HAL
open science

Corpus numériques et production écrite en langue étrangère. Une recherche avec des apprenants d'allemand.

Eva Schaeffer-Lacroix

To cite this version:

Eva Schaeffer-Lacroix. Corpus numériques et production écrite en langue étrangère. Une recherche avec des apprenants d'allemand.. domain_other. Université de la Sorbonne nouvelle - Paris III, 2009. Français. NNT : . tel-00439095v1

HAL Id: tel-00439095

https://theses.hal.science/tel-00439095v1

Submitted on 6 Dec 2009 (v1), last revised 17 Feb 2014 (v2)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

UNIVERSITÉ SORBONNE NOUVELLE – PARIS 3

ÉCOLE DOCTORALE 268

Thèse en sciences du langage

Eva SCHAEFFER-LACROIX

CORPUS NUMÉRIQUES ET PRODUCTION ÉCRITE EN LANGUE ÉTRANGÈRE

UNE RECHERCHE AVEC DES APPRENANTS D'ALLEMAND

Tome 1

Thèse dirigée par Françoise DEMAIZIÈRE et André SALEM

Soutenue le 28 novembre 2009

Jury :

Johannes Angermüller, professeur (Johannes-Gutenberg-Universität Mainz), rapporteur
Martine Dalmas, professeur (université Paris 4), rapporteur
Françoise Demaizière, ingénieur de recherche HDR (université Paris 7), experte invitée
Dominique Macaire, maître de conférences HDR (université Paris 3 et IUFM d'Aquitaine, Bordeaux 4)
André Salem, professeur émérite (université Paris 3), directeur de thèse

Résumé

Cette thèse en didactique des langues étrangères montre le potentiel des corpus numériques pour l'apprentissage de l'allemand au collège. Conformément aux objectifs poursuivis dans une recherche-action, elle renseigne sur l'évolution des acteurs impliqués dans un dispositif spécifique, en tenant compte de facteurs multiples, tels que le type d'*input*, les outils, les modes de travail - en présentiel et à distance - et la relation pédagogique. J'ai observé les effets d'un scénario pédagogique, basé sur l'exploitation de corpus spécialisés pour la création de textes du domaine du tourisme. Les participants, des adolescents de niveau intermédiaire en allemand, langue étrangère, ont exploré les corpus à l'aide de lignes de concordance. J'ai analysé leurs discours et leurs actions face à l'*input*, et j'ai décrit l'évolution de certains aspects linguistiques et discursifs de leurs textes. Le projet et les outils proposés ont soutenu le développement de modes de lecture qui ont favorisé la recherche ciblée d'informations dans l'*input*. Le recyclage d'éléments repérés dans l'*input* a mené, pour une partie des apprenants, à l'obtention de productions écrites dont les caractéristiques linguistiques et discursives sont proches de celles attestées dans les corpus. L'écriture intertextuelle a permis aux participants, à des degrés variables, de se décentrer et de progresser dans l'apprentissage de l'altérité. L'observation de formes linguistiques à l'aide de lignes de concordance et l'intégration de collocations identifiées dans l'*input* ont contribué au développement du système linguistique, en particulier dans le domaine de la flexion des adjectifs.

Mots clés :

corpus, concordances, production écrite, langue étrangère, allemand, ressources pédagogiques, genre textuel, recherche d'informations, apprenant-chercheur

Abstract

This action-research in applied linguistics investigates the evolution of adolescent learners of German as a foreign language during a project involving corpus exploration. It focuses on the role played by resources, tools, face-to-face learning *versus* distance learning and by the pedagogical relationship. My research analyzes how the learners perceived and exploited three small specialized corpora containing texts representing each one sub-genre of the "tourist information" genre. The learners explored them by means of concordance lines. Firstly, I observe what the learners said about the input and what they did with it. Secondly, I describe how the language and the discourse used in the texts changed during the project. The tasks and tools of the project helped some of the learners to develop reading habits which focus on information retrieval more than on complete comprehension of the text. The effects are texts written in a German having linguistic features that resemble those attested in the explored corpora. The intertextual writing method offered by the project supports acculturation to the foreign language. By means of concordancing and the import of collocations, the learners are now able to deal with some language features, especially with the domain of German adjective inflection, better than before.

Keywords:

corpora, concordances, writing, foreign language acquisition, German, pedagogical resources, discourse genre, information retrieval, learner as a researcher

Remerciements

Cette thèse doit son existence tout d'abord à Françoise Demaizière qui m'a encouragée, pendant des phases décisives de mon cursus universitaire, à entreprendre et à poursuivre mon travail de recherche. Ses conseils et remarques pertinents m'ont toujours été d'une grande aide, et je l'en remercie vivement. Je suis également très reconnaissante d'avoir pu échanger avec les membres de l'équipe de doctorantes et doctorants de Paris 3 et Paris 7 que Françoise Demaizière a su rassembler régulièrement pour mener des discussions de fond, empreintes d'ouverture d'esprit et de chaleur humaine. Un grand merci à l'enquêtrice qui a mené, avec tact et talent, les entretiens pour le projet principal de ma recherche.

André Salem a repris la direction de mes travaux fin décembre 2008. Il a donné un éclairage supplémentaire à ma thèse, et il m'a aidée en particulier dans les parties qui nécessitaient un savoir-faire dans le domaine des corpus et concordances. Je lui suis très reconnaissante de m'avoir soutenue avec simplicité et générosité lors de la phase finale de mon travail.

Je remercie l'équipe administrative du Collège Wanda Landowska à Saint-Leu-la-Forêt (95) qui m'a accordé un terrain de recherche pendant trois années de suite. Un grand merci à mes collègues de français et de technologie qui ont participé à une partie des recherches. Le soutien en informatique de François Fix, personne-ressource du collège, a été précieux. Je remercie tout particulièrement les élèves des quatre classes de troisième qui, avec patience et bienveillance, ont bien voulu se prêter à mon "jeu" de recherche, et leurs parents qui m'ont fait confiance dans cette entreprise peu commune au sein d'un collège.

Merci à Janine Burger, Gabriele Helke, Christian Ingelmann, Dirk Petzold, Ludwig Schletzbaum et Gerhard Stahr qui m'ont autorisée à faire des corpus à partir de données pour lesquelles ils gèrent les droits d'auteur.

Mein Sonderdank geht an Jean-Claude, Alice, Rémy und Jacqueline. Ihr wart ganz einfach spitze in diesen intensiven Arbeitsjahren.

Sommaire

INTRODUCTION **13**
1. CORPUS ET CONCORDANCES **18**
1.1. APERÇU HISTORIQUE 18
1.2. QU'EST-CE QU'UN CORPUS ? 20
1.2.1. Définition négative de "corpus" *20*
1.2.2. Distinction entre corpus et autres ressources numériques *20*
1.2.3. Définitions de "corpus" *24*
1.2.4. Types de corpus *27*
1.3. ANNOTER UN CORPUS 29
1.3.1. Pourquoi et comment annoter un corpus ? *29*
1.3.2. Peut-on se passer d'annotations ? *31*
1.4. OBJECTIFS D'UTILISATION D'UN CORPUS 31
1.4.1. Objectifs "théoriques" *31*
1.4.2. Objectifs "pratiques" ou la "théorie appliquée" *33*
1.5. OUTILS D'EXPLORATION DE TEXTES 34
1.5.1. Moteurs de recherche *34*
1.5.2. Fonction "rechercher dans ce document/cette page…" *34*
1.5.3. Requêtes par formulaire *35*
1.6. OUTILS D'EXPLORATION DE CORPUS 35
1.6.1. Calculateur de fréquence *36*
1.6.2. Outil de dispersion *36*
1.6.3. Concordances et concordanciers *36*
1.6.3.1. Concordances 36
1.6.3.2. Concordanciers 38
1.6.4. Concordanciers pour l'allemand *40*
1.7. CORPUS EN LANGUE ALLEMANDE 42
1.7.1. La place des corpus allemands dans le paysage des corpus *42*
1.7.2. Corpus en langue allemande *43*
1.7.2.1. Corpus parallèles 44
1.7.2.2. Corpus unilingues 44
1.7.2.2.1. Corpus d'apprenants 44
1.7.2.2.2. Portails terminologiques, basés sur des corpus 45
1.7.2.2.3. Autres corpus unilingues 45
1.8. POUR RÉSUMER 48
2. APPRENTISSAGE D'UNE LANGUE ÉTRANGÈRE **49**
2.1. LINGUISTIQUE COGNITIVE ET APPRENTISSAGE DE LE 50
2.1.1. Activité langagière *50*

2.1.2. Opérations énonciatives *51*
2.1.2.1. Paraphrases 52
2.1.2.2. Marqueurs 52
2.1.2.3. Invariances 52
2.2. PSYCHOLINGUISTIQUE ET APPRENTISSAGE DE LE 53
2.2.1. Apprentissage de LE, un processus adaptatif *53*
2.2.2. Input, intake, output et apprentissage de LE *54*
2.2.2.1. *Output hypothesis* 54
2.2.2.2. *Input hypothesis* 55
2.2.2.3. Développement de l'*intake* 57
2.2.2.4. Ma définition d'*intake* 58
2.3. DÉFINITIONS D'UNITÉS POLYLEXICALES 59
2.3.1. Perspective des linguistes de corpus *61*
2.3.2. Perspective des lexicologues et des lexicographes *62*
2.3.3. Perspective des phraséologues *62*
2.3.4. Perspective des psycholinguistes *63*
2.3.5. Perspective des didacticiens *63*
2.3.6. Ma définition du phénomène collocatif *64*
2.4. COLLOCATIONS ET APPRENTISSAGE DE LE 67
2.4.1. Collocations et mémorisation *67*
2.4.2. Digression : microsyntaxe et macrosyntaxe *69*
2.4.3. Collocations et microsyntaxe *70*
2.5. ACTIVITÉS BASÉES SUR L'INPUT EN LE 71
2.5.1. Recherche d'informations *71*
2.5.2. Catégorisation *74*
2.6. CONSTRUCTIVISME ET APPRENTISSAGE DE LE 75
2.6.1. Attitude d'apprenant-linguiste *76*
2.6.2. Attitude d'apprenant-détective *77*
2.6.3. Attitude d'apprenant-chercheur *78*
2.7. GUIDAGE ET APPRENTISSAGE DE LE 79
2.7.1. Guidage selon le point de vue de Piaget *79*
2.7.2. Guidage selon le point de vue de Bruner *81*
2.7.3. Guidage et TIC *83*
2.8. APPROCHES PÉDAGOGIQUES 84
2.8.1. Approches privilégiant le sens *84*
2.8.1.1. Pédagogie du projet 84
2.8.1.2. Approche par tâches 85
2.8.2. Approche combinant sens et forme – la conceptualisation *88*
2.9. POUR RÉSUMER 89
3. PRODUCTION ÉCRITE EN LANGUE ÉTRANGÈRE 91
3.1. MODÉLISATION DE LA PRODUCTION ÉCRITE EN LANGUE MATERNELLE 92
3.1.1. Modèle de Hayes et Flower (1980) *92*

3.1.2. Modèle de Bereiter et Scardamalia (1983 ; 1987) *94*
3.1.3. Modèle de Fayol (1996) *94*
3.2. MODÉLISATION DE LA PRODUCTION ÉCRITE EN LANGUE ÉTRANGÈRE 96
3.2.1. Modèles de Hayes (1996) et de Grabe et Kaplan (1996) *96*
3.2.2. Modèle de Zimmermann (2000) *97*
3.3. RÉPERCUSSION DES MODÈLES DE PRODUCTION ÉCRITE SUR MON SCÉNARIO PÉDAGOGIQUE 98
3.4. ENSEIGNEMENT/APPRENTISSAGE DE LA PRODUCTION ÉCRITE EN LANGUE ÉTRANGÈRE 99
3.4.1. Pratiques d'écriture *99*
3.4.2. Input et production écrite *101*
3.4.2.1. Quel *input* pour un projet d'écriture ? 101
3.4.2.2. Genre textuel 103
3.4.2.3. Que faire avec l'*input* ? 104
3.4.2.3.1. Recherche d'informations 104
3.4.2.3.2. Démarche intertextuelle 105
3.5. NOUVEAUX GENRES, NOUVELLES NORMES ? 107
3.6. PRODUIRE DES TEXTES EN LE À L'AIDE DE L'ORDINATEUR 108
3.6.1. Caractéristiques du document numérique *108*
3.6.2. Rôle de l'outil *109*
3.6.3. Savoir-faire numérique *109*
3.7. DES OUTILS D'AIDE À LA RÉDACTION VERS LES RESSOURCES PÉDAGOGIQUES 110
3.7.1. Outils d'aide à la rédaction *110*
3.6.1.1. Traducteur électronique 111
3.6.1.2. Dictionnaire électronique 113
3.7.2. Les ressources *114*
3.6.2.1. Définition générale 114
3.6.2.2. Ressources pédagogiques 115
3.8. RÉPERCUSSION DES MODÈLES ET MÉTHODES SUR MON SCÉNARIO PÉDAGOGIQUE 118
3.8.1. Input et scénario pédagogique *119*
3.8.2. Guidage et scénario pédagogique *120*
4. ENSEIGNER/APPRENDRE UNE LANGUE ÉTRANGÈRE À L'AIDE DE CORPUS ET DE CONCORDANCES 122
4.1. QUELS CORPUS POUR DES TÂCHES DE PRODUCTION ÉCRITE ? 122
4.1.1. Quels objectifs pour une utilisation de corpus ? *123*
4.1.2. Besoins concernant les caractéristiques des corpus *123*
4.1.3. Critique de corpus allemands existants *124*
4.1.4. Création de petits corpus spécialisés *126*
4.1.5. Corpus, outils et objectifs retenus *127*
4.1.6. Mise à l'épreuve du corpus "Informations pratiques" *128*
4.2. TÂCHES BASÉES SUR DES CONCORDANCES 130
4.2.1. Méthodologies convoquées *130*
4.2.2. Macrotâches, microtâches et data-driven learning *133*
4.2.3. Analyse de tâches basées sur des concordances *135*

4.2.4. Quelles microtâches pour soutenir l'observation de la langue ? *142*
4.2.4.1. Atelier de production d'activités pédagogiques 142
4.2.4.2. Je t'invite à mon anniversaire 144
4.3. POUR RÉSUMER 147
5. UNE RECHERCHE MENÉE AU COLLÈGE 149
5.1. TYPES DE RECHERCHE 149
5.1.1. Recherche expérimentale *149*
5.1.1.1. Types d'observation 149
5.1.1.2. Postures de chercheur 149
5.1.2. Recherche-action *150*
5.1.2.1. Pourquoi choisir la méthodologie de la recherche-action ? 150
5.1.2.2. Types d'observation 151
5.1.2.3. Postures de chercheur 151
5.1.2.4. Ma posture de chercheur 152
5.2. LÉGITIMER UNE RECHERCHE AU COLLÈGE 153
5.2.1. S'inspirer de manuels scolaires *153*
5.2.2. Programmes officiels d'enseignement *154*
5.2.2.1. Programme officiel pour la technologie 154
5.2.2.2. Programme officiel pour le français 154
5.2.2.3. Programme officiel pour l'allemand et *Projekt Prospekt* 155
5.2.2.4. Valider certains points du b2i 159
5.2.3. Droit à l'expérimentation *159*
5.2.4. Projet d'établissement *160*
5.3. SCÉNARIO PÉDAGOGIQUE 161
5.3.1. Descriptif général du scénario pédagogique *161*
5.3.2. Phases du processus rédactionnel dans les projets de ma recherche *163*
5.3.3. Création de dépliants touristiques (2006/07) *164*
5.3.4. Projekt Prospekt (2007/08) *164*
5.3.5. Brieffreunde.org (2008/09) *166*
5.3.6. Outils électroniques proposés aux apprenants *167*
5.3.6.1. Corpus et concordanciers 167
5.3.6.1.1. Première et deuxième phase (2006/07 et 2007/08) 167
5.3.6.1.2. Troisième phase (2008/09) 167
5.3.6.2. Outils de communication 168
5.3.6.2.1. Boîte courriel 168
5.3.6.2.2. Blogues pédagogiques 168
Première phase (2006/07) 169
Blogue Projektprospekt 169
Deuxième phase (2007/08) 170
Blogue Tourismus 170
Blogue Projekt Prospekt 170
Troisième phase (2008/09) 171
Blogue esldeutsch 171

Blogue Filmblog 172
5.3.7. Activités préparatoires *172*
5.3.7.1. Préparation aux outils et méthodes 172
5.3.7.2. Séance de présentation de *Projekt Prospekt* 173
5.3.7.3. Activités d'entraînement 175
5.3.7.3.1. Repérage d'informations dans un texte appartenant à un genre textuel 175
Observation des caractéristiques linguistiques de sections d'un genre textuel 175
Assembler des informations appartenant aux mêmes sections (Glisser-déposer) 176
Repérer et surligner des sections 177
Repérer des avis positifs, négatifs et mitigés 178
5.3.7.3.2. Démonstrations de techniques de recherche d'informations 179
5.3.7.3.3. Observation de lignes de concordance 179
5.4. DOCUMENTS ANALYSÉS 182
5.4.1. Entretiens *182*
5.4.2. Questionnaires *182*
5.4.2.1. Première phase 182
5.4.2.2. Deuxième phase 183
5.4.3. Échanges en présentiel *183*
5.4.4. Échanges à distance *183*
5.4.4.1. Courriels 183
5.4.4.2. Articles et commentaires sur le blogue 184
5.4.4.3. Annotations de productions écrites des apprenants 185
5.4.5. Journal de bord *186*
5.5. MÉTHODE D'ANALYSE DES DONNÉES 186
5.5.1. Relevés faits à l'aide de collections d'outils de statistique textuelle *186*
5.5.1.1. Identification des thèmes abordés par les acteurs de la recherche dans les documents recueillis 186
5.5.1.2. Fréquence lexicale dans les productions écrites des apprenants 186
5.5.1.3. Localisation de passages co-présents dans les corpus de référence et les corpus d'apprenants 187
5.5.1.4. Fréquence des prises de parole dans les entretiens 187
5.5.2. Relevés faits à l'aide du traitement de texte *187*
5.5.2.1. Comptage du nombre de versions et du nombre de mots des productions d'apprenants 187
5.5.2.2. Comparaisons de versions 187
5.5.3. Relevé fait manuellement : évolution de caractéristiques linguistiques des productions écrites *188*
5.5.4. Analyse des productions écrites *188*
5.5.4.1. Grille d'analyse des productions (d'après la grille du DCL) 188
5.5.4.2. Grille pour l'identification de traces d'actions vis-à-vis des corpus et concordances (d'après Kennedy et Miceli, 2001) 189
5.5.5. Évaluation des productions écrites *189*
5.5.5.1. Grilles d'évaluation pour des tâches complexes 190
5.5.5.2. Grilles d'évaluation pour *Projekt Prospekt* 191
5.5.5.2.1. The Five requirements (Siepmann) 191
5.5.5.2.2. Niveaux de compétence déterminés par le CECR 191
5.5.5.3. Productions évaluées lors de *Projekt Prospekt* 192
5.5.5.3.1. Production de départ 193

5.5.5.3.2. Production finale 193
5.6. POUR RÉSUMER 193
6. "PLONGÉS DANS LE RÉEL" – APPRENANTS ET CORPUS 194
6.1. THÈMES DANS LE DISCOURS DES ACTEURS DE LA RECHERCHE 195
6.1.1. Identification de thèmes récurrents dans les entretiens et dans mon journal de bord 195
6.1.2. Thèmes retenus pour une analyse approfondie 196
6.2. ATTITUDES DES APPRENANTS FACE À L'INPUT EN LE 197
6.2.1. Que veut dire "avoir une attitude d'apprenant-chercheur" ? 197
6.2.2. Regards des apprenants sur l'input en LE 198
6.2.2.1. Regard exhaustif 198
6.2.2.2. Regard synthétique 198
6.2.2.3. Regard analytique 199
6.2.3. Comment les apprenants désignent-ils l'input ? 200
6.2.4. Comment les apprenants utilisent-ils le terme de "concordances" ? 202
6.2.5. Comment les apprenants décrivent-ils la recherche d'informations ? 203
6.2.6. Obstacles perçus par les apprenants lors de l'utilisation de concordances 205
6.2.6.1. Trop d'informations 205
6.2.6.2. Manque de certains types d'informations 206
6.2.6.3. Mise en forme matérielle comme obstacle 207
6.2.7. Apports des concordances perçus par les apprenants 209
6.2.7.1. Voir des formulations toutes faites 209
6.2.7.2. Voir soi-même ses erreurs 209
6.3. ACTIONS DES APPRENANTS VIS-À-VIS DES CORPUS ET CONCORDANCES 211
6.3.1. Utiliser le corpus comme source pour une écriture intertextuelle 211
6.3.1.1. Segments et passages longs co-présents et quantité d'*input* en LE 213
6.3.1.2. Segments co-présents et erreurs de syntaxe 215
6.3.1.3. Sur-utilisation de segments co-présents 216
6.3.1.4. Incorporer l'*input* en LE dans sa production – Jean et le salon du modélisme 216
6.3.1.5. Formes de plagiat 218
6.3.2. Utiliser la concordance comme aide pour la révision de texte 219
6.3.2.1. Les annonces de Marinette et de Bastien 219
6.3.2.2. Vicky et la joueuse de clarinette 221
6.3.2.3. La découverte de Virgile – révision collaborative 224
6.4. ÉVOLUTION DES PRODUCTIONS ÉCRITES DES APPRENANTS 227
6.4.1. Concordances et évolution des adjectifs 227
6.4.1.1. Pourquoi et comment relever les adjectifs dans les corpus ? 227
6.4.1.2. Remarques concernant l'adjectif allemand 228
6.4.1.3. Comment je procède pour étudier les adjectifs 229
6.4.1.4. Les adjectifs dans les productions d'Alexa 230
6.4.1.5. Les adjectifs dans les productions de Camille 231
6.4.1.6. Les adjectifs dans la production de Chloé 233
6.4.1.7. Les adjectifs dans les productions de Christine 234
6.4.1.8. Les adjectifs dans les productions de Fernand 235

6.4.1.9. Les adjectifs dans les productions de Georges 236
6.4.1.10. Les adjectifs dans les productions de Géraldine 237
6.4.1.11. Les adjectifs dans les productions de Janine 238
6.4.1.12. Les adjectifs dans les productions de Jean 239
6.4.1.13. Les adjectifs dans les productions de Léa 241
6.4.1.14. Les adjectifs dans les productions d'Olivia 243
6.4.1.15. Les adjectifs dans les productions de Simon 244
6.4.1.16. Les adjectifs dans les productions de Vicky 245
6.4.2. Les adjectifs dans les productions du groupe classe 246
6.4.2.1. Observations générales 246
6.4.2.2. Segments répétés, segments uniques 248
6.4.2.3. Évolution des types d'erreurs dans le domaine de l'adjectif 250
6.4.3. Concordances et évolution de la syntaxe 251
6.4.3.1. La syntaxe dans les productions d'Alexa 253
6.4.3.2. La syntaxe dans les productions de Géraldine 254
6.4.3.3. La syntaxe dans les productions de Vicky 256
6.4.3.4. Remarques finales concernant la syntaxe 258
6.5. LE CORPUS COMME ESPACE DES NORMES 258
6.5.1. Normes discursives : points d'exclamation dans les textes de type "Visite guidée" 260
6.5.1.1. Points d'exclamation dans le corpus de référence 260
6.5.1.2. Points d'exclamation dans le corpus d'apprenants 261
6.5.2. Normes linguistiques : Olivia et la date 264
6.6. LITTÉRATIE ÉLECTRONIQUE 267
6.6.1. Maîtrise des outils bureautiques 268
6.6.2. Jouer le rôle de tuteur pour ses pairs 270
6.6.3. Indices métatextuels 272
6.6.4. Remarques finales concernant la littératie électronique 272
6.7. INTERDISCIPLINARITÉ 273
6.8. GUIDAGE 275
6.8.1. Guidage en présentiel – le plongeon de Fernand 275
6.8.1.1. Guidage et repérage 275
6.8.1.2. Rôle de la langue maternelle pour la compréhension 276
6.8.2. Guidage à distance 279
6.8.2.1. Annotations des productions écrites 279
6.8.2.2. Formulation d'items dans les questionnaires distribués aux apprenants 284
6.9. TRACES DE DÉCENTRATION AU FIL DE PROJEKT PROSPEKT 286
6.9.1. "De vraies personnes qui parlent" ? 287
6.9.2. Jeux d'identité 290
6.9.3. Écrire en allemand, une "dimension nouvelle de la parole" ? 292
6.10. TÂCHES CONSTRUCTIVISTES ? 293
6.10.1. Complexité, répétitivité et niveau des tâches 294
6.10.2. Tâches "excitantes et provocantes" ? 295
6.10.3. Pour résumer 296

7. OÙ NOUS A MENÉS L'EXPLORATION DES CORPUS ? 297
7.1. AVOIR UN AUTRE REGARD SUR L'INPUT 297
7.2. DISPOSER D'UNE AIDE À LA FORMULATION 298
7.3. ARRIVER À ÉCRIRE DES TEXTES PEU NATIVISÉS 299
7.4. PROGRESSER EN LANGUE ÉTRANGÈRE 299
7.5. RÉVISER DE FAÇON PLUS INDÉPENDANTE 301
8. DESTINATIONS DE RÊVE 303
8.1. DYNAMISER LES SCÉNARIOS BASÉS SUR DES CORPUS 303
8.2. RÉFLÉCHIR À LA LANGUE ÉTRANGÈRE 305
CONCLUSIONS 308
RÉFÉRENCES GÉNÉRALES 309
RÉFÉRENCES "CORPUS ET CONCORDANCES" 336
LISTE DES FIGURES 345
LISTE DES TABLEAUX 347
INDEX DES NOTIONS 349

Introduction

Les corpus offrent un matériau linguistique qui peut servir dans des situations d'écriture diverses. Le scripteur cherche-t-il à rédiger un texte en tant que spécialiste d'un domaine ? Se contente-t-il de s'exprimer dans un registre peu élaboré, ou aspire-t-il à explorer les limites du dicible, à se comporter en poète ou en joueur ? Tout semble possible, du moment qu'il a l'occasion de s'appuyer, pour ce faire, sur un corpus dont le contenu reflète le discours qu'il choisit de tenir. Un tel corpus offre au scripteur des "modèles sociaux" ou des "exemples-types" (Bronckart 1997 : 79), non dans le sens de modèles ou exemples prescrits, mais dans le sens de productions attestées.

Je propose de considérer les corpus comme outils d'aide à la rédaction en LE (Langue Étrangère). Les aides dont on dispose par ailleurs, à savoir les correcteurs orthographiques et syntaxiques, contribuent à surmonter des difficultés concernant le bas niveau. Concernant le haut niveau, il existe des logiciels qui soutiennent la planification et l'organisation des contenus (cf. 3.7.1.). Selon Rastier (2004), les outils que propose la linguistique de corpus permettent de conjuguer les deux niveaux. Ils mettraient en évidence le lien entre le plan du contenu d'un texte et le plan de l'expression.

> La forme intérieure, loin d'être un mystère esthétique, est constituée par les régularités jusqu'à présent imperceptibles de la forme extérieure, celle de l'expression, que les moyens théoriques et techniques de la linguistique de corpus permettent à présent de mettre en évidence (Rastier, 2004).

L'observation précédente de Rastier a été un des éléments déclencheurs de mes recherches. Celles-ci portent sur l'exploitation de corpus lors d'activités de production écrite en allemand, langue étrangère. Mon public s'est composé de quatre groupes d'apprenants, ayant entre 14 et 15 ans et d'un groupe d'apprenants de 10 à 11 ans. Lors de l'exécution des projets, trois des groupes classe étaient en quatrième année d'apprentissage, un groupe en troisième année et un autre en première année. Les séances en présentiel ont eu lieu dans le collège où je travaille comme enseignante d'allemand titulaire. En complément des séances en présentiel, les apprenants avaient l'occasion de travailler à distance et de me contacter en dehors des cours en présentiel s'ils avaient besoin d'aide.

Jusqu'à présent, peu d'études ont décrit les effets de l'exploitation des corpus lors de projets d'écriture en langue étrangère auxquels participent des scripteurs de niveau intermédiaire (cf.

chapitre 4 ; Braun, 2007). Les données empiriques attestant de l'utilisation des corpus par les apprenants d'allemand, langue étrangère, sont particulièrement rares. Les matériaux d'apprentissage, basés sur des corpus, font également défaut pour cette langue (Lüdeling et Walter, 2009 : 7).

> *Beispielbanken werden bereits seit langem in der Vermittlung von Fachsprachen eingesetzt (...). In der Didaktik sollten diese in der Praxis entwickelten Ansätze einer kritischen Prüfung unterzogen werden. Für DaF wurden bislang keine entsprechenden empirischen Arbeiten vorgelegt. Ein ähnliches Bild liefern auch die Lehrmaterialien. Authentische Beispiele sind zwar in vielen aktuellen Lehrwerken zu finden, aber von einem systematischen Einsatz von Korpusdaten kann in DaF keine Rede sein.*
>
> [Dans le domaine de l'enseignement des langues de spécialité, on utilise depuis longtemps des banques de données qui fournissent des exemples de textes (…). Pour son application en didactique, cette méthodologie qui a été développée par des praticiens devrait être examinée de façon critique. En ce qui concerne l'allemand, LE, il n'existe jusqu'à présent pas de recherches empiriques. Il en est de même pour les matériaux d'apprentissage. Des exemples authentiques se trouvent dans de nombreux manuels d'apprentissage récents, mais on est loin d'une exploitation de données de corpus en allemand, LE.].

La quasi-absence de méthodes et de matériaux conçus pour l'allemand a motivé mon choix de m'inspirer, dans un premier temps, de recherches menées en didactique du français écrit, langue maternelle, concernant l'utilisation de bases de données textuelles comme aide à la rédaction. Crinon, Legros et Marin (2002b : 187-188) nomment deux conditions qui déterminent la qualité de la production écrite en LM (Langue Maternelle) : il convient de mettre à disposition des apprenants un nombre important de textes. Ces textes doivent être compatibles avec ceux que les apprenants sont censés produire, c'est-à-dire qu'ils doivent appartenir au même genre textuel que celui représenté dans la base de données. Les corpus semblent répondre à cette double demande. Crinon, Legros et Marin (2002b : 187) plaident pour une nouvelle attitude de consultation de textes, celle de "chercheur d'informations". Ils dressent un parallèle entre cette attitude et une démarche d'écriture intertextuelle : les scripteurs sont encouragés à chercher des éléments pertinents dans les bases de données afin d'enrichir leur production écrite. Jorro précise que le développement d'une telle attitude dépend du type d'accompagnement qu'un scripteur expert offre à un scripteur novice.

> Ces relations qui font que l'expert cheminera avec le novice sur un autre mode que celui de la correction ou remédiation ouvrent un espace de travail quasi singulier. L'expert cherchant à multiplier les pratiques d'écriture, à faire en sorte que le sujet écrivant travaille plusieurs textes dans un souci

> artisanal en prenant appui sur ses propres écrits, en empruntant ailleurs ce qui lui convient (Jorro, 2004 : 157).

Dans le domaine de la linguistique de corpus, le concept d'apprenant-chercheur, présenté, entre autres, dans les travaux de Johns (1991 ; cf. 2.6.3), fait écho à ce qui précède. J'ai croisé ce concept de type constructiviste avec les apports de la psycholinguistique concernant la recherche d'informations (Rouet *et al*., 2004) et avec la technique de recyclage d'éléments repérés dans les corpus (cf. 3.4.2.3.2.).

Mes questions de recherche sont les suivantes.

Quels effets peuvent être constatés quand on propose à des apprenants adolescents, ayant un niveau intermédiaire en langue étrangère, un scénario de type constructiviste qui conçoit les corpus spécialisés comme aide à la rédaction ? Peut-on observer l'émergence d'une attitude d'apprenant-chercheur, susceptible de soutenir l'exploitation de l'*input* lors de la rédaction ? Le scénario permet-il d'obtenir des productions dont les caractéristiques linguistiques et discursives reflètent les normes attestées dans les corpus ? Le projet, aide-t-il les apprenants à se décentrer de leur langue maternelle et de leurs représentations de la LE ? Mène-t-il au développement de leur système linguistique ?

Afin de pouvoir apporter des éléments de réponse à ces questions, j'ai mis en place une recherche-action d'une durée de trois ans. Les constantes ont été l'âge du public (à part pour un des groupes classe) et l'utilisation de corpus spécialisés et de lignes de concordance lors de l'exécution de tâches de production écrite.

Le premier chapitre de la thèse est consacré aux corpus et concordances. Il les distingue d'objets qui leur sont proches, et il évoque les objectifs liés au travail sur corpus. J'y propose un état des lieux concernant les corpus et concordanciers existants en langue allemande. Le deuxième chapitre s'intéresse aux recherches dans le domaine de l'apprentissage et de l'enseignement des langues étrangères. Il présente une palette de points de vue oscillant entre le bord linguistique et le bord culturel et il décrit des méthodologies d'enseignement/apprentissage qui sont en accord avec ma recherche. Le troisième chapitre traite de la production écrite. J'y décris, entre autres, des modèles d'écriture et le rôle de l'*input* et de l'*output* dans la production écrite. La dernière partie de ce chapitre s'interroge sur le lien entre les outils d'aide à la rédaction et les ressources pédagogiques. Le quatrième chapitre

reprend des pistes proposées dans les trois chapitres précédents afin d'évaluer le potentiel des corpus et concordances pour l'apprentissage de la production écrite en LE. L'analyse d'activités existantes, basées sur des concordances, et la présentation d'une activité que j'ai développée terminent ce chapitre. Le chapitre 5 situe ma recherche par rapport à d'autres types de recherche. Il présente les conditions de mise en place d'une recherche-action dans un établissement du secondaire. J'y décris mon scénario, les outils que j'ai proposés aux apprenants, ainsi que les activités préparatoires et d'entraînement. Le sixième chapitre propose l'analyse des données rassemblées lors de ma recherche. J'y étudie le discours et les actions des apprenants vis-à-vis des corpus et concordances et j'observe l'évolution de certaines caractéristiques de leurs productions écrites. Dans ce chapitre, je m'intéresse également aux traces qui permettent d'avancer que mon projet à amené les apprenants à intégrer de nouveaux aspects linguistiques et culturels dans leur système de représentation de la langue allemande. Le chapitre 7 regroupe les éléments de réponse à mes questions qui débouchent, dans le chapitre 8, sur l'esquisse de chantiers de recherche futurs dans le domaine de l'apprentissage de LE à l'aide de corpus.

Quelques remarques concernant la terminologie que j'emploie dans ma thèse permettront de clarifier certaines facettes de mon positionnement de chercheuse.

Premièrement, les termes d'"apprenant" et d'"enseignant", utilisés dans le milieu de la didactique, soulignent mon point de vue que la relation pédagogique doit avoir des traits d'une relation de compagnonnage (Narcy-Combes, 2005 : 121-122). Malgré la hiérarchie qui s'exprime dans les termes de "professeur" et d'"élève", termes qui sont d'usage dans le milieu de l'enseignement secondaire, je les ai utilisés quand le contexte l'a rendu nécessaire.

Deuxièmement, il convient de clarifier comment on désigne la langue apprise tout de suite après la naissance et la langue apprise ultérieurement, souvent dans un contexte institutionnel. Les termes de "L1" et de "L2", qui sous-entendent "langue apprise en premier/en second", sont utilisés dans le domaine de l'acquisition (Matthey et Véronique, 2004). Ils apparaissent dans ma thèse dans le contexte de citations issues de ce domaine. Les termes de "langue maternelle" et de "langue étrangère" renvoient au domaine de la didactique (Defays, 2003). Je les ai fréquemment employés dans mon texte sous la forme abrégée "LM" et "LE".

Troisièmement, en principe, pour les noms désignant des personnes, je privilégie l'indication aussi bien de la forme féminine que de la forme masculine. J'ai renoncé à cette double dénomination dans ma thèse, au profit de la forme masculine, afin d'alléger le style d'écriture.

1. Corpus et concordances

1.1. Aperçu historique

Le regroupement de textes ayant des caractéristiques communes, est une pratique ancienne, comme en témoigne par exemple l'existence du *Codex Manesse* comportant des chansons de poètes allemands rassemblées au 14ème siècle. Si l'objet résultant d'un tel regroupement est motivé par un objectif d'étude particulier, on l'appelle "corpus". Meyer (2002 : xii) rappelle que certains ouvrages d'historiens de la langue du début du 20ème siècle n'existeraient pas sans l'étude d'un corpus, comme par exemple l'ouvrage *A Modern English Grammar on Historical Principles* de Jespersen (1909 - 1949). Dans l'introduction au numéro 28 de *Linguistik Online*, dédié à la linguistique de corpus, Duffner et Näf (2006a) esquissent le contexte de la "renaissance" assez récente des linguistiques de corpus, c'est-à-dire des diverses approches d'étude linguistique basées sur des collections de textes contenant des énoncés attestés. Jusqu'après la deuxième guerre mondiale, les sciences du langage ont été marquées par des thèmes liés à la diachronie. Les données de corpus ont été observées et classées selon des méthodes de recherche empiriques. Depuis la deuxième moitié du 20ème siècle, la démarche empirique cède le terrain à des recherches d'ordre plus théorique, comme par exemple celles effectuées par les représentants de l'école générative ou le structuralisme (Fuchs et Le Goffic, 1992). L'étude de corpus est remplacée, ou tout au moins secondée, par l'étude de phrases isolées. Depuis les années 60, les corpus numériques ouvrent de nouvelles perspectives d'étude de la langue. Kucera et Francis posent les premiers jalons dans ce domaine. Ils montrent dans leur ouvrage *Computational Analysis of Present-Day American English* (1967) comment on peut analyser de très grands corpus numériques en s'intéressant à des aspects variés de la langue. À partir des années 80, grâce à l'évolution des nouvelles technologies, le potentiel d'exploitation des corpus augmente de façon spectaculaire (Habert, Nazarenko et Salem, 1997 : 7). En effet, la numérisation et la mise en ligne des données permettent de les examiner de façon beaucoup plus confortable. On peut alors observer un regain d'intérêt pour des recherches empiriques. L'observation des données y est considérée comme élément fondamental de la théorisation. On peut s'attendre à ce que cette évolution mène à de nouvelles descriptions linguistiques (Duffner et Näf, 2006a). Le domaine des corpus et des linguistiques de corpus est actuellement en pleine expansion. Les corpus s'agrandissent ou se

spécialisent, et il y a une tendance croissante de mise à disposition en ligne, souvent gratuite, des données.

Afin de mieux cerner le phénomène des corpus, regardons de plus près les diverses utilisations de ce terme. Selon le *TLFi (Trésor de la langue française informatisé)*, le mot "corpus" est attesté en France dans le domaine de la religion chrétienne depuis le 13ème siècle. Il apparaît depuis la fin du 19ème siècle dans le domaine du droit et à partir des années 60 du 20ème siècle dans le domaine de la linguistique.

> Étymol. et Hist. 1. Fin XIIe s. corpus Deu "hostie" (Mort Garin, 132 ds T.-L.) ; 1206 corpus Domini (GUIOT, Bible, 1223 ds GDF.) (…) ; 2. 1863 corpus juris, corpus "collection du droit romain" (LITTRÉ) ; 3. 1890 "collection d'inscriptions de l'Antiquité" (DG) ; 4. 1961 "ensemble d'énoncés servant de base à l'analyse linguistique" (Lar. encyclop.). Au sens 2, empr. au lat. class. corpus juris, v. corps. Sens 3 et 4 développés en fr. à partir de 2. (TLFi).

L'origine du mot "corpus" fait donc appel aux codes, à ce qui "fait foi" en société. Dans quels contextes le terme allemand "*Korpus*", équivalent à "corpus", apparaît-il au 20ème siècle ?

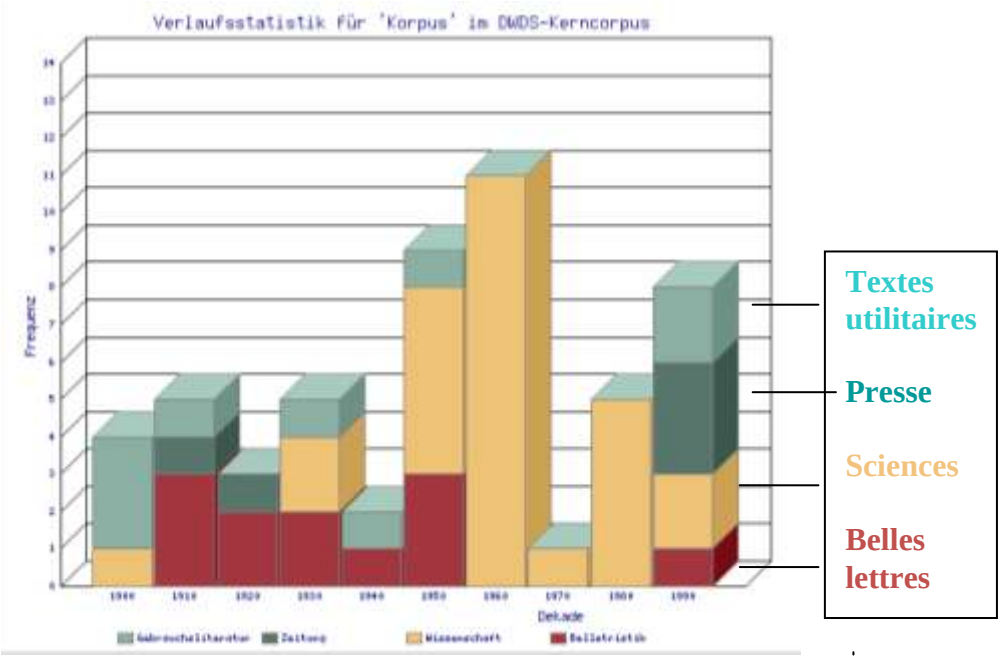


Figure 1 – Évolution de l'utilisation du terme "*Korpus*" entre 1900 et 1990.

Pendant le 20ème siècle, le terme de corpus a d'abord plutôt été réservé à la désignation de bases de données littéraires. Ensuite, il a été utilisé presque exclusivement en rapport avec des textes scientifiques, ceci pendant les années 60 à 80. Il finit par apparaître en relation avec les quatre domaines représentés, à savoir les textes utilitaires, la presse, les sciences et les belles lettres, ceci depuis les années 90. L'apparition des corpus dans des domaines variés entraîne une hétérogénéisation et complexification du terme de "corpus".

1.2. Qu'est-ce qu'un corpus ?

La définition de "corpus" fait défaut dans de nombreux articles présentant des résultats en linguistique de corpus, comme le souligne Meyer (2002 : xi). En 2008, la revue *Corpus* lance un appel à contribution pour le numéro 8 (novembre 2009), concernant des articles destinés à clarifier la terminologie dans ce domaine.

> Dans ce moment et ce mouvement, texte et corpus sont ébranlés jusque dans leurs fondements définitionnels. Deux champs disciplinaires des sciences du langage encore très jeunes, linguistique de corpus et linguistique textuelle, sont dans une redistribution permanente de leurs attendus respectifs et de leurs frontières. (…) Quelle terminologie sont amenés à mettre en place les chercheurs confrontés à la nécessité de définir leurs objets dans le cadre de l'analyse textuelle sur corpus : texte, document, archive, énoncé, discours, base, corpus, sous-corpus, section, unité, segment, partition... ? (Adam et Viprey, 2008).

La complexité de l'objet m'amène à le décrire en procédant d'abord par élimination et ensuite par comparaison à des objets avec lesquels il partage certaines caractéristiques, à savoir les autres types de ressources, pour établir ensuite une définition de "corpus" à proprement parler.

1.2.1. Définition négative de "corpus"

Un corpus n'est pas une simple suite de signes et de lettres, c'est-à-dire, il n'est pas simplement "du texte" (Bommier-Pincemin, 1999), mais il est le résultat d'un regroupement de textes. Il n'est pas non plus une collection aléatoire de textes, telle que la Toile (Sinclair, 2005), mais le résultat d'un regroupement raisonné de textes ou parties de textes.

1.2.2. Distinction entre corpus et autres ressources numériques

Comme d'autres ressources, les corpus contiennent des données, mises à disposition en ligne ou hors ligne. Avant de procéder à une définition concise du terme de "corpus", je propose d'observer sa spécificité par rapport à d'autres ressources, à savoir les portails, banques de données, archives et bases de données.

