs de ces derniers. En didactique des L2, ce phénomène s'explique par le transfert négatif (voir la section 3.3.1), causé par l'interférence de la L1 des apprenants et la surgénéralisation des connaissances acquises en L2 (Doca, 1981 ; Tardif, 1992, 1999). Le transfert constitue un facteur essentiel dans l'apprentissage d'une L2 selon l'approche cognitive. Dans le cadre de notre thèse, nous nous focalisons sur l'interférence causée par l'équivalence interlinguistique. Pour en rendre compte, nous empruntons la méthode élaborée par Pecman (2005) pour donner accès aux CLS via leur forme et leur sens lors de la création du dictionnaire CHINOINFO. D'une part, il est possible de rechercher une CLS à partir de ses deux éléments constitutifs si ces derniers sont répertoriés dans la ressource lexicale. Par exemple, les apprenants peuvent consulter la CLS yìng yòng ruǎn jiàn 应用₁ 软件₁ (« logiciel d'application ») dans les articles yìng yòng 应用₁ (« application ») et ruǎn jiàn 软件₁ (« logiciel »). D'autre part, pour les locuteurs non natifs, comme le souligne Pecman (2005), leur langue première demeure le seul point d'accès possible au sens d'une CLS en L2. Nous proposons de créer un index des termes à partir de leur équivalent français, pour permettre aux apprenants d'accéder à un terme et aux CLS qui le contiennent via l'équivalent du terme dans leur L1. Cette méthode de classement permet de rapprocher les termes dont l'équivalent français est associé à une même forme lexicale. Par exemple, le CHINOINFO répertorie trois articles dont l'équivalent français est *en ligne* (figure 30) : wǎng luò 网络₃ et zài xiàn 在线₁ correspondent à l'adjectif *en ligne* en français, zài xiàn 在线₂ est le dérivé sémantique adverbial de zài xiàn 在线₁ et se traduit par l'adverbe *en ligne*. Les équivalents français sont assortis d'un numéro d'acception (*en ligne*₁ et *en ligne*₂) pour distinguer les sens associés à la même forme lexicale. Nous fournirons une description détaillée de la ressource élaborée dans la section 5.2.

⁶² La source d'une greffe collocationnelle est définie par Polguère (2007, p. 262) comme « la collocation valide qui est à l'origine de l'interférence linguistique ».

Nomenclature (91 entrées) :

triée par lexie triée par catégorie **triée par glose**

| | | | |
|-----------------------------------|---------------------------------------|------------------------------|------------------------------------|
| 远程 2 (adv.) [à distance] | 加密 2 (nom) [cryptage] | 安装 1 (v.tr.) [installer 2] | 备份 1 (v.tr.) [sauvegarder] |
| 访问 1 (v.tr.) [accéder 1] | 加密 3 (adj.) [crypté] | 入侵 2 (nom) [intrusion] | 连接 1 (v.tr.) [se connecter 1] |
| 访问 2 (v.tr.) [accéder 2] | 加密 1 (v.tr.) [crypter] | 软件 1 (nom) [logiciel 1] | 安全 2 (adj.) [sécurisé] |
| 访问 3 (nom) [accès 1] | 恶意 2 (adv.) [de manière malveillante] | 软件 2 (nom) [logiciel 2] | 安全 3 (adv.) [sécuritairement] |
| 访问 4 (nom) [accès 2] | 检测 1 (v.tr.) [détecter 1] | 恶意 1 (adj.) [malveillant 1] | 安全 1 (nom) [sécurité informatique] |
| 移动 2 (adj.) [amovible] | 检测 3 (nom) [détection 1] | 手动 1 (adj.) [manuel] | 存储 3 (nom) [stockage 1a] |
| 应用 1 (nom) [application 1] | 远程 1 (adj.) [distant] | 手动 2 (adv.) [manuellement] | 存储 2 (nom) [stockage 1b] |
| 应用 2 (nom) [application 2] | 网络 3 (adj.) [en ligne 1] | 更新 1 (v.tr.) [mettre à jour] | 存储 1 (v.tr.) [stocker] |
| 应用 4 (v.tr.) [appliquer 2] | 在线 1 (adj.) [en ligne 1] | 迁移 2 (nom) [migration] | 删除 2 (nom) [suppression 1] |
| 攻击 2 (nom) [attaque informatique] | 在线 2 (adv.) [en ligne 2] | 更新 2 (nom) [mise à jour 1.1] | 删除 1 (v.tr.) [supprimer 1] |
| 攻击 1 (v.tr.) [attaquer 1] | 实时 1 (adj.) [en temps réel 1] | 更新 3 (nom) [mise à jour 1.2] | 同步 3 (adj.) [synchrone 1] |
| 自动 1 (adj.) [automatique 1] | 实时 2 (adv.) [en temps réel 2] | 移动 1 (adj.) [mobile 1] | 同步 4 (adv.) [synchroniquement] |
| 自动 2 (adv.) [automatiquement] | 保存 2 (nom) [enregistrement 1.1] | 共享 2 (adj.) [partagé] | 同步 2 (nom) [synchronisation] |
| 设置 2 (nom) [configuration 1.1] | 保存 1 (v.tr.) [enregistrer 1] | 共享 3 (nom) [partage] | 同步 1 (v.tr.) [synchroniser] |
| 设置 3 (nom) [configuration 1.2] | 发送 3 (nom) [envoi 1] | 共享 2 (nom) [partage 1] | 下载 2 (nom) [téléchargement] |
| 设置 1 (v.tr.) [configurer 1] | 发送 4 (nom) [envoi 2] | 共享 1 (v.tr.) [partager 1] | 下载 1 (v.tr.) [télécharger] |
| 连接 2 (v.tr.) [connecter] | 发送 1 (v.tr.) [envoyer 1] | 网络 1 (nom) [réseau 1] | 测试 2 (nom) [test] |
| 连接 3 (nom) [connexion 1.1] | 发送 2 (v.tr.) [envoyer 2] | 网络 2 (nom) [réseau 2] | 测试 1 (v.tr.) [tester] |
| 连接 4 (nom) [connexion 2.1] | 运行 3 (nom) [exécution 1a] | 应用 3 (v.intr.) [s'appliquer] | 传输 1 (v.tr.) [transférer 1] |
| 复制 3 (nom) [copie 1.1] | 运行 4 (nom) [exécution 1b] | 运行 1 (v.intr.) [s'exécuter] | 传输 2 (nom) [transfert 1] |
| 复制 4 (nom) [copie 2] | 运行 2 (v.tr.) [faire exécuter] | 入侵 1 (v.tr.) [s'introduire] | 检测 4 (nom) [vérification] |
| 复制 1 (v.tr.) [copier 1] | 迁移 1 (v.tr.) [faire migrer] | 备份 4 (adj.) [sauvegardé] | 检测 2 (v.tr.) [vérifier] |
| 复制 2 (v.tr.) [copier 2] | 安装 2 (nom) [installation 2] | 备份 2 (nom) [sauvegarde 1] | |
| | | 备份 3 (nom) [sauvegarde 2] | |

Figure 30. – Nomenclature du dictionnaire CHINOINFO

Pour aider les apprenants à différencier ces termes sémantiquement proches et à choisir les collocatifs appropriés pour chaque terme, nous explicitons d'abord les liens sémantiques entre ces termes, ce qui constitue un point de départ pour différencier ces trois termes (tableau 32). 在线₁ (« en ligne₁ ») et 网络₃ (« en ligne₁ ») sont deux adjectifs quasi-synonymes. Bien que l'équivalent français de ces termes semble être le même, le chinois fait une distinction que ne fait pas le français. Les deux termes se combinent avec un nom désignant un type de ressource en ligne (comme 视频 « vidéo » et 游戏 « jeu ») ou un type d'activités en ligne (tel que 直播 « diffusion en direct » et 升级 « mise à niveau »). Seul 网络₃ (« en ligne ») peut s'associer à un nom désignant une activité illicite comme 钓鱼 (« hameçonnage ») ou 攻击₂ (« attaque informatique »). Ainsi, les liens sémantiques mutuels entre ces deux lexies peuvent être représentés au moyen des FL **Syn_▷** (synonyme plus spécifique) et **Syn_◁** (synonyme moins spécifique) comme suit (Mel'čuk et al., 1995, p. 129-130) :

Syn_◁ (在线₁) = 网络₃,

Syn_▷ (网络₃) = 在线₁.

Dans le CHINOINFO, les liens observés sont encodés comme « quasi-synonyme », accompagnés d'une mention précisant s'il s'agit d'un « synonyme plus spécifique » ou d'un « synonyme moins spécifique ».

Les liens de dérivation sémantique entre ^{zài xiàn}在线₂ (« en ligne₂ ») et ^{zài xiàn}在线₁ (« en ligne₁ »), qui peuvent être modélisés par les FL paradigmatiques **Adv**₀ et **A**₀, sont encodés comme « Famille de mots », accompagnés d'une mention de partie du discours auquel appartient le terme relié (voir aussi les sous-sections 3.1.1 et 4.2.1.1). Pour saisir le sens véritable de ces termes, les apprenants devront lire les définitions proposées dans les articles correspondants.

Dans le tableau 32, les CLS contenant ces termes sont présentées de façon structurée : dans chaque article, les CLS sont d'abord organisées en classes génériques en fonction de leur structure syntaxique, celles-ci sont subdivisées en classes intermédiaires pour une description plus fine des propriétés syntaxico-sémantiques des CLS (voir aussi la sous-section 4.2.1.1). La comparaison des CLS contenant les termes dont l'équivalent français est *en ligne* permet de distinguer les deux adjectifs (^{zài xiàn}在线₁ « en ligne₁ » et ^{wǎng luò}网络₃ « en ligne₁ ») de l'adverbe ^{zài xiàn}在线₂ (« en ligne₂ ») puisqu'ils apparaissent dans des CLS de structures syntaxiques différentes. Le regroupement des CLS en classes intermédiaires fait ressortir les nuances sémantiques subtiles entre les deux quasi-synonymes ^{zài xiàn}在线₁ (« en ligne₁ ») et ^{wǎng luò}网络₃ (« en ligne₁ »).

| Termes | Partie du discours | Equivalent | |
|---|----------------------------|---|--|
| wǎngluò 网络 ₃ | adj. | en ligne 1 | |
| zàixiàn 在线 ₁ | adj. | en ligne 1 | |
| zàixiàn 在线 ₂ | adv. | en ligne 2 | |
| Comparaison des termes reliés | | | |
| | Lien sémantique | Terme relié | Equivalent |
| wǎngluò 网络 ₃ | Synonyme plus spécifique | zàixiàn 在线 ₁ | en ligne 1 (adj.) |
| zàixiàn 在线 ₁ | Synonyme moins spécifique | wǎngluò 网络 ₃ | en ligne 1 (adj.) |
| | Famille de mots-adverbe | zàixiàn 在线 ₂ | en ligne 2 (adv.) |
| zàixiàn 在线 ₂ | Famille de mots-adjectif | zàixiàn 在线 ₁ | en ligne 1 (adj.) |
| Comparaison des combinaisons lexicales spécialisées (CLS) | | | |
| | | Classe générique [Modificateur-nom modifié] | Classe générique [Modificateur adverbial-prédicat] |
| En ligne 1 | wǎngluò 网络 ₃ | Classes intermédiaires Types de ressources en ligne wǎngluò zī yuán 网络资源 ressource en ligne wǎngluò yú lè 网络娱乐 divertissement en ligne wǎngluò shì pín 网络视频 vidéo en ligne wǎngluò yóu xì 网络游戏 jeu en ligne Types d'activités illicites en ligne wǎngluò diào yú 网络钓鱼 hameçonnage wǎngluò zhà piàn 网络诈骗 escroquerie en ligne wǎngluò gōng jī 网络攻击 ₂ cyberattaque Types d'activités en ligne wǎngluò yíng xiāo 网络营销 cybermarketing wǎngluò zhí bō 网络直播 diffusion en direct en ligne | Sans objet |
| | zàixiàn 在线 ₁ | Classes intermédiaires Types de ressources en ligne zàixiàn shì pín 在线视频 vidéo en ligne zàixiàn bāng zhù 在线帮助 aide en ligne zàixiàn yóu xì 在线游戏 jeu en ligne zàixiàn kè fú 在线客服 service à la clientèle en ligne zàixiàn kè chéng 在线课程 cours en ligne Types d'activités en ligne zàixiàn cún chū 在线存储 ₂ stockage en ligne zàixiàn shēng jí 在线升级 mise à niveau en ligne zàixiàn jiāo yù 在线教育 éducation en ligne zàixiàn zhī fù 在线支付 paiement en ligne zàixiàn huì yì 在线会议 conférence en ligne | Sans objet |
| En ligne 2 | zàixiàn 在线 ₂ | Sans objet | Classes intermédiaires Types d'activités zàixiàn shēng jí 在线 ₂ 升级 mettre à niveau en ligne zàixiàn fù kuǎn 在线 ₂ 付款 payer en ligne zàixiàn biān jí 在线 ₂ 编辑 éditer en ligne zàixiàn gòu wù 在线 ₂ 购物 faire un achat en ligne |

Tableau 32. – Comparaison des liens sémantiques de trois termes ayant le sens d'*en ligne*

4.2.1.4 Synthèse des solutions proposées

En guise de synthèse, nous pensons que la description des propriétés syntaxico-sémantiques des CLS joue un rôle clé pour aider les apprenants non natifs à surmonter les obstacles d'apprentissage que nous avons anticipés. Cet aspect est pris en considération sous deux angles différents. D'un point de vue linguistique, nous nous appuyons sur la LEC pour modéliser les liens sémantiques des termes décrits. Dans un article, les liens paradigmatiques et syntagmatiques du terme en entrée sont distingués et présentés dans les rubriques « Termes reliés » et « Combinaisons lexicales ». Nous y reviendrons lorsqu'il sera question de présenter notre modèle d'encodage des données terminologiques (sous-section 4.2.2.1).

D'un point de vue didactique, nous adoptons différentes stratégies préconisées par l'approche cognitive, soit la présentation structurée des connaissances lexicales et la prise en compte du transfert d'apprentissage (voir la section 3.3). Lors de l'analyse des données terminologiques recueillies à partir du corpus, nous regroupons les acceptions d'un polysème dans un même vocable, comme le propose la LEC. Chaque acception fait l'objet d'un article dans le CHINOINFO. Lors de l'encodage des données du CHINOINFO, les articles relevant d'un même vocable sont rédigés dans un même document XML⁶³ (nous reviendrons sur l'encodage des données dans la sous-section 4.2.2.2). Au sein d'un article, les CLS contenant un terme sont organisées en classes génériques selon leur structure syntaxique. Elles sont ensuite structurées en classes intermédiaires. Celles-ci regroupent deux types de CLS : les collocations contrôlées par le terme décrit et la combinaison du terme et son ASyntP. L'élaboration des classes intermédiaires rend compte non seulement de la relation sémantique au sein d'une CLS, mais aussi du sens d'un cooccurrent et de son rôle sémantique par rapport au terme vedette. La description systématique des CLS vise à faciliter l'encodage⁶⁴ et l'assimilation des connaissances lexicales par les apprenants. De plus, nous proposons d'indexer les termes décrits par leur équivalent français lors de la création de la ressource CHINOINFO, pour permettre aux apprenants d'accéder aux CLS

⁶³ Dans la version en ligne du CHINOINFO, les articles relevant d'un vocable sont présentés séparément. Ils se trouvent à peu près au même endroit dans la liste des termes grâce à leur forme identique (voir aussi la sous-section 5.1).

⁶⁴ Rappelons que le terme encodage désigne ici la structuration de nouvelles informations en vue de leur archivage dans la mémoire des apprenants. Cette notion est abordée dans la section 3.3 consacrée à l'approche cognitive.

exprimant un sens particulier. Cette démarche vise aussi à mettre les apprenants en garde contre l'interférence de leur L1 sur l'acquisition des CLS, plus précisément, sur le choix de collocations appropriés pour un terme spécifique.

4.2.2 Création de la ressource lexicale

Une méthode qui intègre les pistes de solution esquissées a été appliquée ensuite à l'élaboration du dictionnaire CHINOINFO, une ressource lexicale destinée aux apprenants francophones du COS. La création du CHINOINFO a été réalisée en trois temps : la compilation des données terminologiques (4.2.2.1), la rédaction des articles en langage XML (*eXtensible Markup Language*) (4.2.2.2) et la mise en ligne de la ressource (4.2.2.3).

4.2.2.1 Compilation des données terminologiques

Dans un premier temps, nous avons créé un modèle uniforme pour structurer les données terminologiques dans un article, en vue d'un encodage systématique des données dans le CHINOINFO. À ce stade, nous nous sommes inspirée du modèle d'encodage de deux dictionnaires spécialisés mis au point par l'Observatoire de linguistique Sens-Texte (OLST) : le DiCoInfo (Observatoire de linguistique Sens-Texte, 2022b) et le DiCoEnviro (Observatoire de linguistique Sens-Texte, 2022a). Ces ressources terminologiques multilingues, comme nous l'avons mentionné dans le chapitre III, montrent de quelle manière les principes méthodologiques de l'approche lexico-sémantique guident la description des termes à différents stades, notamment la structuration des données terminologiques. L'annexe 8 illustre l'organisation des renseignements descriptifs dans l'article ^{ruǎn jiàn} 软件₁ (« logiciel₁ »). Comme on peut le constater, notre modèle est composé de quatre grandes sections qui sont organisées de la manière suivante.

1). Renseignements généraux sur le terme vedette

L'organisation du contenu dans cette section est largement inspirée du DiCoInfo (L'Homme, 2008; L'Homme et Alipour, 2009). Dans la première section de l'article ^{ruǎn jiàn} 软件₁ (« logiciel₁ ») (annexe 8), on trouve d'abord l'entrée accompagnée d'un numéro d'acception (^{ruǎn jiàn} 软件₁), l'indication de la partie du discours à laquelle le terme en entrée appartient (nom) et l'équivalent français du terme (logiciel₁). Un équivalent est aussi accompagné d'un numéro d'acception dans le cas où il relève d'un polysème.

Nous présentons ensuite la structure actancielle du terme vedette. Celle-ci est dégagée des contextes d'utilisation du terme que nous avons extraits de notre corpus. La structure actancielle énumère les actants sémantiques du terme et explicite le rôle de chaque actant sémantique par rapport au terme. 软件₁ (^{ruǎnjiàn} « logiciel₁ »), à titre d'exemple, fait appel à deux actants : l'*agent* qui utilise un logiciel, réalisé par des termes typiques comme 用户 (^{yòng hù} « utilisateur), et le *patient* sur lequel l'agent intervient en faisant appel à un logiciel. Le deuxième actant sémantique peut être réalisé par des termes désignant une application informatique, tels que 应用₂ (^{yìng yòng} « application informatique »), 办公 (^{bàn gōng} « bureautique ») et 杀毒 (^{shā dú} « antivirus »). Nous décidons de présenter, d'entrée de jeu, la structure actancielle pour sensibiliser les apprenants aux éléments qui participent à l'expression du sens du terme.

Nous proposons aussi la définition en français pour le terme vedette. Cette démarche vise à permettre aux apprenants non sinophones d'accéder aux connaissances spécialisées véhiculées par le terme en entrée. Nous avons eu recours à des sources secondaires, notamment les dictionnaires spécialisés et les banques de terminologie, pour trouver la grande majorité des définitions. Les outils de référence utilisés comprennent le DiCoInfo (Observatoire de linguistique Sens-Texte, 2022b), le Grand dictionnaire terminologique (GDT) (Office québécois de la langue française, 2022) et TERMIUM PLUS (Travaux publics et services gouvernementaux Canada., 2022). Dans le modèle d'encodage, nous avons indiqué la source de définition à l'aide d'un code (tel que [GDTLOGICIEL]) qui fait référence à un lien Web ou à un ouvrage de référence. Dans le cas où nous ne trouvons pas de définition pour un terme en entrée, nous avons rédigé une définition en français en nous appuyant sur les renseignements définitoires extraits de notre corpus chinois.

2). Contextes d'utilisation

La deuxième section d'un article présente des exemples d'usage du terme en entrée. Nous avons extrait des contextes pour chacun des 91 termes relevés lors des étapes de travail précédentes, c'est-à-dire le repérage des termes à partir des UT polysémiques (sous-section 4.1.2.4) et la collecte des contextes d'utilisation des CLS (sous-section 4.1.3.2). Tous les contextes contenant un terme ne sont pas répertoriés, étant donné que cela risque d'encombrer l'article et rendre la lecture de l'article difficile et chronophage.

Dans un article, le nombre de contextes est fonction du nombre de CLS relevées et varie entre 6 à 15. Nous avons privilégié les contextes dans lesquels apparaissent les CLS décrites dans les articles, afin d'illustrer le comportement combinatoire du terme vedette dans des textes spécialisés authentiques. Les termes vedettes sont surlignés dans les contextes répertoriés lors de la mise en ligne des articles rédigés (voir aussi la sous-section 5.2.3).

3). Termes reliés

Les deux dernières grandes sections d'un article, les plus importantes, sont consacrées à la description des liens lexicaux du terme vedette. La troisième rubrique intitulée « Termes reliés » répertorie les termes qui entretiennent un lien sémantique avec le terme vedette. Dans l'article ^{ruǎn}软件₁ (« logiciel₁ »), par exemple, nous présentons quatre liens paradigmatiques du terme vedette : l'hyponyme de *logiciel*₁ (^{yìng yòng}应用₁ « application »), l'holonyme ^{ruǎn jiàn}软件₂ (« logiciel₂ ») qui désigne l'ensemble de *logiciel*₁, le méronyme ^{chéng xù}程序 (« programme ») qui fait référence à l'élément constitutif d'un *logiciel*₁ et un terme qui décrit l'état d'un *logiciel*₁ à un moment donné de son développement, soit ^{bǎn běn}版本 (« version »).

4). Combinaisons lexicales

La dernière section d'un article présente les CLS contenant le terme décrit. Celles-ci sont structurées en classes génériques selon leur structure syntaxique. Une classe générique se subdivise en différentes classes intermédiaires, soit selon le lien sémantique qui relie le terme à son collocatif, soit selon le sens et le rôle sémantique du cooccurrent qui réalise un ASyntP du terme. La classe intermédiaire « Sortes de » dans l'article ^{ruǎn jiàn}软件₁ (« logiciel₁ ») regroupe les CLS dans lesquelles le terme vedette se combine avec un modificateur, telles que ^{shā dú ruǎn jiàn}杀毒软件₁ (« logiciel antivirus ») et ^{kāi yuán ruǎn jiàn}开源软件₁ (« logiciel libre »). Les classes intermédiaires sont ensuite regroupées en classes génériques selon la structure syntaxique des CLS (voir aussi la sous-section 4.2.1.1). Pour chaque CLS, nous fournissons une brève explication et l'équivalent français de celle-ci. Dans le cas où le collocatif du terme vedette fait partie de la nomenclature du CHINOINFO, nous créons un lien de renvoi vers l'article correspondant lors de la mise en ligne de la ressource.

4.2.2.2 Encodage des données terminologiques

Une fois les données compilées, nous avons procédé à la rédaction des articles. Comme notre modèle d'encodage (voir aussi la sous-section 4.2.2.1) est basé sur le modèle utilisé dans le DiCoInfo (Observatoire de linguistique Sens-Texte, 2022b) et le DiCoEnviro (Observatoire de linguistique Sens-Texte, 2022a), nous avons adapté la méthodologie d'encodage de ces dictionnaires à l'encodage du CHINOINFO.

Les articles du DiCoInfo sont rédigés en langage XML (*eXtensible Markup Language*). Le XML est un langage informatique qui permet de structurer, à l'aide de balises et de règles que l'on peut personnaliser, les données d'un document de manière hiérarchisée et organisée (Roland, 2022). Il s'agit d'un langage de description de documents qui se prête bien à l'organisation de données terminologiques. Quant à l'outil de rédaction des fichiers XML, nous avons eu recours à *Oxygen XML Editor*, un environnement de développement des documents XML. Ce logiciel offre plusieurs fonctionnalités qui nous ont été utiles, outre la création et l'édition des fichiers XML, il permet aussi la création des schémas pour valider le contenu encodé dans un document XML.

Du point de vue de la structure de la ressource CHINOINFO, les articles relevant d'un même vocable sont rédigés dans un même document XML. L'annexe 9 illustre un fichier ouvert dans *Oxygen XML Editor*. Nous avons encodé dans ce fichier les deux acceptions du vocable ^{ruǎn jiàn}软件, soit ^{ruǎn jiàn}软件₁ (« logiciel₁ ») et ^{ruǎn jiàn}软件₂ (« logiciel₂ »). Comme on peut le constater à la partie gauche de l'interface d'*Oxygen XML Editor*, chaque vocable correspond à un fichier XML distinct. Les fichiers sont nommés de façon uniforme ('l'équivalent de la première acception du vocable-forme lexicale du vocable') et classés par ordre alphabétique. Le corps du fichier consacré au vocable ^{ruǎn jiàn}软件 est constitué d'un ensemble de balises qui servent à structurer les données terminologiques encodées de façon hiérarchisée. Le nom des balises indique le contenu consigné dans ces balises. Comme les documents XML ont une structure arborescente, il est possible de plier ou déplier certains éléments d'un document dans l'interface d'*Oxygen XML Editor*. La figure 31 illustre l'affichage plié du fichier ^{ruǎn jiàn}软件 qui met en évidence le nombre d'articles dans le vocable.

```

3 <vocabule xmlns="http://olst.ling.umontreal.ca/dicoinfo/" xmlns:xi="http://www.w3.org/2003/XInclude"
4   xmlns:saxon="http://saxon.sf.net/" xmlns:xs="http://www.w3.org/2001/XMLSchema"
5   xmlns:xlink="http://www.w3.org/1999/xlink" identificateur="软件" xml:lang="zh">
6
7 ▶   <lexie xml:id="_软件1" glose="logiciel1" numero-acceptation="1" statut="3" projet="CHINOINFO" [444 lines]
52 ▶   <lexie xml:id="_软件2" glose="logiciel2" numero-acceptation="2" statut="3" projet="CHINOINFO" [74 lines]
27 </vocabule>

```

Figure 31. – Affichage plié du fichier XML consacré au vocable ruǎn jiàn 软件

Au sein d'un article, la balise *<lexie>* englobe l'ensemble des données descriptives du terme en entrée. Celles-ci sont structurées en quatre sections principales, comme nous l'avons expliqué dans la sous-section 4.2.2.1 : 1) renseignements généraux sur le terme vedette; 2) contextes d'utilisation; 3) termes reliés et 4) combinaisons lexicales. Chaque section contient plusieurs balises. La figure 32 illustre la première section de l'article ruǎn jiàn 软件₁ (« logiciel₁ ») (voir aussi l'annexe 9). Nous présentons ci-dessous les données encodées dans cette section.

- Les deux premières lignes de la balise *<lexie>* contiennent plusieurs attributs. Ces derniers donnent des renseignements supplémentaires sur l'article rédigé, dont l'identifiant du fichier XML (xml : id= “_软件 1”), la glose⁶⁵ (glose=“logiciel1”), le numéro d'acceptation de l'article, le statut de rédaction⁶⁶, le nom de projet (CHINOINFO), le rédacteur du fichier et la date de la dernière mise à jour.
- La balise *<information-grammaticale>* encode la partie du discours du terme en entrée.
- La balise *<équivalence>* contient des attributs qui fournissent des informations requises pour créer un lien de renvoi vers l'article *logiciel₁* dans le dictionnaire DiCoInfo (Observatoire de linguistique Sens-Texte, 2022b).
- La balise *<structure-actancielle>* englobe d'autres balises, y compris *<role>* qui indique le rôle de chaque actant sémantique par rapport au terme en entrée. Sous la balise *<role>* se

⁶⁵ La glose figurant à la première ligne de la balise *<lexie>* sert d'identificateur du terme décrit. Elle diffère de celles que nous avons évoquées pour expliquer les liens sémantiques observés dans les CLS.

⁶⁶ Dans le DiCoInfo, la notation des statuts de rédaction indique l'état d'avancement de la rédaction d'un article : statut 0 signifie = rédaction terminée; statut 1 = rédaction très avancée; statut 2 = rédaction avancée; statut 3 = rédaction en cours. Dans le CHINOINFO, nous utilisons la même méthode de notation. Le statut de rédaction par défaut est de 3 (Observatoire de Linguistique Sens-Texte, 2011, p. 9).

trouve la balise `<tt>` dans laquelle on encode les termes typiques qui réalisent l'actant sémantique décrit. La balise `<tt>` contient en outre un attribut intitulé « glose » qui fournit l'équivalent français du terme typique placé dans ces balises.

- La balise `<définition>` contient la définition du terme en entrée. On y trouve un attribut « source » qui encode la référence de la définition si celle-ci est citée telle quelle.
- La balise `<domaine>` indique le domaine de spécialité auquel le terme appartient, soit l'informatique.

```

<lexie xml:id="_软件1" glose="logiciel1" numero-acceptation="1" statut="3" projet="CHINOINFO"
redacteur="ZWH" mise-a-jour="2022-03-19">date de la dernière mise à jour
rédacteur du fichier
<information-grammaticale>nom</information-grammaticale>
<equivalences>
<equivalence xml:lang="fr" xlink:type="simple" xlink:href="logiciel.xml#_logiciel1"
xlink:show="replace" xlink:actuate="onRequest">logiciel 1</equivalence>
</equivalences>
<structure-actancielle>
<role nom="Agent">
<tt glose="utilisateur">用户</tt>
</role> 使用 (utiliser) <lexie-ref/> 操作 (intervenir sur) <role nom="Patient">
<tt glose="application informatique">应用2</tt>
<tt glose="bureautique">办公</tt>
<tt glose="antivirus">杀毒</tt>
</role></structure-actancielle>
<definition source="GDTLOGICIEL">Ensemble des programmes constituant une unité destinée à
effectuer un traitement particulier sur un ordinateur.</definition>
<domaine>informatique</domaine>

```

Figure 32. – Extrait du fichier XML de l'article 软件₁ (« logiciel₁ »)

La structure arborescente du fichier XML (figure 33) se prête bien à l'encodage des CLS. Dans l'article 软件₁ (« logiciel₁ »). Toutes les combinaisons lexicales du terme sont organisées dans les balises `<combinaisons-lexicales>`. Celles-ci comprennent plusieurs `<structure-syntaxique>` dans lesquelles on peut trouver plusieurs `<lien-sémantique>` qui consignent les classes intermédiaires des CLS telles que « Sortes de ». Une classe intermédiaire peut comprendre un ou plusieurs CLS placé(s) dans les balises `<combinaison-lexicale>`. On y consigne aussi les données descriptives de chaque CLS encodée, c'est-à-dire l'explication de la CLS, sa forme lexicale et son équivalent français. Si un collocatif du terme vedette est répertorié dans le CHINOINFO, nous créons un lien de renvoi vers l'article correspondant, c'est le cas de 应用₁ qui se combine avec le terme vedette dans la CLS 应用₁软件₁ (« logiciel d'application »).

```

<combinaisons-lexicales>
  <structure-syntaxique nom="Modificateur + [Nom modifié]">
    <lien-semantic nom="Sortes de">
      <combinaison-lexicale>
        <explication>Logiciel, ou sous-ensemble de la partie logicielle d'un système
          informatique, qui assure le fonctionnement de l'ordinateur et son
          exploitation [GDTLOGISYST]</explication>
        <combinaison>系统软件1</combinaison>
        <equivalence xml:lang="fr" xlink:type="simple" xlink:href=""
          xlink:show="replace" xlink:actuate="onRequest">logiciel système (=
          logiciel de base)</equivalence>
      </combinaison-lexicale>
      <combinaison-lexicale>
        <explication>Ensemble de programmes dont se servent les utilisateurs afin
          d'accomplir une tâche ou une activité particulière
          [GDTLOGIDAPP]</explication>
        <combinaison><terme-relie xlink:href="application-应用.xml#_应用1"
          xlink:show="replace" xlink:actuate="onRequest"
          >应用1</terme-relie>软件1</combinaison>
        <equivalence xml:lang="fr" xlink:type="simple"
          xlink:href="application.xml#_application1" xlink:show="replace"
          xlink:actuate="onRequest">logiciel d'application</equivalence>
      </combinaison-lexicale>
    </lien-semantic>
  </structure-syntaxique>
</combinaisons-lexicales>

```

Figure 33. – Structure arborescente de la partie « combinaisons lexicales » du fichier XML de l'article ^{ruǎnjiàn} 软件₁ (« logiciel₁ »)

Outre les 33 vocables encodés en format XML, le CHINOINFO contient deux fichiers de schéma : « CHINOINFO.rnc » et « CHINOINFO.sch ». Le premier fichier de format RNC (pour *Relax NG compact*) fournit une description formelle de la structure des fichiers XML. Le deuxième fichier de format SCH sert à valider le contenu d'un document XML, il contient des règles pour vérifier que les valeurs des balises sont cohérentes entre elles dans un document⁶⁷.

4.2.2.3 Mise en ligne de la ressource

À la suite de l'encodage des données terminologiques, nous avons procédé à la mise en ligne du dictionnaire CHINOINFO. Cette tâche a été réalisée en collaboration avec Benoit Robichaud, membre de l'Observatoire de linguistique Sens-Texte (OLST).

⁶⁷ Les deux fichiers ont été adaptés au chinois des fichiers de schéma du DiCoInfo. Nous remercions Benoit Robichaud, membre de l'Observatoire de linguistique Sens-Texte (OLST), pour l'adaptation du schéma XML et l'élaboration de l'interface du CHINOINFO.

L'interface du CHINOINFO est hébergée sur le serveur de l'OLST et accessible au public⁶⁸. L'annexe 10 illustre l'article 软件₁^{ruǎnjiàn} (« logiciel₁ ») dans la version en ligne du dictionnaire. Nous présenterons le contenu et les fonctionnalités du CHINOINFO dans le chapitre 5.

4.2.3 Synthèse

Dans la section 4.2, nous avons présenté la deuxième phase de notre recherche qui comprend deux étapes. Nous nous sommes d'abord attachée à résoudre les trois types de difficultés syntaxico-sémantiques que les CLS sont susceptibles de soulever (4.2.1). À ce stade, nous nous sommes largement inspirée de la LEC et de la méthodologie lexicographique utilisée pour élaborer les ressources lexicales encadrées par la LEC, notamment le LAF, le DiCo et le DiCoInfo. Inspirée des travaux menant à la réalisation du DiCoInfo, nous avons mobilisé différentes stratégies pour représenter les propriétés syntaxico-sémantiques des CLS, notamment la description des liens lexicaux et la structuration hiérarchique des CLS. Une méthode descriptive intégrant les solutions proposées a ensuite été appliquée à la création du dictionnaire CHINOINFO, une ressource d'apprentissage du lexique spécialisé chinois portant sur le domaine de l'informatique (4.2.2). Cette ressource chinois-français est conçue pour les apprenants francophones du COS. Nous avons exposé la démarche méthodologique adoptée pour rédiger les articles du CHINOINFO et mettre la ressource en ligne. Lors de la dernière phase de notre recherche, nous avons mis le CHINOINFO à la disposition d'apprenants francophones du chinois, afin d'évaluer l'efficacité de la description lexicale et de recueillir les commentaires des participants quant à l'utilisation de la ressource lexicale.

4.3 Évaluer l'efficacité de la ressource lexicale mise au point

La dernière phase de notre recherche consiste en une expérimentation comparative auprès d'apprenants francophones du chinois. Dans cette section, nous présentons d'abord la conception de l'expérimentation (4.3.1). Nous exposons ensuite la mise en place de l'expérimentation qui s'est déroulée en trois étapes, à savoir la planification des séances expérimentales, la collecte de

⁶⁸ La version en ligne du CHINOINFO est repérable à l'adresse suivante : <http://olst.ling.umontreal.ca/chinoinfo/>

données et l'analyse des données recueillies (4.3.2). À la fin de la section, nous faisons une synthèse de notre démarche méthodologique (4.3.3).

4.3.1 Conception d'une expérimentation comparative

Cette sous-section présente la mise en place de notre expérimentation. Nous commençons par énoncer les objectifs de l'expérimentation, pour expliquer ensuite le profil des participants recherchés et le processus de recrutement. Nous décrivons par la suite le protocole d'expérimentation. Enfin, nous présentons les outils de collecte de données, y compris un test lexical et trois questionnaires.

4.3.1.1 Objectifs de l'expérimentation

L'objectif principal de notre expérimentation consiste à évaluer l'apport pédagogique d'une ressource lexicale faisant appel à la méthode de description lexicale que nous avons proposée (section 4.2). Nous nous attachons à vérifier, au moyen de cette expérimentation, si la description proposée dans le CHINOINFO assiste les apprenants dans la résolution des difficultés syntaxico-sémantiques que les CLS sont susceptibles de soulever (sous-sections 4.1.3.4 et 4.2.1). Notre objectif secondaire consiste à recueillir certains renseignements sur l'impression des apprenants sur le CHINOINFO et son contenu.

Nous faisons passer un test lexical à deux groupes de participants, à savoir le groupe contrôle (GC) et le groupe expérimental (GE). Nous mettons à leur disposition un ensemble d'outils de référence en ligne qu'ils peuvent consulter lors de la passation du test. Seul le GE a accès au dictionnaire CHINOINFO, dont l'élaboration fait appel à notre méthode de description lexicale. Nous comparons les résultats obtenus par les deux groupes de participants. Nous nous attendons à ce que les résultats fournissent des indices sur l'intérêt pédagogique de notre méthode.

Pour mieux interpréter les données recueillies à l'aide du test lexical, nous invitons les participants à remplir deux questionnaires de sondage : 1) des renseignements sur le profil d'apprentissage des participants et 2) les commentaires des participants sur le déroulement du test et l'utilisation des outils de référence proposés. Les données collectées nous permettent de brosser un portrait des participants, d'identifier leurs besoins spécifiques en matière de ressource d'apprentissage du lexique spécialisé et d'adapter davantage notre méthode de description et notre ressource aux besoins identifiés.

4.3.1.2 Profil des participants recherchés

Nous cherchons des participants de langue maternelle française qui reçoivent une formation de chinois dans un établissement universitaire au Québec ou en Chine.

Le profil des participants recherchés a été défini selon plusieurs critères : outre la langue maternelle des participants, nous avons aussi pris en compte leur milieu de formation, leur objectif d'apprentissage et leur connaissance du chinois. En premier lieu, nous avons ciblé les apprenants qui suivent un cursus universitaire, vu que nous nous sommes focalisée sur l'enseignement du lexique spécialisé chinois en milieu universitaire. Au Québec, nous avons ciblé les étudiants inscrits à l'une des trois universités québécoises : Université de Montréal, Université Laval et Université du Québec à Montréal (UQAM). Les étudiants doivent avoir été inscrits à au moins deux cours de chinois qui sont offerts comme cours optionnels dans ces universités. En Chine, tous les étudiants francophones qui suivent la formation du chinois langue seconde⁶⁹ dans une université chinoise sont admissibles.

De plus, les apprenants universitaires poursuivent des objectifs d'apprentissage similaires. D'une part, les étudiants francophones qui séjournent en Chine doivent suivre des cours de chinois préparatoires avant d'entrer dans un programme d'études universitaires (voir aussi la sous-section 2.3.1). D'autre part, selon notre expérience d'enseignement du chinois dans une université du Québec, un bon nombre d'étudiants universitaires cherchent des opportunités de stage en Chine. D'autres apprenants sont motivés par leur engagement professionnel : au terme de leur formation, ils s'attendent à pouvoir communiquer aisément avec leurs collaborateurs ou clients chinois. L'objectif de ces apprenants est d'acquérir rapidement des compétences linguistiques pour pouvoir étudier ou travailler dans un milieu sinophone, ce qui constitue un des traits communs du public cible du COS (sous-section 2.3.1).

Quant à la compétence linguistique des participants recherchés, nous avons sollicité les apprenants débutants qui font preuve d'un niveau élémentaire de connaissance de la langue chinoise. Pour les étudiants inscrits à une université québécoise, ils doivent avoir suivi au moins

⁶⁹ Rappelons que dans notre thèse, l'appellation du *chinois langue seconde* fait référence à l'enseignement-apprentissage du chinois visant tous les apprenants non natifs.

deux cours du niveau élémentaire dans l'une des trois universités, pour avoir accumulé au moins 270 heures d'apprentissage du chinois⁷⁰ :

– Université de Montréal : les cours du niveau A1 (CHN1901 et CHN1902);

– Université Laval : les cours élémentaires (CHN1010 et CHN1020);

– UQAM : les cours CHN1100 et CHN1200

Pour les apprenants qui reçoivent une formation de chinois dans un établissement universitaire en Chine, ils doivent avoir réussi le test HSK (pour *Hanyu Shuiping Kaoshi* « évaluation du niveau de chinois ») niveau 3 (Chinese Testing International Co., Ltd., 2018), ce qui équivaut à environ 250 heures d'apprentissage. Le HSK est un test standardisé de compétence linguistique s'adressant aux apprenants non natifs du chinois. Il vise à évaluer l'habileté des apprenants à communiquer en chinois dans la vie quotidienne, universitaire et professionnelle. Le HSK comprend six niveaux. La réussite du niveau 3 du HSK est en fait de la condition d'admission des universités chinoises.

4.3.1.3 Outils de cueillette des données

Nous avons eu recours à deux types d'outils pour collecter des données auprès des participants : un test lexical (annexe 11) et trois questionnaires de sondage (annexes 12, 13 et 14). Dans ce qui suit, nous présentons la conception de ces instruments.

⁷⁰ Notons que les apprenants ayant réussi deux cours de chinois élémentaire sont admissibles aux cours du niveau intermédiaire offerts par les universités ciblées. Cela dit, ces étudiants, de même que les apprenants ayant réussi le test HSK niveau 3, sont toujours considérés comme les débutants, étant donné que leurs compétences acquises n'atteignent pas le niveau intermédiaire de la maîtrise du chinois.

Le 1^{er} juillet 2021, le ministère de l'Éducation de la République de Chine a mis en œuvre la première norme d'évaluation du niveau de chinois des apprenants non sinophones (Le ministère de l'Éducation de la République populaire de Chine et Commission d'État pour la langue et l'écriture chinoises, 2021). Selon la norme, le niveau de maîtrise de la langue chinoise des apprenants est classé en niveaux élémentaire, intermédiaire et avancé, chaque niveau comprenant trois échelons. La mesure du niveau de chinois repose sur quatre critères d'évaluation : la syllabe, le caractère, le vocabulaire et la grammaire. On distingue trois dimensions d'évaluation : compétences de communication verbale, contenu des discussions thématiques et indicateurs quantitatifs de la langue (nombre de caractères acquis, etc.),

Test lexical

Le test lexical est l'outil que nous avons utilisé pour évaluer les compétences lexicales des participants. Plus précisément, le test a pour objectif d'évaluer l'habileté des participants à résoudre trois types de problèmes lexicaux soulevés par les difficultés syntaxico-sémantiques des CLS en ayant recours à différents types de ressources : 1) distinguer les acceptions d'un vocable dans différentes CLS; 2) différencier les CLS ayant une forme identique et des sens différents; 3) choisir les cooccurrents appropriés pour un terme donné (sous-sections 4.1.3.4 et 4.2.1). Il s'agit des difficultés pour lesquelles le dictionnaire CHINOINFO propose des solutions. Quant au domaine de spécialité abordé, le test porte sur le lexique de l'informatique.

La conception du test repose sur trois principes de base : la congruence, la validité et la fidélité du test. Il s'agit des qualités essentielles de l'évaluation efficace (Arsenault et Giguère, 2020).

Premièrement, la structuration du test vise à assurer la congruence entre les questions et les compétences évaluées. Le test comprend dix questions. Les questions 1-9 sont divisées en trois sections, dont chacune correspond à l'un des trois types de problèmes lexicaux mentionnés ci-dessus. À cela s'ajoute une question libre (question 10) invitant les participants à identifier les exercices qui leur paraissent particulièrement difficiles.

Deuxièmement, les types de questions ont été choisis en fonction des compétences lexicales que nous nous proposons de mesurer. Pour les deux premières sections du test, nous avons eu recours aux questions à réponse courte pour évaluer l'habileté des participants à distinguer les sens associés à une forme lexicale ou à une combinaison lexicale. Les participants doivent non seulement donner une réponse simple qui est affirmative ou négative, mais aussi justifier leurs réponses. Les éléments justificatifs sont pris en compte lors de la correction du test, ce qui nous permet de suivre le raisonnement des participants. Cette démarche vise à assurer la fidélité de l'évaluation, puisque les participants réussissent parfois à deviner certaines bonnes réponses sans posséder les compétences lexicales requises. Pour la dernière section du test portant sur les choix de collocatifs appropriés, nous avons opté pour les questions à choix de réponses multiples. Les participants doivent choisir toutes les bonnes réponses au lieu d'une seule réponse, ce qui permet aussi de réduire l'interférence des réponses devinées sur la mesure des compétences lexicales.

Pour ce qui est de la rédaction de questions, toutes les CLS faisant l'objet d'une question sont présentées dans leur contexte d'utilisation, soit une phrase ou un segment de texte. Ces énoncés sont extraits de notre corpus (voir les sous-sections 4.1.2.4 et 4.1.3.2) et ne sont pas présentés dans le CHINOINFO. Ils reflètent ainsi les usages des CLS dans les discours spécialisés authentiques. Ce choix méthodologique, fondé sur le principe sémasiologique de l'approche lexico-sémantique (sous-section 3.2.1), permet aux participants de saisir les propriétés sémantiques des CLS dans l'environnement linguistique de celles-ci.

Troisièmement, d'autres facteurs ont été pris en compte pour assurer que les résultats de l'évaluation reflètent fidèlement les compétences lexicales des participants. En premier lieu, nous nous sommes attachée à adapter le degré de difficulté de l'évaluation au niveau d'apprentissage des étudiants. Un questionnaire d'appréciation a été élaboré pour recueillir les commentaires des participants à ce propos. Nous y reviendrons plus loin. En deuxième lieu, nous avons utilisé un vocabulaire accessible aux étudiants lors de la formulation des consignes, pour que celles-ci soient claires et compréhensibles. Plus précisément, nous avons évité l'emploi des termes qui renvoient à des notions linguistiques (*unité lexicale, acception, vocable, base et collocatif* d'une *collocation*) ou terminologiques (*terme, combinaison lexicale spécialisée*)⁷¹. En dernier lieu, chaque question est indépendante des autres questions.

Nous illustrons ci-dessous le contenu du test à l'aide d'exemples précis. La section I du test évalue la compétence des apprenants à distinguer les acceptions d'un vocable dans différentes CLS. Considérons l'exemple ci-dessous :

1). 如果 浏览器 无法 打开 网页, 请 首先 检查 网络 连接。
 Si navigateur ne peut pas ouvrir page Web être prié de d'abord vérifier réseau connexion

⁷¹ Dans le CHINOINFO, nous avons eu recours au métalangage pour nommer les différentes rubriques dans un article (par exemple, structure actancielle, combinaison lexicale, etc.). Ces notions ont été expliquées dans la capsule vidéo transmise au GE avant le test (voir l'annexe 19), de même que dans le tutoriel vidéo figurant sur l'interface de recherche du CHINOINFO (voir aussi la sous-section 5.3.1).

« Si votre navigateur n’ouvre pas une page Web, vérifiez d’abord votre connexion au réseau. »
 [Notre traduction]⁷²

2). 2013年 10月, 清华大学 正式 推出 网络 学习 平台,
 année 2013 octobre Université Tsinghua officiellement mettre en œuvre réseau étudier plateforme
 miànxiàng quán qiú tí gòng zài xiàn kè chéng
 orienté vers monde entier offrir en ligne cours

« En octobre 2013, Université Tsinghua a officiellement lancé sa plateforme d’apprentissage en ligne, afin de proposer des cours en ligne aux apprenants du monde entier. » [Notre traduction]

Les participants doivent indiquer si la forme lexicale 网络 (‘réseau’) a le même sens dans les
 CLS 网络连接 (‘réseau-connecter’ « connexion au réseau ») et 网络学习 (‘réseau-étudier’
 « apprentissage en ligne »), présentées quant à celles-ci dans les phrases 1) et 2).

La section II porte sur la distinction sémantique des CLS ayant une forme identique. Par exemple, les participants doivent indiquer si la combinaison lexicale 备份数据 (‘sauvegarder-données’) véhicule le même sens dans les deux énoncés suivants :

1) 为了 保护 计算机 中 的 重要 数据, 您 可以
 Pour protéger ordinateur milieu DE2 important données vous pouvoir
 shǐ yòng bǎo hù jì suàn jī zhōng de zhòng yào shù jù nín kě yǐ
 utiliser Dropbox 云盘 定期 备份 数据。
 utiliser Dropbox disque infonuagique régulièrement sauvegarder données

« Pour protéger les données importantes de votre ordinateur, vous pouvez utiliser le service infonuagique Dropbox pour sauvegarder régulièrement vos données. » [Notre traduction]

2) 要 恢复 数据, 用户 可以 从 iCloud 云端 硬盘
 Pour restaurer données utilisateur pouvoir de iCloud en nuage disque dur

⁷² Notons que les traductions sont proposées ici pour faciliter la compréhension du lecteur. Elles ne font pas partie des énoncés proposés aux participants lors du test lexical (voir l’annexe 11).

| | | | | | |
|---------------|---------------|--------------|----------|--------------|--------------------|
| xià zài 下载 | bèi fèn 备份 | shù jù 数据 | dào 到 | běn dì 本地 | jì suàn jī 计算机。 |
| télécharger | sauvegarder | données | vers | local | ordinateur |

« Pour restaurer les données, les utilisateurs peuvent télécharger les données de sauvegarde d'iCloud sur leur ordinateur local. » [Notre traduction]

Comme on peut le déduire de la traduction proposée, 备份数据 (‘sauvegarder-données’) véhicule deux sens distincts : dans la phrase 1), 备份数据 (‘sauvegarder-données’) est une combinaison verbale signifiant ‘sauvegarder les données’; dans la phrase 2), cette combinaison lexicale est à base nominale qui signifie ‘données sauvegardées’. Cet exercice mesure l’habileté des participants à différencier les structures syntaxiques associées à la combinaison lexicale en question dans différents contextes.

La section III concerne le choix de collocatifs. Les participants doivent compléter une phrase à trou en utilisant deux des quatre unités lexicales proposées. Examinons l’exemple ci-dessous :

| | | | | | | | | | | |
|----------------|---------------------|------------|---------------|-------------|----------------|---------------------|--------|-------------------|-----|-----------------|
| lián jiē 连接 | hù lián wǎng 互联网 | hòu, 后, | yòng hù 用户 | kě yǐ 可以 | shǐ yòng 使用 | shǒu jī duān 手机端 | Chrome | liú lǎn qì 浏览器 | [] | wǎngzhàn 网站。 |
| Connecter | Internet | après | utilisateur | pouvoir | utiliser | terminal mobile | Chrome | navigateur | [] | site Web |

« Une fois connecté à l’Internet, l’utilisateur peut _____ un site Web en utilisant le navigateur Chrome sur le téléphone mobile. » [Notre traduction]

- | | |
|-------------------|---------------------|
| a) fǎng wèn 访问 | Accéder (à) |
| b) huò qǔ 获取 | Obtenir |
| c) rù qīn 入侵 | S’introduire (dans) |
| d) dēng lù 登陆 | Se rendre (sur) |

Les quatre verbes proposés sont sémantiquement proches : ils peuvent tous être traduits en français par *accéder (à)*. Cet exercice évalue l’aptitude des participants à identifier les collocatifs contrôlés par un terme de base, qui peuvent s’insérer dans une configuration syntaxique donnée.

Outre le test lexical, nous avons conçu trois questionnaires de sondage dans le but de recueillir les renseignements sur le profil des participants et leurs commentaires sur le test et les outils de référence proposés.

Questionnaire 1 : Profil de participants

Dans le domaine de recherche en L2, les questionnaires, tout comme les tests de compétences linguistiques, constituent les instruments de collecte de données les plus courantes (Dörnyei et Taguchi, 2009, p. xiii). Cet outil de recherche présente plusieurs avantages, comme le soulignent Dörnyei et Taguchi (2009, p. xiii) : il est facile à construire, polyvalent et permet de recueillir de façon efficace une grande quantité d'informations qui peuvent être facilement exploitées. Précisons que les questionnaires se définissent comme « les instruments écrits qui présentent aux répondants une série de questions ou d'énoncés auxquels ils doivent réagir soit en rédigeant leurs réponses, soit en choisissant parmi les réponses existantes » [notre traduction] (Brown, 2001, p. 6).

Les questionnaires se distinguent des tests qui visent à mesurer l'atteinte de certains objectifs d'apprentissage ou le niveau de compétences. Il s'agit d'un instrument de collecte de données qui sert à recueillir trois types de données, soit les données *factuelles*, *comportementales* et *attitudinales* (Dörnyei et Taguchi, 2009, p. 5-6). Nous y reviendrons plus loin lorsqu'il sera question de présenter le contenu de nos questionnaires.

Dans le cadre de notre recherche, nous avons élaboré deux sondages. Nous avons d'abord élaboré un questionnaire portant sur le profil des participants (annexe 12). Ce questionnaire comprend 12 questions qui se divisent en deux grandes parties : les questions factuelles et comportementales.

Les questions factuelles (questions 1-7 et 10-12) portent sur les expériences vécues par les répondants, nous avons collecté, plus précisément, cinq types de renseignements : 1) le milieu universitaire des participants, y compris leur affiliation universitaire, leur programme d'études et leur niveau d'études universitaires; 2) leur expérience d'apprentissage du chinois, y compris le(s) milieu(x) où les participants ont reçu leur formation en chinois (école secondaire, université, domicile, entreprise, etc.) et le nombre d'heures consacrées à l'apprentissage du chinois; 3) leur expérience d'apprentissage du chinois sur objectifs spécifiques (COS), s'il y a lieu; 4) leur expérience de formation ou de travail dans le domaine de l'informatique et 5) leur séjour en Chine, s'il y a lieu.

Les questions comportementales (questions 8 et 9) portent sur l'habitude d'apprentissage des répondants, y compris l'utilisation des ressources didactiques du chinois et les difficultés qu'ils ont éprouvées lorsqu'ils cherchaient à résoudre des questions dans les ressources lexicales.

L'analyse des données collectées permet de dresser un portrait des participants du point de vue de leur parcours éducationnel et professionnel, de leur expérience de vie dans un milieu sinophone et de leur habitude d'apprentissage. Les résultats de ce sondage révèlent en outre des facteurs qui sont susceptibles d'influencer le développement des compétences lexicales chez les participants.

Questionnaire 2a : Appréciations et commentaires (Groupe contrôle)

Le second sondage comprend deux questionnaires : le questionnaire 2a s'adresse au GC (annexe 13); le questionnaire 2b s'adresse au GE (annexe 13). Ces questionnaires, composés d'une variété de questions attitudinales et comportementales, visent à recueillir les opinions des participants sur le test lexical et leur degré de satisfaction à l'égard des outils de référence utilisés au cours du test.

Le questionnaire 2a (annexe 13) comprend trois sections, portant respectivement sur l'organisation du test (I), le contenu du test (II) et le recours aux ressources lexicales (III). La section I est composée de 7 questions à échelle d'évaluation. Nous avons opté pour une échelle de Likert, un outil psychométrique couramment utilisé dans les recherches en L2 (Dörnyei et Taguchi, 2009, p. 27-29) pour mesurer les attitudes. Les répondants sont invités à indiquer leur degré d'accord envers chacun des 7 énoncés sur une échelle de 1 à 5 : 1 = fortement en désaccord; 2 = plutôt en désaccord; 3 = plutôt en accord; 4 = fortement en accord; 5 = ne s'applique pas. Les questions 1.1 à 1.4 portent sur l'encadrement des participants avant le test, par exemple : la documentation transmise aux participants était-elle utile; les objectifs et les modalités d'évaluation ont-ils été expliqués de façon claire, les tâches préparatoires étaient-elles utiles pour comprendre le fonctionnement des outils de référence proposés ? Les questions 1.5 à 1.7 mesurent les attitudes des répondants sur le déroulement du test, y compris la durée du test, la communication avec la chercheuse au cours du test et les supports informatiques utilisés (capsule vidéo, fichier PDF et Word).

La section II contient 4 questions à échelle d'évaluation. Ces questions sont conçues à l'aide d'une échelle sémantique différentielle, un outil psychométrique qui s'utilise souvent pour

recueillir l'appréciation des répondants (Dörnyei et Taguchi, 2009, p. 30-31). Dans la section II, les répondants sont invités à porter un jugement sur le niveau de difficulté du test et à indiquer leurs réponses sur une échelle de 1 à 4 : 1 = très difficile; 2 = plutôt difficile; 3 = plutôt facile; 4 = très facile. Les questions 2.1 à 2.3 portent respectivement sur les trois sections du test, celles-ci correspondant aux types de difficultés syntaxico-sémantiques que les CLS sont susceptibles de soulever (sous-sections 4.1.3.4 et 4.2.1). La question 2.4 porte sur le niveau de difficulté du test en général.

La section III collecte une variété de renseignements sur le recours aux outils de référence. Nous demandons aux répondants d'identifier les ressources lexicales qu'ils ont utilisées au cours du test. Ils doivent aussi préciser les motifs qui les ont poussés à consulter une ressource lexicale, par exemple, pour trouver le(s) sens d'un mot inconnu ou choisir un cooccurrent approprié d'un mot. Quant aux résultats de leur recherche, les répondants doivent indiquer, sur une échelle de 1 à 5, le niveau de difficulté qu'ils ont éprouvée pour trouver des éléments de réponses à chaque type de problèmes lexicaux dans les outils de référence utilisés : 1 = très difficile; 2 = plutôt difficile; 3 = plutôt facile; 4 = très facile, 5 = ne s'applique pas (c'est-à-dire les répondants n'ont trouvé aucun élément de réponse dans les ressources lexicales consultées). Les répondants expriment ensuite leur degré de satisfaction globale envers l'utilisation des ressources lexicales sous cinq aspects : 1) la convivialité des ressources; 2) la qualité du contenu des ressources consultées; 3) la méthode de description lexicale; 4) l'utilité des ressources pour résoudre les problèmes lexicaux et 5) l'efficacité pédagogique. Pour aider les répondants à comprendre les cinq aspects à apprécier, nous avons expliqué chaque aspect à l'aide d'un énoncé qui sert à orienter la réflexion des répondants. Par exemple, pour porter un jugement sur la *méthode de description* des ressources lexicales, les répondants doivent indiquer si *les connaissances lexicales sont présentées de façon claire et structurée* dans ces ressources (voir aussi l'annexe 13). À la fin du questionnaire, nous avons élaboré une question ouverte pour inciter les répondants à nous faire part d'autres commentaires jugés pertinents.

Questionnaire 2b : Appréciations et commentaires (Groupe expérimental)

Le questionnaire 2b (annexe 14) s'adresse au groupe expérimental qui a accès au dictionnaire CHINOINFO lors du test. Il est composé de quatre sections. Les sections I et II sont identiques

pour les questionnaires 2a et 2b. Elles portent sur l'organisation du test et le niveau de difficulté du test.

Dans le questionnaire 2b, les sections III et IV portent respectivement sur l'utilisation du CHINOINFO et le recours à d'autres ressources lexicales.

La section III du questionnaire 2b porte sur l'utilisation du CHINOINFO. On y trouve d'abord deux séries de questions à échelle d'évaluation. Les répondants doivent indiquer le niveau de difficulté qu'ils ont éprouvée pour trouver des éléments de réponses à chaque type de problèmes lexicaux dans le CHINOINFO. Ils apprécient ensuite l'utilisation du CHINOINFO sous cinq aspects (la convivialité, la qualité du contenu, la méthode de description lexicale, l'utilité et l'efficacité pédagogique), tout comme le GC qui commente l'utilisation des outils de référence mis à sa disposition. Nous avons aussi élaboré une question ouverte pour inciter les répondants à exprimer ce qu'ils aiment le plus et le moins dans le CHINOINFO et à nous faire part de leurs propositions pour améliorer le dictionnaire. À la fin de cette section, nous avons conçu une question pour avoir une idée globale de l'intérêt que notre méthode de description lexicale a suscité auprès des répondants.

La section IV porte sur le recours à d'autres ressources lexicales. On y trouve presque toutes les questions de la section III du questionnaire 2a, sauf la question 6 (annexe 13) : les répondants n'ont pas à indiquer leur degré de satisfaction globale envers l'utilisation des ressources lexicales autre que le CHINOINFO.

Les données recueillies au moyen des instruments élaborés sont ensuite analysées. Nous reviendrons sur la démarche méthodologique adoptée (sous-section 4.3.3) et les résultats d'analyse (section 6.2).

4.3.1.4 Recrutement des participants

Nous avons entrepris les démarches requises auprès du Comité d'éthique de la recherche en arts et humanités (CERAH) pour obtenir l'approbation éthique avant de lancer le recrutement des

participants⁷³. Le processus de recrutement a entièrement eu lieu en ligne. Nous avons adopté deux démarches différentes pour solliciter les apprenants du chinois au Québec et en Chine.

Pour solliciter des étudiants des trois universités québécoises que nous avons pressenties (sous-section 4.3.1.2), nous avons fait circuler l'annonce de recrutement via deux moyens de diffusion :

1) À l'Université de Montréal, nous avons lancé une invitation au sein du microprogramme en langue et culture chinoises. Nous avons conclu une entente de collaboration avec le responsable du microprogramme, qui a accepté de transmettre l'annonce par liste d'envois électroniques à tous les étudiants inscrits aux cours de chinois offerts par le Centre de langues de l'Université de Montréal.

2) Nous avons diffusé l'annonce de recrutement via les réseaux sociaux pour recruter les étudiants des deux autres universités québécoises (Université Laval et UQAM). L'annonce de recrutement a été publiée sur la page Facebook de l'École de langues de l'Université Laval⁷⁴ et la page Facebook des programmes et cours de langue et culture chinoises de l'École de langues de l'UQAM⁷⁵.

Pour solliciter les étudiants francophones en Chine, nous avons lancé l'invitation à l'expérimentation via les médias sociaux chinois. Nous avons eu recours à deux moyens de diffusion : 1) la plateforme *WeChat* qui est le réseau social le plus utilisé en Chine; 2) un forum public intitulé *BonjourChine*⁷⁶. Ce forum offre une plateforme d'échanges et de discussion pour les apprenants francophones du chinois, notamment ceux qui font leur séjour en Chine.

⁷³ Notre projet de recherche (projet no CERAH-2021-142-D) a été approuvé par le Comité d'éthique de la recherche en arts et humanités de l'Université de Montréal en date du 12 janvier 2021. Cette approbation éthique est valide jusqu'au 12 janvier 2023. L'attestation d'approbation d'éthique est jointe en annexe 15.

⁷⁴ La page Facebook de l'École de langues de l'Université Laval est repérable à l'adresse suivante : <https://www.facebook.com/EcoleLanguesUniversiteLaval/>.

⁷⁵ La page Facebook des programmes et cours de langue et culture chinoises de l'École de langues de l'UQAM est repérable à l'adresse suivante : <https://www.facebook.com/LangueCultureChinoises.UQAM/>.

⁷⁶ Le forum BonjourChine est repérable à l'adresse suivante : <https://www.bonjourchine.com/threads/annonce-de-recrutement-apprenants-du-chinois-recherches-pour-un-projet-de-these.152623/>.

Les apprenants intéressés à participer à l'expérimentation devaient nous contacter par courriel. Un formulaire d'information et de consentement leur a été envoyé par la suite pour signature (voir l'annexe 16).

4.3.2 Protocole d'expérimentation

Notre expérimentation s'est déroulée entièrement en ligne en trois temps : 1) le regroupement des participants et la fixation de l'horaire du test, 2) la collecte de données auprès des participants et 3) l'analyse des données recueillies.

4.3.2.1 Planification des séances expérimentales

Dans un premier temps, les participants recrutés ont été affectés de façon aléatoire à deux groupes à nombre égal (GC et CE). Le test lexical a pris la forme d'une séance en temps réel sur Zoom, d'une durée maximale de 60 minutes. Compte tenu du décalage horaire possible et de la disponibilité variée des participants, il était difficile de fixer un horaire qui convenait à tous les participants. Pour cette raison, nous avons proposé une plage horaire au lieu d'une heure fixe (du 9 mai ou 13 mai 2022, de 9 h à 12 h et de 13 h à 16 h), pour permettre à chaque participant de choisir la date et l'heure qui lui convenait le mieux pour passer le test. En tant que chercheuse et coordinatrice de l'expérimentation, nous avons assisté à toutes les sessions Zoom.

4.3.2.2 Collecte de données

La deuxième phase de l'expérimentation, à savoir la cueillette de données auprès du GC et du GE, s'est réalisée en trois étapes : la réalisation des tâches préparatoires, la passation du test lexical et les sondages par questionnaire.

Collecte de données auprès du groupe contrôle (GC)

Dans un premier temps, nous avons proposé deux tâches préparatoires aux participants du GC. La première tâche consiste à visionner une capsule vidéo descriptive du test lexical, d'une durée de 10 minutes. Il s'agit d'une présentation narrée (annexe 17), soit l'enregistrement d'une présentation PowerPoint accompagnée d'explications verbales. Nous avons communiqué les renseignements suivants aux participants à l'aide de ce document : le contexte de recherche,

l'objectif du test lexical, les habiletés évaluées ainsi que les types de questions correspondant et le nombre de questions, le déroulement du test et les consignes.

La deuxième tâche consiste à prendre connaissance d'un document descriptif portant sur les outils de référence que les participants peuvent utiliser au cours du test (annexe 18). Dans ce document, nous avons proposé une liste de ressources lexicales en ligne. On y trouve deux dictionnaires chinois-français (*Dictionnaire Chine-Nouvelle* et *Dictionnaire Chine-informations*), un dictionnaire français-chinois (*Larousse Français-chinois*), deux dictionnaires chinois-anglais (*LINE Chinese-English Dictionary* et *MDBG Chinese-English Dictionary*) et une banque de terminologie chinois-anglais (*Termonline*). Les cinq dictionnaires proposés relèvent de la langue générale, il s'agit des outils de référence les plus couramment utilisés par les apprenants francophones selon nos observations dans la pratique d'enseignement. La banque de terminologie *Termonline* a été proposée comme la seule ressource en langue de spécialité. À notre connaissance, il n'existe pas encore de dictionnaire spécialisé bilingue en ligne qui traite du lexique chinois de l'informatique. La liste complète des références est présentée dans l'annexe 18. La liste est accompagnée d'un mode d'emploi de certaines ressources proposées : nous avons expliqué, à l'aide des captures d'écran annotées, comment effectuer une recherche à partir de l'interface de consultation de ces ressources. Ces documents, ainsi que le lien de connexion Zoom, ont été envoyés par courriel aux participants une semaine avant la passation du test.

