

Table des matières

Avant-propos	1
Chapitre 1. Comment définir la linguistique de corpus ?	3
1.1. Éléments de définition	3
1.2. Empirisme <i>versus</i> rationalisme en linguistique	5
1.3. Les arguments de Chomsky contre l’empirisme en linguistique	7
1.4. Linguistique de corpus et outils informatiques	11
1.5. Méthodes quantitatives <i>versus</i> qualitatives	13
1.6. Différences entre linguistique de corpus et linguistique expérimentale	16
1.7. Les différents types de corpus	18
1.8. Conclusion	21
1.9. Questions de révision et corrigés	21
1.9.1. Questions	21
1.9.2. Corrigés	22
1.10. Pour aller plus loin	25
Chapitre 2. Comment utiliser des corpus en linguistique théorique ?	27
2.1. Phonétique et phonologie	27
2.2. Morphologie	30
2.3. Syntaxe	33
2.4. Lexique	35
2.5. Analyse du discours	38
2.6. Pragmatique	41

2.7. Sociolinguistique	43
2.8. Diachronie	45
2.9. Conclusion	48
2.10. Questions de révision et corrigés	48
2.10.1. Questions	48
2.10.2. Corrigés	49
2.11. Pour aller plus loin	51

Chapitre 3. Comment utiliser des corpus en linguistique appliquée ? 53

3.1. Acquisition du langage	53
3.2. Pathologies du langage	57
3.3. Acquisition d'une langue seconde	60
3.4. Enseignement des langues	63
3.5. Lexicographie	67
3.6. Stylistique	70
3.7. Linguistique légale	73
3.8. Conclusion	75
3.9. Questions de révision et corrigés	75
3.9.1. Questions	75
3.9.2. Corrigés	76
3.10. Pour aller plus loin	78

Chapitre 4. Comment utiliser des corpus multilingues ? 79

4.1. Corpus comparables et corpus parallèles	79
4.2. À la recherche d'un <i>tertium comparationis</i>	83
4.3. Les traductions comme genre discursif	86
4.4. Corpus multilingues et linguistique contrastive	88
4.5. Corpus parallèles et traductologie	92
4.6. Corpus parallèles et dictionnaires bilingues	95
4.7. Conclusion	97
4.8. Questions de révision et corrigés	97
4.8.1. Questions	97
4.8.2. Corrigés	98
4.9. Pour aller plus loin	100

Chapitre 5. Comment trouver et analyser des corpus en français ?	101
5.1. Formats de mise à disposition des corpus	101
5.2. Les corpus de référence.	104
5.3. Les corpus de français écrit	105
5.4. Les corpus de français oral.	107
5.5. Les corpus de populations spécifiques	109
5.6. Les corpus multilingues incluant le français	112
5.7. Outils de consultation des corpus	115
5.7.1. Les concordanciers	116
5.7.2. Zoom sur le concordancier AntConc	119
5.7.3. Zoom sur le concordancier CLAN	120
5.8. Conclusion	122
5.9. Questions de révision et corrigés	122
5.9.1. Questions	122
5.9.2. Corrigés.	123
5.10. Pour aller plus loin.	128
Chapitre 6. Comment construire un corpus ?	129
6.1. Avant de décider de construire un corpus	129
6.2. Déterminer la taille et la représentativité des données.	131
6.3. Choisir des échantillons à inclure dans le corpus.	135
6.4. Préparer et coder les fichiers du corpus	138
6.5. Enregistrer et transcrire des données orales	143
6.6. Questions éthiques et légales	145
6.7. Conclusion	148
6.8. Questions de révision et corrigés	148
6.8.1. Questions.	148
6.8.2. Corrigés.	149
6.9. Pour aller plus loin	152
Chapitre 7. Comment annoter un corpus ?	153
7.1. Les annotations de corpus	153
7.2. Les différents types d’annotations	156
7.3. Standardisation des schémas d’annotation.	160

7.4. Les étapes d'un processus d'annotation	161
7.5. Les outils d'annotation	166
7.6. Mesurer la qualité et la fiabilité d'une annotation	169
7.7. Partager ses annotations	173
7.8. Conclusion	174
7.9. Questions de révision et corrigés	175
7.9.1. Questions	175
7.9.2. Corrigés	175
7.10. Pour aller plus loin	180
Chapitre 8. Comment analyser les données de corpus ?	181
8.1. Statistiques descriptives pour les données de corpus	181
8.2. Mesurer la richesse lexicale d'un corpus	186
8.3. Mesurer la dispersion d'un mot dans un corpus	187
8.4. Notions de statistiques inférentielles	190
8.5. Les variables typiques dans les études de corpus	193
8.6. Mesurer les différences entre catégories	196
8.7. Conclusion	204
8.8. Questions de révision et corrigés	205
8.8.1. Questions	205
8.8.2. Corrigés	207
8.9. Pour aller plus loin	211
Conclusion. Les étapes de réalisation d'une étude de corpus	213
Bibliographie	219
Index	241