Les portails sont des sites complexes qui donnent accès à des sites plus spécialisés, à des banques de données, des applications et des services. Ils peuvent rassembler une communauté qui partage les mêmes besoins, par exemple les abonnés à un hébergeur de sites. *Le grand dictionnaire terminologique* (Office québécois de la langue française, nd) précise que le portail est un

> Site *Web* dont la page d'accueil propose, en plus d'un moteur de recherche, des hyperliens avec une foule d'informations et de services utiles et attrayants, qui est conçu pour guider les internautes et faciliter leur accès au réseau. (…) Les portails, ces portes d'accès au réseau Internet, peuvent proposer, par exemple, un service de courrier électronique gratuit, une sélection de moteurs de recherche, de bavardoirs (chatrooms), les actualités, la météo, les cotes de la Bourse, un classement de sites par thèmes (ex. : cinéma, voyage, sport), des raccourcis pour les achats en ligne. L'objectif des portails est, du point de vue de leurs créateurs, d'attirer et de fidéliser les internautes au point de devenir leur porte d'entrée dans Internet, c'est-à-dire la page de démarrage du plus grand nombre d'entre eux. Il existe des portails généralistes et des portails thématiques (…).

Voici quelques exemples de portails.

Virtuose, "le portail documentaire du Service commun de documentation de l'université de la Sorbonne nouvelle". Ce portail contient une rubrique qui s'appelle "bases de données".
Tematice, "Portail de la Recherche sur les TIC pour l'Éducation et la Formation". Ce portail contient une section qui s'appelle "Archives".

Leipziger Wortschatz-Portal. Ce portail donne accès à un dictionnaire unilingue et à des dictionnaires multilingues, mais aussi aux données du serveur (*Webservice*) et à la rubrique "*Wörter des Tages*" [Les mots du jour] qui indique les mots utilisés le plus souvent un jour donné.

Le site *deutschland.de* se donne le sous-titre "*das Deutschland-Portal*" [portail de l'Allemagne], mais il n'a pas toutes les caractéristiques d'un portail selon la définition donnée ci-dessus. En effet, ce site ne propose ni l'accès à d'autres sites (sauf par des hyperliens), ni à des services, ni à des applications.

Les termes "banque de données" et "base de données" sont parfois utilisés comme synonymes, comme c'est le cas pour le dictionnaire en ligne *Your Dictionary.com*.

> *databank*
> *Variant of database*
>
> *data•base (-bās')*
> *a large collection of data in a computer, organized so that it can be expanded, updated, and retrieved rapidly for various uses*
> *any large or extensive collection of information*
> *Also data base or databank da'ta•bank' (-baŋk')*.

Selon *Le grand dictionnaire terminologique*, les banques de données se trouvent hiérarchiquement au-dessus des bases de données.

> On confond parfois la base de données et la banque de données. Cette dernière est un ensemble d'informations relatives à un domaine défini de connaissances et organisées pour être accessible par plusieurs utilisateurs. Une banque de données regroupe souvent plusieurs bases de données.

Un exemple de banque de données en ligne est *La banque de données européenne. Les femmes dans la prise de décision.* On peut y interroger plusieurs bases de données. La *Banque de données Santé publique* contient, entre autres, une rubrique qui s'appelle "base documentaire". *VIRTUOSE*, le portail documentaire du "Service commun de documentation de l'université de la Sorbonne nouvelle", place les archives en ligne au niveau des bases de données. J'adhère à ce choix de classification qui est proposé par des spécialistes du domaine de la documentation.

Les corpus se distinguent des portails et des banques de données par leur architecture moins complexe. Ils se situent hiérarchiquement en-dessous de ces deux types de ressources. Ils peuvent être placés au même niveau que les archives. Je considère les corpus et les archives comme des formes spécifiques de bases de données. Les bases de données, archives et corpus contiennent des informations sous forme papier ou électronique. Les trois peuvent contenir des textes écrits et/ou des données de type multimédia. Les corpus se distinguent des bases de données et des archives par leur objectif d'utilisation. Une base de données textuelles n'est pas un corpus si les textes n'y sont pas regroupés selon des critères de genre, de discours et d'application, car "les données sont ce qu'on se donne" (Rastier 2004). Les archives sont faites dans le but principal de conserver des données, tandis que les corpus invitent à étudier les données. Le *Trésor de la langue française informatisé* donné la définition suivante d'"archives".

> Ensemble de documents hors d'usage courant, rassemblés, répertoriés et conservés pour servir à l'histoire d'une collectivité ou d'un individu.

Voici un choix d'archives allemandes à partir desquelles on pourrait créer des corpus susceptibles d'être utiles dans un contexte d'l'enseignement/apprentissage des langues.

sueddeutsche.de Zeit online Spiegel online Stern.de	Archives de quotidiens et revues allemands permettant la consultation – souvent gratuite – d'articles récents.
Project Gutenberg	Le *Project Gutenberg* est un projet américain destiné à diffuser la littérature du domaine public sous forme de livres électroniques. Il s'agit d'une simple collection de textes. En octobre 2008, il y a 549 documents allemands en ligne.
Mannheimer Texte online (MATEO)	L'allemand écrit académique (résumés gratuits, articles récents payants) ; reproductions d'œuvres imprimées (textes jusqu'à 300 ans après l'invention de l'imprimerie).

Tableau 1 – Archives en ligne en langue allemande.

Les distinctions faites dans ce paragraphe m'amènent à proposer la hiérarchie suivante pour les termes "portail", "banque de données", "base de données", "archives", "corpus", "document" et "texte".

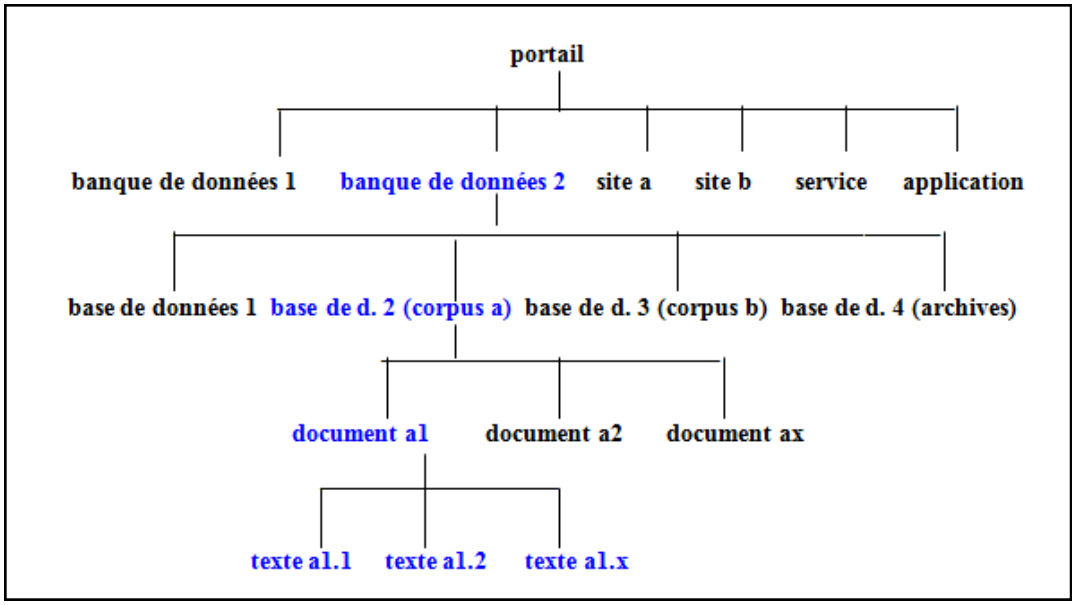


Figure 2 – Position hiérarchique des corpus par rapport à d'autres ensembles publiant du texte.

Dans la figure 2, nous trouvons les termes de "document" et de "texte". J'emprunte ces termes à un modèle de hiérarchisation proposé par l'*IDS* (*Institut für Deutsche Sprache*) de Mannheim. Selon ce modèle, les corpus rassemblés dans le cadre du projet *DeReKo* (*Deutsches ReferenzKorpus*) comportent, en effet, un ou plusieurs documents qui contiennent un ou plusieurs textes. Les textes sont regroupés selon plusieurs types de critères, par exemple source, chronologie, thème ou genre de texte. Sous la dénomination de "texte", nous trouvons des articles de journaux, des extraits d'ouvrages entiers ou des ouvrages entiers. L'*IDS*

propose une hiérarchie à quatre niveaux. L'archive contient des corpus, le corpus contient des documents, le document contient des textes. Voici la hiérarchie entre corpus, document et texte, établie par l'*IDS* au sujet des œuvres de l'auteur Siegfried Lenz.

Anz. Texte [Nombre de textes]	Dokument [Document]	Beschreibung [Description]	Bd. [Vol.]
1	LES/HIL.00000	Es waren Habichte in der Luft. Roman [titre d'un roman]	1
...	...	...	...
1	LES/ALE.00000	Die Auflehnung. Roman [titre d'un roman]	12
77	LES/ERZ.13001 [-16022]	[Erzählungen] [nouvelles]	13-16
3	LES/SCH.17001 [-17003]	[Schauspiele] [Pièces de théâtre]	17
4	LES/HOR.18001 [-18004]	[Hörspiele] [Pièces radiophoniques]	18
98	LES/ESS.19001 [-20032]	[Essays] [Essais]	19+20

Tableau 2 – Hiérarchie entre corpus, document et texte dans *DeReKo*.

Le tableau 2 montre que les documents peuvent contenir entre un et plusieurs textes, ici allant jusqu'à 98 pour les essais de Lenz. Les essais sont regroupés dans le document nommé "LES/ESS.19001 [-20032]" qui, lui, fait partie du corpus nommé "Siegfried Lenz : Œuvres complètes en volumes individuels [20 vol.]".

1.2.3. Définitions de "corpus"

En 1961, le sens de corpus comme "ensemble d'énoncés servant de base à l'analyse linguistique" est premièrement attesté dans le *Larousse encyclopédique*, cité sur le site du *Trésor de la langue française informatisé*. Dans le *Larousse encyclopédique en ligne*, la définition de "corpus" est réactualisée comme suit.

> Recueil de documents relatifs à une discipline, réunis en vue de leur conservation.
>
> Ensemble fini d'énoncés écrits ou enregistrés, constitué en vue de leur analyse linguistique.

Tandis que la première définition semble davantage convenir à une définition du terme "archives" (cf. 1.2.2.), la deuxième est une bonne définition de base pour l'objet que je tente de circonscrire. Elle se recoupe avec la définition de Bußmann, reformulée par Bubenhofer (2006).

> *Corpus (Neutr., Pl. Corpora; lat. corpus 'Körper'). Endliche Menge von konkreten sprachlichen Äußerungen, die als empirische Grundlage für sprachwiss. Untersuchungen dienen. Stellenwert und Beschaffenheit des C.*

> *hängen weitgehend von den je spezifischen Fragestellungen und methodischen Voraussetzungen des theoretischen Rahmens der Untersuchung ab.* [Corpus (…). Une quantité finie d'énoncés langagiers réellement attestés, servant de base pour des recherches linguistiques. Le statut et la nature du corpus dépendent en grande partie des questions de recherche et des choix méthodologiques du cadre théorique de l'étude.] (Définition de Bubenhofer, 2006 ; d'après Bußmann, 2002 : 143).

"Une quantité finie d'énoncés" reste une description très générale de l'objet "corpus". Regardons l'usage contemporain de "corpus", tel que le *Trésor de la langue française informatisé* (*TLFi*) le décrit. Nous y trouvons quatre définitions. La première a trait au domaine de la philologie et des sciences humaines.

> Recueil réunissant ou se proposant de réunir, en vue de leur étude scientifique, la totalité des documents disponibles d'un genre donné, par exemple épigraphiques, littéraires, etc.

La deuxième est une définition du domaine de la linguistique.

> Ensemble de textes établi selon un principe de documentation exhaustive, un critère thématique ou exemplaire en vue de leur étude linguistique.

La troisième concerne le domaine du droit.

> Recueil, collection de droit romain.

La quatrième définition est issue du domaine de l'électronique et de l'informatique.

> Ensemble de données exploitables dans une expérience d'analyse ou de recherche automatique d'informations.

Les quatre définitions ci-dessus contiennent toutes un élément que je retiens pour ma propre définition de "corpus" : genre, étude linguistique, droit et recherches d'information. J'interprète ici le terme de "droit" dans le sens de "code" ou "ce qui indique une norme", ou encore, selon la définition du terme "droit" donné dans le *TLFi*, "Fondement des règles régissant les rapports des hommes en société". Contrairement à ce qui est exprimé dans la première définition du *TLFi*, un corpus ne contient pas forcément "la totalité" de documents représentant un genre donné. Ceci n'est possible et nécessaire que pour certains domaines ou thèmes. Parler d'un "ensemble de données" (quatrième définition) ne me paraît pas suffisamment précis pour définir un corpus. Il convient de distinguer entre corpus et ensemble de mots ou "sac de mots", selon l'expression utilisée par Rastier (2004). Un "sac de mots" est une grande quantité de données rassemblées de façon aléatoire, parfois seulement parce qu'il est possible d'y avoir facilement accès (Habert, 2000). De très grands corpus rassemblant toutes sortes de données peuvent toutefois servir pour tester la performance de logiciels

(Leech, 2003) ou pour apprendre à l'ordinateur à interpréter et à produire du langage humain (Partington, 2006). Tognini Bonelli (2001) définit les corpus comme "*a collection of texts assumed to be representative of a given language*", donc comme un ensemble de textes qui représentent une langue donnée. La représentativité est l'objectif principal des corpus de référence. Habert (2000) explique ce concept comme suit.

> Un corpus de référence est conçu pour fournir une information en profondeur sur une langue. Il vise à être suffisamment grand pour représenter toutes les variétés pertinentes de cette langue et son vocabulaire caractéristique, de manière à pouvoir servir de base à des grammaires, des dictionnaires et d'autres usuels fiables » [Sinclair, 1996, p. 10]. Il s'agit donc d'associer étroitement deux caractéristiques : une taille suffisante et la diversité des usages représentés. Le British National Corpus les allie effectivement.

Selon Bommier-Pincemin (1999), la représentativité est toutefois un concept problématique. Même les très grands corpus ne peuvent donner un aperçu de la langue "générale" ou "moyenne" ou "standard", car, selon elle, une telle langue n'existe pas : "les textes sont tous pris dans des pratiques qui les contextualisent" (Bommier-Pincemin, 1999). Dissimuler les contextes des textes rend donc une analyse linguistique problématique, voire impossible. Bommier-Pincemin définit par conséquent le corpus comme un ensemble qui contient des textes. Elle le différencie d'un ensemble qui contient du texte et qui ne permet pas d'observer la langue en tenant compte du contexte. Selon Bommier-Pincemin (1999), un corpus doit remplir trois conditions. Premièrement, la condition de signifiance ; le corpus doit permettre d'étudier un objet particulier. Deuxièmement, la condition d'acceptabilité ; le corpus doit correspondre à l'objet et le représenter suffisamment. Troisièmement, la condition d'exploitabilité ; le corpus doit contenir un nombre suffisant de textes, et ces textes doivent être comparables entre eux.

Rastier (2004) propose la définition de corpus suivante :

> Un corpus est un regroupement structuré de textes intégraux, documentés, éventuellement enrichis par des étiquetages, et rassemblés : (i) de manière théorique réflexive en tenant compte des discours et des genres, et (ii) de manière pratique en vue d'une gamme d'applications (Rastier, 2004).

En m'appuyant sur les quatre entrées proposées dans le *TLFi*, l'étude linguistique, le genre, la recherche d'information et le code, et en m'inspirant de Rastier, Bommier-Pincemin et Habert, je propose la définition suivante de "corpus".

Un corpus est un ensemble de données linguistiques attestées. Il peut s'agir de données écrites, orales ou de type multimédia. Un corpus est une forme particulière de base de données. Il contient des parties de textes ou des textes entiers qui sont comparables entre eux parce qu'ils ont des caractéristiques communes, comme par exemple l'appartenance à un même genre textuel. La particularité du corpus par rapport à d'autres types de bases de données est l'objectif de théorisation qui sous-tend sa constitution. L'objectif principal n'est pas premièrement le stockage ou la diffusion des données, mais l'étude des données ou l'étude de leur impact dans un domaine d'application. Des objectifs de type théorique peuvent être l'étude linguistique, stylistique ou autre des données. Un des objectifs de type "théorie appliquée" peut être la recherche d'informations dans un contexte d'apprentissage d'une langue étrangère. Un corpus doit contenir un nombre suffisant de textes représentatifs. La quantité de textes est considérée comme suffisante si elle permet de mener de façon satisfaisante les observations qui motivent la constitution du corpus. Les énoncés des textes du corpus ont un statut de norme linguistique, norme spécifique au genre textuel qui y est représenté, le cas échéant.

Figure 3 – Ma définition de "corpus".

1.2.4. Types de corpus

Habert, Nazarenko et Salem (1997 : 144) proposent la typologie de corpus suivante.

– "Corpus de référence" : de tels corpus visent à contenir la globalité des données dont on dispose (concernant une langue, un genre textuel, etc.).
– "Corpus d'échantillons" : de tels corpus sont constitués de phrases ou de parties de phrases. Ils ne donnent pas accès aux textes complets d'où proviennent les mots clés.
– "Corpus de textes complets".
– "Corpus spécialisés" : de tels corpus sont limités à une situation de communication, un domaine, une langue de spécialité, etc.
– "Corpus comparables" : de tels corpus contiennent des textes similaires dans plusieurs langues.
– "Corpus parallèles" : de tels corpus contiennent des textes traduits dans deux langues ou plus (les débats du parlement européen par exemple).

Il est possible d'ajouter à cette liste les "corpus d'apprenants" (Granger *et al.*, 2002) qui rassemblent les textes produits par des apprenants. Il convient aussi de mentionner qu'il n'y a

pas seulement des corpus de langue écrite, mais aussi des corpus de langue orale et que les corpus multimodaux combinent même ces deux entrées. Pour obtenir un aperçu plus complet des types de corpus existants, il est utile de consulter le site *Bookmarks for Corpus Linguists* de Lee (nd).

Quand on choisit un corpus pour une étude donnée, doit-on opter pour la collection d'échantillons de la langue, des parties de textes ou plutôt pour la collection de textes entiers ? Convient-il de chercher à "équilibrer" le contenu des corpus ou plutôt de rassembler des documents qui représentent un genre textuel ? De tels choix ne sont pas anodins ; le positionnement linguistique et les principes épistémologiques du concepteur de corpus peuvent s'y révéler. Rastier (2004) distingue entre deux conceptions qui sous-tendent la création de corpus. Il appelle la première "conception documentaire" ou "logico-grammaticale". Des corpus conçus selon cette logique sont des "corpus de mots" (projet *Paroles*), des "corpus d'attestations", des "corpus d'exemples" (*Frantext*) qui ne donnent pas accès aux sources, et des "corpus de fragments", donc des corpus qui ne contiennent qu'une partie des textes d'où provient le mot recherché (*British National Corpus*). La deuxième conception est une conception philologique-herméneutique. Elle permet de créer des "corpus réfléchis" (Rastier, 2004). Dans de tels corpus, les textes sont regroupés selon le principe critique appliqué à l'étude philologique des textes, donc en tenant compte des rapports de texte à texte. La conception philologique-herméneutique induit un certain type de classement de textes. En effet, Rastier recommande de classer les textes d'un corpus en indiquant dans leur en-tête au moins les trois informations suivantes : le discours (par exemple, littéraire), le champ générique (par exemple, genres narratifs) et le genre proprement dit (par exemple, récit de voyage). Selon cette conception, il est indispensable d'attester des textes dans leur globalité.

La recherche d'informations dans un corpus est facilitée par leur numérisation. Les possibilités d'étude linguistique sur corpus s'en trouvent profondément modifiées. Sinclair (1996) définit les corpus numériques comme suit.

> *A computer corpus is a corpus which is encoded in a standardised and homogenous way for open-ended retrieval tasks. Its constituent pieces of language are documented as to their origins and provenance.*
>
> [Un corpus électronique est un corpus qui est encodé de façon standardisée et homogène, ceci afin de permettre de faire des requêtes ouvertes. On y mentionne l'origine et la provenance des textes qui le constituent.]

Les corpus numériques rassemblent des données textuelles électroniques sous un format standardisé, donc selon des conventions de présentation des données fixées auparavant. Les données sont regroupées et souvent annotées selon de multiples critères, en fonction de la perspective d'utilisation choisie. Ils sont de taille très variable. On peut, par exemple, rassembler tous les textes d'un auteur donné ou tous les témoignages d'une langue de spécialité. Il est possible de se concentrer sur une période donnée ou de regrouper tous les témoignages concernant une langue rare. On peut aussi s'intéresser aux données orales ou multimodales. Les corpus sont souvent codés. Un codage n'est jamais neutre : il convient de s'interroger sur ses implications épistémologiques (cf. 1.3.2.).

1.3. Annoter un corpus

1.3.1. Pourquoi et comment annoter un corpus ?

Un corpus est composé de textes. Il est possible d'étudier les corpus à partir des données textuelles brutes. On peut, par exemple, compter les unités du corpus. Les unités "mots" ou "*tokens*" sont définis ici comme unités délimitées par deux blancs ou par des signes de ponctuation (Siepmann, 1999). Il est également possible de rechercher les occurrences d'un mot ou segment, donc d'une suite ou combinaison de mots. Afin de faciliter les recherches sur corpus, on procède à des annotations. Les annotations sont des informations linguistiques ou des métadonnées, donc des informations extralinguistiques, que l'on "attache" aux éléments dont se compose le corpus. Habert, Nazarenko et Salem (1997 : 7) définissent l'annotation comme "l'ajout d'information, de quelque nature qu'elle soit : morphologique, syntaxique, sémantique, prosodique, critique". L'annotation peut être faite au niveau paradigmatique, donc pour chaque mot du corpus. Ce procédé est appelé "*tagging*". Il peut aussi être effectué au niveau syntagmatique et donc concerner des segments plus longs. Il peut s'agir d'unités de deux mots ou plus, de phrases ou d'unités au-delà de la phrase, par exemple des paragraphes ou des sections, etc. (Pincemin, 2006). Ce deuxième procédé est appelé "*parsing*". Aussi bien le procédé de *tagging* que le procédé de *parsing* se servent d'une base de données qui contient des catégories qui définissent les classes de mots et les types de syntagmes auxquels on peut s'attendre dans un texte d'une langue donnée. Le métalangage utilisé est en général celui de la grammaire traditionnelle, comme par exemple dans le *STTS* (*Stuttgart/Tübinger Tagset* ; Schiller *et al.*, nd). Un autre choix est proposé par Partington (2006) qui s'appuie sur la *SFG* (*Systematic Functional Grammar*). Partington distingue entre deux types d'informations qu'une annotation peut fournir. Premièrement, il est possible de rajouter des informations qui

indiquent le "statut grammatical" du mot dans le contexte où il apparaît (on indique la classe de mot, le genre, nombre, etc.). Deuxièmement, les annotations peuvent avoir un autre contenu (marquage des sections du texte, des informations concernant le sexe, l'âge du locuteur, etc. ; les changements de thème, etc.).

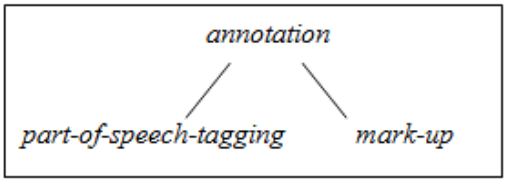


Figure 4 – Types d'annotation d'après Partington (2006).

On peut s'interroger sur le choix fait par Partington de distinguer entre une information qui rappelle une grammaire de type structuraliste et d'autres informations, par exemple de type discursifs, qui se prononceraient entre autres sur le discours, mais qui sont regroupées avec des informations de type extralinguistique. Cette distinction reflète la bipartition proposée par Firth (cité dans Wilcock, 1993) pour décrire deux points de vue permettant d'observer la langue/parole.

"anatomy"	***"physiology"***
chain	*choice*
syntagmatic	*paradigmatic*
structural	*systemic*
formal	*functional*
logical	*rhetorical*

Tableau 3 - *The "anatomy" and "physiology" of language* (Firth).

En effet, Partington se réfère explicitement à la *SFG* (*Systemic Fonctional Grammar*), développée par Halliday d'après les travaux de Firth (Wilcock, 1993). Selon Partington, cette théorie grammaticale, appliquée à l'étude de corpus, aide à mettre en relation la description grammaticale et le sens.

Comme le précisent Habert, Nazarenko et Salem (1997 : 23-24), on peut faire le choix d'étiqueter en profondeur une partie du corpus ou de tout étiqueter de façon plus fruste. Ce choix peut dépendre de l'application que l'on vise, par exemple la création d'un dictionnaire, d'une grammaire descriptive ou l'apprentissage d'une langue étrangère. L'étiquetage est toujours un compromis entre la description linguistique fine et les possibilités qu'offre le système d'annotation automatique. Certaines catégories, comme la notion de *mass noun*, sont difficiles à représenter. Dans le jeu d'étiquetage se révèlent des divergences théoriques réelles. Quels mots convient-il, par exemple, de regrouper sous l'étiquette "adverbe", "déterminant",

"pronoms" ? Selon les langues, il y a de grandes différences d'appréciation sur les catégories (Véronis et Khouri, 1995 : 237, cité dans Habert, Nazarenko et Salem 1997 : 24). Leech (2004) évoque la nécessité d'utiliser des catégories pour lesquelles existe un large consensus parmi les linguistes. Ce dilemme est un argument en faveur de la création de corpus qui, dans un premier temps, sont non étiquetés. Des expérimentations ou recherches-actions peuvent permettre de préciser les besoins d'étiquetage ou, au contraire, montrer que les corpus non étiquetés peuvent convenir à certains types de projets.

1.3.2. Peut-on se passer d'annotations ?

À certains stades de la recherche linguistique, en particulier quand il s'agit de créer de nouvelles descriptions, il convient de travailler avec un corpus non annotée (Teubert, 2005 ; Sinclair, cité dans Leech, 2004). Autrement, les représentations du chercheur qui s'expriment dans les catégories qu'il emploie risquent de l'empêcher de découvrir de nouveaux faits linguistiques ou même de fausser les données. Il paraît également défendable de se passer d'annotations si l'on ne fait pas une étude linguistique sur corpus, mais si l'on propose à des apprenants un travail avec corpus. Certaines tâches d'apprentissage de LE ne nécessitent pas forcément l'étiquetage des unités du corpus. Le métalangage traditionnel est un obstacle pour certains apprenants. Le paragraphe suivant détaillera les deux préoccupations évoquées ci-dessus, à savoir la production de nouvelles descriptions de la langue et l'utilisation de corpus lors d'une tâche d'apprentissage.

1.4. Objectifs d'utilisation d'un corpus

1.4.1. Objectifs "théoriques"

Pourquoi les linguistes utilisent-ils des corpus ? Teubert (2005) considère les données des corpus comme attestations ("*testymony*") de ce qui a été dit ou écrit. Il explique qu'elles relativisent – voire invalident – l'autorité de certaines définitions lexicales que l'on trouve dans des dictionnaires. Ces définitions ont été transmises et acceptées de façon parfois peu critique. Pourtant, l'évolution des langues appelle à une révision régulière des définitions lexicales. L'étude sur corpus donne en plus l'occasion de revoir les catégories grammaticales établies. Les résultats obtenus par l'exploration de grands corpus mettent en cause certaines des catégories, et ils peuvent permettre d'en formuler d'autres, plus adaptées à l'état actuel de la langue ou plus adaptées aux besoins de celui qui les utilise. Contrairement au courant du cognitivisme qui est basé sur des catégories grammaticales prédéfinies, la linguistique de

corpus a tendance à rester ouverte par rapport aux catégories. Travailler sur des corpus ou avec des corpus est donc potentiellement une activité de chercheur – et parfois de créateur – de références. Ceci semble s'appliquer aussi bien aux linguistes qu'aux enseignants et apprenants. Pour Rastier (2004), "la linguistique de corpus peut offrir les moyens théorique et technique d'étudier l'espace des normes". Il considère, comme Teubert (2005), l'espace des normes, qui est offert par les corpus, comme "chaînon manquant" entre langue et parole. Dans la lignée de Rastier, Mayaffre (2005) relie le thème de la conception et de l'utilisation des corpus aux débats de la linguistique contemporaine concernant la distinction entre théorie et empirie, démarche déductive et inductive, langue et parole. Le tableau ci-dessous illustre trois postures scientifiques que l'on peut adopter par rapport aux corpus. Les oppositions y sont présentées de façon plus stricte qu'elles ne le sont en réalité.

corpus sans théorie---------- corpus utilisés pour créer la théorie ---------- théorie sans corpus
parole---------- parole/langue ----------langue

Les trois postures se traduisent par l'utilisation de corpus de niveaux différents :

corpus lexicographiques ---------- corpus textuels ---------- corpus phrastiques
("sacs de mots") (textes entiers) (phrases formulées par des linguistes)

Les postures peuvent être rattachées à des théories et des méthodologies opposées.

Linguistique de corpus :	Linguistique sur corpus :
- linguistique descriptive ;	- linguistique théorique ;
- empirie (les données attestées) ;	- théorie (le système grammatical) ;
- induction ;	- déduction ;
- le corpus sert d'apport pour construire le savoir ;	- le corpus sert à valider et illustrer des hypothèses et théories qui existent déjà ;
- le corpus a une valeur heuristique.	- le corpus a une valeur documentaire.

Tableau 4 – Postures vis-à-vis des corpus et leurs implications théoriques et méthodologiques.

Maingueneau (2008 : 10) oppose un "schéma général de l'énonciation (…), invariant à travers la multiplicité des actes d'énonciation", aux énonciations individuelles qui ont lieu un moment donné, à un endroit donné. Grâce à la juxtaposition d'actes d'énonciation individuels multiples, les corpus peuvent alors être considérés comme expression de ce schéma général. Ceci n'est pas sans rappeler une des "clés pour la communication" proposées par Demaizière et Berthoud (1998) : "Les phénomènes locaux sont pertinents pour la compréhension des phénomènes globaux."

1.4.2. Objectifs "pratiques" ou la "théorie appliquée"

Habert, Nazarenko et Salem (1997 : 28) plaident en faveur d'une constitution de corpus et d'une annotation de ses données motivées par une application précise : "Un étiquetage est toujours orienté par une tâche." (Habert, Nazarenko et Salem, 1997 : 37). Quels peuvent être les objectifs "pratiques" qui motivent la constitution d'un corpus ?

Le *Trésor de la langue française informatisé* mentionne comme objectifs l'étude linguistique de textes et la recherche automatique d'informations.

> (LING.) Ensemble de textes établi selon un principe de documentation exhaustive, un critère thématique ou exemplaire en vue de leur étude linguistique.
>
> (ÉLECTRON., INFORM.) Ensemble de données exploitables dans une expérience d'analyse ou de recherche automatique d'informations.

Bernardini (2004 : 15) évoque deux possibilités d'utilisation des corpus. Premièrement, on peut se servir des descriptions de faits langagiers qu'offre un corpus pour comprendre des phénomènes concernant le fonctionnement linguistique. Deuxièmement, l'enseignant peut utiliser les corpus pour "influencer" les processus d'enseignement/apprentissage. Bernardini s'intéresse particulièrement à la deuxième utilisation. Elle propose de relever le défi de faire entrer les corpus dans les salles de classe et de les faire utiliser lors d'activités communicatives de type "résolution de problème" (*communicative reasoning-gap activities*). Les corpus peuvent, à son avis, favoriser l'apprentissage incident et permettre aux apprenants de créer leurs propres références linguistiques (cf. 4.2.3.).

L'utilisation de corpus peut alors permettre de poursuivre les objectifs suivants :

- faire l'inventaire diachronique et/ou synchronique d'une langue (fonction documentaire) ;
- faire une analyse linguistique de textes ;
- étudier l'espace des normes ;
- rechercher des informations, par exemple dans des situations de type "résolution de problème" ;
- apprendre une langue étrangère de façon incidente.

Un des objectifs principaux est l'étude de la langue et sa description qui résulte d'un travail qui n'est pas basé sur l'introspection, mais sur des énoncés attestés. Cette observation peut, entre autres, mener à la production de grammaires et de dictionnaires. Les objectifs d'analyse linguistique et de recherche d'informations peuvent intéresser aussi bien les linguistes que les apprenants d'une langue étrangère. En raison de ma perspective de didacticienne, le lien entre corpus et tâches m'intéresse particulièrement. En effet, les corpus permettent de proposer des

"*open-end retrieval tasks*" (Sinclair, 1996), donc des tâches d'exploration ou de recherche d'information. Le dernier objectif, étudier l'espace des normes, est un objectif théorique susceptible d'intéresser les chercheurs en linguistique et philologie davantage que les didacticiens. Néanmoins, je pense qu'il concerne aussi ces derniers, car il permet de repenser le statut de notions comme "règle" et "erreur" : leur pertinence peut être interrogée grâce à l'observation d'énoncés attestés. Les représentations des apprenants, mais aussi des enseignants, concernant la langue étrangère peuvent ainsi être mises à l'épreuve et, le cas échéant, être modifiées.

1.5. Outils d'exploration de textes

La numérisation de corpus et la mise à disposition de systèmes de recherche électronique semble avoir, entre autres, motivé l'emploi croissant des corpus qui, avant l'existence de tels outils, ont été utilisés de façon beaucoup plus restreinte (Leech, 2003 ; Duffner et Näf, 2006a). Dans ce qui suit, je vais décrire plusieurs outils d'exploration de textes pour ensuite préciser les spécificités d'un outil d'exploration de corpus, le concordancier.

1.5.1. Moteurs de recherche

Un moteur de recherche permet de rassembler des informations concernant des mots clés. Il sélectionne et classe ensuite les pages Internet contenant les mots clés selon un procédé qui n'est pas toujours transparent (Sinclair, 2005). Ce manque de transparence s'explique par des raisons commerciales et aussi par souci de garder le contrôle sur l'ordre selon lequel les pages Internet s'affichent lors d'une requête (Bubenhofer, 2006). *Google*, *Exalead*, *Ask* et *Yahoo* proposent de tels moteurs de recherche.

1.5.2. Fonction "rechercher dans ce document/cette page…"

Les traitements de texte permettent de retrouver un terme dans un "document", c'est-à-dire dans un fichier donné. Par exemple, le traitement de texte *Word 2003* a une fonction nommée "Rechercher" qui propose d'explorer le document principal, l'en-tête, les notes de bas de page, les commentaires, etc. Sur la Toile, une fonction équivalente existe. Grâce à la fonction "Rechercher dans cette page…", il est possible d'interroger des "pages", donc les informations affichées à l'écran, ceci à l'aide du moteur de recherche que l'on a sélectionné.

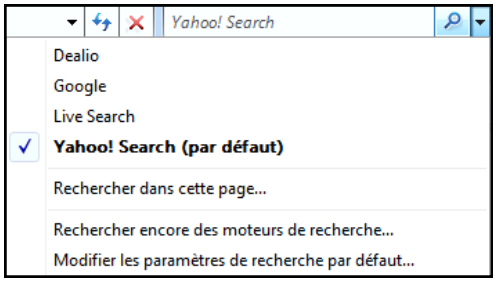


Figure 5 – Fonction en ligne "rechercher dans cette page…".

1.5.3. Requêtes par formulaire

Les banques de données, archives et corpus peuvent être pourvus de systèmes de recherche des informations qu'ils contiennent. Certaines archives de la presse en ligne par exemple proposent uniquement l'option "rechercher un mot clé sur le site". Parfois, le système de recherche est plus élaboré. Voici les options de recherche pour les archives du journal *Le Monde*.

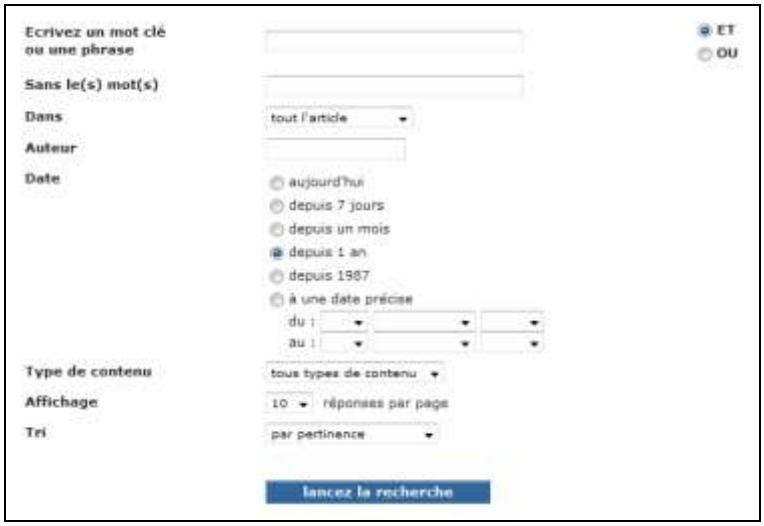


Figure 6 – Système de recherche d'informations sur le site du journal *Le Monde*.

Dans ce formulaire, il est possible de taper un mot clé ou des mots clés, d'exclure certains mots clés, mais aussi de cocher des options prévues par les concepteurs du site ou de préciser le laps de temps concerné par la recherche.

1.6. Outils d'exploration de corpus

Des outils simples ou des collections d'outils peuvent soutenir l'exploration d'un corpus. Un des outils les plus utilisés est le concordancier. Il peut être associé à des outils de calcul et de présentation graphique des mots clés.

1.6.1. Calculateur de fréquence

Le calculateur de fréquence permet de classer les mots selon leur ordre de fréquence. Certains logiciels, comme *AntConc3.2.1w*, donnent l'occasion de comparer la fréquence des mots dans plusieurs corpus différents. La fonction "*keyword*" permet d'établir une liste alphabétique pour chaque corpus et une deuxième liste qui classe les mots par fréquence relative par rapport aux autres mots dans le corpus.

1.6.2. Outil de dispersion

L'outil de dispersion, une des fonctions proposées par *AntConc3.2.1w*, permet de localiser certains mots clés ou segments graphiquement, soit dans le texte qui s'affiche dans son intégralité à l'écran, soit à l'aide d'une forme géométrique qui représente la totalité des textes du corpus en une suite continue. La figure ci-dessous représente la localisation de l'auxiliaire "*werden*" [devenir] dans les deux corpus "Livre d'or" (*Gästebücher*) et "Visite guidée" (*Korpus_Führungen*).

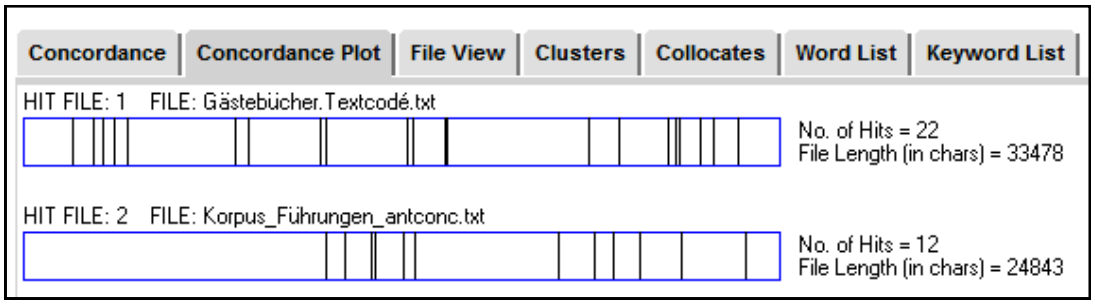


Figure 7 – Localisation du mot clé "*werden*" dans deux corpus de référence.

Les deux corpus sont de taille différente (33 478 caractères versus 24 843 caractères). Pour une étude comparative, il convient d'utiliser des échantillons de même taille.

1.6.3. Concordances et concordanciers

1.6.3.1. Concordances

Les concordances sont le résultat d'une mise en forme matérielle particulière des données d'un corpus. Avant l'existence des TIC (Technologies de l'Information et de la Communication), les concordances étaient fabriquées manuellement afin de soutenir l'étude de textes religieux et littéraires. Le *TLFi* (*Trésor de la Langue Française*) nous apprend qu'en 1564, on définit la concordance comme

> table alphabétique des mots employés dans la Bible, avec indication des textes qui les contiennent.

En 1680, une application pour l'enseignement des langues est évoquée.

> Concordance. Petit rûdiment avec un sintaxe [...] pour instruire les enfans qui commençent à aprendre le latin.

Lamy et Klarskov Mortensen (2006) évoquent la nature des données d'une concordance. Elles proviennent d'ouvrages "authentiques", c'est-à-dire que leur emploi est attesté.

> *So a concordance is a list of words (called keywords, e.g. here "sin"), taken from a piece of authentic language (corpus, e.g. here Romeo and Juliet), displayed in the centre of the page and shown with parts of the contexts in which they occur (here maximum 29 characters to the left of the keyword and to the right). This is also known as a Key Word In Context or KWIC concordance* (Lamy et Klarskov Mortensen, 2006).

La définition de Stevens (1988) précise comment on obtient une concordance à partir d'un corpus et comment les mots clés peuvent être matériellement mis en forme.

> *A concordance is a listing of words and their contexts. A computer-generated concordance typically searches through a body of text for every occurrence of a given word or string and prints the line of text surrounding that string, resulting in a listing of several lines of text. The search-string appears in a column down the middle, with the 25 or 30 character environment either side for each occurrence of the string in the text. The search-string could be several words, or just the first part of a word, so that morphological variants can appear in the same concordance; e.g. comprehen- would concordance comprehend, comprehending, comprehension* (Stevens, 1988).

La définition proposée par Pincemin (2006) nous renseigne sur certaines caractéristiques techniques des concordances.

> Un corpus étant fixé, une concordance (…) est la liste de toutes les occurrences d'un pivot, alignées verticalement en colonne, entourées de part et d'autre par leur contexte, accompagnées d'une référence indiquant de façon pertinente leur positionnement dans le corpus, et triées selon un critère pertinent pour l'analyse.

Je retiens des définitions précédentes les éléments suivants.

Une concordance est une mise en forme particulière des données d'un corpus. Le texte n'y est pas présenté en continu, mais il est matériellement réorganisé suite à une recherche de toutes les occurrences d'un mot clé ou d'un segment dans le corpus. Les occurrences sont regroupées sous forme de liste. Les éléments de cette liste apparaissent sous forme de trois colonnes, à savoir le contexte de gauche du mot clé, le mot clé et le contexte de droite du mot clé. La taille des éléments d'une telle liste dépend des options de recherche définies préalablement. Les concordances aident à observer les mots clés dans leur contexte immédiat. Les applications possibles sont, entre autres, l'étude de texte (linguistique, discursive, etc.) et l'enseignement/apprentissage des langues.

Figure 8 – Ma définition de "concordance".

1.6.3.2. Concordanciers

La mise en forme des données d'un corpus est obtenue à l'aide d'un logiciel, le concordancier. Ce logiciel est un outil d'exploration de textes. Il permet de rechercher et de regrouper dans un corpus toutes les occurrences d'un mot clé donné. Les résultats sont affichés sous forme de lignes contenant toutes le mot clé, qui, lui, est mis en forme d'une façon qui le rend particulièrement visible. D'autres options d'affichage existent, comme par exemple des parties de textes dépassant une ligne, ou bien les textes entiers contenant le mot clé. Ces deux types d'affichage sont privilégiés dans le contexte de corpus parallèles (St John, 2001), donc de corpus permettant de comparer deux ou plusieurs langues entre elles. Dans le contexte de corpus monolingues, l'affichage de type ligne reste la norme, mais selon les types de concordanciers, il est possible d'afficher un contexte plus large pour les lignes auxquelles on s'intéresse particulièrement.

Mit welchen Waffen kämpfte **MAN** ? Am Schluss der Führung könnt ihr es selbst
bauen wir ein Gerät mit dessen Hilfe **MAN** Bilder, Fotos oder Zeichnungen
In der Renaissance sammelte **MAN** in Wunderkammern Kunstgegenstände
aus nur zwei bis drei beweglichen Teilen. **MAN** kann damit ganz einfach
Fußbodenbelag bei dem **MAN** Muster aus gleichartigen Holzplatten bildet.
Auf den Bildern für Kaiserin Maria-Theresia kann **MAN** sehen, wie Wien und das
Doch was sind Radiowellen Wie empfängt **MAN** sie ? Was sagen sie uns?

Figure 9 – Mot clé "*man*", recherché dans le corpus "Visite guidée".