Les participants du GC ont ensuite complété le test en temps réel sur Zoom à la date et l'heure qu'ils ont choisies. Une fois connectés à leur session Zoom, les participants ont reçu trois documents sous format Word : le test lexical (annexe 11), le questionnaire 1 (annexe 12) et le questionnaire 2a (annexe 13). Nous avons demandé aux participants de compléter d'abord le test lexical en inscrivant leurs réponses dans le même document. Ils étaient autorisés à utiliser les outils de référence proposés, les outils de traduction automatique et d'autre matériel didactique de leur choix, y compris des manuels et des notes de cours. Le temps alloué pour le test est de 30 minutes. Les sessions Zoom n'ont pas été enregistrées. Les participants ont ouvert leur caméra durant le test, même si cela n'était pas exigé. Nous avons assisté à chaque séance du test en tant qu'hôtesse de la session Zoom, afin de répondre immédiatement aux questions des participants.

Une fois le test réalisé, les participants ont rempli les deux questionnaires de sondage visant à recueillir les renseignements sur leur profil et leurs commentaires sur le test et les outils de référence proposés. À la fin de leur séance Zoom, ils devaient nous remettre le test lexical complété et les questionnaires remplis. La plupart des participants ont réalisé leurs tâches en environ 40 minutes : 30 minutes pour compléter le test et 10 minutes pour remplir les questionnaires de sondage. Nous avons alloué un temps supplémentaire pour les imprévus, notamment en ce qui concerne l'envoi de documents par voie électronique et la connexion à Internet. La durée maximale d'une séance Zoom est de 60 minutes.

Collecte de données auprès du groupe expérimental (GE)

La collecte des données auprès du GE diffère sur trois points de la démarche que nous avons entreprise auprès du GC : les documents remis aux participants avant le test, les outils de référence autorisés et le questionnaire d'appréciation à remplir.

Avant de passer le test, le GE devait réaliser les mêmes tâches préparatoires que le GC, c'est-à-dire visionner une capsule vidéo et prendre connaissance des outils de référence mis à leur disposition. Comme le GE était invité à utiliser le CHINOINFO lors du test lexical, nous avons fourni des renseignements supplémentaires dans les deux documents préparatifs pour aider le GE à se familiariser avec l'utilisation de cette ressource lexicale. Premièrement, la capsule vidéo destinée au GE (annexe 19) comprend deux grandes parties : outre les renseignements sur le test lexical que l'on trouve dans le document destiné au GC, nous avons préparé un guide d'utilisation du dictionnaire CHINOINFO. Il s'agit d'un tutoriel vidéo, dans lequel on présente les modes de recherche du dictionnaire, la structuration des contenus dans un article et les fonctionnalités de navigation dans le dictionnaire. Nous avons également créé un lien vers ce tutoriel sur l'interface de recherche du CHINOINFO (voir aussi la sous-section 5.3.1). Quant au document descriptif des ressources lexicales proposées (annexe 20), nous y avons intégré le lien d'accès au CHINOINFO. Une mention spéciale y a aussi été ajoutée pour inciter les participants à utiliser le CHINOINFO comme outil de référence principal, car notre objectif secondaire consiste à recueillir les commentaires du GE sur leur utilisation du CHINOINFO. Outre le CHINOINFO, le GE était autorisé à utiliser les outils de référence proposés, les outils de traduction automatique et d'autre matériel didactique, tout comme le GC.

La passation du test lexical suit la même procédure que celle appliquée au GC. Les deux groupes devaient compléter le même test lexical. Seul le GE a eu accès au CHINOINFO. En ce qui concerne le sondage par questionnaire, tous les deux groupes de participants ont rempli le questionnaire 1 portant sur leur profil (annexe 12). Le GE a répondu au questionnaire 2b (annexe 14), ce dernier vise à recueillir les commentaires des participants sur leur utilisation du CHINOINFO (voir aussi la sous-section 4.3.1.3).

4.3.2.3 Analyse des données

Une fois la collecte de données terminée, nous avons procédé à l'analyse des données quantitatives et qualitatives recueillies.

Notre démarche pour traiter les données collectées au moyen du test lexical a été réalisée en quatre étapes. Dans un premier temps, nous avons corrigé les réponses aux questions 1-9 du test, pour analyser par la suite les erreurs repérées en nous inspirant de la méthodologie d'analyse des erreurs (AE) proposée par Corder (1967, 1981) en enseignement-apprentissage de langue étrangère. Ensuite, nous avons synthétisé les réponses à la question 10 qui est une question libre portant sur le niveau de difficulté des exercices (voir aussi la sous-section 4.3.1.3). Enfin, nous avons comparé les résultats obtenus par le GC et le GE selon un ensemble de paramètres (notes obtenues, types des erreurs et fréquence des erreurs, etc.). Nous y reviendrons dans la section 6.1.

Le traitement des réponses aux sondages (questionnaires 1, 2a et 2b) a été effectué en fonction de la forme des questions, à savoir les questions à choix de réponses, les questions à échelle d'évaluation et les questions à réponse courte. Pour chaque type de question, nous avons synthétisé les réponses selon les types de renseignements recueillis. L'interprétation des données synthétisées permet de dresser un portrait du profil des participants, de recueillir les commentaires des participants sur l'utilisation de la ressource lexicale et d'identifier, par le fait même, d'autres facteurs qui sont susceptibles d'influencer la performance des participants.

Les résultats d'analyse des données expérimentales ont fourni des indices sur l'apport pédagogique du CHINOINFO, et l'intérêt pédagogique d'une partie de la méthode de

description lexicale proposée. Nous décrivons en détail les résultats de l'expérimentation dans le chapitre 6.

4.3.3 Synthèse

Dans la section 4.3, nous avons décrit la démarche méthodologique pour réaliser la dernière phase de notre recherche qui consiste à évaluer l'apport pédagogique du CHINOINFO, dont l'élaboration est basée sur la méthode de description lexicale proposée. Pour réaliser cet objectif, nous avons conçu une expérimentation comparative auprès d'apprenants francophones provenant d'un établissement universitaire au Québec ou en Chine (4.3.1). Nous avons fait passer une évaluation des compétences lexicales en ligne à deux groupes de participants (le GC et le GE), afin d'évaluer leurs aptitudes à résoudre trois types de problèmes lexicaux soulevés par les CLS. Seul le GE a eu accès au dictionnaire CHINOINFO. Chaque groupe de participants a rempli deux questionnaires de sondage visant à recueillir les renseignements sur leur profil et leur appréciation sur le test et les outils de référence proposés. Nous avons présenté la collecte de données auprès des deux groupes et l'analyse des données recueillies (4.3.2). Les résultats du traitement des données expérimentales seront abordés dans le chapitre 6.

Chapitre 5 Présentation du CHINOINFO

Ce chapitre fournit une description du dictionnaire CHINOINFO, ressource lexicale élaborée à l'aide de la méthode de description lexicale que nous avons proposée⁷⁷. La section 5.1 donne un aperçu de la structure générale du dictionnaire. Nous y présentons la nomenclature du dictionnaire et l'organisation des articles dans le CHINOINFO. La section 5.2 présente la structuration des données terminologiques au sein des articles. La section 5.3 explique les fonctionnalités de l'interface du CHINOINFO.

5.1 Nomenclature du CHINOINFO

Le CHINOINFO répertorie 91 termes fondamentaux du domaine de l'informatique et chacun est décrit dans un article distinct. Les articles sont classés en 33 vocables lors de l'analyse des données terminologiques⁷⁸. Chaque vocable est encodé dans un fichier XML distinct lors de la rédaction des articles (voir aussi la sous-section 4.2.2). Dans l'interface du CHINOINFO, les articles relevant d'un vocable sont présentés séparément. Ils se succèdent dans la liste des articles grâce à leur forme identique. Chaque article est assorti d'un numéro d'acceptation. Ce dernier est attribué en fonction de la partie du discours à laquelle le terme vedette appartient (voir aussi la nomenclature du CHINOINFO illustrée dans la figure 41).

Dans le CHINOINFO, la grande majorité des vocables sont des mots multicatégoriels, à savoir des polysèmes qui regroupent les termes relevant de différentes parties du discours⁷⁹. Parmi les 33 vocables du CHINOINFO, on compte 30 mots multicatégoriels et 3 polysèmes qui réunissent des termes appartenant à la même partie du discours, comme l'illustre le tableau 33.

⁷⁷ L'élaboration de la version en ligne du CHINOINFO a été réalisée à l'aide de Benoit Robichaud, membre de l'Observatoire de linguistique Sens-Texte (OLST). Nous le remercions chaleureusement pour sa contribution à différentes étapes de réalisation du CHINOINFO, notamment la création de l'interface de recherche, de même que la conception de la mise en page des articles en ligne.

⁷⁸ Voir aussi la sous-section 4.1.2.4 portant sur l'analyse des UT polysémiques. La liste complète des termes regroupés en vocables est présentée dans l'annexe 6.

⁷⁹ Le tableau 15 présente les types de mots multicatégoriels relevés à partir de notre corpus et la répartition de ceux-ci.

| Type de vocable | | | | | Nombre de vocables |
|-----------------|--|--|--|--|--------------------|
| | Acception 1 | Acception 2 | Acception 3 | Acception 4 | |
| Type 1 | Verbe | Nom d'action | | | 9 |
| Exemple | chuán shū 传输 ₁ « transférer » | chuán shū 传输 ₂ « transfert » | | | |
| Type 2 | Verbe | Nom d'action | Adj | | 2 |
| Exemple | gòng xiǎng 共享 ₁ « partager » | gòng xiǎng 共享 ₂ « partage » | gòng xiǎng 共享 ₃ « partagé » | | |
| Type 3 | Verbe | Nom d'action | Adj | Adv | 1 |
| Exemple | tóng bù 同步 ₁ « synchroniser » | tóng bù 同步 ₂ « synchronisation » | tóng bù 同步 ₃ « synchrone » | tóng bù 同步 ₄ « synchroniquement » | |
| Type 4 | Verbe | Nom d'action | Nom de résultat | | 3 |
| Exemple | shè zhì 设置 ₁ « configurer » | shè zhì 设置 ₂ « configuration ₁ » | shè zhì 设置 ₃ « configuration ₂ » | | |
| Type 5 | Verbe | Nom d'action | Nom de résultat | Adj | 1 |
| Exemple | bèi fèn 备份 ₁ « sauvegarder » | bèi fèn 备份 ₂ « sauvegarde ₁ » | bèi fèn 备份 ₃ « sauvegarde ₂ » | | |
| Type 6 | Verbe 1 | Verbe 2 | Nom d'action 1 | Nom d'action 2 | 6 |
| Exemple | jiǎn cè 检测 ₁ « détecter » | jiǎn cè 检测 ₂ « vérifier » | jiǎn cè 检测 ₃ « détection » | jiǎn cè 检测 ₄ « vérification » | |
| Type 7 | Nom 1 | Nom 2 | Verbe 1 | Verbe 2 | 1 |
| Exemple | yìng yòng 应用 ₁ « application ₁ » | yìng yòng 应用 ₂ « application ₂ » | yìng yòng 应用 ₃ « s'appliquer » | yìng yòng 应用 ₄ « appliquer ₂ » | |
| Type 8 | Adj | Adv | | | 6 |
| Exemple | yuǎn chéng 远程 ₁ « distant » | yuǎn chéng 远程 ₂ « à distance » | | | |
| Type 9 | Nom | Adj | Adv | | 1 |
| Exemple | ān quán 安全 ₁ « sécurité » | ān quán 安全 ₂ « sécurisé » | ān quán 安全 ₃ « sécuritairement » | | |
| Type 10 | Adj 1 | Adj 2 | | | 1 |
| Exemple | yí dòng 移动 ₁ « mobile » | yí dòng 移动 ₂ « amovible » | | | |
| Type 11 | Nom 1 | Nom 2 | | | 1 |
| Exemple | ruǎn jiàn 软件 ₁ « logiciel ₁ » | ruǎn jiàn 软件 ₂ « logiciel ₂ » | | | |
| Type 12 | Nom 1 | Nom 2 | Adj | | 1 |
| Exemple | wǎng luò 网络 ₁ « réseau ₁ » | wǎng luò 网络 ₂ « réseau ₂ » | wǎng luò 网络 ₃ « en ligne » | | |

Tableau 33. – Types de vocables dans le CHINOINFO

Dans le CHINOINFO, nous avons encodé au total 282 termes reliés et 644 CLS pour les 91 termes répertoriés, comme l'illustrent les annexes 23 et 24.

5.2 Structure interne des articles

Dans le CHINOINFO, l'encodage des données terminologiques s'est largement inspiré du DiCoInfo (L'Homme et Alipour, 2009; Observatoire de linguistique Sens-Texte, 2022b), comme nous l'avons expliqué dans la sous-section 4.2.2. Lors de la mise en ligne du CHINOINFO, nous nous sommes également inspirée de l'adaptation du DiCoInfo à un dictionnaire d'apprentissage (Alipour, 2014), étant donné que la version repensée du dictionnaire présente des similitudes avec le CHINOINFO en ce qui concerne les types d'utilisateurs et les fonctions lexicographiques (Tarp, 2008, 2010).

Premièrement, le CHINOINFO, tout comme le DiCoInfo, est destiné aux apprenants de la langue de l'informatique, qui ne possèdent qu'une connaissance de base du domaine abordé. Deuxièmement, les deux dictionnaires se veulent être des ressources lexicales qui assistent les utilisateurs dans la compréhension et la production des textes spécialisés dans les situations communicatives⁸⁰ ainsi que dans l'acquisition des connaissances spécialisées véhiculées par les termes dans les situations cognitives (Alipour, 2014, p. 113). Nous nous sommes appuyée sur ce modèle de départ pour envisager une présentation claire et structurée des connaissances lexicales. Quelques adaptations ont été réalisées en ce qui concerne la présentation des liens lexicaux dans le CHINOINFO. Nous les abordons dans les sous-sections 5.2.3 et 5.2.4.

Nous présentons, dans ce qui suit, la structure interne des articles dans la version en ligne du CHINOINFO. Au sein d'un article, les données terminologiques sont présentées dans quatre sections : les renseignements généraux (5.2.1), la définition et la structure actancielle (5.2.2), les contextes (5.2.3), ainsi que les termes reliés et les combinaisons lexicales (5.2.4).

⁸⁰ En lexicographie fonctionnelle (Bergenholtz et Tarp, 2003; Tarp, 2005, 2008, 2010), on distingue deux catégories de situations d'utilisation de dictionnaires : 1) les situations communicatives, dans lesquelles où les utilisateurs consultent des dictionnaires pour résoudre des problèmes dans la lecture, la rédaction ou la traduction de textes spécialisés; 2) les situations cognitives, dans lesquelles les utilisateurs souhaitent obtenir un complément d'informations sur un domaine de spécialité.

5.2.1 Renseignements généraux

Pour aider les apprenants à distinguer rapidement les différentes acceptions d'un vocable, nous offrons un accès rapide aux renseignements de base de chaque terme, comme le fait DiCoInfo (L'Homme et Alipour, 2009). Ces renseignements figurent en haut de chaque article (voir aussi l'annexe 21). Les utilisateurs peuvent les consulter sans avoir à parcourir tout l'article qui pourrait être très long. La figure 34 illustre l'affichage des renseignements généraux de l'article ^{wǎng luò}网络₁ (« réseau₁ »). On y trouve l'entrée assortie d'un numéro d'acception, l'indication de la partie du discours, la mention du domaine de spécialité et l'équivalent français du terme. Si l'équivalent fait l'objet d'un article dans le DiCoInfo, on trouve un hyperlien qui donne accès à l'article du DiCoInfo.

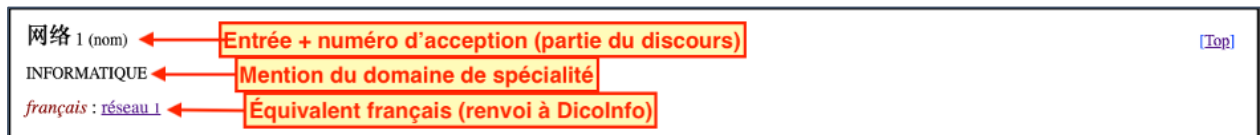


Figure 34. – Affichage des renseignements généraux du terme ^{wǎng luò}网络₁ (« réseau₁ »)

5.2.2 Définition et structure actancielle

Dans un article du CHINOINFO, les autres renseignements descriptifs du terme en entrée sont organisés en différentes rubriques. Celles-ci sont présentées sous forme d'onglets qui s'affichent sur trois rubans figurant en dessous des renseignements généraux du terme (voir aussi l'annexe 21).

Notre choix de présenter les données terminologiques au moyen des rubans et des onglets a été inspiré de l'adaptation du DiCoInfo à un dictionnaire d'apprentissage (Alipour, 2014). Ce mode d'affichage permet aux utilisateurs de consulter le contenu qui les intéresse facilement et rapidement, en activant ou désactivant les onglets selon leurs besoins. Cela leur évite d'avoir à lire l'article au complet pour trouver des éléments de réponse à un problème lexical précis. Par ailleurs, ce mode d'affichage rend les pages d'articles aérées et claires, comme le souligne Alipour (2014, p. 114).

Le premier ruban comprend les rubriques « Définition » et « Structure actancielle », tout comme le DiCoInfo. La définition s'affiche dans un onglet actif par défaut. On indique la source d'une

définition à l'aide d'un code si celle-ci est citée telle quelle, comme le montre la figure 35. La structure actancielle est présentée dans un onglet inactif, tout comme les onglets figurant sur les deux autres rubans. Pour consulter le contenu des onglets masqués, les utilisateurs n'ont qu'à cliquer sur ces onglets pour les activer.

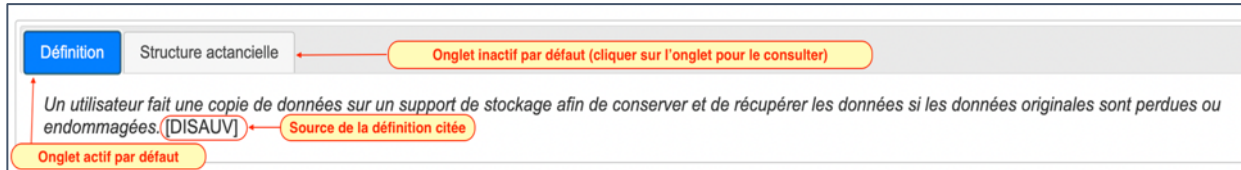


Figure 35. – Ruban affichant « Définition » (l'onglet actif) et « Structure actancielle » (onglet inactif) de 备份₁ (« sauvegarder »)

La structure actancielle énonce les actants sémantiques du terme en entrée. Inspirée du DiCoInfo, nous représentons les actants sémantiques par des termes typiques qui correspondent à des réalisations linguistiques de ces actants. Le rôle des actants par rapport au prédicat sémantique, soit le terme décrit, est également mis en évidence. Par exemple, le terme 备份₁ (« sauvegarder ») fait appel à trois actants : l'agent qui fait l'action de sauvegarder, le patient qui subit l'action et la destination visée par l'action. La structure actancielle du terme est présentée comme suit :

| | | | | |
|-----------------|----------|---------------------------|----------------------------|----------------|
| [Agent] | | [destination] | prédicat | [Patient] |
| yòng hù [用户] | zài 在 | cún chū jiè zhì [存储介质] | bèi fèn 备份 ₁ | shù jù [数据] |
| Utilisateur | sur | support de stockage | sauvegarder | données |

Dans la structure actancielle du terme (figure 36), les termes typiques qui réalisent ces actants sont affichés en trois différentes couleurs. Quand on pointe le curseur sur un terme typique (用户 « utilisateur »), une étiquette jaune apparaît pour indiquer son rôle actanciel (« Agent »). On voit aussi une étiquette bleue qui indique l'équivalent français de ce terme typique.

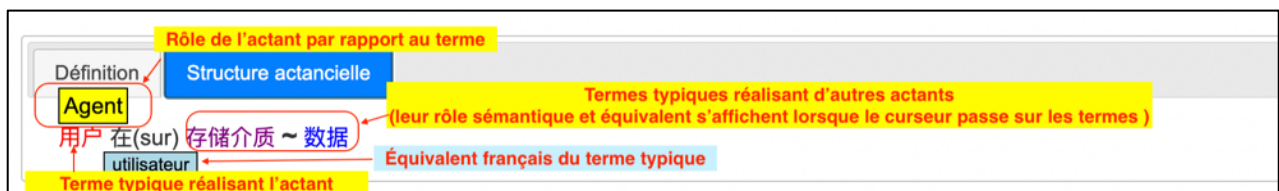


Figure 36. – Ruban affichant « Définition » (l'onglet inactif) et « Structure actancielle » (onglet actif) de 备份₁ (« sauvegarder »)

Les termes typiques qui réalisent un actant sémantique sont présentés sous forme d'une liste. Le terme 入侵₁ (« s'introduire »), à titre d'exemple, fait appel à deux actants : l'agent et la destination. L'actant « Destination » de 入侵₁ (« s'introduire ») est réalisé par trois termes : 计算机 (« ordinateur»), 网络 (« réseau ») et 信息系统 (« système informatique ») dénotant une entité visée par l'action d'intrusion. Seul le premier terme typique (计算机 « ordinateur ») est affiché par défaut dans la structure actancielle (figure 37). Les utilisateurs doivent pointer sur l'icône verte en forme de croix qui s'affiche à côté de ce terme pour avoir accès à la liste complète des réalisations linguistiques de l'actant.

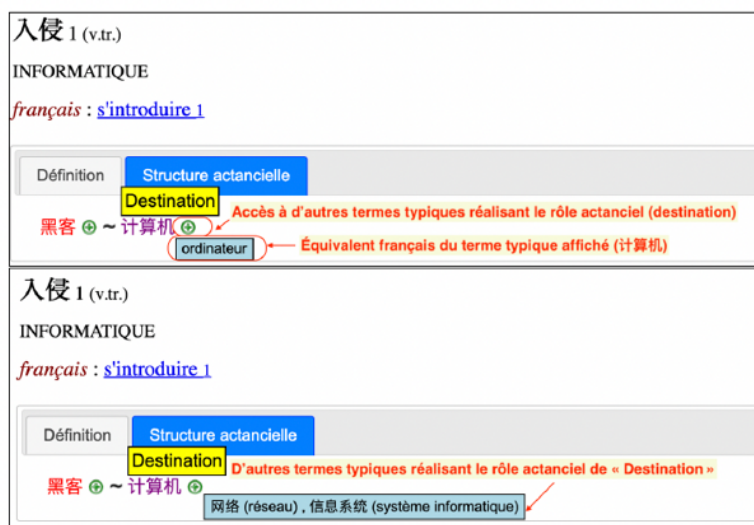


Figure 37. – Ruban affichant « Définition » (l'onglet inactif) et « Structure actancielle » (onglet actif) de 入侵₁ (« s'introduire »)

La description de la structure actancielle d'un terme sensibilise les apprenants aux actants sémantiques, à savoir les éléments impliqués dans le sens du terme ainsi que leur rôle par rapport au terme. L'énumération des termes typiques fait le lien entre le sens et le texte. Du point de vue linguistique, cette présentation a pour objectif d'aider les apprenants, d'une part, à comprendre davantage les propriétés sémantiques du terme et, d'autre part, à constater comment celles-ci déterminent le fonctionnement du terme dans les textes. Les apprenants pourraient être amenés à observer les liens sémantiques dans les CLS composées du terme en entrée et d'un terme typique réalisant un de ses actants sémantiques.

D'un point de vue didactique, la présentation de la structure actancielle vise à expliciter la règle qui régit l'utilisation d'un terme dans la production linguistique. Il s'agit d'une technique

d'enseignement censée favoriser l'acquisition des connaissances selon l'approche cognitive (voir la sous-section 3.3.1).

5.2.3 Contextes

La rubrique « Contextes » qui figure sur le deuxième ruban présente les contextes d'utilisation du terme dans les textes spécialisés. Ces contextes sont extraits de notre corpus lors de l'analyse des UT polysémiques (sous-section 4.1.2.4) et la collecte des contextes des CLS (sous-section 4.1.3.2), comme nous l'avons expliqué la sous-section 4.2.2. Chaque contexte est accompagné d'un code qui indique le numéro de texte original dans notre corpus⁸¹. Quant à la mise en page de cette rubrique, nous avons défini un jeu de couleur pour faire ressortir les informations importantes. La figure 38 illustre la rubrique « Contextes » de l'article 加密₂ (« cryptage »). Comme on peut le constater, le terme en entrée s'affiche en vert. D'autres termes répertoriés dans le CHINOINFO qui apparaissent dans les contextes sont affichés en bleu pâle. Si les contextes contiennent d'autres acceptions d'un même vocable que le terme en entrée, soit 加密₁ (« crypter ») et 加密₃ (« crypté »), un hyperlien dirige les utilisateurs vers les articles correspondants. La création des renvois vise à sensibiliser les apprenants à la distinction des acceptions d'un vocable.

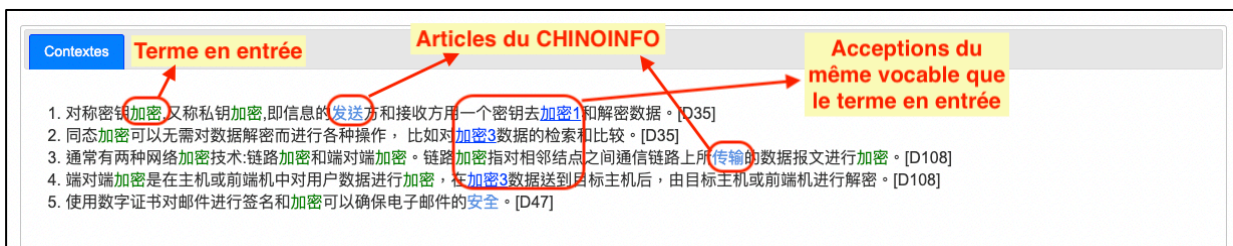


Figure 38. – Ruban contenant « Contextes » de 加密₂ (« cryptage »)

⁸¹ Tous les textes faisant partie de notre corpus sont numérotés (voir aussi la sous-section 4.1.1). Nous présentons la référence complète de ces textes dans l'annexe 2.

5.2.4 Termes reliés et combinaisons lexicales

Sur le dernier ruban s'affichent deux rubriques qui fournissent les liens lexicaux du terme vedette. La rubrique « Termes reliés » répertorie d'autres termes d'informatique qui sont reliés au terme vedette par un lien sémantique. Ces termes sont classés dans différents onglets selon la relation sémantique qu'ils entretiennent avec le terme décrit. Dans l'article 网络₁ (« réseau₁ »), comme l'illustre la figure 39, la rubrique « Termes reliés » comprend trois blocs qui affichent respectivement les synonymes du terme en entrée, les méronymes de celui-ci et le terme désignant l'utilisateur d'un 'réseau', soit 用户₁ (« utilisateur »). Les blocs masqués peuvent être activés en un clic de souris. Si un terme relié est répertorié dans le CHINOINFO, un hyperlien dirige les utilisateurs vers l'article correspondant. L'équivalent français des termes reliés est aussi fourni. Si l'équivalent d'un terme relié est décrit dans le DiCoInfo, un lien cliquable est créé pour donner accès à l'article du terme français.

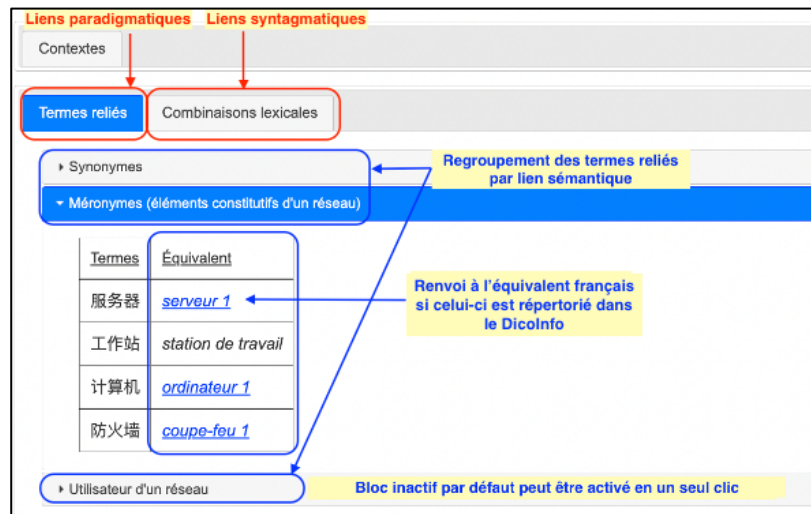


Figure 39. – Rubrique « Termes reliés » de l'article 网络₁ (« réseau₁ »)

Le modèle de présentation des termes reliés dans le CHINOINFO est légèrement différent de celui du DiCoInfo. Ce dernier répertorie les liens paradigmatiques d'un terme décrit dans deux rubans distincts : on distingue les synonymes, les contraires et les voisins du terme décrit des autres liens paradigmatiques reliant le terme vedette à d'autres termes du domaine (voir aussi la sous-section 3.2.2).

La rubrique « Combinaisons lexicales » présente les CLS contenant le terme en entrée. Inspirée du DiCoInfo (L’Homme, 2002, 2008; L’Homme et Alipour, 2009; L’Homme, 2009; Jousse et al., 2011; L’Homme et al., 2012; L’Homme et Jia, 2015) et des recherches menant à sa réalisation (Jousse et al., 2008; Jousse, 2010), nous avons adopté une présentation hiérarchique des CLS (sous-section 4.2.1.1). Dans l’article 网络₁ (^{wǎng luò} « réseau₁ »), la rubrique « Combinaisons lexicales » comprend quatre blocs (figure 40). Chaque bloc présente une classe générique qui regroupe les CLS régies par une même structure syntaxique. La fonction syntaxique du terme décrit est mise en évidence au moyen des crochets. Le troisième bloc, à titre d’exemple, correspond à la classe générique « Prédicat + [Complément d’objet] ». On y trouve toutes les CLS dans lesquelles le terme en entrée 网络₁ (^{wǎng luò} « réseau₁ ») est le COD d’un prédicat verbal. Une classe générique peut être subdivisée en différentes classes intermédiaires, dont chacune correspond à un tableau distinct. Le nom de la classe intermédiaire s’affiche en tête de chaque tableau. Par exemple, la classe intermédiaire « Préparer l’utilisation/le fonctionnement » répertorie deux CLS dans lesquelles 网络₁ (^{wǎng luò} « réseau₁ ») se combine avec un verbe dénotant l’action de ‘préparer l’utilisation d’un réseau ou le fonctionnement d’un réseau’, à savoir 设置₁ (^{shè zhì} « configurer ») et 管理¹¹ (^{guǎn} « gérer »). Si un cooccurrent fait l’objet d’un article dans le CHINOINFO, on voit un hyperlien qui renvoie à l’article en question. Nous proposons également la glose (ou la définition⁸²) et l’équivalent français pour chaque CLS décrite, afin de faciliter la compréhension des CLS par les utilisateurs non sinophones.

⁸² Rappelons que dans notre travail, les liens sémantiques exprimés par les FL non standardisées sont expliqués à l’aide de définitions terminologiques (voir aussi la sous-section 3.1.1).

D'autres termes de l'informatique reliés au terme décrit par un lien sémantique

Termes reliés **Combinaisons lexicales**

► Modificateur + [Nom modifié] **Regroupement des CLS en classes génériques en fonction de leur structure syntaxique**

► [Modificateur] + Nom modifié

▼ Prédicat + [Complément d'objet] **Fonction syntaxique du terme décrit dans les CLS**

| | | |
|---|--|--|
| Préparer l'utilisation/le fonctionnement | | |
| 设置1 网络 | L'administrateur prépare un réseau pour permettre au réseau de fonctionner | configurer un réseau |
| 管理网络1 | L'administrateur assure le bon fonctionnement d'un réseau | gérer un réseau |
| Utiliser/Ne pas utiliser | | |
| 接入网络1 | Un matériel est mis dans un réseau | être branché en réseau |
| 连接1 网络1 | Un matériel est mis dans un réseau | être branché en réseau |
| 连接1 (到) 网络1 | Un matériel est mis dans un réseau | être branché en réseau |
| Nuire au fonctionnement | | |
| 攻击1 网络1 | Endommager, forcer ou détourner un réseau pour commettre un acte préjudiciable | attaquer un réseau |

Regroupement des CLS en classes intermédiaires

Glose : expliquer le lien sémantique entre le terme décrit et ses cooccurrents

Équivalent français des CLS

Renvoi à d'autres articles du CHINOINFO

► [Sujet] + Prédicat

Figure 40. – Rubrique « Combinaisons lexicales » de l'article ^{wǎng luò}网络₁ (« réseau₁ »)

À la différence du DiCoInfo, dans lequel les liens paradigmatiques et syntagmatiques sont présentés dans deux rubans superposés (voir aussi annexe 21), nous présentons les rubriques « Termes reliés » et « Combinaisons lexicales » côte à côte pour raccourcir le temps de navigation entre elles. Comme ces rubriques contiennent souvent de longues listes, la consultation de ces contenus risque d'être chronophage et exige une attention soutenue. Ce mode d'affichage permet aux utilisateurs de repérer rapidement l'information dont ils ont besoin dans un article, sans avoir à parcourir le contenu de chaque rubrique.

5.3 Fonctionnalités du CHINOINFO

Nous avons eu recours à différentes techniques pour assurer la convivialité de l'interface du CHINOINFO, afin de permettre aux utilisateurs de tirer pleinement parti de la ressource élaborée. Les aspects suivants ont été pris en compte lors de la mise en ligne du dictionnaire : la facilité d'accès aux données terminologiques (5.3.1), la simplicité d'utilisation de l'interface (5.3.2), la facilité de navigation dans la page Web (5.3.3) ainsi que le bilan des liens lexicaux encodés (5.3.4).

5.3.1 Recherche de termes

Nous proposons deux modes de recherches dans le CHINOINFO : recherche par index et recherche par caractère, afin d'offrir un accès rapide et facile aux termes recherchés. En premier lieu, on peut rechercher des termes dans la nomenclature qui s'affiche sur l'interface de recherche (figure 41). La nomenclature est assortie de trois onglets, dans lesquels les termes sont triés de façon distincte. Dans l'onglet intitulé « trié par lexie », les termes sont triés par ordre alphabétique selon la transcription phonétique de leur premier caractère. La liste des termes commence par ^{ān quǎn}安全₁ (« sécurité informatique ») puisque sa transcription phonétique en pinyin ([ān quǎn]) se classe en premier. Cet onglet est particulièrement utile pour localiser une forme lexicale dans le cas où les utilisateurs connaissent la prononciation du premier caractère de cette forme lexicale mais ignorent le(s) sens associé(s) à celle-ci. Il s'agit de l'ordre d'affichage des termes par défaut.

Dans l'onglet « trié par catégorie », les termes sont classés selon la partie du discours à laquelle ils appartiennent, dont l'adjectif, l'adverbe, le nom et le verbe. Pour le verbe, on précise s'il s'agit d'un verbe intransitif ou transitif. Cette manière de trier les termes permet aux utilisateurs d'accéder rapidement à une acception d'un vocable qui relève d'une partie du discours spécifique.

Dans l'onglet « trié par glose », les termes sont classés par ordre alphabétique selon leur équivalent français. C'est donc le terme ^{yuǎnchéng}远₂程 (« à distance ») qui s'affiche en haut de la liste des termes. Cet onglet offre un accès aux termes chinois via leur sens, ce qui est utile pour les apprenants francophones qui cherchent un terme pour exprimer un sens spécifique. Le passage d'un onglet à l'autre est facile et simple : les utilisateurs n'ont qu'à cliquer sur l'onglet qui les intéresse pour l'activer.

En deuxième lieu, il est possible de rechercher les termes par caractère. Dans le module de recherche qui se trouve en haut à gauche de l'interface de recherche, nous élaborons une liste de *hanzi* (pour ^{hàn zì}汉字 « caractère chinois ») qui regroupe tous les caractères composant les termes répertoriés. Si un utilisateur pointe le curseur sur un des caractères dans le module, une liste des termes qui contiennent ce caractère s'affiche. L'utilisateur peut sélectionner le terme voulu parmi les termes proposés et cliquer sur le terme choisi pour accéder à l'article. Ce mode de recherche

est notamment utile pour les apprenants qui souhaitent localiser une forme lexicale dont ils ignorent la prononciation et le(s) sens.



ChinoIn Rechercher de termes par caractère :
liste des caractères

Sélectionnez une entrée en fonction du hanzi pointé :

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 安 | 除 | 发 | 共 | 连 | 入 | 试 | 问 | 移 | 在 |
| 保 | 储 | 访 | 击 | 络 | 软 | 手 | 下 | 意 | 载 |
| 备 | 份 | 加 | 密 | 删 | 输 | 送 | 享 | 用 | 置 |
| 步 | 存 | 复 | 迁 | 设 | 同 | 新 | 远 | 装 | |
| 测 | 动 | 更 | 件 | 侵 | 时 | 网 | 行 | 运 | |
| 程 | 恶 | 攻 | 接 | 全 | 实 | | | 自 | |

Le CHINOINFO est un dictionnaire spécialisé bilingue destiné aux apprenants francophones du chinois. Ce dictionnaire répertorie 91 termes fondamentaux du domaine de l'informatique. Tous les renseignements terminologiques sont extraits à partir d'un corpus chinois comptant 341 676 mots.

L'élaboration du CHINOINFO s'inscrit dans le cadre d'un projet de thèse visant à concevoir une méthode pour décrire le lexique spécialisé chinois, plus précisément, les combinaisons lexicales spécialisées (CLS). Nous nous attachons à explorer de quelle manière la représentation sémantique des CLS favorise leur acquisition par les locuteurs non sinophones. L'efficacité de la méthode mise au point sera évaluée au moyen d'une expérimentation.

Ce projet est réalisé par [Zhiwei Han](#), étudiante au doctorat à l'Université de Montréal, sous la supervision de [Marie-Claude L'Homme](#), professeure au [Département de linguistique et de traduction](#) de l'Université de Montréal. Nous remercions chaleureusement [Benoit Robichaud](#), membre de l'[Observatoire de linguistique Sens-Texte \(OLST\)](#), pour la généreuse contribution qu'il a apportée à la mise en ligne du CHINOINFO.

[Lien vers le tutoriel vidéo](#)

Nomenclature (91 entrées) :

[Visionner le tutoriel ici](#)

Rechercher des termes par index:
nomenclature triée de trois façons différentes

triée par lexie
 triée par catégorie
 triée par glose

| | | | |
|---------------------------------------|-----------------------------------|--------------------------------|--------------------------------|
| 安全 1 (nom) [sécurité informatique] | 发送 4 (nom) [envoi 2] | 检测 4 (nom) [vérification] | 同步 4 (adv.) [synchroniquement] |
| 安全 2 (adj.) [sécurisé] | 访问 1 (v.tr.) [accéder 1] | 连接 1 (v.tr.) [se connecter 1] | 网络 1 (nom) [réseau 1] |
| 安全 3 (adv.) [sécuritairement] | 访问 2 (v.tr.) [accéder 2] | 连接 2 (v.tr.) [connecter] | 网络 2 (nom) [réseau 2] |
| 安装 1 (v.tr.) [installer 2] | 访问 3 (nom) [accès 1] | 连接 3 (nom) [connexion 1.1] | 网络 3 (adj.) [en ligne] |
| 安装 2 (nom) [installation 2] | 访问 4 (nom) [accès 2] | 连接 4 (nom) [connexion 2.1] | 下载 1 (v.tr.) [télécharger] |
| 保存 1 (v.tr.) [enregistrer 1] | 复制 1 (v.tr.) [copier 1] | 迁移 1 (v.tr.) [faire migrer] | 下载 2 (nom) [téléchargement] |
| 保存 2 (nom) [enregistrement 1.1] | 复制 2 (v.tr.) [copier 2] | 迁移 2 (nom) [migration] | 移动 1 (adj.) [mobile 1] |
| 备份 1 (v.tr.) [sauvegarder] | 复制 3 (nom) [copie 1.1] | 入侵 1 (v.tr.) [s'introduire] | 移动 2 (adj.) [amovible] |
| 备份 2 (nom) [sauvegarde 1] | 复制 4 (nom) [copie 2] | 入侵 2 (nom) [intrusion] | 应用 1 (nom) [application 1] |
| 备份 3 (nom) [sauvegarde 2] | 更新 1 (v.tr.) [mettre à jour] | 软件 1 (nom) [logiciel 1] | 应用 2 (nom) [application 2] |
| 备份 4 (adj.) [sauvegardé] | 更新 2 (nom) [mise à jour 1.1] | 软件 2 (nom) [logiciel 2] | 应用 3 (v.intr.) [s'appliquer] |
| 测试 1 (v.tr.) [tester] | 更新 3 (nom) [mise à jour 1.2] | 删除 1 (v.tr.) [supprimer 1] | 应用 4 (v.tr.) [appliquer 2] |
| 测试 2 (nom) [test] | 攻击 1 (v.tr.) [attaquer 1] | 删除 2 (nom) [suppression 1] | 远程 1 (adj.) [distant] |
| 传输 1 (v.tr.) [transférer 1] | 攻击 2 (nom) [attaque informatique] | 设置 1 (v.tr.) [configurer 1] | 远程 2 (adv.) [à distance] |
| 传输 2 (nom) [transfert 1] | 共享 1 (v.tr.) [partager 1] | 设置 2 (nom) [configuration 1.1] | 运行 1 (v.intr.) [s'exécuter] |
| 存储 1 (v.tr.) [stocker] | 共享 2 (nom) [partage 1] | 设置 3 (nom) [configuration 1.2] | 运行 2 (v.tr.) [faire exécuter] |
| 存储 2 (nom) [stockage 1a] | 共享 3 (adj.) [partagé] | 实时 1 (adj.) [en temps réel 1] | 运行 3 (nom) [exécution 1a] |
| 存储 3 (nom) [stockage 1a] | 加密 1 (v.tr.) [crypter] | 实时 2 (adv.) [en temps réel 2] | 运行 4 (nom) [exécution 1b] |
| 恶意 1 (adj.) [malveillant 1] | 加密 2 (nom) [cryptage] | 手动 1 (adj.) [manuel] | 在线 1 (adj.) [en ligne 1] |
| 恶意 2 (adv.) [de manière malveillante] | 加密 3 (adj.) [crypté] | 手动 2 (adv.) [manuellement] | 在线 2 (adv.) [en ligne 2] |
| 发送 1 (v.tr.) [envoyer 1] | 检测 1 (v.tr.) [détecter 1] | 同步 1 (v.tr.) [synchroniser] | 自动 1 (adj.) [automatique 1] |
| 发送 2 (v.tr.) [envoyer 2] | 检测 2 (v.tr.) [vérifier] | 同步 2 (nom) [synchronisation] | 自动 2 (adv.) [automatiquement] |
| 发送 3 (nom) [envoi 1] | 检测 3 (nom) [détection 1] | 同步 3 (adj.) [synchrone 1] | |

Figure 41. – Interface de recherche du CHINOINFO

Le fait d'offrir de multiples points d'accès aux articles facilite la consultation du dictionnaire. Cette démarche vise notamment à adapter la ressource aux besoins des apprenants francophones ayant des niveaux de compétence lexicale hétérogènes. Notons qu'un moteur de recherche pourra être ajouté à l'interface du CHINOINFO lorsque la nomenclature du dictionnaire s'enrichira.

5.3.2 Guide d'utilisation du CHINOINFO

Nous avons préparé un tutoriel vidéo afin d'aider les utilisateurs à se familiariser avec l'utilisation du CHINOINFO. Il s'agit d'un enregistrement d'une présentation PowerPoint accompagnée d'explications narrées au sein du même fichier vidéo. Cette capsule vidéo a d'abord été conçue pour les participants à l'expérimentation. Elle peut s'adresser à tout utilisateur de la ressource. Dans ce tutoriel, on fournit des renseignements sur le projet de recherche dans lequel s'inscrit l'élaboration du CHINOINFO, le nombre de termes répertoriés dans le dictionnaire, les modes de recherche du dictionnaire, la structuration des connaissances au sein d'un article, ainsi que les fonctionnalités de navigation dans le CHINOINFO. Un lien vers le tutoriel est affiché sur l'interface de recherche du CHINOINFO⁸³, comme l'illustre la figure 41, pour permettre les utilisateurs de visionner la vidéo à tout moment.

5.3.3 Renvois

La facilité de navigation est un autre facteur important à prendre en compte pour renforcer la convivialité d'une ressource lexicale en ligne.

Dans le CHINOINFO, nous avons créé différents renvois entre les articles, comme cela a été évoqué dans la section 5.2. Premièrement, nous avons établi, dans la mesure du possible, le lien d'équivalence entre les entrées du CHINOINFO et celles du DiCoInfo. Si l'équivalent d'un terme en entrée fait l'objet d'un article du DiCoInfo, un hyperlien est créé pour permettre aux utilisateurs d'y accéder. Dans la rubrique « Termes reliés », l'équivalent français d'un terme relié est signalé sous forme de renvoi si ce terme français est répertorié dans le DiCoInfo⁸⁴. La figure 42 illustre les renvois à deux articles du DiCoInfo dans l'article 入侵^{rù qīn} (« intrusion »), à savoir l'équivalent français du terme en entrée (*intrusion*) et celui du terme dénotant la personne qui réalise une intrusion (*pirate*).

⁸³ Le tutoriel est repérable à l'adresse suivante : <http://olst.ling.umontreal.ca/chinoinfo/>. Nous remercions Benoit Robichaud pour la mise en ligne du tutoriel.

⁸⁴ À terme, le renvoi inverse pourra être ajouté à partir d'un article du DiCoInfo

入侵² (nom)
INFORMATIQUE
français : [intrusion 1](#) ← Lien vers l'article « intrusion 1 » du DiCoInfo

Définition Structure actancielle
Accès non autorisé à un système informatique ou à un réseau, obtenu en contournant ou en désamorçant les dispositifs de sécurité en place. [GDTINTR]

Contextes

Termes reliés Combinaisons lexicales

› Synonymes

▼ Personne qui réalise une intrusion

| Termes | Équivalent |
|--------|---|
| 黑客 | pirate ← Lien vers l'article « pirate » du DiCoInfo |

› Famille de mots-verbe

Figure 42. – Renvois à deux articles du DiCoInfo dans l'article 入侵² (« intrusion »)

Deuxièmement, nous avons créé des renvois aux termes faisant partie de la nomenclature du CHINOINFO si ceux-ci apparaissent dans un autre article du dictionnaire. Dans les rubriques « Termes reliés » et « Combinaisons lexicales », si un terme relié ou un collocatif du terme en entrée est décrit dans le CHINOINFO, un hyperlien est créé pour donner accès à l'article correspondant. Dans la rubrique « Contextes », seules les acceptions d'un même vocable que le terme en entrée sont assorties d'hyperlien. Compte tenu de la densité du contenu dans cette rubrique, cette solution permet d'éviter l'encombrement visuel et rend la rubrique aérée et facile à lire. Par exemple, on trouve dans l'article 迁移² (« migration ») les liens d'accès à deux autres articles du CHINOINFO. Le premier lien qui apparaît dans la section « Combinaisons lexicales » dirige vers le terme 实时¹ (« en temps réel »). Celui-ci est le modificateur adjectival du terme vedette. Le deuxième lien qui figure dans la section « Contextes » renvoie au terme 迁移¹ « faire migrer ». Celui-ci appartient au même vocable que le terme vedette (figure 43).

迁移² (nom) [Tr]
 INFORMATIQUE
français : migration

Définition Structure actancielle

Contextes

1. 数据复制服务的迁移方式分为全量+增量两种。全量模式为数据库一次性迁移，适用于可中断业务的数据库迁移场景。[D44]
2. 数据复制服务 (Data Replication Service, 简称DRS) 是一种易用、稳定、高效、用于数据库实时迁移和数据库实时同步的云服务。[D44]
3. 实时迁移是指在数据复制服务器能够同时连通源数据库和目标数据库的情况下，只需要配置迁移的源、目标数据库实例及迁移对象即可完成整个数据迁移过程，再通过多项指标和数据的对比分析，帮助确定合适的业务割接时机，实现最小化业务中断的数据库迁移。[D44]
4. 实时迁移支持多种网络迁移方式，如：公网网络、VPC网络、VPN网络和专线网络。通过多种网络链路，可快速实现跨云平台数据库迁移、云下数据库迁移上云或云上跨区域的数据迁移等多种业务场景迁移。[D44]
5. 关系型数据库依然是各行业的的主流应用之一，怎样更快的将传统关系型数据中的数据迁移¹上云也是很多行业用户关心的。为此京东云特意打造了一款针对关系型数据库的迁移工具——数据传输DTS。[D28]
6. 通过增量迁移技术，能够最大限度允许迁移过程中业务继续对外提供使用，有效地将业务系统中断时间和业务影响最小化。[D44]
7. 在迁移过程中京东云不仅提供了常规的停机迁移方案，还提供了特殊的不停机迁移方案，保证了物流业务不停服。[D28]

Lien vers l'article 迁移¹ « faire migrer » du CHINOINFO
 (!'acception relevant du même vocable que le terme en entrée)

Termes reliés Combinaisons lexicales

▼ Modificateur + [Nom modifié]

Lien vers l'article 实时¹ « en temps réel (adj) » du CHINOINFO

| | Sortes de | |
|---------------------------------|---|--|
| 数据迁移 ² | Qui porte sur un ensemble de données | <i>migration des données</i> |
| 数据库迁移 ² | Qui porte sur une base de données | <i>migration d'une base de données</i> |
| 实时 ¹ 迁移 ² | Qui se réalise en réduisant le temps d'arrêt opérationnel au minimum | <i>migration en temps réel</i> |
| 全量迁移 ² | Qui se fait sur la totalité des données | <i>migration complète</i> |
| 增量迁移 ² | Qui se fait sur les éléments des données non présents lors de la dernière migration | <i>migration incrémentale</i> |

► [Modificateur] + Nom modifié

Figure 43. – Renvois à deux articles du CHINOINFO dans l'article 迁移² (« migration »)

5.3.4 Bilan des liens lexicaux encodés

Outre l'interface de recherche, nous avons intégré une page de synthèse à la version en ligne du CHINOINFO⁸⁵. Cette page contient deux tableaux qui proposent un bilan statistique des liens lexicaux présentés dans le CHINOINFO⁸⁶. La page de synthèse est conçue pour les personnes qui enrichissent la ressource lexicale.

⁸⁵ La page de synthèse a été créée à l'aide de Benoit Robichaud, membre de l'Observatoire de linguistique Sens-Texte (OLST). Nous le remercions chaleureusement pour la contribution généreuse qu'il a apportée tout au long de l'élaboration de la version en ligne du CHINOINFO.

⁸⁶ La page de synthèse du CHINOINFO est repérable à l'adresse suivante : <http://olst.ling.umontreal.ca/chinoinfo/syntheses.php>

L'annexe 22 illustre le tableau de synthèse des données encodées dans la rubrique « Termes reliés » de tous les articles. Ce tableau fait état du nombre total des termes présentés dans cette rubrique. On y trouve également la liste des liens sémantiques identifiés et le nombre des termes associés à un terme vedette par ces liens sémantiques. Comme on peut le constater, les liens paradigmatiques les plus fréquemment identifiés dans le CHINOINFO comprennent les sens proches, les contraires, les synonymes, ainsi que les liens de dérivation sémantique entre un verbe et un nom relevant d'un même mot multicatégoriel.

Le tableau de synthèse illustré dans l'annexe 23 dénombre les CLS encodées dans le CHINOINFO. On fait état des CLS réparties dans chaque classe intermédiaire. Rappelons que les classes intermédiaires sont élaborées selon les liens syntagmatiques identifiés au sein des collocations restreintes et les liens observés dans les structures actanciennes de termes (voir la sous-section 4.2.1.1).

Le bilan des liens lexicaux est dynamique en ce sens qu'il tient compte de la dernière mise à jour du CHINOINFO. Nous effectuons la mise à jour des données terminologiques dans les fichiers XML. Comme le CHINOINFO est hébergé sur le serveur de l'OLST, les modifications apportées aux articles sous format XML se reflètent immédiatement sur l'interface du CHINOINFO en ligne. L'annexe 10 illustre l'article ^{ruǎnjiàn}软件₁ (« logiciel₁ ») mis en ligne. L'encodage des données en XML du même article est illustré dans l'annexe 9.

En résumé, nous nous attachons à fournir une présentation claire et organisée des renseignements descriptifs terminologiques, dans l'objectif de rendre la consultation du CHINOINFO aisée et efficace.

5.4 Synthèse

Dans cette section, nous avons décrit le contenu et les fonctionnalités du CHINOINFO, dictionnaire spécialisé chinois-français traitant des termes fondamentaux en informatique et technologies de l'information. Cette ressource lexicale est destinée aux apprenants débutant du chinois sur objectifs spécifiques qui ont peu de connaissances du chinois et du domaine spécialisé.

Le CHINOINFO fournit une description des propriétés syntaxico-sémantiques des CLS, afin d'aider les apprenants à surmonter les difficultés faisant obstacle à l'acquisition de la

combinatoire spécialisée. L'élaboration de cette ressource lexicale résulte de l'application de la méthode de description que nous avons proposée dans la section 4.2. Dans la version en ligne du CHINOINFO, la structuration des données terminologiques au sein des articles s'est également inspirée de l'adaptation du DiCoInfo à un dictionnaire d'apprentissage (Alipour, 2014). Quant aux fonctionnalités du CHINOINFO, nous avons fait appel à différents moyens techniques pour assurer la convivialité de la ressource.

L'apport du CHINOINFO à la résolution des problèmes lexicaux soulevés par les CLS a été évalué au moyen d'une expérimentation comparative auprès deux groupes d'apprenants francophones du chinois, comme nous l'avons expliqué dans la section 4.3. Les résultats de cette expérimentation font l'objet du prochain chapitre.

Chapitre 6 Traitement des résultats expérimentaux

Dans ce chapitre, nous présentons en détail notre analyse des données expérimentales et les résultats que nous avons obtenus. Nous analysons en premier lieu le profil des participants (6.1). Nous abordons ensuite le traitement des réponses au test lexical (6.2). La section 6.3 décrit notre démarche pour analyser les données collectées au moyen des questionnaires 2a et 2b visant à recueillir les commentaires du GC et du GE sur le test lexical et leur recours à des ressources lexicales au cours du test (voir la sous-section 4.3.1.3). À la fin du chapitre, nous discutons les résultats de notre analyse pour tracer un portrait de l'intérêt que présente le CHINOINFO dans le contexte pédagogique que nous avons ciblé (6.4).

6.1 Analyse des réponses au questionnaire de profil

À l'issue du processus de recrutement, nous avons recruté huit participants. Ils ont été divisés de façon aléatoire en deux groupes à nombre égal : le groupe contrôle (GC) et le groupe expérimental (GE). Les données recueillies auprès du GC et du GE ont été codées avant d'être analysées, dans le but d'assurer la confidentialité des renseignements personnels des participants. Un code a été attribué à chaque participant pour identifier les documents qu'ils nous ont transmis à la fin de la séance Zoom (voir la sous-section 4.3.2). Les quatre participants du GC ont respectivement été identifiés comme C1, C2, C3 et C4. Les participants du GE ont été identifiés comme E1, E2, E3 et E4.

Nous avons collecté les renseignements sur le profil des participants à l'aide du questionnaire 1. Ce questionnaire comprend 12 questions à choix de réponses (voir aussi la sous-section 4.3.1.3 et l'annexe 12). Il s'agit de questions ouvertes spécifiques (Dörnyei et Taguchi, 2009, p. 37, 98-99) portant sur des informations factuelles, telles que l'expérience d'apprentissage du chinois des participants.

6.1.1 Méthode d'analyse des données collectées

La première étape du traitement des données consiste à coder les réponses reçues. Pour chaque question, nous avons attribué un numéro sans valeur numérique à chaque option de réponse. Par exemple, dans la question 1 portant sur l'affiliation universitaire des participants (annexe 12),

les trois options de réponses sont numérotées ainsi : « Université de Montréal » = 1, « Université Laval » = 2, « Université du Québec à Montréal » = 3. Outre ces options proposées, chaque réponse reçue se voit attribuer un code (par exemple, « l'Université des langues et cultures de Beijing » = 4). Autrement dit, le codage des réponses à une question ouverte comporte autant de catégories que le nombre de réponses différentes. Cette démarche vise à transformer les réponses reçues en catégories quantifiables pour les traiter ensuite comme des données quantitatives (Dörnyei et Taguchi, 2009, p. 98-99).

Nous avons ensuite synthétisé les données selon les types de renseignements recueillis. Dans le questionnaire 1, les questions 1 à 12 (Q1-Q12) peuvent être divisées en six catégories (sous-section 4.3.1.3), comme l'illustre le tableau 34. Les résultats d'analyse nous permettent de broser un portrait descriptif du profil des participants.

| | Type de renseignements recueillis | Questions |
|---|--|-----------|
| 1 | Le milieu universitaire des participants | Q1 – Q3 |
| 2 | Leur expérience d'apprentissage du chinois | Q4 – Q5 |
| 3 | Leur expérience d'apprentissage du chinois sur objectifs spécifiques (COS) | Q6 |
| 4 | Leur expérience de formation ou de travail dans le domaine de l'informatique | Q7 |
| 5 | Leur habitude d'apprentissage du chinois | Q8 – Q9 |
| 6 | Leur séjour en Chine | Q10 – Q12 |

Tableau 34. – Catégorisation des renseignements recueillis au moyen du questionnaire 1

6.1.2 Résultats d'analyse

Selon l'analyse des réponses à la Q1 du questionnaire 1, la majorité des participants recrutés (sept sur huit) provient des universités québécoises, dont quatre étudiants de l'Université Laval (ULaval) et trois étudiants de l'Université de Montréal (UdeM). Nous avons aussi réussi à recruter un étudiant francophone de l'Université des langues et cultures de Beijing (BLCU) (figure 44).

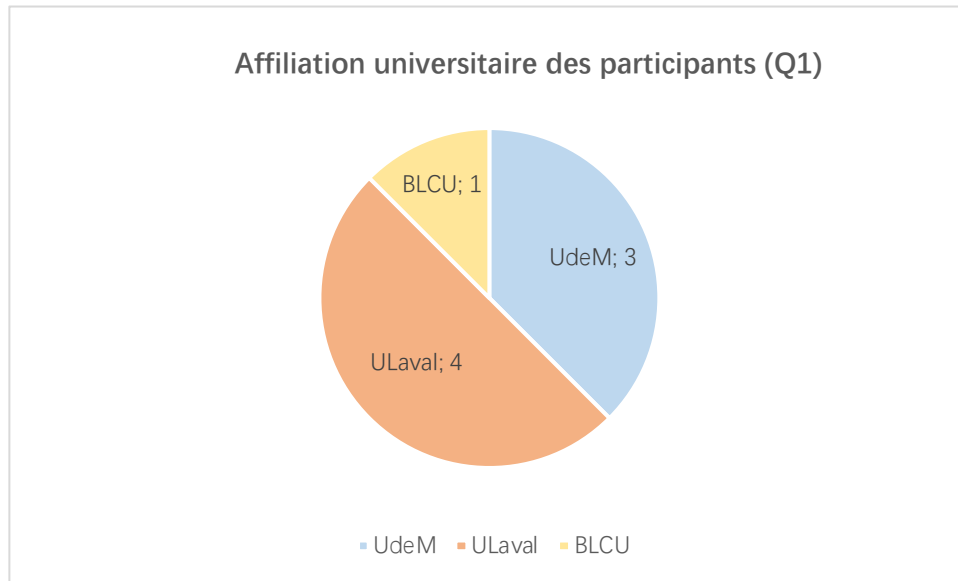


Figure 44. – Répartition des participants par affiliation universitaire

La figure 45 présente le niveau d'études universitaires des participants ainsi que leur programme d'études (Q2-Q3). Parmi les sept participants québécois, cinq étudiants sont inscrits au premier cycle, dont deux étudiants au baccalauréat en études internationales et langues modernes (ULaval), deux étudiants au baccalauréat en études asiatiques (UdeM) et un étudiant inscrit au programme de mineure en études est-asiatiques (UdeM). Les deux autres participants québécois sont titulaires de diplômes d'études supérieures : l'un détient une maîtrise de traduction et d'interprétation (ULaval); l'autre a obtenu son doctorat en philosophie à ULaval. Il convient de mentionner que ce dernier participant est titulaire d'un diplôme d'études collégiales en médecine traditionnelle chinoise. Le seul étudiant francophone recruté en Chine vient de terminer son année préparatoire avec succès et est entré dans le programme de baccalauréat en économie et commerce international (BLCU).

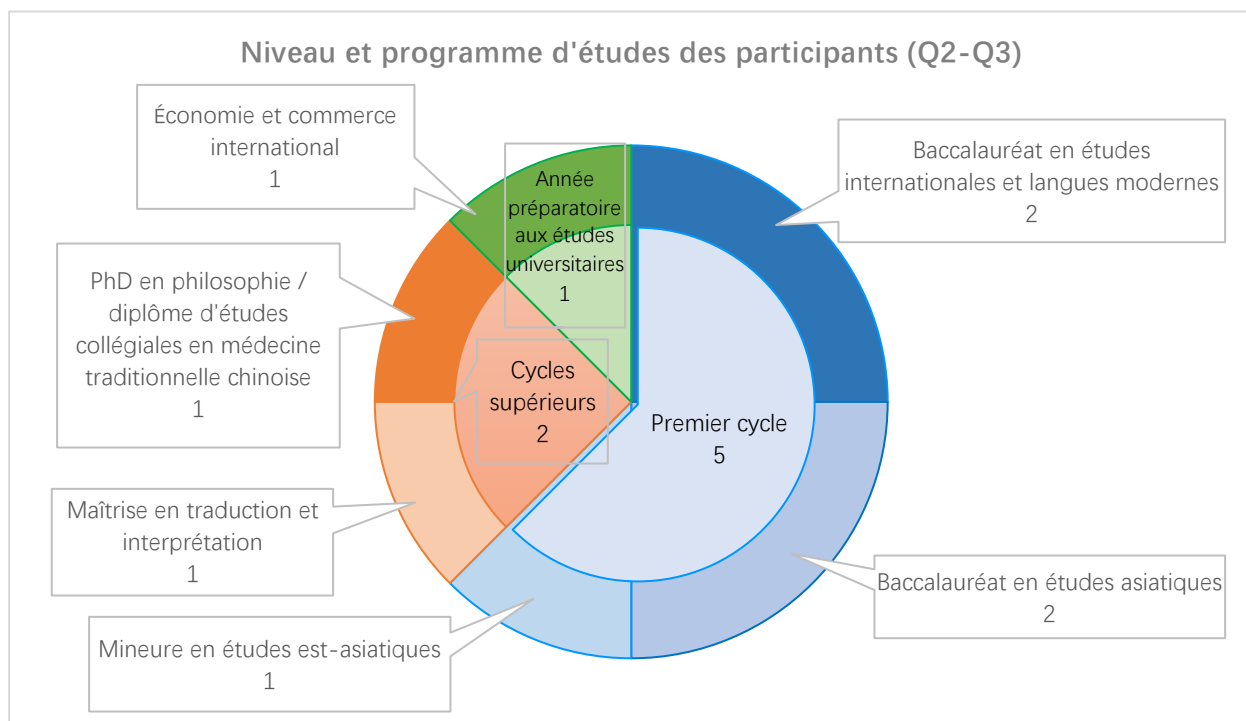


Figure 45. – Répartition des participants par niveau et programme d'études

Quant à l'expérience d'apprentissage du chinois (Q5), quatre des huit participants ont accumulé plus de 500 heures d'apprentissage, ce qui correspond à plus de quatre trimestres de formation à raison de trois heures de cours par semaine. Ces apprenants ont été affectés aléatoirement à nombre égal au GC et au GE. Deux des huit participants ont accumulé entre 400 et 540 heures d'apprentissage, soit trois trimestres de formation. Ils se sont retrouvés dans le GC. Les deux autres participants ont suivi deux cours de chinois et accumulé ainsi entre 270 et 400 heures d'apprentissage. Ces apprenants les moins expérimentés se sont retrouvés dans le GE (figure 46).

Tous les participants ont reçu leur formation en chinois en milieu universitaire (Q4). Il convient de mentionner qu'un des participants qui ont accumulé plus de 500 heures d'apprentissage a suivi son premier cours de chinois au cégep.

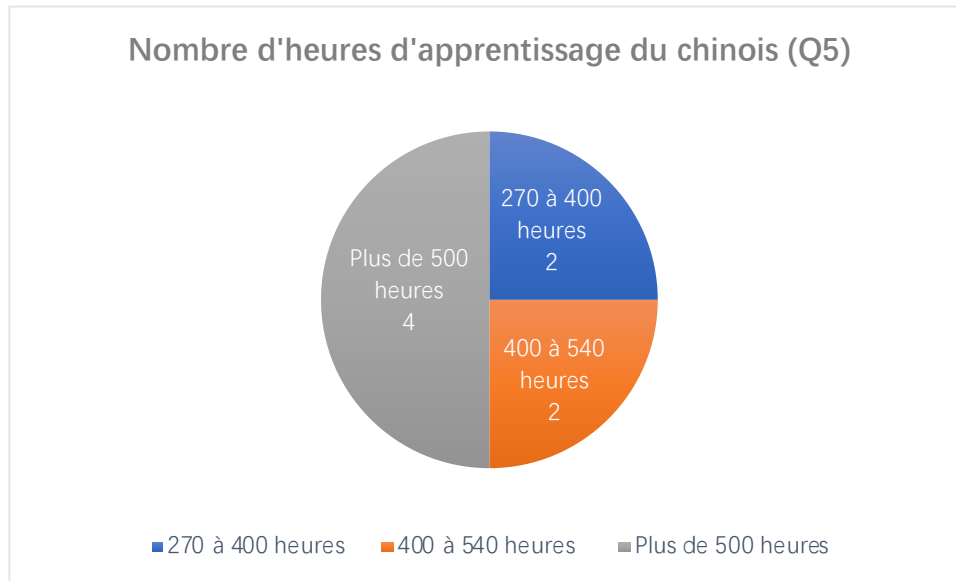


Figure 46. – Répartition des participants par nombre d'heures d'apprentissage du chinois

En ce qui concerne l'expérience de formation en COS (Q6), la majorité des participants (sept sur huit) n'ont jamais suivi de cours du COS, sauf l'étudiant provenant de BLCU qui a suivi un cours de chinois du commerce. Par ailleurs, un participant québécois a affirmé avoir accumulé des connaissances sur la terminologie du chinois médical dans le cadre de sa formation en médecine traditionnelle chinoise et en acupuncture.

Selon l'analyse des réponses à la Q7, six des huit participants n'ont pas d'expérience de formation ou de travail dans le domaine de l'informatique. Seuls deux participants ont affirmé avoir étudié ou travaillé dans le domaine du système informatique.