La figure ci-dessus montre une concordance basée sur un lexème simple. Les lignes ont été "nettoyées" par moi. Cela veut dire que j'ai enlevé des bouts de texte aux extrémités gauche et droite des lignes et que j'ai rajouté des virgules manquantes afin de rendre la lecture du texte plus aisée pour les apprenants.

besten Freundin C-Ko (Bild , blond) auf die	**Schule geht**	. A-Ko zeichnet sich als Tochter von
Die Rüpelschule von Michael Ende In dieser	**Schule geht**	es so ganz und gar anders zu als an allen
das mit einem Lächeln auf den Lippen zur	**Schule geht**	, ist ein Heuchler . " Wer von uns fragt
Nicht mich , eine Schülerin , die noch zur	**Schule geht**	, noch nicht einmal Abitur hat und erst
Südstadt-Bult gemeldet ist oder hier zur	**Schule geht**	. Insgesamt sollen 15 Jugendparlamentarier
adoptiert " und bezahlt dafür , daß es in die	**Schule geht**	. Nach meiner Erfahrung sind diese beiden
Unterordnungsverhältnisse . Wer in eine solche	**Schule geht**	, will zugleich das untergehende Handwerk
Serie heißt " Summer Of 69 " . Wenn man zur	**Schule geht**	, kann man's nicht abwarten , bis die Penne
Tagesplan voraus . Wer ohne Frühstück in die	**Schule geht**	, ist nachher der Dumme : Denn wie

Figure 10 – Segment "*Schule geht*", recherché dans le corpus *DeWac* (Baroni, 2006).

La figure 10 présente la concordance d'un segment, donc d'un groupe de deux lexèmes. Les lignes n'ont pas été retouchées par moi.

Certains outils d'exploration de corpus, comme par exemple un des outils proposés par *AntConc3.2.1w*, permettent une mise en forme matérielle des mots clés appelée *cluster*. Il s'agit d'une forme particulière de concordance. Les *cluster*s sont des combinaisons de mots qui apparaissent ensemble selon un ordre de fréquence significatif. Partington (2006) définit *cluster* comme "*extended collocation*", donc comme une collocation élargie.

Eintritt für
Eintritt für Kinder
Eintritt für Kinder frei
Eintritt für Kinder frei, ermäßigter
Eintritt für Kinder frei, ermäßigter Eintrittspreis
Eintritt für Kinder frei, ermäßigter Eintrittspreis für

Figure 11 – *Cluster* pour "*Eintritt* [entrée]" dans le corpus "Visites guidées" (*AntConc3.2.1w*).

Dans le contexte de l'utilisation de concordances pour l'enseignement/apprentissage des langues étrangères, O'Keeffe *et al.* (2007 : 13) utilisent le terme de *cluster* comme synonyme de *chunk*s. Ils conseillent de s'intéresser à des *cluster*s qui se composent de deux à six mots.

1.6.4. Concordanciers pour l'allemand

Un concordancier présente les mots clés dans leur contexte immédiat plus ou moins large. Selon les possibilités que propose le logiciel, l'utilisateur peut déterminer ou pas la largeur du contexte, et il a ou n'a pas directement accès au texte source. Le tableau ci-dessous présente certaines fonctionnalités de six concordanciers gratuits qui peuvent convenir à un travail sur l'allemand.

Concordancier	**Accès au texte source**	**Déterminer la largeur de contexte du mot clé**	**Importer un corpus de texte**	**Possibilité de comparer plusieurs corpus**	**Créer des listes de fréquence**
AntConc3.2.1w	x	x	x	x	x
Cosmas II (permet d'interroger le corpus *DeReKo*.	x On a accès à un contexte allant d'une à plusieurs phrases ("*Volltext*").	La largeur du contexte est prédéterminée.	On peut regrouper des corpus qui existent dans la base.		
Konkordanzer-Corpus Deutsch v4.1)	x	x			
Lexico3	x On peut afficher le texte source dans une nouvelle fenêtre.	x On peut élargir progressivement le contexte, sans limite.	x	x	x
Sketch Engine Permet d'interroger le corpus *DeWac* (Baroni, 2006).	x Il y a trois niveaux d'affichage du contexte : une ligne, un passage de texte extensible à gauche et à droite, et la page Internet où se trouve le texte source complet.	x	x L'importation d'un corpus est possible, mais cela requiert des connaissances techniques plus importantes que pour *AntConc3.2.1w*, *Lexico3* et *Lextutor*.		x
Text-based concordances (v2)			x		x

Tableau 5 – Comparaison de six concordanciers.

Le logiciel *Text-based concordances* (v2) (Cobb, nd) proposé sur le site *Compleat Lexical Tutor* a été conçu pour l'anglais. Les imperfections qui résultent de son "détournement" pour l'allemand (gestion imparfaite de certaines lettres) ont toutefois été compensées par la facilité d'utilisation du logiciel qui convient à des utilisateurs relativement peu formés en informatique. Pour des utilisateurs plus avertis, *AntConc3.2.1w* (Anthony, 2007) et *Lexico3* (Fleury, Lamalle, W. Martinez et Salem, 2003) sont de bonnes alternatives à *Text-based concordances* (v2) de Cobb (nd). *AntConc3.2.1w* est la version la plus récente de la collection d'outils d'interrogation de corpus *AntConc1.0* (2002). Cette collection a initialement été conçue pour l'enseignement/apprentissage de la rédaction technique (Anthony, 2006), dans l'esprit des principes défendus par Johns (cf. 4.2.1.). Sa prise en main est relativement simple. Il est possible d'interroger ses propres corpus, enregistrés sous forme de texte brut ou sous format "RTF" (*Rich Text Format*). Les fonctionnalités sont variées : création de concordances, de listes de mots (mots classés selon leur fréquence), localisation visuelle d'un terme au sein du corpus ("*concordance plot*"), comparaison de la fréquence de termes dans plusieurs corpus, etc. Pour un aperçu plus complet des fonctionnalités de *Compleat lexical tutor 4.0* – dont *Text-based concordances* (v2) – et d'*AntConc3.2.1w*, il est utile de consulter Diniz (2005). La collection d'outils de statistique textuelle *Lexico3* (Fleury *et al.*) est disponible depuis 2001. Elle est la suite la plus récente de *Lexico1*, créé par Salem (1990). Les outils ont été conçus pour des spécialistes de la langue. *Lexico3* permet, entre autres, de faire segmenter du texte brut et de le présenter sous forme de concordances. Le contexte peut être restreint ou élargi de façon illimitée. Il est possible de créer des statistiques qui permettent de visualiser la fréquence de plusieurs termes dans un corpus.

Konkordanzer-Corpus Deutsch (v5.1) est un concordancier facile d'emploi qui permet d'interroger le *Braun Corpus* (Divsic *et al.*, 2008 ; cf. 1.7.2.). Ce petit corpus non annoté d'environ un million de mots allemands propose plusieurs options d'interrogation, entre autres l'option de cocher les lignes de concordance que l'on souhaite sauvegarder et l'option "*gapped*" [omis] qui offre la possibilité de rechercher des lignes de concordance qui contiennent un même mot clé, mais ce mot clé n'est pas affiché.

User extract for **MUSEUM**

```
007.  s Klassen das Museum besuchen, wird der Preis aus _____spädagogischen und bildungspolitischen Grü
008.   "mehr Museum" Mit der Eröffnung des neuen Bilder_____s im Dezember sollen neue Eintrittspreise ge
010.  ersteinerten Wald. Schon seit 1979 beherbergt das _____ ein umfangreiches Insektarium, in dessen za
019.  pen ist, hat sich die Beethovenstadt Bonn mit der _____smeile in die Herzen kultur-begeisterter Bes
021.   beginnt in Paris: im Louvre liegt die Leiche des _____sdirektors Saunière. Kurz vor seinem Tod ha
022.   die "Theodor Heuss" hat ihren Platz im Deutschen _____ in München, die "Hamburg" dagegen ist wied
032.  bewundern. Gleich nebenan befinden sich das Heimat_____, eine Galerie mit Gemälden des Pirmasenser
052.  eipzig 02.09.2004 Neue Eintrittspreise für "mehr _____" Mit der Eröffnung des neuen Bildermuseums
053.  s im Zentrum der Stadt. Für dieses meistbesuchte _____ in Norddeutschland sind allerdings mehrere
057.  en ihrer Schülerinnen, die sich in einem Plauner _____ befinden, konnte diese Technik wiederentdec
```

Figure 12 – Lignes de concordance sans le mot clé "*Museum*".

Les plus grands corpus de langue allemande sont *DeWac* (Baroni, 2006), un des corpus que l'on peut explorer à l'aide de *Sketch Engine* (Kilgarriff *et al.* , nd) et *DeReKo* (*Deutsches ReferenzKorpus*) de l'*IDS* (*Institut für Deutsche Sprache*, Mannheim), interrogeable à l'aide du concordancier *Cosmas II* (nd). Pour les non-experts en informatique, *Sketch Engine* est un outil suffisamment facile d'emploi, ce qui est moins le cas pour *Cosmas II*. Les commandes booléennes présentées dans le manuel de l'utilisateur de *Sketch Engine* ne semblent toutefois pas toutes être implémentées pour *DeWac* (cf. 1.7.2.2.3.).

L'outil *Linguee* (Fink et Frahling, 2008), accessible au public depuis avril 2009, aligne des textes allemand et anglais qui ont été publiés sur la Toile. Il permet de chercher un lexème ou un segment dont les unités se suivent. Les autres formes du mot que l'on indique sont également prises en compte. On a accès aux sources. Les internautes peuvent contribuer à vérifier les traductions proposées par l'outil. L'évolution de *Linguee* vers une mise en parallèle de textes allemands et français serait une perspective intéressante pour mes recherches.

1.7. Corpus en langue allemande

1.7.1. La place des corpus allemands dans le paysage des corpus

L'évolution quantitative et qualitative des corpus est particulièrement visible dans le domaine de la langue anglaise, comme l'illustrent le site Internet *Bookmarks for Corpus Linguists* de Lee (nd) et le site de Braun (nd) concernant l'utilisation de corpus pour apprendre des langues. Nous disposons actuellement d'une multitude de corpus anglais qui représentent diverses variantes de cette langue et couvrant un large spectre de types de corpus. Certains sont très grands, comme par exemple *ukWac* qui contient au-delà de 2 milliards de mots (en 2008). D'autres sont exemplaires en ce qui concerne leur organisation et leur facilité d'emploi, comme par exemple le *BNC* (*British National Corpus*) qui contient 100 millions de mots (en 2008). Il est annoté, et il classe les données dans des catégories qui représentent des types de

textes variés. On peut l'interroger à l'aide de *Sara*, un outil de recherche qui ne propose pas seulement la recherche par mot clé simple ou par groupes de mots, mais aussi la recherche concernant les annotations non linguistiques (genre de texte, informations concernant les locuteurs, etc.). *Sketch Engine* (Kilgarriff *et al.*, nd) donne accès à un nombre croissant de corpus. Certains des corpus représentant d'autres langues que l'anglais y sont annotés de façon "fruste" (cf. 1.7.2.2.3.). Le corpus allemand *DeWac* (Baroni, 2006) est de très grande taille (1,644 milliards de mots en juillet 2009).

Language	Name	Tokens [?]	
Chinese	Chinese GW, simpl	706 427 624	info
Chinese	Chinese GW, trd	706 428 333	info
English	British Academic Written English Corpus (BAWE)	8 336 262	info
English	British National Corpus	111 244 375	info
English	ukWaC	1 526 599 198	info
French	French web corpus	126 850 281	info
German	deWaC	1 644 785 836	info
Italian	itWaC	1 909 535 984	info
Japanese	JpWaC	409 384 405	info
Russian	Russian Web Corpus	187 965 822	info
Slovenian	Fida PLUS 620m	738 460 168	info
Spanish	Spanish web corpus	116 900 060	info

Tableau 6 – Corpus proposés par *Sketch Engine*.

Un autre méga-corpus allemand est proposé par l'*IDS* (*Institut für Deutsche Sprache*). Il a été conçu de façon exemplaire (Duffner et Näf, 2006b). Le nombre de corpus spécialisés allemands est en expansion (Lemnitzer et Zinsmeister (2006 : 107-125).

1.7.2. Corpus en langue allemande

Les corpus allemands ont une histoire plus jeune que les corpus anglais. Les outils de création et d'interrogation des corpus, conçus dans un premier temps pour l'anglais, ne sont pas tous utilisables tels quels pour l'allemand. Des outils spécifiques à l'allemand ont été élaborés par exemple par l'*IDS* (*Institut für Deutsche Sprache*) de Mannheim ou par l'*IMS* (*Institut für Maschinelle Sprachverarbeitung*) de Stuttgart, mais ces outils ne conviennent pas forcément dans un contexte d'enseignement/apprentissage d'une langue étrangère, et encore moins au collège. Dans ce qui suit, je propose un aperçu général de corpus allemands que j'ai répertoriés en 2008 et 2009. Ne figurent pas dans le tableau ci-dessous les corpus que je considère comme marginaux en raison de leur petite taille ou de leur spécificité trop grande, ni les corpus oraux. De tels corpus sont répertoriés sur le site *Bookmarks for Corpus-based*

Linguists de Lee et chez Lemnitzer et Zinsmeister (2006 : 107-125). Je termine l'aperçu par la présentation de deux portails terminologiques, basés sur des corpus. À l'intérieur d'une rubrique, les corpus sont présentés du plus petit corpus au plus grand. Le tableau est suivi d'une description plus détaillée d'un portail terminologique et la description de deux corpus qui sont actuellement considérés comme les plus grands en langue allemande. Les informations que je présente ci-dessous ont été trouvées chez Duffner et Näf (2006b), sur le site *Schlobi's Linguistic Corner* de Schlobinski (nd), le site de l'*Institut für Kommunikationswissenschaften* [Institut d'Information et Communication] de Bonn, le site *Bookmarks for Corpus-based Linguists* de Lee (nd), déjà évoqué ci-dessus, et le site *Compleat lexical tutor* de Cobb (nd).

1.7.2.1. Corpus parallèles

Chemnitz German-English translation Corpus	Ce corpus parallèle contient environ un million de mots anglais et un million de mots allemands. Quatre domaines sont représentés : discours et articles publics, anglais et allemand académiques, le tourisme et documents de l'union européenne. On peut faire des requêtes par mot clé simple ou par mots clés combinés en sélectionnant un ou plusieurs des quatre corpus.
OPUS	Un regroupement de corpus parallèles en vingt-quatre langues, libres d'accès, géré par Tiedemann (2009). Les données ont été trouvées sur Internet. Elles englobent plusieurs langues de spécialité : la presse, le sous-titrage de films, les manuels en ligne, etc. Tout n'a pas encore été traduit.

1.7.2.2. Corpus unilingues

1.7.2.2.1. Corpus d'apprenants

FALKO	Ce corpus est mis à disposition en ligne sur le site de l'*Institut für Deutsche Sprache und Linguistik* [institut de philologie et linguistique allemande] de l'université Humboldt à Berlin. Il contient des sous-corpus rassemblant des résumés de textes, des essais et des données récoltées lors d'un projet longitudinal. Il existe aussi des sous-corpus contenant des résumés et essais de locuteurs natifs. Les données du corpus sont

annotées selon des hypothèses formulées au sujet de types d'erreur recensés dans les textes des apprenants. La consultation du corpus est gratuite.

1.7.2.2.2. Portails terminologiques, basés sur des corpus

Das Digitale Wörterbuch der Deutschen Sprache (DWDS)	Ce portail dont le nom peut être traduit par "dictionnaire digital de la langue allemande" est basé sur deux corpus oraux et des corpus écrits. Le site est conçu dans un souci d'équilibre entre plusieurs domaines. La consultation est gratuite.
Wortschatz-Portal	Ce portail dont le nom peut être traduit par "portail terminologique" est spécialisé dans le recensement de la langue allemande très récente. Les corpus sur lequel s'appuie ce portail sont créés à partir de données trouvées sur Internet (par *web-crawling*), sans classement particulier. La consultation est gratuite.

Le portail DWDS contient des corpus qui représentent des genres de textes variés. Ses données reflètent l'allemand de façon diachronique. Nous y trouvons deux corpus de référence, quatre corpus du domaine de la presse, deux corpus dans lesquels on a transcrit la langue orale et deux corpus spécialisés (revues juives et langue de la République Démocratique d'Allemagne). Il est possible d'interroger ces corpus séparément. Les deux corpus de référence cherchent à respecter la répartition suivante : belles lettres (26%) ; presse (27%), langue de spécialité (22%) ; textes utilitaires (ca. 20%) et transcriptions de la langue orale (5%). Le *Julliand-"D"-Corpus* est composé de la même façon, mais il ne contient que des textes qui ont été écrits entre 1920 et 1939.

1.7.2.2.3. Autres corpus unilingues

Dortmunder Chat-Korpus	Un corpus de textes de type clavardage, contenant 0,6 millions de mots. Ce corpus a été créé sous la direction d'Angelika Storrer et de Michal Beißwenger de l'université de Dortmund. Certaines parties du corpus sont gratuitement disponibles en ligne.
NEGRA Korpus Version 2	NEGRA Korpus Version 2 est un corpus de l'université de Saarbrücken. Il comporte 354 096 mots (20 602 phrases) qui proviennent du quotidien

Frankfurter Rundschau, rassemblés dans le cédérom *Multilingual Corpus 1 der European Corpus Initiative*. Les phrases sont indexées par l'*IMS* (*Institut für maschinelle Sprachverarbeitung*) de Stuttgart. Il est possible d'utiliser le corpus à des fins non commerciales. L'utilisation de ce corpus demande un temps important de prise en main de l'outil.

TigerCorpus — 900 000 mots. Allemand de la presse. *Frankfurter Rundschau*. Projet de l'*Institut für Phonetik/Phonologie*, Universität des Saarlandes, de l'*Institut für Germanistik*, Potsdam et de l'*Institut für Maschinelle Sprachverarbeitung*, Stuttgart, terminé en 2007.

Braun Corpus — Le *Braun Corpus* a été mis en place en 2006/2007 par Divsic, ceci en collaboration avec Rott et Horst (*Concordia University de Montréal*). Il suit le modèle du *Brown Corpus*, et il contient environ un million de mots non annotés. Plusieurs types de textes y sont représentés.[1] Ce corpus est consultable en ligne à l'aide du concordancier *Konkordanzer Corpus Deutsch*, proposé par Cobb sur le site *Compleat lexical tutor*.

LIMAS — *LIMAS* (*Linguistik und Maschinelle Sprachbearbeitung*) est un corpus proposé par l'université de Bonn. Il contient environ un million de mots. Ses cinq cent documents écrits représentent plusieurs genres (textes pleins ou extraits). Le corpus est clos depuis 1970. La consultation est gratuite. Il fait partie des corpus que l'on peut interroger à l'aide de *Cosmas II* (*IDS*, Mannheim).

DeWac — Le corpus *DeWac* (Baroni, 2006) a été développé pour *Sketch Engine* dans un souci d'équilibre concernant les types de textes représentés. Il contient environ 1,7 milliards de mots allemands qui ont été collectés sur Internet et annotés à l'aide de l'outil *TreeTagger*. L'accès au corpus est payant, mais il y a la possibilité d'essayer gratuitement le corpus pendant trente jours. Il est également possible d'interroger *DeWac* à partir du site *waCky* (*Web-as-Corpus kool ynitiative*).

[1] Informations obtenues grâce à un échange de courriel avec Marlise Horst.

DeReKo — Ce corpus de l'*IDS* (*Institut für Deutsche Sprache*) de Mannheim représente l'allemand parlé et écrit (synchronie, diachronie des dernières décennies). Il regroupe des textes du domaine des belles lettres, des articles de journaux, et il contient 65 corpus (3,6 milliards de mots en février 2009). En raison du *copyright*, certaines données ne sont pas accessibles au public. 2,2 milliards de mots sont gratuitement interrogeables en ligne à l'aide de *Cosmas II*. Certaines parties du corpus peuvent être achetées.

Le groupement de corpus *DeReKo* est le plus grand ensemble attestant l'allemand actuellement disponible. Il contient au-delà de 3,6 milliards de mots (en février 2009).

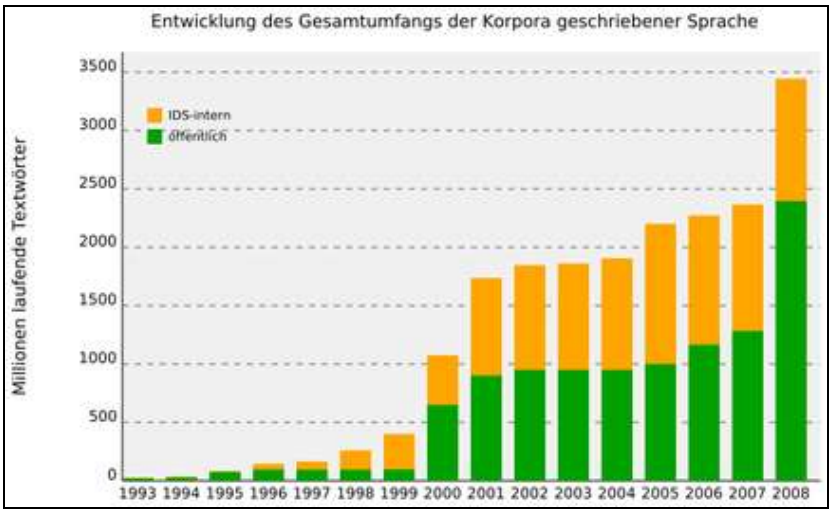


Figure 13 – L'évolution de la taille des corpus de langue allemande écrite dans *DeReKo*.

Une grande partie des données du *DeReKo* peut être consultée à l'aide de l'outil de recherche *COSMAS II* (*Corpus Search, Management and Analysis System*). Il est possible d'interroger certains corpus séparément ou de créer des sous-corpus. L'utilisation de *DeReKo* peut paraître difficile à des non-spécialistes en linguistique de corpus. Toutefois, les possibilités d'exploration et d'exportation des données qu'offre *Cosmas II*, le système d'exploration de *DeReKo*, sont actuellement inégalées par les autres corpus allemands. Le corpus *DeWac* a été développé par Baroni pour *Sketch Engine* (Baroni et Kilgarriff, 2006). Ce corpus peut être assez facilement exploré par des non-spécialistes en matière de linguistique de corpus. Toutefois, quand j'ai essayé d'utiliser des commandes booléennes pour l'exploration du corpus allemand *DeWac*, je n'ai pas obtenu de résultats. Pourtant, je n'ai pas eu de difficultés avec les

commandes booléennes des corpus anglais et français proposés par *Sketch Engine* (Kilgarriff *et al.*, nd). J'ai signalé ces difficultés aux gérants de *Sketch Engine* qui m'ont répondu qu'ils cherchent à rendre l'outil plus facile à utiliser. En effet, le corpus allemand *DeWac* ne propose pas autant d'options que les corpus faits pour d'autres langues (le chinois, le japonais, l'anglais, l'italien, le français, l'espagnol, le portugais, le grec et le slovène). Il manque les options *Thesaurus* et *Sketch Difference*. Nous pouvons constater les mêmes limitations pour le corpus russe et perse qui sont aussi présents sur le portail de *Sketch Engine*. La capture d'écran suivante présente les options d'interrogation qui s'affichent pour le corpus français *French Web Corpus* (Sharoff, nd).

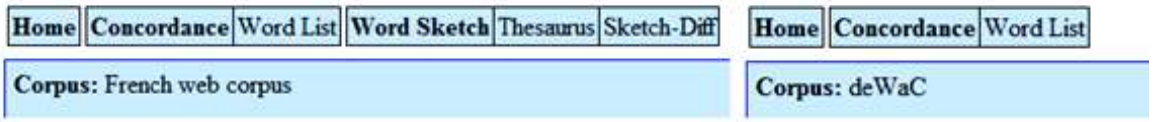


Figure 14 - Options d'exploration pour *French Web Corpus* et pour *DeWac*.

1.8. Pour résumer

Mon étude de l'offre de corpus allemands en 2009 montre qu'il y a peu de corpus dont le contenu semble pouvoir soutenir la production de tâches écrites en allemand pour des apprenants ayant un niveau intermédiaire. Dans le chapitre 4.1.5., je présenterai des corpus créés pour des besoins d'enseignement/apprentissage, et j'étudierai les possibilités d'emploi dans le cadre d'un apprentissage de LE en institution scolaire (4.2). Avant de faire cela, je propose d'aller à la rencontre de divers points de vue concernant l'apprentissage d'une langue étrangère en général (chapitre 2), pour ensuite aborder des questions liées à l'apprentissage de la production écrite (chapitre 3).

2. Apprentissage d'une langue étrangère

Comme le précise Portine (2008), en didactique des langues, une théorie est considérée comme "un ensemble cohérent de réponses visant la complétude". Si l'on se positionne comme didacticien ou didacticienne, il convient donc de se référer à plusieurs théories d'enseignement/apprentissage des langues et à des approches variées qui soient en même temps compatibles entre elles. Portine (2008) propose de convoquer des composantes des paradigmes cognitif, énonciatif et brunéro-vygotskien. Il les considère comme le reflet d'une problématique générale, celle de l'opposition entre la GU (Grammaire Universelle) et l'hypothèse revisitée de Sapir-Whorf. Les adeptes de la GU considèrent les langues comme paramètres d'un même substrat, localisé dans le cerveau sous forme de module réservé à l'activité langagière. Sapir-Whorf crée un lien entre les langues et les représentations du monde. Ces représentations sont façonnées par l'environnement historico-culturel. Selon Sapir-Whorf, les humains perçoivent la langue à travers un "filtre", ou, comme dirait Bruner (1958, cité dans Gaonac'h, 1991 : 90), à travers une "grille". Ceci a comme conséquence qu'ils en remarquent certaines propriétés et qu'ils en ignorent d'autres. Portine (2008) résume les clivages entre la GU et la théorie de Sapir-Whorf en une opposition entre les approches de la langue de type "tout-linguistique" et les approches de type "tout-culturel".

Dans ce qui suit, j'évoquerai quelques apports des théories d'apprentissage des langues qui sont proches du "bord" linguistique, à savoir les recherches en linguistique cognitive et en psycholinguistique. Je définirai les termes, *input*, *intake* et *output*, et je décrirai le rôle qu'on leur accorde pour l'apprentissage d'une langue étrangère. Je donnerai ensuite un aperçu des recherches dans le domaine des collocations, domaine particulièrement étudié en linguistique de corpus (Granger et Meunier, 2008 : 257-258). Au sein d'un chapitre dédié au constructivisme, j'aborderai les attitudes des apprenants par rapport à l'*input*. Le sous-chapitre suivant s'intéressera au guidage offert par l'enseignant dans une perspective constructiviste. Pour finir, je présenterai des méthodologies proches du "bord" culturel, à savoir la pédagogie du projet et l'apprentissage par la tâche, et j'esquisserai la méthodologie de la conceptualisation, une méthodologie susceptible de proposer un compromis prometteur entre "le tout-linguistique" et le "tout-culturel".

2.1. Linguistique cognitive et apprentissage de LE

Robert (2002) explique qu'il y a une répartition "classique" de deux domaines d'étude de la langue. La psycholinguistique s'intéresse aux processus de production et de compréhension, tandis que la linguistique se concentre de façon privilégiée sur l'étude des "structures du langage", pour reprendre le terme utilisé dans ce domaine. Selon Robert (2002), pendant longtemps, les linguistes à tendance fonctionnaliste ou dynamiciste se sont peu intéressés aux processus mentaux qui sous-tendent le fonctionnement du langage. Une évolution récente au sein du courant dynamiciste permet sa contribution à l'élaboration de modèles de production langagière. On s'intéresse toujours au système et aux "structures du langage", mais on considère que celles-ci sont "déterminées par les fonctions qu'il sert" (Robert, 2002). Le linguiste doit "décrire les propriétés fonctionnelles des langues au travers de la variabilité de leurs formes et de leurs usages". La méthode de recherche est inductive ; il s'agit d'une méthode de travail "en reconnaissance" : on part des usages en contexte pour déterminer la valeur des formes et les mécanismes invariants. Je présenterai dans ce qui suit les travaux de Bronckart (2000) et de Culioli (1985 ; 2002) qui prennent en compte la variabilité des formes et usages.

2.1.1. Activité langagière

Bronckart (2000) décrit la langue comme un système dynamique, obtenu grâce à la négociation entre ses utilisateurs. Il se réfère à Saussure (1957 : 25-26) pour expliquer que "le 'système' ou mécanisme est lui-même fondamentalement ouvert sur les processus sociaux qui l'alimentent". Grâce à l'invention d'outils, les humains disposent de la faculté de négocier le sens, ou, dans les termes de Bronckart (2000), ils disposent d'"un mécanisme d'entente sur ce qu'est le contexte même de l'activité et sur la part que les individus instrumentés doivent y prendre". L'activité langagière a pour fonction de "commenter les activités ordinaires (ou non langagières)" et d'aider à les organiser, donc de les planifier, réguler, etc. Elle se matérialise dans des textes oraux ou écrits. Selon Bronckart, les textes sont "des *unités communicatives* complexes qui visent à produire un effet de cohérence sur leurs destinataires". Ils dépendent de la nature des activités qu'ils commentent, du média qui est utilisé pour les transmettre et des "conditions sociologiques de leur élaboration" (Bronckart, 2000). Tous les textes sont singuliers, mais il existe des contraintes au niveau de la syntaxe phrastique, des types de discours, au niveau des mécanismes de textualisation et de la prise en charge énonciative. En produisant, l'humain adopte et adapte un genre textuel. Il s'appuie sur des textes "déjà là".

Dans la citation suivante, Bronckart (2000) décrit la façon dont les humains se créent des représentations, en fonction du contexte linguistique dans lequel ils évoluent.

> La pensée humaine est littéralement "formatée" par les propriétés de la langue acquise, et plus précisément par les propriétés de l'état de langue auquel les enfants ont été confrontés. Ce filtrage des représentations d'objets et de relations du monde par les unités, les catégories et les structures d'une langue se traduit par une sorte d'*indifférenciation* entre mécanismes cognitifs et mécanismes linguistiques, et dès lors que l'enfant conçoit le monde à connaître comme stable, il se construit une image de sa langue comme étant constituée d'unités et de relations elles-mêmes stables et délimitant en conséquence des zones de possibles et des zones d'impossibles.

Ensuite, avec l'éducation formelle, l'enfant quitte cet état d'indifférenciation et commence à généraliser et à raisonner. Il est capable de réfléchir aux formes linguistiques qu'il est censé reproduire, et il est capable d'en modifier certains éléments en adoptant une démarche intertextuelle (cf. 3.4.2.3.2.).

2.1.2. Opérations énonciatives

Selon Fuchs et Le Goffic (1992, 129-130), la rhétorique, la théorie du discours, la théorie de l'argumentation, la grammaire et la logique sont les ancêtres des courants énonciatifs. Ces courants s'intéressent aux traces qui renseignent sur les conditions d'utilisation de la langue. Un de ces courants est la théorie des opérations énonciatives. Elle tient compte du rapport entre pensée et langage. Selon Culioli, l'objet d'étude de cette théorie est l'activité de langage. Cette activité s'exprime "à travers la diversité des langues et des textes" (Culioli, 2002 : 49). À ce sujet, Chini (2004 : 66) pose les questions suivantes. Comment peut-on "articuler l'universalité du langage et la singularité des langues" ? Comment "traiter de la relation entre la matérialité du texte et l'immatérialité de l'activité signifiante des sujets" ? Culioli propose une formalisation de la langue/du langage qui est censée permettre de découvrir des phénomènes linguistiques et de leur donner une forme abstraite pour qu'ils soient traitables (Culioli, 2002 : 103-104). Il préconise la mise en place d'"opérations métalinguistiques construites à partir des observations et non pas à partir d'un système existant" (cité dans Chini 2004 : 104). Le travail sur corpus semble permettre une telle observation de la langue.

> L'analyse sur corpus, si elle est bien menée, représentative, avec des textes variés, en apprend énormément (Culioli, 2002 : 131).

Culioli décrit trois représentants linguistiques qui semblent justifier l'utilisation de corpus et concordances non seulement pour faire des analyses linguistiques, mais aussi pour enseigner et apprendre une langue étrangère. Il s'agit de la paraphrase, du marqueur et de l'invariance.

2.1.2.1. Paraphrases

La paraphrase est une variante possible d'un énoncé. Cette variante introduit un léger changement de sens dans un même contexte d'énonciation. Le linguiste doit "rendre compte aussi de cette capacité que nous avons de faire proliférer" (Culioli, 2002 : 143). Ceci est possible si le linguiste observe les équivalents qui existent pour un terme ou une notion dans une langue donnée, ensuite dans d'autres langues. Il va construire une théorie à partir de ce qu'il a observé, revenir sur les observables, affiner la théorie, etc. : il effectue un "double travail de balancement : entre 'votre' langue et les langues, et aussi entre les phénomènes et la théorie" (Culioli, 2002 : 147). Cette recherche doit permettre, à un niveau d'abstraction très grand, de faire émerger l'organisation de la langue à partir de phénomènes. Un travail conceptuel mis en relation avec un travail empirique peut alors mener à des découvertes sur la langue.

2.1.2.2. Marqueurs

Culioli définit le marqueur comme "une espèce de résumé, de concentré de procédures qui déclenchent et activent des représentations" ; les marqueurs sont les "représentants matériels de ces opérations auxquelles nous n'avons pas accès directement" (Culioli, 2002 : 172), la trace d'"opérations à partir desquelles nous construisons des représentations, des catégories grammaticales, des mises en relation, de telle manière que nous puissions référer" (Culioli, 2002 : 174). Un marqueur est la trace matérielle d'une activité cognitive. Il peut montrer, par exemple, que quelqu'un veut convaincre, prendre de la distance, etc. Si l'on compare dans les langues du monde les différents procédés pour exprimer que l'on veut convaincre, prendre de la distance, etc., on peut parfois trouver des éléments ou procédés que plusieurs langues partagent. On arrive à des invariances.

2.1.2.3. Invariances

Les invariances sont des dénominateurs communs conceptuels. Culioli les appelle "proto-principes" (Culioli, 2002 : 173). Selon Chini, les invariances permettent de créer une base commune entre les langues, "leurs divergences relevant de choix différents parmi l'ensemble des possibles" (Chini, 2004 : 100).

> Cette base commune à tout système linguistique réside dans les diverses opérations d'observation et de structuration mentale du monde que le sujet est amené à gérer (Chini 2004 : 100).

L'apprenant d'une langue étrangère doit se construire un nouveau système, ce qui implique une activité cognitive. Les invariances peuvent être le point de départ d'un tel travail de construction qui peut être soutenu à l'aide d'activités conceptualisantes (cf. 2.8.2.).

2.2. Psycholinguistique et apprentissage de LE

À l'origine, la psycholinguistique a principalement étudié la façon dont les humains comprennent, produisent et acquièrent le langage. Les tendances récentes ont entraîné un certain éloignement entre psycholinguistique et acquisition, malgré le fait que les deux disciplines s'intéressent au traitement du langage (Labelle, 2001). Toutefois, certains chercheurs, comme R. Ellis, travaillent dans les deux domaines. Dans ce qui suit, des excursions vers l'acquisition ne sont donc pas à exclure.

2.2.1. Apprentissage de LE, un processus adaptatif

Selon Tricot (2007 : 10), l'apprentissage est l'expression de la dynamique adaptative des humains au cours de leur propre vie. Il est possible d'apprendre sans en être conscient et sans en avoir l'intention. Bruner (1938 : 264) estime que l'apprentissage "sans conscience" est une exception ; toutefois, il pense que la sérendipité ("faire par accident des découvertes heureuses et inattendues"), est à prendre en compte. Les "organismes vivants", selon l'expression de Tricot (2007 : 10), ont appris quelque chose quand ils sont capables ultérieurement de réutiliser une trace d'un environnement dans le même ou un autre environnement. Ces traces peuvent avoir comme source un événement, une différence, un signe, etc. Elles peuvent se matérialiser en des sensations, émotions, souvenirs, interprétations, concepts, savoir-faire et méthodes. Apprendre est donc "la capacité [des organismes vivants] à s'adapter à leur(s) environnement(s)" (Tricot, 2007 : 10). L'apprentissage peut également être organisé artificiellement ; il y a des "environnements d'apprentissage conçus par des humains" (Tricot, 2007 : 10). Ces environnements sont crées afin d'anticiper des situations futures auxquelles les humains ne pourront faire face que s'ils possèdent certaines capacités. Voici ce que Tricot (2007 : 10) précise à ce sujet.

> Un enseignement est mis en place quand l'apprentissage visé n'est pas (…) immédiatement adaptatif. Pour l'essentiel, nous enseignons des connaissances qui ne permettent pas forcément à l'individu de répondre aux contraintes de son environnement actuel, mais de son environnement futur.

L'apprentissage organisé "artificiellement" doit chercher à compenser l'absence de besoin immédiat en créant d'autres motivations que celles qui existent "naturellement" (2.9). Ce point de vue est partagé par Grabe et Kaplan (1996 : 266) en ce qui concerne l'organisation institutionnelle de tâches de production écrite (cf. 3.).

2.2.2. Input, intake, output *et apprentissage de LE*

En psychologie cognitive et en acquisition des langues, l'environnement linguistique auquel on expose les apprenants est souvent appelé *input*. Ce terme est issu d'un modèle cybernétique de la psychologie qui compare le fonctionnement de l'esprit humain à celui d'un ordinateur. Gaonac'h décrit ce modèle tout en émettant des réserves sur sa validité qu'il présente comme restreinte. Il précise qu'il convient d'étudier le rôle que joue la mémoire dans "les traitements qui aboutissent à l'élaboration de nouvelles connaissances" (Gaonac'h, 1991 : 109). Il n'y a pas de relation linéaire entre *input*, *intake* et *output*.

INPUT	**&**	***INTAKE***	**&**	***OUTPUT***
données		saisie intérimaire		sortie

Figure 15 – *Input*, *intake* et *output*.

L'*input* est ce à quoi on expose les apprenants, donc un texte oral ou écrit. L'*output* est ce que les apprenants produisent, donc un dialogue oral ou une production écrite. La définition d'*intake* varie selon l'importance que l'on accorde soit à l'exposition à un *input*, soit à la production d'un *output*. L'*intake* est soit ce que les apprenants retiennent après avoir été exposés à un *input*, soit ce que les apprenants retiennent après avoir produit de l'*output*. Dans ce qui suit, je détaillerai deux hypothèses concernant le rôle de l'*output* et de l'*input* pour l'apprentissage d'une langue étrangère. Un paragraphe sur l'*intake* terminera cette partie.

2.2.2.1. *Output hypothesis*

Certains chercheurs, en particulier Swain et Lapkin (1995), préconisent l'importance de la production d'un *output* compréhensible pour l'acquisition d'une langue étrangère. Ils précisent leur point de vue comme suit.

> *In producing an L2, learners will on occasion become aware of (i.e. notice) a linguistic problem. Noticing a problem can "push" learners to modify their output* (Swain et Lapkin, 1995).

En effet, selon Swain et Lapkin (1995), la production verbale donne aux apprenants l'occasion de se rendre compte de leurs besoins d'apprentissage. Swain et Lapkin se réfèrent à la formule *noticing the gap*, évoquée également par Schmidt et Frota (1986 : 311) et par R. Ellis (2003 : 149). Cette formule désigne les éléments qui montrent l'écart entre une production linguistique de locuteurs novices et celle de locuteurs experts. Swain et Lapkin (1995) ont étudié le rôle de l'*output* à l'aide d'une expérience menée auprès d'adolescents canadiens qui ont participé à un programme d'apprentissage du français, langue étrangère, en immersion. L'exécution et la révision d'une tâche de production écrite ont été accompagnées d'un protocole expérimental de type *think aloud*. Ce type de protocole invite les apprenants à s'exprimer verbalement sur ce qu'ils font pendant qu'ils exécutent une tâche, par exemple, en rédigeant. Le scénario s'est distingué par un guidage très discret. Parfois, les apprenants ont ressenti le besoin d'observer comment des locuteurs experts ont formulé dans une situation comparable à celle qui leur a posé problème.

> *When learners cannot work out a solution, they may turn to input, this time with more focused attention, searching for relevant input. Or, they may work out a solution, resulting in new, reprocessed output (Swain et Lapkin, 1995).*

Les résultats de l'étude de Swain et Lapkin (1995) confirment que les apprenants ont été en mesure de repérer leurs "lacunes" linguistiques (*gaps*) et d'observer un *input* de façon ciblée, pratiquement sans être guidés, dans le but de réviser leur texte. Toutefois, Swain et Lapkin admettent qu'un guidage plus intense aurait pu réduire le nombre d'hypothèses erronées et de sur-généralisations faites par les apprenants ayant participé à cette étude (cf. 2.2.2.1.). Swain et Lapkin (1995) précisent également que la production d'un *output* compréhensible n'est qu'une des façons susceptibles d'aider les apprenants à améliorer leur interlangue.

2.2.2.2. *Input hypothesis*

Certains chercheurs défendent l'idée que l'*intake* – en d'autres termes, l'apprentissage – est favorisé par l'exposition à un *input* structuré, c'est-à-dire, un *input* qui contient les éléments que les apprenants sont censés remarquer afin de les retenir plus facilement : "*input that has been specially designed to include specific linguistic features*" (Ellis et Fotos, 1999 : 65).

La théorie connexionniste concernant le traitement de l'information suggère que l'on stocke dans le cerveau les représentants d'un fait langagier que l'on rencontre, par exemple, tous les noms français qui se terminent en "-oire". Plus on a rencontré et enregistré de représentants d'un certain type de mots, plus le cerveau établit des connexions neuronales qui créent des

associations entre les représentants de ce type de mots et leurs traits. Ceci permet, par exemple, d'extrapoler que les noms français qui se terminent en "-oire" sont des noms masculins.

> *Learning takes place through the strengthening and weakening of interconnections in response to the frequency of examples encountered in input. The result of such learning is a network of units that acts as if it knew the rules, but the rules themselves exist only as association strengths distributed across an entire network. They are not explicitly represented in the system as abstract principles, and do not guide or control lower level processes* (Schmidt, 1990).

Randall (2007 : 131) énumère plusieurs conditions qui doivent être remplies pour que des connexions entre divers mots puissent être faites et que les mots et leurs caractéristiques collocatives puissent être stockés dans la mémoire à long terme.

- Il faut une grande quantité d'*input*.
- L'*input* doit apparaître dans un contexte susceptible de permettre à l'apprenant de relier forme et sens.
- L'*input* doit être présélectionné par l'enseignant afin de favoriser le repérage.
- Les apprenants doivent diriger leur attention sur des traits caractéristiques des *patterns* qui font l'objet de l'apprentissage.

Un dispositif basé sur des corpus et concordances et sur l'exécution de tâches de production écrite par les apprenants semble pouvoir remplir les conditions évoquées par Randall (2007) : le corpus met à disposition une grande quantité d'*input*. La tâche sociale (cf. 2.8.1.2.) fournit un contexte qui fait sens. L'*input* peut être plus ou moins présélectionné par l'enseignant ; l'action de regrouper dans un corpus des textes appartenant à un même genre textuel est une telle forme de présélection. Il est possible de présélectionner quelques lignes de concordance selon les besoins du moment et d'inviter explicitement les apprenants à diriger leur attention sur l'un ou l'autre mot clé afin qu'ils puissent observer son environnement et en tirer des conclusions concernant la révision de leur texte. Les concordances permettent de réduire le champ de vision des apprenants afin de favoriser la concentration sur des traits pertinents de certains mots clés. Par contre, les concordances peuvent avoir comme conséquence une perte de contexte et donc de sens global ; d'où l'intérêt de veiller à garder un lien entre concordances et corpus et de montrer aux apprenants qu'un tel lien existe : ils doivent pouvoir passer facilement des concordances aux corpus et vice versa. La collection d'outils *Antconc3.2.1w* (Anthony, 2007) donne l'occasion, de façon particulièrement confortable, de naviguer entre corpus et concordance.

2.2.2.3. Développement de l'*intake*

Narcy-Combes (2005b) donne quelques précisions concernant l'*intake* qu'il définit comme résultat d'un traitement local de données.

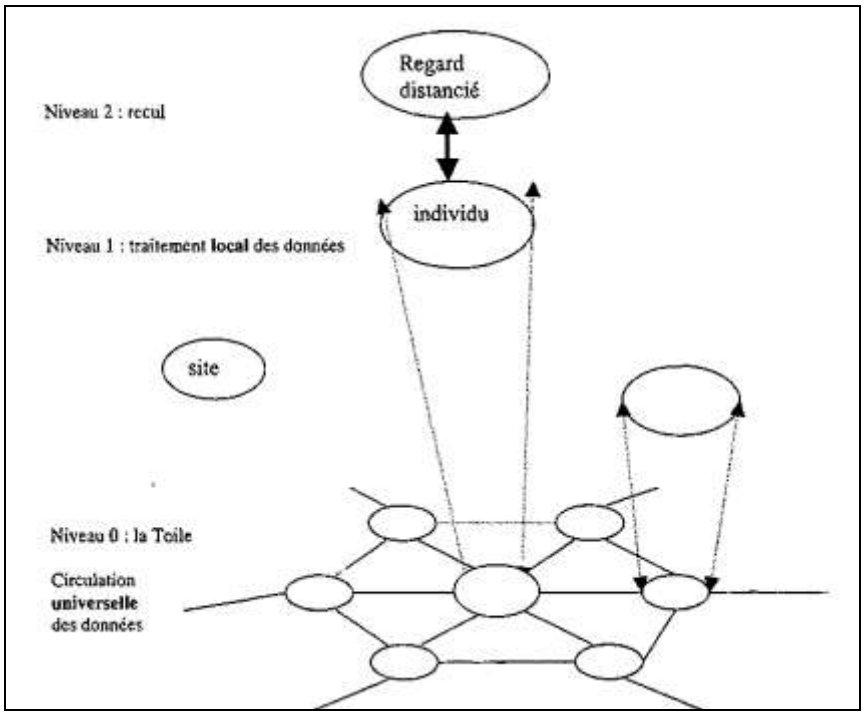


Figure 16 – Traitement local des données.

Selon Narcy-Combes, la diffusion d'informations sur la Toile met en relief "le dilemme universel-local en ce qui concerne le traitement des données". Sur la Toile, l'information circule de façon universelle et globale. Elle est traitée localement par les internautes qui la reçoivent et l'interprètent de façon individuelle, en fonction de ce qu'ils sont et de là où ils sont, donc en fonction de leur environnement socioculturel et de façon subjective. Ceci souligne le fait qu'il n'y a pas de relation linéaire entre *input*, *intake* et *output*. Selon Narcy-Combes (2005 : 108), l'élément qui invalide objectivement la linéarité de cette relation se trouve dans la créativité de l'énonciateur. L'*output* résulte davantage de cette créativité qu'il n'est la reprise de l'*input*. Dans le paragraphe 3.1.2., je reprendrai ce point de vue en décrivant la stratégie de transformation de connaissances et la stratégie de connaissances rapportées proposées par Bereiter et Scardamalia (1987). La créativité par rapport aux données distingue un locuteur ou scripteur humain d'un ordinateur qui ne rend (en les combinant de multiples façons) que les informations qui ont été entrées dans la machine. Bruner (1991 : 20) résume cette problématique comme suit.

> Ainsi, petit à petit, l'accent s'est déplacé de la signification à l'information, et de la construction de la signification au traitement de l'information. Ce sont pourtant des choses bien différentes. À l'origine de ce glissement, une métaphore qui est devenue dominante, celle de l'ordinateur ; c'est à cette aune que l'on a fini par juger qu'un modèle théorique était valable. L'information ne s'intéresse pas à la signification. En termes informatiques, l'information contient un message précodé dans le système. Le sens précède le message. Il n'est pas produit par l'ordinateur, et il n'en relève nullement.

Les utilisateurs d'une langue cherchent à créer du sens dans un contexte précis pour agir dans un environnement social. Le sens obtenu reflète des concepts déjà en place, concepts liés, entre autres, à l'appartenance culturelle et linguistique du locuteur.

> Or, au niveau de cette créativité, des phénomènes de nativisation (…) influent sur la conception même de l'énoncé, au niveau de la pensée hors code (Narcy-Combes, 2005 : 108).

Il convient donc de tenir compte du fait que les apprenants perçoivent et interprètent l'*input* selon des critères personnels qui sont liés à des facteurs internes, donc neurologiques et mentaux, ou externes. Hulstijn (2007) évoque certains facteurs externes qui peuvent entrer en jeu lors de l'acquisition d'une langue étrangère, comme, par exemple, la nature et la quantité de l'*input* et l'aptitude de l'apprenant (cf. 2.9.).

Schmidt (1990) considère l'*intake* comme le résultat de processus mentaux qui permettent de traiter les informations de l'*input*, par exemple, par la généralisation à partir de segments en contexte. Le traitement est également favorisé par la formulation d'hypothèses et par la formulation de caractéristiques linguistiques à partir de l'*input*. Un certain nombre de connexions neuronales semblent être nécessaires pour que les apprenants puissent découvrir des références linguistiques. Celles-ci peuvent être formulées en termes métalinguistiques qui oscillent entre les pôles "métalangage très personnel" et "métalangage très normé" (Faerch, 1986, cité dans Bange, 2003). Je ne dispose pas des moyens techniques permettant l'observation de la création de connexions neurologiques, mais je peux tenter de favoriser de telles créations en donnant accès à des occurrences en grand nombre, et je peux donner aux apprenants l'occasion de verbaliser leurs découvertes concernant le fonctionnement de la langue.

2.2.2.4. Ma définition d'*intake*

Ce qui précède m'amène à définir l'*intake* comme le résultat d'une interaction réussie entre *input* et *output*. L'*input* auquel les apprenants sont exposés doit être adaptable à des niveaux et

besoins individuels. En situation d'apprentissage en institution scolaire, les besoins sont créés par la nature de la tâche de production que l'on confie aux apprenants, donc de l'*output* qu'ils sont censés produire. L'*output* que l'on souhaite obtenir doit avoir des caractéristiques comparables à celles de l'*input*. Cela est possible grâce à la mise à disposition de corpus et concordances lors d'un projet d'écriture.

2.3. Définitions d'unités polylexicales

Pour désigner les unités que les apprenants sont susceptibles de repérer dans l'*input* en LE, Schmidt (1990) emploie les termes d'*instances* ou d'*examples*. Randall (2007) utilise le terme de *patterns* qui sont des groupes de mots composés d'au moins deux éléments distincts, formant une unité de sens. Des travaux sur de telles unités existent depuis les années 50 (Firth, 1951). Ils prennent une ampleur particulière à partir des années 80, entre autres en raison du développement des linguistiques de corpus. Pour un aperçu historique détaillé, il est utile de lire N. Ellis (2008 : 1-13). Dans ce qui suit, je vais présenter cinq points de vue concernant la nature et l'emploi des unités polylexicales. Je terminerai par ma propre définition du phénomène collocatif. Le sous-chapitre 2.4. définira le rôle que les collocations peuvent jouer pour l'apprentissage de la langue.

Si l'on cherche à définir les unités polylexicales, on se trouve devant une multitude de termes, dont chacun est rattaché à une ou plusieurs approches théoriques spécifiques. Voici des termes qui apparaissent en lien avec l'apprentissage de LE depuis les années 70 (N. Ellis, 2008 : 6)[2] :

- *holophrases* (Corder, 1973) ;
- *prefabricated routines and patterns* (Hakuta, 1974) ;
- *formulaic speech* (Wong-Fillmore, 1979) ;
- *memorized sentences and lexicalized stems* (Pawley et Syder, 1983) ;
- *lexical phrases* (Nattinger, 1980) ;
- *formulas* (R. Ellis, 1994 ; McLaughlin, 1995) ;
- *chunks* (N. Ellis, 1996) ;
- *constructions* (N. Ellis, 2003 et 2006).

Les termes de *patterns*, *lexical phrases*, *chunks* et *formulas* sont particulièrement bien représentés dans la littérature récente, comme l'illustre l'index des mots clés dans Granger et

[2] Pour les références concernant ces termes, voir N. Ellis (2008).

Meunier (2008 : 257-258). Le terme qui s'est clairement imposé ces dernières années est celui de collocation ; il est référencé dans cinquante pages de l'ouvrage que je viens de citer. Je vais donc retenir ce terme, même s'il y a aussi des arguments en faveur du terme de "formule" qui semble permettre une interprétation plus ouverte des unités polylexicales.

Le tableau ci-dessous présente des termes concernant la polylexicalité et les courants théoriques qui les emploient. Il met en évidence que certains de ces courants s'intéressent davantage que d'autres à des applications pratiques. Ceci a pour conséquence des degrés variables de formalisation.

Termes désignant des unités polylexicales	**Courants ou domaines**	**Motivations théoriques et applications pratiques**	**Références**
Exemplars *Patterns* *Chunks*	Acquisition	Motivation théorique : expliquer l'acquisition de LM/LE (fluidité).	R. Ellis (2003) Randall (2007 : 131) N. Ellis (2008)
Formule *Instance*	Psycholinguistique	Motivation théorique : expliquer l'apprentissage de LM/LE (accélération du temps de traitement de l'*input* ; mémorisation).	Gaonac'h (2006 : 145) Schmidt (1990)
Collocations	Linguistique de corpus ; phraséologie ; lexicologie ; lexicographie	Motivation théorique : décrire les affinités sémantiques et grammaticales qui existent les mots ; décrire les spécificités de genres et types de textes ; statistique textuelle. Applications pratiques : créer des outils de travail pour des spécialistes de la langue (traducteurs, terminologues, statisticiens, etc.) : corpus réfléchis, bases de données lexicales, dictionnaires.	Firth (1951) Partington (1998) Sinclair (2004) Tutin & Grossmann (2002)
Blocs lexicalisés	Didactique de LE	Motivation théorique : comprendre l'enseignement/apprentissage de LE (collocations en lien avec des situations d'emploi). Applications pratiques : dispositifs d'enseignement/apprentissage de LE ; création de matériaux d'apprentissage de LE.	Chini (2001) ; Demaizière et Narcy-Combes (2005)

Tableau 7 - Unités polylexicales et courants/domaines de recherche.

Selon le tableau ci-dessus, pour les acquisitionnistes, les collocations sont, entre autres, censées expliquer des phénomènes liées à la fluidité de la production. Les psycholinguistes s'intéressent aux collocations en lien avec des questions concernant la mémorisation du lexique et concernant le traitement de l'*input*. Les linguistes de corpus et les lexicologues/lexicographes tentent de formaliser les collocations, entre autres, dans le but de les rendre interprétables par l'ordinateur, en ayant comme objectif la création de bases de données textuelles ou lexicales. En didactique, les applications pratiques peuvent, comme en lexicographie, motiver la recherche théorique. Je pense, par exemple, à la création de logiciels utilisant des bases de données, comme le cédérom *Scripertexte*, distribué sous la dénomination *Écrire en lisant des récits de vie* (Crinon et Vigne, 2002 ; cf. 3.7.2.). Dans le milieu de la didactique, le terme de "bloc lexicalisé" est utilisé depuis un certain temps, certes de façon modeste, pour désigner des phénomènes de locutionnalité (Chini, 2003). Demaizière et Narcy-Combes (2005) définissent les blocs lexicalisés comme "des instances d'utilisation, sortes de modèles tout faits en lien avec leurs contextes d'emploi". Je choisirais peut-être un autre terme que "modèles tout faits" puisqu'il ne semble pas englober les collocations semi-figées. La définition de Demaizière et Narcy-Combes (2005) fait ressortir que parfois, une formulation peut paraître acceptable du point de vue de la forme, mais si elle n'est pas d'usage dans d'un milieu socioculturel donné, elle risque de ne pas "fonctionner". Šimunek (2007 : 21) fait allusion à ce type de cas de figure en utilisant le terme de "*Institutionalisierung eines Ausdrucks*" [institutionnalisation d'une expression].

Dans ce qui suit, je propose de regarder de plus près le phénomène collocatif afin de mieux définir ses multiples facettes dont certaines paraissent plus floues et malléables que d'autres. Pour ce faire, j'exposerai le point de vue de linguistes de corpus, de lexicologues et lexicographes, de phraséologues, de psycholinguistes et de didacticiens.

2.3.1. Perspective des linguistes de corpus

Les contextualistes définissent les collocations de façon large, ceci dans la tradition de Firth (1957) dont les travaux initient le changement d'une approche sémantique conceptuelle vers une approche sémantique formelle et contextuelle (Partington, 1998 : 15). Sinclair et Williams sont des représentants de ce courant. Sinclair, ancien étudiant de Firth, s'intéresse au sens généré par la proximité des mots dans un texte davantage qu'à la valeur sémantique "isolée" des mots (Partington, 1998 : 16). Les contextualistes définissent les collocations comme des cooccurrences statistiquement fréquentes d'éléments linguistiques dans un contexte donné.

Leur rôle est d'assurer la cohésion du texte (Tutin et Grossmann, 2002). Les contextualistes ont des affinités avec les phraséologues, mais leur regard sur les collocations paraît plus "technique" que celui des phraséologues. En statistique textuelle, le terme de segment désigne une forme particulière de collocation (Lebart et Salem, 1994 : 61). Les composantes des segments apparaissent de façon linéaire, comme des blocs plus ou moins lexicalisés.

2.3.2. Perspective des lexicologues et des lexicographes

Depuis quelques années, un deuxième courant s'intéressant aux collocations prend de l'ampleur. Il est initié par des lexicologues et lexicographes. Tutin et Grossmann (2002) adhèrent à ce courant. Ils s'inspirent des théories de Mel'čuk (1994) et de Hausmann (1989). Le phénomène collocatif y est décrit de façon plus étroite afin de rendre possible une formalisation. La combinaison à deux mots est, en l'état actuel des choses, l'unité privilégiée, car l'ordinateur arrive à la gérer de façon satisfaisante. Elle se compose d'un élément de base et d'un élément variable dont le sens est déterminé en fonction du terme de base (Tutin et Grossmann, 2002). Des critères comme la transparence et l'opacité, la relation syntagmatique et sémantique permettent d'obtenir la typologie suivante :

- collocations opaques (association de mots qui paraît arbitraire) ;
- collocations transparentes (association de mots qui paraît interprétable, mais qui est peu prédictible) ;
- collocations régulières (association de mots qui est prédictible en suivant des règles sémantiques).

Un des objectifs des lexicologues et lexicographes est de créer, à l'aide de collocations, des outils de travail numériques, comme des bases de données lexicales, des dictionnaires, etc. Le phénomène collocatif semble bien cerné dans ce domaine, mais sa définition paraît trop étroite pour une transposition dans le domaine de la didactique.

2.3.3. Perspective des phraséologues

La phraséologie est une des branches de la linguistique qui s'intéressent à la polylexicalité. Cette branche est historiquement plus ancienne que le contextualisme. La publication du *Traité de stylistique française* de Bally (1909) est considérée comme date clé pour la naissance de ce courant. L'objet privilégié de la phraséologie est la sémantique conceptuelle, initialement surtout à travers l'étude d'idiomes, d'expressions de type proverbes, de maximes, etc., mais on observe l'ouverture vers d'autres types de collocations (Šimunek, 2007 : 26-27). On se rapproche de la rhétorique et de l'étude du texte. Les phraséologues cherchent à savoir

quelles collocations sont représentées dans quels genres ou types de textes. Le rôle des collocations comme indicateurs d'un genre de texte n'est plus à démontrer.

> *Many kinds of lexical items, including prefabs, function as powerful indicators of register, and in most circumstances it is important for the writer to signal the register to which the text belongs* (Partington, 1998 : 20).

Les phraséologues s'intéressent aux degrés de figement et de transparence des collocations. Partington (1998 : 25-26) rappelle que de nombreuses collocations sont partiellement modifiables ; seules certaines collocations sont réellement figées. Les irrégularités grammaticales peuvent être des indices de figement. En lien avec les recherches dans le domaine de l'analyse de discours, un certain renouveau de l'approche sémantique conceptuelle semble se dessiner : Siepmann (2008 : 185-202) présente un nouveau type de dictionnaires d'apprentissage, basés sur des collocations. Les entrées y sont classées de façon onomasiologique, donc par concepts, au lieu d'être présentées selon un ordre alphabétique ou thématique.

2.3.4. Perspective des psycholinguistes

Gaonac'h (2006 : 145) regroupe les expressions idiomatiques, routines et *patterns* sous le terme de "formules". En situation d'apprentissage précoce d'une LE, les formules (expressions idiomatiques, routines, *patterns*) permettent la mémorisation de "chaînes linguistiques imitées et répétées, sans référence nécessaire à une règle, et non analysées, dans un premier temps tout au moins" (Gaonac'h, 2006 : 145). Les énoncés préconstruits sont considérés comme assez stables. Ils fournissent à l'apprenant une référence aisément disponible en situation d'apprentissage et de communication. Les formules autorisent une certaine flexibilité : elles ne sont pas entièrement ou pas toutes figées. Des automatismes peuvent être créés par le biais de chants, poésies, prières, formules de politesse, proverbes, etc. Si l'on rattache la formule à une situation, la formule peut être plus facilement maîtrisée. Le surentraînement peut mener à une analyse ultérieure et même mener à l'analyse du fonctionnement langagier (Gaonac'h, 2006 : 146).

2.3.5. Perspective des didacticiens

En didactique, il paraît essentiel de décrire les conditions de l'emploi de collocations par les apprenants. Le concepteur d'un scénario pédagogique basé sur la production écrite, tenant compte de genres textuels, gagne alors à connaître la fréquence des collocations dans un

corpus de textes qui représentent le genre textuel du scénario. Ainsi, il est en mesure de faire l'inventaire des collocations pertinentes. Celles-ci peuvent être trouvées à l'aide de méthodes de statistique textuelle et doivent ensuite être revues manuellement par un spécialiste du genre textuel. Il n'est pas inutile de s'adresser à des spécialistes de statistique textuelle (Lebart et Salem, 1994) afin de se faire aider dans une telle entreprise.

Comment définir le phénomène collocatif pour des besoins d'enseignement/apprentissage ? Selon Partington (1998 : 17), il n'y a pas de raison de privilégier une combinaison de deux mots quand on s'intéresse au phénomène de *patterning*, donc au phénomène qui s'intéresse aux combinaisons de mots qui tendent à apparaître ensemble dans un texte donné. Il propose de remplacer le terme de collocation par *combinatorics* afin d'élargir le champ des phénomènes collocatifs.

2.3.6. Ma définition du phénomène collocatif

Contrairement aux lexicologues, je ne restreins pas ma définition de "collocation" aux segments qui ne comportent que deux éléments de base. La fréquence des collocations est intéressante dans la mesure où elle permet de révéler les caractéristiques linguistiques d'un genre textuel donné. Je m'intéresse particulièrement aux collocations potentiellement transparentes, car je m'attends à ce que les apprenants de LE les identifient plus facilement que les autres types d'unités polylexicales, à savoir les expressions idiomatiques et les collocations opaques. Les collocations semi-figées (ou, comme je dirais : partiellement modifiables) sont importantes pour mon travail de recherche ; la façon dont les apprenants les intègrent dans une production écrite est un des critères permettant d'évaluer leur compétence en LE. Ce qui précède m'amène à définir le terme de collocation comme suit.

Une collocation se compose d'au moins deux mots qui se suivent ou qui sont séparés par d'autres mots. Elle ne dépasse pas la limite d'une phrase, la phrase étant définie ici comme mots entre deux signes de ponctuation limitatifs, de type point, point d'exclamation et point d'interrogation. La collocation forme une unité de sens, et elle est soumise à certaines restrictions syntaxiques, sémantiques et pragmatiques. Elle se caractérise par un certain degré de figement et de récurrence au sein d'un corpus.

Figure 17 – Ma définition de "collocation".

Illustrons la définition ci-dessus par des exemples du domaine de la langue française : "dans le" n'est pas une collocation, mais "à la maison" en est une, "Je viendrai demain." n'en est pas une, mais "On s'appelle !" en est une. Le modèle spiralaire de Gläser (1986 : 54ff., cité dans Šimunek, 2007 : 26) permet de visualiser les divers degrés de figement possibles. Les collocations figées opaques, bien représentées par la classe des proverbes ou des maximes, se trouvent dans le centre de la spirale. Ensuite, nous trouvons des collocations moins opaques, mais restreintes, pour arriver aux collocations de moins en moins restreintes du point de vue sémantique et/ou grammatical. La restriction peut être visible, par exemple, au niveau de la rection du verbe ou au niveau de l'affinité sémantique qui existe entre un nom et un adverbe et qui est mesurable statistiquement dans un corpus.

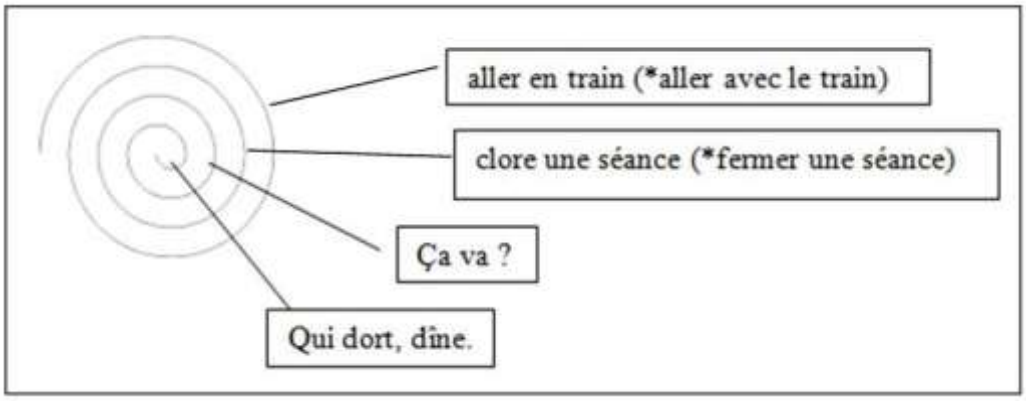


Figure 18 - Illustration du modèle spiralaire de Gläser.

Ce modèle rappelle partiellement le modèle lexicographique de Tutin et Grossmann (2002) : il y a un continuum qui part de la collocation opaque et très restreinte vers la collocation plus libre et de plus en plus "transparente", dont on peut plus facilement reconstituer le sens à partir de ses unités simples. Le modèle de Gläser paraît opérationnel du point de vue descriptif. Toutefois, la notion de "transparence", opérationnelle en lexicologie, est-elle également pertinente en didactique ? Comment savoir ce qu'un apprenant peut trouver transparent ? Les données de Šimunek récoltées lors d'une recherche-action montrent, en effet, qu'il est difficile de prévoir à l'avance quelles unités lexicales les apprenants seront en mesure d'identifier et de réutiliser. Une démarche d'identification individualisée peut aider à résoudre ce dilemme. La citation d'un apprenant, "*Ach, so sagen die das*" [Ah, c'est comme ça qu'ils disent] (Šimunek, 2007 : 11), montre l'intérêt de privilégier la perspective de l'apprenant. Si un apprenant n'est pas en mesure de trouver par lui-même une collocation dont il a besoin dans une tâche de production, on peut l'encourager à la repérer dans un *input* adapté, par exemple dans un corpus réfléchi, présenté sous forme de concordance. Pour un

didacticien, il n'est pas très important d'aller loin dans la formalisation des collocations : les collocations pertinentes sont les combinaisons de lexèmes qu'un apprenant n'aurait pas trouvées sans aide extérieure. Cette démarche peut paraître utilitaire. Doit-on la contester pour autant ? Essayons de la confronter à la démarche inverse, celle que présentent Aguado (2002b : 44) et O'Keeffe, McCarthy et Carter (2007 : 69-70). Aguado propose aux enseignants de faire un inventaire des collocations qu'ils souhaitent enseigner et faire apprendre à un public d'apprenants donné. O'Keeffe, McCarthy et Carter décrivent comment de tels inventaires peuvent être créés pour l'anglais. Ils présentent des statistiques qui illustrent la fréquence de mots simples ou collocations composées de deux, trois ou quatre mots dans un corpus donné. Ils proposent des listes comportant les vingt mots ou collocations les plus fréquents (O'Keeffe, McCarthy et Carter, 2007 : 65-68). De telles listes sont, à mon avis, à utiliser avec prudence. Premièrement, des unités sémantiquement non pertinentes apparaissent dans de telles statistiques ("*in the*" versus "*you know*"). Deuxièmement, les collocations identifiées ainsi ne posent pas forcément problème à tous les apprenants. Admettons qu'une analyse de corpus d'apprenants et une démarche comparative peuvent permettre de lister certaines collocations qui risquent de poser problème pour un apprenant d'une LE donnée, locuteur natif d'une LM donnée (Šimunek, 2007 : 40). Toutefois, une telle entreprise n'est pas forcément pertinente pour tous types de production. En plus, il est possible que les apprenants mémorisent parfois des collocations qui ne correspondent pas à une unité de sens, mais plutôt à une unité phonologique. Il paraît donc justifié de leur donner l'occasion de faire des choix personnels de sélection. Les recherches de Leeke et Shaw (2000 : 284, cités dans Šimunek, 2007 : 80) semblent aller dans ce sens. Leeke et Shaw ont demandé à un groupe d'apprenants de LE de faire une liste d'unités lexicales en vue de réaliser une tâche de production. Un deuxième groupe a été invité à dresser une liste comparable, sans avoir de tâche à produire. Les chercheurs ont observé que les apprenants ayant eu une tâche à réaliser ont sélectionné davantage de collocations dans l'*input* que les apprenants de l'autre groupe. Les listes des derniers comportaient surtout des mots simples. Granger et Meunier (2008 : 248) défendent également le point de vue qu'il peut être justifié que les apprenants s'inspirent de collocations trouvées dans un *input*. Dans le projet présenté par Šimunek, ce procédé n'a pas été utilisé de façon significative. Il semble être peu populaire, peut-être parce qu'il semble être en contradiction avec la consigne scolaire de formuler "avec ses propres mots". Šimunek a toutefois observé que certains apprenants ont repris des éléments dans l'*input* des correspondants, locuteurs experts, pour des raisons affectives (Šimunek, 2007 : 295). Les apprenants ont réussi à utiliser des collocations suite à des négociations ou à des recherches

dans le dictionnaire, mais une acquisition durable de collocations n'a pas pu être documentée de façon satisfaisante (Šimunek, 2007 : 323). Il se trouve que les *peer teachers* participant au projet de Šimunek ont plusieurs fois délibérément ménagé leurs correspondants en décidant de tolérer certaines formulations problématiques : "*Gar nichts sagen wir da zu dem Deutsch erstmal, oder ?*" [On ne dit rien pour l'instant concernant leur allemand, non ?] (Šimunek, 2007 : 327). Une remarque concernant la non-acceptabilité d'une formulation peut, en effet, déstabiliser certains apprenants. Šimunek évoque aussi l'obstacle épistémologique d'une trop grande assurance : les apprenants qu'il a observés sont de futurs enseignants d'anglais. Leurs connaissances ont permis un échange fluide avec les correspondants canadiens. Ils n'ont donc pas forcément eu l'intention d'améliorer leur interlangue, et ils n'ont pas toujours perçu leurs "lacunes" linguistiques.

2.4. Collocations et apprentissage de LE

2.4.1. Collocations et mémorisation

Selon Logan (1988) et Shekan (1997), le langage fonctionne comme un *exemplar-based system*. On apprendrait donc, entre autres, par mémorisation de collocations et de leurs contextes d'emploi (Narcy-Combes et Walski, 2004). N. Ellis met également en avant l'importance des collocations pour la production langagière.

> *Thus native-like speaking is not a matter of syntactic rule alone. Speaking natively is speaking idiomatically, using frequent and familiar collocations* (N. Ellis, 2001 : 45).

Les théories concernant l'apprentissage à l'aide de références et l'apprentissage par mémorisation de collocations ne sont pas contradictoires. Selon Narcy-Combes et Walski (2004), les deux systèmes fonctionnent en parallèle. Ils expliquent qu'il existe "un accord théorique sur le rôle décisif que joue le traitement du sens (*deep-processing*) pour favoriser l'apprentissage et sur la nécessité de mettre en place une observation consciente des données langagières (*noticing*)".

La production langagière relève aussi bien d'une application de références que de la production automatisée de modèles lexicalisés, adaptés à une situation donnée. Dans tous les cas, les apprenants doivent traiter du sens, focaliser leur attention et avoir conscience du fait que leurs activités cognitives sont nativisées (Narcy-Combes et Walski, 2004).

Aguado défend le point de vue que s'exprimer de façon fluide en LE implique une phase de mémorisation préalable.

> *Um überhaupt kreativ werden zu können, ist es zunächst einmal erforderlich, eine Wissensbasis aufzubauen, die es zu memorisieren gilt*. [La condition de l'utilisation créative de la langue nécessite la construction préalable d'une base de savoirs qu'il convient de mémoriser] (Aguado, 2002a : 41).

Une telle "base de savoirs" doit contenir des mots et leurs relations avec d'autres mots. Aguado précise que pour l'acquisition de LM, il a été observé que les enfants emmagasinent d'abord une certaine quantité de collocations sous forme de *patterns* figés. Ils apprennent ensuite à combler les espaces variables au sein des collocations, en s'appuyant sur des références grammaticales de base (Aguado, 2002a : 50). Fillmore (1979) désigne les segments modifiables par "*formulaic frames with slots*". En LE, les collocations peuvent soutenir les apprenants peu avancés, ayant un savoir grammatical peu développé. Comme pour la LM, dans un deuxième temps, la connaissance de collocations peut mener à la découverte de régularités linguistiques (Foster, 2001 : 76, cité dans Aguado, 2002a : 53). Pour des apprenants avancés, ayant des besoins de formulation plus complexes, les collocations semblent alors également jouer un rôle important (Aguado, 2002a : 53).

Comment mémorise-t-on les collocations ? Le modèle de la boucle phonologique ci-dessous, établi par Randall (2007 : 126) propose une formalisation de ce procédé.

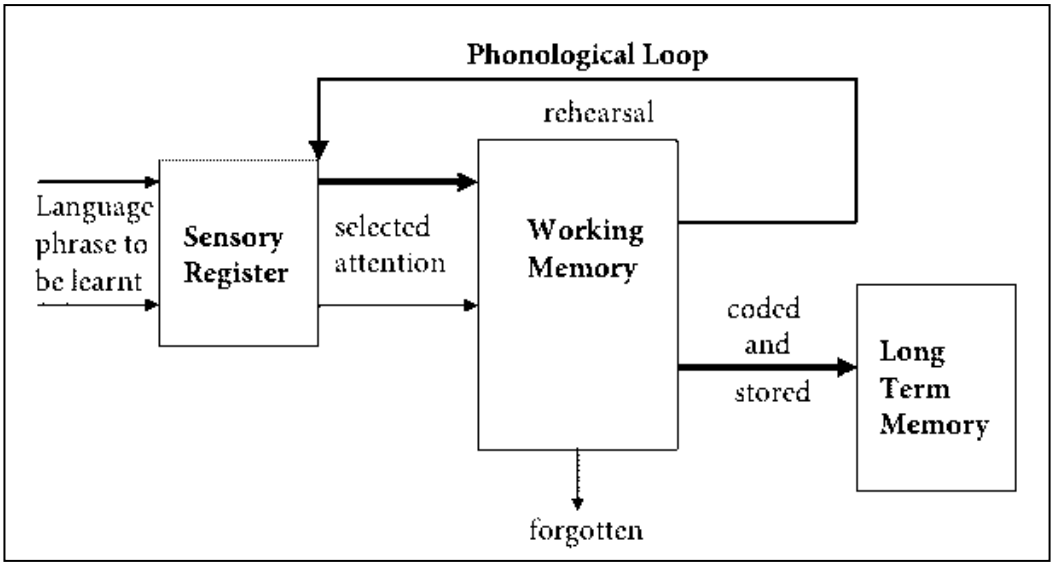


Figure 19 - Modèle d'apprentissage par exploitation de la boucle phonologique.

Selon ce modèle, l'apprentissage par mémorisation de collocations est de type associatif. Plus on rencontre le nouvel élément, mieux on le retient. Ce processus caractérise particulièrement l'acquisition de la LM. Il est favorisé par une très grande quantité d'*input* et de situations de communication rencontrées, largement réparties dans le temps. Les apprenants d'une LE sont exposés à une quantité d'*input* beaucoup moins importante, et ils ont beaucoup moins de temps pour pratiquer la LE (Randall, 2007 : 127). Il convient donc de leur faire rencontrer des collocations de façon plus "artificielle", comme, par exemple, dans des concordances qui donnent accès à un *input* multiple et varié. Si la mémoire de travail d'un apprenant de LE est trop limitée, sa capacité de traiter du texte en LE l'est également (Randall, 2007 : 81). Il n'a alors pas assez d'automatismes dans le domaine de la reconnaissance des mots, ce qui fait qu'il a besoin de se concentrer sur les formes (*bottom up*). Selon Randall, il serait peut-être utile de présenter aux apprenants d'abord un *input* restreint et ensuite un *input* de plus en plus élargi.

> *The solution would seem to be for the second language learner to progressively process larger and larger chunks of text, either written or oral. The focus of attention, the top-down element, will increase in scope as the level of the learner and size of the chunks processed increases* (Randall (2007 : 98).

Randall (2007 : 100) semble défendre le point de vue selon lequel pour les apprenants les moins avancés, le sens doit être construit à partir de petites unités. La quantité d'*input* à laquelle on expose les apprenants de LE doit être adaptée à leur niveau (Randall, 2007 : 98). Les concordances permettent un tel ajustement de la taille du contexte, car on peut choisir le nombre de mots qui doivent se trouver à gauche et à droite d'un mot clé.

2.4.2. Digression : microsyntaxe et macrosyntaxe

Le paragraphe suivant ne traite pas particulièrement des collocations. Il définit les termes de micro- et macrosyntaxe afin de préparer aux informations données dans le paragraphe 2.4.3.

Py (2002) défend l'idée que dans le domaine de l'enseignement/apprentissage des langues, l'institution scolaire se concentre principalement sur le domaine de la microsyntaxe. Il définit ce terme comme un réseau de relations, spécifiées par des prépositions, des terminaisons ou par la position d'éléments dans l'énoncé. Py appelle "clause" un "mouvement discursif" qui se tient grâce un réseau microsyntaxique. Les clauses entrent en relation les unes avec les autres. Les types de relations entre clauses sont des phénomènes de macrosyntaxe. La macrosyntaxe peut être associée au traitement pragmatique et sémantique. Un apprenant qui s'appuie sur la macrosyntaxe fait appel à ses connaissances du monde et à ses connaissances discursives,

donc à son savoir encyclopédique et culturel. S'il ne dispose pas d'assez de ressources microsyntaxiques, il renonce à une spécification des relations entre les clauses qui apparaissent alors de façon juxtaposée. Il peut avoir recours à des expressions préfabriquées (cf. 2.3.), souvent non analysées. Les enseignants de LE ont tendance à considérer un énoncé comme inacceptable quand il comporte une ou plusieurs erreurs de précision, donc au niveau microsyntaxique. Py (2002) propose d'utiliser la macrosyntaxe comme un cadre qui est posé dans le but d'augmenter les ressources microsyntaxiques. La macrosyntaxe serait une "infrastructure préparant l'émergence d'une organisation microsyntaxique" (cf. 4.2.2.). Lors d'une situation exolingue, un locuteur expert reformule parfois un ou des éléments macrosyntaxiques d'un locuteur novice, ce qui peut mener à une SPA (Séquence Potentiellement Acquisitionnelle). L'organisation microsyntaxique peut-elle émerger quand on propose aux apprenants de LE un *input* sous forme de concordances ? Peut-on ainsi recréer une sorte de situation exolingue lors de laquelle la reformulation proposée par un locuteur expert est remplacée par un large choix d'occurrences d'un mot clé présenté avec son co-texte ? Les concordances offrent aux apprenants la possibilité d'observer l'organisation microsyntaxique de la clause contenant un mot clé donné. L'observation de l'organisation macrosyntaxique peut être facilitée par la mise à disposition de corpus réfléchis et par le repérage de collocations pertinentes dans ces corpus.

2.4.3. Collocations et microsyntaxe

Diehl, Christen, Leuenberger, Pelvat et Studer (2000) ont dirigé une expérimentation longitudinale intitulée *DiGS (Deutsch in Genfer Schulen – L'allemand dans les écoles genevoises).* Il s'agit d'un projet qui a été proposé à des apprenants d'allemand LE dans le milieu de l'enseignement secondaire en Suisse romande. L'expérimentation a eu pour but de décrire l'acquisition de l'écrit en milieu guidé. Les prises de position théoriques principales ne correspondent pas toutes aux miennes : la théorie de la grammaire universelle constitue la base de cette expérimentation. Toutefois, d'autres théories d'apprentissage sont évoquées, comme la théorie de l'apprentissage explicite et implicite (Berry et Dienes, 1993). Les productions écrites rédigées lors du projet comportent, entre autres, des segments (appelés *chunks*). Pour certains domaines grammaticaux, comme les prépositions ou les terminaisons de l'adjectif, un fonctionnement par segments mémorisés semble parfois donner de meilleurs résultats que le fonctionnement par règles.

> *[Es] spielt gerade im PP-Bereich [Präpositionalphrasenbereich] das Chunk-Lernen eine besonders wichtige Rolle; und zumindest, was die PfK*

> *[Präpositionen mit festem Kasus] betrifft, verhelfen Chunks den Lernenden offensichtlich dazu, präpositionale Ausdrücke zielsprachenkonform zu verwenden.*
>
> [L'apprentissage à l'aide de chunks joue un rôle particulièrement important dans le domaine des groupes prépositionnels ; et au moins en ce qui concerne les prépositions à cas fixe, les chunks semblent aident les apprenants à utiliser les formules prépositionnelles de façon conforme par rapport aux normes de la langue cible] (Diehl et al. 2000 : *322).*

Les apprenants n'intègrent pas tous de la même façon les collocations dans leur production. Certains semblent les utiliser sans analyse grammaticale. Ils utilisent, par exemple, formule "*Es war einmal*" [Il était une fois] dans le contexte d'un conte, ce qui est conforme du point de vue discursif, mais l'environnement syntaxique ne correspond pas aux normes de ce genre textuel (Diehl *et al.*, 2000 : 342). La capacité d'intégration et de modification de segments dans un nouveau contexte va de pair avec la capacité d'analyser les relations grammaticales qui existent entre les unités des segments.

> *Was erfolgreiche Chunk-Benutzer von erfolglosen unterscheidet, ist offensichtlich die Fähigkeit, die in diesen formelhaften Wendungen enthaltenen grammatischen Beziehungen zu erkennen und sie adäquat in den sprachlichen Kontext einzugliedern. Die Chunks liefern ihnen somit Daten, die sie für ihre Hypothesen über die Regeln der L2 einsetzen können, sobald sie das hierfür erforderliche Sprachwissen erworben haben. So gesehen kann den Chunks im Spracherwerb eine überaus nützliche Funktion zukommen*
>
> [Ce qui distingue les utilisateurs performants de chunks d'utilisateurs non performants est apparemment la capacité de reconnaître les relations grammaticales au sein de formules et de savoir les intégrer dans le contexte linguistique. Les chunks leur fournissent des données qui pourront leur servir pour créer des hypothèses concernant des règles en L2, dès qu'ils auront acquis le savoir linguistique dont ils ont besoin pour ce faire. Vu ainsi, les chunks peuvent avoir une fonction d'une importance cruciale pour l'acquisition d'une langue (Diehl *et al.*, 2000 : 341).

*2.5. Activités basées sur l'*input *en LE*

2.5.1. Recherche d'informations

Rouet, Coutelet et Dinet (2004) précisent que ces dernières années, en raison du nombre croissant de systèmes d'informations qui contiennent des documents complexes, la recherche d'informations est devenue un sujet de recherche à part entière aussi bien en psychologie que dans d'autres disciplines. Elle est aussi devenue un nouveau type d'activité pédagogique – activité dans le sens d'exercice scolaire – dans l'enseignement primaire, secondaire et

supérieur. La recherche d'informations est proposée dans l'espoir qu'elle aiderait à construire activement des connaissances. Rouet, Coutelet et Dinet (2004) relèvent la complexité cette activité. Des questions émergent concernant la façon dont les humains trouvent une information, concernant les facteurs qui conditionnent la rapidité de la recherche et concernant les obstacles qui l'entravent. Exercée à partir de textes, la recherche d'informations implique une activité langagière (donc une activité mentale), au même titre que la lecture, l'écriture et le dialogue. Il convient de bien séparer l'activité de compréhension et l'activité de recherche d'informations. La première activité demande une stratégie de cohérence, susceptible d'aider le lecteur à comprendre le texte dans sa globalité. La deuxième activité demande une stratégie de pertinence qui permet de trouver la combinaison d'informations qui aide le lecteur à résoudre un problème déterminé. Cette stratégie peut conduire à traiter les informations d'un texte à des degrés variables de profondeur, en fonction de l'utilité de l'information lors d'une situation particulière, et non pas en fonction de son importance "globale" au sein du texte (Rouet, Coutelet et Dinet, 2004). Ce mode de traitement que Rouet, Coutelet et Dinet appellent "sélectif et orienté par les objectifs préalables du lecteur", confère un rôle particulier à la gestion d'informations en mémoire de travail et à l'usage des organisateurs textuels. Les processus sélectifs d'activation, d'inhibition et d'inférence que l'on a observés pour l'activité de lecture sont particulièrement actifs pendant la recherche d'informations, car celui qui lit de façon ciblée ne peut se contenter d'une compréhension approximative du sens de l'*input* abordé. Ceci a des conséquences sur la mémoire de travail.

> Pour bien chercher, je dois me rappeler ce que je suis en train de chercher. (…) la recherche d'informations pose le problème du contrôle volontaire de l'activation et de la désactivation d'informations en mémoire de travail (Rouet, Coutelet et Dinet, 2004).

L'utilisation de dispositifs textuels complexes (ouvrages entiers, ressources électroniques, etc.) nécessite un savoir-faire concernant l'interprétation d'éléments de signalement du contenu. Il peut s'agir de marqueurs typographiques, de tables des matières, d'index, de menus, d'en-têtes, etc. L'utilisation de tels éléments peut poser problème à des enfants en dessous de 13 ans et même à certains adultes (Rouet, Coutelet et Dinet, 2004). La stratégie de pertinence qui convient à la recherche d'informations demande la capacité de hiérarchiser les informations en fonction de leur pertinence par rapport à l'objectif de recherche.

De ce qui précède, je conclus que la tâche d'identification de segments dans une concordance est complexe. Il paraît important d'analyser, entre autres, le rôle des indices métatextuels (Coutelet, 2003) qui renseignent les lecteurs sur l'endroit où ils peuvent trouver l'information

dans l'*input* (cf. 6.6.3.). Selon Chanquoy, Tricot et Sweller (2007 : 35), il convient de favoriser "l'habileté à sélectionner une information quelconque dans l'environnement pour en faire une analyse pertinente". Cette habileté semble avoir un lien avec le phénomène d'attention, définie comme "la capacité à sélectionner une information pertinente au milieu d'un flux d'informations différentes" (Chanquoy, Tricot et Sweller, 2007 : 34). Le système cognitif doit arriver à faire un tri parmi ces informations. Il s'agit donc d'aider les apprenants à sélectionner des informations pertinentes lors de la révision de texte à l'aide de concordances.

Schmidt (1990) donne un éclairage supplémentaire sur l'activité de repérage. Il définit le repérage comme "*focal awareness*", donc comme attention consciente et dirigée. Il le considère comme une activité subjective ("*a private experience*") dont on peut rendre compte ("*availability for verbal report*"). Schmidt (1990) propose plusieurs techniques pour observer s'il y a eu repérage ou pas. L'une d'elles est ce que Schmidt appelle "*concurrent verbal report*". Tout de suite après la tâche de repérage, on invite l'apprenant à nommer les éléments qu'il a repérés dans un texte. Selon Jordan (2004 : 218), cette technique ne permet pas de juger si les apprenants sont en train d'effectuer une activité de repérage ou pas : "*It is by no means easy to properly identify when noticing has and has not occurred*". Parler tout de suite après l'exécution de la tâche de ce que l'on a repéré peut permettre de montrer que quelque chose n'a pas été repéré, mais cela ne révèle pas forcément tout ce qui a été repéré. Il convient de construire un ensemble de tâches susceptible de fournir des indices pour le repérage d'informations pertinentes. D'ailleurs, Schmidt (1990) souligne la nécessité de mener des recherches concernant la façon dont les caractéristiques d'une tâche conditionnent la compréhension de l'*input* par les apprenants. Il précise qu'ils n'ont pas toujours la possibilité de repérer par eux-mêmes un élément donné dans l'*input*, mais qu'il est possible de les aider en dirigeant leur attention sur une source d'information particulière et en les invitant à en ignorer d'autres. On peut, par exemple, favoriser davantage le repérage de la forme ou davantage le repérage du sens. Dans la figure ci-dessous, Schmidt (1990) présente plusieurs facteurs qui conditionnent la possibilité de repérer des informations dans l'*input*, à savoir les attentes des apprenants, la "saillance perceptuelle" de l'item et sa fréquence, le niveau d'apprentissage et les besoins liés aux caractéristiques des tâches de production. Les items ont été traduits par moi.

Les attentes :

L'enseignement peut favoriser la création d'attentes qui favorisent à leur tour le repérage.

On ne repère pas les éléments que l'on n'arrive pas à interpréter.

La fréquence :

Plus un item est fréquent dans l'*input*, plus il est probable qu'il soit repéré.

La saillance perceptuelle :

Le contexte immédiat d'un élément peut favoriser ou entraver sa mise en relief.

On risque de le percevoir plus difficilement s'il est assimilé à un autre élément.