Nous avons collecté des renseignements sur l'habitude d'apprentissage des participants, plus précisément, les ressources auxquelles ils ont souvent recours pour résoudre les difficultés lexicales (Q8). Tous les huit participants ont l'habitude de consulter les dictionnaires bilingues (chinois-français ou français-chinois), comme notre expérience d'enseignement nous permet de le prévoir. Trois participants ont mentionné leur recours au matériel didactique du chinois (manuel, lexiques et vocabulaires thématiques, etc.). Parmi les ressources les moins couramment utilisées figurent les dictionnaires monolingues chinois et l'espace de discussion en ligne (forums, blogs et site Web d'apprentissage du chinois), comme l'illustre la figure 47.

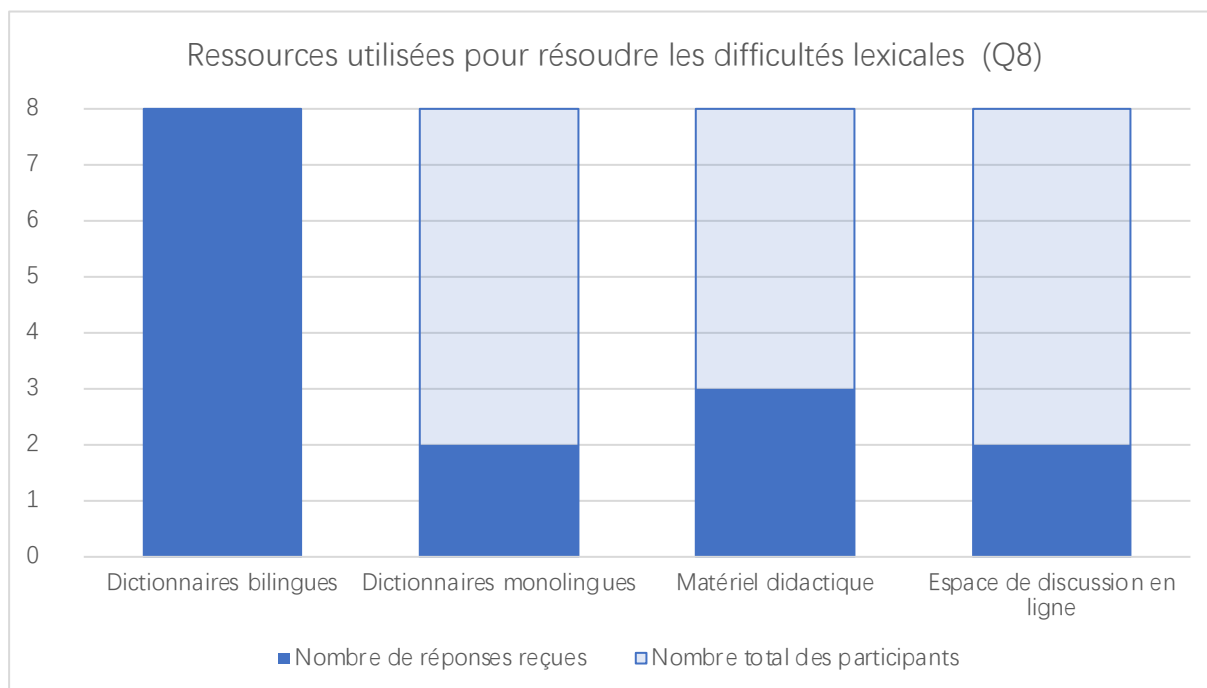


Figure 47. – Résultats du sondage sur les ressources utilisées par les participants pour résoudre les difficultés lexicales (Q8)

Pour faire suite à la Q8, nous avons étudié les besoins d'apprentissage que les ressources lexicales ne permettent pas de combler. Dans la Q9, les participants ont été invités à indiquer la (les) difficulté(s) qu'ils éprouvent lorsqu'ils cherchent à résoudre une question à l'aide de ressources lexicales. Parmi les lacunes les plus signalées par les répondants, on trouve le manque d'exemples d'usage d'un mot cherché (62,5 %) et le manque d'explication sur les propriétés combinatoires d'un mot cherché (62,5 %), comme l'illustre la figure 48. En outre, quatre répondants ont affirmé qu'ils ne savent pas quoi chercher pour comprendre une séquence de mots inconnus, ce qui porte à croire que les ressources lexicales ne fournissent pas suffisamment de renseignements descriptifs sur les combinaisons lexicales.

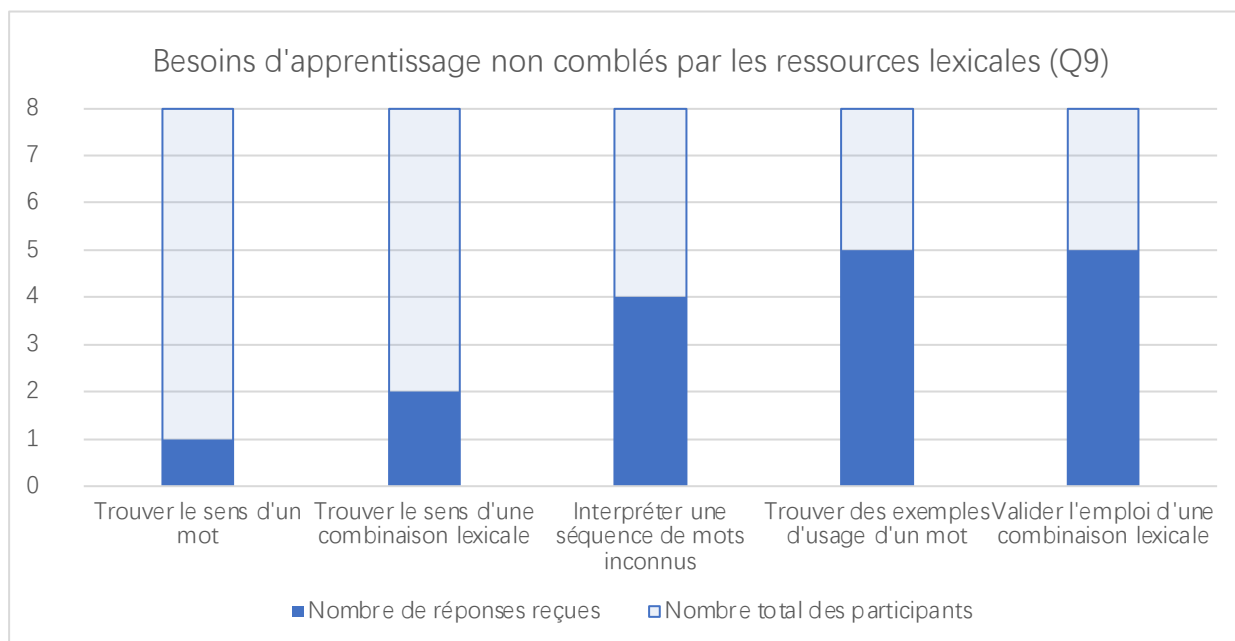


Figure 48. – Résultats du sondage sur les besoins d'apprentissage non comblés par les ressources lexicales (Q9)

Quant à l'expérience de vie des participants dans un milieu sinophone (Q10-Q12), deux étudiants provenant du Québec ont vécu un séjour touristique en Chine, pour une durée de moins d'un mois. Seul un étudiant vit en Chine depuis environ un an pour effectuer une année préparatoire aux études universitaires.

Pour conclure, l'analyse des réponses au questionnaire 1 nous a permis de brosser un portrait des participants. En résumé, tous les apprenants ont reçu leur formation en chinois en milieu universitaire, dont la plupart proviennent de deux universités québécoises. Du point de vue du nombre d'heures consacrées à l'apprentissage du chinois, le GC présente un niveau d'apprentissage légèrement plus élevé que le GE. Tous les participants ont accumulé au moins 270 heures d'apprentissage. La plupart d'entre eux n'ont jamais étudié ou travaillé dans le domaine de l'informatique. Seul un participant parmi les huit fait un séjour régulier en Chine pour des raisons d'études. Quant aux ressources utilisées dans leur apprentissage, les participants utilisent les dictionnaires bilingues comme leurs outils de référence principaux. Ils ont toutefois signalé l'insuffisance de ces ressources pour résoudre les problèmes lexicaux, notamment en raison du manque d'exemples d'usage et du manque de description des combinaisons lexicales.

6.2 Analyse des réponses au test lexical

Notre démarche pour traiter les données collectées à l'aide du test lexical a été réalisée en quatre étapes : 1) corriger les réponses aux questions lexicales du test (6.2.1); 2) comparer les résultats du GC et du GE (6.2.2); 3) catégoriser les erreurs repérées (6.2.3) et 4) synthétiser les réponses à la question libre du test (6.2.4). Nous faisons enfin une synthèse des résultats de notre analyse (6.2.5).

6.2.1 Correction des réponses aux questions lexicales

En premier lieu, nous avons élaboré une grille de correction pour corriger les réponses aux questions 1-9 du test lexical (annexe 24). Comme nous l'avons évoqué dans la sous-section 4.3.1.3, ces questions sont divisées en trois sections. Chaque section vise à mesurer l'habileté des participants à résoudre un type de problèmes lexicaux soulevés par les CLS. Les sections I et II sont composées de six questions à réponse courte. Les participants devaient indiquer si une forme lexicale (section I) ou une combinaison lexicale (section II) véhicule des sens différents dans deux phrases et justifier leur réponse. Chaque question vaut 2 points, dont 1 point est attribué à la correction de la réponse donnée et 1 point à la pertinence de la justification, comme l'illustre le tableau 35.

| | Section 1 (Question 1 à 3) | Section 2 (Question 4 à 6) |
|--|--|--|
| Correction de la réponse (1 point) | Reconnaître que la forme lexicale véhicule deux sens différents dans les phrases 1 et 2 (1 point). | Reconnaître que la même expression lexicale véhicule deux sens différents dans les phrases 1 et 2 (1 point). |
| Pertinence de la justification (1 point) | Identifier le sens de la forme lexicale dans la phrase 1 (0,5 point). OU Identifier la partie du discours de la forme lexicale à l'aide des éléments contextuels dans la phrase 1 (0,5 point). | Identifier le sens de la CLS dans la phrase 1 (0,5 point). OU Identifier la structure syntaxique de la CLS à l'aide des éléments contextuels dans la phrase 1 (0,5 point). |
| | Identifier le sens de la forme lexicale dans la phrase 2 (0,5 point). OU Identifier la partie du discours de la forme lexicale à l'aide des éléments contextuels dans la phrase 1 (0,5 point). | Identifier le sens de la CLS dans la phrase 2 (0,5 point). OU Identifier la partie du discours de la base de la CLS dans la phrase 2 (0,5 point). |

Tableau 35. – Grille de correction des sections I et II du test lexical

La section III du test lexical contient trois questions à choix multiples (sous-section 4.3.1.3). Nous avons demandé aux participants de choisir deux cooccurrents, parmi les quatre réponses proposées, qui peuvent se combiner avec un terme dans une phrase donnée. Chaque question vaut 2 points (1 point par bonne réponse). La grille de correction des questions 1 à 9 du test est illustrée dans l'annexe 24.

Les résultats du GC et du GE ont ensuite fait l'objet d'une analyse comparative que nous présentons en détail dans la sous-section 6.2.2.

6.2.2 Comparaison des résultats obtenus

Notre analyse comparative s'est déroulée en deux étapes. Nous avons d'abord décrit les résultats du GC et du GE (6.2.2.1), pour établir ensuite des comparaisons entre les deux groupes (6.2.2.2). Nous discutons aussi des écarts observés dans nos analyses intragroupe et intergroupe.

6.2.2.1 Analyse intragroupe

Nous présentons les résultats du test au moyen de graphiques à barres regroupées. Le graphique à barres regroupées permet au lecteur de rapidement reconnaître le changement d'amplitude des séries de données (Statistiques Canada, 2021c). Nous avons créé quatre graphiques pour décrire les données ci-dessous :

- 1) les notes finales des participants et leurs notes obtenues pour chaque question du test (figure 49);
- 2) les moyennes des notes du GC et du GE pour chaque section (figure 50);
- 3) le taux d'erreur⁸⁷ par section (figure 51).
- 4) les moyennes⁸⁸ des notes du GC et du GE pour chaque question (figure 52);

⁸⁷ Le taux d'erreur est le rapport entre le nombre de réponses erronées et le nombre total de réponses reçues. Les moyennes des notes et le taux d'erreur reflètent tous les deux les résultats du GC et du GE. Ces deux indicateurs nous permettent d'analyser les résultats des participants dans différentes perspectives. Le taux d'erreur est pris en compte lors de l'analyse des erreurs (voir la sous-section 6.2.3).

⁸⁸ Dans les figures 49 et 50, la moyenne des notes pour une question ou une section est exprimé en pourcentage, il s'agit de la proportion de la note obtenue par rapport au total des points du test.

La figure 49 présente les notes finales des participants ainsi que leurs notes obtenues pour chaque question. Les notes sont affichées sous forme de pourcentage. Les questions 1 à 9 sont identifiées à l'aide des codes Q1 à Q9. Le sommaire des résultats sur l'ensemble du test (figure 49) nous permet de constater que les notes du GE (E1, E2, E3, E4) sont régulièrement plus élevées que celles du GC (C1, C2, C3, C4), sauf pour C2 et E2. Cela dit, les résultats restent assez élevés pour ces deux participants. Il révèle aussi l'écart important entre la note maximale (94,4 %) et la note minimale (19,4 %).

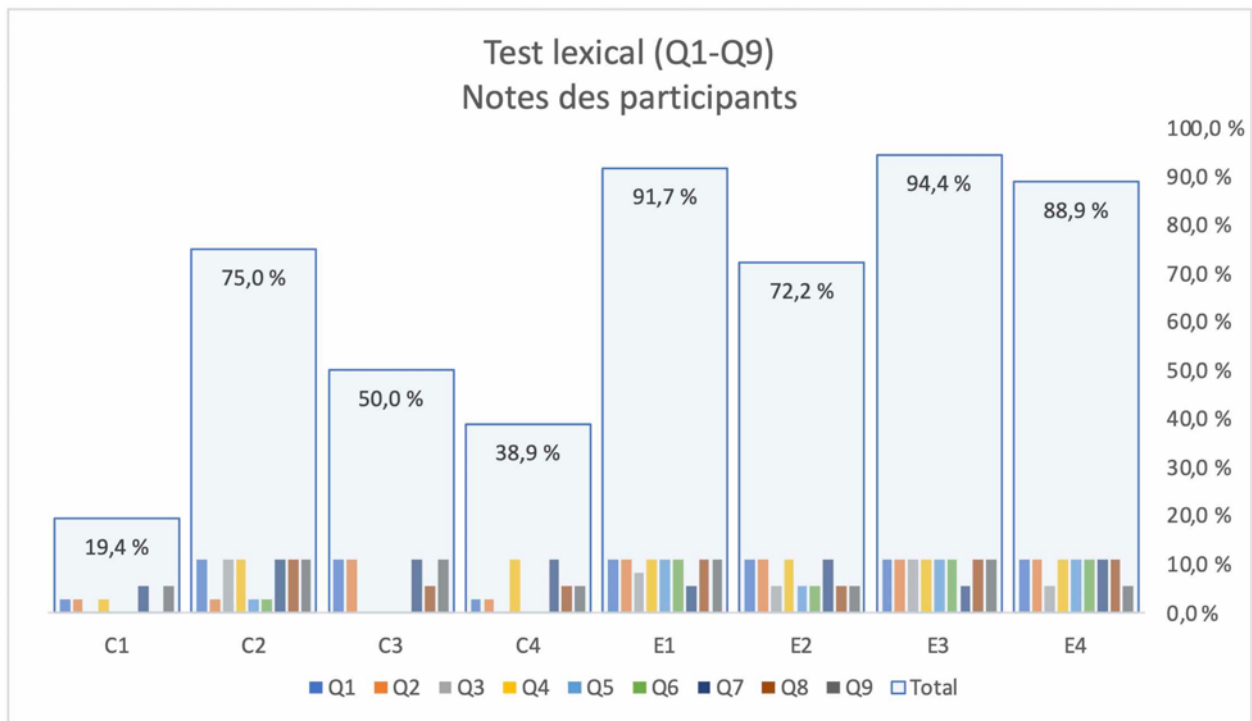


Figure 49. – Notes des participants pour le test lexical (Q1 à Q9)

Nous avons ensuite analysé les moyennes du GC et du GE pour chaque section (figure 50) et leurs taux d'erreur (figure 51), afin d'identifier les sections les mieux réussies et les moins bien réussies par chaque groupe. Pour le GC, la section la mieux réussie est la section III (Q7 à Q9) : le groupe a obtenu une moyenne de 23,6 % sur un total de 33,3 % (figure 50); le taux d'erreur pour cette section est de 50 %, soit le moins élevé parmi les trois sections (figure 51). En revanche, le GC a eu le plus de difficultés à réussir la section II (Q4 à Q6) : la moyenne du groupe pour cette section s'établit à 7,6 % sur un total de 33,3 %, soit la plus basse parmi les trois sections; le taux d'erreur s'élève à 83,3 %.

Nous avons remarqué un contraste intéressant entre les résultats du GC et du GE. Ce dernier a obtenu le meilleur résultat dans la section II, avec une moyenne de 30,6 % sur un total de 33,3 % (figure 50) et un taux d'erreur de 16,7 % (figure 51). La section III représente la plus grande difficulté pour le GE, qui a obtenu une moyenne de 26,4 % sur un total de 33,3 % et un taux d'erreur de 41,7 %.

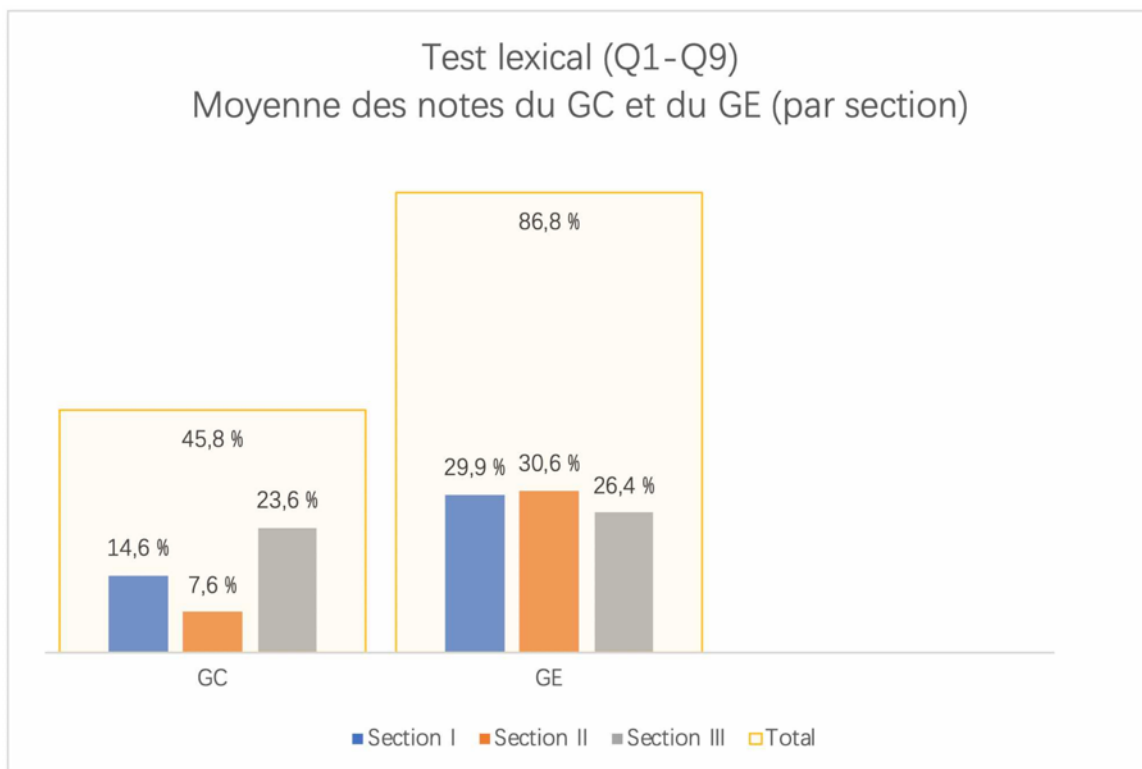


Figure 50. – Moyennes des notes obtenues par le GC et le GE (sections I à III du test lexical)

La figure 51 illustre d'importantes fluctuations lorsqu'on compare les taux d'erreur dans le même groupe. Pour la section III, nous avons observé une baisse du taux d'erreur dans le GC (50 %) par rapport aux sections I et II (66,7 % et 83,3 %), et une augmentation notable dans le GE (41,7 %) comparativement aux sections précédentes (25,0 % et 16,7 %). Par ailleurs, l'écart entre les taux d'erreur du GC et du GE se resserre dans la section III. Ces données montrent que les questions portant sur le choix de collocatifs sont les mieux réussies par le GC et les moins réussies par le GE lorsqu'on les compare aux taux d'erreur des autres questions dans le même groupe.

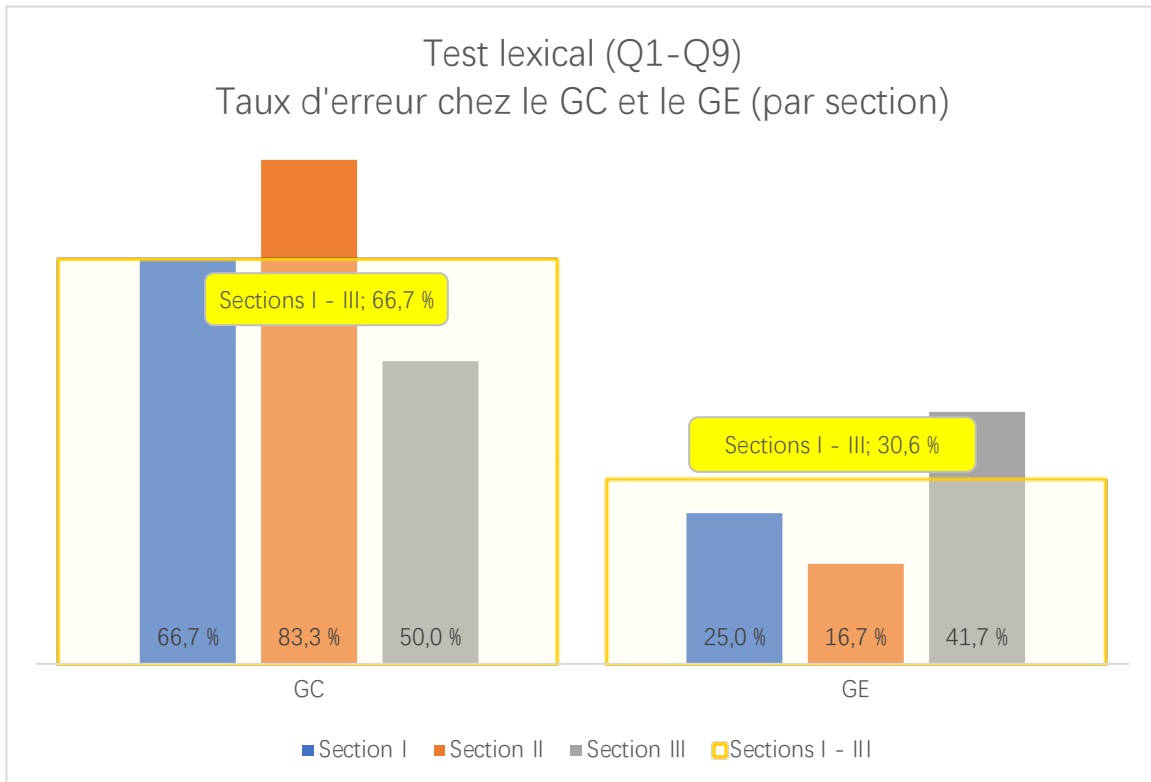


Figure 51. – Taux d’erreur du test lexical chez le GC et le GE

Dans l’ensemble, nous avons constaté une forte disparité des moyennes du GC entre les trois sections, notamment entre la section la moins réussie (7,6 % sur 33,3 %) et les deux autres sections (14,6 % sur 33,3 % et 23,6 % sur 33,3 %). On peut en déduire que les difficultés engendrées par l’ambiguïté syntaxique des CLS représentent un défi de taille pour le GC. Quant au GE, les résultats obtenus pour chaque section sont moins fluctuants par rapport au GC, ce qui suggère que le niveau d’habileté à résoudre différents types de difficultés soulevées par les CLS est relativement homogène dans le GE.

Pour obtenir une meilleure description des résultats, nous avons également comparé la dispersion des notes au sein de chaque groupe, à savoir l’étalement des notes autour de la note la plus représentative de l’ensemble du groupe. Au sein du GC, l’écart entre les notes maximale (75,0 %) et minimale (19,4 %) est de 55,6 % (figure 49). L’étendue des notes dans le GE (94,4 %-72,2 % = 22,2 %) est beaucoup plus réduite par rapport au GC. Les notes du GC sont plus dispersées par rapport à la moyenne du groupe, en comparaison avec les notes du GE.

6.2.2.2 Analyse intergroupe

Nous avons ensuite comparé les résultats obtenus par le GC et le GE en fonction des paramètres suivants : 1) la moyenne des notes obtenues pour l'ensemble du test; 2) le taux d'erreur pour l'ensemble du test; 3) la moyenne des notes cumulées par section; 4) le taux d'erreur par section et 5) la moyenne des notes obtenues pour chaque question.

La définition de ces paramètres tient compte de différents aspects des compétences évaluées au moyen du test lexical. Les paramètres 1 et 2 donnent une idée globale du niveau de compétences lexicales des deux groupes. Les paramètres 3 et 4 rendent compte du niveau d'habileté des participants à résoudre différents types de problèmes lexicaux soulevés par les CLS. Le paramètre 5 révèle les questions les mieux réussies et les moins réussies par chaque groupe.

Comparaison des résultats généraux (paramètres 1 et 2)

En premier lieu, la moyenne des notes finales du GE (86,8 %) est presque deux fois plus élevée que celle du GC (45,8 %), comme l'illustre la figure 50. Le taux d'erreur observé dans le GE (30,6 %) est notablement plus faible par rapport au GC (66,7 %) (figure 51). Le sommaire des résultats des participants (figure 49) permet aussi de visualiser les écarts entre les notes du GE (E1, E2, E3, E4) et celles du GC (C1, C2, C3, C4). Ces données montrent que, dans l'ensemble, le GE a mieux réussi le test lexical que le GC.

Comparaison par section (paramètres 3 et 4)

Nous avons ensuite analysé le niveau d'habileté des deux groupes à résoudre trois types de difficultés soulevées par les propriétés syntactico-sémantiques des CLS. Chaque type de difficulté fait l'objet d'une section du test. D'un point de vue général, le GE a eu les meilleurs résultats pour toutes les trois sections du test. D'une part, la moyenne des notes du GE est plus élevée que celle du GC pour chaque section du test (29,9 % vs. 14,6 % pour la section I; 30,6 % vs 7,6 % pour la section II et 26,4 % vs 23,6 % pour la section III), comme l'illustre la figure 50. L'écart le plus important entre les deux groupes est observé dans la section II portant sur la distinction des CLS de forme identique. La moyenne du GE (30,6 %) est environ quatre fois plus élevée que celle du GC (7,6 %). En revanche, dans la section III portant sur le choix de collocations, l'écart entre le GE (26,4 %) et le GC (23,6 %) est réduit au minimum par rapport aux deux autres sections. D'autre part, l'analyse des taux d'erreur du GC et du GE nous a amenée à dresser les

mêmes constats. Le taux d'erreur du GE est beaucoup plus faible que celui du GC pour chacune des trois sections (figure 51). Les écarts maximal et minimal sont observés respectivement dans la section II et la section III.

Comparaison par question (paramètre 5)

Enfin, nous avons comparé la moyenne des notes obtenues par les deux groupes pour chaque question du test (figure 52).

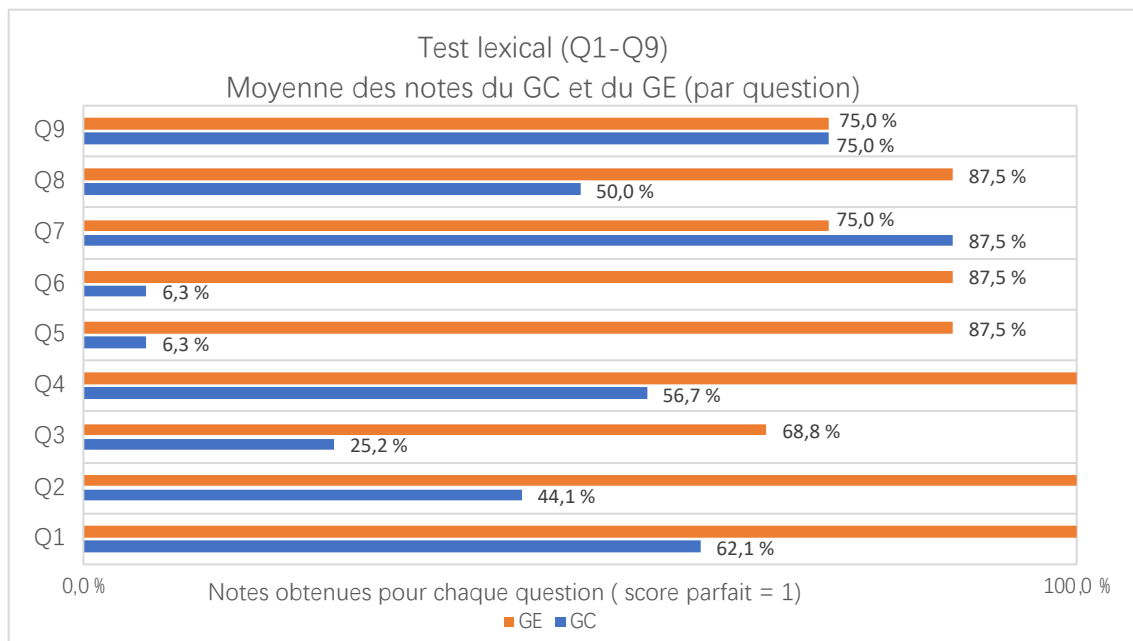


Figure 52. – Moyennes des notes obtenues par le GC et le GE (Q1 à Q9 du test lexical)

Le GE a obtenu la moyenne maximale (100 %, soit le score parfait) pour les Q1, Q2 et Q4, la moyenne minimale est observée dans la Q3 (68,8 %). Le GC a obtenu la moyenne maximale (87,5 %) pour la Q7. Les questions les moins réussies comprennent les Q5 et Q6. Nous en déduisons qu'en général, le GE parvient à mieux distinguer les formes lexicales polysémiques dans les CLS. Cependant, il a moins bien réussi les questions liées au choix de collocations (Q7 à Q9) comparativement aux autres questions. Cela dit, le GE a tout de même mieux réussi ces questions que le GC. Le GC a plus de difficulté à distinguer les sens et les structures syntaxiques associés à des CLS ayant la même forme, ce qui consolide nos observations dans les étapes d'analyse précédentes.

Quant aux écarts entre les moyennes des deux groupes, nous avons observé l'écart maximal entre le GC (6,3 %) et le GE (87,5 %) dans les Q5 et Q6, soit les questions les moins bien réussies par le GC. L'écart minimal est à 0, autrement dit, les deux groupes ont obtenu la même moyenne (74,8 %) pour la Q9. Soulignons que la Q7 est la seule question pour laquelle la moyenne du GC (87,5 %) est supérieure à celle du GE (75 %).

Il convient de rappeler la forte disparité dans les résultats du GC pour les trois sections (figure 50). Quant au GE, les écarts observés sont moins marqués par rapport au GC, ce qui suggère que chez le GE, les niveaux d'habileté à résoudre les trois types de difficultés sont relativement homogènes entre les participants (voir aussi la sous-section 6.2.2.1).

En guise de synthèse, les résultats de notre analyse comparative fournissent des indices sur les points forts de chaque groupe et les points faibles à améliorer dans l'acquisition des compétences lexicales par les participants.

6.2.3 Analyse des erreurs

Dans un troisième temps, nous avons procédé à l'analyse des erreurs repérées. Nous nous sommes inspirée de la méthodologie d'analyse des erreurs (AE) proposée par Corder (1967, 1981) en enseignement-apprentissage de langues étrangères. L'AE comprend trois étapes, comme le souligne Perdu (1980) : 1) relever les erreurs à partir des productions linguistiques des apprenants; 2) les décrire de façon systématique en fonction de différents paramètres, dont la nature des erreurs, leur fréquence et leur probabilité d'apparition et 3) examiner la (les) cause(s) des erreurs.

6.2.3.1 Méthode d'analyse

Inspirée de l'AE, nous avons tenté de fournir une description systématique des erreurs. Le tableau 36 illustre notre grille d'analyse des erreurs. Celles-ci ont été d'abord catégorisées en fonction des compétences lexicales évaluées. Par exemple, les erreurs relevées dans la section 1 (questions 1 à 3) sont répertoriées dans la même catégorie puisque cette section vise à mesurer l'habileté à distinguer les sens associés à une même forme lexicale dans des contextes différents. La nature de chaque catégorie d'erreurs a été identifiée. Nous avons ensuite comptabilisé la fréquence de chaque erreur repérée dans les réponses du GC et celles du GE (tableau 36). Enfin, nous avons tenté d'expliquer la (les) cause(s) des erreurs. À cette étape, nous avons pris en

compte des justifications des participants dans les questions à réponse courte (questions 1 à 6) pour découvrir les sources des erreurs. Les résultats de notre analyse sont présentés dans l'annexe 25.

| | Habilité évaluée | Erreur repérée | | Nature d'erreur | Fréquence-GC (sur 4) | Fréquence-GE (sur 4) | Explication de l'erreur |
|-------------|---|----------------|-------|-----------------|----------------------|----------------------|-------------------------|
| | | Erreur 1 | [...] | | | | |
| Section I | Distinguer les sens associés à une forme lexicale | Erreur 1 | | | | | |
| | | [...] | | | | | |
| | | Erreur n | | | | | |
| Section II | Distinguer les sens associés à une combinaison lexicale | Erreur 1 | | | | | |
| | | [...] | | | | | |
| | | Erreur n | | | | | |
| Section III | Choisir un cooccurrent approprié d'un terme | Erreur 1 | | | | | |
| | | [...] | | | | | |
| | | Erreur n | | | | | |

Tableau 36. – Tableau d'analyse des erreurs repérées dans le test lexical

6.2.3.2 Résultats d'analyse

À l'issue de notre analyse, les erreurs sont regroupées en trois catégories (annexe 25). La première catégorie d'erreurs concerne la compréhension erronée du sens d'une forme lexicale. Ces erreurs ont été relevées dans la section I du test (Q1 à Q3). Elles s'expliquent par trois raisons selon notre analyse des justifications fournies par les participants. Premièrement, certains participants du GC n'étaient pas en mesure d'identifier le sens d'un terme (par exemple, 网络 « en ligne ») dans une CLS qui contient celui-ci (par exemple, 网络学习 « apprentissage en ligne » figurant dans la phrase 2 de la Q1). Ils affirment ne pas avoir trouvé un élément de réponse dans les ressources lexicales consultées.

Deuxièmement, certains participants du GC ont interprété le sens d'un terme de façon erronée. Par exemple, ils ont associé au terme 网络 (« en ligne ») un sens erroné (« réseau ») dans la CLS 网络学习 (« apprentissage en ligne »), car ils ont trouvé les deux sens dans un même article consacré à 网络 dans les ouvrages de référence consultés. En outre, ce type d'erreurs s'explique

par le transfert négatif (Tardif, 1999)⁸⁹. Dans la phrase 2 de la question 2 (Q2), deux participants du GC ont attribué une acception erronée à la forme lexicale 移动^{yí dòng} (« amovible ») figurant dans la CLS 移动硬盘^{yí dòngyìng pán} (« disque dur amovible »). Selon les ouvrages de référence qu'ils ont consultés, 移动^{yí dòng} signifie « mobile » dans la CLS 移动设备^{yí dòng shè bèi} (« appareil mobile »), ce qui les a amenés à croire que 移动 est associé au même sens dans la CLS 移动硬盘^{yí dòngyìng pán} (« disque dur amovible ») puisque 硬盘^{yìng pán} (« disque dur ») et 设备^{shè bèi} (« appareil ») désignent des dispositifs informatiques. Du point de vue didactique, il s'agit d'un cas de transfert négatif causé par la surgénéralisation des connaissances acquises en L2 (Tardif, 1999).

Troisièmement, des participants ont confondu les acceptions associées à un mot multicatégoriel. Par exemple, deux participants du GC ont confondu les deux acceptions nominales de la forme lexicale 更新^{gēng xīn} (phrase 1 de la Q3), soit le nom d'action « mise à jour » et le nom de résultat « mise à jour ». Dans les ressources lexicales auxquelles ils ont eu recours, ils n'ont trouvé qu'un seul sens nominal dans l'article consacré à 更新^{gēng xīn}. Quant aux participants du GE qui ont utilisé le CHINOINFO, certains d'entre eux ont considéré le nom d'action 更新^{gēng xīn} (« mise à jour ») et le nom de résultat 更新^{gēng xīn} (« mise à jour ») comme ayant le même sens. Selon eux, la seule différence entre ces deux termes est leur partie du discours. Un participant du GE s'est trompé de la partie du discours à laquelle appartient 更新^{gēng xīn} (« mise à jour-nom de résultat ») figurant dans la CLS 安装^{ān zhuāng} 更新^{gēng xīn} (« installer une mise à jour »). Il a considéré la CLS comme étant composée de deux verbes d'action (安装^{ān zhuāng} « installer » et *更新^{gēng xīn} « mettre à jour ») coordonnés.

En ce qui concerne la fréquence de la première catégorie d'erreurs, rappelons que le taux d'erreur dans le GC (66,7 %) est plus élevé par rapport au GE (25,0 %), comme l'illustre la figure 51 (voir aussi l'annexe 25).

La deuxième catégorie d'erreurs est définie comme l'interprétation erronée d'une combinaison lexicale sur le plan sémantique ou syntaxique. Ces erreurs ont été repérées dans la section II du test (Q4-Q6). Elles découlent de deux difficultés d'apprentissage. Premièrement, certains participants du GC ont eu de la difficulté à distinguer les acceptions d'un mot multicatégoriel

⁸⁹ Nous reportons le lecteur à la section 3.3 portant sur l'approche cognitive, soit le volet didactique de notre cadre théorique.

figurant dans des CLS. Dans la Q5, ils n'étaient pas en mesure d'identifier les deux sens associés à 备份 dans la séquence 备份数据 qui correspond à deux CLS (« sauvegarder les données » et « données sauvegardées »), car les ressources lexicales qu'ils ont consultées ne distinguent pas les trois acceptions du mot multicatégoriel 备份, soit « sauvegarder », « sauvegarde (nom d'action) » et « sauvegardé ».

En outre, des participants des deux groupes n'étaient pas en mesure d'identifier la structure d'une CLS à l'aide de son contexte d'utilisation. Dans la phrase 2 de la Q6, la CLS 共享文件 « fichier partagé » est précédé d'une série de verbes (查看 « consulter », 编辑 « éditer » et 删除 « supprimer »). On aurait pu en déduire que la CLS est de nature nominale et qu'elle est associée à la structure « modificateur-nom modifié ». Par conséquent, le mot multicatégoriel 共享 véhicule, dans cette phrase, un sens adjectival (« partagé ») et modifie 文件 (« fichier »).

Quant au taux d'erreur, nous avons observé l'écart le plus important entre le GC (83,3 %) et le GE (16,7 %) dans la section II (figure 51). De plus, le taux d'erreur observé chez le GE pour la section II (16,7 %) est plus faible par rapport aux deux autres sections (25,0 % et 41,7 %). Ces données révèlent que le GE a mieux réussi les questions liées à distinction des CLS ayant une forme identique et des structures syntaxiques différentes, tandis que le GC a éprouvé plus de difficulté à résoudre les problèmes lexicaux soulevés par l'ambiguïté syntaxique des CLS.

La troisième catégorie d'erreurs est la construction erronée des combinaisons lexicales. Nous les avons identifiées dans la section III du test (Q7-Q9). Les causes de ces erreurs sont plus variées par rapport aux deux autres catégories. Premièrement, les participants ont eu de la difficulté à identifier les nuances sémantiques entre deux collocatifs d'un terme. Dans la Q7, des participants n'étaient pas en mesure de distinguer 访问 (« accéder ») et 入侵 (« s'introduire ») lorsque ces deux termes se combinent avec 网站 (« site Web »). Bien que les deux termes semblent pouvoir être traduits par le même mot français (*accéder*), 入侵 (« s'introduire ») est chargé d'une connotation négative. 入侵网站 (« s'introduire dans un site Web ») dénote l'activité d'accéder à un site Web de façon illicite.

Deuxièmement, l'erreur de construction d'une CLS est liée à l'interférence de la L1. Dans la Q8, des participants ont confondu le terme 移动 (« mobile ») et l'unité lexicale 移动 (« déplacer ») en langue générale. Ils ont associé 移动 qui semble signifier 'déplacer' avec le terme 文件

« fichier » pour exprimer le sens de ‘faire passer un fichier d’un emplacement à un autre’. Contrairement au français, la combinaison lexicale ainsi formulée (*^{yí dòng wén jiàn} 移动文件 ‘déplacer’-‘fichier’) est inacceptable en chinois. Il s’agit d’un cas de transfert négatif causé par l’interférence de la L1 (Tardif, 1999). Du point de vue linguistique, ce type de constructions erronées est un cas particulier de *greffe collocationnelle interlinguistique* définie par Polguère (voir la sous-section 4.2.1.3).

Troisièmement, le choix de collocatifs inappropriés pour un terme prédicatif s’explique également par le manque de connaissances des actants sémantiques du terme. Dans la Q9, des participants ont associé le terme ^{chuán shū} 传输 (« transférer ») à ^{yìng pán} 硬盘 (« disque dur ») qui dénote un périphérique de stockage. Cette construction (*^{chuán shū yìng pán} 传输硬盘 ‘transférer’-‘disque dur’) est erronée puisque le deuxième actant de ^{chuán shū} 传输 (« transférer ») est réalisé par les termes dénotant un type de données,

6.2.4 Synthèse des réponses à la question libre

Dans un quatrième temps, nous avons analysé les données qualitatives collectées à l’aide de la question 10 (Q10) du test lexical. Il s’agit d’une question libre dans laquelle on invite les participants à identifier les exercices qui leur paraissent particulièrement difficiles (voir aussi la sous-section 4.3.1.3).

Notre démarche d’analyse s’est déployée en trois étapes. Nous avons d’abord classé les questions mentionnées par les participants en trois catégories selon les sections dans lesquelles elles apparaissent. Rappelons que chaque section du test porte sur l’une des trois compétences lexicales évaluées (sous-section 4.3.1.3). À cette étape, nous avons consigné le numéro des questions et le nombre de participants qui les ont mentionnées. Nous avons ensuite résumé les commentaires sur le niveau de difficulté de l’ensemble du test lexical. Enfin, nous avons établi les catégories d’analyse à partir des facteurs identifiés. Les réponses des participants ont été codifiées en fonction des catégories élaborées.

À l’issue de notre analyse, nous avons dressé quelques constats concernant les difficultés que les participants ont éprouvées au cours du test lexical.

En premier lieu, la grande majorité des répondants a signalé le niveau de difficulté élevé du test lexical dans son ensemble. Parmi les huit participants ayant répondu à la Q10, sept répondants ont indiqué que les neuf questions étaient difficiles pour eux.

En deuxième lieu, la section II portant sur la distinction sémantique des CLS ayant la même forme et des structures syntaxiques différentes représente un défi majeur pour le GC. Trois des quatre participants du GC ont estimé que les Q5 et Q6 étaient particulièrement difficiles puisqu'ils ne sont pas parvenus à trouver des éléments de réponse, malgré leur recours à des ouvrages de référence et à la traduction automatique.

En troisième lieu, le GE a éprouvé plus de difficultés à réussir les questions de choix de collocatifs. Deux des quatre participants du GE ont estimé que la section III (Q7 à Q9) était la plus difficile, car ils ont affirmé ne pas avoir le vocabulaire nécessaire pour répondre tout de suite à cette question. De plus, la Q9 était la plus difficile selon un répondant du GC, étant donné que l'énoncé donné dans cette question « comporte beaucoup de termes sans pinyin ».

En général, les deux derniers constats que nous avons dressés consolident les résultats de notre analyse des taux d'erreur du GC et du GE.

L'analyse des commentaires des participants nous a permis d'identifier d'autres facteurs qui ont influencé la performance des participants dans le test. D'abord, la densité élevée des termes dans les contextes proposés constitue une source de difficulté non négligeable. Six des huit répondants (dont tous les quatre participants du GC et deux participants du GE) ont affirmé que le nombre de mots inconnus dans les questions est élevé. Ils ont dû faire beaucoup de recherche dans les ouvrages de référence pour comprendre chaque terme inconnu. De plus, ils ne sont pas parvenus à comprendre tous les énoncés donnés dans les questions, malgré les recherches chronophages, ce qui a amené certains participants à recourir à la traduction automatique pour un décodage rapide. Ensuite, le manque de connaissances du domaine de l'informatique a été évoqué par un participant du GC. Il a dû traduire les termes inconnus à l'aide de Google translate, afin de pouvoir rechercher des renseignements rédigés en français pour comprendre les connaissances véhiculées par ces termes.

6.2.5 Synthèse

En résumé, l'analyse des réponses au test lexical offre un portrait intéressant du niveau d'habileté des participants à résoudre les problèmes lexicaux liés aux CLS d'une part, et d'autre part, à identifier quelques difficultés qui font particulièrement obstacle à la maîtrise des CLS.

La comparaison des résultats obtenus par les deux groupes nous amène à constater que le GE a notablement mieux réussi le test lexical que le GC, tant pour l'ensemble du test que pour chaque type de compétence évaluée. L'écart le plus important entre les deux groupes s'observe dans leur niveau d'habileté à distinguer les CLS de forme identique sur le plan sémantique et syntaxique (section II du test). On peut en déduire que le GC disposait de peu de moyens pour résoudre ce type de difficulté, comparativement au GE qui a eu recours au CHINOINFO lors du test. Ce résultat d'analyse valide en partie notre hypothèse initiale qui suppose que l'ambiguïté syntaxique des CLS contenant une forme lexicale polysémique constitue un défi majeur pour les apprenants non natifs. C'est pour pallier l'absence de matériel didactique du COS à cet égard que nous nous sommes attachée à concevoir une ressource lexicale décrivant les propriétés syntaxico-sémantiques des CLS. Les résultats expérimentaux semblent montrer que ce type de ressource est utile.

En ce qui concerne les écarts observés au sein de chaque groupe, les notes du GE sont moins dispersées par rapport au GC. Au sein du GE, la disparité entre les notes des différentes sections est moins importante que celle observée chez le GC.

6.3 Analyse des réponses aux questionnaires d'appréciation

Nous présentons, dans ce qui suit, les résultats de l'analyse des renseignements recueillis au moyen des questionnaires d'appréciation et de commentaire (6.3.1). Nous formulons enfin quelques remarques sur le traitement des réponses aux questionnaires (6.3.2).

6.3.1 Questionnaire d'appréciation et de commentaire

Les questionnaires 2a (annexe 13) et 2b (annexe 14) visent à recueillir les commentaires du GC et du GE sur leur participation au test lexical et leur recours aux ressources lexicales durant le test. Ces deux questionnaires sont composés de trois types de questions : les questions à échelle

d'évaluation, les questions à choix de réponses et les questions ouvertes. Le traitement des données collectées a été effectué en fonction des types de questions.

6.3.1.1. Questions à échelle d'évaluation

Les questions à échelle d'évaluation dans les questionnaires 2a et 2b ont été catégorisées selon le type des renseignements recueillis (tableau 37). Pour analyser les données recueillies au moyen de ces questions, nous avons d'abord codé les réponses afin de les transformer en données quantifiables. Pour chaque question, un numéro a été attribué à chaque option de réponse dans l'échelle d'évaluation (par exemple, 1 = fortement en désaccord; 2 = plutôt en désaccord; 3 = plutôt en accord; 4 = fortement en accord; 5 = ne s'applique pas). Les données ont ensuite été analysées selon les catégories d'analyse préétablies (tableau 37).

| | Type de renseignements recueillis | Questions (Questionnaire) |
|---|--|--------------------------------------|
| 1 | Appréciation de l'organisation du test | Q1.1 – Q1.7 (2a) Q1.1 – Q1.7 (2b) |
| 2 | Appréciation du niveau de difficulté du test | Q2.1 – Q2.4 (2a) Q2.1 – Q2.4 (2b) |
| 3 | Utilité des outils de référence consultés pour résoudre les problèmes lexicaux | Q5 (2a) Q3 (2b) Q10 (2b) |
| 4 | Niveau de satisfaction globale envers les outils de référence utilisés | Q7 (2a) Q4 (2b) |

Tableau 37. – Catégorisation des données collectées au moyen des questions à échelle d'évaluation dans les questionnaires 2a et 2b

Appréciation de l'organisation du test

Dans questionnaires 2a et 2b, les questions Q1.1 à Q1.7 visent à recueillir l'appréciation des participants sur l'organisation du test. Dans l'ensemble, les deux groupes ont donné des avis plutôt favorables sur cette série de questions. Chez le GC (figure 53), la majorité des participants (trois sur quatre) était « fortement d'accord » avec le fait que les objectifs d'évaluation (Q1.2) et les modalités du test (Q1.3) ont été présentés de façon claire, que les tâches préparatoires étaient instructives (Q1.4), un participant était « plutôt d'accord » avec ces énoncés. Au cours du test, trois des quatre participants du GC ont communiqué avec nous. Ils étaient tous « fortement d'accord » que nous avons fourni des réponses rapides et pertinentes à leurs questions (Q1.6). Tous les participants du GC étaient « fortement d'accord » que les supports informatiques utilisés dans le cadre du test étaient adéquats. Deux participants étaient

« fortement d'accord » avec le fait que la documentation qui leur a été transmise avant le test était utile (Q1.1) et que la durée du test était adéquate (Q1.5). Un répondant était plutôt insatisfait du temps alloué au test.

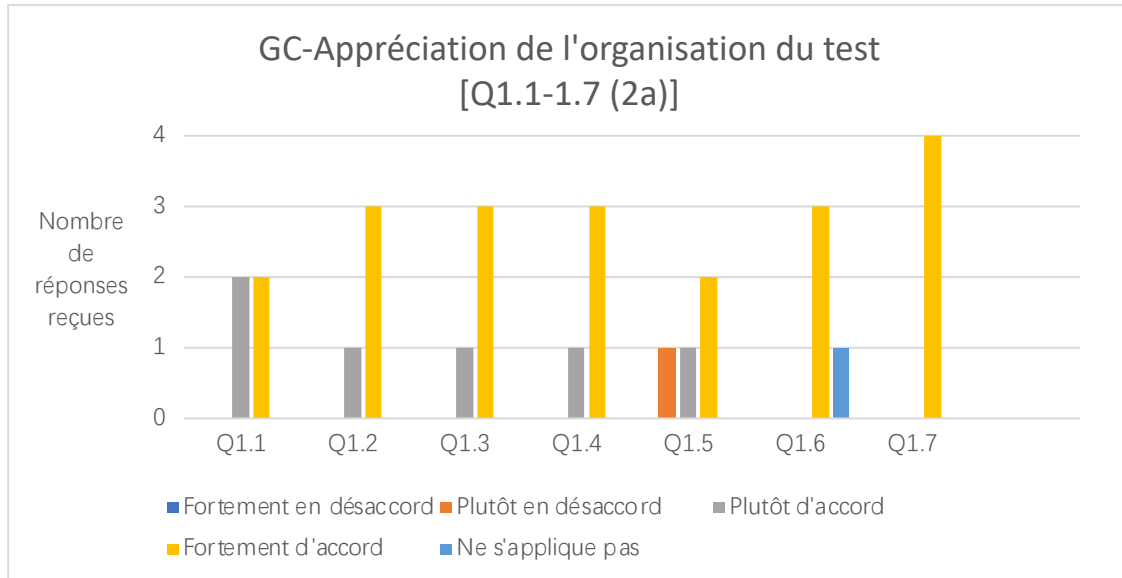


Figure 53. – Appréciation de l'organisation du test par le GC (Q1.1-1.7 du questionnaire 2a)

Le GE présente un taux de satisfaction semblable au GC, comme l'illustre la figure 54. Pour les Q1.1, Q1.2, Q1.4 et Q1.7, trois répondants étaient « fortement d'accord », auxquels s'ajoute un participant qui était « plutôt d'accord ». Pour la Q1.3, les répondants « fortement d'accord » et « plutôt d'accord » étaient en nombre égal. Comme le GC, tous les participants du GE qui nous ont posé des questions durant le test étaient « fortement d'accord » de l'efficacité de notre invention (Q1.6). Les opinions des GE sur la pertinence de la durée du test (Q1.5) sont plus diversifiées par rapport aux autres questions. L'un des quatre participants était « plutôt en désaccord »; deux répondants étaient « plutôt d'accord » et « fortement d'accord », auxquels s'ajoute un autre répondant qui a exprimé un avis neutre. En résumé, les deux groupes ont apprécié positivement l'organisation du test.

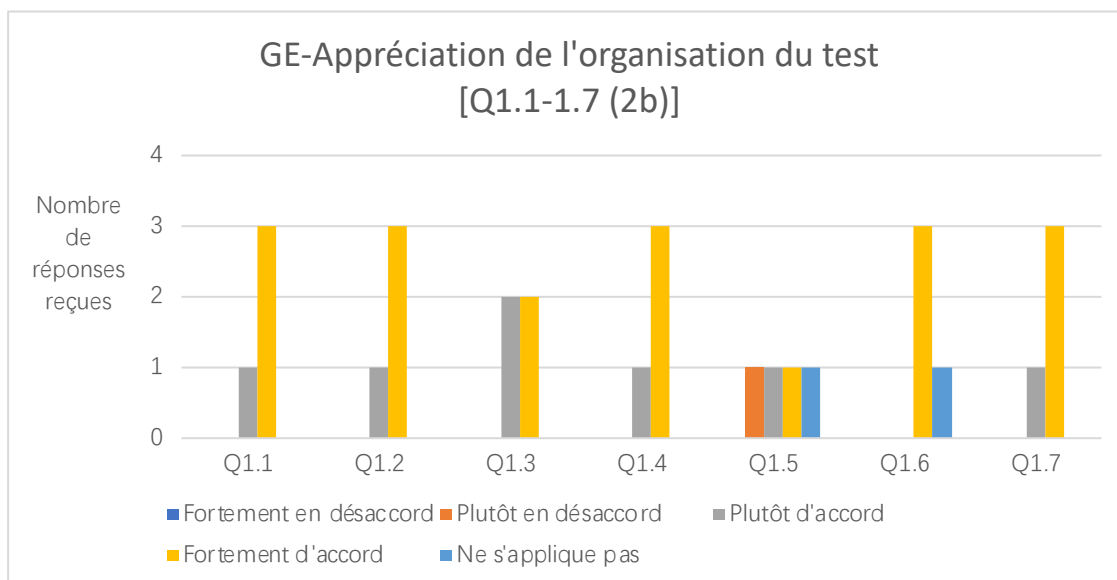


Figure 54. – Appréciation de l'organisation du test par le GE (Q1.1-1.7 du questionnaire 2b)

Appréciation du niveau de difficulté du test

Les questions Q2.1 à Q2.4 des questionnaires 2a et 2b ont pour objectif de recueillir les opinions des participants sur le niveau de difficulté du test. Les Q2.1 à Q2.3 portent sur les trois sections du test. La Q2.4 porte sur l'ensemble du test.

Comme l'illustre la figure 55, les réponses du GC se distribuent entre « très difficile » (68,8 %), « plutôt difficile » (18,8 %) et « plutôt facile » (12,5 %). La distinction sémantique d'une forme lexicale dans différentes CLS (Q2.1) était un défi de taille pour trois répondants du GE. La section I portant sur ce type de difficulté était « très difficile » pour deux étudiants et « plutôt difficile » pour un étudiant. L'autre répondant estimait que ce type de difficulté était « plutôt facile » à résoudre.

La section II portant sur la distinction sémantique des CLS de forme identique (Q2.2) était « très difficile » pour trois répondants du GC et « plutôt facile » pour un répondant.

Le choix des cooccurrents appropriés (Q2.3) faisant l'objet de la section III était « très difficile » selon deux répondants du GC et « plutôt difficile » selon les deux autres répondants du groupe.

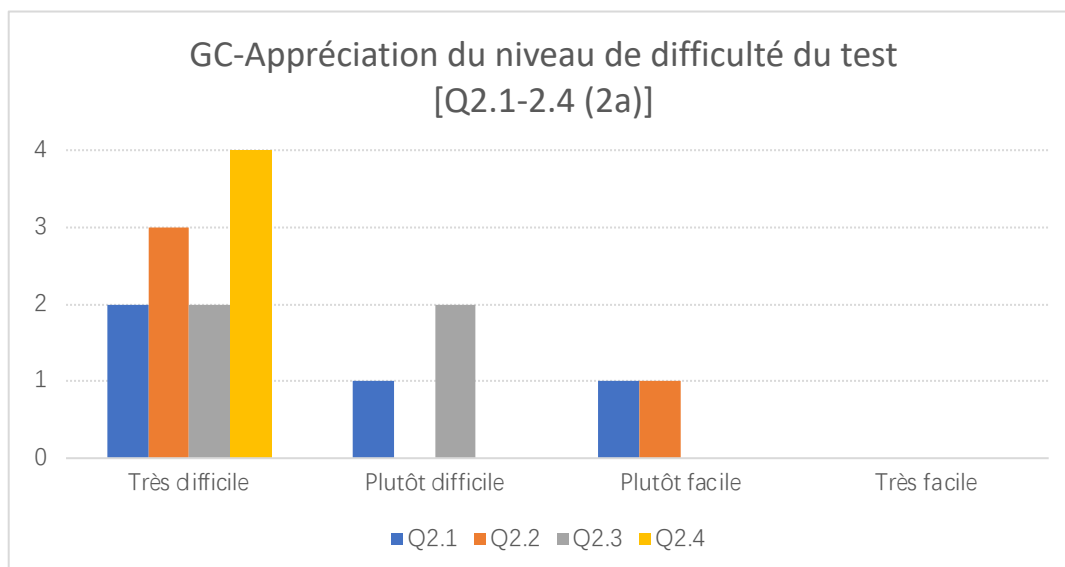


Figure 55. – Appréciation du niveau de difficulté du test par le GC (Q2.1-2.4 du questionnaire 2a)

Les réponses du GE se distribuent entre « plutôt difficile » (56,2 %) et « plutôt facile » (62,5 %), comme on peut le constater dans la figure 56. Trois répondants estimaient les deux premiers types de difficultés (Q2.1 et Q2.2) « plutôt faciles », l’autre répondant du groupe répondant les jugeait « plutôt difficiles ». Pour le troisième type de difficulté (Q2.3), les réponses sont réparties dans une proportion opposée.

En ce qui concerne le niveau de difficulté du test en général (Q2.4), tous les quatre participants du GC ont estimé que le test était « très difficile », tandis que tous les répondants du GE l’ont jugé « plutôt difficile ».

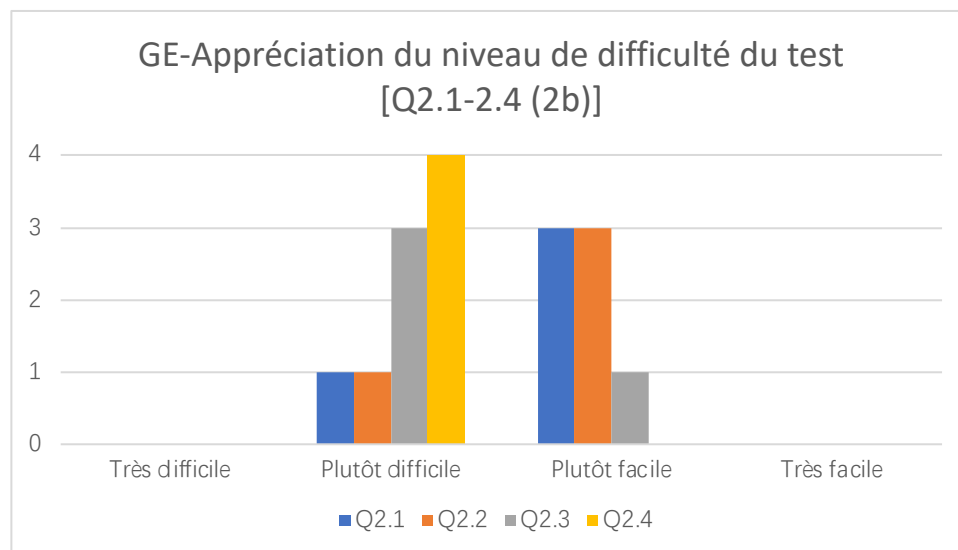


Figure 56. – Appréciation du niveau de difficulté du test par le GE (Q2.1-2.4 du questionnaire 2b)

En guise de synthèse, l'analyse des données suggère que le GE a eu moins de difficulté à réaliser le test lexical. Par ailleurs, nous avons constaté que la section II portant sur la distinction sémantique des CLS de forme identique était la plus difficile pour le GC par rapport aux autres sections, tandis que la section III portant sur le choix des cooccurrents appropriés était la plus difficile pour le GE. Ces constats sont en phase avec les analyses des résultats des deux groupes (voir la sous-section 6.2.2) : ils ont respectivement obtenu la moyenne minimale des notes dans la section qu'ils jugeaient la plus difficile.

Utilité des outils de référence consultés

Dans les questionnaires 2a et 2b, nous avons conçu trois questions pour savoir si les outils de référence que les participants ont consultés au cours du test étaient utiles pour résoudre les différents types de problèmes lexicaux. La Q5 du questionnaire 2a s'adresse au GC. Comme on peut le constater dans la figure 57, la majorité des participants (75 %) pensait qu'il était « plutôt difficile » de trouver des éléments de réponse à chaque type de problèmes lexicaux dans les outils de référence qu'ils ont utilisés. Deux participants ont exprimé un avis plus favorable en ce qui concerne l'utilité des ressources à résoudre le premier (25 %) et le troisième (25 %) type d'exercices. Quant au deuxième type de difficulté, un participant (25 %) estimait qu'il est « très difficile » de distinguer les sens des CLS ayant la même forme à l'aide des ressources lexicales consultées.

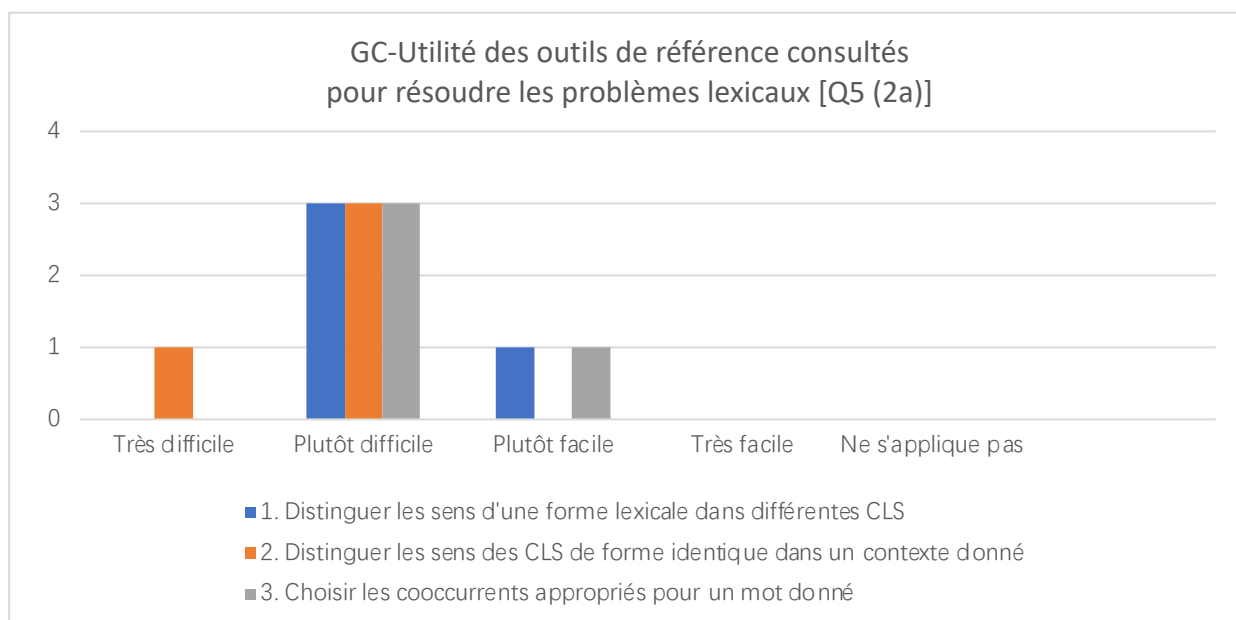


Figure 57. – Opinions du GC sur l'utilité des outils de référence consultés [Q5 du questionnaire 2a]

Les Q3 et Q10 du questionnaire 2b recueillent les appréciations du GE sur l'utilité du CHINOINFO et celle des autres ressources lexicales auxquelles il a eu recours. Nous avons comparé les réponses du GE à ces questions (figure 58), afin d'avoir un premier aperçu de l'apport du CHINOINFO.

Nous avons observé une similarité entre les opinions du GC et GE sur l'utilité des outils de référence élaborés avant le CHINOINFO. Comme l'illustre la partie gauche de la figure 58, les réponses du GE se concentrent sur « plutôt difficile » (75 %), tout comme ce que nous avons observé chez le GC (figure 57). Deux participants du GE pensaient qu'il était « plutôt facile » de distinguer les sens d'une forme lexicale à l'aide des ressources lexicales consultées, soit le premier type de difficulté. Il était « très difficile » de trouver des réponses au deuxième (25 %) ou au troisième (25 %) type de problème lexicaux selon deux répondants.

Le GE a donné un avis plus favorable quant à l'utilité du CHINOINFO. Comme on peut le constater dans la partie droite de la figure 58, les réponses se distribuent entre « plutôt difficile » (33,3 %), « plutôt facile » (41,7 %) et « très facile » (25 %). Trois des quatre participants du GE estimaient qu'il était « plutôt facile » (50 %) ou « très facile » (25 %) de trouver des éléments de réponses au premier et au troisième type de problèmes lexicaux, par rapport au nombre de participants qui ont répondu « plutôt difficile » (25 %). Une moitié des participants trouvaient qu'il était « plutôt facile » (25 %) ou « très facile » de résoudre le deuxième type de difficulté à l'aide du CHINOINFO, par rapport à l'autre moitié des participants qui ont répondu « plutôt difficile ». Ces données laissent croire que, dans le cadre de ce test lexical, le CHINOINFO fournit une aide plus bénéfique aux apprenants du GE par rapport à d'autres outils de référence.