Le niveau d'apprentissage :

Disposer de certains automatismes de traitement de l'*input* peut être la condition pour apprendre certains nouveaux faits langagiers.

Les besoins liés aux caractéristiques des tâches de production :

La nature de la tâche influence la façon dont l'*input* est traité par l'apprenant.

Figure 20 – Repérage d'un élément dans l'*input*.

Repérer une information afin d'exécuter une tâche semble être une des conditions les plus fiables pour que la rencontre entre *input* et apprenant puisse avoir comme résultat un apprentissage (*intake*) : "*the information committed to memory is essentially the information that must be heeded in order to carry out a task*" (Ericsson et H. Simon, 1984 : 118, cité dans Schmidt, 1990). Il n'est pas important de savoir si l'apprenant a eu l'intention ou pas de repérer quelque chose. Ce qui compte, c'est que la tâche favorise le traitement de l'*input* par l'apprenant.

Quelles sont les activités observables qui permettent de supposer qu'il y a traitement de l'*input* ? Bruner (1958), relu par Gaonac'h (1991), donne quelques pistes à ce sujet.

2.5.2. Catégorisation

Selon Bruner (1958, cité dans Gaonac'h, 1991 : 89), percevoir un élément dans un environnement linguistique implique une activité de catégorisation. L'apprenant rattache les indices qu'il perçoit à "certaines classes d'objets". La réponse au "stimulus" est donc catégorielle. Un principe économique opère pendant cette catégorisation : le classement dans une catégorie doit être possible à l'aide de peu d'indices ; leur nombre doit toutefois permettre de limiter, le plus possible, le risque d'erreur. "**Percevoir est donc prendre une décision**" (Gaonac'h, 1991 : 89). Afin de pouvoir classer les indices perçus dans des catégories, ces

catégories doivent préexister. Pour Bruner, elles sont, pour la plupart, construites, comme le rappelle Gaonac'h (1991 : 90). Les catégories en place sont une sorte de grille que le sujet applique en abordant son environnement linguistique. L'activité perceptive a un caractère anticipatoire. L'apprenant sélectionne des éléments en fonction de ses hypothèses concernant la nature et les propriétés de l'environnement qu'il observe. Il peut aborder l'*input* en effectuant les gestes suivants.

> Détection de la présence d'un objet ; recherche d'indices pour une exploration large suivie d'une catégorisation provisoire (hypothèse) ; activité exploratoire plus restreinte aboutissant à la confirmation de l'hypothèse (Gaonac'h, 1991 : 90, se référant à Bruner, 1958).

Les attentes et les besoins des sujets déterminent la recherche d'indices et la formation d'hypothèses. Les catégories convoquées peuvent être plus ou moins accessibles. Si l'apprenant a des attentes fortes, il a besoin de peu d'indices pour percevoir l'objet dans l'environnement ; dans ce cas, "la catégorie correspondante est dite très accessible" (Gaonac'h, 1991 : 90). Le nombre et la nature des indices jouent également un rôle pour la perception. Si l'on s'attend à trouver un représentant d'une certaine catégorie, on a tendance à ne percevoir que les indices liés à cette catégorie et à ne pas percevoir ceux qui plaident en faveur d'une autre catégorie. Il y a des choix stratégiques à faire concernant l'activation de certaines hypothèses et l'application de certains modes de vérification (Gaonach'h, 1991 : 91). Bruner explique qu'une théorie de la perception est également une théorie de la signification.

> La signification d'une chose, c'est la localisation d'un objet dans un réseau d'inférences hypothétiques relatives à ses autres propriétés observables, ses effets, etc. (Bruner, 1958 : 6-7, cité dans Gaonac'h, 1991 : 91-92).

2.6. Constructivisme et apprentissage de LE

Legros et Crinon (2002 : 30) définissent le constructivisme comme suit.

> Le constructivisme s'appuie sur l'idée que la réalité du monde se construit dans la tête de l'individu à partir de son activité perceptive sous forme de représentations mentales ou modèles du monde.

Cette définition semble se référer particulièrement à Piaget, l'un des fondateurs du constructivisme. Comme le rappelle P. Martinez (2004 : 16), le constructivisme piagétien accorde une certaine importance au mental pour l'apprentissage de LE, mais contrairement au point de vue chomskyen, il ne postule pas l'existence d'un module cérébral réservé à l'apprentissage des langues. Il rapproche l'apprentissage de LE du domaine des représentations du monde. Le mental a certaines caractéristiques biologiques, et il évolue en fonction du

contact établi avec l'environnement social (cf. 2.2.1.). Le constructivisme combine donc des éléments mentalistes avec des éléments comportementalistes. La perspective socio-cognitiviste de Vygotski est un élargissement de la théorie établie par Piaget. Le terme vygotskien de "schèmes communicatifs" désigne des types de comportement acquis au contact social. Ce terme est basé sur les "schèmes interactifs", des comportements acquis au contact avec l'environnement physique. Il renvoie à l'idée que nous sommes des membres d'une communauté sociale, dotés d'une subjectivité qui se façonne en échangeant avec d'autres personnes. La langue a donc un ancrage social. Apprendre une nouvelle langue donne l'occasion d'élargir ses schèmes de perception (Perrenoud, 1996). L'apprenant est alors en mesure de reconstruire la vision qu'il a de sa langue maternelle (cf. Chini, 2004 ; 2.8.2.) et la vision des autres langues auxquelles il est exposé.

Cobb (2006) définit le constructivisme comme une théorie du domaine de la psychologie qui stipule que la réalité externe existe, mais que les apprenants n'y ont pas directement accès. Ils doivent créer leurs propres représentations de la réalité. Cobb rappelle qu'une des variantes du constructivisme préconise l'accès direct des apprenants aux données pour les aider à réutiliser leurs connaissances dans d'autres situations.

> *Representations constructed from grappling with raw data, as opposed to representations resulting from someone else's having grappled, are not just generally "better" in some vague way but specifically are more successfully transferred to novel contexts and form a better preparation for further independent learning* (Cobb, 2006).

Sur le terrain, l'accès direct aux données linguistiques "brutes" n'a toutefois pas toujours mené à des résultats probants (Cobb, 2006). Il convient de mettre à disposition des apprenants des outils susceptibles de les aider à se repérer dans un environnement linguistique complexe. Le processus de construction de connaissances peut être plus ou moins lent. Certaines attitudes face au travail et des outils adaptés peuvent aider à l'accélérer. Selon Cobb (2006), l'attitude d'"apprenant-linguiste" et l'utilisation de corpus et concordances sont une combinaison prometteuse. Dans ce qui suit, je vais détailler trois termes utilisés par des chercheurs qui s'intéressent à l'attitude des apprenants par rapport à l'*input* en LE. Il s'agit des termes d'apprenant-linguiste, apprenant-détective et apprenant-chercheur.

2.6.1. Attitude d'apprenant-linguiste

Cobb (2006) défend le point de vue que les méthodes de travail des lexicographes peuvent être transférées dans le domaine de l'apprentissage du lexique. S'appuyer sur des corpus et

concordances afin de reconnaître des *patterns lexicaux* peut faire gagner du temps aussi bien à des lexicographes qu'à des apprenants d'une langue étrangère.

> *For example, the word-learning task facing a learner and a lexicographer are essentially the same, and are essentially 'constructivist.' Both must make sense of an overwhelming amount of raw and distributed information in an artificially brief amount of time. Lexicographers use a corpus and concordance to assemble in moments and force the patterns out of data that would otherwise require years, and it is not inconceivable that learners might be able to do some scaffolded version of this too* (Cobb, 2006).

Comme le rappelle Cobb (2006), ses travaux antérieurs (Cobb, 1997 ; Cobb et Horst, 2000) confirment que le concept d'apprenant-linguiste esquissé ci-dessus et l'utilisation de corpus et concordances permettent de transférer des connaissances acquises dans une nouvelle situation.

> *A strong transfer-of-learning effect was found for constructing word-meanings from learning-adapted corpus and concordance tasks, as compared to learning pre-constructed meanings from dictionary definitions* (Cobb, 2006).

Ceci n'est pas sans rappeler un des éléments de la définition que donne Tricot de l'apprentissage humain (2007 : 10 ; cf. 2.2.1.).

2.6.2. Attitude d'apprenant-détective

Johns (1986 ; 1988) estime également que l'observation de concordances favorise l'apprentissage d'une LE. Selon lui, l'utilisation de concordances aide l'apprenant à adopter une attitude de "détective". Il développe ainsi une attitude réflexive par rapport à la langue, et il est en mesure de considérer les données linguistiques comme des matériaux permettant de trouver des indices qui réconfortent ou invalident ses hypothèses de travail concernant la langue (Johns, 2002 : 108). Un "détective" peut se forger son propre point de vue, ceci à condition qu'il ait l'occasion d'observer les données le plus directement possible. Par conséquent, Johns plaide pour le libre accès aux ressources. Le libre accès doit être, à son avis, une des caractéristiques d'un apprentissage basé sur la pédagogie de la découverte (Johns, 1991 : 30). Ainsi, l'enseignant ne fait pas "écran" entre les apprenants et les données. Il est toutefois peu prudent de laisser les apprenants se perdre dans un "océan" de données sans leur donner des repères. Il convient de leur proposer des outils d'aide et un guidage adaptés (cf. 2.7.).

2.6.3. Attitude d'apprenant-chercheur

Bernardini (2004 : 16) évoque dans le contexte du travail sur corpus l'apprentissage par la découverte et le rôle de l'apprenant comme apprenant-chercheur (*learner as a researcher*), présenté déjà dans les travaux rassemblés par Johns (1991) et évoqué aussi par Kennedy et Miceli (2001), St John (2001) et Cheng *et al*. (2003 : 178). Johns explique qu'à l'aide de concordances, les apprenants sont en mesure de trouver des réponses à leurs propres questions.

> *We simply provide the evidence needed to answer the learner's questions, and rely on the learner's intelligence to find answers* (Johns, 1991 : 2).

Johns et King (1991 : iii) utilisent le terme de *Data Driven Learning (DDL)* pour désigner l'apprentissage par la découverte dans le domaine des corpus et concordances.

> *[DDL can be defined as] the use in the classroom of computer-generated concordances to get students to explore regularities of patterning in the target language, and the development of activities and exercises based on concordance output*.

Plusieurs chercheurs (Cross, 2002 ; Hadley, 2002) défendent le point de vue que le repérage d'informations dans l'*input* aide particulièrement à développer le savoir métalinguistique des apprenants. Gaskell et Cobb (2004) montrent la nécessité d'étudier les effets de l'utilisation des corpus et concordances pour l'acquisition de la grammaire. Pour l'apprentissage d'une langue étrangère à l'aide de corpus et concordances, Hadley (2002) propose comme cadre le *DDL* et la "grammaire pédagogique". Il la définit comme suit.

> *I propose that PG is a modest reference which can be used for course work, which principally draws attention to language patterns. It can be used by teachers, native speakers and second language learners to gain insight into and better utilize the target language. This definition suggests that grammar involves "doing" and "autonomous discovery" much more than "being told about." In DDL learners are not seen simply as recipients of knowledge, but as researchers studying the regularity of the language. Teachers help the learners' research without knowing in advance what patterns they will discover. A DDL approach expects learners to get a "feel" for the language by personally experiencing a focused study of the target language's organic consistencies (Chalker 1994, Johns 1991). (...) Data-driven learning is very much a "research-then-theory" method of studying grammar. Language learners start with a question, and then come to their conclusions after analyzing the corpora with a concordancer program* (Hadley, 2002).

Nous retrouvons dans cette définition des principes déjà évoqués ci-dessus, comme la démarche par la découverte et l'attitude d'apprenant-chercheur. Certaines préoccupations du *Data Driven Learning* (*DDL*) et de la grammaire pédagogique semblent se rejoindre dans les

démarches conceptualisantes qui cherchent à amener les apprenants à comprendre le fonctionnement de la langue (cf. 2.8.2.).

2.7. Guidage et apprentissage de LE

2.7.1. Guidage selon le point de vue de Piaget

Selon Piaget (1969 : 205), éduquer veut dire, "adapter l'enfant au milieu social ambiant". Il présente des "méthodes nouvelles" qui cherchent à exploiter pour l'éducation "les tendances propres de l'enfant" et "l'activité spontanée inhérente au développement mental". L'école dite "traditionnelle" aurait tendance à ignorer les ressources propres de l'enfant.

> L'école traditionnelle impose à l'élève son travail : elle le "fait travailler". (…) dans la logique du système, l'activité intellectuelle et morale de l'élève demeure hétéronome parce que liée à la contrainte continue du maître (…). Au contraire, l'école nouvelle fait appel à l'activité réelle, au travail spontané, fondé sur le besoin et l'intérêt personnel (Piaget, 1969 : 205-206).

Selon Dewey et Claparède (cités dans Piaget, 1969 : 206), "toute activité féconde suppose un intérêt". Les enfants peuvent-ils mener un "travail personnel durable" fondé sur le principe de l'intérêt ? Piaget relate d'une expérience faite à l'école expérimentale *Malting House* à Cambridge. Selon le protocole, les adultes devaient laisser les enfants de 3 à 8 ans faire librement des expériences scientifiques avec du matériel très sophistiqué, sans intervenir. Dans cet environnement exceptionnellement favorable, ne serait-ce que du point de vue matériel, les enfants ont en effet appris à observer et raisonner. Toutefois, Piaget (1969 : 229) observe que "quelque systématisation venant de l'adulte n'aurait peut-être pas été entièrement nuisible aux élèves". La poursuite de l'expérimentation aurait probablement fait ressortir le besoin d'un guidage plus présent.

> (…) la nécessité d'une activité rationnelle, déductive, pour donner un sens à l'expérience scientifique et la nécessité, pour constituer cette raison chez l'enfant, d'une structure sociale englobant, non seulement la coopération entre enfants, mais [aussi] la coopération avec l'adulte (Piaget, 1969 : 229-230).

L'enfant s'adapte peu à peu à son milieu physique et social. Il doit acquérir un équilibre entre assimilation et accommodation. Un organisme biologique est adapté s'il a réussi à "conserver sa structure en lui assimilant les aliments tirés du milieu extérieur" et en même temps à "accommoder cette structure aux diverses particularités de ce milieu" (Piaget, 1969 : 208).

> De même, on peut dire que la pensée est adaptée à une réalité particulière lorsqu'elle a réussi à assimiler à ses propres cadres cette réalité, tout en accommodant ceux-là aux circonstances nouvelles présentées par celle-ci (Piaget, 1969 : 208).

Il faut donc "une interaction entre le sujet et l'objet, telle que le premier puisse s'incorporer le second tout en tenant compte de ses particularités". Une "activité structurante continue" doit aider l'enfant à trouver un tel équilibre. L'adaptation sensorielle au milieu ne peut réussir que si l'enfant arrive à assimiler les propriétés du milieu à son activité propre et à les incorporer (Piaget, 1969 : 209). L'école traditionnelle et l'école nouvelle s'opposent particulièrement au sujet des relations sociales de l'enfant. "L'école traditionnelle ne connaît guère qu'un type de relations sociales : l'action du maître sur l'élève" (Piaget, 1969 : 235). Obliger l'enfant à accepter tout ce que dit le maître, sans y réfléchir, le maintient dans un état d'égocentrisme. Sa croyance individuelle est remplacée par la croyance fondée sur l'autorité. L'école nouvelle favorise la coopération des enfants entre eux (Piaget, 1969 : 242). Quand les enfants jouent ensemble, ils trouvent par eux-mêmes des règles, et ils les respectent (Piaget, 1969 : 244). Les écoles nouvelles tentent de s'appuyer sur ce processus. La coopération entre enfants favorise "l'échange réel de la pensée et de la discussion, c'est-à-dire toutes les conduites susceptibles d'éduquer l'esprit critique, l'objectivité et la réflexion discursive".

Au lieu de se soumettre moralement, les enfants intègrent ainsi des principes de conduite.

> Bien entendu, sauf quelques cas extrêmes, les méthodes nouvelles d'éducation ne tendent pas à éliminer l'action sociale du maître, mais à concilier avec le respect de l'adulte la coopération entre enfants et à réduire, dans la mesure du possible, la contrainte de ce dernier pour la transformer en coopération supérieure (Piaget, 1969 : 245).

À mon avis, les travaux de Piaget de 1969, dans l'édition de 1993, contribuent à trouver des éléments pour un compromis entre le "tout linguistique" et le "tout culturel" (cf. 2.). Ils paraissent étonnamment modernes. Les écrits de Piaget ont actuellement des répercussions dans le monde de l'enseignement des langues, surtout en ce qui concerne la collaboration entre apprenants et la remise en cause du modèle traditionnel de relation entre enseignants et apprenants. J'y retrouve des idées exprimées également par Perrenoud (1994). L'école nouvelle chemine entre configuration mentale héréditaire et activités d'éducation qui tentent de soutenir l'évolution des enfants dans un sens qui accélère leur "maturation naturelle". Pour Piaget, le moteur de l'apprentissage est l'intérêt de l'apprenant pour l'objet d'apprentissage. Il décrit le processus qui aide l'enfant à intégrer les objets de la réalité extérieure sans pour autant renoncer aux traits spécifiques de sa personnalité. Ainsi, il paraît possible de s'adapter à

un nouvel environnement sans s'y perdre. Dans ce qui suit, je ferai appel à Bruner afin de préciser davantage en quoi peut consister le guidage.

2.7.2. Guidage selon le point de vue de Bruner

L'analyse des données de ma recherche-action au chapitre 6 se réfèrera, entre autres, à plusieurs points détaillés dans l'article de Bruner (1983 : 261-280[3]) sur le guidage. C'est pourquoi j'en relèverai ci-dessous des parties substantielles.

Bruner utilise le terme de "tutelle" pour désigner le guidage. Il définit la tutelle comme le déploiement des "moyens grâce auxquels un adulte ou 'un spécialiste' vient en aide à quelqu'un qui est moins adulte ou spécialiste que lui" (1983 : 261). Bruner (1983) étudie les interactions entre une tutrice[4] et des apprenants lors d'une tâche[5] particulière, la construction, par des enfants de trois, quatre et cinq ans, d'un objet en trois dimensions "exigeant un savoir-faire qui les dépasse initialement". Dans ce type de situations, la tutrice "connaît la réponse", et les enfants ne la connaissent pas. Bruner (1983) conçoit le savoir-faire chez l'enfant comme suit.

> (…) un programme hiérarchique dans lequel des savoir-faire constitutifs se constituent en "savoir-faire supérieurs" selon un emboîtement approprié pour faire face à l'exigence d'une tâche nouvelle plus complexe (Bruner, 1973). Le processus est analogue à la résolution de problème dans laquelle la maîtrise des problèmes "d'ordre inférieur" ou constitutifs est une condition sine qua non de réussite dans un problème plus important, chaque niveau affectant l'autre (…) la tâche décisive est souvent celle de combiner (…) l'ensemble des actes constitutifs nécessaires pour aboutir à un résultat déterminé. L'enfant, confronté à de nouveaux défis, doit faire correspondre des moyens à des fins et c'est ce couplage (et la correction des erreurs de ce couplage) qui sont au cœur de la résolution de problème (Bruner, 1983 : 262).

La tutrice peut être simplement présente et exécuter la tâche sous les yeux de l'apprenant. Ce procédé de type "modèle et imitation" peut être secondé ou dépassé par d'autres procédés plus complexes. La tutrice peut étayer l'apprentissage de multiples façons afin d'aider l'apprenant à mener à bien des tâches ou à résoudre des problèmes qui auraient été au-delà de ses possibilités sans le soutien d'un tiers.

---

[3] Les pages se réfèrent à l'édition de 2002.

[4] Par souci de simplicité, j'utiliserai dans ce sous-chapitre de façon privilégiée le terme de "tutrice" au lieu de celui de "tuteur", parfois même dans des situations qui dépassent le cadre de l'étude de Bruner et qui pourraient justifier le genre masculin.

[5] Bruner utilise le terme "tâche" dans le sens d'activité et non dans le sens de "tâche sociale" (cf. 2.8.1.2.).

> [La tutrice prend en main] les éléments de la tâche qui excèdent initialement les capacités du débutant, lui permettant ainsi de concentrer ses efforts sur les seuls éléments qui demeurent en son domaine de compétence et de les mener à terme (Bruner, 1983 : 263).

Une telle assistance n'aide pas seulement l'apprenant à exécuter la tâche en question ; il en tirera des bénéfices plus larges, comme l'exécution de ce type de tâches à un rythme qui dépasse celui qu'il aurait atteint par ses propres efforts. La condition pour la réussite d'un tel étayage est la suivante.

> La compréhension de la solution doit précéder sa production. C'est-à-dire que l'apprenti doit être capable de reconnaître une solution d'une classe déterminée de problèmes avant d'être capable lui-même de produire les démarches qui y conduisent sans aide (Bruner, 1983 : 263).

Les enfants sont capables de choisir parmi les "bonnes questions", les "bonnes stratégies" avant d'être capable d'en produire.

> [Les enfants acquièrent] un sens des résultats possibles aussi bien que des moyens pour atteindre ceux-ci par le jeu de quelque chose qui, en surface, ressemble à un comportement d'essai plutôt "aveugle" (sans qu'on puisse le dire aléatoire cependant) (Bruner, 1983 : 264).

La tutrice que Bruner a observée a laissé le plus possible l'enfant travailler, et elle est intervenue seulement quand il semblait en avoir besoin.

> Quand c'était possible, enfin, elle [la tutrice], laissait l'enfant agir à sa guise. C'était seulement s'il s'arrêtait de construire ou s'il s'embarrassait dans des difficultés que la tutrice intervenait. Son dessein était de laisser autant que possible l'enfant régler tout seul le rythme de son travail (Bruner, 1983 : 268).

Les interventions de la tutrice ont été plus nombreuses et plus variées avec les enfants les plus jeunes qu'avec les plus âgés. Ses formes d'intervention ont été les suivantes. Elle est intervenue directement, ou elle a fait repérer des erreurs par l'apprenant en lui posant des questions, comme par exemple celle-ci : "Est-ce que cela (…) ressemble à cela (…) ?". La tutrice a parfois proposé à l'enfant l'exécution d'une action précise : "Peux-tu en faire d'autres [des constructions] comme celle-ci ?" (Bruner, 1983 : 269).

Bruner (1983 : 267) conseille de concevoir la tâche en ayant à l'esprit les objectifs suivants.

- La tâche doit être "excitante et provocante" pour l'enfant.
- La tâche doit être suffisamment complexe pour permettre le développement de l'enfant et la modification de sa conduite.

- La tâche doit être "à caractéristiques multiples" : on doit pouvoir prendre en compte plusieurs de ses composants.
- La tâche doit contenir une structure sous-jacente répétitive afin que l'expérience acquise puisse être mise à profit dans d'autres situations.
- La tâche ne doit pas être totalement au-delà des possibilités de tous les enfants.

À partir de ses observations des interactions entre tutrice et apprenants, Bruner (1983 : 278-279) établit six fonctions d'étayage du tuteur, à savoir l'enrôlement de l'apprenant dans la tâche, la réduction des degrés de liberté (réduire les possibilités de mouvement, jusqu'à ce que l'apprenant reconnaisse si ce qu'il a fait colle aux exigences de la tâche ou non), le maintien de l'orientation, la signalisation des caractéristiques déterminantes (faire comprendre les écarts entre ce que l'apprenant fait et ce qui est attendu), le contrôle de la frustration et la démonstration. L'étude présentée par Bruner (1983) fait ressortir l'importance de l'enrôlement des apprenants dans la tâche. Le lien affectif entre tutrice et apprenants est également un moteur d'action non négligeable, comme le montre l'observation suivante.

> La tutrice (…) a créé un climat d'approbation tel que les enfants paraissaient impatients de réussir leurs constructions – souvent, semble-t-il, pour les lui montrer autant que pour atteindre le but per se. Une procédure d'épreuve et un tuteur créent une atmosphère soit de stimulation soit de découragement (Bruner, 1983 : 268).

2.7.3. Guidage et TIC

Le principe de cohérence qui ne signifie pas l'uniformité.
Ne pas user de guidages contradictoires.
Résister à la tentation de faire cours.
Éviter les arguments d'autorité non nécessaires, sans oublier son devoir d'ingérence pédagogique et la nécessité d'intervenir et de guider pour que certaines ressources soient pleinement exploitées.
Savoir gérer le temps et les priorités.
Savoir "les laisser seuls".
Ne pas demander à l'apprenant ce que peut seul faire le formateur.

Tableau 8 – Éléments pour un précis du bon usage du guidage.

Demaizière (2004) évoque des principes de guidage applicables à un travail pédagogique basé sur les TIC (Technologies de l'Information et de la Communication). Ces principes se recoupent avec certaines des fonctions du tuteur évoquées en 2.7.2., et ils en précisent

d'autres, comme, par exemple, la cohérence de ce que l'on propose aux apprenants et la "tentation de faire cours".

2.8. Approches pédagogiques

2.8.1. Approches privilégiant le sens

2.8.1.1. Pédagogie du projet

Depuis la publication d'ouvrages de Freinet et d'autres représentants de la pédagogie active, la notion de projet est connue et exploitée en milieu scolaire, particulièrement à l'école maternelle et à l'école primaire. La pédagogie Freinet contient des éléments constructivistes piagétiens, mais aussi certains éléments que Minuth et Schlemminger (2000) considèrent comme behavioristes, à savoir l'apprentissage individualisé à l'aide de fiches d'autocorrection. Minuth et Schlemminger se distancient de ces éléments. Ils mettent en avant d'autres aspects, comme l'apprentissage par exploration. Ils précisent qu'un guidage adapté doit accompagner l'apprentissage en institution. Tout en étant capables d'apprendre par eux-mêmes, les apprenants ont besoin d'un échange social qui soutient leur travail d'apprentissage. Les adeptes de la pédagogie Freinet s'intéressent principalement à l'échange qui a lieu autour du produit, créé de façon collaborative. Cet échange a lieu aussi bien à l'intérieur qu'à l'extérieur de la communauté scolaire qui crée ce produit. Le processus d'apprentissage se caractérise par les actions suivantes des apprenants : formulation de questions, prise de conscience ("*Erkennen*"), apprentissage (Minuth et Schlemminger, 2000). Il convient de partir des questions des enfants, de leur donner l'occasion de faire des essais et d'avancer "en tâtonnant", de vivre succès et échecs, de s'auto-évaluer, de réfléchir. Il y a des lieux (journaux, cercles de discussion, exposés, etc.) où les apprenants peuvent exprimer la perception qu'ils ont d'eux-mêmes et où ils peuvent évaluer comment les autres les perçoivent. Pour compléter cette liste établie par Minuth et Schlemminger, je propose l'utilisation de blogues pédagogiques comme lieux de publication et d'échange. Mangenot (1998) présente également comme une des caractéristiques les plus intéressantes du projet la possibilité de "faire sortir des murs de l'institution, [à] mettre (…) en relation les apprenants entre eux (ce que la tâche fait également), mais en outre à amener à communiquer avec l'extérieur en vue de la réalisation d'un but bien précis".

Une description de l'approche par tâches précisera les apports possibles d'un projet en classe de langue.

2.8.1.2. Approche par tâches

L'approche par tâches est une méthodologie qui cherche à motiver le travail par une production finale. Elle nécessite l'identification d'objectifs et de besoins et la mise en place d'étapes de travail susceptibles de permettre d'atteindre un objectif donné. Springer (2009 : 25) évoque le terme de "tâche sociale". Il se réfère ici, à mon avis, à ce que les auteurs du *CECR (Cadre Européen Commun de Référence pour les langues*, 2001 : 121) appellent "tâche proche de la vie réelle". Je définis le terme de "tâche sociale" comme suit : quand l'exécution d'une tâche vise à faire produire quelque chose qui a des caractéristiques de produits que l'on trouve aussi en dehors du cadre scolaire, il s'agit d'une "tâche sociale". Elle se distingue des tâches que les auteurs du *CECR* (2001 : 121) appellent "tâches pédagogiques communicatives" et qui tiennent compte des besoins immédiats des apprenants en situation d'apprentissage de LE.

En France, depuis la prise en compte du *CECR* (2001), l'approche par tâches a fait son entrée dans l'enseignement-apprentissage des langues dans l'enseignement primaire et secondaire. En ce qui concerne l'allemand au niveau collège, le terme de "tâche" figure dans le *Projet de programme de l'enseignement d'allemand* de mars 2005, dans la *partie Allemand palier 1* (2005 : 1). Certains manuels d'allemand pour le collège, comme, par exemple, *Spontan* (Halberstadt *et al*., 2008) et *Gute Fahrt* (Bernardy *et al*., 2009) se réfèrent explicitement à cette méthodologie. "Comment l'approche par tâches permet de tout intégrer ou presque" ; ce titre d'une communication de Nicol-Benoît (2004b) montre que l'approche par tâches est une méthodologie bien "accueillante". Il convient de définir ce que l'on souhaite y intégrer. R. Ellis présente dans l'ouvrage *Taskbased Language Learning and Teaching* (2003 : 4-5) plusieurs définitions de "tâche", allant de conceptions selon lesquelles l'échange verbal n'est pas une condition nécessaire pour la réalisation d'une tâche jusqu'à des conceptions selon lesquelles la négociation de sens en interaction verbale joue un rôle important. Certaines définitions privilégient des résultats (*outcomes*) plus concrets que d'autres qui estiment qu'une tâche peut aussi mener à un résultat immatériel (un savoir-faire, une prise de décision commune). D'autres définitions montrent que c'est le processus (la façon dont procèdent les apprenants pour réaliser la tâche) et non le résultat lui-même qui est central. R. Ellis définit le terme de "tâche" comme suit.

> Une tâche est un plan de travail ("*workplan*") qui cherche à obtenir la procéduralisation pragmatique du langage par les apprenants et qui a pour but un résultat qui permet d'évaluer si le contenu des propos a été rendu de

façon appropriée. Les apprenants doivent focaliser leur attention avant tout sur le sens des productions et ils puisent librement dans leurs propres ressources linguistiques, sauf si la tâche les oblige à utiliser une forme particulière. Une tâche cherche à faire utiliser un langage qui ressemble de près ou de loin au langage utilisé en dehors de la salle de classe ("*in the real world*"). Comme d'autres activités linguistiques, une tâche peut inclure une ou plusieurs des quatre compétences et elle déclenche également des processus cognitifs (R. Ellis, 2003 : 16 ; ma traduction).

R. Ellis (2003 : 21) conseille de tenir compte des facteurs suivants pour construire une tâche.

1 Le but	L'objectif général d'une tâche (par exemple, s'entraîner à décrire des objets de façon précise).
2 *Input*	L'information verbale ou non verbale nécessitée par la tâche (un texte écrit, des images, etc.).
3 Conditions	La façon dont on présente l'information (morcelée ou entière) et dont on la fait utiliser (trouver les similitudes ou les différences, etc.).
4 Choix méthodologiques	L'organisation du groupe (travail à deux, en petits groupes, etc.), le temps (non) limité, etc.
5 Résultats prévus, le produit	Le produit qu'on obtient après avoir effectué la tâche (une carte complétée, un tableau rempli, un nouveau savoir-faire, etc.). La tâche peut être "ouverte" (permettre plusieurs solutions ou réalisations) ou "fermée" (ne permettre qu'une seule solution).
6 Processus	Le processus linguistique et cognitif que la tâche est censée créer.

Tableau 9 - Caractéristiques d'une tâche.

Perrenoud (1994 : 109) présente les tâches dans le contexte des "nouvelles didactiques". Les nouvelles didactiques apparaissent comme une alternative à la didactique dite "traditionnelle". Elles ont un héritage dans les "pédagogies alternatives, actives, coopératives, institutionnelles ou modernes" (Perrenoud, 1994 : 109). Dans le tableau suivant, Perrenoud (1994 : 110) montre les caractéristiques des tâches nouvelles en les juxtaposant aux tâches "traditionnelles".

Tâches traditionnelles	**Tâches nouvelles**
Les tâches sont identiques.	Les élèves ne font pas tous toujours la même chose en même temps.
Elles sont fermées (permettant une procédure unique pour les réaliser).	Tâches ouvertes (réponses vraiment inconnues avant ; démarches au moins partiellement personnalisées).
Elles sont fragmentées (tâches élémentaires sans contexte).	Tâches plus globales.

Elles sont standardisées (consignes, type de données, genre d'opérations et démarches qui sont familiers aux élèves).	Tâches moins stéréotypées.
Elles sont souvent à faire à l'écrit.	Elles font fréquemment appel à l'oral.
Elles sont à faire individuellement, mais sans être individualisées.	Elles sont souvent à faire à plusieurs.
Elles sont quantifiables.	Les rendements sont difficiles à comparer.
Elles sont courtes et apparaissent en alternance rapide.	Tâches de longue haleine dans lesquelles toute la classe s'engage.
Elles sont relativement faciles.	Tâches qui paraissent utiles et ayant un intérêt pour les élèves.
Elles sont assorties de consignes qui sont peu interactives.	Tâches qui se définissent en concertation avec les élèves.

Tableau 10 – Tâches traditionnelles, tâches nouvelles (Perrenoud).

Devant les nouvelles tâches, les apprenants ont des choix plus ouverts à faire. Ce genre de tâches entraîne parfois les apprenants dans un "tourbillon de projets et de possibilités" ce qui fait qu'elles peuvent être ressenties comme "complexes et pour certains angoissantes" (Perrenoud, 1994 : 111). Perrenoud (1994 : 113) explique qu'instaurer la nouvelle didactique peut avoir comme conséquence un affaiblissement du rôle de l'enseignant et un élargissement du champ d'action des apprenants. Afin de pouvoir faire face à de telles modifications relationnelles, l'enseignant doit tenir compte des nouvelles stratégies que développent les apprenants face aux nouvelles tâches (Perrenoud, 1994 : 113).

L'approche par tâches est une méthodologie qui permet à l'apprenant de s'impliquer dans le processus d'apprentissage. Elle favorise une démarche lors de laquelle les apprenants se concentrent sur le sens davantage que sur la forme quand ils échangent entre pairs ou avec d'autres personnes associées à la formation. L'approche par tâches renverse l'ordre habituel des priorités : c'est la tâche sociale, au lieu d'un programme grammatical, qui détermine le déroulement des activités d'apprentissage. Le travail sur les formes n'est ainsi pas un point de départ, mais un outil de travail parmi d'autres. Il n'est pas occulté, mais il est motivé par son lien avec la tâche sociale.

2.8.2. Approche combinant sens et forme – la conceptualisation

Culioli (1990 : 17) souligne la nécessité d'une théorie de l'application des opérations énonciatives (cf. 2.1.2.) pour l'apprentissage de LE. La méthodologie de la conceptualisation répond à cette demande. Elle a été mise en pratique par l'équipe Charlirelle dans les années 70. La méthode d'anglais *Behind the Words* (1977, 1978), développée par ce collectif d'auteures, a été conçue selon les principes de la conceptualisation. La deuxième de couverture du livret de présentation du manuel pour la classe de sixième (1977) explique que cette méthode cherche à faire comprendre aux enfants ce qu'ils font. Les apprenants sont censés "saisir un nouveau système de langage et sa façon de fonctionner en rapport avec les situations qu'on vit et qui amènent à parler". Bailly, la fondatrice du groupe Charlirelle, définit le terme de conceptualisation comme suit.

> Par convention, au sens restrictif du terme : recours à des procédures pédagogiques d'ordre métalinguistique (analyse verbalisée auprès des élèves de l'objet-langue et de son statut discursif/pragmatique). Le but de cette procédure est de faire prendre conscience aux apprenants du fonctionnement langagier, en rapportant de manière explicative la spécificité des marqueurs de la L2 à des valeurs invariantes (repérages cognitifs et psycholinguistiques) (Bailly, 1998 : 49).

La conceptualisation sollicite la "participation très active des élèves" ; elle prévoit de les "faire réfléchir sur la langue (principalement la grammaire) de façon simplifiée, mais complète" et de leur faire prendre conscience des rapports entre formes et valeurs (Bailly 1998 : 49). Cette réflexion est nécessairement reliée à la pratique de la langue, et le travail doit être contextualisé. Chini (2004 : 65) pose également la question de la transposition didactique des concepts de la théorie des opérations énonciatives, développée par Culioli. Selon Chini (2004 : 68), il serait conforme à cette théorie que

> le guidage enseignant crée les conditions d'une véritable approche métalinguistique où l'apprenant est conduit à élaborer le système de la langue étrangère à partir d'observations ancrées dans sa pratique en contexte.

Il convient alors de proposer aux apprenants des activités qui font sens pour eux, par exemple la formulation individuelle de références. Pour être en mesure de faire de telles activités, ils ont besoin de connaître quelques opérations énonciatives de base. Chini (2004 : 80) explique comme suit ce qui peut entraver la mise en pratique d'activités conceptualisantes.

> (…) la difficulté qu'il y a non seulement à sortir du schéma traditionnel de transmission d'un savoir prédéfini pour laisser à l'élève un espace de liberté

où construire du sens, mais aussi à adapter les concepts et discours de la théorie pour en faire des outils utiles à l'apprentissage.

Chini conseille de prendre comme point de départ les références personnelles, donc les connaissances antérieures des apprenants. Il convient de considérer la LE non comme un objet d'apprentissage, mais comme un "outil cognitif indispensable à tout processus de raisonnement et de structuration de la pensée" (Chini, 2004 : 87). Pour Chini, apprendre une langue étrangère est lié à une "redéfinition du rapport au langage". Écrire en langue étrangère permet de prendre doublement une distance ce qui soutient la réflexion sur le langage. Pour que les références dont disposent les apprenants en lien avec leur langue maternelle ne fassent pas obstacle, il faut les démonter, les problématiser, afin d'amener l'élève à percevoir leur spécificité, et donc leur relativité. Ce n'est qu'alors qu'il pourra accepter la possibilité d'un autre parcours (Chini, 2004 : 94).

2.9. Pour résumer

Hulstijn (2007) évoque des facteurs externes qui peuvent entrer en jeu lors de l'apprentissage d'une langue étrangère. Ces facteurs reprennent plusieurs des sujets traités dans le chapitre 2.

- La nature et la quantité de l'*input*.
- L'environnement social (*social setting*).
- Les facteurs qui dépendent de l'apprenant (*learner variables*), comme sa motivation pour apprendre la LE, son attitude par rapport à la LE et par rapport aux locuteurs natifs de la LE et du pays de la LE.
- L'aptitude de l'apprenant : sa capacité à traiter de l'information (liée à sa capacité de mémoire de travail), ses connaissances métalinguistiques (dont la littératie) et ses stratégies de résolution de problème.

Les points évoqués par Hulstijn peuvent être reliés à mon projet de recherche. Je propose aux apprenants un *input* dont la nature est en accord avec la nature des tâches d'apprentissage. Il s'agit d'un *enhanced input* : certaines données sont matériellement mises en forme à l'aide d'une présentation sous forme de concordances. Celles-ci permettent de guider les apprenants vers au moins une partie de ce qu'ils sont censés repérer. Les concordances donnent aux apprenants l'occasion d'observer les mots clés dans de nombreux co-textes. Une des questions que je poursuivrai lors de l'analyse des données sera liée à l'identification de collocations ou, en utilisant un terme plus neutre utilisé en statistique textuelle, l'identification de segments. Le public avec lequel j'ai travaillé a été composé d'adolescents ayant entre 14 et 15 ans. Lors du

projet de production écrite, les apprenants ont-ils repéré et repris des segments pertinents ? Les ont-ils modifiés selon le nouveau contexte ? Les ont-ils analysés ? Ont-ils utilisé trop ou pas assez de segments repérés dans l'*input* par rapport au reste de la quantité de texte qu'ils ont produite eux-mêmes ? Siepmann (2008 : 194) rappelle plusieurs recherches qui confirment les tendances suivantes concernant la réutilisation d'éléments venant d'un *input* en langue étrangère : les apprenants semblent privilégier la réutilisation de lexèmes simples. Ils utilisent quelques rares types de collocations, et ils les utilisent à haute fréquence. Ils ont tendance à surgénéraliser à partir de collocations qu'ils connaissent déjà.

> *Learners use fewer phrases more often; they have lexical 'teddy bears' for particular pragmatic functions which they use much more frequently than other phrases of similar frequence in native discourse; many of these are found in textbooks (in my opinion; for example, etc.) ; they have disproportionate preference for one-word markers such as however over multi-word markers fulfilling similar pragmatic functions such as that said; there are collocational error that can be traced back to overgeneralization in the L2 (...) or to L1 influence* (Siepmann, 2008 : 194).

Ma recherche vise, entre autres, à confirmer ou infirmer ces résultats de recherches présentés par Siepmann. La capacité des apprenants à traiter des informations et à résoudre des problèmes peut influencer la façon dont ils utilisent des corpus et concordances. La manière dont ils abordent et, éventuellement, résolvent des problèmes, peut être illustrée par les données de la recherche-action. L'observation de la capacité des apprenants à traiter des informations nécessiterait un protocole de recherche expérimental ; ce sujet dépasse donc le cadre de mes recherches. Mes données permettent de rendre compte du paramètre "littératie", plus précisément de la littératie électronique, définie comme capacité à s'orienter dans un environnement électronique et à utiliser les TIC (Technologies de l'Information et de la Communication) à des fins d'apprentissage de LE. J'examinerai les attitudes des apprenants par rapport au scénario et du guidage afin d'évaluer dans quelle mesure il y a eu apprentissage lors de l'exécution de mon projet de recherche.

Après avoir présenté, d'un point de vue général, dans le chapitre 1 le domaine des corpus et dans le chapitre 2 l'apprentissage d'une langue étrangère, je proposerai dans ce qui suit un chapitre sur les recherches faites dans le domaine de la production écrite. Dans le chapitre 4, je reprendrai le thème des corpus et concordances, en me concentrant sur ses aspects didactiques et pédagogiques.

3. Production écrite en langue étrangère

Comment apprend-on à écrire en langue étrangère ? Des modèles de production écrite existants, peuvent-ils donner des pistes d'observation de l'apprentissage de l'écriture, ou convient-il de trouver d'autres points d'entrée ? Qu'est-ce qui fait qu'un apprenant s'implique dans le travail d'écriture en langue étrangère, qu'il arrive à réviser son texte et à obtenir un résultat conforme aux normes qui d'un genre textuel donné ? Pourquoi certains scripteurs ont-ils massivement recours à un traducteur électronique ? Quel rôle jouent les interactions entre membres du groupe-classe et entre enseignant(s) et apprenant(s) pour l'exécution d'une tâche de production écrite ? L'étendue de ces questions montre que l'enseignement/apprentissage de la production écrite en langue étrangère est un objet d'étude très vaste. Ce domaine est en même temps assez peu étudié, en particulier en ce qui concerne le public d'apprenants de l'enseignement secondaire (Krings, 1992 ; Leki, Cumming et Silva, 2008). Leki, Cumming et Silva (2008 : 26) présentent une synthèse des recherches sur l'enseignement/apprentissage de la production écrite en anglais, langue seconde, dans le contexte de la *highschool* en Amérique du Nord.[6] Ils expliquent la quantité faible d'études par la grande complexité de ce contexte. Dans le cadre d'une institution scolaire, l'activité d'écriture peut être ressentie comme contraignante, surtout si les apprenants ont une obligation de présence. L'apprentissage de la production écrite en langue étrangère est alors parfois la scène de questionnements, de conflits et d'émotions diverses, souvent difficiles à identifier et à gérer par les enseignants. Selon Barbier (2003), "l'absence dans la littérature de modélisations spécifiques de la production écrite de textes en L2" se justifie par les parallélismes que de nombreux chercheurs ont observés dans le domaine du processus de production écrite en langue maternelle et en langue étrangère. Grabe et Kaplan (1996 : 1) soulignent, en effet, que, dans de nombreux sous-domaines, les recherches concernant les deux domaines se superposent. Ils trouvent néanmoins nécessaire de différencier entre production écrite en langue maternelle et en langue étrangère. À défaut de pouvoir me référer exclusivement à des spécialistes de la production écrite en langue étrangère dans le milieu du secondaire, je choisis, en toute prudence, l'apprentissage de la production écrite en langue maternelle, en particulier à l'école primaire et au collège, comme un de mes domaines de référence. Je proposerai des ajustements là où les besoins spécifiques de la langue étrangère le rendent nécessaire.

[6] Le niveau de la *highschool* correspond au niveau de l'enseignement secondaire en France.

3.1. Modélisation de la production écrite en langue maternelle

Vygotski (1934 : 343) présente le "langage écrit" comme "la forme la plus difficile et la plus complexe de l'activité verbale". Afin de réduire la complexité du domaine de la production écrite, je propose de regarder dans ce qui suit un choix de modèles de production écrite susceptibles de donner des pistes pour l'enseignement/apprentissage. Plane (2004) cite trois types de modèles que je distingue ici selon la centration sur le produit et/ou le processus d'écriture.