D'un point de vue général, les résultats d'analyse suggèrent que le GE était mieux outillé que le GC pour résoudre les différents types de problèmes lexicaux soulevés par les CLS.

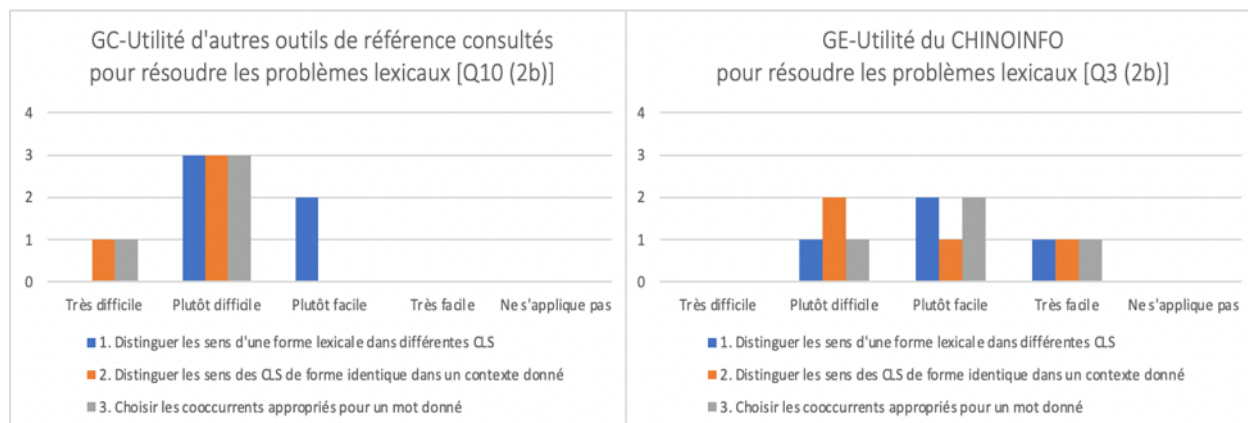


Figure 58. – Opinions du GE sur l'utilité des outils de référence consultés [Q3 et Q10 du questionnaire 2b]

Niveau de satisfaction globale envers les outils de référence utilisés

La Q7 du questionnaire 2a et la Q4 du questionnaire 2b invitent les participants à exprimer leur degré de satisfaction par rapport aux ressources lexicales sous cinq aspects : 1) la convivialité des ressources; 2) la qualité du contenu des ressources consultées; 3) la méthode de description lexicale; 4) l'utilité des ressources pour résoudre les problèmes lexicaux et 5) l'efficacité pédagogique.

En général, le taux de satisfaction du GE est plus élevé que celui observé chez le GC. Les réponses du GC à la Q7 du questionnaire 2a (figure 59) se concentrent autour de « neutre », auxquelles s'ajoute une proportion importante de « insatisfait(e) ». Parmi les 20 réponses reçues (cinq réponses par participant), on trouve une seule réponse indiquant que le participant était « satisfait » de la convivialité des outils de référence utilisés, ce qui révèle un taux de satisfaction relativement faible de la part du GC.

En outre, les réponses « insatisfait(e) » se concentrent sur le quatrième et le cinquième aspect évalué, soit l'utilité des ressources pour résoudre les problèmes lexicaux et l'efficacité pédagogique. On peut en déduire que les ressources lexicales auxquelles le GC a eu recours ne répondent pas de façon adéquate à leurs besoins d'apprentissage en ce qui concerne la résolution des difficultés liées aux CLS.

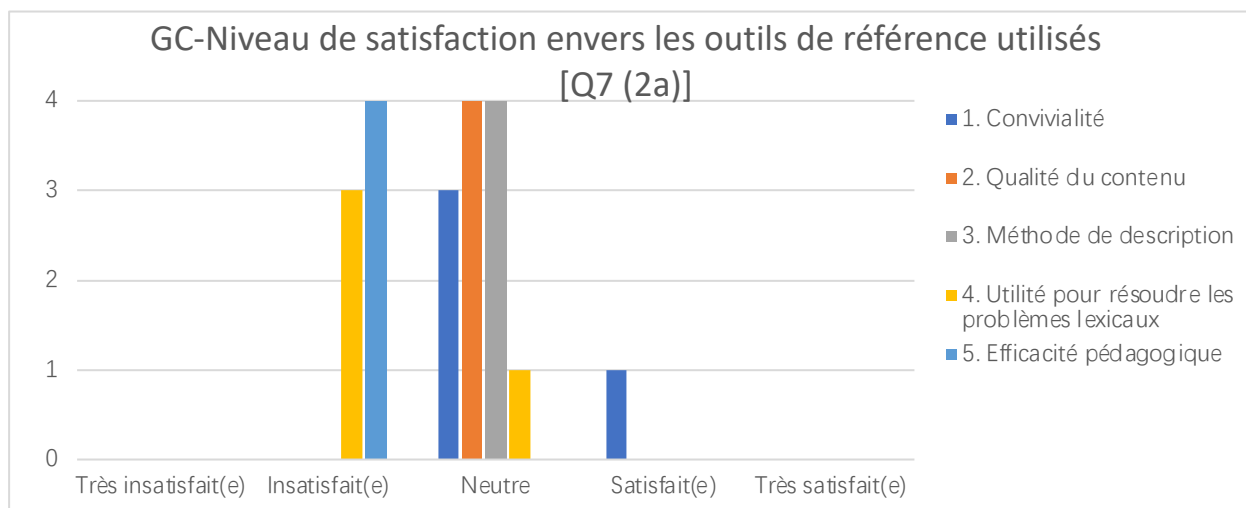


Figure 59. – Niveau de satisfaction du GC envers les outils de référence utilisés (Q7 du questionnaire 2a)

Le GE a exprimé des opinions plus favorables que le GC. Les réponses du GE à la Q4 du questionnaire 2b (figure 60) se distribuent entre « neutre », « satisfait(e) » et « très satisfait(e) ». Il convient de noter que les réponses « satisfait(e) » se distribuent dans tous les cinq aspects évalués. Nous avons observé le taux de satisfaction le plus élevé dans les aspects 2 et 4, soit la qualité du contenu et l'utilité des ressources : trois des quatre participants étaient « satisfaits » et un autre était « très satisfait » pour chacun de ces deux aspects. Les participants ont donné un avis neutre en ce qui concerne la convivialité des ressources, la méthode de description lexicale et l'efficacité pédagogique. Tout compte fait, le GE était plus satisfait de son recours aux outils de référence.

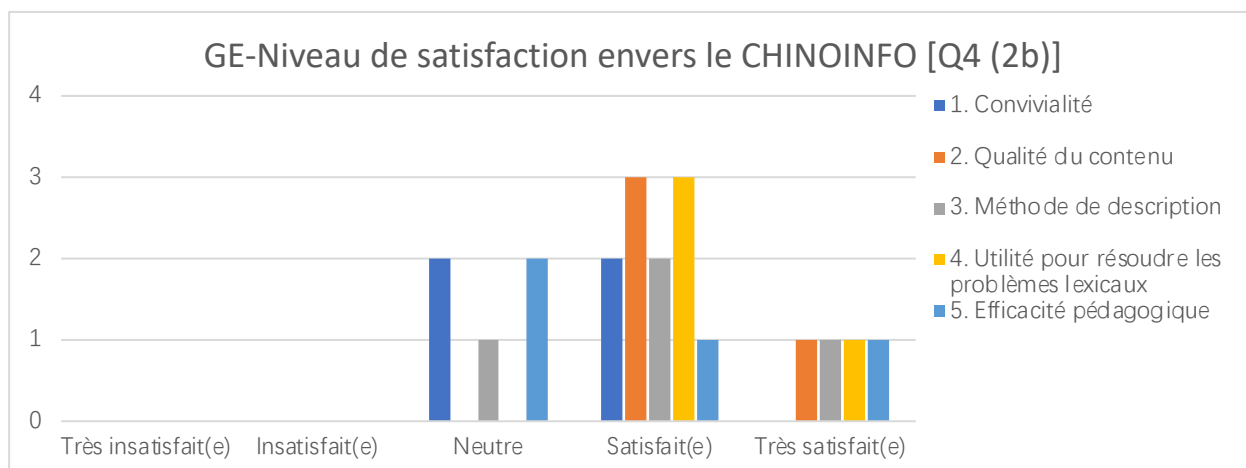


Figure 60. – Niveau de satisfaction du GE envers les outils de référence utilisés (Q4 du questionnaire 2b)

6.3.1.2. Questions à choix de réponses

Le traitement des questions à choix de réponses suit la même procédure appliquée à l'analyse du questionnaire 1. Nous avons d'abord effectué le codage des réponses. Dans le cas où une question contient une option de réponse ouverte, chaque réponse est analysée avant d'être codifiée. Les données ont été ensuite analysées et décrites selon le type de renseignements recueillis (tableau 38).

| | Type de renseignements recueillis | Questions (Questionnaire) |
|---|---|-------------------------------|
| 1 | Motifs de recours aux ressources lexicales | Q4 (2a) Q7 (2b) Q9 (2b) |
| 2 | Intérêt suscité par la méthode de description lexicale proposée | Q6 (2b) |

Tableau 38. – Catégorisation des données collectées au moyen des questions à choix

La Q4 du questionnaire 2a vise à identifier les raisons pour lesquelles le GC a fait appel aux ressources lexicales. La question et les options de réponses sont formulées comme suit :

Dans quelle(s) situation(s) avez-vous utilisé une ressource lexicale ?

1. Pour trouver le(s) sens d'un mot inconnu
2. Pour choisir un cooccurrent approprié d'un mot
3. Pour trouver des exemples d'usage d'un mot
4. Pour d'autres raisons (veuillez préciser)

Si un répondant choisissait « Pour d'autres raisons », il devait préciser le motif de son recours à des ressources lexicales au cours du test. Nous avons examiné toutes les réponses afin de regrouper certaines réponses similaires dans la même catégorie. Par exemple, la réponse « pour trouver le sens de la phrase au complet » a été classée dans la première catégorie, soit « pour trouver le(s) sens d'un mot inconnu ». Selon les résultats d'analyse, tous les participants du GC ont choisi les trois premières options de réponse, à savoir « trouver le(s) sens d'un mot inconnu », « choisir un cooccurrent approprié » et « trouver des exemples d'usage d'un mot » (figure 61).

Les Q7 et Q9 visent à connaître si le GE a eu recours à des ressources lexicales autres que le CHINOINFO et, dans le cas de réponse affirmative, les motifs de leur utilisation de ces ressources. L'analyse des réponses ouvertes du GE met en évidence le recours massif à la traduction automatique. Bien qu'ils aient eu accès au CHINOINFO lors du test, tous les participants du GE ont utilisé Google Translate ou DeepL pour trouver un sens d'un mot inconnu

ou pour trouver des exemples d'usage d'un mot. Trois des quatre participants ont signalé que les outils de traduction automatique étaient plus rapides pour avoir une idée du sens d'une phrase dans son ensemble. Il s'agit d'un point de départ de leurs recherches. Ces participants ont par la suite utilisé CHINOINFO pour effectuer des recherches approfondies afin de vérifier leur réponse à certaines questions. Un participant a indiqué qu'il a utilisé Google Translate pour connaître la transcription phonétique d'un mot inconnu.

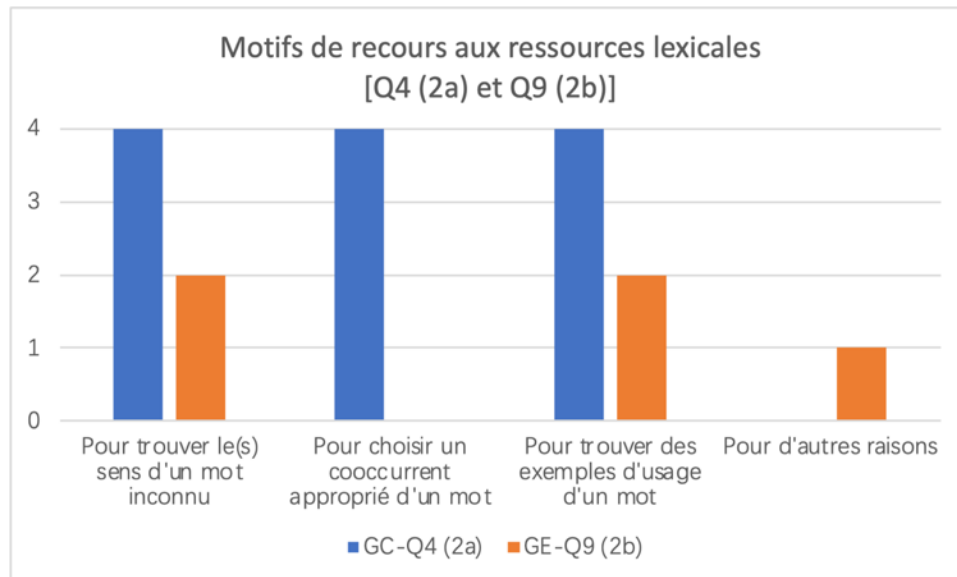


Figure 61. – Motifs de recours aux ressources lexicales autres que le CHINOINFO

Quant aux impressions du GE sur la méthode description lexicale proposée (Q6 du questionnaire 2b), tous les participants ont confirmé leur intérêt à utiliser une ressource didactique intégrant la méthode de description utilisée dans le CHINOINFO.

6.3.1.3. Questions à réponse courte

Les questions à réponse courte figurant dans les questionnaires 2a et 2b ont été regroupées en deux catégories : les questions factuelles et attitudinales (tableau 39). Les questions factuelles portent sur les ressources lexicales utilisées. Nous avons traité les réponses à l'aide de la même méthode de codification qui a été appliquée à l'analyse du questionnaire 1. La deuxième catégorie des questions porte sur les opinions des répondants sur leur utilisation des ressources lexicales. Les réponses à ces questions ont été codifiées en deux grandes catégories, soit les commentaires positifs et négatifs. Nous avons procédé par la suite à la synthèse des réponses catégorisées.

| | Types de questions | Questions (Questionnaire) | |
|---|-------------------------|---------------------------|--|
| 1 | Questions factuelles | Q3 (2a) | Quelle(s) ressource(s) lexicale(s) avez-vous utilisée(s) au cours du test ? |
| | | Q8 (2b) | Indiquez les outils que vous avez utilisés au cours du test. |
| 2 | Questions attitudinales | Q5 (2b) | Qu'est-ce que vous aimez le plus et le moins dans le CHINOINFO ? Quelles sont vos propositions pour l'améliorer ? |
| | | Q11 (2b) | Autres commentaires sur le recours à ces outils de référence. |
| | | Q8 (2a) | Autres commentaires sur le recours à ces outils de référence. |

Tableau 39. – Catégorisation des données collectées au moyen des questions à réponse courte

Ressources lexicales utilisées

Les participants du GC ont eu recours à trois types d'outils de référence : les dictionnaires bilingues, la banque de terminologie et les outils de traduction automatique.

Parmi les outils de référence que nous avons proposés (voir aussi l'annexe 18), deux dictionnaires bilingues ont été utilisés par tous les participants, soit le dictionnaire chinois-anglais MDBG (*MDBG Chinese-English dictionary*, 2022), ainsi que le dictionnaire chinois-français Chine-Nouvelle (VerticalScope Inc., 2022). Certains participants ont fait appel à leur outil de référence habituel, dont le dictionnaire chinois-français HantrainerPro (Stahlmann, 2022) qui est intégré au logiciel d'apprentissage de vocabulaire chinois Han Trainer Pro, ainsi que Pleco Chinese Dictionary (Pleco Inc., 2022). Ce dernier est une application mobile qui intègre un ensemble d'outils d'apprentissage du chinois, dont un dictionnaire chinois-anglais et un dictionnaire chinois-français intitulé *Grand dictionnaire Ricci de la langue chinoise* (Institut Ricci, 2001). Un participant du GC a consulté la banque de terminologie chinois-anglais Termonline (China National Committee for Terminology in Science and Technology, 2022), mais « sans grands résultats ».

Outre les dictionnaires, tous les participants du GC ont opté pour les outils de traduction automatique, dont Google Translate et DeepL, afin de comprendre le sens une phrase contenant plusieurs mots qui leur étaient inconnus.

Quant au GE, tous les participants ont utilisé le CHINOINFO comme outil de référence principal au cours du test, tel que nous l'avons demandé (voir l'annexe 19). Deux participants ont consulté le MDBG (*MDBG Chinese-English dictionary*, 2022) pour trouver les mots qui n'étaient pas

répertoriés dans le CHINOINFO. Un participant a utilisé l'application Pleco Chinese Dictionary (Pleco Inc., 2022) installée sur son téléphone mobile.

Tous les participants du GE ont utilisé les outils de traduction automatique, y compris Google Translate et DeepL, comme ce que nous avons constaté chez le GC.

En résumé, les dictionnaires bilingues étaient les outils de référence les plus utilisés dans le cadre de notre expérimentation. Aucun participant n'a utilisé le matériel didactique, probablement en raison de contraintes de temps, comme l'ont signalé deux participants dans leurs commentaires.

Commentaires sur le CHINOINFO

En premier lieu, le GE était satisfait de la qualité de contenu. Les opinions sont unanimes en ce qui concerne la richesse des connaissances lexicales présentées. Deux d'entre eux estimaient que les sections « Termes reliés » et « Combinaisons lexicales » étaient particulièrement utiles pour obtenir des connaissances approfondies sur l'emploi d'un terme dans les textes spécialisés. Par rapport aux autres dictionnaires bilingues, le CHINOINFO est un outil de référence pour « faire un travail de fond », selon le dire d'un répondant. En outre, la structure actancielle des termes paraissait fort intéressante selon un participant. Celui-ci a indiqué que la compréhension de la structure actancielle l'a beaucoup aidé à distinguer les acceptions d'un même vocable. Ce participant a également souligné l'originalité du CHINOINFO, du fait qu'aucune ressource d'apprentissage du lexique chinois, à sa connaissance, ne présente la structure actancielle.

En deuxième lieu, trois participants du GE ont formulé des commentaires positifs sur la structuration des connaissances dans le CHINOINFO. Deux d'entre eux ont souligné la clarté de la structure interne des articles. Ils ont notamment aimé la classification hiérarchique des liens lexicaux dans les sections « Termes reliés » et « Combinaisons lexicales ». L'autre répondant s'est plutôt intéressé au regroupement des termes en vocable. Il estimait que la « numérotation de différents sens d'un mot polysémique » était fort utile pour les apprenants débutants « qui avaient souvent de la misère à distinguer les mots polysémiques à l'aide des dictionnaires bilingues, ou même de Google Translate ».

Tous les participants se sont exprimés favorablement quant à l'apport du CHINOINFO à la résolution des problèmes lexicaux dans le cadre de notre expérimentation. Un répondant était d'avis que « le test était parfaitement approprié pour l'usage du CHINOINFO et permettait bien

de comprendre le fonctionnement de l’outil ». D’autres estimaient que cette ressource lexicale était d’une grande utilité pour les apprenants du lexique spécialisé. Cela dit, la plupart des répondants n’a jamais reçu de formation en COS et se sont focalisés sur l’apprentissage de la langue générale. Un répondant a indiqué qu’il aurait aimé avoir accès à un dictionnaire « de ce genre », c’est-à-dire un dictionnaire élaboré à l’aide de la même méthode de description, qui aborde le lexique en langue générale.

Quant aux commentaires négatifs, le GE a soulevé plusieurs points à améliorer pour le CHINOINFO en ce qui concerne la convivialité. Le premier bémol, mentionné par deux participants, est l’affichage de l’interface de recherche du CHINOINFO. Dans la version en ligne qui a été rendue disponible au GE, la description du dictionnaire affichée en haut à droite de la page d’accueil du CHINOINFO pouvait être superposée au module de recherche par caractère chinois, et ce, selon le fureteur que les participants utilisaient. Pour régler ce problème d’affichage, il faudrait ajuster la taille de la page Web en HTML à l’aide de la fonctionnalité Zoom, dans le cas où on utilise Mozilla Firefox, Google Chrome et Safari (voir la figure 62).

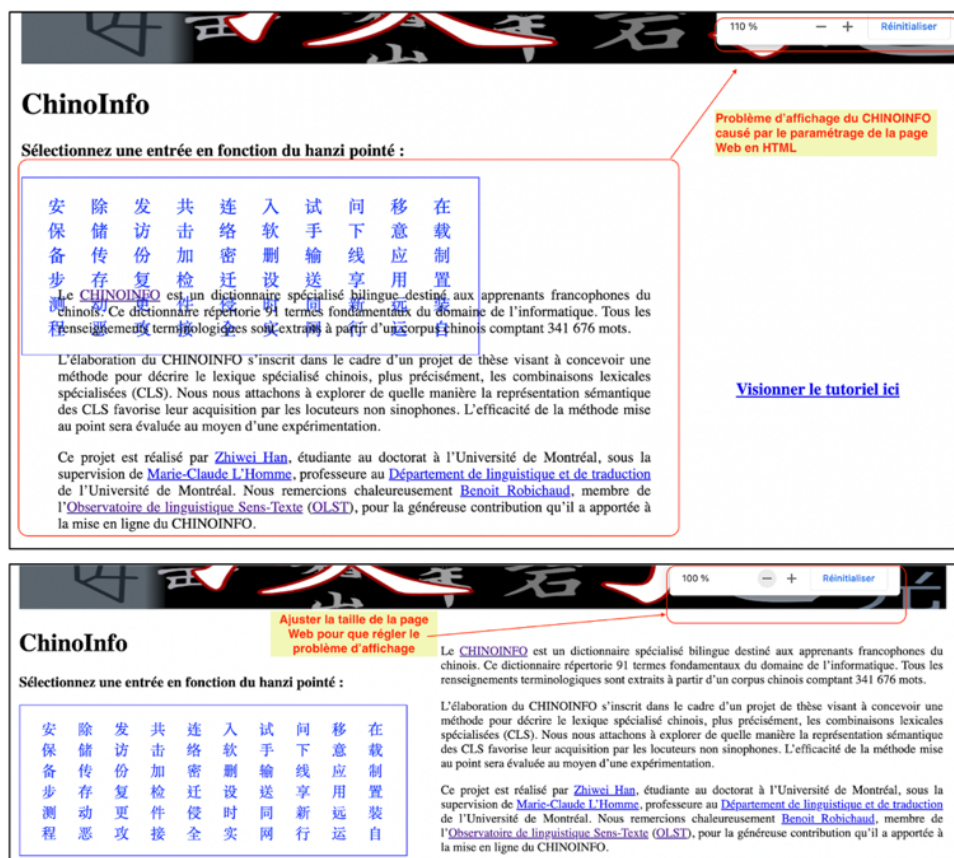


Figure 62. – Problème d’affichage de l’interface du CHINOINFO et solution proposée

Le deuxième défaut que les participants ont reproché au CHINOINFO est l'efficacité de la consultation. Un participant a indiqué que l'efficacité était un critère essentiel pour apprécier la qualité d'un dictionnaire. Malgré les efforts que nous avons déployés pour assurer la facilité d'accès aux connaissances lexicales (voir la sous-section 5.3.1), ce participant a affirmé devoir « zieuter tous les mots avant de trouver le mot qu'on cherche » lorsqu'il effectuait une recherche dans la nomenclature des termes. Il aurait souhaité avoir accès à un moteur de recherche pour pouvoir saisir le mot cherché. Nous pourrions intégrer cette fonctionnalité à l'interface du CHINOINFO lorsque la nomenclature du dictionnaire s'enrichira.

Plusieurs participants ont formulé des propositions pour améliorer notre ressource lexicale. Nous avons reçu deux commentaires en ce qui concerne la facilité d'accès à l'information. Outre la création d'un moteur de recherche, un participant a suggéré l'intégration de la traduction automatique au CHINOINFO au moyen de l'effet de survol. De cette manière, les utilisateurs auront accès à une infobulle affichant la traduction automatique d'un mot ou d'une phrase au passage du curseur de la souris. Cette fonction lui semblait particulièrement utile et pour les apprenants débutants, notamment pour les aider à comprendre rapidement une phrase contenant plusieurs mots inconnus dans la section « Contextes » de chaque article. Quant au contenu du dictionnaire, deux autres participants nous ont proposé d'élargir la nomenclature du CHINOINFO. Un participant aurait aimé avoir accès à la transcription phonétique (le pinyin) des termes répertoriés. Selon lui, il était très important de connaître le pinyin d'un terme même avant de comprendre son sens et son emploi dans les textes.

Enfin, les opinions divergent quant à l'intérêt pédagogique du CHINOINFO. Trois des quatre répondants ont donné des avis très favorables à ce sujet. Cependant, le CHINOINFO semblait davantage utile pour un apprentissage approfondi du lexique spécialisé, selon l'autre répondant. Sans nier le potentiel didactique du CHINOINFO, ce répondant l'estimait mieux adapté aux besoins de traducteurs et de linguistes. Il convient de mentionner qu'un participant du GE est effectivement un interprète professionnel anglais-français qui travaille parfois dans la combinaison chinois-français. Selon ce participant, le CHINOINFO était un excellent outil de travail, en comparaison avec d'autres outils qui ne permettaient qu'un décodage rapide, comme l'application Pleco et Google Translate. Tout bien considéré, nous pouvons conclure que le taux de satisfaction envers le CHINOINFO était très élevé.

Commentaires sur d'autres outils de référence

Les participants du GC et du GE se sont exprimés à propos d'autres ressources lexicales auxquelles ils ont eu recours.

Dans le GC, seul un répondant a commenté son utilisation des outils de référence que nous avons proposés avant le test. Il a eu l'impression « d'avoir réussi à (se) débrouiller à tâtons, sans être certain d'aucune des réponses ». Ce répondant a signalé plusieurs lacunes que nous avons nous-même constatées dans les dictionnaires bilingues chinois-anglais ou chinois-français. Premièrement, le répondant n'a trouvé aucune explication sur les équivalents proposés dans les dictionnaires consultés. Les utilisateurs ne savent pas dans quelle mesure les termes anglais ou français sont équivalents au terme chinois décrit. Il en va de même pour les logiciels de traduction automatique. Deuxièmement, on trouve peu d'exemples d'usage dans les dictionnaires consultés. Le répondant aurait souhaité avoir accès à une « ressource pédagogique complète avec cooccurrents et explication des nuances à partir de la deuxième langue (autre que le chinois) », qui « serait difficile à créer, mais extrêmement utile ».

Quant au GE, un participant a souligné que la rapidité à trouver les éléments de réponse dans un dictionnaire était très importante pour les débutants. Par conséquent, il préférait utiliser les outils de traduction automatique, tels que Google Translate et DeepL, pour comprendre le sens d'une phrase, plutôt que d'aller chercher chaque mot inconnu dans les dictionnaires. Il a toutefois reconnu que par rapport à la traduction automatique, il était plus bénéfique d'utiliser un dictionnaire bien conçu pour étudier un nouveau mot de façon approfondie.

6.3.2 Synthèse

Les résultats de l'analyse des questionnaires d'appréciation (2a et 2b) révèlent que les participants étaient plutôt satisfaits de l'organisation du test lexical. Les taux de satisfaction observés chez le GC et le GE sont similaires, bien que la documentation et les tâches préparatoires qui leur ont été proposées ne soient pas les mêmes. Quant au niveau de difficulté du test, le GE a eu l'impression d'avoir moins de difficultés à réaliser le test par rapport au GC. La section II portant sur la distinction sémantique des CLS de forme identique était la plus grande difficulté pour le GC, tandis que la section III portant sur le choix des cooccurrents appropriés était la plus difficile pour le GE.

Dans le cadre de notre expérimentation, le GE était mieux outillé que le GC pour résoudre les différents types de problèmes lexicaux soulevés par les CLS. D'un point de vue général, le GE était plus satisfait de son recours aux ressources lexicales au cours du test. Le GC, par contre, était plutôt insatisfait de l'utilité et de l'efficacité pédagogique des outils de référence utilisés.

En ce qui concerne l'appréciation du CHINOINFO, le GE a donné des commentaires très positifs sur la qualité de contenu et la structuration des connaissances. Dans l'ensemble, les participants étaient satisfaits de l'apport pédagogique du CHINOINFO à la résolution des problèmes lexicaux. Ils ont démontré un fort intérêt à utiliser d'autres ressources d'apprentissage faisant appel à la même méthode descriptive, ce qui laisse entrevoir le potentiel pédagogique de la méthode de description lexicale que nous avons proposée.

Le CHINOINFO est une ressource en construction, dont la convivialité reste à améliorer. Les participants ont formulé des commentaires et des suggestions en ce qui concerne l'affichage de l'interface de recherche et la facilité d'accès aux termes décrits. Dans le futur, nous envisageons d'intégrer d'autres fonctionnalités à la version en ligne du CHINOINFO, telles que le moteur de recherche, l'effet de survol permettant d'afficher la traduction automatique d'un mot ou d'une phrase, selon les propositions des participants. En outre, nous allons continuer d'enrichir la nomenclature du dictionnaire.

Par ailleurs, un constat plutôt inattendu qui ressort de notre analyse est le recours massif aux logiciels de traduction automatique par les participants. Bien qu'ils ramènent rapidement des réponses à certaines questions lexicales, l'utilisation de ces outils n'est pas toujours bénéfique pour les apprenants débutants. Nous allons y revenir dans la sous-section 6.4.

6.4 Interprétation des résultats et discussion

Dans ce qui suit, nous évaluons, à la lumière des résultats du traitement de données, l'efficacité du CHINOINFO en tant que ressource d'apprentissage du lexique spécialisé (6.4.1). Nous discutons ensuite d'autres facteurs susceptibles d'influencer les résultats expérimentaux (6.4.2) et des limites de notre expérimentation (6.4.3).

6.4.1 Efficacité pédagogique du CHINOINFO

En général, l'analyse comparative des résultats du test lexical (voir aussi la sous-section 6.2.2) montre que le GE a mieux réussi à résoudre les trois types de difficultés soulevées par les CLS. Les moyennes du GE sont plus élevées par rapport au GC, tant pour l'ensemble du test que pour chaque section, malgré le fait que les participants du GE ont un niveau d'apprentissage moins élevé que le GC (voir aussi la sous-section 6.1.2). Ces données nous amènent à penser que le recours au CHINOINFO a contribué à la réussite du GE.

Partant de ce premier constat, nous avons analysé de plus près la dispersion des notes au sein de chaque groupe. Du point de vue de la performance individuelle des participants, les notes du GE sont moins dispersées par rapport au GC, ce qui révèle que le taux de réussite des participants du GE est plus homogène. Du point de vue des types de compétences évaluées, les écarts entre les moyennes des notes pour chaque section sont plus faibles dans le GE que dans le GC. À la lumière de ces résultats, nous estimons que l'utilisation du CHINOINFO a aidé les participants du GE à atteindre un niveau de compétences adéquat pour résoudre les problèmes lexicaux liés aux propriétés syntaxico-sémantiques des CLS, et ce, peu importe leur expérience d'apprentissage du chinois.

Il ressort de notre étude que la deuxième section du test était la plus grande difficulté pour le GC. Cette section vise à évaluer l'habileté des participants à distinguer les sens et les structures syntaxiques des CLS de forme identique. Selon l'analyse des réponses à la question libre (Q10) du test lexical (sous-section 6.2.4), les deux premières questions de la section II étaient les plus difficiles pour la majorité des participants du GC. De plus, l'analyse des justifications du GC pour les Q4 à Q6 (section II) du test montre que les apprenants ont de la difficulté à identifier la structure syntaxique d'une CLS à l'aide du contexte d'utilisation de celle-ci, en raison de leur vocabulaire limité et de leur faible niveau de compétence linguistique en chinois. Par conséquent, les apprenants ne sont pas en mesure de déduire le sens de la CLS et d'identifier le sens d'une forme lexicale polysémique figurant dans la CLS, ce qui valide en partie notre hypothèse initiale sur les difficultés d'ordre syntaxiques des CLS.

Nous avons observé, dans cette section, le plus grand écart entre les résultats du GC et du GE : la moyenne du GE (30,6 %) est environ quatre fois plus élevée que celle du GC (7,6 %), comme nous l'avons évoqué dans la sous-section 6.2.2.1 (figure 50). Cet écart s'explique en partie par le

fait que le GC n'a pas réussi à trouver des éléments de réponses à leurs questions dans les outils de référence utilisés, comme l'ont signalé plusieurs participants dans leur réponse à la question libre (Q10) du test. Pour le GE qui a utilisé le CHINOINFO comme outil de référence principal, la section II est la section la mieux réussie (figure 50) et l'une des sections les plus faciles à réaliser (figure 56), ce qui nous laisse supposer que le recours au CHINOINFO a aidé le GE à résoudre ce type de difficulté de façon efficace.

Par ailleurs, les opinions des participants sur l'utilité des ressources lexicales utilisées, qui ont été recueillies au moyen des questionnaires de sondage (Q5 du questionnaire 2a pour le GC, Q3 et Q10 du questionnaire 2b pour le GE) consolident le constat dressé à partir des résultats du test⁹⁰. Tout compte fait, nous sommes d'avis que le CHINOINFO a contribué à bonifier la méthode descriptive des CLS, notamment en ce qui concerne la description des propriétés syntaxiques des CLS dans les ressources lexicales.

6.4.2 Autres facteurs sur les résultats expérimentaux

Le traitement des données collectées auprès des participants nous a permis d'identifier plusieurs facteurs qui sont susceptibles d'influencer les résultats de notre expérimentation, dont le recours à la traduction automatique, la recherche d'information sur le Web et le profil d'apprentissage des participants.

D'abord, les logiciels de traduction automatique, tels que Google Translate et DeepL, sont considérés comme des outils essentiels, voire indispensables par tous les participants. La majorité des participants ont fait appel à ces outils pour trouver le sens d'un mot, décoder une phrase complexe ou trouver l'équivalent d'une combinaison lexicale, comme nous l'avons indiqué dans la sous-section 6.3.1.3. Le recours à ces outils risque d'avoir un effet à double tranchant, étant donné que les apprenants débutants ne sont pas en mesure de vérifier l'exactitude des résultats proposés. Par exemple, dans la Q8 du test, on demande aux participants de choisir, parmi les quatre verbes proposés, deux collocations appropriés pour exprimer le sens de 'déplacer un fichier vers un nouvel emplacement', à savoir 上传 ^{shàngchuán} « téléverser (dans) » et 保存 ^{bǎo cún} « enregistrer (dans) ».

⁹⁰ Le traitement des réponses à ces questions est abordé dans la sous-section 6.3.1.1.

Nous avons constaté qu'un participant a choisi le verbe ^{yí dòng} 移动 (« déplacer ») puisque celui-ci a été proposé par DeepL. En fait, pour traduire la CLS *déplacer un fichier* du français en chinois, DeepL propose trois différents verbes de sens proches qui se combinent avec ^{wén jiàn} 文件, soit l'équivalent chinois du terme *fichier* : ^{yí dòng} 移动 (« déplacer »), ^{bān dòng} 搬动 (« acheminer ») et ^{bān yí} 搬移 (« transmettre »). Aucun de ces regroupements lexicaux n'est acceptable en chinois. En outre, trois participants du GC ont indiqué que les outils de traduction automatique ne distinguent pas les sens des CLS de forme identique (Q10 du test lexical).

En deuxième lieu, plusieurs participants ont affirmé avoir effectué des recherches sur le Web pour répondre aux questions du test lexical, car aucun élément de réponse n'a été trouvé dans les outils de référence qu'ils ont consultés. Bien qu'une requête dans Google permette d'obtenir rapidement des résultats, elle risque d'être une source d'erreur étant donné que l'information trouvée sur le Web est parfois non pertinente, voire trompeuse. Par exemple, certains participants ont vérifié leur choix de collocatif avec le moteur de recherche Google. Un participant du GE a justifié sa réponse à la Q9 du test, même si cela n'a pas été demandé. Il a tenté de valider si ^{chuán shū} 传输 (« transférer ») peut se combiner avec ^{yìng pán} 硬盘 (« disque dur »). Il a trouvé la séquence lexicale ^{gāo sù} 高速 ^{chuán shū yìng pán} 传输硬盘 ('haute vitesse' - 'transfert' - 'disque dur' « disque dur haute vitesse ») qui a une occurrence élevée sur Google, ce qui l'a porté à croire que *^{chuán shū yìng pán} 传输硬盘 'transférer'- 'disque dur' est une CLS attestée. Par contre, dans la séquence lexicale, ^{chuán shū} 传输 est un nom (« transfert ») qui est modifié par ^{gāo sù} 高速 (« haute-vitesse »). Il ne s'agit pas d'un collocatif verbal du terme ^{yìng pán} 硬盘 (« disque dur »).

Pour conclure, la recherche d'information sur le Web, tout comme le recours à la traduction automatique, a pu avoir des effets négatifs sur la performance des participants dans le test lexical si ceux-ci se fient entièrement aux résultats de recherche qui contiennent souvent des erreurs.

Enfin, le profil des participants peut influencer leur appréciation des ressources lexicales utilisées. Quant au niveau de satisfaction du GE envers le CHINOINFO, nous avons constaté un écart entre le résultat attendu et le résultat réel. Comme la plupart des participants ont formulé des commentaires très positifs sur leur utilisation du CHINOINFO (Q5 du questionnaire 2b)⁹¹, nous

⁹¹ Nous reportons le lecteur à la sous-section 6.3.1.3 pour la synthèse des réponses à la Q5 du questionnaire 2b.

nous attendions à un taux de satisfaction très élevé sur l'efficacité pédagogique du dictionnaire. Par contre, une moitié des participants du GE a donné une opinion neutre à ce sujet plutôt que d'être « satisfait(e) » ou « très satisfait(e) » (figure 60). Ce résultat, qui était moins satisfaisant que prévu, peut être expliqué par le fait que la plupart des participants n'ont jamais reçu de formation en COS. Pour cette raison, un dictionnaire spécialisé qui ne traite que du lexique en informatique peut sembler moins adapté à leurs besoins d'apprentissage. Cela dit, les opinions des participants sont unanimes quant à leur intérêt à utiliser une ressource didactique intégrant la méthode de description proposée dans le CHINOINFO (Q6 du questionnaire 2b), ce qui tend à montrer l'intérêt pédagogique de notre méthode de description lexicale.

6.4.3 Limites de l'expérimentation

Nous soulignons ici deux limites de notre expérimentation : le nombre limité des participants et le manque d'expérience d'apprentissage en COS des participants.

D'abord, le nombre de participants recrutés est restreint en raison de deux contraintes majeures. Premièrement, les critères que nous avons définis pour recruter les participants dans les trois universités québécoises étaient très stricts (voir la sous-section 4.3.1.2). Les participants recherchés doivent avoir suivi au moins deux cours de chinois élémentaires, ce qui exclut une proportion importante de la population étudiante des universités sollicitées. Deuxièmement, nous ne disposons pas de moyens suffisants pour lancer un recrutement d'envergure auprès des étudiants francophones qui font leurs études en Chine. Peu d'étudiants ont répondu à l'annonce de recrutement publiée sur le forum de discussion. Un seul étudiant nous a fait part de son intérêt à participer au projet via la plateforme WeChat. Si nous avions eu les moyens, nous aurions collaboré avec des universités chinoises pour pouvoir lancer une invitation à tous les étudiants francophones inscrits aux cours de chinois.

Ensuite, nous avons remarqué un manque d'expérience d'apprentissage en langue de spécialité. En fait, la majorité des participants est formée en chinois langue générale. Parmi les huit participants, seul un étudiant a suivi un cours de chinois du commerce offert par la BLCU. Un participant du GE a affirmé avoir étudié la terminologie en médecine traditionnelle chinoise, sans pour autant recevoir une formation systématique en COS. Faute de moyens nécessaires, nous n'avons pas réussi à joindre les étudiants qui suivent les cours du COS en Chine. Par conséquent,

le CHINOINFO pourrait sembler moins adapté aux besoins de ces participants, ce qui a un effet négatif sur l'évaluation de l'efficacité pédagogique du CHINOINFO. Il serait bénéfique de réaliser une deuxième phase de notre expérimentation qui s'adresse aux apprenants du COS, et dans la mesure du possible, aux apprenants qui font leur étude dans le domaine de l'informatique.

Chapitre 7 Conclusion

Dans ce dernier chapitre, nous faisons d'abord une synthèse de notre thèse (7.1). Nous rappelons les principaux résultats obtenus en lien avec nos objectifs de recherche et nos hypothèses, afin de répondre à notre problématique. Nous présentons ensuite les apports de notre recherche à l'enseignement de la combinatoire spécialisée chinoise et aux recherches dans ce domaine (7.2). À la fin du chapitre, nous proposons des pistes de recherche qui se dégagent de notre étude (7.3).

7.1 Synthèse des résultats de la thèse

Notre thèse se situe dans le contexte de l'enseignement-apprentissage du lexique spécialisé chinois. Pour les apprenants non natifs, l'acquisition de la combinatoire spécialisée représente un défi majeur compte tenu des difficultés syntaxico-sémantiques que celle-ci soulève. Dans les ressources d'apprentissage en COS, peu de solutions sont proposées pour rendre compte des propriétés combinatoires du lexique spécialisé.

Pour pallier les lacunes observées, nous nous attachons à explorer de quelle manière une méthode de description lexicale basée sur une représentation sémantique et syntaxique permet de trouver des solutions aux difficultés qui font obstacle à l'apprentissage des CLS. Nous avons formulé trois objectifs de recherche en fonction de cette problématique : 1) concevoir une méthode de description des CLS en vue de la résolution de difficultés lexicales par les locuteurs non sinophones; 2) appliquer la méthode mise au point à l'élaboration d'une ressource didactique du lexique spécialisé et 3) évaluer l'efficacité pédagogique de la ressource élaborée.

Le chapitre 2 dresse un état de l'art de la didactique du lexique spécialisé. Nous avons d'abord recensé les travaux en rapport avec notre problématique et notre objet d'étude, soit l'enseignement du lexique basé sur une approche sémantique et l'étude de la combinatoire lexicale à des fins didactiques. Nous avons ensuite passé en revue les travaux axés sur l'enseignement du lexique spécialisé chinois auprès des locuteurs non natifs. La revue de littérature nous a permis de constater que l'enseignement de la combinatoire spécialisée a occupé jusqu'ici une place périphérique en COS, ce qui met en évidence l'originalité de notre thèse par rapport aux travaux antérieurs s'inscrivant dans le même contexte de recherche.

Le chapitre 3 présente notre cadre théorique qui se situe à la conjonction de la lexicologie, de la terminologie et de la didactique des langues secondes. Premièrement, nous avons abordé la Lexicologie explicative et combinatoire (LEC) qui nous sert d'assise théorique pour définir les principes de base de la description du lexique chinois. Deuxièmement, nous avons examiné l'approche lexico-sémantique de la terminologie qui fournit des repères théoriques et méthodologiques pour les tâches liées à la description du lexique spécialisé. Enfin, nous avons présenté l'approche cognitive en didactique des langues secondes qui explore les facteurs favorisant l'acquisition des connaissances linguistiques. Nous nous appuyons sur cette approche pour envisager une présentation à la fois structurée et intuitive des CLS.

Le chapitre 4 décrit notre démarche méthodologique qui comprend trois phases. Nous avons d'abord assemblé un corpus spécialisé portant sur le domaine de l'informatique, pour en extraire un échantillon de CLS ainsi que leurs renseignements descriptifs. Le traitement des données collectées à partir du corpus nous a permis d'anticiper trois types de difficultés syntaxico-sémantiques soulevées par les CLS. Nous avons ensuite conçu une méthode de description qui propose des solutions aux difficultés anticipées. Cette méthode a ensuite été appliquée à l'élaboration du CHINOINFO, un dictionnaire spécialisé chinois-français en ligne. Enfin, nous avons mené une expérimentation comparative auprès d'apprenants francophones pour évaluer l'efficacité pédagogique du CHINOINFO. Nous avons fait passer un test lexical en ligne à deux groupes d'apprenants (le GC et le GE), afin d'évaluer leurs aptitudes à résoudre trois types de problèmes lexicaux soulevés par les CLS. Les deux groupes étaient autorisés à utiliser une variété de ressources lexicales au cours du test, y compris un nombre d'outils de référence que nous leur avons proposés. Seul le GE a eu accès à CHINOINFO. Nous avons aussi recueilli, au moyen de trois questionnaires de sondage, le profil des participants ainsi que leur appréciation sur le test et les outils de référence proposés.

Le chapitre 5 fournit une description du CHINOINFO. Nous avons présenté le contenu et les fonctionnalités de cette ressource. Le CHINOINFO répertorie 91 termes fondamentaux en informatique et technologies de l'information, pour lesquels nous avons encodé au total 282 termes reliés et 644 CLS. La structuration des données au sein des articles s'est largement inspirée de l'adaptation du DiCoInfo à un dictionnaire d'apprentissage (Alipour, 2014). Différents moyens techniques ont été mis en œuvre pour assurer la convivialité de la ressource.

Le chapitre 6 expose les résultats de notre expérimentation. Nous avons recruté huit participants dont la majorité provient de deux universités québécoises. L'analyse des réponses au questionnaire 1 nous a permis de connaître les parcours universitaire et professionnel des participants, leur expérience d'apprentissage du chinois et leur expérience vécue dans un milieu de vie sinophone. En outre, les participants ont indiqué que le recours à des dictionnaires bilingues ne leur permet pas toujours de résoudre les problèmes lexicaux, étant donné le manque d'exemples d'usage et le manque de description des combinaisons lexicales dans ces outils de référence.

De façon générale, l'analyse des résultats du test lexical montre que le GE a notablement mieux réussi les trois sections du test par rapport au GC, bien que le GE présente un niveau d'apprentissage moins élevé que le GC selon les renseignements recueillis à l'aide du questionnaire 1 (voir la sous-section 6.1.2).

L'écart le plus important entre les deux groupes s'observe dans la section II qui porte sur la distinction des CLS de forme identique. Le GC a moins bien réussi cette section et l'a jugée plus difficile comparativement aux deux autres sections. Les participants du groupe ont indiqué qu'ils n'ont pas trouvé d'éléments de réponse dans les outils de référence consultés. On peut en déduire que la résolution des difficultés engendrées par l'ambiguïté syntaxique des CLS représente un défi de taille pour le GC, ce qui valide en partie notre première hypothèse au sujet des difficultés faisant obstacle à l'apprentissage des CLS.

Pour le GE, la section II est la mieux réussie et l'une des sections les plus faciles à réaliser, ce qui nous laisse supposer que le recours au CHINOINFO a aidé le GE à distinguer les sens et les structures syntaxiques associés à des CLS de forme identique.

En revanche, l'écart le moins marqué entre les deux groupes s'observe dans la section III portant sur le choix des cooccurrents appropriés. Il s'agit de la section la mieux réussie par le GC et la moins réussie par le GE, comparativement à deux autres sections. La section III représente la plus grande difficulté pour le GE qui a dû effectuer plusieurs recherches pour valider leur choix de cooccurrents pour un terme, ce qui suggère que des améliorations sont à apporter au CHINOINFO pour qu'il fournisse une aide plus adéquate à la résolution de ce type de problèmes lexicaux.

L'analyse intragroupe montre que les notes des participants dans le GE sont moins dispersées par rapport au GC. Dans le GE, la disparité entre les moyennes pour chaque section est également moins marquée comparativement au GC. Nous en déduisons que l'utilisation du CHINOINFO a contribué à la réussite du GE pour l'ensemble du test lexical.

L'analyse des renseignements recueillis à l'aide des questionnaires 2a et 2b révèle que les participants étaient plutôt satisfaits de l'organisation du test lexical. En ce qui concerne le niveau de difficulté du test, le GE a eu moins de difficultés à réaliser le test par rapport au GC. Le GE se sentait mieux outillé pour trouver des éléments de réponses aux questions du test par rapport GC. Ce dernier, par contre, était plutôt insatisfait de l'utilité et de l'efficacité pédagogique des outils de référence utilisés.

Quant à l'appréciation du CHINOINFO, le GE a donné des commentaires très positifs sur la qualité de contenu et la structuration des connaissances. Il s'est exprimé favorablement quant à l'apport du CHINOINFO à la résolution des problèmes lexicaux dans le cadre de notre expérimentation. Toutefois, certains participants ont formulé des commentaires négatifs quant à la convivialité du CHINOINFO et l'efficacité de la consultation. Ils ont fourni des suggestions intéressantes pour améliorer la ressource.

Pour conclure, même si le nombre de participants recrutés était moins élevé que ce que nous aurions souhaité, les résultats sont prometteurs et indiquent l'apport du CHINOINFO à la résolution de certaines difficultés soulevées par les propriétés syntaxico-sémantiques des CLS. Ces résultats nous permettent de valider en partie nos trois hypothèses portant respectivement sur les difficultés d'apprentissage des CLS, les solutions aux difficultés identifiées et les facteurs favorisant l'acquisition des connaissances lexicales. En fin de compte, cette recherche apporte des éléments de réponse à notre problématique en ce qui concerne l'intérêt de la méthode de description lexicale proposée dans un contexte pédagogique.

Les résultats de notre recherche devront être confirmés par un plus grand nombre de participants. Par ailleurs, les données que nous avons recueillies au moyen des questionnaires de sondage (voir la sous-section 6.3.1.1), notamment les commentaires des participants sur l'organisation du test et le niveau de difficulté du test, permettront de bonifier une éventuelle deuxième phase de l'expérimentation.

7.2 Apports de la thèse

En premier lieu, notre recherche a débouché sur des résultats tangibles qui sont susceptibles de bonifier concrètement l'enseignement des CLS auprès des locuteurs non sinophones.

Nous avons élaboré le dictionnaire CHINOINFO, une ressource d'apprentissage ouverte aux apprenants du chinois et à toute la communauté scientifique (enseignants et chercheurs) œuvrant dans le domaine de la didactique du lexique spécialisé. Cette ressource lexicale vise à aider les apprenants non natifs à résoudre des problèmes lexicaux soulevés par les propriétés syntaxico-sémantiques des CLS extraites de textes spécialisés authentiques.

Nous avons confectionné, aux fins de cette étude, un corpus spécialisé chinois portant sur le domaine de l'informatique et les technologies de l'information. Ce corpus peut être mis à la disposition des enseignantes et des chercheurs qui souhaitent l'explorer de façon différente ou récupérer une partie des textes compilés pour construire un nouveau corpus.

L'analyse des résultats de l'expérimentation nous a permis de comprendre davantage les besoins des participants dans leur apprentissage des CLS. Les commentaires et suggestions des participants permettront d'améliorer le CHINOINFO et de bonifier une éventuelle deuxième phase de l'expérimentation.

En deuxième lieu, notre démarche méthodologique est susceptible d'apporter des contributions à l'avancement des recherches sur la combinatoire spécialisée chinoise. D'abord, notre méthode de collecte et d'analyse des CLS peut enrichir les travaux sur le traitement des CLS en terminologie. Ensuite, la méthode descriptive des CLS que nous avons conçue peut être appliquée à d'autres domaines de spécialité et mener à la création de ressources terminologiques. Quant à l'encodage des CLS, nous avons proposé un schéma d'encodage XML adapté à la description des CLS en chinois. Ce schéma est facilement exploitable pour créer d'autres ressources lexicales. Enfin, la méthode que nous avons utilisée pour associer des CLS chinoises à leur équivalent français répertorié dans une autre ressource terminologique peut inspirer les travaux sur l'équivalence de la combinatoire lexicale.

7.3 Perspectives de recherche

Notre recherche ouvre des perspectives qui pourraient être développées sur différents plans. En premier lieu, les résultats de notre première expérimentation pourraient être enrichis.

Il y aurait lieu de mener notre expérimentation auprès d'un nombre plus important d'apprenants du COS. Il serait opportun de joindre les étudiants qui reçoivent une formation du COS en Chine, et dans la mesure du possible, les apprenants qui ont une expérience d'études ou de travail dans le domaine de l'informatique.

Quant à la ressource élaborée, nous continuons d'enrichir la nomenclature du CHINOINFO et d'apporter des améliorations à la ressource selon les commentaires que nous avons reçus dans le cadre de cette étude. Dans le futur, il serait intéressant d'intégrer d'autres fonctionnalités à la version en ligne du CHINOINFO, telles que le moteur de recherche, l'effet de survol permettant d'afficher la traduction automatique d'un mot ou d'une phrase, afin d'améliorer la convivialité de la ressource. Il serait aussi bénéfique d'améliorer la présentation des CLS dans le CHINOINFO pour qu'il fournisse une aide plus efficace aux apprenants dans la résolution des problèmes liés au choix des cooccurrents.

En deuxième lieu, il serait opportun de poursuivre la recherche sur l'application pédagogique de nos résultats de recherche. Plus précisément, le corpus spécialisé compilé à des fins de cette étude, la méthode de description que nous avons proposée, de même que le dictionnaire CHINOINFO pourraient être mis à contribution pour : 1) créer du matériel pédagogique, comme une plateforme numérique d'exercices (Dechamps, 2013a); 2) élaborer des activités d'apprentissage en classe, telles que les tâches basées sur un corpus spécialisé (Alipour, 2018); et 3) concevoir des dispositifs pédagogiques pour des cours à distance.

Une autre piste de recherche s'ouvre sur la pédagogie renversée (Lebrun, 2007; Lebrun et al., 2014; Cailliez et Lebrun, 2019) qui se présente comme un complément de la pédagogie magistrale. En pédagogie renversée, les étudiants sont davantage responsabilisés dans l'acquisition des connaissances. Au lieu de suivre un cours, ils prennent le rôle d'enseignant. Ils définissent les connaissances dont ils ont besoin pour construire un cours et les assimilent pour pouvoir les présenter de façon structurée. De cette manière, les étudiants jouent un rôle actif dans leur apprentissage. Il serait intéressant d'explorer la possibilité d'appliquer nos résultats de

recherche à cette formule pédagogique innovante pour soutenir un apprentissage actif et autonome.

En troisième lieu, notre thèse ouvre des pistes de recherche sur l'utilisation de la traduction automatique par les apprenants du COS. Il ressort de notre étude que le recours à des logiciels de traduction automatique risque d'être à double tranchant comme les apprenants ne sont pas en mesure de vérifier l'exactitude des résultats proposés. En cela, il serait utile de mener des études comparatives pour examiner si ces outils peuvent être aussi efficaces qu'une ressource lexicale conçue pour soutenir un apprentissage systématique des CLS.

À l'heure actuelle, le débat bat son plein sur la pertinence d'intégrer la traduction automatique dans l'enseignement des langues secondes. Pour enrichir la littérature scientifique déjà très riche à ce sujet (Chang, 2022; Chon et al., 2021; Clifford et al., 2013; Josselin-Leray et al., 2019; Kol et al., 2018), nous souhaitons explorer comment former les apprenants non natifs à utiliser la traduction automatique de façon raisonnée afin de bonifier leur apprentissage du lexique spécialisé.

Les pistes de prolongement qui se dégagent de ce travail pourront être approfondies lors de recherches ultérieures.

Références bibliographiques

Académie française. (2019). Envoyer. Dans *Dictionnaire de l'Académie française*. <https://www.dictionnaire-academie.fr/article/A9E2066>

Alipour, M. (2014). *Méthodologie de conversion de dictionnaires spécialisés en dictionnaires d'apprentissage : application au domaine de l'informatique* [thèse de doctorat]. Université de Montréal.

Alipour, M. (2018). *Approche socioconstructiviste pour l'enseignement-apprentissage du lexique spécialisé : apport du corpus dans la conception d'activités lexicales* [thèse de doctorat]. Université de Montréal.

Alonso Ramos, M. (2000, 8 décembre). *Critères heuristiques pour l'encodage des collocations au moyen de fonctions lexicales*. C. R. Ulrich Heid Stefan Evert, Egbert Lehmann (dir.), Stuttgart, Germany (p. 463-473).

Alonso Ramos, M. (2005a). *Semantic description of collocations in a lexical database* (vol. 2005, p. 17-27).

Alonso Ramos, M. (2005b). *Semantic Description of Collocations in a Lexical Database*. F. Kiefer, G. Kiss et J. Pajzs (dir.), Computational Lexicography, COMPLEX 2005, Budapest (p. 17-27).

Alonso Ramos, M., Nishikawa, A. et Vincze, O. (2010). *DiCE in the web: An online Spanish collocation dictionary*. eLexicography in the 21st Century: New Challenges, New applications (p. 369-374).

Alonso Ramos, M., Rambow, O. et Wanner, L. (2008). *Using Semantically Annotated Corpora to Build Collocation Resources*. Sixth International Conference on Language Resources and Evaluation. http://www.lrec-conf.org/proceedings/lrec2008/pdf/294_paper.pdf

An, H. (2005). 从两种词表看名、动、形兼类的处理 [Traitement de mot multicatégoriel de type nom/verbe/adjectif dans deux listes de mots]. *语言教学与研究 [Language Teaching and Linguistic Studies]*, (4), 31-39.

An, R. et Shan, Y. (2005). *科技汉语：中级阅读教程 [Scientific Chinese: An Intermediate Reading Course]*. 北京大学出版社 [Peking University Press].

An, R. et Shan, Y. (2008). *科技汉语：高级阅读教程 [Scientific Chinese: An advanced Reading Course]*. 北京大学出版社 [Peking University Press].

Anctil, D. (2005). *Maîtrise du lexique chez les étudiants universitaires : typologie des problèmes lexicaux et analyse des stratégies de résolution de problèmes lexicaux* [mémoire de maîtrise]. Université de Montréal.

Anctil, D. (2010). *L'erreur lexicale au secondaire : analyse d'erreurs lexicales d'élèves de 3e secondaire et description du rapport à l'erreur lexicale d'enseignants de français* [thèse de doctorat]. Université de Montréal.

Anthony, E. M. (1963). Approach, Method, and Technique. *English Language Teaching (ELT) Journal*, XVII(2), 63-67.

Arcodia, G. F. et Basciano, B. (2021). *Chinese Linguistics: An Introduction*. Oxford University Press.

Arsenault, L. et Giguère, M. (2020, 18 juin). *Construire un examen de qualité administré à distance* [formation]. Formation offerte par le Service de soutien à l'enseignement de l'Université Laval.

ARTES : Aide à la Rédaction de Textes Scientifiques [CLILLAC-ARP]. (s. d.). <http://www.clillac-arp.univ-paris-diderot.fr/projets/artes>

Ausubel, D. P. (1963). Cognitive Structure and the Facilitation of Meaningful Verbal Learning1.

Ausubel, D. P. (1968). *Educational Psychology: A Cognitive View*. Holt.

Bai, X. hong. (2014). *科技汉语读写教程 [Chinese for Science and Technology: Reading and Writing]*. 北京语言大学出版社 [Beijing Language and Culture University Press].

Bellassen, J. (2010). La didactique du chinois et la malédiction de Babel. Émergence, dynamique et structuration d'une discipline. *Études chinoises*, 1(1), 27-44.

Bellassen, J. et Arslangul, A. (2014). *Bescherelle: le chinois pour tous*. Hatier.

Bellassen, J., Medhat-Lecocq, H. et Ouvrard, L. (2012). *Ecritures, politiques linguistiques et didactique des langues*. Archives contemporaines.

Bellassen, J. et Wong, W. (1995). *Les idéogrammes chinois ou l'empire du sens*. You-Feng.

Bellassen, J. et Zhang, P. (1990). *Méthode d'initiation à la langue et à l'écriture chinoises*. la Compagnie.

Bellassen, J. et Zhang, P. (1991). *Perfectionnement à la langue et à l'écriture chinoises*. la Compagnie.

Benson, M. (1989). The Structure of the Collocational Dictionary. *International Journal of Lexicography*, 2(1), 1-14.

Bergenholtz, H. et Tarp, S. (1995). *Manual of Specialised Lexicography*. btl.12. John Benjamins Publishing Company.

Bergenholtz, H. et Tarp, S. (2003). Two opposing theories. On HE Wiegand's recent discovery of lexicographic functions. *HERMES-Journal of Language and Communication in Business*, (31), 171-196.

Binon, J., Dancette, J. et Verlinde, S. (1998). *Comment améliorer le traitement des synonymes dans un dictionnaire de langue*. EURALEX'98 (Huitième Congrès International de Lexicographie), Liège (vol. 1, p. 77-86).

Binon, J., Selva, T. et Verlinde, S. (2004). *Tendances et innovations récentes en lexicographie pédagogique. La contribution des dictionnaires d'apprentissage DAFA et DAFLES* (p. 53-81).

Binon, J. et Verlinde, S. (2003). Les collocations : clef de voûte de l'enseignement et de l'apprentissage du vocabulaire d'une langue étrangère ou seconde. *La Lettre de l'AIRDF*, 33(2), 31-36.

Binon, J. et Verlinde, S. (2004). L'enseignement/apprentissage du vocabulaire et la lexicographie pédagogique du français sur objectifs spécifiques (FOS) : le domaine du français des affaires. *Ela. Études de linguistique appliquée*, 135(3), 271-283.

Binon, J., Verlinde, S. et Selva, T. (2005). Influences internationales sur la lexicographie pédagogique du FLE. *Trabalhos em Linguística Aplicada*, 44(2), 215-231.

- Binon, J., Verlinde, S., Van Dyck, J. et Bertels, A. (2000). *Dictionnaire d'apprentissage du français des affaires*. Didier.
- Bourigault, D. et Slodzian, M. (1999). Pour une terminologie textuelle. *Terminologies nouvelles*, 19, 29-32.
- Bowker, L. et Pearson, J. (2002). *Working with Specialized Language: A Practical Guide to Using Corpora* (1^{re} éd.). Routledge.
- Brown, J. D. (2001). *Using Surveys in Language Programs*. Cambridge University Press.
- Bruner, J. S. (1960). *The Process of Education* (p. xvi, 97 pages ;). Harvard University Press.
- Byram, M. (2002). *Routledge Encyclopedia of Language Teaching and Learning*. Routledge.
- Cabré, M. T. (2003). Theories of terminology: Their description, prescription and explanation. *Terminology*, 9(2), 163-199.
- Caignon, P. (2001). *Vocabulaire et cooccurrents de la comptabilité*. Brossard, Québec: Liguattech.
- Cailliez, J.-C. et Lebrun, M. (2019). *La classe renversée : l'innovation pédagogique par le changement de posture* (2e éd.) (2e éd.). Ellipses.
- Cajolet-Laganière, H., Martel, P. et Masson, C.-É. (2022). Amovible. Dans *Le dictionnaire USITO*. <https://usito.usherbrooke.ca/d%C3%A9finitions/amovible>
- Calvi, S. (2022). *Collocations et terminologie : extraction semi-automatique et classement des collocations dans le domaine du commerce international*. Università Degli Studi Di VERONA.
- Carlisle, J. F. (1995). Morphological Awareness and Early Reading Achievement. Dans L. Feldman (dir.), *Morphological Aspects of Language Processing* (p. 189-209). Erlbaum.
- Carroll, J. B. (1964). *Language and thought*. Prentice-Hall international.
- Carroll, J. B. (1966). The contribution of psychological theory and educational research to the teaching of foreign languages. Dans A. Valdman (dir.), *Trends in Language Teaching* (McGraw-Hill).
- Cauchon, J. (2003). *Enseignement de la structure prédicative du verbe au primaire : étude*

expérimentale à partir du verbe DÉCIDER [mémoire de maîtrise]. Université de Montréal.

Cavalla, C. (2008a). Les collocations dans les écrits universitaires : un lexique spécifique pour les apprenants étrangers. Dans O. B. et I. Schaffner (dir.), *Apprendre une langue de spécialité : enjeux culturels et linguistiques* (p. 93-104). Editions Ecole Polytechnique.

Cavalla, C. (2008b). Propositions didactiques pour l'enseignement d'éléments phraséologiques en FLE. Dans O. Galatanu, M. Pierrard, D. Van Raemdonck et M.-E. Damar (dir.), *Enseigner les structures langagières en FLE* (Peter Lang, p. 147-158).

Cavalla, C. (2013). *Le lexique des émotions : réflexions pour l'enseigner en FLE*. Louvain la Neuve, Belgium.

Cavalla, C. (2015a). Collocations transdisciplinaires dans les écrits de doctorants FLS/FLE. *Linx*, (72), 95-110.

Cavalla, C. (2015b). L'enseignement des affects en FLE : Essai autour des collocatifs d'intensité. *Le Langage et l'Homme*, 50(1), 117-130.

Cavalla, C. (2016). *Les apprentissages lexicaux : des unités linguistiques à l'enseignement du FLE*. [synthèse en vue de l'obtention de l'habilitation à diriger des recherches]. Université Grenoble Alpes.

Cavalla, C. (2018a). Exemple d'enseignement de la phraséologie transdisciplinaire à l'aide de corpus numériques en FLE. *La lettre de l'AIRDF*, (64), 43-47.

Cavalla, C. (2018b). Lexique scientifique transdisciplinaire et enseignement aux étudiants allophones. Dans M.-P. Jacques et A. Tutin (dir.), *Lexique transversal et formules discursives des sciences humaines* (p. 191-212). ISTE Group.

Cavalla, C., Loiseau, M., Lascombe, V. et Socha, J. (2014). Corpus, base de données, cartes mentales pour l'enseignement. Dans P. Blumenthal, I. Novakova et D. Siepmann (dir.), *Les émotions dans le discours- Emotions in Discourse* (p. 327-341). Peter Lang.

Cellier, M. (2008). *Guide pour enseigner le vocabulaire à l'école primaire*. Retz.

Cellier, M. (2011). *Le vocabulaire et son enseignement : des outils pour structurer l'apprentissage du vocabulaire*. Eduscol : portail national des professionnels de l'éducation.