Modèle de la rhétorique classique	Écrire veut dire "reproduire un modèle", ceci selon un procédé linéaire.	Centration sur le produit.
Modèle de la psycholinguistique	Écrire est un processus récursif.	Centration sur le processus d'écriture.
Modèle de la critique génétique	Les versions successives des textes d'un scripteur ont une valeur heuristique.	Centration sur le processus d'écriture et sur les diverses variantes du produit.

Tableau 11 - Produit et processus dans des modèles d'écriture.

Je détaillerai ci-dessous plus particulièrement certains modèles de la psycholinguistique auxquels de nombreux spécialistes de la production écrite se réfèrent depuis une trentaine d'années. Dans la partie 3.2.3., consacrée à la répercussion des modèles sur mon scénario, je reviendrai sur le modèle de la rhétorique classique et sur le modèle de la critique génétique.

Dans le domaine de la psycholinguistique, il existe de nombreuses recherches concernant la lecture et la production écrite en langue maternelle, dans une moindre mesure concernant la langue étrangère (Barbier, 2003). Regardons de plus près quelques modèles qui y sont proposés, en gardant à l'esprit les questions suivantes.

- Quelle est la nature de l'*input* proposé aux apprenants, et quel rôle lui accorde-t-on ?
- Comment les scripteurs planifient et révisent-ils leur texte ?
- Quelles sont les aides dont dispose le scripteur ?

3.1.1. Modèle de Hayes et Flower (1980)

En psycholinguistique, le modèle d'écriture de référence est celui de Hayes et Flower (1980). Ce modèle continue à constituer la toile de fond pour de nombreux modèles élaborés ultérieurement (Hayes, 1996 ; Barbier, 2003 ; Olive et Piolat, 2003). Comme le précisent

Marin et Legros (2008 : 98), ce modèle a été élaboré à partir de protocoles verbaux, faits par des scripteurs experts dont il tente de décrire le fonctionnement cognitif. Hayes et Flower (1980) inscrivent la production écrite dans un environnement de tâche ("*task environment*"). Dans ce modèle, l'écriture est considérée comme un processus récursif, donc non linéaire. Le modèle prend en compte la mémoire à long terme, les connaissances du scripteur concernant le sujet qu'il traite, sa connaissance du public, ses connaissances concernant des plans de rédaction, stockés dans la mémoire à long terme ("*stored writing plans*"), et des processus comme la planification, la formulation, la révision et l'exécution. Ces processus ou phases de rédaction sont coordonnées par une instance que Hayes et Flower (1980) appellent "*monitor*".

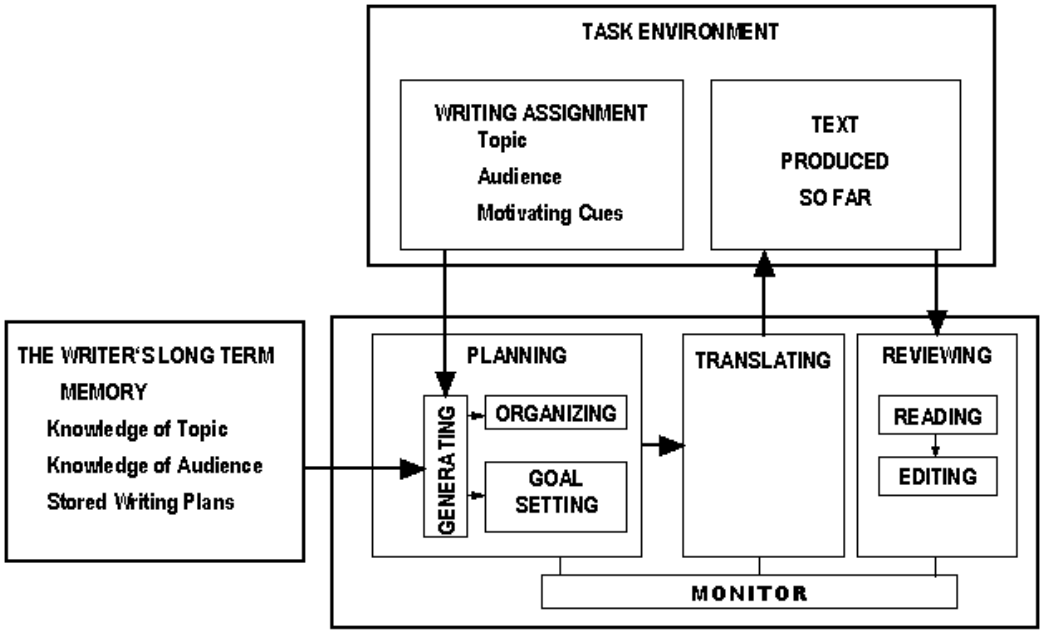


Figure 21 - Structure du modèle d'écriture de Hayes et Flower.

Dans la lignée de Hayes et Flower (1980), Alamargot, Lambert et Chanquoy (2005) décrivent les processus rédactionnels comme suit.

Planification

Cette phase mobilise les connaissances du domaine et les connaissances pragmatiques. Elle doit permettre de récupérer, élaborer et organiser les contenus.

Formulation

Cette phase doit permettre d'élaborer le message linguistique en articulant les niveaux sémantique et linguistique. Le scripteur y effectue le traitement orthographique, lexical et grammatical.

Révision

Cette phase est la phase d'évaluation du texte. Le scripteur est censé l'améliorer, le réduire ou l'étoffer. Il doit le comparer au produit initialement prévu.

Exécution

Lors de cette phase a lieu la réalisation physique de la trace écrite.

3.1.2. Modèle de Bereiter et Scardamalia (1983 ; 1987)

Marin et Legros (2008 : 104-22) décrivent deux contributions de Bereiter et Scardamalia, celle concernant la modélisation de la révision et celle concernant la prise en compte non seulement des rédacteurs experts, mais aussi des rédacteurs novices.

Bereiter et Scardamalia (1983) présentent le modèle de révision "CDO" (*Compare, Diagnostic, Operate*). Selon ce modèle, le scripteur compare le texte qu'il a écrit au texte qu'il a eu l'intention d'écrire, donc au texte planifié. Ensuite, il tente de savoir d'où vient l'écart qu'il a constaté entre le texte produit et le texte planifié. En fonction de cela, il modifie son texte.

Bereiter et Scardamalia (1987) proposent deux modèles de production écrite en langue maternelle. Le premier modèle est adapté à des scripteurs entre 9 et 16 ans. Pour rédiger, les scripteurs de cette classe d'âge ont tendance à utiliser la stratégie de connaissances rapportées[7] (*knowledge telling strategy*). Ils ne transforment pas les informations qu'ils trouvent, mais ils les juxtaposent. Peu à peu, ils vont se rapprocher du fonctionnement qui correspond au deuxième modèle, celui des scripteurs experts. Les scripteurs experts procèdent selon la stratégie de transformation de connaissances (*knowledge transforming strategy*). Ils traitent les informations identifiées au niveau sémantique et ils les adaptent au nouveau contexte. Les modifications correspondent à la situation d'écriture qui a ses contraintes linguistiques, rhétoriques et thématiques (d'après Marin et Legros, 2008 : 122).

3.1.3. Modèle de Fayol (1996)

En se référant à l'apprentissage de la langue maternelle à l'école primaire et au collège, Fayol (1996 : 10) évoque le facteur "connaissances" comme décisif pour la production écrite, Il détaille en particulier la phase de révision.

[7] Cette traduction de *knowledge telling* est utilisée dans le milieu de la psychologie cognitive (Chanquoy et Alamargot, 2003).

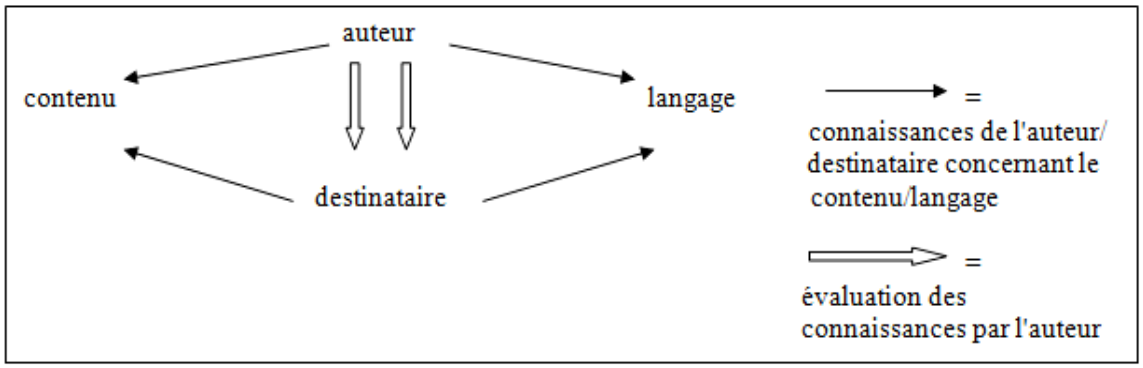


Figure 22 - Modèle de production écrite de Fayol.

Selon ce modèle, la qualité de la production écrite dépend, entre autres, des connaissances thématiques et linguistiques du scripteur et de son estimation de la disponibilité de ces connaissances auprès du destinataire (Fayol, 1996 : 10). Fayol retient du modèle de Hayes et Flower (1980) les processus suivants : planification, mise en texte, révision. Il décrit la révision comme un processus non spontané et rare, même auprès de scripteurs experts (Fayol, 1996 : 33). Les modifications effectuées lors de la révision d'une production écrite entraînent parfois une baisse de qualité (Fayol, 1996 : 34).

> Dès qu'il faut revenir en arrière, réfléchir, réorganiser, des erreurs surviennent. Celles-ci correspondent à la réapparition, dans le texte, d'associations qui existent depuis longtemps (Fayol, 1996 : 29-30).

Lors du processus de révision, il s'agit de reconnaître d'abord qu'une formulation ne convient pas et d'analyser pourquoi elle ne convient pas. R. Ellis (2003 : 149) appelle ce type de prise de conscience "*noticing the gap*". Suite au diagnostic, l'apprenant doit trouver des solutions alternatives. Fayol propose deux stratégies susceptibles d'aider le scripteur à contrôler sa propre production. Le scripteur peut relire sont texte plus tard, afin de prendre de la distance par rapport à la version à réviser, ou bien il peut fait relire son texte par quelqu'un d'autre. Lors de ma recherche, les apprenants sont invités à confronter leur production à celles de scripteurs experts dont les écrits ont été regroupés dans un corpus réfléchi. Ainsi, ils ont l'occasion de constater en quoi leur formulation ne correspond pas à certaines normes linguistiques et/ou discursives attestées. Ils peuvent découvrir que leur formulation n'existe pas dans le corpus, qu'elle y est trop rarement employée ou que le contexte d'emploi dans lequel elle apparaît n'est pas celui qui convient à la tâche de production écrite.

3.2. Modélisation de la production écrite en langue étrangère

Le modèle de Hayes et Flower (1980) qui décrit le processus rédactionnel en langue maternelle est fréquemment employé pour interpréter la façon dont on apprend à écrire en langue étrangère (Barbier, 2003). Pourtant, il convient de tenir compte de certaines différences entre la production écrite en langue maternelle et en langue étrangère. Barbier (2004 : 183-186) en énumère certaines. En langue étrangère, la production écrite est considérée comme plus laborieuse et coûteuse en temps. La phase de planification a tendance à être plus courte qu'en langue maternelle, et les scripteurs y développent souvent moins d'idées qu'en langue maternelle. La phase de révision est généralement plus longue. Pendant cette phase, les scripteurs se concentrent davantage sur la forme (traitements lexicaux et morpho-syntaxiques) que sur les dimensions textuelles et rhétoriques de la production. Les productions en langue étrangère ont tendance à comporter davantage de répétitions que les productions en langue maternelle.

Deux réajustements du modèle de Hayes et Flower (1980) paraissent nécessaires pour l'adapter à la production écrite en LE. Le premier réajustement concerne la prise en compte de l'environnement social et physique de la production. Le deuxième concerne la phase de formulation qui ne semble pas assez détaillée dans ce modèle.

3.2.1. Modèles de Hayes (1996) et de Grabe et Kaplan (1996)

Le premier réajustement du modèle de Hayes et Flower (1980) concerne la prise en compte de l'environnement social et physique de la tâche d'écriture et de certains aspects liés à l'individu. Un tel réajustement est proposé par Hayes (1996). Weigle (2002 : 23-25) précise qu'il évoque l'environnement social, à savoir les co-auteurs, le public (imaginaire ou réel) et l'environnement physique, à savoir les divers stades du texte disponibles à un moment donné et le moyen de mise en texte, comme l'écriture à la main ou le traitement de texte, ce deuxième entraînant une modification des caractéristiques de la phase de planification et d'exécution. Selon Hayes (1996), la motivation et l'affect sont des facteurs qui déterminent, entre autres, la réalisation de la tâche. Les objectifs, prédispositions, croyances, attitudes, l'estimation du rapport entre l'investissement de temps et le résultat ("*cost/benefit estimates*") peuvent influencer la façon dont un scripteur aborde sa tâche. Une des croyances qu'un scripteur peut avoir est l'idée qu'il n'est pas doué pour écrire. Cette croyance peut représenter un obstacle d'apprentissage important. Grabe et Kaplan (1996) ajoutent au modèle de Hayes

(1996) des détails concernant les variables situationnelles. Ils présentent les notions de "*participants, setting, task, text, topic*", et ils précisent les connaissances qui sont nécessaires pour la rédaction, par exemple des connaissances linguistiques et des connaissances discursives concernant la façon dont on construit un texte cohésif en langue étrangère (Grabe et Kaplan, 1996, cité dans Weigle, 2002 : 29).

3.2.2. Modèle de Zimmermann (2000)

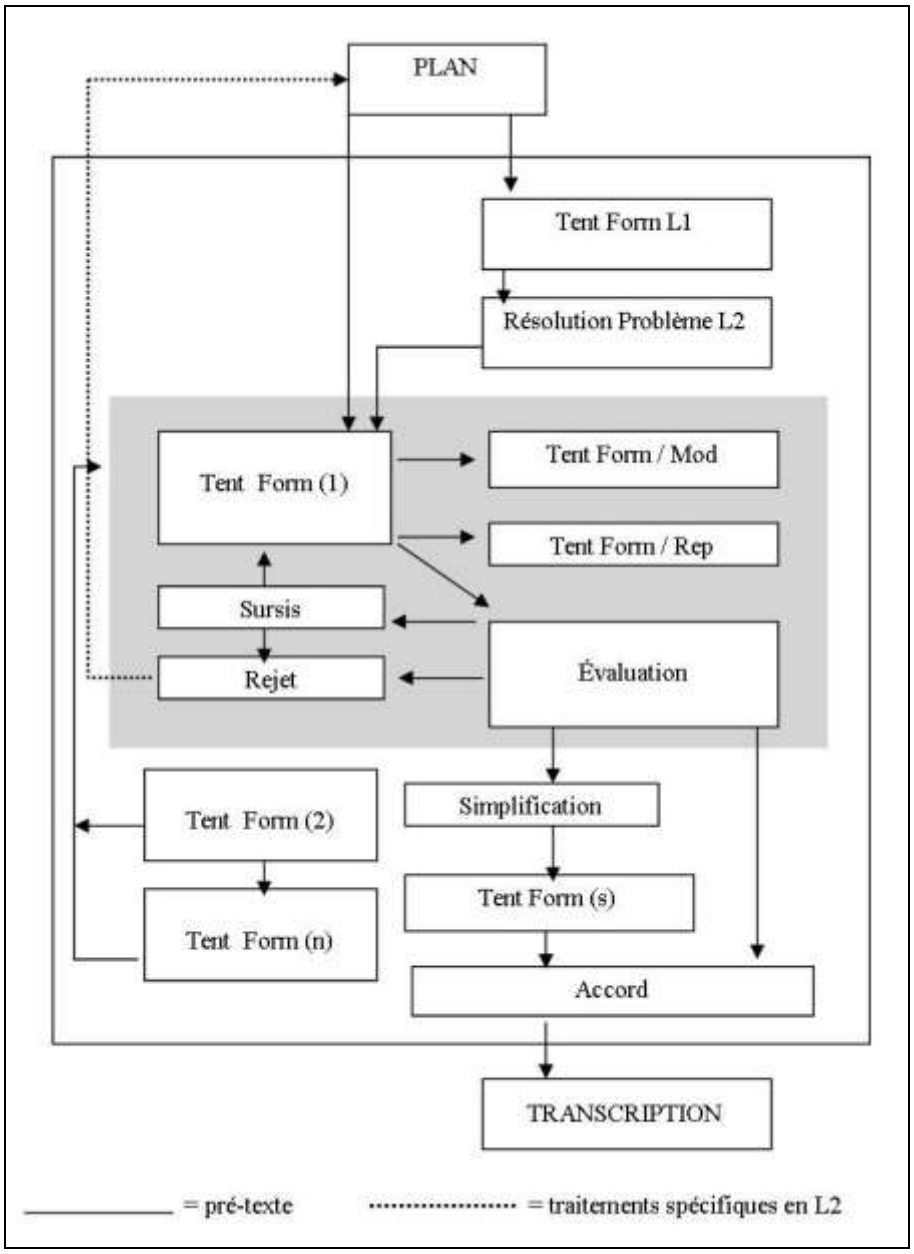


Figure 23 - Modélisation du processus de formulation en LE par Zimmermann.

Le deuxième réajustement du modèle de Hayes et Flower (1980) concerne la formulation, donc le processus de mise en mots. Le modèle de Zimmermann (2000) est basé sur son observation des procédés d'écriture d'apprenants d'anglais, ayant comme langue maternelle l'allemand. Selon le modèle, les scripteurs font des tentatives de formulation (nommées "*Tent form*" dans le modèle) qui peuvent mener à des modifications du texte. Ces modifications sont ensuite évaluées, parfois abandonnées et remplacées par des formulations alternatives qui, souvent, sont plus simples que les formulations auquel le scripteur a décidé de renoncer. Ceci rappelle l'observation de Fayol concernant la possibilité d'une baisse de qualité lors de la phase de révision (3.1.3). Zimmermann (2000) précise que la phase de formulation est en relation étroite avec la phase de révision. Ceci motive mon choix de me concentrer principalement sur les procédés de révision, sans négliger l'analyse fine des caractéristiques des différentes versions de texte produites par les apprenants.

3.3. Répercussion des modèles de production écrite sur mon scénario pédagogique

Quels éléments des modèles que je viens de présenter peuvent être exploités pour l'enseignement/apprentissage de l'écrit en LE ? Les modèles de la psycholinguistique se prononcent sur les quatre phases du processus de production écrite qui constituent une sorte de fil conducteur pour des tâches écrites. Certains modèles décrivent des facteurs contextuels et/ou des différences entre novices et experts. Ces facteurs peuvent fournir des critères pour évaluer les productions écrites. Hayes (1996) considère la lecture comme un facteur déterminant pour la production écrite. Toutefois, les modèles de production écrite que je viens de présenter ne se prononcent ni sur la nature de l'*input*, ni sur la façon dont les apprenants abordent l'*input*. Il convient donc de les compléter par des études qui s'intéressent à la nature des textes susceptibles de soutenir un projet d'écriture et à la possibilité de les exploiter à des fins de rédaction. Les modèles psycholinguistiques ne nous renseignent pas non plus sur le concept de scripteur, ni sur l'aide à la rédaction, qu'elle soit humaine ou logicielle. Dans les sous-chapitres suivants, je vais présenter le concept de scripteur et trois types de pratiques d'enseignement de la production écrite. Ensuite, j'évoquerai la nature de l'*input* que l'on peut proposer lors de tâches d'écriture. Puis, j'aborderai les thèmes de genre textuel et d'intertextualité, en précisant leur lien avec la possibilité de l'exploitation de l'*input* par les scripteurs. Le sous-chapitre 3.6. traitera de l'aide logicielle et humaine.

3.4. Enseignement/apprentissage de la production écrite en langue étrangère

3.4.1. Pratiques d'écriture

Que veut dire "écrire" ? Dans la littérature concernant l'enseignement/apprentissage de la production écrite, la personne qui écrit un texte est fréquemment appelée "scripteur" (Plane, 1996 ; Mangenot, 2000 ; Legros et Crinon, 2002). Il existe d'autres termes qui correspondent à l'activité d'écriture : auteur, écrivain, rédacteur. Les termes d'"écrivain" et de "rédacteur" renvoient à des professions et sont rarement utilisés pour désigner des apprenants. Regardons quelques définitions des termes "auteur" et "scripteur" afin de comprendre les implications liées à leur l'utilisation dans le milieu de l'enseignement/apprentissage. Les définitions du mot "auteur" ci-dessous évoquent la proximité entre écriture, création et invention.

Trésor de la langue française informatisé [auteur]

> I. Celui ou celle qui est la cause principale d'une chose. Syn. créateur, instigateur, inventeur, responsable (…).
>
> II. Domaine des arts, des sc. et des lettres. Celui et celle qui, par occasion ou par profession, écrit un ouvrage ou produit une œuvre de caractère artistique.
>
> éthymol. : 1160-74 : auctur, "écrivain, celui qui a écrit un livre".
> 1174 : autur "celui qui est à l'origine de quelque chose".

Les définitions suivantes du mot "scripteur" présentent l'écriture comme un acte technique, parfois explicitement manuel. L'idée de créativité ou d'invention n'y apparaît quasiment pas.

Trésor de la langue française informatisé [scripteur]

> A. Relig. Cath. Officier de la chancellerie pontificale, qui écrit les bulles.
>
> B. [P. oppos. à locuteur] Celui qui écrit, qui met par écrit. Empr. au latin scriptor : "secrétaire, rédacteur, écrivain".

Dans le milieu de l'enseignement/apprentissage de la production écrite, la vision de l'acte d'écriture semble évoluer depuis un certain temps, au bénéfice d'un intérêt grandissant concernant sa technicité (Plane, 2006 ; Bouchard, J.-P. Simon et Vourzay, 2004). Il n'est pas inintéressant de voir comment ce changement se traduit dans les programmes officiels pour

l'enseignement du français au lycée. Voici un extrait du tableau récapitulatif du document d'accompagnement des programmes de la classe de seconde et de première (2001).[8]

> Travail d'écriture :
> Comprendre que l'écriture est un travail.
> **Devenir sensible** aux questions de style.
>
> Les réécritures :
> Réfléchir, sur les processus, réfléchir sur les processus de création.
> Comprendre que l'écriture suppose des modèles, des références.
>
> **Être sensible** aux questions de style, à l'originalité d'une écriture sur un thème. Améliorer la maîtrise de la langue écrite.

Même si les termes "sensible" et "modèle" renvoient encore à une conception romantique ou classique de l'acte d'écriture, il est évoqué ici que "l'écriture est un travail" qui peut être appris si l'on réfléchit aux "processus de création". Vu ainsi, les apprenants-scripteurs peuvent être considérés comme des "auteurs" ce qui est valorisant. Ils ne sont plus appelés à reproduire le style ou la "plume" d'un auteur considéré comme "grand" et "exemplaire". Ceci ne veut pas dire qu'il ne faut plus donner d'*input* aux apprenants ou ne plus s'intéresser à l'observation de genres et de styles. Il convient de proposer aux apprenants une grande quantité d'*input*, donc un corpus de textes, et de les encourager à le considérer comme un matériau à retravailler. La question du "don d'écriture" est ainsi relativisée.

Plane (1996 : 51-53) présente des pratiques qui existent dans l'enseignement/apprentissage de l'écrit en français, langue maternelle. La portée générale de ses propos justifie de les prendre en compte pour le domaine de la langue étrangère. La première pratique qu'elle décrit consiste à inviter les apprenants à faire des rédactions, ceci sans les initier ni aux genres textuels, ni à des techniques d'écriture, et à évaluer les productions dans leur globalité. Le "Grand Tout" est noté. Les enseignants qui adhèrent à la deuxième pratique décomposent le processus d'écriture de façon analytique. Les apprenants sont censés cumuler les composantes pour acquérir des savoirs et savoir-faire. La production écrite sera exécutée au terme de cette phase d'apprentissage. Le texte obtenu par la première pratique est évalué de façon holistique, celui de la deuxième pratique de façon analytique (cf. 4.2.2. pour la distinction entre perception holistique et analytique). Dans la lignée du modèle de Hayes et Flower (1980), Plane propose une troisième pratique qui paraît contourner les inconvénients des deux premières. Selon cette méthode, la production écrite est considérée comme une situation de résolution de problème.

[8] Les termes en gras le sont dans l'original.

Le travail de l'apprenant consiste alors à surmonter certains obstacles. Ces obstacles peuvent être de nature épistémologique, psycho-cognitive, ou didactique et pédagogique. Les obstacles de nature épistémologique concernent les représentations des apprenants concernant l'écriture, les thèmes, etc. Il s'agit de déconstruire des représentations erronées, comme, par exemple, l'idée qu'une production écrite doit être faite selon un processus linéaire et irréversible. Il convient de remplacer de telles représentations, dont certaines sont ancrées dans l'histoire de la littérature (le concept du poète inspiré par exemple), par des représentations adéquates, comme, par exemple, l'idée que l'on peut retravailler un texte et revenir sur ce que l'on a écrit à un moment donné. Les obstacles de nature psycho-cognitive peuvent avoir trait à la surcharge cognitive. Par exemple, si un scripteur ne connaît pas bien son sujet, la cohésion du texte peut s'en ressentir. Il peut également y avoir un conflit entre un investissement affectif lors de la production d'une version initiale et la nécessité de prendre de la distance afin de la réviser (cf. 3.1.3 et 3.2.2). Les obstacles de nature didactique et pédagogique peuvent provenir de représentations erronées de l'enseignant concernant la production écrite. Il y a des pratiques d'enseignement qui favorisent particulièrement certaines composantes du texte, comme l'orthographe, par exemple. Le traitement de texte peut apporter des aides à la rédaction. Toutefois, s'il est proposé sans guidage adapté, les apprenants n'en tireront pas parti.

3.4.2. Input *et production écrite*

3.4.2.1. Quel *input* pour un projet d'écriture ?

Un rapide aperçu des modèles de la rhétorique classique et de la critique génétique, déjà évoqués en 3.1., permettra d'estimer l'apport éventuel de ces deux types de modèle pour évaluer le rôle de l'*input* pour l'apprentissage de la production écrite.

Le modèle de la rhétorique classique est parfois associé à des termes comme "reproduction d'un modèle" et "centration sur le produit". À mon avis, la prescription de modèles et la linéarité du processus de rédaction ne sont pas des composantes obligatoires de la rhétorique. Ces caractéristiques sont peut-être davantage liées à certaines pratiques d'enseignement/apprentissage de la production écrite qui ont eu cours en France et en Allemagne avant leur remise en cause autour des années 70 (Plane, 1996 ; Schenk, 1998). Il paraît judicieux de "revisiter" le modèle rhétorique pour la didactique de la production écrite en modifiant certains de ses aspects, comme par exemple la mise à disposition d'un *input* plus varié et en quantité plus importante. Avec l'émergence de bases de données pédagogiques (Develotte et Potolia, 2005), il semble intéressant d'enseigner la production écrite en mettant

l'accent sur les genres textuels (Bronckart, 1997) et sur les techniques d'écriture (Mangenot, 1997 ; Bouchard, J.-P. Simon et Vourzay, 2004).

Le modèle de la génétique critique a été employé par Doquet-Lacoste (2004) dans un contexte d'enseignement/apprentissage de la production écrite en langue maternelle. Je retiens de cette approche l'intérêt pour les brouillons, donc les différentes versions produites avant la version considérée comme définitive. Le modèle de la génétique critique permet, davantage que les deux autres modèles de production écrite, la prise en compte simultanée du processus d'écriture et des diverses variantes du produit. Selon Olive et Piolat (2003), un tel regard double semble prometteur.

Rastier (2004) préconise le texte comme unité minimale linguistique. Un projet d'écriture s'appuiera alors utilement sur un corpus qui regroupe des textes ayant des caractéristiques communes (par exemple, le genre textuel). Le corpus doit correspondre au thème ou à d'autres aspects du projet d'écriture. De tels corpus (nommés ici "bases de données") sont proposées dans des projets décrits par Crinon, Legros et Marin (2002a) pour le français, langue maternelle et par Mangenot (1997 ; 2000) pour le français, langue maternelle ou étrangère. Dans le contexte de l'enseignement/apprentissage de la production écrite en anglais, langue étrangère, Henry et Roseberry (1996 ; 2001 : 93-133) et Rohrbach (2003) font utiliser des corpus réfléchis (cf. 1.2.4). Les corpus choisis par Henry et Roseberry sont découpés en sous-corpus qui correspondent à des parties "obligatoires" des documents (*obligatory moves*). À titre d'exemple, dans leur corpus regroupant des documents du domaine du tourisme (Henry et Roseberry, 1996), la partie "Informations pratiques" est obligatoire, tandis que "Histoire de la région" est non obligatoire. Lors du projet de Henry et Roseberry (2001), des apprenants adultes explorent des sous-corpus afin de se préparer à la rédaction d'une lettre d'embauche et du support écrit d'un discours introduisant des conférenciers lors d'un colloque. À l'aide d'un concordancier, ils désignent la fréquence d'unités lexicales et de segments. Ils sont censés en tirer des conclusions susceptibles d'avoir des répercussions sur leur production écrite (Henry et Roseberry, 2001 : 100). Celle-ci est exécutée à la fin du projet, donc après une phase d'exploration du corpus et après l'observation des caractéristiques des divers genres textuels des parties (cf. l'approche analytique, décrite en 3.4.). Les auteurs n'évoquent pas l'exploitation des corpus pour la révision, et il ne semble pas y avoir de va-et-vient entre productions provisoires et corpus. Néanmoins, je retiens l'idée de faire observer les caractéristiques linguistiques et discursives des textes. Je m'inspire en particulier du projet de

Henry et Roseberry (1996) concernant la production de textes du domaine du tourisme. Je retiens des travaux de Henry et Roseberry que des techniques d'écriture peuvent être enseignées/apprises à l'aide d'outils d'interrogation de corpus.

3.4.2.2. Genre textuel

Les titres des articles de Henry et Roseberry comportent les termes "*linguistic patterns of one genre*" (1996) et "*teaching a genre*" (2001). Ils définissent le terme de "genre" comme "a *cultural and interpersonal event making use of language*", donc comme un "phénomène" qui est lié à une culture et une société et qui implique l'emploi de la langue (Henry et Roseberry, 2001 : 94). Cette définition peut être complétée par celle de Rastier (2001).

> Un genre se définit en effet par (i) la cohésion d'un faisceau de critères, tant au plan du signifié qu'à celui du signifiant, et par (ii) son incidence sur la textualité, sur ces deux plans également.

Henry et Roseberry s'intéressent aux critères qui déterminent un genre textuel en prenant la perspective de didacticiens. Ils utilisent le terme de *moves* pour désigner les parties qui constituent le texte. Il paraît nécessaire de clarifier à quel terme français correspond le terme anglais. Si nous suivons la logique proposée par Henry et Roseberry (2001), reprise d'ailleurs par Rohrbach (2003), le terme de *move* équivaut à ce que Bronckart appellerait "séquence d'un texte". Le genre "Lettre d'embauche" comporte, par exemple, le *move* "présenter quelqu'un sous un jour positif" (Henry et Roseberry, 2001 : 98). En se référant à Adam (1992), Bronckart (1997 : 220) définit le terme de "séquence" comme suit.

> [Les séquences sont des] unités structurelles relativement autonomes qui intègrent et organisent des macro-propositions, ces dernières combinant elles-mêmes diverses propositions.

Bronckart (1997 : 220-221) reprend les cinq prototypes de séquences proposés par Adam (1992), à savoir la séquence narrative, la séquence descriptive, la séquence argumentative, la séquence explicative et la séquence dialogale. Il complète cette série par la séquence injonctive et le script (Bronckart, 1997 : 240-241). Admettons que les parties d'un dépliant ou site touristique peuvent être considérées comme genres textuels.[9] Dans un dépliant qui présente un musée, le genre "Visite guidée" comporte obligatoirement une séquence descriptive. Elle peut aussi comporter une séquence narrative, pour parler par exemple de l'invention d'une machine ou de la biographie d'un artiste, et éventuellement une séquence injonctive, du type "Venez en famille".

[9] Henry et Roseberry (1996) semblent plutôt les considérer comme parties du genre "dépliant touristique".

Bronckart définit les genres et types discursifs en vigueur comme des "modèles sociaux" ou des "exemples-type" adaptés à une situation de communication-type donnée (Bronckart 1997 : 79). Selon Bronckart (1997 : 103), un scripteur doit "choisir, parmi les genres de textes disponibles dans l'intertextualité, celui qui lui paraît le plus adapté à sa situation d'action spécifique". Le scripteur a certaines connaissances concernant les caractéristiques d'un texte du même genre textuel que celui qu'il est en train de rédiger. Les connaissances concernant les genres aident le lecteur à interpréter une œuvre (Jauss, 1978). Elles aident aussi le scripteur à évaluer quels moyens discursifs et linguistiques sont conformes à une situation d'écriture donnée. Bronckart (1997 : 110) précise que

> l'agent qui entreprend une action langagière doit nécessairement mettre en interface la connaissance de sa situation d'action et celle des genres de textes, tels qu'ils sont indexés dans l'intertexte et tels qu'ils mobilisent les ressources et les préconstruits particuliers d'une langue naturelle.

Dans ce qui suit, j'utiliserai le terme de section que l'on rencontre dans le domaine de la linguistique de corpus (voir l'interface de *Lexico3* de Fleury *et al.*, 2003) et de la documentation.[10] J'ai fait ce choix en ayant conscience du fait que dans ces domaines, le découpage en sections n'obéit pas aux mêmes règles que le découpage en séquences décrit chez Bronckart (1997 : 220). Le terme de "section" paraît moins restreignant que celui de "séquence", tout en permettant de désigner, comme le terme de *moves*, les parties d'un texte.

3.4.2.3. Que faire avec l'*input* ?

3.4.2.3.1. Recherche d'informations

Crinon, Legros et Marin (2002b : 188) rappellent que la lecture joue un rôle important dans le processus d'écriture et de réécriture. Les textes sélectionnés pour soutenir un projet d'écriture peuvent être abordés de plusieurs manières. Il est possible d'en lire certains intégralement, ou bien, il est possible que le lecteur développe un regard sélectif afin d'y repérer les éléments qui paraissent pertinents en vue de la réalisation de son projet. Comme le précisent Rouet *et al.* (2004), le premier type de lecture mobilise la compétence de compréhension globale, tandis que le second type de lecture nécessite un regard qui cible les éléments susceptibles d'aider le lecteur à résoudre un problème (cf. 4.2.3.). La recherche d'informations convient, à mon avis, davantage que la lecture exhaustive, à des apprenants d'allemand de niveau

[10] L'article "Numériser : conversion en mode texte" sur le site de la BnF (Bibliothèque nationale de France) contient ce terme. http://www.bnf.fr/pages/infopro/numerisation/num_technique_tdm.htm

intermédiaire en situation de révision de texte. Elle a des chances de leur convenir encore mieux si la mise en forme matérielle du texte renforce la visibilité des éléments potentiellement pertinents. Les concordances sont un bon exemple d'une telle mise en forme matérielle (cf. 1.6.3.). Elles permettent de présenter un mot clé ou un segment de façon particulièrement saillante, ceci grâce à sa position dans le document et grâce à sa mise en forme particulière (en gras, en rouge, en capitales, etc.). Les informations que les apprentis-scripteurs trouvent lors de la lecture peuvent ensuite être mises à profit lors de la phase de révision et lors du processus de replanification (Crinon, Legros et Marin, 2002a).

3.4.2.3.2. Démarche intertextuelle

Les textes ne sont jamais créés à partir de rien, mais ils s'insèrent dans d'autres textes écrits auparavant (Bouchard, J.-P. Simon et Vourzay, 2004), donc dans des intertextes. Suite à des importations volontaires ou involontaires d'unités ou de segments, de figures de style, de situations, etc., les intertextes peuvent laisser des traces dans les nouvelles productions. Le concept d'intertextualité évoque la dimension historique et sociologique des textes. Selon Bronckart (1997 : 111),

> tout texte empirique est réalisé par emprunt à un genre, et il relève donc lui-même toujours d'un genre ; tout texte empirique procède cependant aussi d'une adaptation du genre-modèle aux valeurs attribuées par l'agent à sa situation d'action, et dès lors, outre les caractéristiques communes au genre emprunté, il exhibe également des propriétés singulières, qui définissent son style propre.

En écrivant, nous reprenons dans une certaine mesure les caractéristiques des intertextes, mais nous les façonnons également en y ajoutant notre propre style, en proposant de nouveaux types de combinaisons de segments, d'éléments discursifs, etc. (Bronckart 1997 : 79). Teubert (2006) considère également qu'un nouveau texte est toujours joint à un discours existant. Il est une recombinaison et permutation de textes existants.

Dans le contexte de l'acquisition de la production orale, Aguado (2002a : 215) présente l'imitation, définie comme la reprise de ce que dit un interlocuteur, comme une preuve que l'attention de l'apprenant-locuteur est dirigée vers l'*input* oral. Une telle attention peut également être mise à profit dans le contexte de l'apprentissage de la production écrite. Le scripteur qui a comme tâche l'identification d'unités ou de segments dans un texte afin d'en emprunter certains doit aller à la rencontre du texte. L'attention que l'apprenant accorde à l'*input* est susceptible de l'aider à assimiler ou "re-traiter" au moins une partie de ce qu'il a observé et d'en retenir certains éléments. Doughty et Long (2003 : 63) décrivent la sélection

de segments dans un corpus de textes à des fins d'exécution d'une tâche écrite comme une forme de "plagiat nécessaire". Une telle conception de l'écriture peut surprendre : de nombreux enseignants regrettent la tendance au plagiat de certains apprenants à qui Internet offre l'occasion d'obtenir des textes en faisant du "copier-coller" extensif au lieu d'en produire eux-mêmes. Toutefois, rappelons que l'imitation, aussi appelée mimésis, a pendant longtemps été un critère pour la classification générique de textes (Jenny, 2003). Selon Bouillaguet (1996 : 4), au 17ème siècle,

> [L'imitation était] une pratique courante et encouragée en Europe – la valeur des œuvres tenant non pas à une originalité à laquelle on ne croit guère, mais à l'habileté des auteurs à adapter un modèle ancien à une réalité nouvelle (Bouillaguet 1996 : 4).

À la fin du 19ème siècle, Anatole France crée le lien entre une telle doctrine de l'imitation et le concept qui précède la théorie de l'intertextualité telle que nous la connaissons depuis le 20ème siècle, grâce, entre autres, aux travaux de Julia Kristeva (Bouillaguet 1996 : 5). La "renaissance" du procédé d'imitation peut être particulièrement observée dans le contexte d'une écriture ludique (pastiche, parodie, collage, centon). L'intertextualité comme procédé littéraire est alors motivée par l'intention d'être drôle ou de vouloir créer un style incongru. Ce procédé se distingue radicalement de "l'imitation-culte" du 17ème siècle (Bouillaguet 1996 : 5). L'encouragement à l'emprunt dans le cadre de l'enseignement/apprentissage de la production écrite en LE n'est pas forcément motivé par l'intention de détourner des textes à des fins humoristiques. Toutefois, cet usage d'un intertexte n'est pas exclu. Citons l'exemple de deux apprenantes qui ont participé à ma recherche-action et qui ont tourné en dérision le titre d'une critique de film en remplaçant les cinq étoiles distribuées par l'auteur au film par deux étoiles.

Titre sur lequel Olivia et Géraldine se sont appuyées :
Diesmal haben alle 5 Rezensenten Recht!* [Cette fois-ci, tous ceux qui ont donné cinq étoiles au film ont raison].
Nouveau titre proposé par Olivia et Géraldine :
Diesmal haben alle 2 Rezensenten Recht!* [Cette fois-ci, tous ceux qui ont donné deux étoiles au film ont raison].

Selon St John et Cash (1995 : 50), le procédé de réutilisation de segments trouvés dans des intertextes peut améliorer la qualité de productions écrites dans le domaine de la syntaxe (cf. 6.4.3.). Šimunek (2007) a étudié la reprise de segments dans l'*input* par des étudiants allemands, futurs enseignants d'anglais, lors d'un projet collaboratif impliquant des étudiants

canadiens. Le projet a permis aux étudiants de créer ensemble des tâches pour l'enseignement de la langue étrangère. Les échanges de courriels des deux groupes à des fins de création de tâches ont servi d'*input* pouvant permettre le repérage de segments. Lors de ce projet, les apprenants semblent avoir eu assez peu recours au procédé de reprise de segments dans les courriels des partenaires canadiens. Parfois, des raisons affectives ont mené à une reprise de segments trouvés dans un courriel (Šimunek, 2007 : 295).

Le manque de "popularité" du procédé de reprise d'éléments dans l'*input* peut s'expliquer, entre autres, par le fait qu'il ne correspond pas à une pratique scolaire répandue. Souvent, les scripteurs sont invités à formuler "avec leurs propres mots". Citons, à titre d'exemple, un extrait du programme d'enseignement pour l'allemand, palier 2, au collège (2007).

> Reformulation : L'élève sera capable de transmettre une information en utilisant ses propres mots.

Dans un contexte de psychosociologie, Linard (1996) considère l'imitation de comportements humains comme un processus qui aide à intégrer des informations nouvelles.

> Chez l'individu en développement, l'imitation permet de dériver activement sa structure du dedans d'une structure du dehors qui lui ressemble tout en étant autre (...). Cette imitation (...) ne peut en aucun cas se réduire à une simple copie-transfert entre imitant et imité. C'est une forme d'action auto-organisée et hétéro-régulée qui implique des capacités d'intégration et d'interprétation, totalement étrangères aux automates (Linard, 1996 : 97-98).

Ces paroles trouvent un écho en Vygotski (1931), relu par Brossard (2004). Ce dernier explique que, selon Vygotski, l'imitation est fondamentale pour l'apprentissage ; elle n'est toutefois jamais une reprise à l'identique : "l'activité en même temps qu'elle s'intériorise, elle se réorganise." (Brossard, 2004 : 104).

Encourager les apprenants à identifier des unités dans l'*input* et à les réutiliser afin de réviser leur production écrite est une méthode qui implique le traitement et parfois la modification des données identifiées. Ce sont des activités qui ont un coût cognitif plus ou moins élevé. À la lumière des perspectives développées par Aguado (2002a : 215), Linard (1996 : 97-100) et Vygotski (1934 : 353-355), cette méthode peut soutenir l'apprentissage de la production écrite en LE, me semble-t-il.

3.5. Nouveaux genres, nouvelles normes ?

L'écriture en ligne est un champ d'observation privilégié d'évolutions dans le domaine des normes linguistiques. Afin d'illustrer ces propos, je décrirai ci-dessous certaines

caractéristiques d'un des genres textuels qui font partie du projet *Projekt Prospekt*, proposé lors de ma recherche-action. Le genre "Commentaire dans le livre d'or électronique d'un musée" a des caractéristiques linguistiques, discursives et typographiques spécifiques. Dans le corpus "Livre d'or", nous pouvons observer un procédé de mise en relief récurrent, à savoir l'occupation de la première place dans la phrase par un élément autre que le sujet. Il peut s'agir d'un adverbe ou d'une collocation à sens intensifiant : "*Insbesondere meine Kinder waren begeistert*" [Ce sont surtout mes enfants qui ont été enthousiastes] ou au superlatif : "*Am besten hat mir das digitale Beethovenhaus gefallen.*" [C'est la maison de Beethoven qui m'a plu le plus.] Nous trouvons dans cette position marquée également des compléments d'objet indirect : "*Mir hat es gefallen*" [Moi, j'ai aimé]. La répétition d'intensificateurs, comme "*total*" ou "*sehr*" [totalement ou beaucoup], est également fréquente dans les commentaires. L'orthographe est parfois personnalisée ; l'omission des majuscules des noms est fréquente. Les commentaires se caractérisent par un degré de formalité peu élevé. De nombreux scripteurs renoncent à des formules de politesse marquant le début et la fin de leur commentaire. Ces formules peuvent être oralisantes : "*Hallo*!" [salut] et courtes : "*Gruß*" [salutations]. Dans certains des commentaires, nous pouvons observer la répétition de points d'exclamation, des points de suspension et l'absence de points et virgules. Avec l'émergence du genre "commentaire en ligne", une nouvelle norme linguistique semble se dessiner (cf. 6.5.). Cette évolution est également observable pour le clavardage, par exemple. La rédaction d'un commentaire dans un livre d'or électronique peut permettre aux apprenants de se rendre compte de ces nouvelles normes qui risquent, d'ailleurs, d'être moins ressenties comme "nouvelles" par les scripteurs adolescents que par les adultes.