- Chang, L.-C. (2022). Chinese language learners evaluating machine translation accuracy. *The JALT CALL Journal*, 18(1), 110-136. <https://doi.org/10.29140/jaltcall.v18n1.592>
- Chao, Y. R. (1968). *A Grammar of Spoken Chinese*. University of California Press.
- Chastain, K. (1970). Behavioristic and Cognitive Approaches in Programmed Instruction. *Language Learning*, 20(2), 223-235.
- Chastain, K. (1971). *The Development of Modern Language Skills: Theory to Practice*. The Center for Curriculum Development, Inc.
- Chastain, K. (1990). La théorie cognitive de l'apprentissage et son influence sur l'apprentissage et l'enseignement des langues secondes. *Études de Linguistique Appliquée*, 77, 21-27.
- Cheng, N. et Polguère, A. (1993). *The DiCo Approach to Computational Lexicography and Its Applicability to Chinese*. Fukuoka, Japon (p. 292-301).
- China National Committee for Terminology in Science and Technology. (2022). *Termonline*. Termonline: a public terminology knowledge service platform. <https://www.termonline.cn/index>
- Chinese Testing International Co., Ltd. (2018). *Chinese Proficiency Test (HSK)*. 汉语考试服务网[Chinesetest]. <http://www.chinesetest.cn/gosign.do?id=1&lid=0>
- Chomsky, N. (1957). *Syntactic Structures* (p. 1 vol. (118 p.) ;). Mouton.
- Chomsky, N. (1959). Review of Verbal behavior. *Language*, 35(1), 26-58.
- Chon, Y. V., Shin, D. et Kim, G. E. (2021). Comparing L2 learners' writing against parallel machine-translated texts: Raters' assessment, linguistic complexity and errors. *System*, 96, 102408. <https://doi.org/10.1016/j.system.2020.102408>
- Clifford, J., Merschel, L. et Munné, J. (2013). Surveying the landscape: What is the role of machine translation in language learning? *@ tic. revista d'innovació educativa*, (10), 108-121.
- Cohen, B. (1986). *Lexique de cooccurrents : bourse et conjoncture économique*. Linguattech.
- Condamines, A. (2003). *Sémantique et corpus spécialisés: Constitution de Bases de Connaissances Terminologiques* [mémoire présenté en vue de l'obtention de l'habilitation à diriger les recherches en linguistique]. Université Toulouse Le Mirail.

Corder, S. P. (1967). The Significance of Learner's Errors. *International Review of Applied Linguistics*, 5, 161-170.

Corder, S. P. (1981). *Error analysis and interlanguage*. Oxford University Press.

Corino, E. et Onesti, C. (2019). Data-Driven Learning: A Scaffolding Methodology for CLIL and LSP Teaching and Learning. *Frontiers in Education*, 4. <https://www.frontiersin.org/articles/10.3389/educ.2019.00007>

Coxhead, A. (2017). *Vocabulary and English for Specific Purposes Research: Quantitative and Qualitative Perspectives* (1^{re} éd.). Routledge. <https://doi.org/10.4324/9781315146478>

Coyle, D. (2005). CLIL: Planning tools for teachers. *Nottingham: University of Nottingham*.

Coyle, D. (2007). Content and Language Integrated Learning: Towards a Connected Research Agenda for CLIL Pedagogies. *International Journal of Bilingual Education and Bilingualism*, 10(5), 543-562. <https://doi.org/10.2167/beb459.0>

Coyle, D. (2008). CLIL—A Pedagogical Approach from the European Perspective. Dans N. H. Hornberger (dir.), *Encyclopedia of Language and Education* (p. 1200-1214). Springer US. https://doi.org/10.1007/978-0-387-30424-3_92

Coyle, D., Hood, P. et Marsh, D. (2010). *CLIL: Content and Language Integrated Learning*. Cambridge University Press. <https://doi.org/10.1017/9781009024549>

Cruse, D. A. (1986). *Lexical semantics*. Cambridge University Press.

Dalton-Puffer, C. (2007). *Discourse in Content and Language Integrated Learning (CLIL) Classrooms*. John Benjamins Publishing.

Dalton-Puffer, C., Nikula, T. et Smit, U. (2010). *Language Use and Language Learning in CLIL Classrooms*. John Benjamins Publishing.

Dancette, J. et Réthoré, C. (1997). Le Dictionnaire bilingue (anglais-français) de la distribution : entre dictionnaire de langue et encyclopédie. *Meta : journal des traducteurs / Meta: Translators' Journal*, 42(2), 229-243.

Dechamps, C. (2004). Enseignement/apprentissage des collocations d'une langue de spécialité à un public allophone : l'exemple de la langue juridique. *Ela. Etudes de linguistique appliquée*, no

135(3), 361-370.

Dechamps, C. (2012). Problématiques de l'enseignement/apprentissage du français juridique dans la formation de traducteurs juridiques. *La traduction Juridique*, 89-105.

Dechamps, C. (2013a). L'enseignement/apprentissage du français juridique : une proposition de démarche terminologique et didactique. *Equivalences*, 40(1), 191-224.

Dechamps, C. (2013b). Traduction juridique et étude des collocations : quelles perspectives? *Parallèles*, (25), 5-18.

Dechamps, C. (2015). Corpus et étude des collocations verbales. *POLISSEMA—Revista de Letras do ISCAP*, 15, 13-30.

Dechamps, C. (2017). Quelques propositions pour un dictionnaire d'apprentissage des collocations en français et en portugais. *Synergies Brésil : revue du GERFLINT*, 12, 103-118.

Département d'informatique et de recherche opérationnelle. (2020). *Expertises de recherche*. Université de Montréal. <https://diro.umontreal.ca/recherche/interets/>

Doca, G. (1981). *Analyse psycholinguistique des erreurs faites lors de l'apprentissage d'une langue étrangère: applications au domaine franco-roumain*. Publications de la Sorbonne.

Dörnyei, Z. et Taguchi, T. (2009). *Questionnaires in Second Language Research: Construction, Administration, and Processing*. Taylor & Francis Group. <http://ebookcentral.proquest.com/lib/umontreal-ebooks/detail.action?docID=465410>

Drouin, P. (2003). Term extraction using non-technical corpora as a point of leverage. *Terminology*, 9(1), 99-115.

Du, H. (2011a). *基础科技汉语教程: 听说课本 (上) [An Elementary Course in Scientific Chinese: Listening and Speaking]* (vol. I). 华语教学出版社 [Sinolingua Press].

Du, H. (2011b). *基础科技汉语教程: 听说课本 (下) [An Elementary Course in Scientific Chinese: Listening and Speaking]* (vol. II). 华语教学出版社 [Sinolingua Press].

Du, H. (2012a). *基础科技汉语教程: 阅读课本 (上) [An Elementary Course in Scientific Chinese: Reading Comprehension (volume 1)]* (vol. I). 华语教学出版社 [Sinolingua Press].

- Du, H. (2012b). *基础科技汉语教程: 阅读课本 (下)* [An Elementary Course in Scientific Chinese: Reading Comprehension (volume 2)] (vol. II). 华语教学出版社 [Sinolingua Press].
- Envoyer. (2020). Dans P. Robert, J. Rey-Debove et A. Rey (dir.), *Le Petit Robert de la langue française*. <https://petitrobert-lerobert-com.acces.bibl.ulaval.ca/robert.asp>
- Felber, H. (1980). In memory of Eugen Wüster, founder of the general theory of terminology. *International Journal of the Sociology of Language*, 23, 7-14.
- Feng, L. (2002). *词汇结构在中外汉语学习者合成词加工中的作用* [The Effect of Lexical Structure in the Processing of Compound Words by Chinese Learners from Home and Abroad] [thèse de doctorat]. Beijing Normal University.
- Feng, L. (2003). 中级汉语水平留学生的词汇结构意识与阅读能力的培养 [The intermediate Chinese learners awareness of lexical structure and the development of reading competence]. *世界汉语教学* [Chinese Teaching in the World], (2), 66-71.
- Feng, L. (2009). 外国学生汉语词素的形音义加工与心理词典的建构模式研究 [A study of the orthographic, phonological and semantic processing of Chinese morphemes by foreign students and the constructive pattern of their mental lexicon]. *世界汉语教学* [Chinese Teaching in the World], (1).
- Feng, L. et Song, Z. (2004). 词素性质与构词能力对留学生中文词汇识别的影响 [The influence of the nature and productivity of Chinese morphemes on the morphological recognition by foreign students]. *Journal of Yunnan Normal University (Teaching and Research of Chinese as a Foreign Language)*, 2(6), 33-38.
- Fillmore, C. J. (1982). Frame Semantics. Dans The Linguistic Society of Korea (dir.), *Linguistics in the Morning Calm* (p. 111-137). Hanshin.
- Fillmore, C. J. et Baker, C. (2015). A frames Approach to Semantic Analysis. Dans B. Heine et H. Narrog (dir.), *The Oxford Handbook of Linguistic Analysis* (2^e éd., p. 791-816). Oxford University Press.
- Fillmore, C. J., Petruck, M. R. L., Ruppenhofer, J. et Wright, A. (2003). Framenet in Action: The Case of Attaching. *International Journal of Lexicography*, 16(3), 297-332.

- Firth, J. R. (1957). *Papers in linguistics, 1934-1951*. Oxford University Press.
- Gabillon, Z. (2020a). Éditorial. *Contextes et didactiques. Revue semestrielle en sciences de l'éducation*, (15). <https://journals.openedition.org/ced/2047>
- Gabillon, Z. (2020b). Revisiting CLIL: Background, Pedagogy, and Theoretical Underpinnings. *Contextes et Didactiques. Revue Semestrielle En Sciences de l'éducation*, (15). <https://doi.org/10.4000/ced.1836>
- Gagne, R. M. (1965). *The Conditions of Learning*. Holt, Rinehart and Winston,.
- Gan, H. (2009). 词语结构及其识别对汉语阅读中词汇学习的影响 [The Efforts of Word Formation and Its Recognition on Vocabulary Learning in Chinese Reading]. *语言文字应用 [Applied Linguistics]*, (03), 120-128.
- Garcia, M., García-Salido, M. et Alonso Ramos, M. (2019). *Towards the Automatic Construction of a Multilingual Dictionary of Collocations using Distributional Semantics*. *Electronic lexicography in the 21st century (eLex 2019)* (p. 17).
- Gaudin, F. (1993). *Pour une socioterminologie : des problèmes sémantiques aux pratiques institutionnelles*. Publications de l'Université de Rouen.
- Gaudin, F. (2003). *Socioterminologie. Une approche sociolinguistique de la terminologie*. De Boeck/Duculot.
- Grossmann, F. (2011). Didactique du lexique : état des lieux et nouvelles orientations. *Pratiques. Linguistique, littérature, didactique*, (149-150), 163-183.
- Grossmann, F. et Tutin, A. (dir.). (2003). *Les collocations: analyse et traitement*. Éditions De Werelt.
- Guo, R. (2002). 汉语词类研究 [Studies on Word Classes in Chinese]. Dans T. Lin (dir.), *世纪中国学术大典：语言学 [A Collection of Academic Research in China: Linguistics]*. 福建教育出版社 [Fujian Education Press].
- Halliday, M. A. K. (1966). Lexis as a Linguistic Level. Dans C. E. Bazell, J. C. Catford et R. H. Robin (dir.), *In memory of J.R. Firth*. (p. 148-162). Longmans.

Han, Zhigang. (2012). *科技汉语听说教程 [Chinese For Science and Technology: Listening and Speaking]*. 北京语言大学出版社 [Beijing Language and Culture University Press].

Han, Zhiwei et L'Homme, M.-C. (2022). Adding Chinese to a multilingual terminological resource : A focus on collocations. *Lexicography: Journal of ASIALEX*, 9(1), 21-43.

Hausmann, F. J. (1979). Un dictionnaire des collocations est-il possible? *Travaux de Linguistique et de Littérature*, 17(1), 187-195.

Hausmann, F. J. (1989). Le dictionnaire de collocations. Dans F. J. Hausmann, O. Reichmann, H. E. Wiegand et L. Zgusta (dir.), *Wörterbücher/Dictionaries/Dictionnaires, Ein internationales Handbuch zur Lexikografie/An International Encyclopedia of Lexicography/Encyclopédie internationale de lexicographie*. (p. 1010-1019). De Gruyter.

Hinkel, E. (2012). Cognitive-Code Learning. Dans N. M. Seel (dir.), *Encyclopedia of the Sciences of Learning* (p. 625-626). Springer US.

Hu, M. (1996a). 兼类问题 [The Issue of Multicategoriality]. Dans M. Hu (dir.), *词类问题考察 [An Investigation of the Issue of Word Classes]* (p. 215-257). 北京语言文化大学出版社 [Beijing Language and Culture University Press].

Hu, M. (1996b). 动名兼类的计量考察 [A Quantitative Investigation of Verb-Noun Multicategory Word]. Dans M. Hu (dir.), *词类问题考察 [An Investigation of the Issue of Word Classes]* (p. 258-285). 北京语言文化大学出版社 [Beijing Language and Culture University Press].

Hu, M. (1996c). 现代汉语词类问题考察 [An Investigation of the Issue of Word Classes in Modern Chinese]. Dans M. Hu (dir.), *词类问题考察 [An Investigation of the Issue of Word Classes]* (p. 1-21). 北京语言文化大学出版社 [Beijing Language and Culture University Press].

Huong, D. A. (2010). L'approche cognitive dans l'apprentissage des langues étrangères, sa conception et ses apports à l'acquisition des compétences linguistiques, 2, 127-139.

Hyland, K. (2007). English for Specific Purposes. Dans J. Cummins et C. Davison (dir.), *International Handbook of English Language Teaching* (p. 391-402). Springer US.

https://doi.org/10.1007/978-0-387-46301-8_28

Hyland, K. (2022). English for Specific Purposes: What is it and where is it taking us? *ESP Today*, 10(2), 202-220. <https://doi.org/10.18485/esptoday.2022.10.2.1>

Hyland, K. et Tse, P. (2007). Is there an “academic vocabulary”? *TESOL quarterly*, 41(2), 235-253.

Jiang, Y. (2016). *对外汉语教学中语素教学法使用情况调查分析 [The investigation and analysis of Morpheme Pedagogy in Teaching Chinese as a Second Language]* [mémoire de maîtrise, Xinjiang University].

Jin, Y. et Guo, L. (2017). *中医汉语听说教程 [Chinois de la médecine traditionnelle chinoise : compréhension et expression orales]*. 北京语言大学出版社 [Beijing Language and Culture University Press].

Josselin-Leray, A., Rossi, C. et Sauron, V. (2019, mai). *Corpus et mémoires de traduction : deux approches du lexique et de la phraséologie en traduction spécialisée*. Nantes, France. <https://hal-univ-tlse2.archives-ouvertes.fr/hal-02132686>

Jousse, A.-L. (2010). *Modèle de structuration des relations lexicales fondé sur le formalisme des fonctions lexicales* [thèse de doctorat]. Université de Montréal et Université Paris Diderot.

Jousse, A.-L., L’Homme, M.-C., Leroyer, P. et Robichaud, B. (2011). *Presenting collocates in a dictionary of computing and the internet according to user needs*. The 5th International Conference on Meaning-Text Theory, Barcelona (p. 134-144).

Jousse, A.-L. et Polguère, A. (2005). Le DiCo et sa version DiCouèbe. Document descriptif et manuel d’utilisation. Version du rapport 1.0 – 19 avril 2005. Observatoire de linguistique Sens-Texte (OLST). <http://idefix.ling.umontreal.ca/dicouebe/DiCoDOC.pdf>

Jousse, A.-L., Polguère, A. et Tremblay, O. (2008). Du dictionnaire au site lexical pour l’enseignement/apprentissage du vocabulaire. Dans F. Grossmann et S. Plane (dir.), *Les apprentissages lexicaux : Lexique et production verbale*. (p. 141-157). Presses universitaires du Septentrion.

Kahane, S. et Polguère, A. (2001). *Formal foundation of lexical functions*. 39th Annual Meeting

and 10th Conference of the European Chapter of the Association for Computational Linguistics, Toulouse (p. 8-15).

Kilgarriff, A., Baisa, V., Bušta, J., Jakubíček, M., Kovář, V., Michelfeit, J., Rychlý, P. et Suchomel, V. (2014). The Sketch Engine: ten years on. *Lexicography*, 1(1), 7-36.

Kirby, J. (2004). Morphological awareness: Just “more phonological”? The roles of morphological and phonological awareness in reading development. *Applied Psycholinguistics*, 25, 223-238.

Kol, S., Schcolnik, M. et Spector-Cohen, E. (2018). Google Translate in academic writing courses? *The EuroCALL Review*, 26(2), 50-57.

Kromann, H.-P., Rüiber, T. et Rosbach, P. (1991). Principles of bilingual lexicography. *Wörterbücher/Dictionaries/Dictionnaires. Ein internationales Handbuch zur Lexikographie/An International Encyclopedia of Lexicography/Enciclopédie internationale de lexicographie*, 3, 2711-2728.

Lavoie, C. (2015). Trois stratégies efficaces pour enseigner le vocabulaire : une expérience en contexte scolaire innu. *The Canadian Journal of Applied Linguistics*, 18(1), 1-20.

Lavoie, C. et Hoa, L. thi. (2014). S'appropriier le vocabulaire : la stratégie « réseau de mots ». *Vivre le primaire*, 27(1), 26-27.

Le ministère de l'Éducation de la République populaire de Chine et Commission d'État pour la langue et l'écriture chinoises. (2021). *国际中文教育中文水平等级标准 [Chinese proficiency grading standards for international Chinese language education]*.

Lebrun, M. (2007). Quelques méthodes pédagogiques actives. Dans *Théories et méthodes pédagogiques pour enseigner et apprendre : Quelle place pour les TIC dans l'éducation ?* (vol. 2e éd., p. 123-168). De Boeck Supérieur.

Lebrun, M., Peltier, C., Peraya, D., Burton, R. et Mancuso, G. (2014). Un nouveau regard sur la typologie des dispositifs hybrides de formation. Propositions méthodologiques pour identifier et comparer ces dispositifs. *Éducation et formation*, (e-301), 55-74.

Lerat, P. (1995). *Les langues spécialisées* (1re éd). Presses universitaires de France.

Lexical Computing. (2020). *CQL – Corpus Query Language*. Sketch Engine. <https://www.sketchengine.eu/documentation/corpus-querying/>

L'Homme, M.-C. (1997). Organisation des classes conceptuelles pour l'accès informatisé aux combinaisons lexicales spécialisées verbe+ terme. Dans *Actes des deuxièmes rencontres Terminologie et intelligence artificielle, TIA-97* (p. 161-174). Université Toulouse-le-Mirail.

L'Homme, M.-C. (2000). Understanding specialized lexical combinations. *Terminology. International Journal of Theoretical and Applied Issues in Specialized Communication*, 6(1), 89-109.

L'Homme, M.-C. (2002). Fonctions lexicales pour représenter les relations sémantiques entre termes. *Traitement automatique de la langue*, 43, 19-41.

L'Homme, M.-C. (2003). Les combinaisons lexicales spécialisées (CLS): Description lexicographique et intégration aux banques de terminologie. Dans F. Grossmann et A. Tutin (dir.), *Les collocations: analyse et traitement (Travaux et Recherches en linguistique appliquée.) Collection associée à la Revue Française de Linguistique Appliquée. Série E-Lexicologie et lexicographie*. De Werelt.

L'Homme, M.-C. (2005). Sur la notion de « terme ». *Meta : journal des traducteurs / Meta: Translators' Journal*, 50(4), 1112-1132.

L'Homme, M.-C. (2008). Le DiCoInfo - Méthodologie pour une nouvelle génération de dictionnaires spécialisés. *Traduire*, (217), 78-103.

L'Homme, M.-C. (2009). A methodology for describing collocations in a specialized dictionary. Dans S. Nielsen et S. Tarp (dir.), *Lexicography in the 21st century: in honour of Henning Bergenholtz* (p. 237-256). John Benjamins Pub.

L'Homme, M.-C. (2016). Terminologie de l'environnement et Sémantique des cadres. *SHS Web of Conferences*, 27, 1-14.

L'Homme, M.-C. (2017). Combinatoire spécialisée : trois perspectives et des enseignements pour la terminologie. *TTR*, 30(1-2), 215-241.

L'Homme, M.-C. (2019). *TRA-6603 Module 1 Unité 3 : Aborder les textes en fonction de leur*

niveau de spécialisation [note de cours]. Note de cours communication présentée au Département de linguistique et de traduction, Université de Montréal.

L'Homme, M.-C. (2020a). *La terminologie : principes et techniques* (2^e éd.). Les Presses de l'Université de Montréal.

L'Homme, M.-C. (2020b). *Lexical semantics for terminology: an introduction*. John Benjamins Publishing Company.

L'Homme, M.-C. (2020c). *Lexical Semantics for Terminology: An introduction* (vol. 20). John Benjamins Publishing Company.

L'Homme, M.-C. (2020d, 21 novembre). *Revisiting Polysemy in Terminology*. Euralex 2020. *Lexicography for Inclusion* (vol. 1, p. 415-424).

L'Homme, M.-C. et Alipour, M. (2009). DiCoInfo. Le dictionnaire fondamental de l'informatique et d'Internet. Dictionnaire élaboré par l'équipe ÉCLECTIK. Observatoire de Linguistique Sens-Texte (OLST). <http://olst.ling.umontreal.ca/dicoinfo/Manuel%20DiCoInfo%20-%202015.pdf>

L'Homme, M.-C. et Bertrand, C. (2000). *Specialized Lexical Combinations: Should they be Described as Collocations or in Terms of Selectional Restrictions*. Stuttgart, Germany (p. 497-506). <http://olst.ling.umontreal.ca/pdf/berlhom.pdf>

L'Homme, M.-C. et Gemme, R. (1997). Modèle d'accès informatisé aux combinaisons lexicales spécialisées: verbe+ nom (terme) et extension au nom (déverbal)+ préposition+ nom (terme). Dans R. Kocourek, I. Oore, H. R. Runte et L. Lapierre (dir.), *Mélanges de linguistique offerts à Rostislav Kocourek* (p. 89-103). Presses d'ALFA.

L'Homme, M.-C. et Jia, Z. (2015). Classement des combinaisons lexicales spécialisées à base nominale dans un dictionnaire d'informatique. *Cahiers de lexicologie*, (106), 229-251.

L'Homme, M.-C. et Laneville, M.-E. (2009). DiCoEnviro. Le dictionnaire fondamental de l'environnement. Dictionnaire élaboré par l'équipe ÉCLECTIK. Observatoire de Linguistique Sens-Texte (OLST). <http://olst.ling.umontreal.ca/dicoenviro/manuel-DiCoEnviro.pdf>

L'Homme, M.-C. et Leroyer, P. (2009). Combining the semantics of collocations with situation-

driven search paths in specialized dictionaries. *Terminology*, 15(2), 258-283.

L'Homme, M.-C. et Meynard, I. (1998). Le point d'accès aux combinaisons lexicales spécialisées : présentation de deux modèles informatiques. *TTR : traduction, terminologie, rédaction*, 11(1), 199-227.

L'Homme, M.-C., Robichaud, B. et Leroyer, P. (2012). Encoding Collocations in DiCoInfo: From formal to user-friendly representations. Dans S. Granger et M. Paquot (dir.), *Electronic lexicography* (First edition, p. 211-236). Oxford University Press.

L'Homme, M.-C. et Vandaele, S. (dir.). (2007). *Lexicographie et terminologie: comptabilité des modèles et des méthodes*. Presses de l'Université d'Ottawa.

Li, K. (2002). 对外汉语教学中的词汇教学与设计 [Vocabulary Teaching and Pedagogical design in Teaching Chinese as a Foreign Language]. *语言教学与研究 [Language Teaching and Linguistic Studies]*, (5), 55-58.

Li, L. et Ding, A. (2002). *公司汉语 [Business Chinese]*. 北京大学出版社 [Peking University Press].

Li, Q. (2011). 论专门用途汉语教学 [Teaching Chinese for specific purposes and its textbook compilation]. *语言文字应用 [Applied Linguistics]*, 3, 110-117.

Li, X. (2008a). *新丝路中级速成商务汉语I [New Silk Road Business Chinese. Intermediate 1]* (édité par H. Cui). 北京大学出版社 [Peking University Press].

Li, X. (2008b). *新丝路中级速成商务汉语II [New Silk Road Business Chinese. Intermediate 2]* (édité par H. Cui). 北京大学出版社 [Peking University Press].

Li, X. (2008c). *新丝路高级速成商务汉语I [New Silk Road Business Chinese. Advanced 1]* (édité par H. Li). 北京大学出版社 [Peking University Press].

Li, X. (2008d). *新丝路高级速成商务汉语II [New Silk Road Business Chinese. Advanced 2]* (édité par H. Li). 北京大学出版社 [Peking University Press].

Liu, J. et Wei, W. (2013). 对外汉语教学中的词汇语义网络构建研究 [A Study on the

Construction of Vocabulary Semantic Network in Teaching Chinese as a Foreign Language]. *Journal of Bohai University (Philosophy & Social Science Edition)*, 1, 68-71.

Liu, X. (2000). *对外汉语教育学引论 Dui wai han yu jiao yu xue yin lun [Introduction to Pedagogy in Teaching Chinese as a Foreign Language(Chinese Edition)]* (Di 1 ban). Beijing yu yan wen hua da xue chu ban she [Beijing Language and Culture University Press].

Liu, Y., Pan, W. et Gu, W. (2004). *实用现代汉语语法 [Practical Modern Chinese Grammar]*. 商务印书馆 [The Commercial Press].

logDice. (2016, 22 juin). *Sketch Engine*. https://www.sketchengine.eu/my_keywords/logdice/

Lü, B. (1990). *对外汉语教学发展概要 [Development of Teaching Chinese as a Foreign Language]* (1^{re} éd.). 北京语言大学出版社 [Beijing Language and Culture University Press].

Lu, Jianming. (1994). 关于词的兼类问题 [On the Issue of Multicategorical Words]. *中国语文 [Chinese Linguistics]*, (1), 28-34.

Lu, Jianming. (2003a). *现代汉语语法研究教程 [Modern Chinese Grammar Research Course]*. 北京大学出版社 [Peking University Press].

Lu, Jianming. (2003b). *现代汉语语法研究教程 [Un cour d'études sur la grammaire du chinois moderne]*. 北京大学出版社 [Peking University Press].

Lu, Jianming. (2011). 我关于“字本位”的基本观点 My Position on Chinese Character as the Very Basic Unit of Chinese Grammar. *Yuyan Kexue*, 52, 225-230.

Lu, Jiansheng, Wang, L. et Wu, H. (2017). *科技汉语简明教程 [A Concise Course of Scientific and Technological Chinese]*. 科学技术文献出版社 [Scientific and Technical Documentation Press].

Lü, W. (1999). *对外汉语教学语法体系研究 [A Study on the Grammar Teaching System of Chinese as a Foreign Language]*. 北京语言文化大学出版社 [Beijing Language and Culture University Press].

Lu, Zhiwei. (1964). *汉语的构词法 [Word Formation in Chinese]*. 科学出版社 [Scientific

Press].

Lu, Zhiying. (2006). 商贸类汉语教材编写和研究的基本情况述评 [A Summary Review of the Compilation and Study of Business Chinese Textbooks]. *云南师范大学学报(对外汉语教学与研究版)* [Journal of Yunnan Normal University: Teaching and Research of Chinese as A Foreign Language], 4(5), 28-33.

Lux-Pogodalla, V. et Polguère, A. (2011, août). *Construction of a French Lexical Network: Methodological Issues*. Ljubljana, Slovenia (p. 54-61).

MDBG Chinese-English dictionary. (2022). [MDBG.NET. https://www.mdbg.net/chinese/dictionary?page=worddict&wdrst=0&wdqb=](https://www.mdbg.net/chinese/dictionary?page=worddict&wdrst=0&wdqb=)

Mel'čuk, I. (1981). Meaning-Text Models: A Recent Trend in Soviet Linguistics. *Annual Review of Anthropology*, 10(1), 27-62. <https://doi.org/10.1146/annurev.an.10.100181.000331>

Mel'čuk, I. (1992). Paraphrase et lexique: la théorie Sens-Texte et le Dictionnaire explicatif et combinatoire. Dans *Dictionnaire explicatif et combinatoire du français contemporain. Recherches lexico-sémantiques III*. (p. 9-59). Les Presses de l'Université de Montréal.

Mel'čuk, I. (1993). *Cours de morphologie générale. Volume 1*. Presse de l'Université de Montréal/Éditions du Centre National de la Recherche Scientifique.

Mel'čuk, I. (1996). Lexical Functions: A Tool for the Description of Lexical Relations in a Lexicon. Dans L. Wanner (dir.), *Lexical functions in lexicography and natural language processing* (vol. 31, p. 37-102). John Benjamins Publishing.

Mel'čuk, I. (1997). *Vers une linguistique Sens-Texte* [leçon inaugurale faite le vendredi 10 janvier 1997.]. Leçon inaugurale faite le vendredi 10 janvier 1997. communication présentée au Collège de France, Chaire internationale.

Mel'čuk, I. (2003). Collocations dans le dictionnaire. Dans T. Szende (dir.), *Les écarts culturels dans les Dictionnaires bilingues* (Honoré Champion, p. 19-64). <http://olst.ling.umontreal.ca/pdf/Collocations-szende.pdf>

Mel'čuk, I., Arbatchewsky-Jumarie, N. et Clas, A. (1999). *Dictionnaire explicatif et combinatoire du français contemporain: recherches lexico-sémantiques IV*. Presses de

l'Université de Montréal.

Mel'čuk, I. et Clas, A. (1988). *Dictionnaire explicatif et combinatoire du français contemporain: recherches lexico-sémantiques II*. Presses de l'Université de Montréal.

Mel'čuk, I. et Clas, A. (1992). *Dictionnaire explicatif et combinatoire du français contemporain: recherches lexico-sémantiques III*. Presses de l'Université de Montréal.

Mel'čuk, I., Clas, A. et Arbatchewsky-Jumarie, N. (1984). *Dictionnaire explicatif et combinatoire du français contemporain: recherches lexico-sémantiques I*. Presses de l'Université de Montréal.

Mel'čuk, I., Clas, A. et Polguère, A. (1995). *Introduction à la lexicologie explicative et combinatoire*. Duculot.

Mel'čuk, I. et Collège de France. (1997). *Leçon inaugurale faite le vendredi 10 janvier 1997: vers une linguistique sens-texte*. Collège de France.

Mel'čuk, I., Jumarie, Nadia. A., Iordanskaja, L. et Mantha, S. (1992). *Dictionnaire explicatif et combinatoire du français contemporain. V. III: Recherches lexico-sémantiques* (édité par A. Clas). Presses de l'Université de Montréal.

Mel'čuk, I. et Polguère, A. (2007). *Lexique actif du français: l'apprentissage du vocabulaire fondé sur 20 000 dérivations sémantiques et collocations du français* (1re éd.). De Boeck.

Mel'čuk, I. et Polguère, A. (2008). Prédicats et quasi-prédicats sémantiques dans une perspective lexicographique. *Lidil. Revue de linguistique et de didactique des langues*, (37), 99-114.

Mel'čuk, I. et Polguère, A. (2021). Les fonctions lexicales dernier cri. Dans S. Marengo (dir.), *La Théorie Sens-Texte. Concepts-clés et applications* (p. 75-155). L'Harmattan.

Mel'čuk, I., Polguère, A., Kahane, S. et Steinlin, J. (2013, 25 juillet). *DiCouèbe. Dictionnaire en ligne de combinatoire du français*. <http://olst.ling.umontreal.ca/dicouebe/index.php>

Meyer, I. et Mackintosh, K. (1994). Phraseme analysis and concept analysis: Exploring a symbiotic relationship in the specialized lexicon. Dans W. Martin, W. Meijs, M. Moerland, E. ten Pas, P. van Sterkenburg et P. Vossen (dir.), *Euralex 1994: proceedings of the 6th EURALEX International Congress on Lexicography*. (p. 339-348). Euralex.

Meynard, I. (1998). Combinaisons lexicales en langue de spécialité : problèmes et perspectives. Dans Z. Guével et E. Valentine (dir.), *Traduction et langues de spécialité : approches théoriques et considérations pédagogiques* (p. 123-138). CIRAL.

Milićević, J. (2006). A SHORT GUIDE TO THE MEANING-TEXT LINGUISTIC THEORY. *Journal of Koralex*, 8, 187-233.

Milićević, J. (2008). *Paraphrase as a Tool for Achieving Lexical Competence in L2*. Bruxelles (p. 153-167).

Milićević, J. et Hamel, M.-J. (2007). Un dictionnaire de reformulation pour les apprenants du français langue seconde. *Revue de l'Université de Moncton*, (Numéro hors-série), 145-167.

Most frequent or most typical collocations? (2017, 21 août). *Sketch Engine*. <https://www.sketchengine.eu/blog/most-frequent-or-most-typical-collocations/>

Nguyen, É. V. T. (2006). *Unité lexicale et morphologie en chinois mandarin. Vers l'élaboration d'un Dictionnaire Explicatif et Combinatoire du chinois* [thèse de doctorat]. Université de Montréal.

Observatoire de linguistique Sens-Texte. (2007). Détecter. Dans *DiCoInfo. Dictionnaire fondamental de l'informatique et de l'Internet*. http://olst.ling.umontreal.ca/dicoinfo/dicohtml-bilingue/dicohtml-fr/D-fr.html#_d%C3%A9tecterfr

Observatoire de linguistique Sens-Texte. (2008). Accéder 1. Dans *DiCoInfo. Dictionnaire fondamental de l'informatique et de l'Internet*. http://olst.ling.umontreal.ca/cgi-bin/dicoinfo/search.cgi?ui=fr&lang=fr&mode=id&prec=exact&equi=1&rq=id:_acc%C3%A9der1

Observatoire de Linguistique Sens-Texte. (2011). Guide pour la création de fiches terminologiques dans le DiCoInfo et le DiCoEnviro. <http://olst.ling.umontreal.ca/dicoinfo/documentation/Tutoriel-Oxygen-Dicos.pdf>

Observatoire de linguistique Sens-Texte. (2013a). Antivirus. Dans *DiCoInfo. Dictionnaire fondamental de l'informatique et de l'Internet*. <http://olst.ling.umontreal.ca/cgi-bin/dicoinfo/search.cgi?ui=fr&lang=fr&mode=terme&prec=exact&equi=1&rq=antivirus>

Observatoire de linguistique Sens-Texte. (2013b). Exécuter. Dans *DiCoInfo. Dictionnaire fondamental de l'informatique et de l'Internet*. http://olst.ling.umontreal.ca/cgi-bin/dicoinfo/search.cgi?&ui=fr&mode=terme&lang=fr&prec=exact&ract=0&equi=1&rq=ex%C3%A9cuter#_ex%C3%A9cuter1bfr

Observatoire de linguistique Sens-Texte. (2013c). S'exécuter. Dans *DiCoInfo. Dictionnaire fondamental de l'informatique et de l'Internet*. http://olst.ling.umontreal.ca/cgi-bin/dicoinfo/search.cgi?ui=fr&lang=fr&mode=id&prec=exact&equi=1&rq=id:_ex%C3%A9cuter1a

Observatoire de linguistique Sens-Texte. (2013d, 11 février). Logiciel. Dans *DiCoInfo. Dictionnaire fondamental de l'informatique et de l'Internet*. <http://olst.ling.umontreal.ca/cgi-bin/dicoinfo/search.cgi?ui=fr&lang=fr&mode=terme&prec=exact&equi=1&rq=logiciel>

Observatoire de linguistique Sens-Texte. (2015a). Accéder 2. Dans *DiCoInfo. Dictionnaire fondamental de l'informatique et de l'Internet*. http://olst.ling.umontreal.ca/cgi-bin/dicoinfo/search.cgi?ui=fr&lang=fr&mode=id&prec=exact&equi=1&rq=id:_acc%C3%A9der2

Observatoire de linguistique Sens-Texte. (2015b). Amovible. Dans *DiCoInfo. Dictionnaire fondamental de l'informatique et de l'Internet*. <http://olst.ling.umontreal.ca/cgi-bin/dicoinfo/search.cgi?ui=fr&lang=fr&mode=terme&prec=exact&equi=1&rq=amovible>

Observatoire de linguistique Sens-Texte. (2015c). Matériel 1. Dans *DiCoInfo. Dictionnaire fondamental de l'informatique et de l'Internet*. http://olst.ling.umontreal.ca/dicoinfo/dicohtml-bilingue/dicohtml-fr/M-fr.html#_mat%C3%A9rielfr

Observatoire de linguistique Sens-Texte. (2015d). Programme. Dans *DiCoInfo. Dictionnaire fondamental de l'informatique et de l'Internet*. <http://olst.ling.umontreal.ca/cgi-bin/dicoinfo/search.cgi?ui=fr&lang=fr&mode=terme&prec=exact&equi=1&rq=programme>

Observatoire de linguistique Sens-Texte. (2022a). *DiCoEnviro. Dictionnaire fondamental de l'environnement*. http://olst.ling.umontreal.ca/cgi-bin/dicoenviro/search_enviro.cgi

Observatoire de linguistique Sens-Texte. (2022b). *DiCoInfo. Dictionnaire fondamental de l'informatique et de l'Internet*. <http://olst.ling.umontreal.ca/cgi-bin/dicoinfo/search.cgi?ui=fr>

Office québécois de la langue française. (2005). Application de l'informatique. Dans *Grand dictionnaire terminologique*. https://gdt.oqlf.gouv.qc.ca/ficheOqlf.aspx?Id_Fiche=8355610

Office québécois de la langue française. (2016). Mobile. Dans *Grand dictionnaire terminologique*. https://gdt.oqlf.gouv.qc.ca/ficheOqlf.aspx?Id_Fiche=8396805

Office québécois de la langue française. (2020a). Intrusion informatique. Dans *Grand dictionnaire terminologique*. https://gdt.oqlf.gouv.qc.ca/ficheOqlf.aspx?Id_Fiche=8384979

Office québécois de la langue française. (2020b). Logiciel. Dans *Grand dictionnaire terminologique*. https://gdt.oqlf.gouv.qc.ca/ficheOqlf.aspx?Id_Fiche=8355564

Office québécois de la langue française. (2020c). Logiciel espion. Dans *Grand dictionnaire terminologique*. https://gdt.oqlf.gouv.qc.ca/ficheOqlf.aspx?Id_Fiche=8369709

Office québécois de la langue française. (2021a). Attaque par déni de service. Dans *Grand dictionnaire terminologique*. https://gdt.oqlf.gouv.qc.ca/ficheOqlf.aspx?Id_Fiche=26557495

Office québécois de la langue française. (2021b). Attaque par interception. Dans *Grand dictionnaire terminologique*. https://gdt.oqlf.gouv.qc.ca/ficheOqlf.aspx?Id_Fiche=26542694

Office québécois de la langue française. (2021c). réseau. Dans *Grand dictionnaire terminologique*. https://gdt.oqlf.gouv.qc.ca/ficheOqlf.aspx?Id_Fiche=8391527

Office québécois de la langue française. (2022). *Grand dictionnaire terminologique*. Grand dictionnaire terminologique. <https://gdt.oqlf.gouv.qc.ca/>

Packard, J. L. (2000). *The Morphology of Chinese: A Linguistic and Cognitive Approach*. Cambridge University Press.

Pavel, S. (1993a). La phraséologie en langue de spécialité. Méthodologie de consignation dans les vocabulaires terminologiques. *Terminologies nouvelles*, 10, 67-82.

Pavel, S. (1993b). Vers une méthode de recherche phraséologique en langue de spécialité. *L'actualité terminologique/Terminology update*, 26(2), 9-13.

Pavel, S. (1994). *Guide de la recherche phraséologique en langue de spécialité*. Division Montréal-Québec, Direction de la terminologie et de la documentation, Bureau de la traduction.

Pecman, M. (2004). *Phraséologie contrastive anglais-français : analyse et traitement en vue de l'aide à la rédaction scientifique*. [thèse de doctorat]. Université de Nice-Sophia Antipolis.

Pecman, M. (2005). Les apports possibles de la phraséologie à la didactique des langues étrangères. *Apprentissage des langues et systèmes d'information et de communication*, 8(1), 109-122.

Pecman, M. (2007). Approche onomasiologique de la langue scientifique générale. *Revue française de linguistique appliquée*, 12(2), 79-96.

Pecman, M. (2012). Étude lexicographique et discursive des collocations en vue de leur intégration dans une base de données terminologiques. *The Journal of specialised translation (JoSTrans)*, (18), 113-138.

Pecman, M., Juilliard, C., Kübler, N. et Volanschi, A. (2009). Processing collocations in a terminological database based on a cross-disciplinary study of scientific texts. *eLEX2009*, 177.

Pecman, M. et Kübler, N. (2011, 1 août). *ARTES: an online lexical database for research and teaching in specialized translation and communication*. Ljubljana, Slovenia (p. 87-93).

Perdue, C. (1980). L'analyse des erreurs : un bilan pratique. *Langages*, 14(57), 87-94.

Pérez-Cañado, M. L. (2012). CLIL research in Europe: past, present, and future. *International Journal of Bilingual Education and Bilingualism*, 15(3), 315-341. <https://doi.org/10.1080/13670050.2011.630064>

Picoche, J. (1992). *Précis de lexicologie française : l'étude et l'enseignement du vocabulaire*. Éditions Vigdor.

Picoche, J. (1993a). *Didactique du vocabulaire français*. Nathan.

Picoche, J. (1993b). La cohérence des polysèmes : un outil pour débloquent l'enseignement du vocabulaire. *Repères. Recherches en didactique du français langue maternelle*, 8(1), 11-28.

Picoche, J. (1999). Dialogue autour de l'enseignement du vocabulaire. *Étude de Linguistique Appliquée*, (116), 421-434.

Picoche, J. (2016). *L'enseignement du vocabulaire*. Vocanet - Didactique du vocabulaire. <http://www.vocanet.fr/index.php/2-non-categorise/190-didactique-du-vocabulaire>

Picoche, J. et Rolland, J.-C. (2002). *Dictionnaire du français usuel: 15000 mots utiles en 442 articles*. De Boeck : Duculot.

Picoche, J. et Souhaité, S. (2004). L'utilisation du Dictionnaire du français usuel pour l'enseignement du vocabulaire. Dans E. Calaque et J. David (dir.), *Didactique du lexique : contextes, démarches, supports* (p. 215-225). De Boeck Supérieur.

Pimentel, J. (2012). *Identifying equivalents of specialized verbs in a bilingual comparable corpus of judgments: A frame-based methodology*. The eighth international conference on Language Resources and Evaluation (p. 1791-1798).

Pimentel, J., L'Homme, M.-C. et Laneville, M.-È. (2011). General and specialized lexical resources: A study on the potential of combining efforts to enrich formal lexicons. *International Journal of Lexicography*, 25(2), 152-190.

Pleco Inc. (2022). Pleco Chinese Dictionary. <https://www.pleco.com/products/pleco-for-iphone-ipad/>

Poizat, G. H. (2011). 构词法与对外汉语教学 [La morphologie chinoise et l'enseignement du chinois langue étrangère]. *Chinesisch Unterricht (CHUN)*, 26, 33-44.

Poizat, G. H. (2017). 词汇扩展与构词法 [Vocabulary Expansion and the Morphology]. Dans *Applied Chinese Language Studies IX* (p. 58-64). Sinolingua.

Polguère, A. (1998). La théorie Sens-Texte. *Dialangue*, 8-9, 9-30.

Polguère, A. (2000a). *Towards a theoretically-motivated general public dictionary of semantic derivations and collocations for French*. Proceedings of the 9th EURALEX International Congress (p. 517-527).

Polguère, A. (2000b). Une base de données lexicales du français et ses applications possibles en didactique. *Revue de Linguistique et de Didactique des Langues (LIDIL)*, 21, 75-97.

Polguère, A. (2003). Collocations et fonctions lexicales : pour un modèle d'apprentissage. Dans F. Grossmann et A. Tutin (dir.), *Les collocations: analyse et traitement (Travaux et Recherches en linguistique appliquée.)* (p. 117-133). De Werelt.

Polguère, A. (2007). Soleil insoutenable et chaleur de plomb: le statut linguistique des greffes

collocationnelles. Dans M.-C. L'Homme et S. Vandaele (dir.), *Lexicographie et terminologie : compatibilité des modèles et méthodes* (p. 247-291). Presses de l'Université d'Ottawa.

Polguère, A. (2014). From Writing Dictionaries to Weaving Lexical Networks. *International Journal of Lexicography*, 27(4), 396-418.

Polguère, A. (2016). *Lexicologie et sémantique lexicale : notions fondamentales* (3^e éd.). Les Presses de l'Université de Montréal.

Polguère, A. (2018). Traitement lexicographique des collocations à collocatif actanciel. Dans I. Sfar et P.-A. Buvet (dir.), *La phraséologie entre fixité et congruence. Hommage à Salah Mejri* (p. 265-287). Academia-L'Harmattan.

Pottier, B. (1992). *Sémantique générale*. Presses Universitaires de France.

Redmond, L. et Emirkanian, L. (2018). *Analyse de la polysémie verbale : apports à la didactique du français L2*. 6e Congrès Mondial de Linguistique Française (CMLF 2018), Mons, Belgique.

Richards, J. C. et Rodgers, T. S. (2001). *Approaches and Methods in Language Teaching*. Cambridge University Press.

Roland, L. (2022). *Qu'est-ce que le XML ?* Structurez vos données avec XML. <https://tutoriel-xml.rolandl.fr/partie-1-les-bases-du-xml/quest-ce-que-le-xml.html#:~:text=Enfin%2C%20les%20langages%20de%20description,d'un%20CD%2C%20etc>

Rousseau, L.-J. (1993). Terminologie et phraséologie, deux composantes indissociables des langues de spécialités. *Terminologies nouvelles*, (10), 9-11.

Ruppenhofer, J., Ellsworth, M., Petruck, M. R. L., Johnson, C. R. et Scheffczyk, J. (2010). *FrameNet II: Extended Theory and Practice*. <https://framenet2.icsi.berkeley.edu/docs/r1.5/book.pdf>

Rychlý, P. (2008). *A Lexicographer-Friendly Association Score*. P. Sojka et A. Horák (dir.), Second Workshop on Recent Advances in Slavonic Natural Language Processing, RASLAN 2008, Karlova Studánka, Czech Republic (p. 6-9).

Sardier, A. (2015). *Construire la compétence lexicale : quelle place en didactique pour le cotexte ?* [thèse de doctorat]. Université Grenoble Alpes.

Shan, Y. (2008). Issues on the Compiling of Chinese Textbooks for Specific Purpose: Illustrated by Compiling Chinese Reading on Science and Technology. *Journal of College of Chinese Language and Culture of Jinan University*, 2, 31-37.

Shan, Y. et An, R. (2009). 专门用途汉语课程设置探析——以《科技汉语》课程为例 [Exploring curriculum design of Chinese for specific purposes: The case of Chinese for technology]. *Journal of Southwest University of Nationality (Humanities and Social Science Edition)*, (8), 258-263.

Sheng, Y. (1990). *语言教学原理 [Principles of Language Teaching]*. 重庆出版社 [Chongqing Publishing House].

Sinclair, J. (1987). Collocation: a progress report. Dans T. Threadgold et R. Steel (dir.), *Language Topics: Essays in Honour of Michael Halliday*. (p. 319-331). John Benjamins Publishing.

Sinclair, J. (1991). *Corpus, concordance, collocation*. Oxford University Press.

Skinner, B. F. (1957). *Verbal Behavior*. Appleton-Century-Crofts.

Stahlmann, R. (2022). *Dictionnaire Hantrainerpro*. Han Trainer Pro. <https://www.hantrainerpro.com/chinois/dictionnaire/>

Statistiques Canada. (2021a). [4.5 Mesures de la dispersion] 4.5.3 Calculer la variance et l'écart-type. Statistique Canada. <https://www150.statcan.gc.ca/n1/edu/power-pouvoir/ch12/5214876-fra.htm#toc-tdm>

Statistiques Canada. (2021b). [4. Exploration des données] 4.5 Mesures de la dispersion. Statistique Canada. <https://www150.statcan.gc.ca/n1/edu/power-pouvoir/ch12/5214876-fra.htm#toc-tdm>

Statistiques Canada. (2021c). [5. Visualisation des données] 5.2 Graphique à barres. Statistique Canada. <https://www150.statcan.gc.ca/n1/edu/power-pouvoir/ch9/bargraph-diagrammeabarres/5214818-fra.htm>

Tao, H. et Chen, H. H.-J. (dir.). (2019). *Chinese for Specific and Professional Purposes: Theory, Pedagogical Applications, and Practices*. Springer Singapore. <https://doi.org/10.1007/978-981->

13-9505-5

Tardif, J. (1992). *Pour un enseignement stratégique: l'apport de la psychologie cognitive*. Éditions Logiques.

Tardif, J. (1999). *Le transfert des apprentissages*. Les Éditions Logiques.

Tardif, M., Bélanger, N. et Grenier, D. (1994). L'idéologie cognitiviste et l'éducation. *Canadian Journal of Education / Revue canadienne de l'éducation*, 19(3), 316-330.

Tarp, S. (2005). The pedagogical dimension of the well-conceived specialised dictionary. *Ibérica: Revista de la Asociación Europea de Lenguas para Fines Específicos (AELFE)*, (10), 7-21.

Tarp, S. (2008). *Lexicography in the Borderland Between Knowledge and Non-Knowledge: General Lexicographical Theory with Particular Focus on Learner's Lexicography*. De Gruyter, Inc.

Tarp, S. (2010). Functions of specialised learners' dictionaries. Dans P. A. Fuertes Olivera (dir.), *Specialised dictionaries for learners* (p. 39-53). De Gruyter.

Temmerman, R. (2000). *Towards new ways of terminology description: The sociocognitive-approach*. John Benjamins Publishing.

Thoiron, P. et Béjoint, H. (1989). Pour un index évolutif et cumulatif de cooccurrents en langue techno-scientifique sectorielle. *Meta : journal des traducteurs / Meta: Translators' Journal*, 34(4), 661-671.

Travaux publics et services gouvernementaux Canada. (2022). *Termium Plus*. Termium Plus. <http://www.btb.termiumplus.gc.ca>

Tremblay, O. (2003). *Une approche structurée de l'enseignement du lexique en français langue maternelle basée sur la lexicologie explicative et combinatoire* [mémoire de maîtrise]. Université de Montréal.

Tremblay, O. (2004). Pour une approche structurée de l'enseignement-apprentissage du lexique. Dans E. Calaque et J. David (dir.), *Didactique du lexique : Contextes, démarches, supports* (p. 127-139). De Boeck Supérieur.

Tremblay, O. (2009). *Une ontologie des savoirs lexicologiques pour l'élaboration d'un module de cours en didactique du lexique* [thèse de doctorat]. Université de Montréal.

Tutin, A. (2007). Autour du lexique et de la phraséologie des écrits scientifiques. *Revue française de linguistique appliquée*, XII(2), 5.

Tutin, A. (2010a). *Le traitement des collocations dans les dictionnaires monolingues de collocations du français et de l'anglais*. 2ème Congrès Mondial de Linguistique Française, La Nouvelle-Orléans, Etats-Unis (p. 071).

Tutin, A. (2010b). *Sens et combinatoire lexicale: de la langue au discours* [dossier en vue de l'habilitation à diriger des recherches, Université de Grenoble 3].

Tutin, A. (2013). Les collocations lexicales : une relation essentiellement binaire définie par la relation prédicat-argument. *Langages*, 189(1), 47.

Tutin, A. (2014). La phraséologie transdisciplinaire des écrits scientifiques: des collocations aux routines sémantico-rhétoriques. Dans A. Tutin et F. Grossmann (dir.), *L'écrit scientifique: du lexique au discours. Autour de scientext* (Presses Universitaires de Rennes, p. 27-44).

Tutin, A. (2018). Les expressions polylexicales transdisciplinaires dans les articles de recherche en sciences humaines : retour d'expérience. Dans M.-P. Jacques et A. Tutin (dir.), *Lexique transversal et formules discursives des sciences humaines* (p. 73-89). ISTE Group.

Tutin, A. et Grossmann, F. (2002). Collocations régulières et irrégulières : esquisse de typologie du phénomène collocatif. *Revue française de linguistique appliquée*, Vol. VII(1), 7-25.

Tutin, A. et Grossmann, F. (dir.). (2014). *L'écrit scientifique: du lexique au discours. Autour de scientext* (Presses Universitaires de Rennes).

Tutin, A. et Kraif, O. (2016). Routines sémantico-rhétoriques dans l'écrit scientifique de sciences humaines : l'apport des arbres lexico-syntaxiques récurrents. *Lidil. Revue de linguistique et de didactique des langues*, (53), 119-141.

Tutin, A., Novakova, I., Grossmann, F. et Cavalla, C. (2006). Esquisse de typologie des noms d'affect à partir de leurs propriétés combinatoires. *Langue française*, 150, 32-49.

Université de Montréal. (2020). *Répertoire des cours en informatique*. Université de Montréal.

https://admission.umontreal.ca/repertoire-des-cours/filtres/cycle_1/matiere_IFT/

Urmeneta, C. E. et Sola, A. S. (2009). Language learning through tasks in a Content and Language Integrated Learning (CLIL) science classroom. *Porta Linguarum*, 11(11), 65-83.

Verlinde, S., Selva, T. et Binon, J. (2003). Les collocations dans les dictionnaires d'apprentissage: repérage, présentation et accès. Dans F. Grossmann et A. Tutin (dir.), *Les collocations : analyse et traitement* (p. 105-115).

Verlinde, S., Selva, T. et Binon, J. (2005). Dictionnaires électroniques et environnement d'apprentissage du lexique. *Revue française de linguistique appliquée*, Vol. X(2), 19-30.

VerticalScope Inc. (2022). *Dictionnaire chinois français*. Chine-Nouvelle. <https://www.chine-nouvelle.com/outils/dictionnaire.html>

Vincze, O., Mosqueira, E. et Alonso Ramos, M. (2011). *An online collocation dictionary of Spanish* (p. 275-286).

Wang, J. et Xing, H. (2010). 留学生单音节多义语素构词习得过程的实验研究 [The Investigation of Overseas Students' Acquisition of Monosyllabic Polysemous Morphemes]. *语言教学与研究 [Language Teaching and Linguistic Studies]*, (2), 10-16.

Wang, Z. et Qing, X. (2004). 语素教学是对外汉语词汇教学的基础 [Morpheme teaching, the basis of vocabulary teaching in TCFL (Teaching Chinese as a Foreign Language)]. *云南师范大学学报 [Journal of Yunnan Normal University]*, (5).

Wanner, L., Alonso Ramos, M., Vincze, O., Nazar, R., Ferraro, G., Mosqueira, E. et Prieto, S. (2013). *Annotation of collocations in a learner corpus for building a learning environment*. Twenty years of learner corpus research. Looking back, moving ahead (p. 493-503).

Wierzbicka, A. (1996). *Semantics : Primes and Universals*. (Oxford University Press).

Williams, G. (1999). *Les réseaux collocationnels dans la construction et l'exploitation d'un corpus dans le cadre d'une communauté de discours scientifique*. [thèse de doctorat, Université de Nantes].

Williams, G. (2001). Mediating between lexis and texts: collocational networks in specialised

corpora. *ASp. La Revue Du GERAS*, (31-33), 63-76.

Wüster, E. (1974). Die allgemeine Terminologielehre - ein Grenzgebiet zwischen Sprachwissenschaft, Logik, Ontologie, Informatik und den Sachwissenschaften. *Linguistics*, 199(1), 61-106.

Wüster, E. (1976). *La théorie générale de la terminologie—un domaine interdisciplinaire impliquant la linguistique, la logique, l'ontologie, l'informatique et les sciences des objets*. Dupuis (dir.), (p. 49-57).

Xiao, X. (2002). 对外汉语词汇教学中“语素法”的几个问题 [Several Problems of « Morpheme Method » in Teaching Chinese as a Foreign Language]. *汉语学习 [Chinese Language Learning]*, 6, 68-73.

Xie, Y. (2004). “语素教学法”初探 [A Discussion on the « Teaching Methodology of Morphemes » in TCSL] [mémoire de maîtrise, 华中科技大学 [Huazhong University of Science and Technology]].

Xing, H. (2003). 留学生偏误合成词的统计分析 [An analysis of the error compound words used by foreign learners of Chinese]. *世界汉语教学 [Chinese Teaching in the World]*, 4, 67-78.

Xing, H. (2006). 《(汉语水平)词汇等级大纲》双音合成词语素统计分析 [A Statistic Analysis on Morphemes from the Disyllabic Words in Graded Vocabulary for HSK]. *世界汉语教学 [Chinese Teaching in the World]*, (3), 63-71.

Xu, J. (2013). *中医汉语综合教程 [Chinois de la médecine traditionnelle chinoise : cours intégré]*. 北京语言大学出版社 [Beijing Language and Culture University Press].

Xu, L., Jia, L. et Wu, C. (2012). *电力科技汉语阅读教程 [Chinois des science et technologie de l'énergie électrique : manuel de compréhension écrite]*. 中国水利水电出版社 [China Water & Power Press].

Xu, T. (1994a). “字”和汉语研究的方法论 [“Character” and the Methodology of Chinese Studies]. *世界汉语教学 [Chinese Teaching in the World]*, (3), 1-14.

- Xu, T. (1994b). “字”和汉语的句法结构 [“Words” and Chinese Syntactic Structure]. *世界汉语教学* [*Chinese Teaching in the World*], 2.
- Xu, T. (2005). 字本位和语言研究 [Characters-based Theory and Language Studies]. *语言教学与研究* [*Language Teaching and Linguistic Studies*], 6.
- Xu Z. et Wu R. (2005). *实用对外汉语教学法 Shi yong dui wai han yu jiao xue fa* [*Practical Method of Teaching Chinese as a Foreign Language*]. Peking University Press.
- Yang, T. (2011). *立足于对外汉语词汇教学的“语素法”* [*Morpheme Method in Teaching Chinese as a Foreign Language*] [mémoire de maîtrise]. Hubei University.
- Ying, X. (2013). *Approche cognitive en didactique des langues : analyse et interprétation d’erreurs écrites prototypiques en français langue étrangère par des apprenants chinois et remédiation* [thèse de doctorat]. Université de Wuhan & Université Bordeaux III.
- Yu, S., Duan, H. et Zhu, X. (2005). *词语兼类暨动词向名词漂移现象的计量分析* [*A Quantitative Analysis on Multi-class Words and Shift from Verbs to Nouns in Chinese*] (p. 70-76).
- Zhai, N. (2009). *试论专门用途汉语体系的构建* [On the Construction of Chinese for Specific Purposes Teaching System]. *上海工程大学教育研究* [*Educational Research of Shanghai University of Engineering Science*], 2, 39-45.
- Zhang, L., Zhang, Y. et Gao, Y. (2016). *专门用途汉语教学：对外汉语/汉语国际教育研究生系列教材* [*Teaching Chinese for Special Purposes : Graduate Textbook for Teaching Chinese as a Foreign Language/Chinese International Education*]. 北京语言大学出版社 [Beijing Language and Culture University Press].
- Zhang, P. (1992). *词本位教学法和字本位教学法的比较* [Comparison of word-based teaching method and character-based teaching method]. *世界汉语教学* [*Chinese Teaching in the World*], (3), 222-223.
- Zhang, S. (2014). *IT汉语与中文电脑操作* [*Chinois de l’informatique et des opérations informatiques*]. 华南理工大学出版社 [South China University of Technology Press].

- Zhang, T. (2007). *法律汉语：商事篇 [Chinois juridique : droit des affaires]*. 北京大学出版社 [Peking University Press].
- Zhang, Y. (2014). 近30年科技汉语教材编写情况的回顾与思考. *Publishing Research*, (11), 66-68.
- Zhao, W. (2017). “语素法”和“语境法”汉语二语词汇教学效果的对比研究 [A Comparative Study on the Morphemic and Contextual Vocabulary Teaching Methods to Learners of Chinese as a Second Language]. *语言教学与研究 [Language Teaching and Linguistic Studies]*, 186(4), 16-25.
- Zhao, X. (2018). *Les collocations du champ sémantique des émotions en mandarin* [mémoire de maîtrise]. Université de Montréal.
- Zhu D. (1982). *语法讲义 [Grammar Notes]*. 商务印书馆 [The Commercial Press].
- Zinglé, H. et Brobeck-Zinglé, M.-L. (2003). *Dictionnaire combinatoire du français: expressions, locutions et constructions*. Maison du Dictionnaire.
- Žolkovskij, A. et Mel'čuk, I. (1965). O Vozmoznom Metode i Instrumentax Semanticeskogo Sinteza. [À propos d'une méthode possible pour la synthèse sémantique et des instruments correspondants]. *Naucno-TeXniceskaja Informacija*, 5, 23-28.

Annexe 1 – Article *admiration* du LAF

ADMIRATION, NOM, FÉM; PAS DE PL

SENTIMENT POSITIF

Admiration de l'individu X [= de N, A_{poss}] envers Y [= «à l'égard de» N, devant N, envers N, pour N] pour Z
[= «à l'égard de» N, «à propos de» N, devant N, pour N, pour V_{inf-passé}]



émerveillement < adoration **a**; enthousiasme; ravissement

Ant. aversion; mépris **Verbe** admirer **I** [N_Y pour N_Z] **Nom pour X qui éprouve de l'A. pour tout ce que fait Y** admirateur [de N_{Y/Z}] **Nom pour Y objet** [de (ART) ~] **Nom pour Z objet** [de (ART) ~], source [d'~] **[X] qui éprouve de l'A.** en [-] //admiratif < plein, rempli, soulevé, transporté [de (ART) ~] **[X faire qqch.] en éprouvant de l'A.** avec [-] **[Y/Z] envers qui/quoi on tend à éprouver de l'A.** digne [de (ART) ~] //admirable **[Z] tendre à être un objet d'A.** porter [à l'~] **[Z être fait par Y] d'une façon qui tend à provoquer de l'A.** admirablement [Il s'en est sorti admirablement.] **[Z] en provoquant l'A. de X** à [l'~ de N_X] **Intense** beaucoup [d'~], grande, vive | antépos, profonde < folle | antépos, immense < absolue | postpos, «la plus complète», «sans bornes», «sans réserve» //adoration **a**; indéfectible **Exagérée** aveugle, béate; immodérée | postpos **A. est très intense** «ne pas connaître de bornes» **A. devient plus intense** grandir **Éprouvée par beaucoup de X** générale, universelle | postpos **[X] éprouver de l'A. pour Y** avoir, éprouver, ressentir [ART ~ envers/pour N_Y], **litt** concevoir [ART ~ pour N_Y], vouer [ART ~ à N_Y], être [en ~ devant N_Y]; nourrir [ART ~ envers/pour N_Y] **[X] éprouver de l'A. pour Z de Y** avoir, **litt** concevoir, éprouver, ressentir [ART ~ pour N_Z], être [en ~ devant N_Z] **[X] éprouver une intense A. pour Y/Z** se pâmer [d'~ devant N_{Y/Z}] **[X] commencer à éprouver de l'A.** tomber [en ~ devant N_{Y/Z}]; être soulevé, être frappé [de (ART) ~] **[Y] être l'objet de l'A. de X** faire [l'~ de N_X], «être l'objet» [de art ~] **[Z] être l'objet de l'A. de X** forcer, inspirer [l'~], faire [l'~ de N_X], «être l'objet» [de art ~] **[Y] commencer à être l'objet de l'A. de X** s'attirer, emporter [l'~ de N_X avec/par N_Z <en Vpart-Z>] **[Z] commencer à être l'objet d'une A.** provoquer, soulever, susciter [ART ~] < frapper, ravir, saisir, soulever [N_X d'~] **A. a pour objet Z** se porter [sur N_Z][Son admiration se porte plus particulièrement sur les toiles de Chagall.] **[X] manifester son A.** montrer [ART ~]; témoigner [de l'~ à N_Y]/[A_{poss-X} ~ à N_Y] **Manifestation par X de son A.** témoignage II.3 [de (ART) ~] **Sensation qui est une manifestation de l'A.** frisson **I** [d'~] [Elle ne put réprimer un frisson d'admiration. Un frisson d'admiration parcourut l'assistance.

Nous avons tous une admiration sans bornes pour ton talent. Son dernier film s'est attiré l'admiration de la critique. Le jeune comédien a emporté l'admiration du public. Les nombreux échanges que j'eus avec elle me firent concevoir une profonde admiration pour cette femme intelligente, épanouie et généreuse.

Annexe 2 – Article *programme* du DiCoInfo

▼ **programme** ₁, n. m.
Fiche co

Définition
Structure actancielle
i

Texte organisé sous forme d'instructions qu'un **informaticien** écrit dans un **langage** compréhensible pour un ordinateur et qui sert à faire exécuter des **tâches**.

Synonyme(s)
Voir aussi
i

programme informatique

Contextes
i

*Il existe des **programmes** traducteurs (compilateurs) de ces langages qui font partie du système d'exploitation et qui transforment les instructions* [Source : fr [MSPCDOS1](#)]

*Un système d'exploitation est donc un groupe de **programmes** qui gèrent les moyens matériels (l'ordinateur et les périphériques) et logiciels (vos **programmes**) et les mettent en interface avec l'utilisateur.* [Source : fr [DELAVI1](#)]

*Puisque votre **programme** d'impression est rédigé entièrement en " codage maison ".* [Source : fr [IMPRI](#)]

Famille de mots
Sortes de
Autres liens
i

| Explication | Terme relié |
|------------------------------------|---|
| L' informaticien crée un p. | programmer ₁ |
| | programmation ₁ |
| Division | sous-programme ₁ |

Créer ou supprimer
Fonctionner ou faire ce qu'on attend de
Mettre quelque part
i

Transformer / Augmenter / Réduire

Utiliser / Ne pas utiliser
Tout

▼ Préparer l'utilisation / Le fonctionnement

| Explication | Terme relié |
|---|---|
| Qqn ou qqch. prépare un p. afin de le faire fonctionner | appeler ₁ _un ~ charger ₁ _un ~ |
| | chargement ₁ _d'un ~ appel ₁ _d'un ~ |
| Qqn ou qqch. prépare un p. afin de le faire fonctionner | appeler ₁ _un ~ charger ₁ _un ~ |

► Utiliser / Faire fonctionner

► Cesser d'utiliser / De faire fonctionner

anglais : [program](#) ₁
espagnol : [programa](#) ₁

Annexe 3 – Liste de références de notre corpus

| Code | Référence complète | Thème abordé | Genre textuel | Nombre de caractères | Nombre de mots |
|------|---|--------------|---------------|----------------------|----------------|
| D1 | Sun, R. (2006, 20 avril). <i>互联网 [Internet]</i> . 科普中国 [China Science Communication]. https://cloud.kepuchina.cn/newSearch/imgText?id=6792974618812325888 | T5 | G1 | 19243 | 9726 |
| D2 | Wan, Z. (2017). 信息化时代下计算机网络安全研究 [Research on Computer Network Security in the Era of Informatization]. <i>现代职业教育 [Modern Vocational Education]</i> , (16), 56. | T2 | G3 | 1870 | 1064 |
| D3 | 华为技术有限公司 [Huawei Technologies Co., Ltd.]. (2022, 11 janvier). <i>创建云硬盘备份 [Creating a Cloud Disk Backup]</i> . 华为云 [Huawei Cloud]. https://support.huaweicloud.com/qs-cbr/cbr_02_0007.html | T4 | G5 | 775 | 444 |
| D4 | 警惕 8 种不同类型的恶意软件 [Protecting against eight different types of malicious software]. (2020, 8 septembre). <i>安全内参 [Information Security RSS]</i> . https://www.secrss.com/articles/25384 | T2 | G3 | 3206 | 1807 |
| D5 | Zhang, Y. et Huang, Y. (2018). 基于机器学习的恶意软件分类识别研究 [Research on Malware Classification and Identification Based on Machine Learning]. <i>科技资讯 [Science & Technology Information]</i> , (30), 15-17. | T2 | G2 | 1093 | 599 |
| D6 | Tang, X., Liang, H., Zhe, F. et Tang, Z. (2014). 第十二章：保护与安全 [Chapitre 12 : Protection et sécurité]. Dans <i>计算机操作系统 [Systèmes d'exploitation des ordinateurs]</i> (4e éd., p. 380-405). 西安电子科技大学出版社 [Xidian University Press]. | T1 | G1 | 14945 | 8659 |
| D7 | Zhang, B. (2006). 操作系统的概念、功能和用户接口 [Concept, fonctions et interfaces utilisateur des systèmes d'exploitation]. Dans <i>计算机操作系统教程 [Introduction aux systèmes d'exploitation des ordinateurs]</i> (p. 6-13). 清华大学出版社 [Tsinghua University Press]. | T1 | G1 | 4937 | 2787 |
| D8 | 数据处理基础知识 [Notions de base sur le traitement des données]. (2021, 24 juillet). <i>腾讯云 [Cloud Tencent]</i> . https://cloud.tencent.com/developer/inventory/287/article/1089096 | T4 | G1 | 5995 | 3249 |
| D9 | 网易公司 [NetEase, Inc.]. (2021, juillet). <i>渗透测试服务 [Service de tests d'intrusion]</i> . 网易易盾安全解决方案 [NetEase Yidun-One-Stop Security Solution]. https://dun.163.com/product/penetration-test#test | T2 | G4 | 480 | 244 |

| | | | | | |
|-----|--|----|----|-------|-------|
| D10 | Kang, H., Wei, D. et Mei, F. (2005). 第九章：网络安全与网络管理 [Chapitre 9 : Sécurité des réseaux et gestion des réseaux]. Dans <i>计算机网络基础教程 [Introduction aux réseaux informatiques]</i> (p. 295-332). 清华大学出版社 [Tsinghua University Press]. | T5 | G1 | 28340 | 15212 |
| D11 | Wang, D. (2013). 计算机网络的基本组成 [Composants de base d'un réseau informatique]. Dans <i>深入理解计算机网络 [Computer Network Technologies in Depth]</i> (p. 32-36). 机械工业出版社 [China Machine Press]. https://yuedu.163.com/source/f0c2cac3133c410f8ef4c8ce6e264b2c4 | T5 | G1 | 4086 | 2238 |
| D12 | Li, L. et Li, J. (2018). 软件的定义及分类方法 [Définition et typologie des logiciels]. Dans <i>软件测试实用技术与常用模板 [Practical Technology and Common Template for Software Testing]</i> (2e éd., p. 8-12). 机械工业出版社 [China Machine Press]. https://yuedu.163.com/source/f0c2cac3133c410f8ef4c8ce6e264b2c4 | T3 | G1 | 1688 | 963 |
| D13 | Yan, Q. (2002). 当前各种软件类型和模式概述. <i>Internet: 共创软件 [Internet: Software Co-creation]</i> , (2), 57 58. | T3 | G3 | 1513 | 802 |
| D14 | 百度公司 [Baidu, Inc.]. (2021, 31 décembre). <i>云数据库 RDS 产品文档 [Cloud Database RDS Product Documentation]</i> . 百度智能云 [Baidu AI Cloud]. https://bce-cdn.bj.bcebos.com/p3m/pdf/bce-doc/online/RDS/RDS.pdf?timeStamp=1648344760356 | T4 | G4 | 2280 | 1125 |
| D15 | 华为技术有限公司 [Huawei Technologies Co., Ltd.]. (2021). <i>深度 Linux 系统如何更新和升级系统 [How to Update and Upgrade Deepin Linux System]</i> . 华为消费者业务官网 [Huawei Consumer Business Official Website]. https://consumer.huawei.com/cn/support/content/zh-cn00759066/ | T3 | G5 | 676 | 379 |
| D16 | 华为技术有限公司 [Huawei Technologies Co., Ltd.]. (2020, 10 juillet). <i>华为手机备份的三种方式 [Back up Your Phone in Three Ways]</i> . 华为消费者业务官网 [Huawei Consumer Business Official Website]. https://consumer.huawei.com/honor/cn/support/how-to/detail-troubleshooting/?resourceId=zh-cn11911018 | T3 | G5 | 3055 | 1629 |
| D17 | 华为技术有限公司 [Huawei Technologies Co., Ltd.]. (2022). <i>HiSuite 华为手机助手 [HiSuite Huawei Mobile Assistant]</i> . 华为消费者业务官网 [Huawei Consumer Business Official Website]. https://consumer.huawei.com/honor/cn/support/how-to/detail-troubleshooting/?resourceId=zh-cn11911018 | T3 | G5 | 837 | 425 |
| D18 | 华为技术有限公司 [Huawei Technologies Co., Ltd.]. (2022). <i>华为帐号 [Huawei Account ID]</i> . 华为消费者业务官网 [Huawei Consumer Business Official Website]. https://consumer.huawei.com/cn/mobileservices/id/ | T3 | G5 | 3380 | 1932 |

| | | | | | |
|-----|--|----|----|-------|------|
| D19 | Zoom 视频通讯有限公司 [Zoom Video Communications, Inc.]. (2022, 30 janvier). <i>Zoom 新用户入门指南 [Getting started guide for new users]</i> . Zoom 支持 [Zoom Support]. https://support.zoom.us/hc/zh-cn/articles/360034967471 | T3 | G5 | 5636 | 2977 |
| D20 | Teng, L. (2008). 浅谈软件测试的几种类型 [Types of Software Testing]. <i>科教文汇 [Journal of Science and Education]</i> , (12), 269. | T3 | G2 | 2057 | 1169 |
| D21 | Hu, W., Zhao, Z., Han, Y. et Chu, H. (2019). 信息化时代计算机网络安全防护的技术研究 [Research on Computer Network Security Prevention Technology in the Information Age]. <i>中国管理信息化 [China Management Informationization]</i> , 22(1), 170-172. | T2 | G2 | 3721 | 2099 |
| D22 | Microsoft Corporation. (2011). Microsoft 微软安全研究报告第 10 卷 (2010 年 7 月到 12 月) [Microsoft Security Intelligence Report. Volume 10 (July to December 2010)]. https://www.microsoft.com/zh-cn/download/details.aspx?id=17030 | T2 | G3 | 16221 | 7698 |
| D23 | Shao, S., Gao, Q., Ma, S., Duan, F., Ma, X., Zhang, S. et Hu, J. (2018). 缓冲区溢出漏洞分析技术研究进展 [Progress in Research on Buffer Overflow Vulnerability Analysis Technologies]. <i>软件学报 [Journal of Software]</i> , 29(5), 1179-1198. | T2 | G3 | 5830 | 3159 |
| D24 | 开发移动应用程序的三种选择 [Three Options for Developing Mobile Applications]. (2017, 27 décembre). <i>CSDN 专业开发者社区 [CSDN Professional Developer Community]</i> . https://blog.csdn.net/laizixingxingdewo/article/details/78913067 | T3 | G3 | 4985 | 2756 |
| D25 | 操作系统简介 [Introduction of Operating System]. (2020, 12 septembre). <i>CSDN 专业开发者社区 [CSDN Professional Developer Community]</i> . https://blog.csdn.net/Jackson_mvp/article/details/108154591 | T1 | G3 | 14921 | 7584 |
| D26 | ABBY Production LLC. (2013). ABBY FineReader 12 用户指南 [Getting started guide for new users of ABBY FineReader 12]. https://www.abby.com/media/6646/guide_chinesesimplified.pdf | T3 | G5 | 11679 | 5864 |
| D27 | Kang, H., Wei, D. et Mei, F. (2005). 第七章：应用层 [Chapitre 7 : Couche d'application du réseau informatique]. Dans <i>计算机网络基础教程 [Introduction aux réseaux informatiques]</i> (p. 208 220). 清华大学出版社 [Tsinghua University Press]. | T5 | G1 | 13101 | 7179 |
| D28 | 挣脱数据沼泽 [Get out of the data swamp]. (2021, 19 mars). <i>CSDN 专业开发者社区 [CSDN Professional Developer Community]</i> . https://blog.csdn.net/jdcdev/article/details/115005147 | T4 | G3 | 3452 | 1777 |

| | | | | | |
|-----|---|----|----|-------|------|
| D29 | Jingdong Zhilian Cloud Database. (2020, 20 octobre). 京东云新一代分布式数据库 TiDB 架构揭秘 [The architecture of TiDB, a new generation of distributed database in Jingdong Cloud]. <i>CSDN 专业开发者社区 [CSDN Professional Developer Community]</i> . https://blog.csdn.net/jdcdev/article/details/109193079 | T4 | G4 | 2651 | 1400 |
| D30 | Kang, H., Wei, D. et Mei, F. (2005). 4.8 网络操作系统 [4.8 : Système d'exploitation réseau]. Dans <i>计算机网络基础教程 [Introduction aux réseaux informatiques]</i> (p. 135-148). 清华大学出版社 [Tsinghua University Press]. | T1 | G1 | 14743 | 7275 |
| D31 | 人民政协网 [The Chinese People's Political Consultative Conference (CPPCC)]. (2020, 15 octobre). <i>全国移动 App 第三季度安全研究报告 [National Report on Mobile App Security Research (Third Quarter)]</i> . 人民政协网 [The Chinese People's Political Consultative Conference (CPPCC)]. http://www.rmzxb.com.cn/c/2020-10-15/2687444.shtml | T2 | G3 | 4038 | 1979 |
| D32 | Alibaba JAQ Security. (2015, 15 novembre). <i>2015 阿里移动安全第三季度报告 [Alibaba JAQ 2015 Third Quarter Mobile Security Report]</i> . 阿里云开发者社区 [Aliyun Developer Community]. https://developer.aliyun.com/article/86120 | T2 | G3 | 6235 | 3049 |
| D33 | 恒安嘉新（北京）科技股份有限公司 [Eversec (Beijing) Technology Co., Ltd.]. (2021). <i>2020 年移动互联网安全态势报告 [2020 Mobile Internet Security Trend Report]</i> . EversecTechInc. https://mp.weixin.qq.com/s?__biz=MzIwMTUzNDY0NA==&mid=2649114324&idx=4&sn=156428980db924ff149926ffca53484e&chksm=8efe9b0fb989121998339e6c5c5a86afe63f12862f3eec6e12146eb008f34ecc3af7609fc0ff&scene=21#wechat_redirect | T2 | G3 | 3958 | 2079 |
| D34 | Fang, B. (2018). 定义网络空间安全 [Define Cyberspace Security]. <i>网络与信息安全学报 [Chinese Journal of Network and Information Security]</i> , 4(1), 1-5. | T2 | G2 | 4728 | 2466 |
| D35 | Wei, K., Weng, J. et Ren, K. (2016). 大数据安全保护技术综述 [Data Security and Protection Techniques in Big Data]. <i>网络与信息安全学报 [Chinese Journal of Network and Information Security]</i> , 2(4), 1-11. | T4 | G2 | 9034 | 5369 |
| D36 | Li, B. (2020, 14 juillet). <i>中国移动互联网发展报告(2020)正式发布 [Official Release of China Mobile Internet Development Report (2020)]</i> . 人民网 [people.com.cn]. http://media.people.com.cn/n1/2020/0714/c433219-31783080.html | T5 | G3 | 2886 | 1475 |
| D37 | 国家计算机网络应急技术处理协调中心 [The National Computer Network Emergency Response Technical Team/Coordination Center of China]. (2017). <i>2016 年中国互联网网络安全报告 [2016 China Internet Security Report]</i> (人民邮电出版社 [The People's Posts and Telecommunications Press]). | T2 | G3 | 16487 | 8680 |

| | | | | | |
|-----|---|----|----|------|------|
| D38 | Zheng, B. (2005). <i>Windows 平台下网络直播系统的实现 [Implementation of Webcasting System under Windows platform]</i> [master's thesis]. 华东师范大学 [East China Normal University]. | T5 | G3 | 2384 | 1261 |
| D39 | Amazon Web Services (AWS). (2022). <i>缓存概览 [Caching Overview]</i> . 亚马逊网络服务 [Amazon Web Services]. https://aws.amazon.com/cn/caching/ | T5 | G4 | 5997 | 3260 |
| D40 | 华为技术有限公司 [Huawei Technologies Co., Ltd.]. (2020, 3 janvier). <i>云计算 [Cloud Computing]</i> . 华为云官网 [Huaweicloud Official Website]. https://www.huaweicloud.com/zhishi/191128.html | T5 | G4 | 581 | 1808 |
| D41 | 华为技术有限公司 [Huawei Technologies Co., Ltd.]. (2021, 6 août). <i>数据安全 [Data Security]</i> . 华为云官网 [Huaweicloud Official Website]. https://www.huaweicloud.com/zhishi/edit-582200.html | T2 | G4 | 2561 | 323 |
| D42 | 华为技术有限公司 [Huawei Technologies Co., Ltd.]. (2021, 6 août). <i>桌面管理系统 [Desktop Management System]</i> . 华为云官网 [Huaweicloud Official Website]. https://www.huaweicloud.com/zhishi/edit-582211.html | T3 | G4 | 851 | 1448 |
| D43 | 华为技术有限公司 [Huawei Technologies Co., Ltd.]. (2021, 6 août). <i>漏洞扫描系统 [Vulnerability Scanning System]</i> . 华为云官网 [Huaweicloud Official Website]. https://www.huaweicloud.com/zhishi/edit-582198.html | T2 | G4 | 2041 | 472 |
| D44 | 华为技术有限公司 [Huawei Technologies Co., Ltd.]. (2022). <i>数据复制服务 [Data Replication Service]</i> . 华为云官网 [Huaweicloud Official Website]. https://www.huaweicloud.com/zhishi/edit-582198.html | T4 | G4 | 2370 | 1281 |
| D45 | 华为技术有限公司 [Huawei Technologies Co., Ltd.]. (2022). <i>视频直播 [Live video streaming]</i> . 华为云官网 [Huaweicloud Official Website]. https://www.huaweicloud.com/solution/Onlinevideolive/ | T5 | G4 | 1986 | 1021 |
| D46 | Asiainfo Security. (2022). <i>亚信发布第三季度网络安全报告 [AsiaInfo Releases Third Quarter Cybersecurity Report]</i> . http://www.yitianzixun.com/yitianzixun/vip_doc/3101129.html?p_m_page=200-5 | T2 | G3 | 1760 | 930 |
| D47 | 沃通电子认证服务有限公司 [WoTrus CA Limited]. (2022). <i>安全电子邮件系统解决方案 [Secure Email System Solution]</i> . 沃通解决方案 [wosign.com]. https://www.wosign.com/solution/securemail.htm | T2 | G4 | 2036 | 1067 |
| D48 | Guo, S., Jia, H. et Xiong, J. (2009). <i>互联网海量数据存储和处理技术综述 [A Review of Big Data Storage and Processing Technologies]</i> . <i>信息技术学报 [Journal of Information Technology]</i> , 7(5), 1-29. | T4 | G2 | 4707 | 2436 |
| D49 | Lin, D., Tian, Y. et Tian, C. (2014). <i>移动安全技术研究综述 [A Review of Mobile Security Technology Research]</i> . <i>保密科学技术 [Secrecy Science and Technology]</i> , (3), 4-25. | T2 | G4 | 1455 | 770 |

| | | | | | |
|-----|--|----|----|-------|-------|
| D50 | 华为技术有限公司 [Huawei Technologies Co., Ltd.]. (2021, 16 juin). 分布式缓存服务 [Distributed Cache Service]. 华为云官网 [Huaweicloud Official Website]. https://support.huaweicloud.com/swp-dds/dds_swp_0001.html | T4 | G3 | 1306 | 711 |
| D51 | 华为技术有限公司 [Huawei Technologies Co., Ltd.]. (2018). 华为云数据安全白皮书 [Huawei Cloud Security White Paper]. https://res-static1.huaweicloud.com/content/dam/cloudbusite/archive/china/zh-cn/securecenter/security_doc/Security_cn_20181008.pdf | T4 | G3 | 6504 | 3656 |
| D52 | 猎鹰安全科技有限公司 [Falcon Security]. (2022). 安全应用场景 [Security Application Scenarios]. 猎鹰安全 [Falcon Security]. https://www.ejinshan.net/lywz/emrlnet#cpjz | T2 | G4 | 9473 | 5330 |
| D53 | Gu, Q., Cai, Z. et Jia, C. (2003). 第三章：网络通信安全 [Network Communication Security]. Dans 计算机网络安全 [Computer Network Security] (p. 44-65). 科学出版社 [Scientific Press]. | T2 | G1 | 29000 | 15787 |
| D54 | 艾瑞咨询 [iResearch Consulting Group]. (2012). 移动应用产业发展洞察：多样化趋势中的机遇 [Insight into Mobile App: Opportunities in the Trend of Diversification]. https://www.microsoft.com/zh-cn/download/details.aspx?id=17030 | T3 | G3 | 3167 | 1764 |
| D55 | Li, S. et Wang, W. (2017). Office 2010 的安装与基本操作 [Installation and Use of Office 2010]. Dans Office 2010 办公软件实用教程 (p. 1-16). 清华大学出版社 [Tsinghua University Press]. | T3 | G1 | 4657 | 2140 |
| D56 | 科极网拓信息技术有限公司 [TechTarget China]. (2021). 电子邮件安全手册 [Email Security Handbook]. TechTarget. https://searchsecurity.techtarget.com.cn/guide/11-24905/ | T2 | G3 | 2454 | 1378 |
| D57 | 科极网拓信息技术有限公司 [TechTarget China]. (2021). Web 应用程序安全手册 [Web Application Security Handbook]. TechTarget. https://searchsecurity.techtarget.com.cn/guide/11-24901/ | T2 | G3 | 8712 | 4894 |
| D58 | 深圳爱加密科技有限公司 [Shenzhen Ijiami Technology Co., Ltd.]. (2022). iOS 移动应用安全加固 [IOS Mobile Application Security Enhancement]. Ijiami Official Site. https://www.ijiami.cn/iosProtect | T2 | G4 | 1872 | 959 |
| D59 | 中国科学技术协会. China Association for Science and Technology. (2022). 流媒体 [Streaming media]. 科普中国 [China Science Communication]. https://baike.baidu.com/item/%E6%B5%81%E5%AA%92%E4%BD%93/98740?fromtitle=%E7%BD%91%E7%BB%9C%E6%B5%81%E5%AA%92%E4%BD%93&fromid=4954359 | T5 | G1 | 4149 | 2125 |
| D60 | 堆栈科技有限公司 [Stack Technology Co. Ltd.]. (2022). 计算机网络 [Computer Network]. 思否中文技术交流平台 [SegmentFault technical communication platform]. https://segmentfault.com/a/1190000040405731 | T5 | G4 | 4431 | 2156 |

| | | | | | |
|-----|---|----|----|-------|-------|
| D61 | Dong, J. et Ding, B. (2007). 数据库系统结构 [Database System Architecture]. Dans <i>数据库实用教程 [Practical Database Tutorial]</i> (3e éd., p. 18-30). 清华大学出版社 [Tsinghua University Press]. | T4 | G1 | 3053 | 1682 |
| D62 | 火山引擎 [Volcengine Inc.]. (2021, 14 décembre). <i>访问密钥 [Access Key]</i> . 火山引擎官方网站 [Volcengine Official Site]. https://www.volcengine.com/docs/6291/65568 | T5 | G4 | 2248 | 1153 |
| D63 | Wu, J., Li, W. et Huang, J. (2015). 移动互联网研究综述 [A Review of Mobile Internet Research]. <i>中国科学: 信息科学 [SCIENCE CHINA Information Sciences]</i> , (45), 45-69. | T5 | G2 | 18325 | 10162 |
| D64 | Fu, Y. et Wei, J. (2012). 2011 年移动互联网研究综述 [A Review of Mobile Internet Research]. Dans J. Guang et S. Tang (dir.), <i>中国移动互联网发展报告 (2012) [Annual Development Report on China's Mobile Internet (2012)]</i> (社会科学文献出版社 [Social Sciences Academic Press (China)], p. 35-51). | T5 | G2 | 8992 | 4636 |
| D65 | Xiao, Q., Wang, J. et Zhu, Y. (2017). 物联网智能终端设备识别方法 [Intelligent Terminal Device Identification Method of Internet of Things]. <i>电信科学 [Telecommunications Science]</i> , 33(2), 3-8. | T5 | G2 | 3260 | 1736 |
| D66 | Wang, S., Wang, L., Jin, H. et Shen, J. (2015). 网络安全分析中的大数据技术应用 [Big Data Application in Network Security Analysis]. <i>电信科学 [Telecommunications Science]</i> , 31(7), 139-144. https://doi.org/10.11959/j.issn.1000-0801.2015176 | T4 | G2 | 7005 | 3539 |
| D67 | 华为技术有限公司 [Huawei Technologies Co., Ltd.]. (2021, 6 août). <i>防病毒软件 [Anti-virus software]</i> . 华为云官网 [Huaweicloud Official Website]. https://www.huaweicloud.com/zhishi/edit-582220.html | T3 | G4 | 1125 | 599 |
| D68 | 华为技术有限公司 [Huawei Technologies Co., Ltd.]. (2022, 30 août). <i>云防火墙 [Cloud Firewall]</i> . 华为云官网 [Huaweicloud Official Website]. https://support.huaweicloud.com/productdesc-cfw/cfw_01_0002.html | T4 | G4 | 1635 | 894 |
| D69 | Jiang, H., Ren, Z. et Nie, L. (2014). 面向移动应用大数据的软件工程研究 [Software Engineering Research for Big Data of Mobile Applications]. <i>中国计算机学会通讯 [Communications of China Computer Federation]</i> , 10(3), 24-30. | T3 | G2 | 4456 | 2547 |
| D70 | IBM 中国 [IBM China]. (2021, 17 mars). <i>移动应用程序开发简介 [Introduction to Mobile Application Development]</i> . IBM 中国 [IBM China]. https://www.ibm.com/docs/zh/mpf/7.1.0?topic=overview-introduction-mobile-application-development | T3 | G4 | 2125 | 1059 |

| | | | | | |
|-----|---|----|----|-------|------|
| D71 | Xu, Y., Ma, Z., Wang, Z., Niu, X. et Yang, Y. (2016). Android 智能终端安全综述 [Survey of Security For Android Smart Terminal]. <i>通信学报 [Journal on Communications]</i> , 37(6), 169-184. https://doi.org/10.11959/j.issn.1000-436x.2016127 | T1 | G2 | 13583 | 7626 |
| D72 | 润雪科技有限公司 [Runxuekeji.Co.Ltd]. (2021, 29 juillet). <i>移动程序开发 [Mobile Application Development]</i> . 润雪科技官方网站 [Runxuekeji Official Site]. https://www.runxuekeji.cn/shoujiappkaifa/id/913.html | T3 | G4 | 2459 | 1410 |
| D73 | 润雪科技有限公司 [Runxuekeji.Co.Ltd]. (2021, 27 juillet). <i>手机 App 开发中的回归测试 [Regression test in Mobile Application Development]</i> . 润雪科技官方网站 [Runxuekeji Official Site]. https://www.runxuekeji.cn/shoujiappkaifa/id/898.html | T3 | G4 | 3109 | 1820 |
| D74 | 上海亦存网络科技有限公司 [Shanghai YiCun Network Tech Co., Ltd]. (2022, 17 décembre). <i>企业网盘云存储工作原理 [How Does Enterprise Cloud Storage Work]</i> . 坚果云官方网站 [Nutstore Official Website]. https://content.jianguoyun.com/30411.html | T4 | G4 | 5921 | 3477 |
| D75 | Yu, H. (2021, 29 juillet). <i>在多台云服务器之间共享数据 [Share Data between Multiple Cloud Servers]</i> . 腾讯云开发者社区 [Tencent Cloud Developer Community]. https://cloud.tencent.com/developer/article/1854531 | T4 | G3 | 1473 | 808 |
| D76 | Tencent Cloud. (2020, 15 mai). <i>自动化测试与手动测试有什么区别? [What is the Difference between Automated Testing and Manual Testing?]</i> . 腾讯云开发者社区 [Tencent Cloud Developer Community]. https://cloud.tencent.com/developer/article/1854531 | T3 | G3 | 2285 | 1261 |
| D77 | Lenovo Group Ltd. (2022, 2 avril). <i>联想 M90-1 用户指南 [Getting started guide for new users of Lenovo ThinkCentre M90n-1]</i> . https://think.lenovo.com.cn/htmls/advancedsearch.aspx?doccatid=1249014520765 | T1 | G5 | 11555 | 5799 |
| D78 | Lenovo Group Ltd. (2022, 2 avril). <i>使用 Vantage 或联想电脑管家应用程序 [Use Vantage or Lenovo PC Manager application]</i> . https://think.lenovo.com.cn/htmls/advancedsearch.aspx?doccatid=1249014520765 | T1 | G5 | 8693 | 4097 |
| D79 | Micron Technology, Inc. (2017). <i>及时更新计算机的重要性 [The Importance of Update Your Computer in a Timely Manner]</i> . Crucial 英睿达官方网站 [Crucial Official Site by Micron]. https://www.crucial.cn/articles/pc-users/keeping-your-computer-up-to-date | T1 | G5 | 8614 | 4693 |
| D80 | <i>计算机系统优化的几种方法 [Several Methods of Computer System Optimization]</i> . (2021, 24 juillet). CSDN 专业开发者社区 [CSDN Professional Developer Community]. https://blog.csdn.net/weixin_29251131/article/details/119073423 | T1 | G3 | 928 | 502 |

| | | | | | |
|-----|---|----|----|-------|------|
| D81 | Microsoft Corporation. (2022). 在 Windows 中提高电脑性能的提示 [Tips for Improving PC Performance in Windows]. Microsoft Official Website. https://support.microsoft.com/zh-cn/windows/%E5%9C%A8-windows-10-%E4%B8%AD%E6%8F%90%E9%AB%98%E7%94%B5%E8%84%91%E6%80%A7%E8%83%BD%E7%9A%84%E6%8F%90%E7%A4%BA-b3b3ef5b-5953-fb6a-2528-4bbed82fba96 | T1 | G5 | 6548 | 3375 |
| D82 | Pan, J. (2008). 华东师范大学讲义-Linux 操作系统 [Lecture Notes of East China Normal University-Linux Operating System]. https://math.ecnu.edu.cn/~jypan/Teaching/Linux/Linux08/lect01.pdf | T1 | G1 | 8961 | 4668 |
| D83 | Pan, J. (2020, 21 avril). 慕课教程-Linux 操作系统 [Imooc Tutorial-Linux Operating System]. 慕课网 [imooc.com]. | T1 | G1 | 6235 | 3113 |
| D84 | Xu, X. (2018, 9 mai). 加密传输 [Encrypted Data Transmission]. 知乎网 [zhihu.com]. https://zhuanlan.zhihu.com/p/36603247 | T2 | G3 | 2214 | 1210 |
| D85 | Sun, Z. (2008). 操作系统概论 [Introduction to Operating Systems]. Dans 操作系统教程 [Operating System Tutorial] (vol. 4, p. 10-38). 高等教育出版社 [Higher Education Press]. | T1 | G1 | 14522 | 7979 |
| D86 | Sun, Z. (2008). 操作系统存储管理 [Operating System Storage Management]. Dans 操作系统教程 [Operating System Tutorial] (vol. 4, p. 40-48). 高等教育出版社 [Higher Education Press]. | T1 | G1 | 4926 | 2760 |
| D87 | 提高电脑运行速度 [Improve Your Computer Performance]. (2020, 4 mars). 知乎网 [zhihu.com]. https://zhuanlan.zhihu.com/p/110976776 | T1 | G3 | 1356 | 752 |
| D88 | Wan, Z. (2018). 基于网站制作的 Web 前端开发技术与优化 [Web front-end Development Technology and Optimization Based on Website Production]. 电脑知识与技术 [Computer Knowledge and Technology], 14(33), 198-199. | T3 | G2 | 3734 | 2085 |
| D89 | 软件开发的四种类型 [Four Types of Software Development]. (2017, 18 novembre). CSDN 专业开发者社区 [CSDN Professional Developer Community]. https://blog.csdn.net/vipers_/article/details/78570158 | T3 | G3 | 2033 | 1044 |
| D90 | APP 软件开发方法的类型 [Types of APP software development methods]. (2021, 14 juillet). 网易 [163.com]. https://www.163.com/dy/article/GESMTT7M0531AJGK.html | T3 | G3 | 798 | 385 |
| D91 | Emerson Electric Company. (2022). 监控信息系统 202 功能设计 [Monitoring Information System (SIS) 202 Function Design]. https://www.emerson.com/documents/automation/training-sis202-zh-cn-41692.pdf | T3 | G1 | 1668 | 946 |

| | | | | | |
|------|--|----|----|-------|------|
| D92 | APICloud. (2019, 9 décembre). 安卓 APP 开发中常见的 API 类型 [Four Types of Software Development]. <i>APICloud 移动低代码开发平台 [APICloud Mobile Low Code Development Platform]</i> . https://www.apicloud.com/blogDetails/2539 | T3 | G3 | 10323 | 4787 |
| D93 | 什么是软件开发 [What is Software Development]. (2019, 26 décembre). 知乎网 [zhihu.com]. https://zhuanlan.zhihu.com/p/99539914 | T3 | G3 | 2309 | 1344 |
| D94 | Salesforce Inc. (2022). 云计算—随时随地可供使用的应用程序 [Cloud Computing - Applications Available Anywhere, Anytime]. Salesforce Official Website. https://www.salesforce.com/cn/cloudcomputing/?d=70130000000YsaZ | T4 | G4 | 3194 | 1734 |
| D95 | Li, J. (2015). 云计算在移动互联网中的应用探析 [Exploring the Application of Cloud Computing in Mobile Internet]. <i>电子技术与软件工程 [Electronic Technology & Software Engineering]</i> , (14), 174. | T4 | G2 | 1950 | 1056 |
| D96 | 前端开发分类 [Types of Front-end Development]. (2018, 26 février). <i>CSDN 专业开发者社区 [CSDN Professional Developer Community]</i> . https://blog.csdn.net/qq_37255910/article/details/79378112?utm_medium=distribute.pc_relevant.none-task-blog-2%7Edefault%7ECTRLIST%7Edefault-3.no_search_link&depth_1-utm_source=distribute.pc_relevant.none-task-blog-2%7Edefault%7ECTRLIST%7Edefault-3.no_search_link | T3 | G3 | 3061 | 1396 |
| D97 | Su, Q. (2021, 17 mars). 专有软件 [Proprietary Software]. <i>MBA 智库百科 [MBALIB.COM]</i> . https://wiki.mbalib.com/wiki/%E4%B8%93%E6%9C%89%E8%BD%AF%E4%BB%B6 | T3 | G1 | 742 | 427 |
| D98 | 华为技术有限公司 [Huawei Technologies Co., Ltd.]. (2022). <i>软件工程 [Software Engineering]</i> . 华为云官网 [Huaweicloud Official Website]. https://www.huaweicloud.com/articles/55e265196c320379fcd0a596b20ba191.html | T3 | G3 | 4215 | 2353 |
| D99 | Fan, Y. et Ling, X. (2013). Android 程序安全检测模型 [Modeling of Android Application Security Testing]. <i>第 28 次全国计算机安全学术交流会论文集 [Proceedings of the 28th National Computer Security Academic Exchange Conference]</i> , 84-86. | T3 | G2 | 4129 | 1911 |
| D100 | Chen, Z., Zeng, F., Chen, G., Zhang, Y. et Li, X. (2019). 物联网安全测评技术综述 [A Survey for IoT Security Assessment Technologies]. <i>信息安全学报 [Journal of Cyber Security]</i> , 4(3), 2-16. | T5 | G2 | 16224 | 8810 |
| D101 | Feng, H. (2016). 云计算在移动互联网上的应用 [Application of Cloud Computing on Mobile Internet]. <i>中文科技期刊数据库: 工程技术 [Chinese Science and Technology Journal Database : Engineering Technology]</i> , (66), 15-16. | T5 | G2 | 2638 | 1450 |

| | | | | | |
|------|---|----|----|-------|------|
| D102 | Gao, L. (2014). 云计算在移动互联网上的应用 [Application of Cloud Computing on Mobile Internet]. <i>计算机光盘软件与应用 [Computer CD, Software and Applications]</i> , (7), 99-101. | T5 | G3 | 2958 | 1699 |
| D103 | Shen, W., Zhang, C. et Zeng, Z. (2012). 软件工程概述 [Overview of Software Engineering]. Dans <i>软件工程基础与实用教程: 基于架构与MVC模式的一体化开发 [Software Engineering Fundamentals and Practical Tutorial: Integrated Development Based on Architecture and MVC Patterns]</i> (p. 3-26). 清华大学出版社 [Tsinghua University Press]. | T3 | G1 | 15473 | 7917 |
| D104 | Liu, J. (2012). 软件工程概述 [Overview of Software Engineering]. Dans <i>实用软件工程教程 [Practical Software Engineering]</i> (p. 7-29). 高等教育出版社 [Higher Education Press]. | T4 | G1 | 11269 | 6238 |
| D105 | <i>软件开发项目类型 [Software Development Project Types]</i> . (2021). Limnor-Studio Official Website. http://www.limnor.com/support/ProjectTypes_CN.pdf | T3 | G5 | 979 | 536 |
| D106 | 清华大学软件学院 [School of Software, Tsinghua University]. (2021). 清华大学讲义-软件工程 [Lecture Notes of Tsinghua University-Software Engineering]. | T3 | G1 | 2496 | 1395 |
| D107 | Sun, Z. (2008). 信息系统的安全模型 [Information System Security Model]. Dans <i>操作系统教程 [Operating System Tutorial]</i> (vol. 4, p. 411-425). 高等教育出版社 [Higher Education Press]. | T2 | G1 | 15507 | 9333 |
| D108 | Sun, Z. (2008). 安全操作系统的设计 [Design of secure operating system]. Dans <i>操作系统教程 [Operating System Tutorial]</i> (vol. 4, p. 437-461). 高等教育出版社 [Higher Education Press]. | T1 | G1 | 15376 | 8792 |
| D109 | Zhang, C. et Wang, J. (2018). 云计算技术及应用 [Cloud Computing Technology and Applications]. <i>信息与电脑 [China Computer & Communication]</i> , (7), 169-171. | T4 | G2 | 2699 | 1565 |
| D110 | Liu, J. (2013). 云计算应用模式下的移动互联网安全问题 [Mobile Internet Security in the Cloud Computing Application Model]. <i>硅谷 [Silicon Valley]</i> , (16), 138-141. | T5 | G2 | 2233 | 1238 |
| D111 | 迅捷通信技术有限公司 [Fast Communication Technology Co. Ltd.]. (2020). 无线安全—您的无线网络被非法接入了吗? [Wireless Security-Is Your Wireless Network Being Accessed Illegally?]. 迅捷网络官方网站 [Fast Official Website]. https://service.fastcom.com.cn/article-153.html | T5 | G4 | 1261 | 680 |
| D112 | 百度公司 [Baidu, Inc.]. (2022). 网络营销 [Online Marketing]. 百度百科 [Baidu Baike]. https://baike.baidu.com/item/%E7%BD%91%E7%BB%9C%E8%90%A5%E9%94%80/175416 | T5 | G1 | 3126 | 1647 |
| D113 | Cloudflare 公司 [Cloudflare Inc.]. (2022). 什么是远程访问策略? [What is Remote Access Policy?]. Cloudflare 官方网站 [Cloudflare Official Website]. https://www.cloudflare.com/zh-cn/learning/access-management/gdpr-remote-access/ | T3 | G1 | 2374 | 1256 |

Annexe 4 – Script Python pour nettoyer un texte de notre corpus

```
Python 3.9.6 (v3.9.6:db3ff76da1, Jun 28 2021, 11:14:58)
[Clang 12.0.5 (clang-1205.0.22.9)] on darwin
Type "help", "copyright", "credits" or "license()" for more information.
>>> import re
>>> with open('/Users/ZhiweiHan/Desktop/D1.txt') as f:
    result=f.read()
    r="[_.!+-=—,$%^, ° ? ~#@#¥%.....&* 《》 <> 「」 {} 【】 ()/]"
    result=re.sub(r, ' ', result)
    result1=result.encode('utf-8')
    with open('/Users/ZhiweiHan/Desktop/D1nett.txt', 'wb') as f2:
        f2.write(result1)
        f.close()
        f2.close()

55153
>>>
>>> import os
>>> import sys
>>> with open('/Users/ZhiweiHan/Desktop/D1nett.txt') as f:
    document=f.read()
    document2=document.replace(" ", "")
    with open ('/Users/ZhiweiHan/Desktop/D1nett.txt', 'wb') as f2:
        document3=document2.encode('utf-8')
        f2.write(document3)
        f.close()
        f2.close()

52582
>>> |
```


Annexe 5 – Liste de CT extraits de notre corpus

Légende

CT : candidat-terme extrait à partir du corpus

Frequency (focus) : nombre d’occurrences d’un CT dans le corpus spécialisé

Frequency (reference) : nombre d’occurrences d’un CT dans le corpus de référence

DOCF (focus) : nombre de documents contenant un CT dans le corpus spécialisé

DOCF (reference) : nombre de documents contenant un CT dans le corpus de référence

UT : unité terminologique (candidat-terme doté d’au moins un sens spécialisé)

CT rejeté : candidat-terme non terminologique

| | CT | Frequency (focus) | Frequency (reference) | DOCF (focus) | DOCF (reference) | Score de spécificité | Statut terminologique |
|----|---------|-------------------|-----------------------|--------------|------------------|----------------------|-----------------------|
| 1 | 程序 | 2174 | 2288748 | 90 | 1128705 | 33,373 | UT |
| 2 | 加密 | 704 | 209574 | 45 | 90302 | 32,968 | UT |
| 3 | 漏洞 | 694 | 300041 | 38 | 160695 | 29,908 | UT |
| 4 | 恶意 | 611 | 246845 | 44 | 152350 | 27,724 | UT |
| 5 | 计算机 | 1254 | 1361171 | 69 | 593697 | 27,561 | UT |
| 6 | 存储 | 924 | 793947 | 74 | 327993 | 27,537 | UT |
| 7 | 访问 | 977 | 897234 | 79 | 554640 | 27,348 | UT |
| 8 | 软件 | 2584 | 3949558 | 93 | 1432705 | 25,85 | UT |
| 9 | 攻击 | 985 | 1113755 | 48 | 549486 | 24,496 | UT |
| 10 | Windows | 502 | 182 894 | 40 | 87 041 | 24,363 | CT rejeté |
| 11 | Web | 425 | 92 747 | 34 | 40 630 | 22,78 | CT rejeté |
| 12 | 服务器 | 736 | 774619 | 68 | 252853 | 22,304 | UT |
| 13 | 邮件 | 669 | 664448 | 40 | 327776 | 21,819 | UT |
| 14 | 移动 | 1390 | 2681982 | 66 | 1125443 | 19,034 | UT |
| 15 | 应用 | 2467 | 5400702 | 103 | 2261364 | 18,937 | UT |
| 16 | 用户 | 2710 | 6038557 | 102 | 1959216 | 18,859 | UT |
| 17 | 传输 | 520 | 540411 | 59 | 259131 | 18,63 | UT |
| 18 | 文件 | 1307 | 2692119 | 83 | 1158985 | 17,86 | UT |
| 19 | 数据库 | 516 | 623774 | 54 | 274953 | 17,43 | UT |
| 20 | 密钥 | 301 | 45768 | 20 | 15245 | 17,276 | UT |

| | | | | | | | |
|----|-------|------|-----------|-----|---------|--------|-----------|
| 21 | 权限 | 385 | 280495 | 44 | 167051 | 17,216 | UT |
| 22 | 数据 | 2980 | 7555719 | 99 | 2784988 | 16,976 | UT |
| 23 | 密码 | 489 | 598564 | 45 | 251109 | 16,84 | UT |
| 24 | 终端 | 543 | 796312 | 43 | 400736 | 16,369 | UT |
| 25 | 备份 | 330 | 203951 | 33 | 81103 | 15,964 | UT |
| 26 | 网络 | 2653 | 7270623 | 91 | 2807651 | 15,656 | UT |
| 27 | 操作 | 1526 | 3987107 | 96 | 2167377 | 15,217 | UT |
| 28 | 病毒 | 484 | 858529 | 43 | 285721 | 14,106 | UT |
| 29 | 代码 | 386 | 537432 | 58 | 253409 | 14,015 | UT |
| 30 | 系统 | 3752 | 11943456 | 106 | 4147982 | 14,014 | UT |
| 31 | 攻击者 | 232 | 30483 | 28 | 14036 | 13,773 | UT |
| 32 | 云计算 | 257 | 166182 | 25 | 68713 | 13,089 | UT |
| 33 | Linux | 229 | 84 459 | 24 | 34 442 | 12,804 | CT rejeté |
| 34 | 缓存 | 226 | 89649 | 24 | 43827 | 12,577 | UT |
| 35 | 文档 | 314 | 438915 | 42 | 205642 | 12,409 | UT |
| 36 | 安全 | 2743 | 9876095 | 88 | 3371496 | 12,245 | UT |
| 37 | 安全性 | 305 | 427519 | 53 | 310540 | 12,181 | UT |
| 38 | 通信 | 457 | 1071221 | 55 | 440835 | 11,853 | UT |
| 39 | 协议 | 607 | 1716704 | 47 | 874019 | 11,646 | UT |
| 40 | 进程 | 342 | 649958 | 34 | 485561 | 11,537 | UT |
| 41 | 算法 | 288 | 431191 | 32 | 160408 | 11,505 | UT |
| 42 | 运行 | 772 | 2476051 | 84 | 1364070 | 11,341 | UT |
| 43 | 云 | 567 | 1 634 431 | 43 | 660 884 | 11,264 | UT |
| 44 | 连接 | 467 | 1266106 | 67 | 693308 | 10,978 | UT |
| 45 | 主机 | 255 | 357905 | 33 | 157994 | 10,896 | UT |
| 46 | 检测 | 649 | 2076200 | 46 | 861851 | 10,892 | UT |
| 47 | 远程 | 295 | 535263 | 39 | 273096 | 10,872 | UT |
| 48 | 硬件 | 353 | 799894 | 61 | 463240 | 10,796 | UT |
| 49 | 互联网 | 997 | 3635795 | 55 | 1385652 | 10,789 | UT |
| 50 | 测试 | 667 | 2189159 | 38 | 888348 | 10,767 | UT |
| 51 | 开发 | 1180 | 4683355 | 69 | 2234224 | 10,315 | UT |
| 52 | 删除 | 238 | 380527 | 44 | 216170 | 10,025 | UT |
| 53 | 内存 | 255 | 461921 | 31 | 195523 | 10,018 | UT |
| 54 | 威胁 | 305 | 714857 | 40 | 509867 | 9,915 | UT |
| 55 | 联网 | 280 | 604593 | 24 | 248381 | 9,849 | UT |
| 56 | 验证 | 268 | 563714 | 42 | 359101 | 9,729 | UT |
| 57 | 组件 | 194 | 242321 | 31 | 116331 | 9,368 | UT |
| 58 | 发送 | 246 | 506567 | 41 | 321142 | 9,364 | UT |
| 59 | 黑客 | 181 | 184319 | 27 | 78323 | 9,295 | UT |

| | | | | | | | |
|----|----------|------|----------|-----|---------|-------|-----------|
| 60 | 基于 | 440 | 1551898 | 67 | 937776 | 9,122 | UT |
| 61 | 选项 | 215 | 383899 | 24 | 186450 | 9,097 | UT |
| 62 | 节点 | 215 | 384867 | 23 | 226072 | 9,09 | UT |
| 63 | 请求 | 225 | 461845 | 43 | 274161 | 8,916 | UT |
| 64 | 登录 | 263 | 668610 | 39 | 409492 | 8,89 | UT |
| 65 | Internet | 145 | 44 418 | 30 | 25 334 | 8,827 | CT rejeté |
| 66 | 界面 | 249 | 629476 | 48 | 340289 | 8,673 | UT |
| 67 | 入侵 | 169 | 206961 | 32 | 134475 | 8,542 | UT |
| 68 | 浏览器 | 188 | 315041 | 36 | 123016 | 8,524 | UT |
| 69 | 接入 | 184 | 307524 | 28 | 166512 | 8,413 | UT |
| 70 | 共享 | 287 | 952718 | 42 | 528183 | 8,112 | UT |
| 71 | 模型 | 304 | 1078146 | 32 | 404056 | 8,002 | UT |
| 72 | 保存 | 227 | 634485 | 36 | 429438 | 7,929 | UT |
| 73 | 实时 | 236 | 694935 | 44 | 445475 | 7,895 | UT |
| 74 | 磁盘 | 134 | 93963 | 32 | 36572 | 7,788 | UT |
| 75 | 调用 | 135 | 105410 | 37 | 58535 | 7,743 | UT |
| 76 | 扫描 | 186 | 435945 | 29 | 247427 | 7,635 | UT |
| 77 | 更改 | 147 | 197807 | 29 | 146444 | 7,602 | UT |
| 78 | 命令 | 184 | 432020 | 29 | 243240 | 7,583 | UT |
| 79 | 隐私 | 147 | 203311 | 29 | 126419 | 7,561 | UT |
| 80 | 木马 | 129 | 93924 | 24 | 42846 | 7,531 | UT |
| 81 | 接口 | 228 | 723701 | 46 | 309374 | 7,504 | UT |
| 82 | 识别 | 243 | 835966 | 41 | 422115 | 7,426 | UT |
| 83 | 修改 | 270 | 1012186 | 59 | 578792 | 7,412 | UT |
| 84 | 定义 | 221 | 726852 | 55 | 441004 | 7,276 | UT |
| 85 | 通常 | 322 | 1399596 | 63 | 1022878 | 7,232 | UT |
| 86 | 完整性 | 123 | 98256 | 26 | 76423 | 7,189 | UT |
| 87 | 工具 | 418 | 2068176 | 63 | 1162138 | 7,136 | UT |
| 88 | 允许 | 245 | 917083 | 48 | 671884 | 7,136 | UT |
| 89 | 使用 | 1668 | 10397826 | 109 | 4820822 | 7,129 | UT |
| 90 | 按钮 | 163 | 374226 | 22 | 191542 | 7,119 | UT |
| 91 | 客户端 | 172 | 436635 | 40 | 227976 | 7,106 | UT |
| 92 | 设置 | 448 | 2321704 | 60 | 1256590 | 7,014 | UT |
| 93 | 分布式 | 129 | 164966 | 31 | 73505 | 6,993 | UT |
| 94 | API | 114 | 65 330 | 25 | 37 065 | 6,976 | CT rejeté |
| 95 | 授权 | 200 | 651795 | 42 | 412958 | 6,971 | UT |
| 96 | 输入 | 237 | 907766 | 49 | 451974 | 6,955 | UT |
| 97 | 信息 | 1643 | 10579679 | 94 | 4408304 | 6,911 | UT |
| 98 | 架构 | 175 | 515094 | 28 | 278033 | 6,797 | UT |

| | | | | | | | |
|-----|----|-----|---------|----|---------|-------|-----------|
| 99 | 设备 | 999 | 6289758 | 74 | 2392552 | 6,78 | UT |
| 100 | 版本 | 234 | 948388 | 56 | 548688 | 6,716 | UT |
| 101 | 对象 | 325 | 1599602 | 44 | 980933 | 6,695 | UT |
| 102 | 静态 | 120 | 154627 | 27 | 104614 | 6,633 | UT |
| 103 | 硬盘 | 155 | 420137 | 27 | 139077 | 6,553 | UT |
| 104 | IP | 157 | 442 566 | 30 | 174 898 | 6,513 | CT rejeté |
| 105 | 特定 | 171 | 546337 | 52 | 400088 | 6,505 | UT |
| 106 | 类型 | 331 | 1719134 | 78 | 1129595 | 6,493 | UT |
| 107 | 功能 | 865 | 5717611 | 93 | 2709823 | 6,401 | UT |
| 108 | 窃取 | 101 | 52530 | 27 | 38078 | 6,377 | UT |
| 109 | 扩展 | 156 | 500476 | 44 | 364908 | 6,193 | UT |
| 110 | 格式 | 168 | 613536 | 33 | 316228 | 6,103 | UT |
| 111 | 网页 | 170 | 629447 | 32 | 353112 | 6,102 | UT |
| 112 | 云端 | 102 | 102982 | 20 | 65669 | 6,083 | UT |
| 113 | 无线 | 193 | 822402 | 21 | 304077 | 6,05 | UT |
| 114 | 传送 | 116 | 219332 | 24 | 136030 | 6,042 | UT |
| 115 | 本地 | 171 | 652317 | 47 | 470225 | 6,039 | UT |
| 116 | 安装 | 406 | 2506524 | 59 | 1094162 | 6,028 | UT |
| 117 | 执行 | 438 | 2765778 | 59 | 1525826 | 6,016 | UT |
| 118 | 例如 | 254 | 1326409 | 63 | 915915 | 5,979 | UT |
| 119 | 列表 | 118 | 260792 | 31 | 162687 | 5,9 | UT |
| 120 | 页面 | 159 | 607440 | 32 | 331020 | 5,831 | UT |
| 121 | 交互 | 116 | 260291 | 41 | 147839 | 5,815 | UT |
| 122 | 文本 | 119 | 308561 | 24 | 156291 | 5,694 | UT |
| 123 | 桌面 | 116 | 293736 | 34 | 154485 | 5,642 | UT |
| 124 | 控制 | 699 | 5254082 | 79 | 2484389 | 5,593 | UT |
| 125 | 外部 | 142 | 556582 | 37 | 388288 | 5,463 | UT |
| 126 | 物理 | 167 | 781178 | 40 | 392446 | 5,438 | UT |
| 127 | 生成 | 125 | 414658 | 37 | 257649 | 5,437 | UT |
| 128 | 电脑 | 334 | 2272697 | 36 | 1062444 | 5,38 | UT |
| 129 | 获取 | 176 | 915211 | 59 | 662803 | 5,266 | UT |
| 130 | 防护 | 127 | 474709 | 26 | 250295 | 5,26 | UT |
| 131 | 打开 | 232 | 1435188 | 32 | 1013675 | 5,231 | UT |
| 132 | 泄露 | 94 | 182992 | 25 | 123223 | 5,227 | UT |
| 133 | 地址 | 291 | 1973227 | 37 | 1196543 | 5,226 | UT |
| 134 | 配置 | 253 | 1646624 | 53 | 914049 | 5,187 | UT |
| 135 | 包含 | 178 | 960029 | 58 | 743573 | 5,187 | UT |
| 136 | 默认 | 95 | 208291 | 29 | 143633 | 5,146 | UT |
| 137 | 屏幕 | 159 | 802380 | 23 | 423854 | 5,135 | UT |

| | | | | | | | |
|-----|-----|------|----------|-----|---------|-------|-----------|
| 138 | 编程 | 100 | 257166 | 27 | 119422 | 5,133 | UT |
| 139 | 端口 | 88 | 151874 | 24 | 66065 | 5,103 | UT |
| 140 | 框架 | 133 | 573328 | 28 | 373366 | 5,093 | UT |
| 141 | 自动 | 321 | 2328755 | 67 | 1201409 | 5,089 | UT |
| 142 | 日志 | 89 | 165061 | 24 | 102239 | 5,083 | UT |
| 143 | 认证 | 203 | 1253944 | 33 | 579296 | 5,025 | UT |
| 144 | 管理员 | 99 | 276240 | 24 | 171955 | 5,001 | UT |
| 145 | 错误 | 182 | 1078176 | 39 | 677444 | 4,965 | UT |
| 146 | 创建 | 213 | 1379537 | 50 | 717194 | 4,954 | UT |
| 147 | 模块 | 139 | 705265 | 31 | 281936 | 4,841 | UT |
| 148 | 用于 | 248 | 1797785 | 58 | 1191038 | 4,798 | UT |
| 149 | 篡改 | 71 | 48834 | 22 | 34499 | 4,782 | UT |
| 150 | 机器 | 147 | 810382 | 36 | 443153 | 4,762 | UT |
| 151 | 路由 | 75 | 93287 | 21 | 35878 | 4,758 | UT |
| 152 | 目录 | 126 | 603410 | 24 | 358141 | 4,754 | UT |
| 153 | CPU | 89 | 234 240 | 24 | 103 946 | 4,752 | CT rejeté |
| 154 | 计算 | 274 | 2083820 | 49 | 987442 | 4,751 | UT |
| 155 | 动态 | 194 | 1288194 | 45 | 872164 | 4,742 | UT |
| 156 | 更新 | 257 | 1934142 | 53 | 1316322 | 4,716 | UT |
| 157 | 复制 | 110 | 458657 | 32 | 308139 | 4,699 | UT |
| 158 | 某些 | 141 | 776229 | 41 | 611329 | 4,686 | CT rejeté |
| 159 | 同步 | 131 | 676494 | 28 | 475292 | 4,681 | UT |
| 160 | 面向 | 143 | 807847 | 30 | 614709 | 4,654 | CT rejeté |
| 161 | 流 | 116 | 542330 | 27 | 392714 | 4,62 | CT rejeté |
| 162 | 逻辑 | 126 | 648620 | 34 | 388008 | 4,609 | UT |
| 163 | 特征 | 190 | 1308632 | 44 | 791228 | 4,608 | UT |
| 164 | 处理 | 558 | 5144790 | 72 | 2600713 | 4,574 | UT |
| 165 | 资源 | 645 | 6052656 | 76 | 2717377 | 4,571 | UT |
| 166 | 提供 | 1094 | 10745029 | 102 | 5877176 | 4,56 | CT rejeté |
| 167 | 身份 | 182 | 1247882 | 34 | 861752 | 4,56 | UT |
| 168 | 语言 | 256 | 2021278 | 41 | 963639 | 4,555 | UT |
| 169 | 开源 | 74 | 125252 | 24 | 62571 | 4,549 | UT |
| 170 | 方法 | 724 | 6922960 | 69 | 3305971 | 4,542 | CT rejeté |
| 171 | 开发者 | 79 | 180311 | 23 | 85677 | 4,536 | UT |
| 172 | 破坏 | 138 | 798450 | 30 | 560872 | 4,535 | UT |
| 173 | 编译 | 71 | 98726 | 22 | 72210 | 4,525 | UT |
| 174 | 如图 | 93 | 336014 | 23 | 139768 | 4,501 | CT rejeté |
| 175 | 恢复 | 217 | 1644890 | 38 | 1054320 | 4,5 | UT |
| 176 | 示 | 101 | 428001 | 29 | 214591 | 4,473 | CT rejeté |

| | | | | | | | |
|-----|-----|------|------------|-----|-----------|-------|-----------|
| 177 | 启动 | 239 | 1894699 | 39 | 1249340 | 4,471 | UT |
| 178 | 性能 | 263 | 2157643 | 61 | 1077759 | 4,459 | UT |
| 179 | 指令 | 86 | 282922 | 27 | 157491 | 4,416 | UT |
| 180 | 利用 | 481 | 4544992 | 73 | 2720950 | 4,404 | UT |
| 181 | 分析 | 754 | 7502328 | 68 | 3142114 | 4,397 | UT |
| 182 | 自动化 | 121 | 668026 | 23 | 322132 | 4,391 | UT |
| 183 | 防止 | 167 | 1171927 | 51 | 857182 | 4,377 | CT rejeté |
| 184 | 保护 | 443 | 4196427 | 63 | 2061110 | 4,351 | UT |
| 185 | 函数 | 93 | 380773 | 21 | 82046 | 4,334 | UT |
| 186 | 响应 | 107 | 538017 | 34 | 384951 | 4,322 | UT |
| 187 | 编码 | 87 | 321801 | 26 | 180999 | 4,309 | UT |
| 188 | 账号 | 94 | 407294 | 22 | 223047 | 4,28 | UT |
| 189 | 指 | 275 | 2428089 | 62 | 1463111 | 4,263 | CT rejeté |
| 190 | 策略 | 190 | 1509300 | 36 | 839304 | 4,212 | UT |
| 191 | 指定 | 119 | 719092 | 31 | 502418 | 4,185 | CT rejeté |
| 192 | 实现 | 656 | 6926778 | 86 | 3622579 | 4,123 | CT rejeté |
| 193 | 包括 | 574 | 6009339 | 91 | 3838405 | 4,107 | CT rejeté |
| 194 | 任务 | 327 | 3161371 | 54 | 1749538 | 4,099 | UT |
| 195 | 或 | 1403 | 15 660 280 | 102 | 6 702 428 | 4,091 | CT rejeté |
| 196 | 构建 | 142 | 1024581 | 38 | 695122 | 4,084 | UT |
| 197 | 图形 | 84 | 353901 | 24 | 184527 | 4,071 | UT |
| 198 | 虚拟 | 104 | 588658 | 40 | 282724 | 4,067 | UT |
| 199 | 手动 | 81 | 328140 | 22 | 200475 | 4,039 | UT |
| 200 | 机制 | 278 | 2653930 | 40 | 1392571 | 4,028 | UT |
| 201 | 内部 | 225 | 2032903 | 49 | 1232515 | 4,022 | UT |
| 202 | 存放 | 72 | 227132 | 25 | 158150 | 4,021 | UT |
| 203 | 它 | 857 | 9 558 920 | 76 | 4 447 884 | 3,997 | CT rejeté |
| 204 | 分为 | 165 | 1376342 | 49 | 1043146 | 3,928 | CT rejeté |
| 205 | 特性 | 119 | 820615 | 46 | 547252 | 3,927 | UT |
| 206 | 通过 | 1190 | 13790312 | 107 | 7239226 | 3,922 | CT rejeté |
| 207 | 维护 | 203 | 1856645 | 54 | 1206065 | 3,898 | UT |
| 208 | 接收 | 86 | 437811 | 28 | 274864 | 3,877 | UT |
| 209 | 支持 | 514 | 5683248 | 82 | 3214887 | 3,875 | UT |
| 210 | 业务 | 509 | 5703127 | 54 | 2224433 | 3,827 | CT rejeté |
| 211 | 修复 | 109 | 756851 | 27 | 387679 | 3,786 | UT |
| 212 | 技术 | 1326 | 16071542 | 94 | 5401331 | 3,775 | UT |
| 213 | 然后 | 359 | 3923189 | 61 | 2466765 | 3,762 | CT rejeté |
| 214 | 隔离 | 71 | 287676 | 22 | 168727 | 3,76 | UT |
| 215 | 搜索 | 163 | 1455355 | 31 | 724462 | 3,751 | UT |

| | | | | | | | |
|-----|----|------|------------|-----|-----------|-------|-----------|
| 216 | 流量 | 124 | 962891 | 33 | 417488 | 3,748 | UT |
| 217 | 非法 | 110 | 786955 | 24 | 388551 | 3,745 | UT |
| 218 | 从而 | 217 | 2176141 | 70 | 1594422 | 3,704 | CT rejeté |
| 219 | 管理 | 1378 | 17083702 | 87 | 5074642 | 3,699 | UT |
| 220 | 电子 | 370 | 4163758 | 45 | 1773028 | 3,685 | UT |
| 221 | 步骤 | 97 | 653193 | 31 | 434927 | 3,666 | CT rejeté |
| 222 | 可 | 1408 | 17 676 514 | 105 | 8 565 542 | 3,658 | CT rejeté |
| 223 | 故障 | 101 | 714756 | 28 | 301416 | 3,643 | UT |
| 224 | 它们 | 181 | 1770327 | 49 | 1067008 | 3,625 | CT rejeté |
| 225 | 服务 | 1239 | 15675008 | 86 | 5316146 | 3,615 | UT |
| 226 | 敏感 | 82 | 477496 | 28 | 372339 | 3,614 | CT rejeté |
| 227 | 过程 | 541 | 6510587 | 84 | 3703985 | 3,613 | CT rejeté |
| 228 | 属性 | 100 | 716479 | 24 | 382963 | 3,608 | UT |
| 229 | 复杂 | 147 | 1340159 | 52 | 1029588 | 3,6 | CT rejeté |
| 230 | 集成 | 103 | 760037 | 33 | 388241 | 3,599 | UT |
| 231 | 下载 | 216 | 2256205 | 47 | 976291 | 3,593 | UT |
| 232 | 转换 | 92 | 629503 | 25 | 391375 | 3,563 | UT |
| 233 | 划分 | 78 | 446019 | 29 | 313145 | 3,554 | CT rejeté |
| 234 | 智能 | 283 | 3190284 | 42 | 1064911 | 3,55 | UT |
| 235 | 驱动 | 107 | 836316 | 25 | 456373 | 3,548 | UT |
| 236 | 称为 | 128 | 1131851 | 45 | 855703 | 3,522 | CT rejeté |
| 237 | 方式 | 628 | 7926715 | 91 | 4615337 | 3,501 | CT rejeté |
| 238 | 使得 | 153 | 1486519 | 47 | 1124006 | 3,495 | CT rejeté |
| 239 | 引擎 | 83 | 543981 | 22 | 278796 | 3,473 | UT |
| 240 | 进行 | 1568 | 20968578 | 108 | 9265181 | 3,454 | CT rejeté |
| 241 | 选择 | 616 | 7893851 | 66 | 4393401 | 3,448 | CT rejeté |
| 242 | 所有 | 455 | 5685318 | 78 | 3819503 | 3,444 | CT rejeté |
| 243 | 结构 | 338 | 4087890 | 57 | 1972137 | 3,433 | UT |
| 244 | 模式 | 411 | 5099609 | 58 | 2520381 | 3,432 | UT |
| 245 | 不同 | 573 | 7442739 | 88 | 4105162 | 3,39 | CT rejeté |
| 246 | 针对 | 236 | 2736850 | 61 | 2042391 | 3,375 | CT rejeté |
| 247 | 交换 | 75 | 472002 | 20 | 297117 | 3,374 | UT |
| 248 | 包 | 136 | 1 341 592 | 29 | 765 497 | 3,357 | UT |
| 249 | 分配 | 104 | 903621 | 37 | 516744 | 3,328 | UT |
| 250 | 微软 | 74 | 477690 | 26 | 166563 | 3,323 | CT rejeté |
| 251 | 需要 | 873 | 11921577 | 103 | 6372579 | 3,316 | CT rejeté |
| 252 | 记录 | 174 | 1916250 | 44 | 1125685 | 3,311 | CT rejeté |
| 253 | 手段 | 149 | 1559820 | 46 | 1127689 | 3,308 | CT rejeté |
| 254 | 路径 | 70 | 430194 | 26 | 291647 | 3,297 | UT |

| | | | | | | | |
|-----|----|------|------------|-----|------------|-------|-----------|
| 255 | 库 | 72 | 463 790 | 29 | 274 179 | 3,285 | CT rejeté |
| 256 | 可以 | 1933 | 27394461 | 110 | 10997821 | 3,282 | CT rejeté |
| 257 | 限制 | 131 | 1319081 | 54 | 923483 | 3,281 | CT rejeté |
| 258 | 存在 | 385 | 5034094 | 76 | 3076813 | 3,259 | CT rejeté |
| 259 | 之间 | 337 | 4348854 | 71 | 2820128 | 3,251 | CT rejeté |
| 260 | 端 | 105 | 970 614 | 39 | 604 268 | 3,231 | UT |
| 261 | 数量 | 184 | 2140756 | 38 | 1396927 | 3,221 | CT rejeté |
| 262 | 确保 | 208 | 2503443 | 52 | 1663756 | 3,212 | CT rejeté |
| 263 | 用来 | 86 | 700092 | 36 | 560596 | 3,212 | CT rejeté |
| 264 | 开放 | 170 | 1958478 | 29 | 1178458 | 3,193 | UT |
| 265 | 可靠 | 83 | 665455 | 42 | 518009 | 3,191 | UT |
| 266 | 您 | 525 | 7 282 485 | 36 | 3 244 653 | 3,176 | CT rejeté |
| 267 | 任何 | 286 | 3721112 | 62 | 2564975 | 3,167 | CT rejeté |
| 268 | 目的 | 183 | 2197093 | 57 | 1546355 | 3,145 | CT rejeté |
| 269 | 场景 | 88 | 783850 | 28 | 500854 | 3,104 | UT |
| 270 | 平台 | 514 | 7312793 | 61 | 2884703 | 3,1 | UT |
| 271 | 采集 | 71 | 541287 | 23 | 284415 | 3,065 | UT |
| 272 | 网站 | 375 | 5267993 | 48 | 1921734 | 3,057 | UT |
| 273 | 需求 | 360 | 5039863 | 68 | 2723083 | 3,054 | UT |
| 274 | 周期 | 99 | 986817 | 27 | 618321 | 3,045 | UT |
| 275 | 点击 | 179 | 2236217 | 26 | 1458583 | 3,043 | UT |
| 276 | 部署 | 115 | 1240784 | 37 | 834509 | 3,038 | UT |
| 277 | 关闭 | 76 | 634356 | 25 | 432999 | 3,032 | UT |
| 278 | 台 | 201 | 2 592 583 | 42 | 1 387 941 | 3,032 | CT rejeté |
| 279 | 规则 | 113 | 1223315 | 44 | 679625 | 3,018 | UT |
| 280 | 数字 | 147 | 1763192 | 35 | 887357 | 3,012 | CT rejeté |
| 281 | 手机 | 377 | 5389309 | 44 | 1961758 | 3,012 | CT rejeté |
| 282 | 收集 | 95 | 953013 | 27 | 659699 | 2,996 | CT rejeté |
| 283 | 传播 | 120 | 1357417 | 31 | 705269 | 2,985 | UT |
| 284 | 状态 | 193 | 2558470 | 50 | 1707814 | 2,95 | CT rejeté |
| 285 | 客户 | 373 | 5469792 | 58 | 2022894 | 2,944 | UT |
| 286 | 快速 | 201 | 2700674 | 56 | 1863600 | 2,939 | CT rejeté |
| 287 | 相同 | 100 | 1068790 | 42 | 801106 | 2,938 | CT rejeté |
| 288 | 提出 | 299 | 4287399 | 44 | 2695193 | 2,937 | CT rejeté |
| 289 | 种 | 1639 | 25 963 425 | 108 | 10 686 334 | 2,936 | CT rejeté |
| 290 | 采用 | 346 | 5051669 | 63 | 2673071 | 2,935 | CT rejeté |
| 291 | 其他 | 470 | 7115493 | 78 | 4522742 | 2,914 | CT rejeté |
| 292 | 大量 | 197 | 2675384 | 64 | 1980469 | 2,906 | CT rejeté |
| 293 | 如 | 444 | 6 733 968 | 82 | 4 012 073 | 2,897 | CT rejeté |

| | | | | | | | |
|-----|----|------|------------|-----|-----------|-------|-----------|
| 294 | 显示 | 284 | 4124592 | 53 | 2548706 | 2,89 | CT rejeté |
| 295 | 混合 | 74 | 674995 | 26 | 385747 | 2,887 | CT rejeté |
| 296 | 评估 | 123 | 1493585 | 24 | 711209 | 2,871 | UT |
| 297 | 保证 | 218 | 3065672 | 59 | 2090909 | 2,871 | CT rejeté |
| 298 | 可能 | 572 | 8932479 | 89 | 4997795 | 2,868 | CT rejeté |
| 299 | 效率 | 138 | 1768230 | 47 | 1124331 | 2,842 | UT |
| 300 | 方案 | 280 | 4149017 | 55 | 2111724 | 2,837 | UT |
| 301 | 所 | 695 | 11 100 180 | 83 | 6 028 428 | 2,836 | CT rejeté |
| 302 | 感染 | 93 | 1021914 | 20 | 493526 | 2,833 | UT |
| 303 | 类 | 422 | 6 537 908 | 65 | 3 589 762 | 2,833 | CT rejeté |
| 304 | 并 | 1069 | 17 425 240 | 104 | 9 446 720 | 2,826 | CT rejeté |
| 305 | 无法 | 235 | 3425116 | 66 | 2392900 | 2,818 | CT rejeté |
| 306 | 即 | 301 | 4 552 222 | 72 | 2 904 403 | 2,81 | CT rejeté |
| 307 | 组 | 168 | 2 314 256 | 40 | 1 200 713 | 2,802 | CT rejeté |
| 308 | 需 | 216 | 3 131 308 | 65 | 2 195 700 | 2,799 | CT rejeté |