3.6. Produire des textes en LE à l'aide de l'ordinateur

3.6.1. Caractéristiques du document numérique

La numérisation de documents présente certains avantages pour les utilisateurs. Elle facilite l'accès aux documents, et elle permet de présenter les données de multiples façons. Grâce à des liens hypertextuels, plusieurs textes peuvent être mis en relation les uns avec les autres. Intertexte "par excellence", le document hypertextuel nourrit d'autres textes, et il s'en nourrit. Ma recherche ne s'appuie pas explicitement sur des hypertextes. Certains concordanciers, comme celui que propose *AntConc3.2.1w* (Anthony, 2007), permettent de naviguer entre le corpus et les lignes de concordance. Il paraît alors défendable de considérer leur relation comme hypertextuelle.

3.6.2. Rôle de l'outil

Des recherches ont été faites afin de déterminer le rôle de l'utilisation de l'ordinateur pour la production écrite (David et Plane, 1996 ; Crinon et Legros, 2002 : 107-120). Les comparaisons entre production écrite papier et électronique ont montré que les performances des scripteurs tiennent moins au média qu'à la médiation. Selon Crinon et Legros (2002 : 115), "l'effet du traitement de texte est d'installer dans la classe de nouvelles modalités de travail, bénéfiques à l'apprentissage de la production de texte." Ces nouvelles modalités peuvent se traduire par des "médiations qui cassent la relation pédagogique duelle et permettent aux enfants d'évoluer" (Crinon et Legros, 2002 : 115). Il serait intéressant de développer ce point de vue esquissé par Crinon et Legros en étudiant le domaine de l'écriture collaborative en ligne, un domaine que je ne puis détailler ici.[11] J'ai fait le choix de proposer aux apprenants une autre forme d'écriture collaborative "indirecte", en les dirigeant vers des corpus réfléchis. Hallet (2008) définit l'acte d'écrire comme "*markierte Intervention eines Individuums in einem fortlaufenden Diskurs*", donc comme la participation explicite d'un individu à un discours continu. Le scripteur s'insère dans un discours existant, "hébergé" dans un corpus, et il façonne ce discours à son tour (cf. 3.4.2.3.2. et 6.3.1.4.).

3.6.3. Savoir-faire numérique

La présentation sous forme numérique permet au lecteur de consulter en peu de temps un grand nombre de documents (Tricot, 2007 : 105). La multimodalité et la possibilité de moduler la quantité de redondance à l'intérieur d'un document contribuent à améliorer la compréhension. Toutefois, le coût cognitif du traitement de documents numériques peut être important. Il paraît nécessaire que les concepteurs et utilisateurs de tels documents développent de nouvelles compétences (cf. 6.6.). Dans la citation suivante, Tricot relève l'importance des compétences en lecture et en compréhension.

> Les individus en difficulté avec la lecture et la compréhension sont en difficulté avec les documents numériques. Si ceux-ci prennent plus de place dans nos sociétés, ce sera au détriment des personnes ne sachant pas lire ou ayant des difficultés de compréhension (Tricot, 2007 : 105).

Selon Tricot (2007 : 18), le document en ligne modifie l'attitude du lecteur – scripteur potentiel – face à l'*input*. Son utilisation implique une activité de recherche d'informations et

[11] Il est utile de consulter à ce sujet le numéro 1 du volume 13 de la revue *ZIF (Zeitschrift für Interkulturellen Fremdsprachenunterricht)*, revue concernant l'enseignement interculturel des langues étrangères, qui a comme thème principal la production écrite électronique.

la compréhension de la relation entre plusieurs sources (Tricot, 2007 : 31). Hallet (2008) encourage à développer la littératie électronique auprès des apprenants (cf. le terme de "*electronic literacy approach*" proposé par Shetzer et Warschauer, 2000 : 174). Il précise que dans le contexte d'un travail de production écrite à l'aide de documents numériques, les trois savoirs suivants paraissent nécessaires : savoir sélectionner les informations dont on a besoin (trouver des sites pertinents), savoir sauvegarder ces informations et savoir transformer ces informations pour les adapter au nouveau contexte. Dans la figure suivante, la flèche représente une activité de la part de l'apprenant.

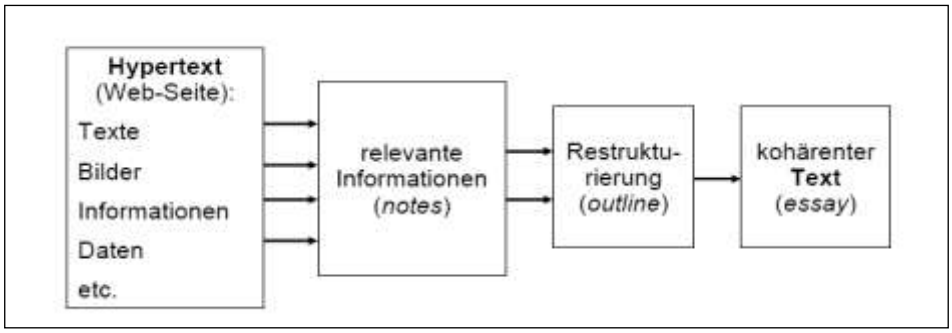


Figure 24 – De l'hypertexte au texte (Hallet).

3.7. Des outils d'aide à la rédaction vers les ressources pédagogiques

3.7.1. Outils d'aide à la rédaction

Legros et Crinon (2002 : 116) rappellent que l'utilisation des machines doit être accompagnée par des médiateurs humains et qu'il y a une interaction entre les médias, les apprenants et le contexte. Essayons de déterminer en quoi les logiciels d'aide se distinguent de l'aide humaine qui peut être apportée en situation d'enseignement/apprentissage de la production écrite.

Dans son article concernant les aides à l'apprentissage, Demaizière (2007) distingue entre aide humaine et aide logicielle. Elle présente la première comme un service fourni par un enseignant ou formateur et la deuxième comme composante d'un produit. Lors d'une formation, les deux ont un rôle à jouer. Ce rôle peut être plus ou moins accentué selon les postures d'enseignement/apprentissage et les ressources numériques adoptées. Les deux types d'aide peuvent consister en une réaction qui fait suite à une action de la part de l'apprenant. On le félicite par exemple pour une solution qui correspond à ce qui est attendu ou on lui indique la première lettre d'une solution possible. Parfois, l'enseignant ou le logiciel commentent la proposition de l'apprenant. Il s'agit d'une aide réactive. L'aide peut aussi être proactive, c'est-à-

dire qu'elle suggère à l'apprenant des actions futures, par exemple elle l'incite à consulter telle ou telle ressource ou à travailler autrement, etc. Un logiciel peut-il offrir une aide proactive ? Mangenot (1998) décrit des logiciels semi-tutoriels qui semblent aller dans ce sens, par exemple *Gammes d'écriture* (Mangenot, 1996). Un logiciel peut-il suggérer des pistes d'action qui laissent à l'apprenant une certaine liberté d'action et qui ne lui fournissent pas des réponses toutes faites ? Afin d'apporter des éléments de réponse à cette question, nous regarderons dans ce qui suit quelques aides logicielles destinées à soutenir la rédaction, en nous interrogeant sur la nécessité ou la possibilité de les articuler avec une aide humaine.

Certains types de production écrite sont relativement peu complexes, comme par exemple les tâches d'entraînement lors desquelles les apprenants doivent compléter ou transformer du texte de manière très limitée. Si les apprenants sont invités à réaliser une production écrite plus élaborée, ils ont besoin d'outils d'aide à la rédaction. Lors de ma recherche-action, j'ai pu constater que les apprenants se dirigent soit vers un traducteur électronique (par exemple celui proposé sur le site de *free* ou *Reverso*), ou vers un dictionnaire papier ou électronique, quand ils cherchent une aide à la rédaction (cf. 6.2.6.2.). Regardons de plus près les caractéristiques et les potentialités de ces outils d'aide à la rédaction.

3.6.1.1. Traducteur électronique

Le traducteur électronique est un logiciel (en ligne ou hors ligne) qui traduit du texte d'une langue vers une autre. Il est intéressant de voir que son nom est un homonyme du métier de traducteur. Ce qui est fait ailleurs par une personne est proposé ici par un outil électronique. Le *TLFi* (*Trésor de la Langue Française*) propose des définitions qui illustrent un des dilemmes liés à l'émergence de la traduction automatisée : comment repérer et traduire les unités de sens composés de deux unités lexicales ou plus ? Dans le *TLFi*, nous apprenons qu'un traducteur humain est une "Personne effectuant des traductions, transposant des textes d'une langue vers une autre". Le traducteur humain interprète – dans le sens littéral du terme latin interpretari[12], en prenant plus ou moins de liberté, le message qu'il doit transmettre. L'évolution historique des traducteurs électroniques montre que ces logiciels doivent, encore aujourd'hui, relever le défi de traduire autrement qu'en associant simplement un "mot" d'une langue donnée à un "mot" d'une autre langue. En 1959, le terme "traductrice électronique" ou "traductrice automatique" est utilisé pour désigner une "machine à traduire d'une langue dans une autre" (*TLFi*). Le terme finit par se stabiliser au masculin. En 1972, l'emploi de

[12] "Expliquer, éclaircir, prendre dans tel ou tel sens". TLFi.

"traducteur" est attesté pour désigner une "machine qui fournit la traduction de mots d'une langue dans une autre". L'idée d'une correspondance exacte entre "mots" de plusieurs langues qui est exprimée ici est contestable d'un point de vue linguistique (Chiss, Mainguenau et Filliolet, 2001). Nous observons donc une régression qui accompagne le progrès technique d'automatisation de la traduction. La définition que propose le *TLFi* de "traducteur télégraphique" semble mieux décrire ce que pourrait être une traduction acceptable. Un traducteur télégraphique est une "partie d'un récepteur télégraphique qui restitue le texte d'un message à partir des signaux reçus" (*TLFi*). Malgré les imperfections liées à la difficulté (peut-être plus conceptuelle que technique) de faire traduire par une machine non des unités lexicales simples, mais des segments de taille variable qui correspondent entre eux dans plusieurs langues, les traducteurs dont nous disposons actuellement sont des outils de travail qui peuvent être des aides précieuses, comme l'a démontré Poiarkova (2006) pour le contexte de la formation d'enseignants de langue. Toutefois, leur utilisation pertinente requiert une bonne maîtrise de la langue, car ils nécessitent un travail d'évaluation et de vérification des résultats.[13] Grâce au traducteur électronique, les apprenants obtiennent certes très rapidement un texte en langue étrangère à partir d'un texte qu'ils ont saisi en langue maternelle, mais son emploi présente deux inconvénients majeurs. Premièrement, les résultats sont peu sûrs, surtout si le texte saisi en langue maternelle contient des erreurs d'orthographe. À titre d'exemple, lors de deux phases différentes de ma recherche, deux des apprenants n'ont pas remarqué que le traducteur a traduit le nom d'une ville. La ville de Saint-Leu-la-Forêt a ainsi été transformée en "perforeuse sainte". Deuxièmement, l'utilisation du traducteur électronique ne donne pas l'occasion de traiter la LE, sauf si l'apprenant prend la peine de vérifier la validité du résultat. Les apprenants qui ont recours à ce type d'outil adoptent parfois une posture d'apprentissage qui les incite peu à effectuer de telles vérifications. Ils semblent attendre du traducteur qu'il formule à leur place, peut-être parce que l'obstacle lié à une rédaction en LE leur paraît trop important. L'ordinateur est utilisé ici de façon excessive comme "dispenseur de savoir". Un apprenant qui utilise (presque) exclusivement un traducteur signale à mon avis qu'il a besoin d'aide, mais peut-être d'aide humaine davantage que technique, dans un premier temps tout au moins. Il serait éventuellement possible de répondre à cette demande par la mise à disposition de corpus parallèles qui sont susceptibles de rassurer un apprenant qui se sentirait dépassé par le "tout allemand".

[13] Cordier-Gauthier et Dion (2003) évoquent d'ailleurs également l'expertise comme facteur décisif pour l'utilisation appropriée de correcteurs grammaticaux par des apprenants de niveau intermédiaire.

3.6.1.2. Dictionnaire électronique

Comme nous l'avons déjà constaté pour le traducteur électronique, le dictionnaire électronique est un outil qui demande un certain degré d'expertise, et son utilisation a un certain coût cognitif. Selva et Verlinde (2002) concluent leur étude de cas sur l'utilisation d'un dictionnaire électronique par des apprenants de FLE à l'université en disant que

> le point crucial reste à notre avis l'effort que veut bien consentir l'utilisateur pour lire et analyser les informations contenues dans le dictionnaire. L'apprentissage du dictionnaire reste donc le point décisif.

Les dictionnaires font partie des grands "classiques" conseillés aux apprenants de langues étrangères. Pourtant, les apprenants collégiens ont rarement l'occasion d'apprendre à les utiliser, qu'il s'agisse de versions papier ou électroniques. Le choix qu'ils doivent effectuer parmi plusieurs solutions proposées n'est toutefois pas toujours aisé. Demaizière (2007) relève la difficulté que peut avoir un apprenant quand il doit exploiter les informations proposées dans un dictionnaire. Une définition de dictionnaire est un texte d'un haut niveau d'abstraction ; savoir la lire et savoir trier entre les différents sens est d'autant plus délicat que l'on a un niveau peu avancé. Les erreurs de choix dans un dictionnaire peuvent mener à des erreurs de sens, comme l'illustre l'emploi du terme "*Bergarbeiter*" [mineurs, dans le sens d'hommes qui travaillent dans des mines], utilisé au lieu du terme "*Minderjährige*" [mineurs, dans le sens de jeunes n'étant pas majeurs] par Yvan pendant la première phase de ma recherche-action. Parfois, la solution qui serait adaptée à une situation précise fait défaut dans un dictionnaire. Toutefois, depuis quelques années, de nouveaux dictionnaires, présentant les unités lexicales dans des contextes variés, voient le jour. Citons, parmi les dictionnaires en ligne les plus connus, *Das Digitale Wörterbuch der Deutschen Sprache (DWDS)*, un dictionnaire en ligne basé sur deux corpus oraux et des corpus écrits, et *Wortschatz-Portal*, un dictionnaire en ligne spécialisé dans la représentation de la langue allemande très récente qui trouve ses données sur Internet (par *web-crawling*). Les corpus et concordances permettent de s'appuyer sur des données attestées afin d'élaborer des dictionnaires qui paraissent plus fiables et plus complets que ceux élaborés à l'aide de l'intuition des auteurs, une pratique encore répandue pendant les années 90.[14] Tutin (2005) discute plus amplement l'utilité de dictionnaires de collocations. Pour l'apprentissage d'une langue étrangère, nous trouvons depuis quelques années des dictionnaires de fréquence basés sur des corpus électroniques, tel que *A Frequency Dictionary*

[14] En 1992, j'ai pratiqué moi-même la méthode de l'introspection lors d'une co-rédaction d'un dictionnaire franco-allemand de taille moyenne, publié par Larousse.

of German (Tschirner *et al.*, 2006). Les auteurs présentent ce dictionnaire, basé sur un corpus de référence (cf. 1.2.4.), comme un outil d'apprentissage parmi d'autres.

> *This German frequency dictionary is intended to be used as a learning tool in conjunction with other learning tools, i.e. a good German/English dictionary, a good textbook, and other support material that can be helpful for learning German* (Tschirner *et al.*, 2006 : 8).

Notons que le site *LEO*, surtitré "un service en ligne", propose non seulement un dictionnaire électronique qui permet de traduire de l'allemand vers plusieurs langues (l'anglais, le français, l'italien, l'espagnol et le chinois) et vice versa, mais aussi un forum de discussion qui complète, le cas échéant, les informations du dictionnaire. Il propose aussi un espace de travail collaboratif pour apprendre le vocabulaire, seul ou en groupe, et pour faire des activités de type lecture de textes, dictée, traduction, questionnaire à choix multiple. Nous nous dirigeons vers des ressources que l'on peut appeler pédagogiques. *LEO* est géré par une entreprise privée. Ce site ne semble pas avoir été créé en partenariat avec le *Kultusministerium*, un organisme allemand, comparable à l'Éducation Nationale. Pourtant, j'entrevois ici la possibilité de réduire le "hiatus fort entre les pratiques à l'École et celles qui se diffusent dans la société" (Baron et Dané, 2007). Un tel rapprochement peut être soutenu par la combinaison entre nouvelles fonctionnalités (dictionnaire en contexte et lien entre le dictionnaire et les données obtenues grâce au forum), pratiques scolaires (apprentissage du vocabulaire, dictée, etc.) et pratiques répandues en société (s'exprimer sur un forum). Par conséquent, j'ai proposé aux apprenants l'utilisation du dictionnaire proposé sur le site *LEO* comme outil de "dépannage linguistique" lors de *Brieffreunde.org*, le dernier projet de la recherche-action. Ce projet a encouragé les apprenants à s'appuyer sur un corpus d'annonces pour trouver un correspondant (cf. 5.3.5.).

3.7.2. Les ressources

3.6.2.1. Définition générale

Selon le *TLFi* (*Trésor de la langue française informatisé*), une ressource est initialement un "moyen de se tirer d'embarras ou d'améliorer une situation difficile". Comme synonyme, nous trouvons "recours". Une ressource est donc ce vers quoi une personne se tourne quand elle estime que ses propres moyens (par exemple, ses moyens linguistiques) ne sont pas suffisants et qu'elle a besoin d'avoir recours à une aide extérieure. Selon une acception attestée plus récemment, cette aide peut être matérielle : le *TLFi* parle de "moyens financiers", "moyens matériels dont dispose un pays, une région, une collectivité", "ressources naturelles" et de

"moyens militaires". Les derniers évoquent la superposition entre moyens matériels et humains que nous pouvons parfois observer. En effet, les ressources peuvent aussi désigner une aide venant d'une personne, comme le suggère le terme de "personne-ressource". Il désigne dans le domaine de l'Éducation Nationale des personnes qui ont une fonction d'aide, comme l'illustre cette concordance faite à l'aide de *Sketch Engine*.

on considère l' enseignant comme une	**personne ressource**	éventuelle dont la compétence
réaliser nos objectifs , nous avions besoin de	**personnes ressources**	pour compléter notre formation
psychologues conseillers en orientation , autres	**personnes ressources**	. L' importance sociale du rôle des
aux ordinateurs et à l' Internet , les	**personnes ressources**	locales dans le cadre du sujet de stage

Figure 25 – Mot clé "personne ressource/personnes ressources" dans *French Web Corpus*.

Une personne-ressource peut être un enseignant, spécialiste des nouvelles technologies, comme le confirme aussi la définition ci-dessous, publiée sur le site de l'académie de Rouen. L'accent y semble être mis sur les compétences techniques davantage que pédagogiques.

> Dans chaque établissement, la personne ressource est reconnue comme le collègue auquel on s'adresse si l'on rencontre une difficulté quelconque concernant l'utilisation des TICE dans sa pratique pédagogique. L'équipe administrative fait appel à ses compétences pour le choix et l'installation de nouveaux équipements (Site de d'académie de Rouen, 2001).

Si l'aide humaine et matérielle peut se superposer dans des concepts comme "ressources militaires" et "personnes-ressources", pouvons-nous en conclure que les apprenants attendent des logiciels le même type d'aide que celui auquel ils s'attendent de la part de leurs enseignants ? Ces attentes se reflètent-elles dans l'emploi qu'ils font des outils logiciels ? Le recours des apprenants les moins avancés aux traducteurs électroniques semble confirmer cela. Dans ce qui suit, je présente deux logiciels d'aide à la production écrite. Je m'intéresserai particulièrement au type d'aide proposé par les deux et au rôle que peut jouer l'enseignant qui fait utiliser ces logiciels par les apprenants.

3.6.2.2. Ressources pédagogiques

Dans un contexte d'enseignement/apprentissage, le sens d'aide matérielle de "ressource" peut s'étendre de "manuel scolaire" jusqu'à "site Internet". Le point commun est l'apport attendu pour l'apprenant : "on s'intéresse à ce qui peut être offert à chaque apprenant et on singularise, on discrétise la ressource également." (Demaizière, 2004). L'ordinateur peut-il aider à réduire la complexité de tâches de production écrite en prenant en charge certaines parties d'une tâche ? Ceci semble être confirmé par les études qui analysent l'emploi de certains logiciels

d'aide à la rédaction (Mangenot, 2000 ; Crinon et Legros, 2000). Dans ce qui suit, je présenterai deux logiciels d'aide à la production écrite en français qui contiennent tous les deux un corpus de textes, donc des ressources pédagogiques. Le premier, *Gammes d'écriture* (Mangenot, 1996) a été adapté de l'italien pour le contexte français, langue maternelle et étrangère (collège, lycée et formation continue). Le deuxième, *Scripertexte* (Crinon et Vigne, 2002) ou – selon sa désignation pour la commercialisation, *Écrire en lisant des récits de vie* -, a été développé pour un contexte de production écrite en français, langue maternelle, à l'école primaire (cycle 3) et au collège.

Gammes d'écriture est un logiciel semi-tutoriel. Il permet aux scripteurs de lire et manipuler des textes et d'en écrire en se servant d'éléments préfabriqués. Ce logiciel s'adresse à un public de collégiens, lycéens et adultes (formation continue). Mangenot l'a aussi fait utiliser par des étudiants en FLE (Mangenot, 2000). Le logiciel se compose de trois modules, nommés "Assistants", "Bibliothèque" et "Gymnase". Le module "Assistants" fonctionne selon le modèle d'un dialogue personne/machine. L'ordinateur joue ici le rôle de tuteur. Le module "bibliothèque" est un petit corpus d'extraits de textes. Les textes sont classés dans les rubriques "actes de langage", "types de textes", "ton" et "figures rhétoriques". Il est possible de combiner plusieurs rubriques entre elles. Le module "gymnase" offre des activités de manipulation de textes et d'écriture sous contrainte. Je retiens de ce logiciel deux idées, celle de proposer un scénario qui invite les apprenants à lire des textes d'un corpus et à proposer des activités de manipulation de texte. Par contre, j'aimerais laisser une marge de manœuvre plus grande à l'apprenant que celle que propose *Gammes d'écriture* (Mangenot, 1996).

Scripertexte (Crinon et Vigne, 2002) est un outil d'aide à la réécriture. Il opère avec moins de contraintes que *Gammes d'écriture* (Mangenot, 1996). Les apprenants sont invités à construire un nouveau texte à partir de textes existants : "Il s'agit de construire un parcours en fonction de son projet et non, comme dans un texte linéaire, de reconstruire un sens déjà largement préétabli". Les auteurs proposent la démarche suivante : les apprenants lisent trois textes concernant le thème choisi ; ensuite, ils rédigent un nouveau texte en s'inspirant de cette lecture ; l'enseignant retravaille ce texte pour le normaliser au plan du bas niveau. La lecture de textes supplémentaires, choisis par les apprenants dans le corpus proposé par le logiciel, doit permettre d'améliorer la production dans le but de "la rendre plus intéressante pour les futurs lecteurs" (Crinon et Legros, 2000). La dernière étape consiste à reprendre la première version, revue par l'enseignant, et de l'enrichir à l'aide des éléments trouvés dans le corpus.

Les auteurs montrent qu'une telle utilisation de *Scripertexte* a permis d'obtenir des productions écrites enrichies au plan du haut niveau (ajouts, propositions sémantiques, éléments évaluatifs). Des apprenants qui ont utilisé ce logiciel sur trois ans semblent avoir progressé dans plusieurs domaines.

> On peut donc supposer que ces élèves n'ont pas acquis seulement des compétences liées à un contexte de travail et à un genre textuel particuliers, mais des connaissances plus larges (des métaconnaissances) sur les processus d'écriture et les procédures à mettre en œuvre lors de l'élaboration d'un texte en activant leurs connaissances et en mobilisant des ressources extérieures (Crinon et Legros, 2000).

Une observation comparable a été faite par Kennedy et Miceli (2001) concernant l'utilisation de corpus et concordances lors d'un projet d'écriture (cf. 4.2.3.). Dans l'expérience de Crinon et Legros (2000), il n'a pas été aisé de faire admettre aux apprenants l'intérêt d'épaissir les textes au lieu de rajouter des éléments nouveaux. Ils ont également hésité à admettre comme légitime le procédé de copie d'expressions issues d'autres textes. Un travail explicatif doit accompagner le travail du logiciel. *Scripertexte* (Crinon et Vigne, 2002) semble, davantage encore que *Gammes d'écriture* (Mangenot, 1996), nécessiter la présence aidante d'un enseignant, présence jugée indispensable par Mangenot (2000). L'enseignant est censé expliquer le fonctionnement de *Scripertexte*. Il donne ensuite des conseils concernant la nécessité d'épaissir le texte, c'est-à-dire de sélectionner des segments dans le corpus susceptibles d'améliorer le texte au plan du haut niveau. L'enseignant fait même une partie du travail, à savoir la révision du texte au plan du bas niveau. Je retiens de *Scripertexte* l'idée de faire rédiger un premier jet et de l'épaissir et réorganiser à l'aide des textes du corpus. Je rajoute la composante "réviser son texte au bas niveau" qui me paraît importante dans un contexte d'apprentissage d'une LE. Cette composante fait d'ailleurs également défaut à *Gammes d'écriture* (Mangenot, 1996), un logiciel qui, pourtant, est, entre autres, utilisé par des étudiants en FLE (Français Langue Étrangère).

Crinon et Legros (2000) expliquent que les logiciels d'écriture sont "des outils pour mieux écrire" à condition d'être des environnements ouverts qui s'insèrent dans un "dispositif d'apprentissage ou d'usage cohérents avec leurs fonctionnalités" et qui sont "incorporés à la culture écrite de leurs utilisateurs", c'est-à-dire s'ils servent à exécuter des tâches qui sont clairement expliquées aux apprenants et à l'exécution desquelles ils sont régulièrement entraînés. Crinon et Legros (2000) précisent que les résultats que permet d'obtenir l'utilisation de logiciels d'aide à la production écrite dépendent fortement de facteurs liés au contexte

pédagogique, à savoir un enseignement frontal ou d'autres formes de travail, et de facteurs liés aux sujets apprenants, par exemple leurs connaissances préalables et leur pratique de l'outil. Les auteurs constatent que l'outil d'aide à la production de l'écrit n'est pas en soi une aide – "le média en lui-même est neutre", disent-ils -, mais il peut favoriser une telle production quand il est utilisé au sein de dispositifs pédagogiques et d'approches didactiques permettant "une nouvelle manière d'écrire, plus 'publique', plus collaborative, plus intense, plus 'méta'". Il convient donc de créer un scénario d'apprentissage basé sur un ou des corpus et qui met en accord le contexte d'utilisation, l'outil matériel et l'outillage cognitif.

3.8. Répercussion des modèles et méthodes sur mon scénario pédagogique

Quelles répercussions les théories et méthodes exposées ci-dessus ont-elles sur le scénario que j'ai créé pour la recherche-action ? J'intègre à mon scénario les quatre phases du processus rédactionnel de Hayes et Flower (1980). Les travaux de Zimmermann (2000) et de Fayol (1996 : 34) montrent que la révision n'a pas toujours comme conséquence l'amélioration du produit. Il convient donc de proposer des aides qui permettent à l'apprenant de vérifier si la version modifiée correspond aux normes établies par le genre textuel qui fait référence pour la tâche. Une formulation non acceptable peut s'expliquer par la nativisation, donc par le fait que l'apprenant se laisse guider par des critères qui conviennent à d'autres langues ou interlangues du scripteur (Demaizière et Narcy-Combes, 2005). Si les connaissances linguistiques et discursives des scripteurs sont trop limitées, ils ont besoin d'un apport externe. L'accès à une base de données, par exemple à un corpus, est une des solutions possibles. Les corpus peuvent permettre aux scripteurs de modifier leur texte à l'aide d'unités lexicales ou de segments pertinents qu'ils y ont identifiés. Le contenu, terme évoqué par Fayol (1996) et par Bereiter et Scamardalia (1987), est une des composantes de la production écrite. Mon scénario propose une graduation des tâches en prenant comme critère la prise en charge du contenu. Certaines tâches, comme par exemple la production du texte "Informations pratiques" d'un musée, demandent peu de connaissances ou idées concernant le contenu. La manipulation des données linguistiques du corpus "Informations pratiques", donc la formulation – d'ailleurs très contrainte pour ce genre textuel -, est prioritaire ici. Il s'agit davantage de juxtaposer des informations que de les transformer (cf. *knowledge telling*, 3.1.2.). D'autres tâches, comme l'écriture du texte "Visite guidée" demandent un degré plus important de génération de connaissances et d'idées. Par contre, les contraintes au niveau de la formulation sont moins

importantes que pour le genre "Informations pratiques". Cette deuxième tâche sollicite les quatre phases de rédaction de façon équilibrée. La tâche "Commentaire dans le livre d'or électronique d'un musée" comporte peu de contraintes linguistiques. La phase de révision peut être plus courte pour cette tâche que pour les deux tâches précédentes. Les deux dernières tâches demandent une transformation des données identifiées dans les corpus par l'apprenant (cf. *knowledge transforming*, 3.1.2.).

Précisons enfin quelques caractéristiques du scénario pédagogique que j'ai développé pour un public d'apprenants d'allemand de niveau intermédiaire. Ce scénario vise à respecter certaines conditions, évoquées précédemment, concernant l'*input* qui est mis à disposition des apprenants et concernant le guidage qui leur est offert lors de l'exécution du projet.

3.8.1. Input *et scénario pédagogique*

Le tableau suivant regroupe les caractéristiques de l'*input* que je propose lors de mon projet. J'y nomme aussi les outils qui permettent de fournir et les phases de rédaction lors desquelles cet input a un rôle à jouer.

Caractéristiques de l'*input*	**Outils**	**Phases identifiables**
Le scénario doit s'appuyer sur une quantité importante d'*input*.	Ceci est possible grâce aux corpus.	
L'*input* doit ressembler au résultat de la tâche de production que l'apprenant doit exécuter.	Ceci est possible grâce aux corpus réfléchis qui regroupent des textes d'un même genre textuel.	Planification.
La complexité de l'*input* doit pouvoir être réduite.	Ceci est possible grâce aux concordanciers.	Formulation, révision.
L'*input* doit pouvoir être exploré.	Ceci est possible grâce aux outils d'exploration des concordanciers et grâce à la fonction du traitement de texte ou du navigateur "rechercher dans ce texte/dans cette page...".	Planification, formulation, révision.
L'*input* doit pouvoir être manipulé.	Ceci est possible grâce à la fonction "copier-coller" du traitement de texte.	Formulation, révision, exécution.
L'*input* doit être libre d'accès sans restriction ni de temps, ni d'espace.	Ceci est possible grâce à la mise à disposition du corpus et de concordances sur une plateforme numérique de travail ou sur un blogue. Un corpus en ligne peut également convenir.	Planification, formulation, révision.

Tableau 12 – *Input*, outils et phases de rédaction.

3.8.2. Guidage et scénario pédagogique

Il est possible, dans une certaine mesure, d'apprendre par observation et imitation (Bruner, 1983 : 263). Dans le domaine de la production écrite, "activité culturellement complexe" (Vygotski, 1931 : 145, cité dans Brossard, 2004 : 100), un guidage humain paraît bénéfique. Peu de personnes réussissent à apprendre à écrire sans être accompagnées pendant plusieurs années par un expert, la plupart du temps en institution (Fayol, 1996 : 35). Demaizière (2004) évoque des principes de guidage applicables à un travail pédagogique basé sur les TIC (Technologies de l'Information et de la Communication), et Bruner (1983 : 261-280) décrit les caractéristiques d'un guidage qui favorise un apprentissage de type constructiviste (cf. 2.6.). Les principes de Demaizière (2004) et de Bruner (1983) sous-tendent le scénario que j'ai développé pour la recherche-action. Je confronte ci-dessous quatre composantes du processus de soutien proposé par Bruner (1983 : 278-279) aux caractéristiques de mon scénario.

Il s'agit d'engager les apprenants dans la tâche.

Pour le projet principal de la recherche (cf. 5.3.4.), un tel engagement est recherché par le biais de la liberté des apprenants par rapport au thème du musée pour lequel ils sont censés créer des textes pouvant s'insérer dans un dépliant ou site touristique. Les tâches proposées lors de la dernière phase de la recherche-action (cf. 5.3.5.) se caractérisent par une prise en charge énonciative forte du discours. Il s'agit d'écrire une annonce comme celles qui aident à trouver un correspondant. Dans un cadre d'approche par tâches, proposer aux apprenants de publier leurs textes en ligne peut être un facteur de motivation.

Le guidage doit réduire les degrés de liberté des apprenants, si nécessaire.

La mise en forme matérielle de l'*input* à l'aide de concordances permet de réduire sa complexité. L'emprunt d'éléments dans l'*input* permet de s'appuyer sur des formulations qui sont déjà là.

L'enseignant doit orienter les apprenants quand ils risquent de se diriger dans des directions peu favorables à la réalisation de la tâche.

Lors de la recherche-action, les apprenants ont l'occasion de publier des commentaires sur le blogue du projet ou de m'envoyer des courriels. J'annote les diverses versions des productions écrites, et je donne des conseils concernant la démarche de révision.

Le guidage doit viser à faire comprendre aux apprenants les caractéristiques de la tâche et les écarts entre celle-ci et le produit réalisé.

Des démonstrations à l'aide d'un vidéoprojecteur ou à l'aide de feuilles de travail permettent de montrer comment on peut identifier, sauvegarder et modifier un passage pertinent. Il convient également de montrer au groupe-classe comment réviser un texte à l'aide de concordances.

Les chapitres suivants donneront un aperçu théorique et méthodologique sur le domaine des corpus et concordances.

4. Enseigner/apprendre une langue étrangère à l'aide de corpus et de concordances

Mukherjee (2002) et O'Keeffe *et al.* (2007) incitent à une utilisation de corpus par des enseignants et apprenants de langue étrangère du secondaire. Les corpus seraient potentiellement porteurs d'évolutions des pratiques et des attitudes d'enseignement/apprentissage. En effet, la linguistique de corpus remet en question non seulement les intuitions des enseignants et apprenants concernant les normes sémantiques, grammaticales, discursives, etc., mais aussi leurs représentations concernant la façon dont il convient d'enseigner ou d'apprendre une langue étrangère (O'Keeffe, McCarthy & Carter, 2007 : 20 ; Braun, 2007).

L'utilisation de corpus dans le secondaire est encore un phénomène rare. Braun (2007) et Rohrbach (2003) présentent des études empiriques faites auprès d'un public d'apprenants d'anglais du secondaire. Boulton (2007) cite deux autres études faites auprès de jeunes apprenants d'anglais, celle de Ciezielska-Ciupek (2001) et celle de Sun et Wang (2003). Les recherches de St John (2001) et de Chambers (2005) décrivent l'utilisation de corpus par des apprenants d'allemand, langue étrangère, à l'université. Yoon (2008) a fait utiliser des corpus lors d'un projet de production de textes en anglais académique. Je n'ai pas eu connaissance d'études concernant ma combinaison "utilisation de corpus et concordances par des apprenants d'allemand dans l'enseignement secondaire lors de tâches de production écrite". Dans ce chapitre, je décrirai d'abord les caractéristiques des corpus pouvant convenir à ma recherche. Ces caractéristiques dépendent des objectifs d'utilisation (cf. 1.4.). Je présenterai ensuite des corpus allemands existants, et j'expliquerai pourquoi et comment j'ai créé de très petits corpus réfléchis. La confrontation entre les performances d'un de ces petits corpus et des performances de trois corpus existants clora le premier sous-chapitre. La deuxième partie de ce chapitre s'intéressera aux tâches basées sur des concordances.

4.1. Quels corpus pour des tâches de production écrite ?

Les caractéristiques des corpus et des outils d'exploration de corpus doivent être adaptées au public, à l'utilisateur et aux objectifs d'apprentissage (Braun, 2007). Dans le cas d'un public d'adolescents de 13 à 14 ans, la prise en main des outils doit être particulièrement facile afin

de réduire les obstacles liés à une littératie électronique parfois limitée.[15] Les connaissances linguistiques des apprenants peuvent aussi être encore relativement modestes, ce qui risque d'empêcher certains de trouver leurs repères dans un environnement textuel d'apparence complexe, nécessitant l'adoption de nouveaux modes de lecture. Le contexte de l'enseignement secondaire impose également une gestion des coûts et nécessite des outils gratuits ou peu onéreux qui protègent suffisamment les données.

4.1.1. Quels objectifs pour une utilisation de corpus ?

Le corpus est un outil parmi d'autres que l'on peut proposer à des apprenants d'une langue étrangère. Il doit répondre à un objectif d'utilisation précis, comme le suggère Sinclair (1996). Lors de ma recherche-action, je privilégie la recherche d'informations, l'étude de l'espace des normes et l'analyse linguistique (cf. 1.4.2.). Ces objectifs sont susceptibles de soutenir une démarche d'observation et d'exploration de la langue. Quels types de corpus convient-il de solliciter dans le cadre d'une tâche d'écriture ? J'écarte pour mon projet les corpus de référence, les corpus d'apprenants, les corpus oraux et les corpus d'échantillons (cf. 1.2.4.). Je m'intéresse particulièrement aux corpus spécialisés, contenant des textes complets, représentant le même genre textuel. Ceci permet de travailler avec des données dont on sait avec une certitude assez grande qu'elles recouvrent un domaine précis. Les corpus comparables ou les corpus parallèles, souvent sollicités dans les domaines de la traduction et de la lexicologie, pourraient également trouver leur place au sein d'un projet d'écriture en langue étrangère (Narita, Kurokawa et Utsuro, 2003). Je n'ai toutefois pas eu accès à des corpus mettant en parallèle des textes français et allemands et représentant les domaines que j'ai choisis pour la recherche-action (tourisme, critiques de film, annonces pour trouver un correspondant). Le corpus parallèle *English-German Translation Corpus* a trait au domaine du tourisme. Toutefois, il juxtapose l'allemand et l'anglais, et il a certains inconvénients que je détaillerai en 4.1.2.

4.1.2. Besoins concernant les caractéristiques des corpus

Quand on planifie un projet nécessitant un ou des corpus, il convient de clarifier les priorités concernant la taille, le contenu et la facilité d'utilisation. Quel est l'avantage de très grands corpus par rapport à des petits corpus spécialisés ? Chambers (2005), Bernardini (2000),

[15] J'ai pu observer un réel changement en ce qui concerne la *media literacy* à partir de 2008/09. Les élèves qui ont participé à ma recherche-action cette année-là ont presque tous été très à l'aise avec l'outil informatique, contrairement à la majorité des élèves des groupes de 2006/07 et 2007/08 pour lesquels la prise en main des outils a été beaucoup plus longue.

Cheng, Warren et Xun-feng (2003) privilégient les très grands corpus, donc ceux qui dépassent cinq millions de mots. À leur avis, en-dessous d'une certaine quantité de données, un corpus n'est pas utilisable pour l'étude de la langue. Une taille trop petite ne permettrait pas de travailler selon la méthode inductive. Il est, par contre, admis que les très grands corpus conviennent moins aux débutants que les petits corpus (Chambers, 2005). Dodd (1997 : 131) plaide pour l'utilisation de "petits" corpus généraux, contenant toutefois au moins un million de mots. Tribble (1997) et Aston (2002) donnent des arguments en faveur de l'utilisation et de la création de corpus de petite taille. Aston (2002) conseille de travailler avec des corpus tout faits de très grande taille et d'en extraire des sous-corpus, selon les besoins. Il paraît, en effet, judicieux de tirer profit du travail effectué par des professionnels du domaine de la création de corpus, tout en ayant la possibilité de cocher des sous-corpus, d'éliminer ou de retravailler certains textes, d'en rajouter d'autres, etc., selon le public d'apprenants et la tâche que l'on lui propose.

4.1.3. Critique de corpus allemands existants

Il y a plusieurs corpus anglais qui correspondent aux exigences décrites ci-dessus, comme par exemple le *BNC Corpus*, accessible par l'interface de *Sketch Engine* (Kilgarriff *et al.*, nd). Actuellement, peu de corpus allemands donnent l'occasion de sélectionner des sous-corpus au sein d'un corpus rassemblant des textes de plusieurs domaines de spécialité. Ceci est possible, dans une certaine mesure, avec le *DeReKo* (*Deutsches Referenzkorpus*) de *l'Institut für Deutsche Sprache*. Pour des raisons évoquées en 1.8., ce corpus ne convient pas pour mon scénario. *LIMAS* (*Linguistik und Maschinelle Sprachbearbeitung*), maintenant un sous-corpus de *DeReKo*, regroupe environ 500 documents écrits de tous genres (textes pleins ou extraits). Trois types de recherche sont proposés : recherche de lexèmes, de segments et d'une combinaison entre un mot et un autre, sans obligation que les deux mots ne se suivent. Les sources d'où proviennent les mots ou *patterns* recherchés peuvent être consultées également. L'emploi de ce corpus paraît suffisamment simple pour une utilisation pédagogique. Toutefois, ce corpus est clos depuis 1970. L'allemand très récent n'y est donc pas attesté. Il représente une multitude de genres (langues de spécialité, textes utilitaires, belles lettres). Toutefois, il n'est pas possible de faire des recherches limitant les résultats à un genre textuel ou à un domaine donné. Le *Braun Corpus* (Divsic *et al.*, 2008) a été mis en place par Divsic, en collaboration avec Rott et Horst de *Concordia University* de Montréal. Ce "mini-corpus de référence" suit le modèle du *Brown Corpus* (nd) en ce qui concerne la variété de genres représentés, mais il n'est pas annoté. Il contient un million de mots environ, et il est

consultable à l'aide du concordancier *Konkordanzer Corpus Deutsch* (Cobb, nd). Une grande variété de types de textes y est représentée. Nous y trouvons les catégories "classiques", telles que la presse, le monde du travail et la fiction, mais il y a également une rubrique nommée "*Humour*", les rubriques "*Skills and hobbies*" et "*Popular lore*". Ces dernières semblent particulièrement utiles pour le domaine de l'enseignement/apprentissage d'une langue étrangère dans le secondaire. La rubrique "*Skills and hobbies*" contient, entre autres, des données provenant du site *Die Zoo-Datenbank für Deutschland* (Petzold, nd). Dans le *Braun Corpus* (Divsic *et al.*, 2008) il n'est pas possible de trier les résultats des requêtes selon les genres textuels. On peut cocher les lignes contenant un même mot clé afin de les exporter. En cliquant sur un mot clé donné, son contexte plus large (environ 160 – 230 mots) s'affiche, mais contrairement à *deWac*, la source n'est pas indiquée, et l'on n'a pas accès aux textes entiers. Le corpus parallèle *English-German Translation Corpus* de l'université de Chemnitz propose un classement fonctionnel des textes qu'il contient. Les quatre sous-corpus (*Public Speeches and Articles, Academnic Texts, Tourist brochures, EU documents)* rassemblent des discours et articles publics, des textes en anglais et allemand académiques, des brochures touristiques et des documents de l'union européenne. Le corpus peut être partiellement interrogé à partir du site *Chemnitz Internet Grammar*. On peut cocher et décocher des sous-corpus. L'accès aux sources n'est pas proposé. En 2009, il est uniquement possible de faire des requêtes de lexèmes qui s'affichent ensuite dans un contexte de phrases entières. L'option *keywords in context*, ici définie comme l'affichage d'un mot clé dans un contexte de trois à neuf mots, semble encore en développement. On ne peut pas non plus faire des requêtes à l'aide de commandes booléennes.

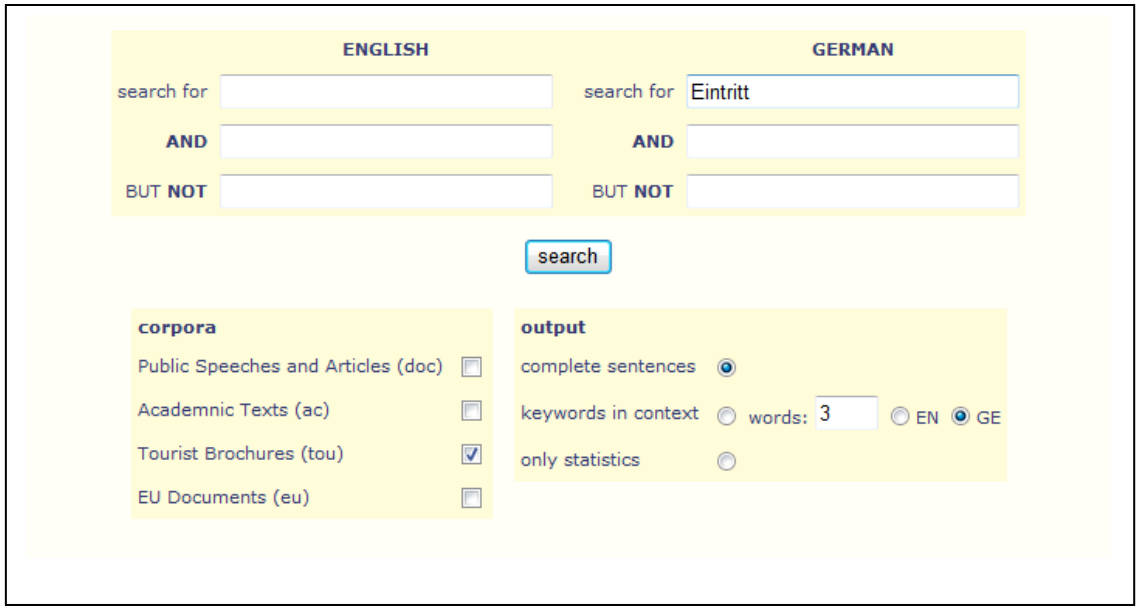


Figure 26 – Interface de *English-German Translation Corpus*.