Annexe 6 – Liste de termes regroupés en vocables

| Numéro de vocable | Vocable | Acception (Terme en entrée) | Partie du discours | Équivalent français |
|-------------------|-----------------|------------------------------|--------------------|-------------------------|
| 1 | 安全 ān quán | 安全 ₁ ān quán | nom | sécurité informatique |
| | | 安全 ₂ ān quán | adj. | sécurisé |
| | | 安全 ₃ ān quán | adv. | sécuritairement |
| 2 | 安装 ān zhuāng | 安装 ₁ ān zhuāng | v. tr. | installer |
| | | 安装 ₂ ān zhuāng | nom | installation |
| 3 | 保存 bǎo cún | 保存 ₁ bǎo cún | v. tr. | enregistrer |
| | | 保存 ₂ bǎo cún | nom | enregistrement |
| 4 | 备份 bèi fèn | 备份 ₁ bèi fèn | v. tr. | sauvegarder |
| | | 备份 ₂ bèi fèn | nom | sauvegarde1 |
| | | 备份 ₃ bèi fèn | nom | sauvegarde2 |
| | | 备份 ₄ bèi fèn | adj. | sauvegardé |
| 5 | 测试 cè shì | 测试 ₁ cè shì | v. tr. | tester |
| | | 测试 ₂ cè shì | nom | test |
| 6 | 传输 chuán shū | 传输 ₁ chuán shū | v. tr. | transférer1 |
| | | 传输 ₂ chuán shū | nom | transfert1 |
| 7 | 存储 cún chǔ | 存储 ₁ cún chǔ | v. tr. | stocker |
| | | 存储 ₂ cún chǔ | nom | stockage1b |
| | | 存储 ₃ cún chǔ | nom | stockage1a |
| 8 | 恶意 è yì | 恶意 ₁ è yì | adj. | malveillant1 |
| | | 恶意 ₂ è yì | adv. | de manière malveillante |
| 9 | 发送 fā sòng | 发送 ₁ fā sòng | v. tr. | envoyer1 |
| | | 发送 ₂ fā sòng | v. tr. | envoyer2 |
| | | 发送 ₃ fā sòng | nom | envoi1 |
| | | 发送 ₄ fā sòng | nom | envoi2 |

| | | | | |
|----|------------------|-------------------------------|--------|----------------------|
| 10 | fǎng wèn 访问 | fǎng wèn 访问 ₁ | v. tr. | accéder1 |
| | | fǎng wèn 访问 ₂ | v. tr. | accéder2 |
| | | fǎng wèn 访问 ₃ | nom | accès1 |
| | | fǎng wèn 访问 ₄ | nom | accès2 |
| 11 | fù zhì 复制 | fù zhì 复制 ₁ | v. tr. | copier1 |
| | | fù zhì 复制 ₂ | v. tr. | copier2 |
| | | fù zhì 复制 ₃ | nom | copie1.1 |
| | | fù zhì 复制 ₄ | nom | copie2 |
| 12 | gēng xīn 更新 | gēng xīn 更新 ₁ | v. tr. | mettre à jour |
| | | gēng xīn 更新 ₂ | nom | mise à jour1.1 |
| | | gēng xīn 更新 ₃ | nom | mise à jour1.2 |
| 13 | gōng jī 攻击 | gōng jī 攻击 ₁ | v. tr. | attaquer1 |
| | | gōng jī 攻击 ₂ | nom | attaque informatique |
| 14 | gòng xiǎng 共享 | gòng xiǎng 共享 ₁ | v. tr. | partager1 |
| | | gòng xiǎng 共享 ₂ | nom | partage1 |
| | | gòng xiǎng 共享 ₃ | adj. | partagé |
| 15 | jiā mì 加密 | jiā mì 加密 ₁ | v. tr. | crypter |
| | | jiā mì 加密 ₂ | nom | cryptage |
| | | jiā mì 加密 ₃ | adj. | crypté |
| 16 | jiǎn cè 检测 | jiǎn cè 检测 ₁ | v. tr. | détecter1 |
| | | jiǎn cè 检测 ₂ | v. tr. | vérifier |
| | | jiǎn cè 检测 ₃ | nom | détection1 |
| | | jiǎn cè 检测 ₄ | nom | vérification |
| 17 | lián jiē 连接 | lián jiē 连接 ₁ | v. tr. | se connecter1 |
| | | lián jiē 连接 ₂ | v. tr. | connecter |
| | | lián jiē 连接 ₃ | nom | connexion1.1 |
| | | lián jiē 连接 ₄ | nom | connexion2.1 |
| 18 | qiān yí 迁移 | qiān yí 迁移 ₁ | v. tr. | faire migrer |
| | | qiān yí 迁移 ₂ | nom | migration |

| | | | | |
|----|----------------|-----------------------------|----------|------------------|
| 19 | rù qīn 入侵 | rù qīn 入侵 ₁ | v. tr. | s'introduire |
| | | rù qīn 入侵 ₂ | nom | intrusion |
| 20 | ruǎnjiàn 软件 | ruǎnjiàn 软件 ₁ | nom | logiciel1 |
| | | ruǎnjiàn 软件 ₂ | nom | logiciel2 |
| 21 | shān chú 删除 | shān chú 删除 ₁ | v. tr. | supprimer1 |
| | | shān chú 删除 ₂ | nom | suppression1 |
| 22 | shè zhì 设置 | shè zhì 设置 ₁ | v. tr. | configurer1 |
| | | shè zhì 设置 ₂ | nom | configuration1.1 |
| | | shè zhì 设置 ₃ | nom | configuration1.2 |
| 23 | shí shí 实时 | shí shí 实时 ₁ | adj. | en temps réel1 |
| | | shí shí 实时 ₂ | adv. | en temps réel2 |
| 24 | shǒudòng 手动 | shǒudòng 手动 ₁ | adj. | manuel |
| | | shǒudòng 手动 ₂ | adv. | manuellement |
| 25 | tóng bù 同步 | tóng bù 同步 ₁ | v. tr. | synchroniser |
| | | tóng bù 同步 ₂ | nom | synchronisation |
| | | tóng bù 同步 ₃ | adj. | synchrone1 |
| | | tóng bù 同步 ₄ | adv. | synchroniquement |
| 26 | wǎng luò 网络 | wǎng luò 网络 ₁ | nom | réseau1 |
| | | wǎng luò 网络 ₂ | nom | réseau2 |
| | | wǎng luò 网络 ₃ | adj. | en ligne |
| 27 | xià zài 下载 | xià zài 下载 ₁ | v. tr. | télécharger |
| | | xià zài 下载 ₂ | nom | téléchargement |
| 28 | yí dòng 移动 | yí dòng 移动 ₁ | adj. | mobile1 |
| | | yí dòng 移动 ₂ | adj. | amovible |
| 29 | yìngyòng 应用 | yìngyòng 应用 ₁ | nom | application1 |
| | | yìngyòng 应用 ₂ | nom | application2 |
| | | yìngyòng 应用 ₃ | v. intr. | s'appliquer |
| | | yìngyòng 应用 ₄ | v. tr. | appliquer2 |

| | | | | |
|----|-----------------|------------------------------|----------|-----------------|
| 30 | yuǎnchéng 远程 | yuǎnchéng 远程 ₁ | adj. | distant |
| | | yuǎnchéng 远程 ₂ | adv. | à distance |
| 31 | yùn xíng 运行 | yùn xíng 运行 ₁ | v. intr. | s'exécuter |
| | | yùn xíng 运行 ₂ | v. tr. | faire exécuter |
| | | yùn xíng 运行 ₃ | nom | exécution1a |
| | | yùn xíng 运行 ₄ | nom | exécution1b |
| 32 | zài xiàn 在线 | zài xiàn 在线 ₁ | ad | en ligne1 |
| | | zài xiàn 在线 ₂ | adv. | en ligne2 |
| 33 | zì dòng 自动 | zì dòng 自动 ₁ | adj. | automatique1 |
| | | zì dòng 自动 ₂ | adv. | automatiquement |

Annexe 7 – Exemples de paires de CLS contenant les polysèmes relevés

| Regroupement des CLS contenant un polysème | Forme des CLS | Structure syntaxique des CLS | Exemples | Structure syntaxique des CLS | CLS | Glose | Équivalent de la CLS | Acception du polysème | Partie du discours | Équivalent de l'acception | Fonction syntaxique du polysème |
|---|---------------|------------------------------|-----------|------------------------------|-------------------------------------|--|---|--------------------------------------|--------------------|--|---------------------------------|
| 2 CLS ⇒ 1 forme ⇒ 1 structure syntaxique | Identique | Identique | Exemple 1 | [Modificateur] + Nom modifié | 网络1攻击 | Endommager, forcer ou détourner un réseau pour commettre un acte préjudiciable | attaque envers un réseau | 网络1 | nom | réseau | Modificateur |
| | | | | | 网络3攻击 | Actions malveillantes qui ciblent différents dispositifs informatiques par l'intermédiaire d'un réseau | cyberattaque | 网络3 | adj | en ligne | |
| | | | Exemple 2 | Modificateur + [Nom modifié] | 系统设置2 | Activité réalisée par un utilisateur qui définit les paramètres d'un système (soit un logiciel particulier d'exploitation), afin d'en assurer le bon fonctionnement | configuration d'un système | 设置2 | nom | configuration (nom d'action) | Nom modifié |
| | | | | | 系统设置3 | Ensemble des paramètres associés par un utilisateur à un système pour que ce système fonctionne en conformité avec du matériel ou un logiciel spécifique | configuration d'un système | 设置3 | nom | configuration (nom de résultat) | |
| Regroupement des CLS contenant un polysème | Forme des CLS | Structure syntaxique des CLS | Exemples | Structure syntaxique des CLS | CLS | Glose | Équivalent de la CLS | Acception du polysème | Partie du discours | Équivalent de l'acception | Fonction syntaxique du polysème |
| 2 CLS ⇒ 1 forme ⇒ 2 structures syntaxiques | Identique | Différente | Exemple 1 | [Prédicat] + COD | 加密1数据 | Transformer des données en un cryptogramme, de manière à la rendre inintelligible à toute personne non autorisée et à en assurer ainsi la confidentialité | crypter des données | 加密1 | verbe | crypter | Prédicat |
| | | | | | [Modificateur] + Nom modifié | 加密3数据 | Un ensemble de données chiffrées | données cryptées | 加密3 | adj | crypté |
| | | | Exemple 2 | [Modificateur] + Nom modifié | 自动1下载2 | Activité réalisée sans intervention humaine qui consiste à déplacer un fichier ou un logiciel à partir d'un réseau ou d'un ordinateur distant et place le fichier ou le logiciel sur l'ordinateur d'un utilisateur | téléchargement automatique | 自动1 | adj | automatique | Modificateur |
| | | | | | [Modificateur adverbial] + Prédicat | 自动2下载1 | Déplacer, sans intervention humaine, un fichier ou un logiciel à partir d'un réseau ou d'un ordinateur distant et place le fichier ou le logiciel sur l'ordinateur d'un utilisateur | télécharger automatiquement | 自动2 | adv | automatiquement |
| Regroupement des CLS contenant un polysème | Forme des CLS | Structure syntaxique des CLS | Exemples | Structure syntaxique des CLS | CLS | Glose | Équivalent de la CLS | Acception du polysème | Partie du discours | Équivalent de l'acception | Fonction syntaxique du polysème |
| 2 CLS ⇒ 2 formes ⇒ 1 structure syntaxique | Différente | Identique | Exemple 1 | [Modificateur] + Nom modifié | 移动1应用1 | Application conçue pour être téléchargée et fonctionner sur un appareil mobile | application mobile | 移动1 | adj | mobile | Modificateur |
| | | | | | 移动2硬盘 | Unité de disque dur montée dans une cartouche, que l'utilisateur insère dans un ordinateur hôte ou dans l'un de ses périphériques et qu'il peut retirer par la suite, afin de la remplacer par une autre | disque dur amovible | 移动2 | adj | amovible | |
| | | | Exemple 2 | Modificateur + [Nom modifié] | 原生应用1 | Application écrite avec un code compilé spécialement pour le processeur qui va l'exécuter | application native | 应用1 | nom | application1 (=logiciel d'application) | Nom modifié |
| | | | | | 商业应用2 | Utilisation de moyens informatiques pour répondre à un besoin déterminé et faciliter les processus de travail des utilisateurs | application commerciale | 应用2 | nom | application2 (= application de l'informatique) | |
| | | | Exemple 3 | [Prédicat] + COD | 检测1攻击2 | Un programme détecte la présence d'une attaque informatique, soit une action réalisée par un pirate ou un virus qui entre dans un réseau ou un système informatique illégalement dans le but d'accéder à des données et, éventuellement, de nuire au fonctionnement du matériel ou à l'intégrité des données | détecter une attaque informatique | 检测1 | verbe | détecter | Prédicat |
| | | | | | 检测2软件1 | Tester, à l'aide d'un programme approprié, l'état d'un logiciel afin d'en évaluer le niveau de sécurité et d'en déterminer les vulnérabilités | mettre un logiciel à l'essai | 检测2 | verbe | mettre à l'essai | |
| Regroupement des CLS contenant un polysème | Forme des CLS | Structure syntaxique des CLS | Exemples | Structure syntaxique des CLS | CLS | Glose | Équivalent de la CLS | Acception du polysème | Partie du discours | Équivalent de l'acception | Fonction syntaxique du polysème |
| 2 CLS ⇒ 2 formes ⇒ 2 structures syntaxiques | Différente | Différente | Exemple 1 | [Prédicat] + COD | 测试1软件1 | Contrôler, à l'aide de tests, la bonne exécution d'un logiciel | tester un logiciel | 测试1 | verbe | tester | Prédicat |
| | | | | | Modificateur + [Nom modifié] | 软件1测试2 | Opération destinée à contrôler la bonne exécution d'un logiciel | test logiciel | 测试2 | nom | test |
| | | | Exemple 2 | [Prédicat] + COD | 传输1数据 | Acheminer des données d'un emplacement de mémoire vers un autre | transférer des données | 传输1 | verbe | transférer | Prédicat |
| | | | | | Modificateur + Nom modifié | 传输2介质 | Élément matériel ou virtuel permettant d'emmagasiner et de transférer l'information | support de transmission (de données) | 传输2 | nom | transfert |

Annexe 8 – Modèle d’encodage du CHINOINFO (exemple de l’article ^{ruǎn jiàn} 软件₁ « logiciel₁ »)

| | |
|--|--|
| Entrée | 软件 ₁ |
| Partie du discours | Nom |
| Équivalent français | Logiciel ₁ |
| Structure actancielle | 用户(utilisateur) 使用(utilise) 软件 ₁ 操作 (intervenir sur) 应用 ₂ (application informatique) 办公 (bureautique) 杀毒 (antivirus) |
| | [agent] [patient] |
| Définition en français | Ensemble des programmes constituant une unité destinée à effectuer un traitement particulier sur un ordinateur. [GDTLOGICIEL] |
| Contextes | |
| <p>1.按软件的功能可以将软件分为：系统软件、支撑软件、应用软件。应用软件是在特定领域内开发、为特定目的服务的一类软件。[D12]</p> <p>2.软件工程是开发、运行、维护和修复软件的系统方法。[D98]</p> <p>3.专有软件，又称非自由软件、专属软件、私有软件、封闭性软件等，是指在使用、修改上有限制的软件。与专有软件对应的是自由软件。专有软件并不等同于商业软件。专有软件和自由软件都可以免费或者收费分发。它们之间的区别在于：专有软件的所有者可以决定是否可以分发该软件，以及费用的数额；而自由软件可以被任何持有者随意分发。[D97]</p> <p>4.我国国家标准《计算机软件开发规范》（GB8566-88）把软件生命周期划分为可行性研究、需求分析、概要设计、详细设计、实现、组装测试、确认测试、使用和维护 8 个阶段。通常我们把可行性研究、需求分析称为软件定义时期，把概要设计、详细设计、实现、组装测试、确认测试称为软件开发时期，而把使用与维护阶段称为软件运行维护时期。[D104]</p> <p>5.支撑软件是在系统软件和应用软件之间，提供应用软件设计、开发、测试、评估、运行检测等辅助功能的软件。[D103]</p> <p>6.移动互联网终端软件主要包括操作系统和第三方应用软件，其特点是以智能终端操作系统为基础，结合各种层次或类别的中间件实现对应用服务的支持。[D63]</p> <p>7.手机安全解决方案通过将加密、验证、恶意软件防护、防火墙和入侵检测结合在一起，提供了防病毒、防垃圾短信、防间谍软件、防数据丢失、防黑客等一系列功能。[D63]</p> <p>8.为了加强客户端的安全，避免或减少通过浏览器实施的攻击，除了安装防火墙和杀毒软件等安全工具外，一般还应当采取以下措施：[...] 不要下载安装未经安全认证的软件和插件。[D53]</p> <p>9.NetBeans 是用 Java 编写的开源和免费软件开发工具，可轻松，快速地开发世界一流的 Web，移动和桌面应用程序。[D92]</p> <p>10.软件测试工具是用于对系统，子系统，模块或单元进行测试的工具。[D92]</p> <p>11.及时更新系统能够更好地避免病毒和其他恶意软件的入侵，因为黑客通常都是利用操作系统的漏洞来发</p> | |

送病毒软件或程序，及时更新已修复漏洞的系统能够更好地保护个人信息安全和使用安全。[D79]

12.Windows 10 自带的杀毒软件会在开机时自动更新病毒库。[D43]

13.应用商店是恶意软件传播的渠道之一。恶意软件通常在应用商店中伪装成合法的应用程序进行分发。恶意程序在获得超级用户权限的情况下，可通过将恶意子包释放到系统目录下并重启手机以自动启动恶意子包 这样用户将无法卸载恶意软件。[D33]

14.木马、间谍软件通常以远程控制的方式来操控用户设备。[D33]

15.通常软件安装完成后，会在桌面上建立快捷方式图标，双击这些图标即可启动对应的软件。[D55]

16.为了保证信息安全，用户应为计算机安装可靠的杀毒软件和防火墙软件，并定期更新这些软件。[D94]

Termes reliés

| Lien sémantique | Explication | Terme relié | Équivalent |
|-----------------|--|-----------------|--------------------------|
| Hyponyme | Un type spécifique de logiciel ₁ | 应用 ₁ | application ₁ |
| Hyperonyme | Ensemble de logiciel ₁ s | 软件 ₂ | logiciel ₂ |
| Méronyme | Élément constitutif d'un logiciel ₁ | 程序 | programme ₁ |
| Autre lien | État d'un logiciel ₁ à un moment donné de son développement | 版本 | version ₁ |

Combinaisons lexicales

| Classe générique | Classe intermédiaire | Glose | CLS | Équivalent français |
|--|----------------------|---|--------------------|---------------------------------------|
| Modificateur- nom modifié | Sortes de | Logiciel, ou sous-ensemble de la partie logicielle d'un système informatique, qui assure le fonctionnement de l'ordinateur et son exploitation [GDTLOGISYST] | 系统软件 ₁ | logiciel système (= logiciel de base) |
| | | Ensemble de programmes dont se servent les utilisateurs afin d'accomplir une tâche ou une activité particulière [GDTLOGIDAPP] | 应用软件 ₁ | logiciel d'application |
| | | Logiciel de sécurité informatique qui procède à l'analyse de données afin de détecter les virus, de bloquer leur intrusion ou de les supprimer d'un système infecté [GDTLOGIANTI] | 杀毒软件 ₁ | logiciel antivirus |
| | | | 防病毒软件 ₁ | logiciel antivirus |
| | | Logiciel fourni avec son code source de manière à ce qu'il puisse être étudié, copié, modifié et redistribué [GDTLOGILIBR] | 开源软件 ₁ | logiciel libre |
| | | | 自由软件 ₁ | logiciel libre |
| | | Qui a été conçu par un propriétaire à des fins lucratives | 商业软件 ₁ | logiciel commercial |
| Logiciel protégé par des droits d'auteur et distribué sans son code source par la personne physique ou morale l'ayant créé, pour lequel il est nécessaire d'acquiescer une licence d'utilisation [GDTLOGIPROP] | 专有软件 ₁ | logiciel propriétaire | | |

| | | | | |
|--------------|---|---|---------------------------------|-----------------------------------|
| | | Logiciel sur lequel le programmeur conserve ses droits d'auteur, mais ne réclame pas de redevances, et qui peut donc être copié et distribué gratuitement [GDTLOGIGRAT] | 免费软件 ₁ | logiciel gratuit (= gratuiticiel) |
| | | Logiciel visant à voler, à altérer ou à détruire des données afin de porter préjudice à leur propriétaire ou de nuire au fonctionnement d'un système informatique [GDTLOGIMALV] | 恶意软件 ₁ | logiciel malveillant (= maliciel) |
| | | Logiciel généralement indétectable qui s'infiltré dans un système informatique lors d'un téléchargement ou d'une installation dans le but d'employer la connexion Internet d'un utilisateur pour recueillir et transmettre ses renseignements personnels à son insu [GDTLOGIESPI] | 间谍软件 ₁ | logiciel espion (= espiogiciel) |
| | | Logiciel qui informatise les travaux courants dans un bureau [GDTLOGIBURE] | 办公软件 ₁ | logiciel bureautique |
| Prédicat-COD | Préparer l'utilisation/le fonctionnement d'un ~ | Un utilisateur place un logiciel sur un ordinateur et définit des paramètres pour que le logiciel puisse bien fonctionner [DIINST] | 安装 ₁ 软件 ₁ | installer un logiciel |
| | Mettre un ~ quelque part | Un utilisateur déplace un fichier ou un logiciel à partir d'un réseau ou d'un ordinateur distant et place le fichier ou le logiciel sur l'ordinateur de l'utilisateur | 下载 ₁ 软件 ₁ | télécharger un logiciel |
| | Transformer un ~ | Un utilisateur actualise un logiciel [= adapter le logiciel de manière qu'il soit conforme aux développements les plus récents] [GDTMETTAJ] | 更新 ₁ 软件 ₁ | mettre à jour un logiciel |
| | Créer ou supprimer un ~ | Concevoir, réaliser, évaluer et mettre à jour un logiciel [GDTDEVE] | 开发 ₁ 软件 ₁ | développer un logiciel |
| | Commencer à utiliser un ~ | Un utilisateur ou un programme place un logiciel en mémoire de manière à pouvoir utiliser le logiciel [DICHARGER] | 加载 ₁ 软件 ₁ | charger un logiciel |
| | | Un utilisateur charge un programme, un logiciel ou un logiciel d'application en mémoire afin qu'il s'exécute [GDTLANC] | 启动 ₁ 软件 ₁ | lancer un logiciel |
| | Empêcher l'utilisation d'un ~ | Exécuter la procédure visant à supprimer un logiciel d'application installé sur un ordinateur, ainsi que les fichiers secondaires qui y sont associés [GDTDESINST] | 卸载 ₁ 软件 ₁ | désinstaller un logiciel |

| | | | | |
|-------------------------------------|---|--|---------------------------------|------------------------------|
| <u>Modificateur-</u> nom modifié | Moyen et instrument | Ensemble des connaissances, des méthodes et des acquis scientifiques et techniques servant à la conception, au développement, à la vérification et à la documentation de logiciels, afin d'en rationaliser la production, le suivi et la qualité [GDTGENILOGICIEL] | 软件 ₁ 工程 | génie logiciel |
| | Fonctionnement un ~ | Processus pendant lequel un logiciel fonctionne sur un ordinateur ou sous un système d'exploitation | 软件 ₁ 运行 ₃ | exécution d'un logiciel |
| | Opération visant à préparer l'utilisation d'un ~ | Activité réalisée par un utilisateur qui place un logiciel sur un ordinateur et définit des paramètres pour que le logiciel puisse bien fonctionner. | 软件 ₁ 安装 ₂ | installation d'un logiciel |
| | Opération visant à mettre un ~ quelque part | Action réalisée par un utilisateur qui consiste à déplacer un fichier ou un logiciel à partir d'un réseau ou d'un ordinateur distant et place le fichier ou le logiciel sur l'ordinateur de l'utilisateur. | 软件 ₁ 下载 ₂ | téléchargement d'un logiciel |
| | Opération visant à créer un ~ | Activité informatique qui consiste à concevoir, à réaliser, à évaluer et à mettre à jour des logiciels ou des programmes. | 软件 ₁ 开发 ₂ | développement logiciel |
| | | Ensemble des activités consistant à réaliser la conception architecturale et la conception détaillée d'un logiciel. | 软件 ₁ 设计 | conception logicielle |
| | Opération visant à transformer un ~ | Opération consistant à déployer une version d'un logiciel qui comporte des corrections ou des modifications mineures par rapport à la version précédente. | 软件 ₁ 更新 ₂ | mise à jour d'un logiciel |
| | | Opération consistant à améliorer les fonctionnalités d'un logiciel, ou encore à en ajouter ou à en retirer, et à y apporter des modifications majeures. | 软件 ₁ 升级 | mise à niveau d'un logiciel |
| | Opération visant à assurer le bon fonctionnement d'un ~ | Activité qui consiste à mettre en œuvre différents tests visant à s'assurer qu'un logiciel est conforme aux besoins du client, à détecter les erreurs qu'il pourrait comporter ainsi qu'à vérifier son fonctionnement et sa maintenabilité. | 软件 ₁ 测试 ₂ | test logiciel |
| | | Ensemble des opérations qui assurent en tout temps le bon fonctionnement d'un système informatique, conformément à des spécifications définies. | 软件 ₁ 维护 | maintenance |
| <u>Sujet-Prédicat</u> | Fonctionner/ Ne pas fonctionner | un logiciel fonctionne sur un ordinateur ou sous un système d'exploitation | 软件 ₁ 运行 ₁ | un logiciel s'exécute |
| | | un logiciel cesse de fonctionner | 软件 ₁ 闪退 | un logiciel plante |

Annexe 9 – Fiche 软件 « logiciel » en XML

ruǎn jiàn

Explorateur de source de don... • logiciel-%E8%BD%AF%E4%BB%B6.xml x

Chaque vocable correspond à un fichier XML

La première acception du vocable 软件

```

2
3 <lexie xml:id="软件1" glose="logiciel1" numero-acception="1" statut="3" projet="CHINOINFO" redacteur="ZWH" mise-a-jour="2022-03-19">
4
5 <information-grammaticale>nom</information-grammaticale>
6 <equivalences> [2 lines]
7
8 <structure-actancielle> [7 lines]
9
10 <definition source="GDTLOGICIEL">Ensemble des programmes constituant une unité destinée à [1 line]
11
12 <domaine>informatique</domaine>
13
14 <contextes> [26 lines]
15
16 <liens-lexicaux> [17 lines]
17
18 <combinaisons-lexicales>
19
20 <structure-syntaxique nom="Modificateur + [Nom modifié]">
21 <lien-semantic nom="Sortes de"> [96 lines]
22 </structure-syntaxique>
23
24 <structure-syntaxique nom="Prédicat + [Complément d'objet]">
25 <lien-semantic nom="Préparer l'utilisation/le fonctionnement"> [9 lines]
26 <lien-semantic nom="Mettre un logiciel quelque part"> [9 lines]
27 <lien-semantic nom="Transformer"> [8 lines]
28 <lien-semantic nom="Créer ou supprimer"> [8 lines]
29 <lien-semantic nom="Commencer à utiliser un logiciel"> [15 lines]
30 <lien-semantic nom="Empêcher l'utilisation d'un logiciel"> [9 lines]
31 </structure-syntaxique>
32 <structure-syntaxique nom="[Modificateur] + Nom modifié"> [111 lines]
33 <structure-syntaxique nom="[Sujet] + Prédicat"> [16 lines]
34 </combinaisons-lexicales>
35 <informations-complementaires/>
36 </lexie>
37
38 <lexie xml:id="软件2" glose="logiciel2" numero-acception="2" statut="3" projet="CHINOINFO" redacteur="ZWH" mise-a-jour="2022-03-19">
39 <information-grammaticale>nom</information-grammaticale>
40 <equivalences> [2 lines]
41 <structure-actancielle> [8 lines]
42 <definition source="DILOGICIEL2">Ensemble des ressources logiques utilisées par un [2 lines]
43 <domaine>informatique</domaine>
44 <contextes> [6 lines]
45 <liens-lexicaux> [9 lines]
46 <combinaisons-lexicales>
47 <structure-syntaxique nom="Construction appositive = [Modificateur] + Nom modifié apposé">
48 <lien-semantic nom="Logiciel envisagé du point de vue de ses utilisations"> [16 lines]
49 </structure-syntaxique>
50 </combinaisons-lexicales>
51 <informations-complementaires/>
52 </lexie>
53 </vocable>

```

Les balises sont signalées par des chevrons.

Valeur de la balise : données encodées

Ensemble des programmes constituant une unité destinée à

Rubrique 1: Renseignements généraux sur le terme vedette

Rubrique 2: Contextes d'utilisation

Rubrique 3: Termes reliés (liens paradigmatiques)

Rubrique 4: Combinaisons lexicales (liens syntagmatiques)

Classes intermédiaires

Classe générique

La deuxième acception du vocable 软件

images
dicoLilex
dicoRH

Annexe 10 – Exemple d'article du CHINOINFO

软件 1 (nom)
[Top]

INFORMATIQUE

français : [logiciel 1](#)

Définition

Structure actancielle

Ensemble des programmes constituant une unité destinée à effectuer un traitement particulier sur un ordinateur. [GDTLOGICIEL]

Contextes

1. 按软件的功能可以将软件分为：系统软件、支撑软件、应用软件。应用软件是在特定领域内开发、为特定目的服务的一类软件。[D12]
2. 软件工程是开发、运行、维护和修复软件的系统方法。[D98]
3. 专有软件，又称非自由软件、专属软件、私有软件、封闭性软件等，是指在使用、修改上有限制的软件。与专有软件对应的是自由软件。专有软件并不等同于商业软件。专有软件和自由软件都可以免费或者收费分发。它们之间的区别在于：专有软件的所有者可以决定是否可以分发该软件，以及费用的数额；而自由软件可以被任何持有者随意分发。[D97]
4. 我国国家标准《计算机软件开发规范》（GB8566-88）把软件生命周期划分为可行性研究、需求分析、概要设计、详细设计、实现、组装测试、确认测试、使用和维修8个阶段。通常我们把可行性研究、需求分析称为软件定义时期，把概要设计、详细设计、实现、组装测试、确认测试称为软件开发时期，而把使用与维护阶段称为软件运行维护时期。[D104]
5. 支撑软件是在系统软件和应用软件之间，提供应用软件设计、开发、测试、评估、运行检测等辅助功能的软件。[D103]
6. 移动互联网终端软件主要包括操作系统和第三方应用软件。其特点是以智能终端操作系统为基础，结合各种层次或类别的中间件实现对应用服务的支持。[D63]
7. 手机安全解决方案通过将加密、验证、恶意软件防护、防火墙和入侵检测结合在一起，提供了防病毒、防垃圾短信、防间谍软件、防数据丢失、防黑客等一系列功能。[D63]
8. 为了加强客户端的安全，避免或减少通过浏览器实施的攻击，除了安装防火墙和杀毒软件等安全工具外，一般还应当采取以下措施：[...] 不要下载安装未经安全认证的软件和插件。[D53]
9. NetBeans是用Java编写的开源和免费软件开发工具，可轻松、快速地开发世界一流的Web、移动和桌面应用程序。[D92]
10. 软件测试工具是用于对系统、子系统、模块或单元进行测试的工具。[D92]
11. 及时更新系统能够更好地避免病毒和其他恶意软件的入侵，因为黑客通常都是利用操作系统的漏洞来发送病毒软件或程序，及时更新已修复漏洞的系统能够更好地保护个人信息安全和使用安全。[D79]
12. Windows 10自带的杀毒软件会在开机时自动更新病毒库。[D43]
13. 应用商店是恶意软件传播的黑道之一。恶意软件通常在应用商店中伪装成合法的应用程序进行分发。恶意程序在获得超级用户权限的情况下，可通过将恶意安装包释放到系统目录下并重启手机以自动启动恶意安装包。这样用户将无法卸载恶意软件。[D33]
14. 木马、间谍软件通常以远程控制的方式来操控用户设备。[D33]
15. 通常软件安装完成后，会在桌面上建立快捷方式图标，双击这些图标即可启动对应的软件。[D55]
16. 为了保证信息安全，用户应该为计算机安装可靠的杀毒软件和防火墙软件，并定期更新这些软件。[D94]

Termes reliés

Combinaisons lexicales

+ Modificateur + [Nom modifié]

| Sortes de | | |
|-----------|---|---|
| 系统软件 1 | Logiciel, ou sous-ensemble de la partie logicielle d'un système informatique, qui assure le fonctionnement de l'ordinateur et son exploitation [GDTLOGISYST] | logiciel système (= logiciel de base) |
| 应用 1 软件 1 | Ensemble de programmes dont se servent les utilisateurs afin d'accomplir une tâche ou une activité particulière [GDTLOGIDAPP] | logiciel d'application |
| 杀毒软件 1 | Logiciel de sécurité informatique qui procède à l'analyse de données afin de détecter les virus, de bloquer leur intrusion ou de les supprimer d'un système infecté [GDTLOGIANTI] | logiciel antivirus |
| 防病毒软件 1 | Logiciel de sécurité informatique qui procède à l'analyse de données afin de détecter les virus, de bloquer leur intrusion ou de les supprimer d'un système infecté [GDTLOGIANTI] | logiciel antivirus |
| 开源软件 1 | Logiciel fourni avec son code source de manière à ce qu'il puisse être étudié, copié, modifié et redistribué [GDTLOGILIBR] | logiciel libre |
| 自由软件 1 | Logiciel fourni avec son code source de manière à ce qu'il puisse être étudié, copié, modifié et redistribué [GDTLOGILIBR] | logiciel libre |
| 商业软件 1 | Qui a été conçu par un propriétaire à des fins lucratives | logiciel commercial |
| 专有软件 1 | Logiciel protégé par des droits d'auteur et distribué sans son code source par la personne physique ou morale l'ayant créé, pour lequel il est nécessaire d'acquiescer une licence d'utilisation [GDTLOGIPROP] | logiciel propriétaire |
| 免费软件 1 | Logiciel sur lequel le programmeur conserve ses droits d'auteur, mais ne réclame pas de redevances, et qui peut donc être copié et distribué gratuitement [GDTLOGIGRAT] | logiciel gratuit (= gratuiticiel) |
| 黑客 1 软件 1 | Logiciel visant à voler, à altérer ou à détruire des données afin de porter préjudice à leur propriétaire ou de nuire au fonctionnement d'un système informatique [GDTLOGIMALV] | logiciel malveillant (= maliciel) |
| 间谍软件 1 | Logiciel généralement indétectable qui s'infiltre dans un système informatique lors d'un téléchargement ou d'une installation dans le but d'employer la connexion Internet d'un utilisateur pour recueillir et transmettre ses renseignements personnels à son insu [GDTLOGIESPI] | logiciel espion (= espioiciel) |
| 办公软件 1 | Logiciel qui informatise les travaux courants dans un bureau [GDTLOGIBURE] | logiciel bureautique |

▶ Prédicat + [Complément d'objet]

▶ [Modificateur] + Nom modifié

▶ [Sujet] + Prédicat

Annexe 11 – Test lexical portant sur le lexique de l’informatique

I. Distinguer les sens d’une forme lexicale

1. Dites si 网络 a le même sens ou des sens différents dans les phrases suivantes. Justifiez votre réponse (2 points).

- 1). 如果浏览器无法打开网页，请首先检查网络连接。
- 2). 2013 年 10 月，清华大学正式推出网络学习平台，面向全球提供在线课程。

Réponse et justification

2. Dites si 移动 a le même sens ou des sens différents dans les phrases suivantes. Justifiez votre réponse (2 points).

- 1). 移动互联网是指用户使用移动终端（如手机、平板电脑），通过无线通信网络访问互联网。
- 2). 移动硬盘是一种常见的便携式存储设备，用于存储和访问数据。

Réponse et justification

3. Dites si 更新 a le même sens ou des sens différents dans les phrases suivantes. Justifiez votre réponse (2 points).

- 1). 如果有新的软件版本发布，手机会提示您下载并安装更新。
- 2). 为了保护计算机的安全，请及时更新和升级您的操作系统。

Réponse et justification

II. Appréhender le sens d'une combinaison lexicale

4. Dites si l'expression « 安装程序 » véhicule le même sens dans les phrases suivantes. Justifiez votre réponse (2 points).

- 1). 如果应用程序无法正常运行，用户可以卸载并重新安装程序。
- 2). 安装程序的文件名一般为 Setup.exe。用户双击文件名之后，Windows 7 操作系统将自动运行这一程序。

Réponse et justification

5. Dites si l'expression « 备份数据 » véhicule le même sens dans les phrases suivantes. Justifiez votre réponse (2 points).

- 1). 为了保护计算机中的重要数据，您可以使用 Dropbox 云盘定期备份数据。
- 2). 要恢复数据，用户可以从 iCloud 云端硬盘下载备份数据到本地计算机。

Réponse et justification

6. Dites si l'expression « 共享文件 » véhicule le même sens dans les phrases suivantes. Justifiez votre réponse (2 points).

- 1). Google Drive 允许用户通过互联网与其他用户共享文件。
- 2). 在局域网中，管理员可以授权指定用户查看、编辑或删除共享文件。

Réponse et justification

III. Choisir des cooccurrents appropriés pour un mot donné

7. Identifiez deux verbes qui peuvent s'associer avec 网站 (« site web ») dans la phrase suivante (2 points).

连接互联网后，用户可以使用手机端 Chrome 浏览器_____网站。

- A). 访问
- B). 获取
- C). 入侵
- D). 登陆

8. Identifiez deux verbes qui peuvent s'associer avec 文件 (« fichier ») dans la phrase suivante (2 points).

在局域网内，用户可以把本地文件_____到 FTP 服务器。

- A). 上传
- B). 移动
- C). 保存
- D). 记录

9. Identifiez deux noms qui peuvent s'associer à 传输 (« transférer ») dans la phrase suivante (2 points).

云盘是一种数据存储和管理工具。连接网络后，用户可以把本地_____传输到云端数据库。

- A). 文档
- B). 硬盘
- C). 视频
- D). 设备

10. Question libre : certains exercices sont-ils particulièrement difficiles pour vous? Si oui, veuillez préciser lesquels et pourquoi.

Annexe 12 – Questionnaire 1 : Profil des participants

1. Votre affiliation universitaire :

- Université de Montréal
 - Université Laval
 - Université du Québec à Montréal
 - Autre (inscrivez le nom de votre université)
-

2. Votre programme d'études (inscrivez le nom du programme) :

3. Votre niveau d'études universitaires :

- 1^{er} cycle universitaire (baccalauréat)
- Cycles supérieurs (DESS, maîtrise, doctorat)
- Année préparatoire aux études universitaires

4. Où avez-vous appris le chinois?

- École secondaire
 - Université
 - Milieu de travail (entreprise)
 - Autre (précisez votre lieu de formation, par exemple, institution de formation privée)
-

5. Combien de temps avez-vous consacré à l'apprentissage du chinois?

- 270 à 400 heures¹ (correspondant à deux trimestres de formation à raison de 3 heures de cours par semaine)
- 400 à 540 heures (trois trimestres de formation)
- Plus de 500 heures (quatre trimestres de formation et plus)

6. Avez-vous reçu une formation en chinois sur objectifs spécifiques?

- Non
 - Oui (précisez la matière : chinois du commerce, chinois scientifique et technique, etc.)
-

¹ Un cours de 3 heures exige 6 heures de travail individuel par semaine

7. Avez-vous étudié ou travaillé dans un ou plusieurs domaines suivants?

- Système informatique
- Sécurité informatique
- Génie logiciel
- Traitement de données (bases de données, infonuagique)
- Réseaux informatiques (réseaux mobiles, internet des objets)

8. Quelle(s) ressource(s) utilisez-vous pour résoudre des difficultés lexicales?

- Dictionnaires bilingues (chinois-français/français-chinois)
- Dictionnaires monolingues chinois
- Matériel didactique (manuel, lexiques et vocabulaires thématiques, etc.)
- Espace de discussion en ligne (forums, blogues, sites web d'apprentissage, etc.)

9. Quelle(s) difficulté(s) éprouvez-vous lorsque vous cherchez à résoudre des questions à l'aide d'une ressource lexicale? (Cochez plusieurs réponses, s'il y a lieu)

- Je ne trouve pas le sens d'un mot inconnu
- Je ne trouve pas le sens d'une combinaison lexicale, même si je connais déjà les mots qui composent celle-ci
- Je ne sais pas quoi chercher pour comprendre une séquence de mots inconnus
- Je ne trouve pas d'exemples d'usage du mot cherché
- Je n'arrive pas à vérifier si l'emploi d'une combinaison lexicale est approprié

10. Avez-vous vécu en Chine?

- Non
- Oui

11. Indiquez la durée de votre séjour en Chine (si vous répondez oui à la question 10) :

- Moins d'un mois
- 1 à 3 mois
- 3 à 6 mois
- 6 à 12 mois
- Plus d'un an

12. Indiquez le motif principal de votre séjour en Chine (si vous répondez oui à la question 10) :

- Études
- Travail
- Tourisme
- Autre

Annexe 13 – Questionnaire 2a : Appréciations et commentaires (Groupe contrôle)

I. ORGANISATION DU TEST

1. Indiquez votre degré d'accord envers chacun des énoncés ci-dessous sur une échelle de 1 à 5 :

1= Fortement en désaccord; 2= Plutôt en désaccord; 3= Plutôt d'accord; 4=Fortement d'accord;

5= Ne s'applique pas. [Cliquez sur la case à cocher correspondante]

| | | 1 | 2 | 3 | 4 | 5 |
|-----|---|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|
| 1.1 | La documentation qui m'a été transmise avant le test était utile. | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| 1.2 | Les objectifs d'évaluation ont été présentés de façon claire. | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| 1.3 | Les modalités du test ont été expliquées de façon claire et précise. | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| 1.4 | Les tâches préparatoires m'ont aidé à bien comprendre le fonctionnement des ressources lexicales proposées. | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| 1.5 | La durée du test était adéquate. | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| 1.6 | J'ai pu obtenir des réponses rapides et pertinentes à mes questions concernant le test. | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| 1.7 | Les supports informatiques utilisés (capsule vidéo, fichier PDF et Word) étaient adéquats. | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |

II. CONTENU DU TEST

2. Indiquez le niveau de difficulté du test sur une échelle de 1 à 4 (évaluez chaque type d'exercices et le test en général) :

1= Très difficile; 2= Plutôt difficile; 3= Plutôt facile; 4= Très facile;

| | | 1 | 2 | 3 | 4 |
|-----|---|--------------------------|--------------------------|--------------------------|--------------------------|
| 2.1 | Distinguer les sens d'une forme lexicale dans différentes CLS | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| 2.2 | Appréhender le sens d'une combinaison lexicale dans un contexte donné | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| 2.3 | Choisir les cooccurrents appropriés pour un mot donné | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| 2.4 | Le niveau de difficulté du test en général | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |

III. RECOURS AUX RESSOURCES LEXICALES

3. Quelle(s) ressource(s) lexicale(s) avez-vous utilisée(s) au cours du test?

4. Dans quelle(s) situation(s) avez-vous consulté une ressource lexicale?

- Pour trouver le(s) sens d'un mot inconnu
 Pour choisir un cooccurrent approprié d'un mot
 Pour trouver des exemples d'usage d'un mot
 Pour d'autres raisons (veuillez préciser)

5. Indiquez, sur une échelle de 1 à 5, le niveau de difficulté que vous avez éprouvée pour trouver des éléments de réponse à chaque type de problèmes lexicaux :

1= Très difficile; 2= Plutôt difficile; 3= Plutôt facile; 4= Très facile;

5= Ne s'applique pas [aucun élément de réponse n'a été trouvé]

| | | 1 | 2 | 3 | 4 | 5 |
|-----|---|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|
| 5.1 | Distinguer les sens d'une forme lexicale dans différentes CLS | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| 5.2 | Appréhender le sens d'une combinaison lexicale dans un contexte donné | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| 5.3 | Choisir les cooccurrents appropriés pour un mot donné | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |

6. De façon globale, dans quelle mesure êtes-vous satisfait(e) des ressources proposées?

1= Très insatisfait(e); 2= Insatisfait(e); 3= Neutre; 4= Satisfait(e); 5= Très satisfait(e)

| | | 1 | 2 | 3 | 4 | 5 |
|-----|--|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|
| 6.1 | Convivialité : les ressources sont-elles faciles à consulter? | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| 6.2 | Qualité du contenu : les connaissances lexicales dans ces ressources sont-elles riches et précises ? | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| 6.3 | Méthode de description : les connaissances lexicales sont-elles présentées de façon claire et structurée? | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| 6.4 | Utilité : les ressources sont-elles utiles pour résoudre les problèmes lexicaux soulevés par les combinaisons lexicales spécialisées? | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| 6.5 | Efficacité pédagogique : les ressources répondent-elles à vos besoins en apprentissage du lexique? | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |

7. Autres commentaires sur le recours aux ressources proposées

Annexe 14 – Questionnaire 2b : Appréciations et commentaires (Groupe expérimental)

I. ORGANISATION DU TEST

1. Indiquez votre degré d'accord envers chacun des énoncés ci-dessous sur une échelle de 1 à 5 :

1= Fortement en désaccord; 2= Plutôt en désaccord; 3= Plutôt d'accord; 4=Fortement d'accord;

5= Ne s'applique pas. [Cliquez sur la case à cocher correspondante]

| | | 1 | 2 | 3 | 4 | 5 |
|-----|---|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|
| 1.1 | La documentation qui m'a été transmise avant le test était utile. | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| 1.2 | Les objectifs d'évaluation ont été présentés de façon claire. | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| 1.3 | Les modalités du test ont été expliquées de façon claire et précise. | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| 1.4 | Les tâches préparatoires m'ont aidé à bien comprendre le fonctionnement des ressources lexicales proposées. | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| 1.5 | La durée du test était adéquate. | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| 1.6 | J'ai pu obtenir des réponses rapides et pertinentes à mes questions concernant le test. | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| 1.7 | Les supports informatiques utilisés (capsule vidéo, fichier PDF et Word) étaient adéquats. | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |

II. CONTENU DU TEST

2. Indiquez le niveau de difficulté du test sur une échelle de 1 à 4 (évaluez chaque type d'exercices et le test en général) :

1= Très difficile; 2= Plutôt difficile; 3= Plutôt facile; 4= Très facile;

| | | 1 | 2 | 3 | 4 |
|-----|---|--------------------------|--------------------------|--------------------------|--------------------------|
| 2.1 | Distinguer les sens d'une forme lexicale dans différentes CLS | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| 2.2 | Appréhender le sens d'une combinaison lexicale dans un contexte donné | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| 2.3 | Choisir les cooccurrents appropriés pour un mot donné | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| 2.4 | Le niveau de difficulté du test en général | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |

III. RECOURS À LA RESSOURCE LEXICALE CHINOINFO

3. Indiquez, sur une échelle de 1 à 5, le niveau de difficulté que vous avez éprouvée pour trouver des éléments de réponse à chaque type de problèmes lexicaux à l'aide de CHINOINFO :

1= Très difficile; 2= Plutôt difficile; 3= Plutôt facile; 4= Très facile;
5= Ne s'applique pas [aucun élément de réponse n'a été trouvé]

| | | 1 | 2 | 3 | 4 | 5 |
|-----|---|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|
| 3.1 | Distinguer les sens d'une forme lexicale dans différentes CLS | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| 3.2 | Appréhender le sens d'une combinaison lexicale dans un contexte donné | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| 3.3 | Choisir les cooccurrents appropriés pour un mot donné | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |

4. De façon globale, dans quelle mesure êtes-vous satisfait(e) de cette ressource lexicale ?

1= Très insatisfait(e); 2= Insatisfait(e); 3= Neutre; 4= Satisfait(e); 5= Très satisfait(e)

| | | 1 | 2 | 3 | 4 | 5 |
|-----|--|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|
| 4.1 | Convivialité : la ressource est-elle facile à consulter? | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| 4.2 | Qualité du contenu : les connaissances lexicales dans cette ressource sont-elles riches et précises ? | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| 4.3 | Méthode de description : les connaissances lexicales sont-elles présentées de façon claire et structurée? | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| 4.4 | Utilité : cette ressource est-elle utile pour résoudre les problèmes lexicaux soulevés par les combinaisons lexicales spécialisées? | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| 4.5 | Efficacité pédagogique : cette ressource est-elle adaptée à vos besoins en apprentissage du lexique? | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |

5. Qu'est-ce que vous aimez le plus et le moins dans cette ressource? Quelles sont vos propositions pour l'améliorer ?

6. Serez-vous intéressé(e) à utiliser une ressource didactique intégrant la méthode de description lexicale proposée dans CHINOINFO?

- Oui
 Sans opinion
 Non

IV. RECOURS À D'AUTRES OUTILS DE RÉFÉRENCE

7. Avez-vous utilisé d'autres outils de référence au cours du test?

- Non
 Oui (Si vous répondez oui, passez aux questions 8-11)

8. Indiquez les outils que vous avez utilisés au cours du test.

9. Dans quelle(s) situation(s) avez-vous utilisé ces outils de référence?

- Pour trouver le(s) sens d'un mot inconnu
 Pour choisir un cooccurrent approprié d'un mot
 Pour trouver des exemples d'usage d'un mot
 Pour d'autres raisons (veuillez préciser)

10. Indiquez, sur une échelle de 1 à 5, le niveau de difficulté que vous avez éprouvée pour trouver des éléments de réponse à chaque type de problèmes lexicaux à l'aide de ces outils :

1= Très difficile; 2= Plutôt difficile; 3= Plutôt facile; 4= Très facile;
5= Ne s'applique pas [aucun élément de réponse n'a été trouvé]

| | | 1 | 2 | 3 | 4 | 5 |
|------|---|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|
| 10.1 | Distinguer les sens d'une forme lexicale dans différentes CLS | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| 10.2 | Appréhender le sens d'une combinaison lexicale dans un contexte donné | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| 10.3 | Choisir les cooccurrents appropriés pour un mot donné | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |

11. Autres commentaires sur le recours à ces outils de référence

Annexe 15 – Attestation d’approbation d’éthique

Projet : CERAH-2021-142-D_Description du lexique spécialisé chinois pour locuteurs non sinophones : apport de la Lexicologie explicative et combiantoire pour l'acquisition de compétences lexicales

Chercheuse étudiante : Zhiwei Han – candidat au doctorat – Département de linguistique et de traduction – Faculté des arts et des sciences – Université de Montréal

Directrice : Marie-Claude L’Homme – professeure titulaire – Département de linguistique et de traduction – Faculté des arts et des sciences – Université de Montréal

Financement: Non-financé

Bonjour,

Suite à l’approbation conditionnelle du projet de recherche cité en objet, vous avez transmis le 23 décembre 2021 une réponse au Comité d’éthique de la recherche en arts et humanités (CERAH) de l’Université de Montréal. Au nom du Comité, je vous remercie pour les précisions apportées dans vos réponses ainsi que les modifications apportées aux documents. Suite à l’évaluation de cette réponse et des documents transmis, le tout ayant été jugé satisfaisant, j’ai le plaisir de vous informer que le (CERAH) considère que **le projet de recherche susmentionné répond aux normes en vigueur au chapitre de l’éthique de la recherche** et est en conséquence approuvé.

Cette approbation éthique est **valide pour un an**, à compter de la date des présentes **jusqu’au 12 janvier 2023**, et pourra être renouvelée de la manière prévue ci-après aux mesures de suivi éthique.

Mesures de suivi éthique continu

Le CERAH demeure responsable de l’acceptabilité éthique des activités de recherche menées sous son autorité. Une fois l’approbation éthique initiale obtenue, une évaluation éthique minimalement annuelle est requise. L’évaluation éthique continue sera effectuée par le CERAH à partir des notifications qui lui seront transmises par l’équipe de recherche chercheur pendant le déroulement de la recherche. À cette fin, le CERAH fixe les mesures suivantes de suivi éthique continu de votre projet de recherche :

- La soumission d’un **rapport d’étape annuel**, à soumettre un mois avant l’échéance de la date d’approbation afin de renouveler l’approbation éthique.
- La soumission de toute **modification au projet de recherche qui touche les participants**; une modification ne peut être mise en œuvre sans l’approbation du CERAH.
- La soumission dans les meilleurs délais d’un rapport de tout **évènement indésirable, de tout accident ou de tout incident** lié à la réalisation du projet de recherche.
- La soumission d’un rapport sur toute **déviaton au protocole** de recherche susceptible d’augmenter le niveau de risque ou susceptibles d’influer sur le bien-être du participant ou son consentement.
- La notification de toute **cessation prématurée, interruption temporaire ou suspension**, qu’elle soit temporaire ou permanente.
- La soumission d’un **rapport de fin de projet**.

Ces notifications doivent être transmises au CERAH en complétant le questionnaire de suivi disponible sur la [page web du CERAH](#) à la section « Modifications envisagées à un projet de recherche » et en le retournant par courriel à suivi-ethique@umontreal.ca avec la mention « Suivi éthique – [no d’approbation éthique] – date de complétion » dans le champ « Objet »

Tout défaut de respecter une de ces mesures de suivi éthique pourrait résulter en une suspension ou une révocation de l’approbation.

Le CERAH de l’Université de Montréal est désigné par le ministre de la Santé et des Services Sociaux aux fins de l’application de l’article 21 du Code civil du Québec. Il exerce ses activités en conformité avec la *Politique sur la recherche avec des êtres humains* (60.1) de l’Université de Montréal ainsi que l’Énoncé de politique des trois conseils (EPTC). Il suit également les normes et règlements applicables au Québec et au Canada.

Le présent courriel est la décision officielle du CERAH.

Annexe 16 – Formulaire d’information et de consentement



FORMULAIRE D’INFORMATION ET DE CONSENTEMENT

« Description du lexique spécialisé chinois pour locuteurs non sinophones : apport de la Lexicologie explicative et combinatoire pour l’acquisition de compétences lexicales »

Qui dirige ce projet?

Moi, Zhiwei Han. Je suis étudiante au doctorat à l’Université de Montréal au Département de linguistique et de traduction. Ma directrice de recherche est Marie-Claude L’Homme, professeure au Département de linguistique et de traduction.

Décrivez-moi ce projet

L’apprentissage du lexique spécialisé chinois est un chemin semé d’embûches. Parmi les nombreux défis à relever, l’acquisition des combinaisons lexicales soulèvent des difficultés de nature variée pour les apprenants non sinophones. On recense, toutefois, peu de propositions pour résoudre ces difficultés en didactique du chinois sur objectifs spécifiques.

Mon projet vise à concevoir une méthode de description des combinaisons lexicales spécialisées (CLS), afin de favoriser leur acquisition par les locuteurs non sinophones. Je m’attache à résoudre, plus précisément, les difficultés syntaxico-sémantiques soulevées par les CLS. La méthode sera appliquée au domaine de l’informatique. Elle aboutira à l’élaboration d’une ressource didactique bilingue destinée aux apprenants francophones. L’efficacité pédagogique de cette méthode sera évaluée au moyen d’une expérimentation comparative.

Si je participe, qu’est-ce que j’aurai à faire?

Vous aurez à réaliser un test lexical en ligne. Ce test vise à évaluer vos habiletés à résoudre trois catégories de problèmes lexicaux soulevés par les CLS: 1) distinguer les différents sens associés à une forme lexicale dans des CLS; 2) identifier, dans un contexte donné, le sens d’une combinaison lexicale qui est syntaxiquement ambiguë; 3) choisir les cooccurrents appropriés pour un mot donné.

Avant de passer le test, vous aurez à réaliser deux tâches préparatoires durant d’environ 15 minutes : 1). Visionner une capsule vidéo (d’une durée de 10 minutes) portant sur le déroulement du test; 2). Prendre connaissance d’un document descriptif portant sur les outils de référence mis à votre disposition. Cette tâche dure d’environ 5 minutes.

La passation du test se déroulera en mode synchrone via Zoom. La séance synchrone ne sera pas enregistrée. Vous ne serez pas obligés d’ouvrir votre caméra durant le test. La durée du test lexical est de 30 minutes.

Une fois le test réalisé, vous remplirez deux questionnaires visant à recueillir : 1) les renseignements sur votre profil d’apprentissage; 2) vos commentaires sur le test et les outils de référence mis à votre disposition. La durée du sondage est de 10 minutes.

Y a-t-il des risques ou des avantages à participer à cette recherche?

Il n’y a pas de risque particulier à participer à ce projet. Il existe, toutefois, des limites à la confidentialité dans le cas où le test lexical se fait en ligne via Zoom. Dans ce contexte, il n’est pas possible de garantir une confidentialité totale du test pour les raisons mentionnées ci-après. Premièrement, le serveur de Zoom pourrait être situé aux États-Unis, et en conséquence soumis aux dispositions légales américaines qui autorisent les services de sécurité américains à accéder aux données informatiques détenues par les particuliers et les entreprises, sans autorisation préalable et sans en informer les utilisateurs. De plus, en passant un test en ligne, il est possible que quelqu’un

de votre entourage voie vos réponses. Il est donc important de bien planifier la passation du test dans un endroit qui vous assure un minimum d'intimité.

Il n'y a pas d'avantage particulier à participer à ce projet. Votre contribution pourrait cependant aider à bonifier l'enseignement du lexique spécialisé chinois auprès des locuteurs non natifs.

D'ailleurs, votre participation vous donnera l'occasion de pratiquer une nouvelle activité pédagogique et d'accéder à un nombre de ressources lexicales chinois-français, ce qui pourrait favoriser votre apprentissage du chinois.

Que ferez-vous avec mes réponses?

Je vais analyser l'ensemble de vos réponses. Je vais corriger le test lexical et examiner les erreurs repérées. Vos réponses aux questionnaires seront aussi analysées avec minutie. Les résultats obtenus me permettront d'évaluer l'efficacité de la méthode proposée et d'adapter celle-ci aux besoins spécifiques des apprenants. Les résultats feront partie de ma thèse de doctorat.

Est-ce que mes données personnelles seront protégées?

Oui! Aucune information permettant de vous identifier d'une façon ou d'une autre ne sera publiée. De plus, les renseignements recueillis seront conservés de manière confidentielle. Les fichiers informatiques seront enregistrés sur l'ordinateur sécurisé et seules ma directrice de recherche et moi-même en prendrons connaissance. Toute information permettant de vous identifier sera détruite 7 ans après la fin du projet. Seules les réponses transcrites seront conservées après cette période, sans aucune information concernant les personnes qui me les auront fournies.

Les résultats généraux de mon projet pourraient être utilisés dans des publications ou des communications, mais toujours de façon anonyme, c'est-à-dire sans jamais nommer ou identifier les participants.

Est-ce que je suis obligé de répondre à toutes les questions et d'aller jusqu'au bout?

Non! Vous pouvez décider de ne pas répondre à une ou plusieurs questions. Vous pouvez aussi à tout moment décider que vous ne voulez plus participer à l'expérience et que vous abandonnez le projet. Dans ce cas, vous pourrez même me demander de ne pas utiliser vos réponses pour ma recherche et de les détruire. Cependant, une fois que le processus de publication des données sera mis en route, je ne pourrai pas détruire les analyses et les résultats portant sur vos réponses, mais aucune information permettant de vous identifier ne sera publiée.

À qui puis-je parler si j'ai des questions durant l'étude?

Pour toute question, vous pouvez me contacter au numéro suivant 418-948-1801 ou à l'adresse suivante zhiwei.han@umontreal.ca. Plusieurs ressources sont à votre disposition.

Ce projet a été approuvé par le Comité d'éthique de la recherche en arts et en sciences de l'Université de Montréal. Pour toute préoccupation sur vos droits ou sur les responsabilités des chercheurs concernant votre participation à ce projet, vous pouvez contacter le comité par courriel à l'adresse cerah@umontreal.ca ou encore consulter le site Web : <http://recherche.umontreal.ca/participants>.

Si vous avez des plaintes concernant votre participation à cette recherche, vous pouvez communiquer avec l'ombudsman (c'est un « protecteur des citoyens ») de l'Université de Montréal, au numéro de téléphone 514-343-2100 ou à l'adresse courriel ombudsman@umontreal.ca (l'ombudsman accepte les appels à frais virés).

Comment puis-je donner mon accord pour participer à l'étude ?

En signant ce formulaire de consentement et en me le remettant. Je vous laisserai une copie du formulaire que vous pourrez conserver afin de vous y référer au besoin.

CONSENTEMENT

Déclaration du participant

- Je comprends que je peux prendre mon temps pour réfléchir avant de donner mon accord ou non à ma participation.
- Je peux poser des questions à l'équipe de recherche et exiger des réponses satisfaisantes.
- Je comprends qu'en participant à ce projet de recherche, je ne renonce à aucun de mes droits ni ne dégage les chercheurs de leurs responsabilités.
- J'ai pris connaissance du présent formulaire d'information et de consentement et j'accepte de participer au projet de recherche.

Signature du participant : _____ Date : _____

Nom : _____ Prénom : _____

Engagement de la chercheuse

J'ai expliqué les conditions de participation au projet de recherche au participant. J'ai répondu au meilleur de ma connaissance aux questions posées et je me suis assuré de la compréhension du participant. Je m'engage, avec l'équipe de recherche, à respecter ce qui a été convenu au présent formulaire d'information et de consentement.

Signature de la chercheuse : _____ Date : _____

Nom : Han _____ Prénom : Zhiwei _____

Annexe 17 – Renseignements sur le test lexical (présentation narrée destinée au GC)



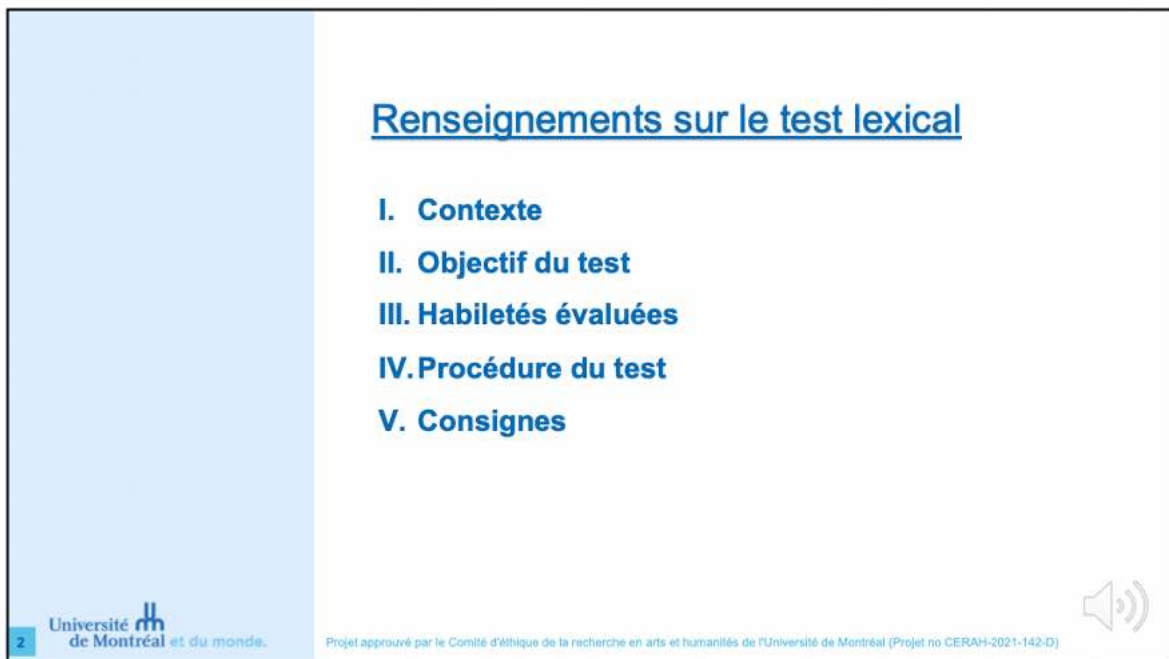
Projet de recherche
Description du lexique spécialisé chinois pour locuteurs non sinophones :
apport de la Lexicologie explicative et combinatoire pour l'acquisition de compétences lexicales

Renseignements sur le test lexical

Chercheuse étudiante : Zhiwei HAN
Le 20 avril 2022



Université  de Montréal et du monde 

1



Renseignements sur le test lexical

- I. Contexte**
- II. Objectif du test**
- III. Habiletés évaluées**
- IV. Procédure du test**
- V. Consignes**

Université  de Montréal et du monde. 

Projet approuvé par le Comité d'éthique de la recherche en arts et humanités de l'Université de Montréal (Projet no CERAH-2021-142-D)

2

I. Contexte

Pourquoi ce test ?

Le test lexical s'inscrit dans le cadre d'un projet de thèse visant à concevoir une méthode de description des **combinaisons lexicales spécialisées (CLS)**, afin de favoriser leur acquisition par les apprenants non natifs.

3

Projet approuvé par le Comité d'éthique de la recherche en arts et humanités de l'Université de Montréal (Projet no CERAH-2021-142-D)

3

II. Objectif du test

Quel est l'objectif de ce test ?

Le test lexical a pour objectif d'évaluer les habiletés des participants à résoudre les difficultés syntaxico-sémantiques soulevées par les CLS.

Quelles habiletés seront évaluées ?

4

Projet approuvé par le Comité d'éthique de la recherche en arts et humanités de l'Université de Montréal (Projet no CERAH-2021-142-D)

4

III. Habiletés évaluées

Types de problèmes lexicaux à résoudre

- ❖ Distinguer les sens associés à une forme lexicale dans différentes CLS
- ❖ Identifier, dans un contexte donné, le sens d'une combinaison lexicale qui est syntaxiquement ambiguë
- ❖ Choisir les cooccurents appropriés pour un mot donné.

Nombre de questions

- ❖ Le test comprend 10 questions :
 - 6 questions à réponse courte
 - 3 questions à choix multiples
 - 1 question libre

5

Projet approuvé par le Comité d'éthique de la recherche en arts et humanités de l'Université de Montréal (Projet no CERAH-2021-142-D)

5

IV. Procédure du test

Une semaine avant le test



À la date et l'heure prévues du test



À la fin du test

La chercheuse envoie aux participants, par courriel :

- ⇒ la capsule vidéo présentant le test lexical
- ⇒ la liste d'outils de référence accompagnée d'un mode d'emploi (format Word)
- ⇒ le lien Zoom pour participer au test

La chercheuse envoie aux participants, via Zoom chat, trois documents Word :

- ⇒ le test lexical
- ⇒ le questionnaire de sondage sur le profil des participants
- ⇒ le questionnaire d'appréciations et commentaires sur le test

Les participants envoient à la chercheuse, via Zoom chat, trois documents Word remplis :

- ⇒ le test lexical
- ⇒ le questionnaire de sondage sur le profil des participants
- ⇒ le questionnaire d'appréciations et commentaires sur le test

6

Projet approuvé par le Comité d'éthique de la recherche en arts et humanités de l'Université de Montréal (Projet no CERAH-2021-142-D)

6

V. Consignes

- 1). Les participants seront autorisés à utiliser les outils de référence proposés et d'autre matériel didactique du chinois, y compris des manuels et des notes de cours.
- 2). La passation du test se déroulera en mode synchrone via Zoom. La séance ne sera pas enregistrée. Les participants n'aurons pas à ouvrir leur caméra durant le test.
- 3). Une fois le test commence, toute communication sera effectuée par chat dans le logiciel Zoom.
- 4). La durée du test lexical est d'environ 40 minutes : 30 minutes pour compléter le test et 10 minutes pour remplir les questionnaires de sondage.
- 5). La chercheuse sera disponible pour répondre à des questions au tout long du test. La durée maximale de la session Zoom est de 60 minutes.

7

Projet approuvé par le Comité d'éthique de la recherche en arts et humanités de l'Université de Montréal (Projet no CERAH-2021-142-D)

7

8

谢谢！
Merci pour votre participation!

Pour toute question, veuillez
me contacter par courriel :

Zhiwei HAN

zhiwei.han@umontreal.ca

Projet approuvé par le Comité d'éthique de la recherche en arts et humanités de l'Université de Montréal (Projet no CERAH-2021-142-D)

8

Annexe 18 – Liste d’outils de référence proposés au GC

Description du lexique spécialisé chinois pour locuteurs non sinophones : apport de la Lexicologie explicative et combinatoire pour l’acquisition de compétences lexicales

Liste d’outils de référence

| Titre abrégé | Référence | Description |
|-----------------------------------|--|--|
| Dictionnaire Chine-Nouvelle | VerticalScope Inc. (2022). <i>Dictionnaire chinois français</i> . Chine-Nouvelle. https://www.chine-nouvelle.com/outils/dictionnaire.html | Dictionnaire chinois-français |
| Dictionnaire Chine-informations | Chine Informations. (2022). <i>Apprendre le chinois: dictionnaire chinois-français</i> . Chine-informations. https://chine.in/mandarin/dictionnaire/ | Dictionnaire chinois-français |
| LINE Chinese-English Dictionary | NAVER Corp. (2022). <i>LINE Dictionary : Chinese-English</i> . Dict.naver.com. https://dict.naver.com/linedict/#/cnen/home | Dictionnaire chinois-anglais |
| Larousse Français-chinois | Éditions Larousse. (2022). <i>Dictionnaire Français-chinois</i> . Larousse.fr. https://www.larousse.fr/dictionnaires/francais-chinois | Dictionnaire français-chinois |
| MDBG Chinese-English Dictionary | <i>MDBG Chinese-English dictionary</i> . (2022). MDBG.NET. https://www.mdbg.net/chinese/dictionary?page=worddict&wdrst=0&wdgb= | Dictionnaire chinois-anglais |
| Banque de terminologie Termonline | China National Committee for Terminology in Science and Technology. (2022). <i>Termonline</i> . Termonline: a public terminology knowledge service platform. https://www.termonline.cn/index | Banque de terminologie chinois-anglais |

Chercheuse étudiante : Zhiwei HAN (2022-04-21)

Projet approuvé par le Comité d’éthique de la recherche en arts et humanités de l’Université de Montréal (Projet no CERAH-2021-142-D)

1

Guide d’utilisation : comment consulter ces dictionnaires ?

Cliquer pour visionner les vidéos démonstratives

[Dictionnaire Chine-Nouvelle](#)

The screenshot shows the homepage of Chine-Nouvelle.com. At the top, there is a search bar with the text 'dictionnaire fr-cn' and a 'site' button. Below the search bar, there are navigation links for 'Accueil', 'La Chine', 'La Saison', 'Dico', 'Outils', 'Préparer chinois', 'Livre', and 'Forum'. A prominent advertisement for Samsung Galaxy Tab S7 FE is visible. The main content area is titled 'Dictionnaire chinois français' and includes a search bar with the word 'Expression' entered. Below the search bar, there is a list of recent searches: 'Étonné', 'motifs', 'motif', 'chou', '嘉', '紫', '立', '在', '島', 'milan'. At the bottom, there is a section for 'Tous les dictionnaires' with links to various language pairs: Dictionnaire chinois français (Français), Chinese-English dictionary (English), Dicionario chino-español (Español), Chinesisch-Deutsches Wörterbuch (Deutsch), Dicionario cinese-italiano (Italiano), and Dicionário chinês-português (Português).

Chercheuse étudiante : Zhiwei HAN (2022-04-21)

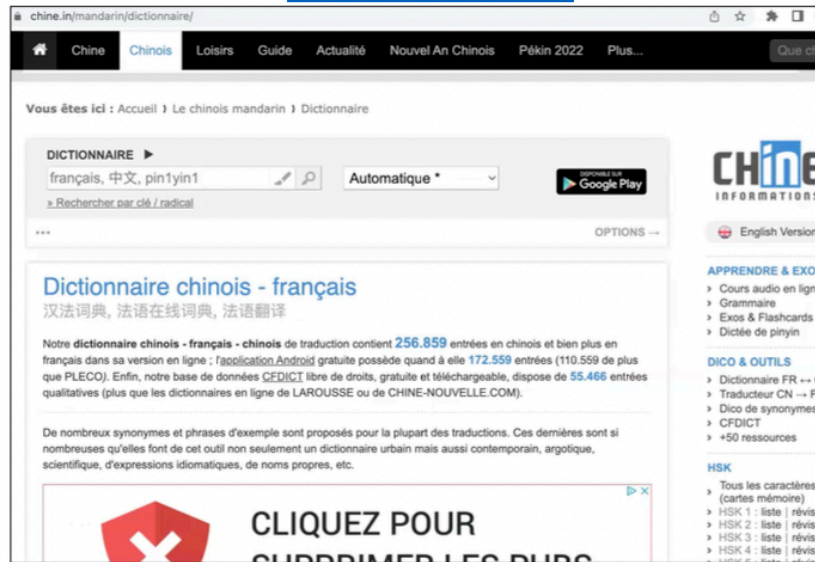
Projet approuvé par le Comité d’éthique de la recherche en arts et humanités de l’Université de Montréal (Projet no CERAH-2021-142-D)

2

Guide d'utilisation : comment consulter ces dictionnaires ?

Cliquer pour visionner les vidéos démonstratives

[Dictionnaire Chine-informations](#)



Chercheuse étudiante : Zhiwei HAN (2022-04-21)

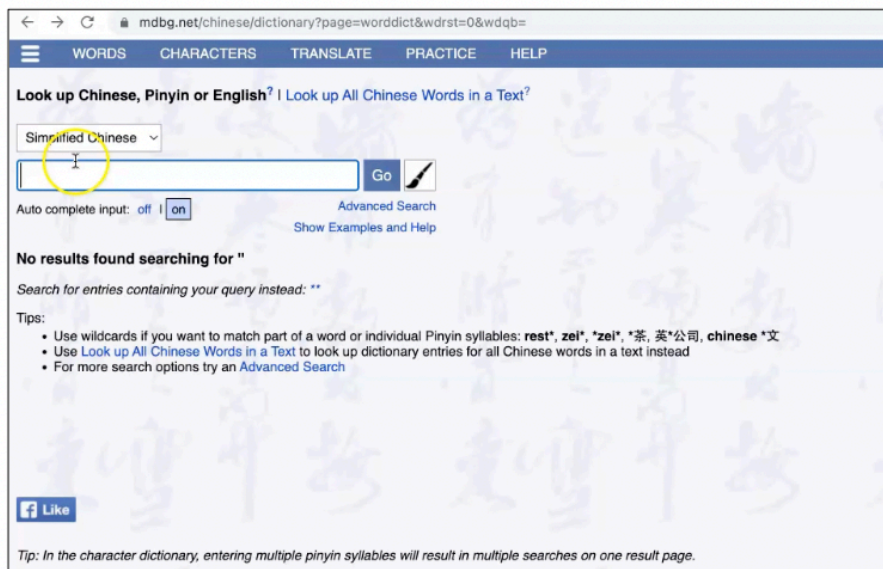
Projet approuvé par le Comité d'éthique de la recherche en arts et humanités de l'Université de Montréal (Projet no CERAH-2021-142-D)

3

Guide d'utilisation : comment consulter ces dictionnaires ?

Cliquer pour visionner les vidéos démonstratives

[MDBG Chinese-English Dictionary](#)




Chercheuse étudiante : Zhiwei HAN (2022-04-21)

Projet approuvé par le Comité d'éthique de la recherche en arts et humanités de l'Université de Montréal (Projet no CERAH-2021-142-D)

4

Annexe 19 – Renseignements sur le test lexical (présentation narrée destinée au GE)



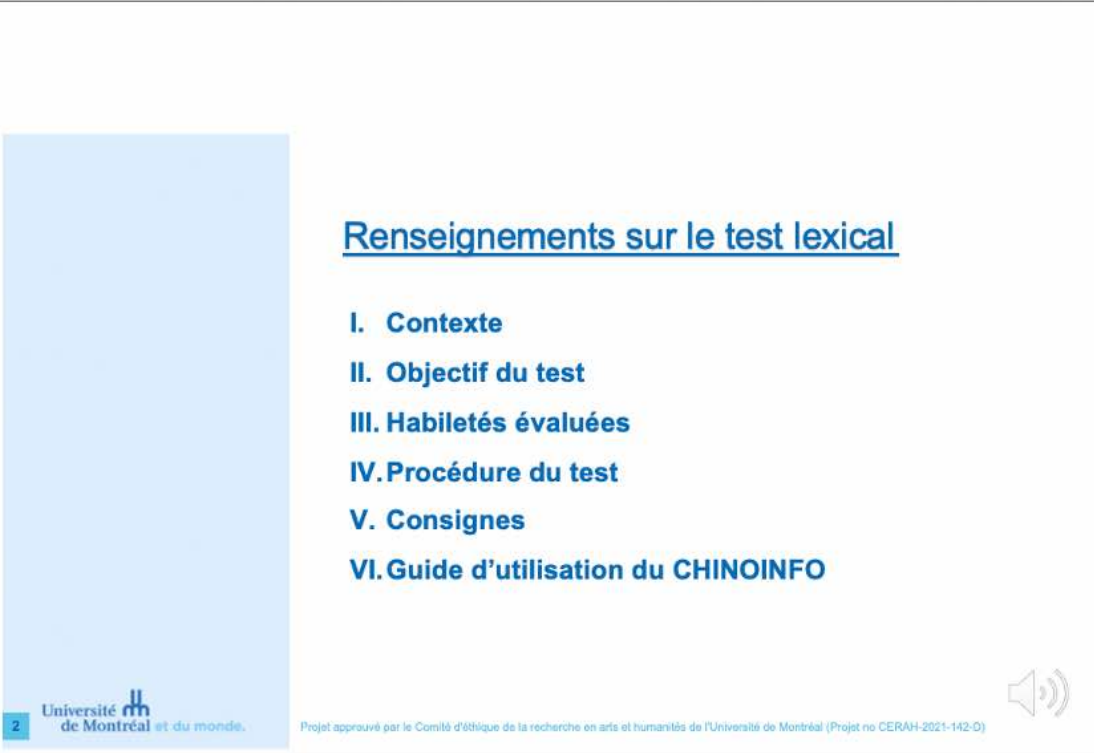
Projet de recherche

Description du lexique spécialisé chinois pour locuteurs non sinophones :
apport de la Lexicologie explicative et combinatoire pour l'acquisition de compétences lexicales

Renseignements sur le test lexical

Chercheuse étudiante : Zhiwei HAN
Le 20 avril 2022

Université de Montréal et du monde



Renseignements sur le test lexical

- I. Contexte
- II. Objectif du test
- III. Habiletés évaluées
- IV. Procédure du test
- V. Consignes
- VI. Guide d'utilisation du CHINOINFO

2 Université de Montréal et du monde. Projet approuvé par le Comité d'éthique de la recherche en arts et humanités de l'Université de Montréal (Projet no CERAH-2021-142-D)

I. Contexte

Pourquoi ce test ?

Le test lexical s'inscrit dans le cadre d'un projet de thèse visant à concevoir une méthode de description des **combinaisons lexicales spécialisées (CLS)**, afin de favoriser leur acquisition par les apprenants non natifs.

3

Projet approuvé par le Comité d'éthique de la recherche en arts et humanités de l'Université de Montréal (Projet no CERAH-2021-142-D)

Université  de Montréal  et du monde.

II. Objectif du test

Quel est l'objectif de ce test ?

Le test lexical a pour objectif d'évaluer les habiletés des participants à résoudre les difficultés syntaxico-sémantiques soulevées par les CLS.

Quelles habiletés seront évaluées ?

4

Projet approuvé par le Comité d'éthique de la recherche en arts et humanités de l'Université de Montréal (Projet no CERAH-2021-142-D)

Université  de Montréal  et du monde.

III. Habiletés évaluées

Types de problèmes lexicaux à résoudre

- ❖ Distinguer les sens associés à une forme lexicale dans différentes CLS
- ❖ Identifier, dans un contexte donné, le sens d'une combinaison lexicale qui est syntaxiquement ambiguë
- ❖ Choisir les cooccurrents appropriés pour un mot donné.

Nombre de questions

- ❖ Le test comprend 10 questions :
 - 6 questions à réponse courte
 - 3 questions à choix multiples
 - 1 question libre

5

Projet approuvé par le Comité d'éthique de la recherche en arts et humanités de l'Université de Montréal (Projet no CERAH-2021-142-D)

IV. Procédure du test

Une semaine avant le test



À la date et l'heure prévues du test



À la fin du test

La chercheuse envoie aux participants, par courriel :

- ➔ la capsule vidéo présentant le test lexical
- ➔ la liste d'outils de référence accompagnée d'un mode d'emploi (format Word)
- ➔ le lien Zoom pour participer au test

La chercheuse envoie aux participants, via Zoom chat, trois documents Word :

- ➔ le test lexical
- ➔ le questionnaire de sondage sur le profil des participants
- ➔ le questionnaire d'appréciations et commentaires sur le test

Les participants envoient à la chercheuse, via Zoom chat, trois documents Word remplis :

- ➔ le test lexical
- ➔ le questionnaire de sondage sur le profil des participants
- ➔ le questionnaire d'appréciations et commentaires sur le test

6

Projet approuvé par le Comité d'éthique de la recherche en arts et humanités de l'Université de Montréal (Projet no CERAH-2021-142-D)

V. Consignes

- 1). Les participants seront autorisés à utiliser les outils de référence proposés et d'autre matériel didactique du chinois, y compris des manuels et des notes de cours.
- 2). La passation du test se déroulera en mode synchrone via Zoom. La séance ne sera pas enregistrée. Les participants n'aurons pas à ouvrir leur caméra durant le test.
- 3). Une fois le test commence, toute communication sera effectuée par chat dans le logiciel Zoom.
- 4). La durée du test lexical est d'environ 40 minutes : 30 minutes pour compléter le test et 10 minutes pour remplir les questionnaires de sondage.
- 5). La chercheuse sera disponible pour répondre à des questions au tout long du test. La durée maximale de la session Zoom est de 60 minutes.

7

Projet approuvé par le Comité d'éthique de la recherche en arts et humanités de l'Université de Montréal (Projet no CERAH-2021-142-D)

VI. Guide d'utilisation du CHINOINFO

Élaboration du CHINOINFO

❖ Le CHINOINFO



Langue

Chinois-français



Domaine ciblé

L'informatique et Internet



Locuteurs ciblés

Apprenants francophones

❖ Contexte de recherche

Projet de thèse de doctorat en traduction (option terminologie)

Chercheuse: - doctorante (Université de Montréal)

- chargée de cours du chinois (Université Laval)

8



ChinoInfo

Sélectionnez une entrée en fonction du hanzhi pointé :

安 除 发 共 连 入 试 同 移 在
保 储 访 击 格 款 手 下 盘 载
备 传 份 加 密 删 输 线 应 制
步 存 复 检 迁 设 送 享 用 置
测 动 更 件 侵 时 同 新 延 装
程 恶 攻 接 全 实 网 行 运 自

Le CHINOINFO est un dictionnaire spécialisé bilingue destiné aux apprenants francophones du chinois. Ce dictionnaire répertorie 91 termes fondamentaux du domaine de l'informatique. Tous les renseignements terminologiques sont extraits à partir d'un corpus chinois comptant 341 676 mots.

L'élaboration du CHINOINFO s'inscrit dans le cadre d'un projet de thèse visant à concevoir une méthode pour décrire le lexique spécialisé chinois, plus précisément, les combinaisons lexicales spécialisées (CLS). Nous nous attachons à explorer de quelle manière la représentation sémantique des CLS favorise leur acquisition par les locuteurs non-linguistes. L'efficacité de la méthode mise au point sera évaluée au moyen d'une expérimentation.

Ce projet est réalisé par **Zhiwei Han**, étudiante au doctorat à l'Université de Montréal, sous la supervision de **Marie-Claude L'Homme**, professeure au Département de linguistique et de traduction de l'Université de Montréal. Nous remercions chaleureusement **Benwei Reibichang**, membre de l'Observatoire de linguistique Sens-Texte (OLST), pour la généreuse contribution qu'il a apportée à la mise en ligne du CHINOINFO.

Nomenclature (91 entrées) :

| trier par lexie | trier par catégorie | trier par glose |
|-----------------------------------|---------------------------------------|--------------------------------|
| 远程 2 (adv.) [à distance] | 加密 2 (nom) [cryptage] | 安装 1 (v.tr.) [installer 2] |
| 访问 1 (v.tr.) [accéder 1] | 加密 3 (adv.) [crypter] | 入侵 2 (nom) [invasion] |
| 访问 2 (v.tr.) [accéder 2] | 加密 1 (v.tr.) [crypter] | 软件 1 (nom) [logiciel 1] |
| 访问 3 (nom) [accès 1] | 恶意 2 (adv.) [de manière malveillante] | 软件 2 (nom) [logiciel 2] |
| 访问 4 (nom) [accès 2] | 检测 1 (v.tr.) [détecter 1] | 恶意 1 (adv.) [malveillant 1] |
| 移动 2 (adv.) [amovible] | 检测 3 (nom) [détaction 1] | 手动 1 (adv.) [manuel] |
| 应用 1 (nom) [application 1] | 远程 1 (adv.) [à distance] | 手动 2 (adv.) [manuellement] |
| 应用 2 (nom) [application 2] | 网络 3 (adv.) [en ligne] | 更新 1 (v.tr.) [mettre à jour] |
| 应用 4 (v.tr.) [appliquer 2] | 在线 1 (adv.) [en ligne 1] | 迁移 2 (nom) [migration] |
| 攻击 2 (nom) [attaque informatique] | 在线 2 (adv.) [en ligne 2] | 更新 2 (nom) [mise à jour 1.1] |
| 攻击 1 (v.tr.) [attaquer 1] | 实时 1 (adv.) [en temps réel 1] | 更新 3 (nom) [mise à jour 1.2] |
| 启动 1 (adv.) [à l'automatique 1] | 实时 2 (adv.) [en temps réel 2] | 移动 1 (adv.) [mobile 1] |
| 启动 2 (adv.) [à l'automatique 2] | 保存 2 (nom) [enregistrement 1.1] | 同步 4 (adv.) [synchroniquement] |
| 设置 2 (nom) [configuration 1.1] | 保存 1 (v.tr.) [enregistrer 1] | 共享 3 (adv.) [partager] |
| 设置 3 (nom) [configuration 1.2] | 发送 3 (nom) [envoi 1] | 共享 2 (nom) [partage 1] |
| | | 共享 1 (v.tr.) [partager 1] |
| | | 备份 1 (nom) [sauvegarde] |
| | | 备份 2 (nom) [sauvegarde 2] |
| | | 备份 3 (nom) [sauvegarde 3] |
| | | 备份 4 (adv.) [à l'automatique] |
| | | 测试 1 (v.tr.) [tester] |
| | | 测试 2 (nom) [test] |
| | | 传输 1 (v.tr.) [transférer 1] |
| | | 传输 2 (nom) [transfert 1] |
| | | 存储 1 (v.tr.) [stocker] |
| | | 检测 4 (nom) [vérification] |
| | | 连接 1 (v.tr.) [se connecter 1] |
| | | 连接 2 (v.tr.) [connecter] |
| | | 连接 3 (nom) [connexion 1.1] |
| | | 连接 4 (nom) [connexion 2.1] |
| | | 迁移 1 (v.tr.) [faire migrer] |
| | | 迁移 2 (nom) [migration] |
| | | 入侵 1 (v.tr.) [s'introduire] |
| | | 入侵 2 (nom) [invasion] |
| | | 软件 1 (nom) [logiciel 1] |
| | | 软件 2 (nom) [logiciel 2] |
| | | 删除 1 (v.tr.) [supprimer 1] |
| | | 删除 2 (nom) [suppression 1] |
| | | 删除 1 (v.tr.) [configurer 1] |
| | | 设置 2 (nom) [configuration 1.1] |
| | | 设置 3 (nom) [configuration 1.2] |
| | | 同步 4 (adv.) [synchroniquement] |
| | | 网络 1 (nom) [réseau 1] |
| | | 网络 2 (nom) [réseau 2] |
| | | 网络 3 (adv.) [en ligne] |
| | | 下载 1 (v.tr.) [télécharger] |
| | | 移动 2 (nom) [déchargement] |
| | | 移动 1 (adv.) [mobile 1] |
| | | 移动 2 (adv.) [amovible] |
| | | 应用 1 (nom) [application 1] |
| | | 应用 2 (nom) [application 2] |
| | | 应用 3 (nom) [application 3] |
| | | 应用 4 (v.tr.) [appliquer 2] |
| | | 远程 1 (adv.) [à distance] |
| | | 设置 1 (v.tr.) [configurer 1] |
| | | 运行 1 (v.tr.) [faire exécuter] |
| | | 运行 2 (v.tr.) [faire exécuter] |

VI. Guide d'utilisation du CHINOINFO

Le CHINOINFO

<http://olst.ling.umontreal.ca/chinoinfo/>

⇒ Nomenclature :

91 termes fondamentaux du domaine de l'informatique regroupés en 33 vocables (termes ayant la forme lexicale)

⇒ Sources de données :

un corpus spécialisé chinois comptant 341 676 mots

Comment rechercher des termes dans le CHINOINFO?

⇒ Par index

La nomenclature est triée de trois façons différentes (par lexie/ partie du discours/glose)

⇒ Par caractère

Un module est conçu pour rechercher les termes en fonction des leurs éléments constitutifs



VI. Guide d'utilisation du CHINOINFO

Quels renseignements peut-on trouver dans un article?

⇒ Entrée
Numéro d'acception
Partie du discours
Mention du domaine
Équivalent en français

⇒ Définition

⇒ Structure actancielle

⇒ Contextes

⇒ Termes reliés
Classement par lien
sémantique

⇒ Combinaisons lexicales
Regroupement par structure
syntaxique
Classe sémantique
Renvoi à d'autres articles
Glose
Équivalent en français



ChinoInfo

Sélectionnez une entrée en fonction du hanzhi pointé :

安 除 发 共 连 入 试 同 移 在
保 储 访 击 格 款 手 下 盘 载
备 传 份 加 密 删 输 线 应 制
步 存 复 检 迁 设 送 享 用 置
测 动 更 件 侵 时 同 新 延 装
程 恶 攻 接 全 实 网 行 运 自

Le CHINOINFO est un dictionnaire spécialisé bilingue destiné aux apprenants francophones du chinois. Ce dictionnaire répertorie 91 termes fondamentaux du domaine de l'informatique. Tous les renseignements terminologiques sont extraits à partir d'un corpus chinois comptant 341 676 mots.

L'élaboration du CHINOINFO s'inscrit dans le cadre d'un projet de thèse visant à concevoir une méthode pour décrire le lexique spécialisé chinois, plus précisément, les combinaisons lexicales spécialisées (CLS). Nous nous attachons à explorer de quelle manière la représentation sémantique des CLS favorise leur acquisition par les locuteurs non-linguistes. L'efficacité de la méthode mise au point sera évaluée au moyen d'une expérimentation.

Ce projet est réalisé par **Zhiwei Han**, étudiante au doctorat à l'Université de Montréal, sous la supervision de **Marie-Claude L'Homme**, professeure au Département de linguistique et de traduction de l'Université de Montréal. Nous remercions chaleureusement **Benwei Reibichang**, membre de l'Observatoire de linguistique Sens-Texte (OLST), pour la généreuse contribution qu'il a apportée à la mise en ligne du CHINOINFO.

Nomenclature (91 entrées) :

| trier par lexie | trier par catégorie | trier par glose |
|------------------------------------|-----------------------------------|--------------------------------|
| 安全 1 (nom) [sécurité informatique] | 发送 4 (nom) [envoi 2] | 检测 4 (nom) [vérification] |
| 安全 2 (adv.) [sécuriser] | 访问 1 (v.tr.) [accéder 1] | 连接 1 (v.tr.) [se connecter 1] |
| 安全 3 (adv.) [sécurité] | 访问 2 (v.tr.) [accéder 2] | 连接 2 (v.tr.) [connecter] |
| 安装 1 (v.tr.) [installer 1] | 访问 3 (nom) [accès 1] | 连接 3 (nom) [connexion 1.1] |
| 安装 2 (nom) [installation 2] | 访问 4 (nom) [accès 2] | 连接 4 (nom) [connexion 2.1] |
| 保存 1 (v.tr.) [enregistrer 1] | 复制 1 (v.tr.) [copier 1] | 迁移 1 (v.tr.) [faire migrer] |
| 保存 2 (nom) [enregistrement 1.1] | 复制 2 (v.tr.) [copier 2] | 迁移 2 (nom) [migration] |
| 备份 1 (nom) [sauvegarde 1] | 复制 3 (nom) [copie 1.1] | 入侵 1 (v.tr.) [s'introduire] |
| 备份 2 (nom) [sauvegarde 2] | 复制 4 (nom) [copie 1.2] | 入侵 2 (nom) [invasion] |
| 备份 3 (nom) [sauvegarde 3] | 更新 1 (v.tr.) [mettre à jour] | 软件 1 (nom) [logiciel 1] |
| 备份 4 (adv.) [à l'automatique] | 更新 2 (nom) [mise à jour 1] | 软件 2 (nom) [logiciel 2] |
| 测试 1 (v.tr.) [tester] | 更新 3 (nom) [mise à jour 2] | 删除 1 (v.tr.) [supprimer 1] |
| 测试 2 (nom) [test] | 攻击 1 (v.tr.) [attaquer 1] | 删除 2 (nom) [suppression 1] |
| 传输 1 (v.tr.) [transférer 1] | 攻击 2 (nom) [attaque informatique] | 删除 1 (v.tr.) [configurer 1] |
| 传输 2 (nom) [transfert 1] | 共享 1 (v.tr.) [partager 1] | 设置 2 (nom) [configuration 1.1] |
| 存储 1 (v.tr.) [stocker] | 共享 2 (nom) [partage 1] | 设置 3 (nom) [configuration 1.2] |

谢谢！
Merci pour votre participation!

Pour toute question, veuillez
me contacter par courriel :

Zhiwei HAN

zhiwei.han@umontreal.ca



Annexe 20 – Liste d’outils de référence proposés au GE

Description du lexique spécialisé chinois pour locuteurs non sinophones : apport de la Lexicologie explicative et combinatoire pour l’acquisition de compétences lexicales

Liste d’outils de référence

| Titre abrégé | Référence | Description |
|-----------------------------------|---|--|
| Dictionnaire CHINOINFO | Je vous invite à utiliser cette ressource comme outil de référence principal durant le test http://olst.ling.umontreal.ca/chinoinfo/ | Dictionnaire chinois-français |
| Dictionnaire Chine-Nouvelle | VerticalScope Inc. (2022). <i>Dictionnaire chinois français</i> . Chine-Nouvelle. https://www.chine-nouvelle.com/outils/dictionnaire.html | Dictionnaire chinois-français |
| Dictionnaire Chine-informations | Chine Informations. (2022). <i>Apprendre le chinois: dictionnaire chinois-français</i> . Chine-informations. https://chine.in/mandarin/dictionnaire/ | Dictionnaire chinois-français |
| LINE Chinese-English Dictionary | NAVER Corp. (2022). <i>LINE Dictionary : Chinese-English</i> . Dict.naver.com. https://dict.naver.com/linedict/zhendict/#/cnen/home | Dictionnaire chinois-anglais |
| Larousse Français-chinois | Éditions Larousse. (2022). <i>Dictionnaire Français-chinois</i> . Larousse.fr. https://www.larousse.fr/dictionnaires/francais-chinois | Dictionnaire français-chinois |
| MDBG Chinese-English Dictionary | <i>MDBG Chinese-English dictionary</i> . (2022). MDBG.NET. https://www.mdbg.net/chinese/dictionary?page=worddict&wdrst=0&wdab= | Dictionnaire chinois-anglais |
| Banque de terminologie Termonline | China National Committee for Terminology in Science and Technology. (2022). <i>Termonline</i> . Termonline: a public terminology knowledge service platform. https://www.termonline.cn/index | Banque de terminologie chinois-anglais |

Chercheuse étudiante : Zhiwei HAN (2022-04-21)
Projet approuvé par le Comité d'éthique de la recherche en arts et humanités de l'Université de Montréal (Projet no CERAH-2021-142-D)

1

Guide d’utilisation : comment consulter ces dictionnaires ?

Cliquer pour visionner les vidéos démonstratives

Dictionnaire Chine-Nouvelle

The screenshot shows the homepage of Chine-Nouvelle.com. At the top, there is a navigation menu with links like 'ACCUEIL', 'LA CHINE', 'LE CHINOIS', 'DICO', 'OUTILS', 'PRÉFOM CHINOIS', 'LOISIR', and 'FORUM'. Below the navigation, there are several promotional banners, including one for 'BEST BUY' on Samsung Galaxy Tab S7 FE. The main section is titled 'Dictionnaire chinois français' and features a search bar with the placeholder text 'Expression'. A yellow circle highlights the search bar, and a mouse cursor is pointing at it. Below the search bar, there are links for 'DERNIÈRES RECHERCHES DANS LE DICTIONNAIRE' and 'TOUS LES DICTIONNAIRES', which lists various language pairs like Chinese-French, Chinese-English, Chinese-Spanish, Chinese-Italian, and Chinese-Portuguese.

Chercheuse étudiante : Zhiwei HAN (2022-04-21)
Projet approuvé par le Comité d'éthique de la recherche en arts et humanités de l'Université de Montréal (Projet no CERAH-2021-142-D)

2

Guide d'utilisation : comment consulter ces dictionnaires ?
Cliquer pour visionner les vidéos démonstratives

[Dictionnaire Chine-informations](#)

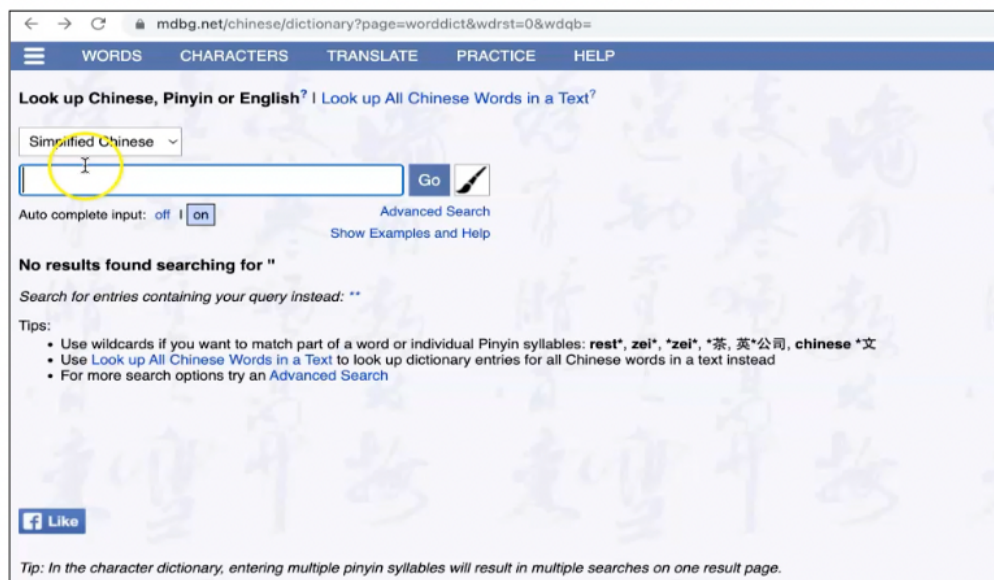


Chercheuse étudiante : Zhiwei HAN (2022-04-21)
Projet approuvé par le Comité d'éthique de la recherche en arts et humanités de l'Université de Montréal (Projet no CERAH-2021-142-D)

3

Guide d'utilisation : comment consulter ces dictionnaires ?
Cliquer pour visionner les vidéos démonstratives

[MDBG Chinese-English Dictionary](#)



Chercheuse étudiante : Zhiwei HAN (2022-04-21)
Projet approuvé par le Comité d'éthique de la recherche en arts et humanités de l'Université de Montréal (Projet no CERAH-2021-142-D)

4

Annexe 21 – Exemple de vocable du CHINOINFO

The screenshot displays three entries for the term '网络' (network) in the CHINOINFO interface. Each entry includes a definition, a domain, and a French equivalent. Red arrows point from text boxes to specific UI elements, and a yellow box highlights a note about inactive tabs.

Network 1 (nom) [Top]

INFORMATIQUE

français : [réseau 1](#)

Annotations for Network 1:

- Entrée + numéro d'acception (partie du discours)
- Mention du domaine de spécialité
- Équivalent français (renvoi à DicoInfo)
- Source de la définition citée (GDTRESE)
- Onglet actif par défaut
- Onglets inactif par défaut : cliquer sur les onglets masqués pour les consulter

Network 2 (nom) [Top]

INFORMATIQUE

français : [réseau 2](#)

Network 3 (adj.) [Top]

INFORMATIQUE

français : en ligne 1

Annexe 22 – Synthèse des termes reliés répertoriés dans le CHINOINFO

| Synthèses des termes reliés et combinaisons lexicales du ChinoInfo | |
|---|-------------|
| Termes reliés Combinaisons | |
| Relation | total = 282 |
| Sens proche | 69 |
| Contraire | 28 |
| Famille de mots-nom | 27 |
| Famille de mots-verbe | 26 |
| Synonyme | 18 |
| Famille de mots-nom de résultat | 13 |
| Famille de mots-adverbe | 10 |
| Méronymes (éléments constitutifs d'un réseau) | 7 |
| Instrument | 6 |
| Famille de mots-adjectif | 5 |
| Famille de mots-nom d'action | 4 |
| Personne qui réalise l'action d'attaquer | 4 |
| Dispositif de stockage | 3 |
| Hyponyme | 3 |
| Nom de sens voisin | 3 |
| Conversif | 2 |
| Lieu où on fait une connexion | 2 |
| Personne qui effectue un envoi | 2 |
| Personne qui envoie | 2 |
| Quasi-synonyme (synonyme plus spécifique) | 2 |
| Utilisateur | 2 |
| Action de causer un stockage | 1 |
| Action de créer une sauvegarde | 1 |
| Action de faire en sorte que quelque chose s'applique (à) | 1 |
| Action de mettre un logiciel (un programme) dans la forme d'une mise à jour | 1 |
| Action de rendre des données est cryptées | 1 |
| Action de rendre une tâche automatique | 1 |
| Cause qu'une configuration existe | 1 |
| Causer que les données ou le matériel soit partagé(es) | 1 |
| Causer un stockage | 1 |
| Causer une migration | 1 |
| Ce qui fait en sorte que quelque chose (un réseau ou un transfert de données) soit sécurisé | 1 |
| Co-méronyme | 1 |

| | |
|--|---|
| Créateur d'une application | 1 |
| D'une manière qui rend quelque chose (un réseau ou un transfert de données) sécurisé | 1 |
| Destinataire | 1 |
| Destinataire d'un envoi | 1 |
| Faire en sorte qu'un programme fasse l'objet d'une exécution | 1 |
| Faire en sorte qu'un programme s'exécute | 1 |
| Faire en sorte qu'une sauvegarde existe | 1 |
| Faire en sorte que des données est cryptées | 1 |
| Faire en sorte que quelque chose s'applique (à) | 1 |
| Faire l'objet d'une application informatique | 1 |
| Holonyme | 1 |
| Hyperonyme | 1 |
| Hyperonyme (ensemble de logiciels) | 1 |
| Hyperonyme (un logiciel fait partie de) | 1 |
| L'action de causer qu'une configuration existe | 1 |
| L'action qui cause que les données ou le matériel soit partagé(es) | 1 |
| L'utilisateur met un logiciel (un programme) dans la forme d'une mise à jour | 1 |
| L'état des données ou du matériel qui fait l'objet d'un partage | 1 |
| L'état des données ou du matériel qui subit l'action de partager | 1 |
| Méronyme (élément constitutif d'un logiciel) | 1 |
| Méronyme (élément constitutif du logiciel) | 1 |
| Personne qui réalise une intrusion | 1 |
| Quasi-synonyme (synonyme moins spécifique) | 1 |
| Rendre une tâche automatique | 1 |
| Réussir à se connecter (à) | 1 |
| Subir l'action d'appliquer | 1 |
| Utilisateur d'un réseau | 1 |
| Utiliser le réseau Internet | 1 |
| Verbe de sens voisin | 1 |
| État d'un logiciel à un moment donné de son développement | 1 |
| État des activités qui se réalisent sécuritairement | 1 |
| État qui résulte des mesures de sécurité | 1 |

Annexe 23 – Synthèse des CLS répertoriées dans le CHINOINFO

| Synthèses des termes reliés et combinaisons lexicales du ChinoInfo | | |
|--|---|-------------|
| Termes reliés Combinaisons | | |
| Structure | Combinaison | total = 644 |
| Construction appositive = [Modificateur] + Nom modifié apposé | Logiciel envisagé du point de vue de ses utilisations | 2 |
| | Application envisagée du point de vue de ses éléments constitutifs | 2 |
| Modificateur + [Nom modifié] | Sortes de | 160 |
| | Transformer | 1 |
| | Mettre quelque part | 1 |
| Modificateur adverbial + [Prédicat] | D'une certaine manière | 33 |
| Prédicat + [Complément de direction] | Dans une direction spécifique | 5 |
| Prédicat + [Complément d'objet] | Faire / Avoir lieu | 18 |
| | Accéder à un objet spécifique | 13 |
| | Envoyer un objet spécifique | 9 |
| | Télécharger un objet spécifique | 8 |
| | Utiliser/Ne pas utiliser | 7 |
| | Partager un objet spécifique | 7 |
| | Copier un objet spécifique | 7 |
| | Arrêter / Empêcher / Cesser d'avoir lieu | 6 |
| | Détecter un objet spécifique | 6 |
| | Se connecter à une destination spécifique | 5 |
| | connecter un objet spécifique | 4 |
| | Mettre à jour un objet spécifique | 4 |
| | Préparer l'utilisation/le fonctionnement | 3 |
| | Faire exécuter une entité spécifique | 3 |
| | Sauvegarder un objet spécifique | 3 |
| | Synchroniser un type de données | 3 |
| | Crypter un objet spécifique | 3 |
| | Configurer un paramètre spécifique (d'un matériel, d'un logiciel ou d'un ensemble matériel ou logiciel) | 3 |
| | Faire fonctionner | 3 |
| | Transformer | 3 |
| | Mettre quelque part | 2 |
| | Configurer un ensemble de matériel ou de logiciel | 2 |
| | Configurer un matériel | 2 |
| | Utiliser | 2 |
| | Commencer à utiliser une application | 2 |
| | Créer ou supprimer | 2 |
| | Empêcher d'avoir lieu | 2 |
| | Commencer à utiliser un logiciel | 2 |
| | Empêcher l'utilisation d'une configuration | 1 |
| | Permettre l'utilisation d'une configuration | 1 |
| | Empêcher l'utilisation d'un logiciel | 1 |
| | Mettre un logiciel quelque part | 1 |
| | Faire/avoir lieu | 1 |
| | Mettre une application quelque part | 1 |
| | Préparer l'utilisation/le fonctionnement d'une application | 1 |
| | Créer/Supprimer | 1 |
| Empêcher l'utilisation d'une application | 1 | |
| Créer | 1 | |
| Nuire au fonctionnement | 1 | |
| Vérifier un objet spécifique | 1 | |
| Arrêter de faire/cesser d'avoir lieu | 1 | |
| Configurer un logiciel | 1 | |

| | | |
|--|---|----|
| Structure coordonnée = [Nom1] + Nom2 | Opérations successives | 1 |
| Structure coordonnée = [Précat1] + Précat2 | Effectuer successivement l'action de copier et une autre action | 1 |
| [Modificateur adverbial] + Précat | Réaliser une tâche automatiquement | 8 |
| | Réaliser une activité à distance | 5 |
| | Réaliser une activité de manière malveillante | 4 |
| | Réaliser un traitement de données en temps réel | 4 |
| | Réaliser une activité en ligne | 4 |
| | Réaliser une activité sécuritairement | 3 |
| | Réaliser une tâche manuellement | 2 |
| | Réaliser une tâche synchroniquement | 2 |
| [Modificateur] + Nom modifié | Fonctionnement | 29 |
| | Moyen et instrument | 11 |
| | Types de ressources en ligne | 9 |
| | Types d'activités en ligne | 7 |
| | Appareil / technologie / réseau mobile | 7 |
| | Matériel et propriétés | 7 |
| | Traitement de données en temps réel | 6 |
| | Types de tâche automatique | 6 |
| | Types d'activités distantes | 5 |
| | Intervention sur le fonctionnement | 5 |
| | Opération visant à créer un logiciel | 4 |
| | Instrument | 4 |
| | Propriétés et phénomène liés à la sécurité informatique | 3 |
| | Types de données partagées | 3 |
| | Objets de nature malveillante | 3 |
| | Composants et propriétés | 3 |
| | Périphérique / Support de stockage amovible | 3 |
| | Types d'activités illicites en ligne | 3 |
| | Activités visant à assurer la sécurité informatique | 3 |
| | Types de données cryptées | 3 |
| | Types de tâche manuelle | 2 |
| | Traitement de données cryptées | 2 |
| | Activités de nature malveillante | 2 |
| | Arrêter / Empêcher / Cesser d'avoir lieu | 2 |
| | Types de matériels distants | 2 |
| | Types de ressources distantes | 2 |
| | Types de données sauvegardées | 2 |
| | Expéditeur (= entité qui effectue un envoi) | 2 |
| | Types de tâche synchrone | 2 |
| | Types d'intervenants | 2 |

| | | |
|--------------------------------------|--|----|
| | Un environnement virtuel qui est sécurisé | 2 |
| | Activités dont le déroulement est sécurisé | 2 |
| | Opération visant à transformer un logiciel | 2 |
| | Opération visant à assurer le bon fonctionnement d'un logiciel | 2 |
| | Source de téléchargement | 2 |
| | Opération visant à préparer l'utilisation ou le fonctionnement d'un logiciel | 1 |
| | Opération visant à mettre un logiciel quelque part | 1 |
| | Destination | 1 |
| | Ensemble de données et de matériel partagé | 1 |
| | Types de matériel partagé | 1 |
| | Propriétés | 1 |
| | Moyen | 1 |
| | Qui fait une synchronisation | 1 |
| | Programme conçu pour réaliser des traitements en temps réel | 1 |
| | Général | 1 |
| | Action de préparer l'utilisation/le fonctionnement | 1 |
| | Résultat | 1 |
| | Création d'une application | 1 |
| | Utilisateur | 1 |
| [Prédicat] + Complément de direction | Dans une direction spécifique | 19 |
| [Prédicat] + Complément de manière | D'une certaine manière | 1 |
| [Prédicat] + Complément de résultat | Détecter pour aboutir à un résultat spécifique | 2 |
| [Prédicat] + Complément d'objet | Installer un objet spécifique | 7 |
| | Appliquer une technique ou un ensemble de techniques | 6 |
| | Supprimer un objet spécifique | 6 |
| | S'introduire dans un objet spécifique | 5 |
| | Enregistrer un objet spécifique | 4 |
| | Faire migrer un objet spécifique | 4 |
| | Tester un objet spécifique | 4 |
| | Stocker un objet spécifique | 3 |
| | Transférer un objet spécifique | 3 |
| | Attaquer un objet spécifique | 2 |
| [Sujet] + Prédicat | Avoir lieu/Ne pas avoir lieu | 9 |
| | Fonctionner/Ne pas fonctionner | 8 |
| | Une entité s'exécute | 4 |
| | Avoir lieu / Ne pas avoir lieu | 2 |
| | Arrêter d'avoir lieu | 1 |

Annexe 24 – Grille de correction des questions 1 à 9 du test lexical

| | | |
|------------|--|--|
| Question 1 | Correction de la réponse (1 point) | Reconnaître que la forme lexicale véhicule deux sens différents dans les deux phrases (1 point). |
| | Pertinence de la justification (1 point) | Identifier le sens de la forme lexicale dans la phrase 1 (0,5 point). OU Identifier la partie du discours de la forme lexicale à l'aide des éléments contextuels dans la phrase 1 (0,5 point). |
| | | Identifier le sens de la forme lexicale dans la phrase 2 (0,5 point). OU Identifier la partie du discours de la forme lexicale à l'aide des éléments contextuels dans la phrase 1 (0,5 point). |
| Question 2 | Correction de la réponse (1 point) | Reconnaître que la forme lexicale véhicule deux sens différents dans les deux phrases (1 point). |
| | Pertinence de la justification (1 point) | Identifier le sens de la forme lexicale dans la phrase 1 (0,5 point). OU Identifier la partie du discours de la forme lexicale à l'aide des éléments contextuels dans la phrase 1 (0,5 point). |
| | | Identifier le sens de la forme lexicale dans la phrase 2 (0,5 point). OU Identifier la partie du discours de la forme lexicale à l'aide des éléments contextuels dans la phrase 1 (0,5 point). |
| Question 3 | Correction de la réponse (1 point) | Reconnaître que la forme lexicale véhicule deux sens différents dans les deux phrases (1 point). |
| | Pertinence de la justification (1 point) | Identifier le sens de la forme lexicale dans la phrase 1 (0,5 point). OU Identifier la partie du discours de la forme lexicale à l'aide des éléments contextuels dans la phrase 1 (0,5 point). |
| | | Identifier le sens de la forme lexicale dans la phrase 2 (0,5 point). OU Identifier la partie du discours de la forme lexicale à l'aide des éléments contextuels dans la phrase 1 (0,5 point). |
| Question 4 | Correction de la réponse (1 point) | Reconnaître que la même expression lexicale véhicule deux sens différents dans les phrases 1 et 2 (1 point). |

| | | |
|------------|--|--|
| | Pertinence de la justification (1 point) | Identifier le sens de la CLS dans la phrase 1 (0,5 point). OU Identifier la structure syntaxique de la CLS à l'aide des éléments contextuels dans la phrase 1 (0,5 point). |
| | | Identifier le sens de la CLS dans la phrase 2 (0,5 point). OU Identifier la structure syntaxique de la CLS à l'aide des éléments contextuels dans la phrase 2 (0,5 point). |
| Question 5 | Correction de la réponse (1 point) | Reconnaître que la même expression lexicale véhicule deux sens différents dans les phrases 1 et 2 (1 point). |
| | Pertinence de la justification (1 point) | Identifier le sens de la CLS dans la phrase 1 (0,5 point). OU Identifier la structure syntaxique de la CLS à l'aide des éléments contextuels dans la phrase 1 (0,5 point). |
| | | Identifier le sens de la CLS dans la phrase 2 (0,5 point). OU Identifier la structure syntaxique de la CLS à l'aide des éléments contextuels dans la phrase 2 (0,5 point). |
| Question 6 | Correction de la réponse (1 point) | Reconnaître que la même expression lexicale véhicule deux sens différents dans les phrases 1 et 2 (1 point). |
| | Pertinence de la justification (1 point) | Identifier le sens de la CLS dans la phrase 1 (0,5 point). OU Identifier la structure syntaxique de la CLS à l'aide des éléments contextuels dans la phrase 1 (0,5 point). |
| | | Identifier le sens de la CLS dans la phrase 2 (0,5 point). OU Identifier la structure syntaxique de la CLS à l'aide des éléments contextuels dans la phrase 2 (0,5 point). |
| Question 7 | Correction de la réponse (2 points) | Identifier le 1er cooccurrent approprié du terme donné (1 point). |
| | | Identifier le 2e cooccurrent approprié du terme donné (1 point). |
| Question 8 | Correction de la réponse (2 points) | Identifier le 1er cooccurrent approprié du terme donné (1 point). |
| | | Identifier le 2e cooccurrent approprié du terme donné (1 point). |
| Question 9 | Correction de la réponse (2 points) | Identifier le 1er cooccurrent approprié du terme donné (1 point). |
| | | Identifier le 2e cooccurrent approprié du terme donné (1 point). |

Annexe 25 – Analyse des erreurs repérées dans le test lexical

| | Habilité évaluée | Erreur repérée | | Nature de l'erreur | Fréquence | | Explication de l'erreur | | | | |
|-----------|---|----------------|---|--|-----------|----|---|--|---|---|--|
| | | | | | GC | GE | | | | | |
| Section I | Distinguer les sens associés à une forme lexicale | Erreur 1 | Q1-Phrase 2 wǎngluò xué xī 网络学习 « apprentissage en ligne » | Compréhension erronée du sens d'une forme lexicale | 2 | 0 | 1. Ne pas être en mesure d'identifier le sens de la forme lexicale 网络 (« en ligne ») figurant dans la CLS 网络学习 (« apprentissage en ligne »), car aucun élément de réponse n'a été trouvé dans les ressources lexicales consultées. 2. Attribuer une acception erronée (« réseau ») à la forme lexicale 网络 (« en ligne ») figurant dans la CLS 网络学习 (« apprentissage en ligne »). | | | | |
| | | | Erreur 2 | | | | Q2-Phrase 2 yí dòng yìng pán 移动硬盘 « disque dur amovible » | Compréhension erronée du sens d'une forme lexicale | 3 | 0 | 1. Ne pas être en mesure d'identifier le sens de la forme lexicale 移动 (« amovible ») figurant dans la CLS 移动硬盘 (« disque dur amovible »), car aucun élément de réponse n'a été trouvé dans les ressources lexicales consultées. 2. Attribuer une acception erronée (« mobile ») à la forme lexicale 移动 (« amovible ») figurant dans la CLS 移动硬盘 (« disque dur amovible »), car 移动 signifie « mobile » dans la CLS 移动设备 (« appareil mobile »). Il s'agit d'un cas de transfert négatif causé par la surgénéralisation des connaissances acquises en L2 (Tardif, 1999). |
| | | | Erreur 3 | | | | Q3-Phrase 1 ān zhuāng gēng xīn 安装更新 « installer une mise à jour » | | | | Compréhension erronée du sens d'une forme lexicale |

| | | | | | | | |
|------------|---|----------|--|--|---|--|--|
| | | | | | | <p>2. Identifier de façon erronée la partie du discours à laquelle appartient 更新 (^{gēngxīn} « mise à jour-nom de résultat ») figurant dans la CLS 安装更新 (^{ān zhuānggēngxīn} « installer une mise à jour »).</p> <p>3. Confondre les deux acceptions nominales de la forme lexicale 更新 (^{gēngxīn} (le nom d'action « mise à jour » et le nom de résultat « mise à jour »), car les ressources lexicales consultées ne font pas la distinction entre ces deux acceptions.</p> <p>4. Considérer le nom d'action 更新 (^{gēngxīn} « mise à jour ») et le nom de résultat 更新 (^{gēngxīn} « mise à jour ») comme ayant le même sens.</p> | |
| Section I | Distinguer les sens associés à une forme lexicale | Erreur 4 | <p>Q3-Phrase 2</p> <p>^{gēngxīn cāo zuò xì tǒng} 更新操作系统</p> <p>« mettre à jour un système d'exploitation »</p> | Compréhension erronée du sens d'une forme lexicale | 3 | 1 | <p>1. Ne pas être en mesure d'identifier la partie du discours à laquelle appartient 更新 (^{gēng xīn} « mettre à jour ») figurant dans la CLS 更新操作系统 (^{gēngxīn cāo zuò xì tǒng} « mettre à jour un système d'exploitation »), car les ressources lexicales consultées ne font pas la distinction entre le sens verbal et les sens nominaux associés à la forme lexicale 更新 (^{gēngxīn}).</p> <p>2. Considérer le verbe 更新 (^{gēngxīn} « mettre à jour ») et son dérivé nominal 更新 (^{gēngxīn} « mise à jour-nom de résultat ») comme ayant le même sens.</p> |
| Section II | Distinguer les sens associés à une combinaison lexicale | Erreur 1 | <p>Q4-Phrase 1</p> <p>^{ān zhuāngchéng xù} 安装程序</p> <p>« installer un programme »</p> | Interprétation erronée d'une combinaison lexicale | 2 | 0 | <p>1. Ne pas être en mesure d'identifier le sens de 安装 (^{ān zhuāng} « installer ») figurant dans la CLS 安装程序 (^{ān zhuāngchéng xù} « installer un programme ») car les ressources lexicales consultées ne font pas la distinction entre le sens verbal (« installer ») et le sens nominal (« installation ») associés à la forme lexicale 安装 (^{ān zhuāng}).</p> <p>2. Ne pas être en mesure d'identifier la structure syntaxique de la CLS 安装程序 (^{ān zhuāngchéng xù} « installer un programme ») à l'aide de son contexte d'utilisation. Dans la phrase 1 de la</p> |

| | | | | | | |
|------------|---|----------|--|---|-----|--|
| | | | | | | <p>Q4, la CLS est modifiée par l'adverbe 重新 (chóng xīn) (« de nouveau »). On aurait pu en déduire que la CLS est composée d'un verbe (安装 ān zhuāng « installer ») et de son complément d'objet (程序 chéng xù « programme »).</p> |
| Section II | Distinguer les sens associés à une combinaison lexicale | Erreur 2 | <p>Q4-Phrase 2 ān zhuāng chéng xù 安装程序 « programme d'installation »</p> | Interprétation erronée d'une combinaison lexicale | 1 0 | <p>1. Ne pas être en mesure d'identifier le sens de 安装 (ān zhuāng « installation ») figurant dans la CLS 安装程序 (ān zhuāng chéng xù « programme d'installation ») car les ressources lexicales consultées ne font pas la distinction entre le sens verbal (« installer ») et le sens nominal (« installation ») associés à la forme lexicale 安装 (ān zhuāng).</p> <p>2. Ne pas être en mesure d'identifier la structure syntaxique de la CLS 安装程序 (ān zhuāng chéng xù « programme d'installation ») à l'aide de son contexte d'utilisation. Dans la phrase 2 de la Q4, la CLS est reprise à l'aide de l'anaphore 这一程序 (zhè yī chéng xù « ce programme »). On aurait pu en déduire que 程序 (chéng xù « programme ») est la base nominale de la CLS et que 安装 (ān zhuāng « installation ») constitue le modificateur nominal de la base.</p> |
| Section II | Distinguer les sens associés à une combinaison lexicale | Erreur 3 | <p>Q5-Phrase 1 bèi fèn shù jù 备份数据 « sauvegarder les données »</p> | Interprétation erronée d'une combinaison lexicale | 4 0 | <p>1. Ne pas être en mesure d'identifier le sens de 备份 (bèi fèn « sauvegarder ») figurant dans la CLS 备份数据 (bèi fèn shù jù « sauvegarder les données ») car les ressources lexicales consultées ne distinguent pas les trois acceptions du mot multicatégoriel 备份 (bèi fèn), soit « sauvegarder », « sauvegarde (nom d'action) » et « sauvegardé ».</p> <p>2. Ne pas être en mesure d'identifier la structure syntaxique de la CLS 备份数据 (bèi fèn shù jù « sauvegarder les données ») à l'aide de son contexte d'utilisation. Dans la phrase 1 de la</p> |

| | | | | | | | |
|------------|---|----------|---|---|---|---|---|
| | | | | | | Q5, la CLS est modifié par l'adverbe 定期 (dìng qī « régulièrement »). On aurait pu en déduire que la CLS est composé d'une base verbale 备份 (bèi fèn « sauvegarder ») et de son complément d'objet 数据 (shù jù « données »). | |
| Section II | Distinguer les sens associés à une combinaison lexicale | Erreur 4 | Q5-Phrase 2 备份数据 (bèi fèn shù jù « données sauvegardées ») | Interprétation erronée d'une combinaison lexicale | 3 | 0 | <p>1. Ne pas être en mesure d'identifier le sens de 备份 (bèi fèn « sauvegardé ») figurant dans la CLS 备份数据 (bèi fèn shù jù « données sauvegardées ») car les ressources lexicales consultées ne distinguent pas les trois acceptions du mot multicatégoriel 备份 (bèi fèn), soit « sauvegarder », « sauvegarde (nom d'action) » et « sauvegardé ».</p> <p>2. Ne pas être en mesure d'identifier la structure syntaxique de la CLS 备份数据 (bèi fèn shù jù « données sauvegardées ») à l'aide de son contexte d'utilisation. Dans la phrase 2 de la Q5, la CLS est le complément d'objet du verbe 下载 (xià zài « télécharger »).</p> <p>3. Identifier à tort 下载 (xià zài « télécharger ») et 备份 (bèi fèn « sauvegardé ») comme une combinaison lexicale composée de deux verbes coordonnés.</p> |
| Section II | Distinguer les sens associés à une combinaison lexicale | Erreur 5 | Q6-Phrase 1 共享文件 (gòng xiǎng wén jiàn « partager un fichier ») | Interprétation erronée d'une combinaison lexicale | 3 | 1 | <p>1. Ne pas être en mesure d'identifier le sens de 共享 (gòng xiǎng « partager ») figurant dans la CLS 共享文件 (gòng xiǎng wén jiàn « partager un fichier ») car les ressources lexicales consultées ne distinguent pas les trois acceptions du mot multicatégoriel 共享 (gòng xiǎng), soit « partager », « partage (nom d'action) » et « partagé ».</p> <p>2. Ne pas être en mesure d'identifier la structure syntaxique de la CLS 共享文件 (gòng xiǎng wén jiàn « partager un fichier ») à l'aide de son contexte d'utilisation. Dans la phrase 1 de la Q6, la CLS est associée à la structure « verbe-complément d'objet ». Le mot multicatégoriel</p> |

| | | | | | | | |
|-------------|---|----------|---|---|---|--|--|
| | | | | | | <p>gòngxiāng 共享 véhicule un sens verbal puisqu'il est modifié par un complément de manière 与其 他用户 (« avec d'autres utilisateur »).</p> | |
| Section II | Distinguer les sens associés à une combinaison lexicale | Erreur 5 | <p>Q6-Phrase 2 gòngxiāng wén jiàn 共享文件 « fichier partagé »</p> | Interprétation erronée d'une combinaison lexicale | 4 | 0 | <p>1. Ne pas être en mesure d'identifier le sens de gòngxiāng (« partagé ») figurant dans la CLS gòngxiāng wén jiàn (« fichier partagé ») car les ressources lexicales consultées ne distinguent pas les trois acceptions du mot multicatégoriel gòngxiāng, soit « partager », « partage (nom d'action) » et « partagé ».</p> <p>2. Ne pas être en mesure d'identifier la structure syntaxique de la CLS gòngxiāng wén jiàn (« fichier partagé ») à l'aide de son contexte d'utilisation. Dans la phrase 2 de la Q6, La CLS est précédé d'une série de verbes (chá kàn « consulter », biān jí « éditer » et shān chú « supprimer »). On aurait pu en déduire que la CLS est de nature nominale et qu'elle est associée à la structure « modificateur-nom modifié ». Le mot multicatégoriel gòngxiāng wén jiàn véhicule un sens adjectival (« partagé ») et modifie wén jiàn (« fichier »).</p> |
| Section III | Choisir un cooccurrent approprié d'un terme | Erreur 1 | <p>Q7-Réponse A rù qīnwǎngzhàn *入侵网站 'accéder à' - 'site Web'</p> | Construction erronée d'une combinaison lexicale | 4 | 0 | <p>1. Ne pas être en mesure de distinguer fǎngwèn (« accéder ») et rù qīn (« s'introduire ») lorsque ces deux termes se combinent avec wǎngzhàn (« site Web »). Bien que les deux termes semblent pouvoir être traduit par le même mot français (accéder), le deuxième terme fait référence à une activité illicite et s'utilise avec une connotation négative.</p> <p>2. Ne pas être en mesure de choisir le collocatif verbal approprié pour le terme wǎngzhàn (« site Web ») peut former une CLS signifiant « accéder à un site Web » car aucun élément</p> |

| | | | | | | |
|-------------|---|----------|---|---|-----|--|
| | | | | | | de réponse n'a été proposé dans les ouvrages de référence consultés. |
| Section III | Choisir un cooccurrent approprié d'un terme | Erreur 2 | Q7-Réponse B *入侵网站 rù qīnwǎngzhàn 'accéder à' - 'site Web' | Construction erronée d'une combinaison lexicale | 0 2 | 1. Ne pas être en mesure de distinguer 访问 (fǎngwèn « accéder ») et 入侵 (rù qīn « s'introduire ») lorsque ces deux termes se combinent avec 网站 (wǎngzhàn « site Web »). Bien que les deux termes semblent pouvoir être traduit par le même mot français (accéder), le deuxième terme fait référence à une activité illicite et s'utilise avec une connotation négative. |
| Section III | Choisir un cooccurrent approprié d'un terme | Erreur 3 | Q8-Réponse A *移动文件 yí dòng wén jiàn 'déplacer' - 'fichier' | Construction erronée d'une combinaison lexicale | 2 0 | 1. Confondre le terme 移动 (yí dòng « mobile ») et l'unité lexicale 移动 (yí dòng « déplacer ») en langue général. On attribue à tort un sens verbal (« déplacer ») à l'adjectif 移动 (yí dòng « mobile »). 2. Ne pas être en mesure de choisir le collocatif verbal approprié pour construire la CLS signifiant 'transférer un fichier'. On combine à tort une unité lexicale qui semble signifier 'déplacer' avec le terme 文件 (wén jiàn « fichier » pour exprimer le sens de 'faire passer un fichier d'un emplacement à un autre'. Contrairement au français, la combinaison lexicale est inacceptable en chinois. Il s'agit d'un cas de transfert négatif causé par l'interférence de la L1 (Tardif, 1999). 3. Se fier à la traduction automatique erronée (déplacer un fichier est traduit par 移动文件 (yí dòng wén jiàn)). |

| | | | | | | |
|-------------|---|----------|--|---|---|---|
| Section III | Choisir un cooccurrent approprié d'un terme | Erreur 4 | <p>Q8-Réponse B</p> <p><small>jì lù wénjiàn</small> * 记录文件</p> <p>'consigner' - 'fichier'</p> | Construction erronée d'une combinaison lexicale | 1 | <p>1</p> <p>1. Ne pas être en mesure de distinguer <small>bǎo cún</small> 保存 (« enregistrer ») et <small>jì lù</small> 记录 (« consigner ») lorsque ces deux termes se combinent avec <small>wén jiàn</small> 文件 (« fichier »). Ces deux verbes semblent pouvoir être traduits par le même mot français enregistrer. En chinois, <small>jì lù</small> 记录 (« consigner ») dénote l'action de prendre une information en note. Il ne s'agit pas d'une unité lexicale en langue générale et non pas un terme en informatique.</p> <p>2. Ne pas être en mesure de choisir le collocatif verbal approprié pour construire la CLS signifiant 'enregistrer un fichier'.</p> <p>On combine à tort un verbe qui semble signifier 'enregistrer' avec le terme <small>wén jiàn</small> 文件 « fichier ». Contrairement au français, la combinaison lexicale est inacceptable en chinois. Il s'agit d'un cas de transfert négatif causé par l'interférence de la L1 (Tardif, 1999).</p> <p>3. Se fier à la traduction automatique erronée (consigner un fichier est traduit par <small>jì lù wénjiàn</small> 记录文件).</p> |
| Section III | Choisir un cooccurrent approprié d'un terme | Erreur 5 | <p>Q9-Réponse B</p> <p><small>chuán shū yìng pán</small> * 传输硬盘</p> <p>'transférer' - 'disque dur'</p> | Construction erronée d'une combinaison lexicale | 2 | <p>2</p> <p>1. Ne pas être en mesure de choisir le collocatif approprié pour le terme <small>chuán shū</small> 传输 (« transférer »).</p> <p>En informatique, <small>chuán shū</small> 传输 (« transférer ») a comme complément d'objet un nom désignant un type de données. Il ne peut se combiner avec <small>yìng pán</small> 硬盘 (« disque dur ») qui dénote un périphérique de stockage.</p> <p>2. Se fier aux résultats de recherche sur le Web. On vérifie la cooccurrence sur Google et trouve la séquence lexicale <small>gāo sù chuán shū yìng pán</small> 高速传输硬盘</p> |

| | | | | | | | |
|--|--|--|--|--|--|--|--|
| | | | | | | | <p>(‘haute vitesse’-‘transfert’-‘disque dur’ « disque dur haute vitesse »). Par contre, 传输^{chuán shū} est un nom (« transfert ») qui est modifié par 高速^{gāo sù} (« haute-vitesse »). Il ne s’agit pas d’un collocatif verbal du terme 硬盘^{yìng pán} (« disque dur »).</p> |
|--|--|--|--|--|--|--|--|