Le corpus *deWac*, interrogeable à l'aide de *Sketch Engine*, est un outil suffisamment simple d'emploi. Il est possible de créer des sous-corpus à partir de noms de sites sollicités pour la création de *deWac*. Il est toutefois difficile de maîtriser le contenu. En plus, l'accès à ce corpus est payant.

Chacun des corpus que je viens de passer en revue a certains inconvénients. La langue qui y est représentée n'est pas assez actuelle, ou le corpus est trop général. Les requêtes sont soit techniquement trop difficiles à faire, soit elles sont trop restreintes. Si l'on n'a pas besoin de faire des requêtes basées sur des données annotées, la création de petits corpus spécialisés, peut donc être justifiée. Quel procédé choisir afin d'éviter que la création de tels corpus ne demande ni trop de compétences techniques, ni trop de temps ?

4.1.4. Création de petits corpus spécialisés

Je m'intéresse aux corpus qui représentent l'allemand écrit récent. Il serait utile d'avoir à disposition un très grand corpus qui permettrait de sélectionner des sous-corpus concernant les genres textuels adaptés au public de la recherche-action. Comme un tel corpus n'existait pas au moment de la planification de ma recherche, j'ai choisi de créer des corpus à partir de textes disponibles sur Internet, protégés par un *copyright*, mais gracieusement mis à ma disposition par les propriétaires des sites concernés. Pour créer un corpus réfléchi, il convient d'utiliser un moteur de recherche permettant de trouver des sites qui cadrent bien avec un genre textuel donné. Les textes ainsi trouvés peuvent être récupérés par l'option "copier-coller" (option que j'ai retenue) ou par l'utilisation d'un aspirateur de sites, comme par exemple le logiciel *HHTrack*. Une option intéressante est la possibilité de créer des corpus spécialisés à l'aide de l'outil *BootCat* (Baroni et Bernardini, 2003) proposé par *Sketch Engine*, qui rassemble des données de la Toile en fonction de mots clés spécifiés par le créateur de corpus. Les sites contenant les mots clés peuvent être cochés ou décochés. Les lignes de concordance peuvent être exportées grâce à la fonction "copier-coller". Il n'est pas possible de cocher des lignes en particulier. Quand on clique sur le numéro en début de ligne, l'adresse de la page d'où provient le passage s'affiche. En cliquant sur le mot clé, on a accès à un contexte plus large que l'on peut encore élargir davantage soit à gauche soit à droite du mot clé (pour une description détaillée, voir Baroni et Bernardini, 2004). La création d'un corpus spécialisé à l'aide de *BootsCat* a des avantages indéniables, comme la rapidité, la grande quantité de données et la possibilité de décider quelles données on souhaite garder ou éliminer. Lors de la première phase de *Projekt Prospekt*, j'ai créé un corpus du domaine du tourisme en me servant

de cet outil. J'ai choisi des mots clés que j'ai estimé pertinents pour le domaine du tourisme, comme, par exemple "*Tourismus, Eintritt, Kinder, Erwachsene, Behinderte, Führung*" [tourisme, entrée, enfants, adultes, handicapés, visite guidée]. Après une période d'essai, j'ai toutefois renoncé à son utilisation, car la prise en main des outils d'interrogation de ce corpus et la gestion des occurrences très nombreuses ont été un obstacle assez important pour les apprenants. L'appartenance au genre touristique n'a pas été garantie pour toutes les occurrences. J'ai été contrainte de faire un travail important de présélection d'occurrences pertinentes. J'ai donc abandonné ce type de corpus au courant de ma recherche.

4.1.5. Corpus, outils et objectifs retenus

Lors de la deuxième phase du projet *Projekt Prospekt*, j'ai fait le choix de créer trois corpus non annotés qui sont de très petite taille (entre 3 400 et 9 800 mots environ). Ceci m'a permis de travailler avec des données relativement "maîtrisées". J'ai pu vérifier manuellement le contenu et l'orthographe. J'ai obtenu les droits d'utilisation des textes en m'adressant aux gestionnaires des sites concernés. J'ai créé des concordances pour ces corpus à l'aide du concordancier *Text-based concordances (v2)* de Cobb (nd), conçu pour l'anglais. Avant fin 2008, cet outil n'a pas suffisamment bien géré les voyelles allemandes ayant un tréma (ä, ö, ü). Elles ont été remplacées soit par des sigles, soit par une majuscule (Ä, Ö, Ü). Ceci a posé des problèmes de lisibilité de la concordance. Ces désagréments ont été largement compensés par les avantages de la prise en main qui m'a paru particulièrement facile.

Rastier (2004) recommande l'indication d'une triple information dans l'en-tête des corpus, à savoir discours, champ générique et genre proprement dit. Les textes des trois corpus de *Projekt Prospekt* peuvent être décrits selon ces critères de typologie. Le champ générique des corpus de *Projekt Prospekt* est celui du tourisme. Les genres et discours représentés dans ce projet sont les "Informations pratiques" (discours informationnel), la "Visite guidée pour jeunes" (discours publicitaire) et le "Commentaire dans un livre d'or électronique" (discours de type épistolaire). Les corpus présentent une graduation qui part d'un discours avec désengagement énonciatif fort pour arriver à un discours avec engagement énonciatif fort (d'après Bronckart, cité dans Portine, 2008). Pour d'autres projets réalisés lors de ma recherche-action, j'ai créé des corpus dont les textes représentent des genres textuels avec prise en charge énonciative forte, à savoir un corpus de critiques de film et un corpus d'annonces pour trouver un correspondant.

"Informations pratiques" (dépliant d'un musée)	Prise en charge énonciative faible
"Visite guidée" (dépliant d'un musée)	
"Critique de film"	
"Commentaire dans le livre d'or d'un musée"	
"Annonce pour trouver un correspondant"	Prise en charge énonciative forte

Tableau 13 – Genres textuels et prise en charge énonciative.

4.1.6. Mise à l'épreuve du corpus "Informations pratiques"

Afin de tester la fiabilité d'un des corpus spécialisés que j'ai créés, le corpus "Informations pratiques", je présente ci-dessous la comparaison des résultats que l'on obtient quand on fait une requête pour le mot clé *Eintritt* [entrée], à partir du *English-German Translation Corpus*, du *Braun Corpus*, du corpus *deWac* et à partir du corpus "Informations pratiques". J'ai créé ce dernier à l'aide du site *Die Zoo-Datenbank für Deutschland*. [Banque de données pour les zoos en Allemagne].

Dans le *Braun Corpus*, la requête du mot clé "*Eintritt*", option "*contains Eintritt*" [contient *Eintritt*] permet de trouver 38 occurrences. 13 occurrences correspondent au genre "Informations pratiques". Il est possible de cocher les lignes que l'on souhaite retenir et de les exporter ensuite. Sept de ces lignes sont particulièrement utiles en vue de notre projet d'écriture. On peut y repérer comment on dit "Prix d'entrée : 5 € pour les enfants" ou comment on dit "ticket d'entrée", etc. Certaines des 38 occurrences proviennent de textes qui appartiennent à un autre genre que celui de "Informations pratiques". Le temps qui y est utilisé est le passé, le contexte paraît narratif davantage qu'informatif, comme l'illustrent les deux lignes suivantes.

und Kassierer zugleich. 20 Pfennig kostete der EINTRITT für Kinder, für Erwachsene 50.
[et en même temps vendeur. L'ENTRÉE a coûté 20 Pfennig pour les enfants, 50 pour les adultes].
murmelte er in seinen Bart, "wer hat denen die EINTRITTskarten gegeben?"
[murmura-t-il dans sa barbe, "mais qui est-ce qui leur a donné les tickets d'ENTRÉE ?"].

Dans le *English-German Translation Corpus*, la requête "*Eintritt*" permet de trouver 11 occurrences. Toutes – sauf peut-être la première qui se réfère à l'entrée physique d'un bâtiment – sont pertinentes en vue de *Projekt Prospekt*. Nous trouvons parmi les occurrences

un mot composé, "*Eintritt*spreise" [prix d'entrée]. Le corpus *deWac* affiche 46 130 occurrences du mot clé *Eintritt* et de mots contenant ce mot clé. Nous obtenons trop d'occurrences pour un observateur humain. Une requête plus restrictive, "*Eintritt*" en combinaison avec "*frei*" [libre] dans le contexte de droite, permet d'obtenir 5 849 d'occurrences pertinentes, mais il y a redondance. Une telle quantité de données pertinentes, très appréciable lors d'une étude linguistique sur corpus, peut paraître prohibitive dans un contexte scolaire. Le corpus "Informations pratiques" permet de trouver 117 occurrences du mot clé "*Eintritt*". Elles sont toutes pertinentes, mais il y a redondance. Après suppression des lignes redondantes, il reste 14 occurrences appartenant toutes au genre "Informations pratiques".

	Braun Corpus	*English-German Translation Corpus*	"Informations pratiques"	*Sketch Engine*
nombre d'occurrences de *Eintritt*	38	11	117	46130 pour *Eintritt* 5849 pour *Eintritt + frei (right context*), mais il y a redondance
occurrences utiles de *Eintritt*	11 après élimination des occurrences non pertinentes	10 après élimination d'une occurrence non pertinente	14 après élimination des lignes redondantes	au moins 5000

Tableau 14 – Nombre d'occurrences pertinentes pour "*Eintritt*" [entrée] dans quatre corpus.

La comparaison du *English-German Translation Corpus*, *Braun Corpus*, du corpus *deWac* et de mon corpus "Informations pratiques" permet de tirer les conclusions suivantes. Le *English-German Translation Corpus* a l'avantage de contenir un corpus spécialisé dans le domaine du tourisme. Ceci permet d'obtenir des résultats pertinents pour notre requête. Toutefois, les types de requêtes applicables à ce corpus paraissent limités en 2009. Le *Braun Corpus* propose des options de recherche élaborées. Par contre, pour obtenir des occurrences qui représentent un genre textuel donné, les lignes pertinentes du *Braun Corpus* doivent être présélectionnées par l'enseignant. Le corpus *deWac* permet de trouver une quantité impressionnante d'occurrences pertinentes, mais à condition de choisir une combinaison de mots clés restreignant suffisamment le nombre de lignes. Une présélection des lignes par l'enseignant semble également indispensable pour une utilisation de ce corpus. Le corpus "Informations pratiques" permet de trouver un nombre raisonnable d'occurrences de *Eintritt*

qui sont toutes pertinentes. Certes, il y a redondance, mais la redondance peut être porteuse de signification. La fréquence renseigne les utilisateurs de la concordance sur l'importance d'un mot clé ou d'une formulation. Afin de créer une concordance complète à partir du corpus "Informations pratiques", j'ai utilisé le concordancier *Text-Concordancer (v2)*, proposé par Cobb (nd) sur *Compleat lexical tutor*. Cet outil permet, après importation du corpus, de faire apparaître tous les mots du corpus comme mots clés, classés dans l'ordre alphabétique. Ils s'affichent dans un contexte d'environ 90 caractères. L'unique outil d'exploration est la fonction "rechercher dans cette page" du traitement de texte. Bien sûr, une exploration des mots clés en suivant l'ordre alphabétique est également possible. Au courant de la phase finale de la recherche-action, j'ai découvert la collection d'outils *AntConc3.2.1w* (Anthony, 2007) qui permet, comme le logiciel de Cobb, d'interroger son propre corpus. *AntConc3.2.1w* propose, de plus, plusieurs options intéressantes, comme, par exemple, l'affichage du texte source par simple clic de souris et la possibilité de comparer deux corpus entre eux. J'opterai pour cet outil lors d'un futur un projet basé sur l'utilisation d'un concordancier.

4.2. Tâches basées sur des concordances

Depuis les années 80, nous disposons d'un certain nombre d'études qui décrivent le potentiel des concordances numériques et, dans une moindre mesure, des corpus, pour enseigner et apprendre une langue étrangère. Quels concepts méthodologiques sous-tendent le travail avec ces outils ?

4.2.1. Méthodologies convoquées

Plusieurs méthodologies sont évoquées en lien avec l'utilisation de corpus et de concordances pour l'apprentissage des langues. Il s'agit de l'apprentissage par la découverte, de la grammaire pédagogique, de l'approche par tâches et de l'apprentissage basé sur des données, généralement appelé *Data Driven Learning (DDL).* Ce sont des approches constructivistes dont certaines, comme la grammaire pédagogique, encouragent les apprenants à observer la langue au lieu de partir de règles grammaticales toutes faites. Hadley (2002) parle dans ce contexte d'"'*autonomous discovery' much more than 'being told about [grammar]'*". L'observation de la langue est également proposée par la démarche conceptualisante. Cette démarche invite les apprenants à observer un environnement langagier et à créer leurs représentations individuelles de la langue, en remettant en cause leurs représentations antérieures et en les modifiant, si nécessaire. Dans les articles et ouvrages que j'ai pu consulter, la conceptualisation n'a pas été mentionnée en rapport avec un travail basé sur des

concordances. Pourtant, la combinaison "concordances et conceptualisation" paraît prometteuse. D'autres méthodologies, comme l'apprentissage par tâches, incitent les apprenants à appréhender l'*input* dans le but de créer un produit. Nous trouvons ici un lien avec l'idée d'intertextualité. Les concordances donnent l'occasion d'interroger, de manipuler et de modifier le texte ; elles soutiennent l'idée l'on a "le droit" de se servir de l'*input* pour écrire (Mpartusa *et al.*, 1991).

<table>
<tr><td>Concept</td><td colspan="2">Constructivisme</td></tr>
<tr><td rowspan="2">Méthodologie</td><td>Plutôt orienté bottom up :

noticing</td><td>Plutôt orienté top down :

approche par tâches
apprentissage incident</td></tr>
<tr><td colspan="2">Va-et-vient entre des mouvements bottom up et top down :

Data Driven Learning
apprentissage par la découverte
grammaire pédagogique/conceptualisation
apprenant-chercheur</td></tr>
</table>

Tableau 15 – Méthodologies associées à l'utilisation de corpus et concordances pour l'enseignement/apprentissage d'une langue étrangère.

Johns et King (1991) associent étroitement l'utilisation de corpus et de concordances pour l'enseignement/apprentissage des langues étrangères et le *DDL* (*Data Driven Learning*). Cette méthodologie, déjà évoquée en 2.6.3., semble aller de pair avec le recours aux concordances, comme le suggère la citation suivante, trouvée chez Mukherjee (2005).

> *Johns and King (1991: iii) define DDL as use in the classroom of computer-generated concordances to get students to explore regularities of patterning in the target language, and the development of activities and exercises based on concordance output*.

Le *DDL* s'appuie sur un *input* qui, en général, a été produit par des locuteurs experts. L'accès à cet *input* peut être facilité par les corpus et concordances. Il doit être le plus direct possible si l'on se place dans une perspective d'apprentissage par la découverte (Bruner, années 60, Kaltenböck et Mehlmauer-Larcher, 2005, Chambers, 2005). Dans ce contexte, certains attribuent aux apprenants le rôle d'"apprenants-chercheurs" (Johns et King, 1991 ; Stevens, 1995 ; Bernardini, 2004 ; Chambers, 2005). En 2002, Johns affine ce concept en proposant le terme de "*detective, learning to recognise and draw conclusions from clues in the data*" (Johns, 2002 : 108). L'utilisation de concordances favorise parfois un apprentissage incident

(Bernardini, 2004) ; c'est-à-dire, les apprenants font des découvertes inattendues (cf. 2.2.1.). Siepmann (1999) précise que les concordances sont un terrain d'exploration de "sujets" très variés qui peuvent émerger au fil de l'exécution d'une tâche d'apprentissage.

> On ne soulignera jamais assez que, outre des sujets évidents tels que la position des adjectifs ou celle des adverbes, les recherches de l'apprenant peuvent porter sur n'importe quel phénomène rencontré au hasard d'une lecture ou d'un exercice : la combinatoire de tel verbe, l'usage de telle préposition, l'emploi des pronoms relatifs, etc. Le concordancier s'avère donc un outil précieux favorisant l'éveil au langage (language awareness) et permettant aux apprenants avancés de trouver réponse à toutes leurs questions (Siepmann, 1999).

Pour certains, le *DDL* va de pair avec un regain d'intérêt pour le travail sur la forme (et non sur "les formes", comme dirait Long, 1997). Ce domaine a été mis à l'écart lors de l'approche communicative (Hadley, 2002). Hadley considère le *DDL* comme une application de la grammaire pédagogique (cf. 2.6.3.) : "*Special attention will then be given to one practical application of PG, which is known as Data-Driven Learning (DDL)*" (Hadley, 2002). Siepmann (1999) précise comme suit le lien entre linguistique de corpus et grammaire pédagogique.

> Pour terminer, rappelons que la linguistique de corpus, tout en permettant de résoudre des problèmes séculaires liés à l'enseignement des langues vivantes, ne saurait avoir la prétention de constituer un nouveau paradigme, ni à plus forte raison de poser, à elle seule, les fondements d'une nouvelle grammaire pédagogique. Cela, avons-nous dit, parce qu'elle n'est pas une discipline linguistique autonome, mais un ensemble de techniques et de méthodes permettant une nouvelle approche des faits de langue. On sait d'ailleurs que, vu la diversité des facteurs à prendre en compte, la prudence est de mise lorsqu'il s'agit d'établir un rapport entre description linguistique et description pédagogique d'une langue (cf. Bausch 1979, Dirven 1990, Mindt 1995b). Ces mises au point étant bien entendues, il est clair que le lexique-grammaire présente l'avantage d'être compatible avec tous les grands systèmes conceptuels et donc, a priori, avec une description de la langue jugée appropriée par les didacticiens. La grammaire pédagogique, croyons-nous, voit ainsi s'ouvrir des horizons plus larges.

L'utilisation de corpus et de concordances et l'accès direct aux données semblent être motivés par l'espoir de pouvoir ainsi établir une relation plus égalitaire que hiérarchisée entre apprenants et enseignants. La manipulation des données d'un corpus, l'utilisation d'un concordancier pour présenter les données dans des contextes plus ou moins restreints et la réutilisation des données dans une tâche sont des activités qui favorisent, en effet, une attitude réflexive aussi bien de l'enseignant que des apprenants. Selon Stevens (1995), les explorateurs de corpus ont l'occasion de mettre leur savoir sur la langue à l'épreuve. Pour ce faire,

l'enseignant doit accepter de se mettre en retrait par rapport à de nouvelles sources de savoir. Une équation entre le *DDL*, l'utilisation de corpus et des approches exploratoires et réflexives semble se dessiner dans les propos présentés ci-dessus. Malgré l'intérêt que semble présenter une telle combinaison, les corpus ne sont pas (encore ?) très présents dans le milieu du secondaire. Braun (2007) explique cela, entre autres, le fait que de nombreux corpus ont été créés pour des linguistes et non pour des apprenants. Les corpus sont simplement des données qui peuvent être interrogées à l'aide d'outils, permettant de varier la mise en forme matérielle. Ils ne sont pas d'emblée des ressources pédagogiques et doivent être adaptés aux situations d'enseignement/apprentissage spécifiques.

Quelles activités basées sur des concordances sont compatibles avec les principes méthodologiques que je viens d'esquisser ? Dans ce qui suit, je développerai particulièrement le lien entre l'utilisation de corpus et de concordances et l'approche par tâches, une combinaison qui semble encore peu étudiée jusqu'à présent.

4.2.2. Macrotâches, microtâches et data-driven learning

Dans sa thèse sur l'approche par tâches, Nicol-Benoît (2004a) distingue entre exercice, activité et tâche. Selon sa définition, l'exercice est un travail non-contextualisé sur les formes. L'activité permet une utilisation de la langue en situation contextualisée, et la tâche favorise l'utilisation autonome de la langue dans une "situation authentique ou du moins vraisemblable" (Nicol-Benoît, 2004a : 342). Les tâches sont censées favoriser l'apprentissage incident. Les apprenants se concentrent sur le sens davantage que sur la forme quand ils échangent entre pairs ou avec d'autres personnes associées à la formation. Durant l'exécution de ce projet, des besoins individuels d'apprentissage linguistique peuvent se faire sentir. L'enseignant guidera alors les apprenants vers des activités plus pointues, des microtâches, susceptibles de les aider à surmonter un obstacle qui risque d'entraver la communication (Narcy-Combes et Walski, 2004). Il peut s'agir d'activités de repérage, de mémorisation, etc. (Demaizière & Narcy-Combes, 2004).

Les microtâches correspondent à ce que Portine (2008) appelle "activités pédagogiques". Ces activités sont censées soutenir l'apprentissage. Portine les distingue clairement de l'activité langagière, donc de ce qui se passe dans le cerveau des apprenants lors du processus d'apprentissage. Pour décrire certains éléments du processus cognitif, Portine (2008) se réfère aux travaux de Norman et Rumelhart (1975). Ces travaux décrivent, entre autres, les

opérations de haut et de bas niveau. Les opérations de haut niveau sont des opérations holistiques lors desquelles le traitement de l'information est dirigé par des concepts. On parle de *conceptually-driven processing* et de mouvement *top down*. Les opérations de bas niveau sont des opérations lors desquelles le traitement de l'information est dirigé par les données. On parle de *data-driven processing* et de mouvement *bottom up*. Ces opérations sont analytiques. Elles portent sur une partie isolée de l'objet "langue" (une lettre, syllabe, un mot, un segment, etc.). Selon Rumelhart (1977), lors de l'apprentissage, les mouvements de *top down* et de *bottom up* interviennent en alternance. Portine (2008) précise que le mouvement *top down* est premier. Les apprenants se concentrent particulièrement sur le bas niveau (*bottom*) quand ils se trouvent en situation d'insécurité (cf. 6.3.1.4. et 6.8.1.). Ce niveau évolue lors du processus d'apprentissage. Ce qui évolue dépend du niveau de l'apprenant : lettres ou syllabes, mots, relations entre les mots, etc. Afin d'illustrer le va-et-vient entre ces deux mouvements, je m'appuie sur le modèle suivant de Portine (2008) qui décrit trois types de perception de l'objet "langue".

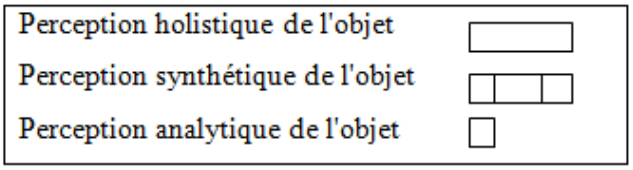


Figure 27 – Perception holistique, synthétique et analytique de l'objet "langue".

Une perception synthétique permet de se rendre compte des constituants de l'*input*, tout en gardant en mémoire sa présentation globale. Lors de la deuxième phase de la recherche-action (2007/08), j'ai proposé une activité qui vise à soutenir un tel regard synthétique (cf. 6.2.2.2.).

La possibilité qu'offrent les concordanciers de restreindre la taille du contexte va dans le sens d'une démarche analytique. Vu ainsi, observer des concordances revient à se placer au niveau du mouvement *bottom up*. Par ailleurs, Chambers (2005), Bernardini (2004) et Rüschoff (nd) évoquent la pédagogie du projet et l'approche par tâches comme arrière-plans méthodologiques pour justifier l'utilisation de concordances. Ces deux approches accentuent le mouvement *top-down* : elles cherchent à contextualiser le travail d'apprentissage et à créer un lien fort entre les tâches pédagogiques et l'environnement social des apprenants. La combinaison entre la pédagogie du projet ou l'approche par tâches et le travail sur corpus permet de mobiliser les représentations des apprenants (mouvement *top down*) pour les confronter aux traces matérielles de la langue (mouvement *bottom up*). L'observation des concordances, motivée par un projet de production, peut provoquer des situations de

déstabilisation et de réajustement des représentations. On crée ainsi des occasions d'apprentissage.

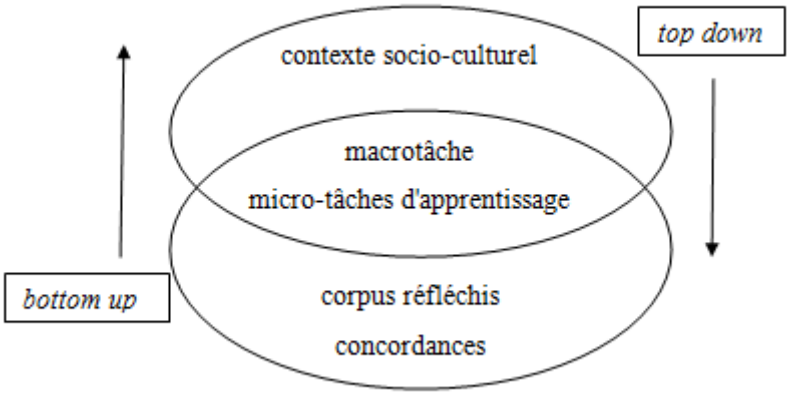


Figure 28 - Mouvements *top down* et *bottom up* et tâches basées sur des corpus et concordances.

Quels types de microtâches sont susceptibles d'alimenter le va-et-vient entre le mouvement *top down* et le mouvement *bottom up* ? Si la macrotâche est complexe, il convient de proposer des microtâches d'apprentissage qui font appel aussi bien aux corpus qu'aux concordances. Ceci permet d'observer les caractéristiques du produit que l'on souhaite obtenir. En effet, les corpus réfléchis contiennent les traces matérielles de produits réalisés par d'autres scripteurs. Si la tâche est peu complexe, les concordances peuvent suffire (cf. 4.2.4.2.).

4.2.3. Analyse de tâches basées sur des concordances

Selon Johns et King (1991 : iii), il y a un lien étroit entre la consultation de concordances et une approche de découverte ou d'exploration de l'*input*. Ce point de vue est également évoqué par Kaltenböck et Mehlmauer-Larcher (2005) et Chambers (2005). Toutefois, dans leurs articles, nous n'obtenons pas de précisions concernant ni l'objet de la découverte, ni le guidage. Johns et King (1991 : iii) précisent l'objet de la découverte comme suit : "*the use in the classroom of computer-generated concordances to get students to explore regularities of patterning in the target language*". Il s'agit donc de reconnaître des *patterns* que l'on peut définir comme regroupement d'unités simples en un segment, porteur d'un sens global grammatical ou discursif par exemple. Si l'on encourage à repérer des segments, on peut le faire en ayant des objectifs divergents. Premièrement, il est possible de chercher à favoriser une vision globale de l'objet. Le segment sera donc repéré et mémorisé comme une unité de sens, sans qu'il y ait forcément une analyse des parties qui le composent. Chini (2001) trouve

cette démarche problématique, surtout quand on dépasse le stade de l'apprenant débutant. Deuxièmement, on peut faire repérer le segment et encourager à une vision synthétique de l'objet. La compréhension du sens global sera donc accompagnée par une réflexion concernant l'agencement des parties qui composent le segment. La deuxième démarche est certainement cognitivement plus exigeante. Elle est toutefois plus proche du concept d'apprenant-chercheur, et elle paraît prometteuse en termes d'apprentissage.

Comment peut-on amener les apprenants à adopter une attitude exploratoire par rapport aux corpus et concordances ? Qu'est-ce que les apprenants peuvent y découvrir ? Selon Stevens (1995), grâce aux concordances, la relation de l'apprenant au texte change ; le texte devient manipulable. La lecture devient moins linéaire et plus ciblée. Cette vision est compatible avec l'activité de recherche d'information décrite par Rouet *et al.* (2004 ; cf. 2.5.1.).

J'analyserai ci-dessous des tâches existantes en me posant les questions suivantes. Apprend-on dans la description des tâches pourquoi les apprenants sont dirigés vers les corpus et/ou concordances ? Les tâches sont elles contextualisées ? Encouragent-elles un regard holistique, synthétique ou analytique ? Quelles tâches favorisent une attitude exploratoire, c'est-à-dire, quelles tâches donnent aux apprenants l'occasion de formuler leurs questions ? Comment aide-t-on les apprenants à exploiter les indices métatextuels ou autres qui leur permettent de savoir où ils peuvent trouver les informations qui correspondent à leur question ?

L'évolution technique des concordanciers et la possibilité d'accéder à une grande quantité de données en ligne permettent d'envisager, depuis quelques années, l'utilisation des concordanciers non seulement par des enseignants ou des étudiants particulièrement habitués aux ordinateurs, mais aussi par des personnes moins expertes en informatique. La facilitation de l'accès direct aux données en ligne semble en plus avoir donné un nouveau souffle au domaine de l'utilisation de corpus et de concordances pour l'enseignement/apprentissage des langues. Dans les études qui traitent de l'utilisation de ces outils, l'accent est en général mis sur l'apport soit pour les enseignants, soit pour les apprenants, soit pour les deux. Les enseignants peuvent s'en inspirer pour créer des matériaux d'apprentissage (Flowerdew, 1996). Enseignants et apprenants peuvent les utiliser pour vérifier et élargir leurs connaissances sur la langue (Stevens, 1995). Dans ce qui suit, je présenterai un échantillon de tâches lors desquelles les apprenants sont invités à utiliser des lignes de concordance. Cette restriction aux concordances s'explique par le fait que je n'ai pas eu accès à des documents

dans lesquels l'observation des textes complets d'un corpus est décrite de façon détaillée. Henry et Roseberry (1996 ; 2001) et Rohrbach (2003) donnent des pistes à ce sujet (cf. 3.4.2.2.).

La plupart des tâches dont j'ai eu connaissance ont été rédigées pour l'apprentissage de l'anglais, langue étrangère. Pour un aperçu plus complet concernant les langues cible citées en relation avec un travail sur des corpus, il est utile de consulter Boulton (2007). Un bon nombre des tâches invitent à une observation analytique d'unités, comme par exemple le repérage de prépositions régies par un verbe ou un adjectif (Johns, nd ; De Szendeffy, nd) ou l'observation de l'utilisation de verbes de modalité (Lamy et Klarskov Mortensen, 2006 ; Axtell, nd). Il y a une littérature relativement abondante concernant l'apprentissage du lexique à l'aide de concordances. Il s'agit de trouver les mots clés qui ont été supprimés dans des lignes de concordance (Stevens, 1988), de définir des mots auparavant inconnus (De Szendeffy, nd), de distinguer le sens de mots sémantiquement proches (Axtell, nd), ou de mémoriser du lexique après l'avoir observé dans des contextes variés (Stevens, 1991). Kaur et Hegelheimer (2005) proposent à des étudiants de mémoriser une liste de mots auparavant observés dans une concordance afin de les réutiliser ensuite lors d'une tâche de production écrite. Selon Goodwin-Jones (2001), les concordances peuvent aider à améliorer la compétence en lecture de textes en langue étrangère. Il explique qu'elles fournissent un contexte qui peut ensuite aider à choisir parmi les sens d'un terme proposés dans un dictionnaire électronique.

> *An interesting example of this use of concordances is providing contextual help in the reading of second-language texts. This approach seems to work best when students try computer-aided contextual inferences first (through the concordance) which can then be confirmed through on-line dictionary access* (Goodwin-Jones, 2001).

Selon Gaskell et Cobb (2004), les concordances peuvent soutenir l'apprentissage basé sur des *exemplars* (*exemplar-based learning*). Ils conseillent de construire le programme d'apprentissage en fonction des besoins exprimés par les apprenants.

> *It [the syllabus] should not present sequences of new structures that learners are hypothesized to be ready for, but rather shed light on structures they have already shown a need for. (...) the instructional presentation should be as non-declarative as possible; for example, it should high-light repeated patterns rather than collapse them into abstract terms. In other words, it should be example (data) driven* (Gaskell et Cobb, 2004).

Tandis que Gaskell et Cobb plaident pour une utilisation rare du métalangage grammatical, Ruthven-Stuart (nd), Tognini Bonelli (2000) et Chambers (2005) en font un objet d'apprentissage. Ils proposent des tâches qui donnent l'occasion de tester des références grammaticales existantes et d'en formuler par soi-même. Chez Ruthven-Stuart (nd), nous trouvons la tâche suivante.

> *Find the rule*
>
> *Rather than tell students a particular rule, students can be given a set of concordances and asked to discover for themselves what rules can be deduced from the evidence. This could be something simple such as discovering that the verb form changes [in] the third person. In addition, this task can be done with more complex patterns such as collocation* (Ruthven-Stuart, nd).

Tognini Bonelli (2000) décrit plusieurs exemples d'exploration de concordances. L'exploration est faite à partir d'une expression courte ou à partir d'une notion grammaticale. Elle propose de faire découvrir aux apprenants les caractéristiques grammaticales et sémantiques des expressions et de généraliser les observations afin de formuler des règles. Ces règles sont mises en concurrence avec celles formulées dans des grammaires qui ne sont pas basées sur des données de corpus. Dans les activités proposées par Tognini Bonelli (2000), le métalangage grammatical normé est toutefois très présent. Chambers (2005) précise que le travail grammatical peut être effectué de façon inductive ou déductive. Il est possible de faire trouver des occurrences et d'en déduire des règles grammaticales que l'on comparera ensuite avec les règles et occurrences fournies dans des grammaires ou manuels (démarche inductive). La démarche déductive propose aux apprenants de travailler sur un point linguistique à partir d'une grammaire ou à partir d'explications trouvées dans un manuel. Ils cherchent ensuite des occurrences qui correspondent à ce point linguistique, et ils comparent les données du corpus à celles des autres ressources. Chambers (2005) a démontré que la méthode inductive est préférée par plus d'apprenants que la méthode déductive.

Certaines (composantes) des tâches que j'ai décrites ci-dessus ont plus d'affinités avec le mouvement *top down* que d'autres (la contextualisation chez Goodwin-Jones, 2001 ; l'entrée par notions chez Tognini Bonelli, 2000). Les activités de Gaskell et Cobb (2004) font explicitement appel à l'*exemplar based learning*. D'autres activités, celles qui se réfèrent davantage au *rule based learning*, ont plus d'affinités avec le mouvement *bottom up* (Tognini Bonelli, 2000). Le lien entre les deux mouvements est rarement établi. Un tel lien semble être présent chez Kaur et Hegelheimer (2005). Toutefois, la tâche qu'ils décrivent est assez

contraignante. La réutilisation de mots d'une liste dans une production écrite est en fait imposée. Nous n'apprenons pas quelle place est laissée à l'exploration et à la réflexion lors de l'exécution de cette tâche. Johns (nd), Pickard (1993), Ma (1993), Flowerdew (1996) et Kennedy et Miceli (2001) combinent corpus, concordances et projets d'écriture. Cette combinaison donne l'occasion de créer le lien entre une approche holistique et une approche analytique. Les concordances y sont souvent utilisées lors de la phase de révision des premiers jets de production écrite rédigés par les apprenants. En se référent à Swain (1985 ; 1985), Gaskell et Cobb (2004) mettent en avant le rôle de la production d'un *output* compréhensible pour la construction du savoir grammatical.

> *Input may be the source of new grammatical forms, but this is not where the details are worked out; it is rather in production that learners have the opportunity to explore new forms and match them to communicants' expectations, particularly in the offline opportunities provided by writing*.

Rézeau (2007) décrit le potentiel de l'utilisation de concordanciers par des apprenants d'anglais, langue étrangère, à l'université lors d'activités de production écrite sur des forums de discussion. Kennedy et Miceli présentent une étude qualitative faite en Australie auprès d'étudiants de disciplines variées, apprenant l'italien comme non-spécialistes ("*undergraduate Italian program*"). Cette étude décrit l'utilisation de concordances lors d'un projet d'écriture par un public qui apprend une autre langue que l'anglais et qui a un niveau intermédiaire à avancé. Mon propre public apprend une autre langue que l'anglais, et il a un niveau intermédiaire. Comme j'ai eu accès à très peu d'études ayant cette configuration, j'estime qu'une présentation exhaustive des travaux de Kennedy et Miceli (2001) se justifie, d'autant plus que le thème du guidage y est abordé. Ce thème est rarement traité par les autres chercheurs du domaine. Kennedy et Miceli (2001) proposent l'utilisation de corpus et d'autres aides à la rédaction lors d'un cours de deux heures hebdomadaires de production écrite et de grammaire : "*Interrogation of the corpus was not presented as an end in itself but rather as an integral part of the writing and grammar work being undertaken*". Un corpus spécialisé d'environ 570 000 mots a été créé pour le projet. Il contient des textes entiers (lettres et courriels privés et officiels, courriers des lecteurs et réponses des experts, articles dans des magasines, critiques de film, etc.). Les textes ont été classés dans deux rubriques, à savoir la rubrique des textes écrits par des scripteurs professionnels et la rubrique des textes écrits par des scripteurs non-professionnels. Kennedy et Miceli ont opté pour le modèle d'apprentissage du compagnonnage qui donne un statut particulier à l'observation de la façon dont les personnes plus expérimentées procèdent : "*promote learning by example and by*

experience".[16] Ceci se traduit par une initiation à l'utilisation de corpus et de concordances et par le choix de faire travailler les étudiants en binômes ou en petits groupes. Un des objectifs du projet de recherche est la formulation de conseils, basés sur les observations des chercheurs lors du projet. Le guidage est donc au centre de cette étude. Les étudiants ont été encouragés à utiliser le corpus pour réviser leur production écrite ou celle de scripteurs anonymes. Il s'agit d'une activité de type résolution de problème. Les étudiants ont également fait des activités de type "chasse au trésor" (*treasure hunting*). Il s'agit d'activités de découverte de formulations dans les textes du corpus. Ces activités peuvent, par exemple, faire découvrir comment commencer ou terminer une lettre, comment formuler une excuse ou un remerciement, etc. De telles activités permettent d'aborder les textes du corpus en choisissant une entrée par le sens. Elles traduisent une approche du corpus davantage discursive que grammaticale ou lexicale. Nous sommes ici dans le mouvement *top down*. L'observation de listes de fréquence lors de séances de chasse au trésor dans le domaine des collocations semble avoir particulièrement motivé les apprenants (Kennedy et Miceli, 2001).

> *These [frequency lists of collocations] proved to be interesting and entertaining to the students, not only as an alternative to overused expressions, but also as triggers for further searches*.

Les résultats de l'étude montrent que ce n'est pas toujours le niveau de langue qui pose problème lors de l'exploration de corpus et concordances. Ce sont plutôt des aptitudes dans le domaine de la recherche d'information, donc des aptitudes cognitives, qui sont en jeu. Il s'agit de raisonner, de procéder de façon logique, de changer de stratégie s'il n'y a pas assez ou trop d'occurrences qui s'affichent, et en général, de poser les questions susceptibles de mener à un résultat pertinent qui peut être parfois inattendu. Bref, il s'agit de former de futurs "*corpus researchers*", donc des apprenants capables de faire des recherches sur corpus. Nous reconnaissons ici l'héritage de Johns (1988 ; 2002). Kennedy et Miceli (2001) proposent quatre étapes à suivre lors de l'exploration d'un corpus.

– Formuler la question.
– Déterminer une stratégie de recherche.
– Observer les occurrences et sélectionner celles qui sont pertinentes.
– Tirer des conclusions à partir des occurrences pertinentes.

À partir des questions formulées par les étudiants, Kennedy et Miceli (2001) ont établi cinq catégories concernant la difficulté de faire des recherches pertinentes.

[16] On peut également consulter Narcy-Combes (2005 : 121) à ce sujet.

– Formulation de questions trop générales ou trop spécifiques.
– Formulation de questions fermées, ce qui peut mener à l'obligation de reformuler une autre question.
– Manque de formulation d'une question de départ ; observation d'occurrences et parfois abandon de la recherche quand les occurrences ne conviennent pas.
– Observation de prépositions et des mots qui suivent, sans regarder les mots précédents.
– Trop s'attacher au sens des prépositions.
– Observation de caractéristiques grammaticales davantage que de caractéristiques lexicales.

Souvent, la première difficulté a été la formulation d'une question de recherche pertinente. Kennedy et Miceli (2001) donnent aux apprenants des conseils susceptibles de les aider à formuler de telles questions. Les conseils concernent des domaines dans lesquels j'ai moi-même formulé des conseils très explicites pour les apprenants lors de la recherche-action : quel mot clé il convient de rechercher, rechercher une partie d'un mot clé seulement, utiliser le dictionnaire pour trouver tel ou tel mot, etc.

Suite à leurs résultats, Kennedy et Miceli (2001) proposent un scénario qui débute par une phase de "chasse au trésor", donc une phase d'exploration libre. Si les apprenants débutent par une phase d'observation libre, il y a des chances qu'ils soient plus sensibles aux phénomènes de collocation. Le mouvement premier qui est préconisé ici est donc un mouvement *top down*. Ceci est en accord avec les recommandations de Portine (2008). Dans un deuxième temps, les apprenants font un travail de résolution des problèmes qui se sont manifestés lors de la rédaction de textes. La révision des textes peut être soutenue par des concordances. Kennedy et Miceli (2001) expliquent pourquoi une préparation aux méthodes de travail sur corpus leur semble indispensable.

> *Although we do not go so far as to suggest that learners need formal training in logic in preparation for corpus work, our evaluation of the ways students go about problem-solving with CWIC has convinced us of the importance of an awareness of logical principles applicable to this kind of operation.*

Selon Kennedy et Miceli (2001), apprendre à travailler en procédant de façon logique peut aider les apprenants à progresser non seulement dans le domaine de l'apprentissage de la langue étrangère, mais aussi dans d'autres domaines, en particulier dans des domaines qui nécessitent la manipulation d'outils. Nous n'apprenons pas ici comment on pourrait amener les apprenants à travailler "de façon plus logique". La conceptualisation s'intéresse à un travail

cognitif. Il serait à voir si des activités conceptualisantes pourraient aider à concrétiser la proposition de Kennedy et Miceli (2001). Cette étude est celle que je trouve la plus proche de mes préoccupations parmi les articles et ouvrages que j'ai eu l'occasion de consulter. Le public, la tâche rattachée à l'exploration de corpus et la façon dont Kenney et Miceli envisagent l'apprentissage (compagnonnage réfléchi) ressemblent à ma configuration de recherche. J'analyserai mes données en m'inspirant des cinq catégories de recherches non pertinentes présentées par Kennedy et Miceli (2001). Je reviendrai sur ces catégories en 5.5.4.2.

4.2.4. Quelles microtâches pour soutenir l'observation de la langue ?

Braun (2007) relève le fait que, jusqu'à présent, le grand nombre de travaux de recherche sur les corpus et concordances n'a pas eu beaucoup de retombées sur le terrain de l'enseignement/apprentissage des langues. Comment créer des tâches qui reflètent les principes méthodologiques évoquées par les chercheurs en linguistique de corpus appliquée (cf. 1.4.1.) et qui soient compatibles avec les possibilités du terrain d'application ? Afin de donner des éléments de réponse à cette question, j'examinerai d'abord quelques-unes des tâches créées par des personnes ayant participé à un atelier proposé par Johns en 2000. Ensuite, je présenterai la genèse et l'analyse d'une activité d'observation de lignes de concordance que j'ai proposée en février 2009 à des apprenants en classe de sixième. J'examinerai le degré de contextualisation, et j'évaluerai les occasions offertes à l'apprenant d'être en situation de recherche, donc d'adopter une attitude d'apprenant-chercheur (cf. 2.6.3.). A-t-il des chances de repérer des éléments pertinents grâce aux consignes et grâce aux données qu'il est censé consulter ?

4.2.4.1. Atelier de production d'activités pédagogiques

En 2000, Tim Johns a proposé à Ustí nad Labem (Bohême du Nord) un atelier de production d'activités pédagogiques, basées sur des corpus et concordances. Le site qui publie les activités créées lors de l'atelier à Ustí nad Labem ne précise pas le public qui participe à la formation, mais il est possible d'inférer qu'il s'agit d'enseignants d'anglais, langue étrangère (Johns, 2000). Nous n'apprenons pas non plus si les concepteurs des tâches prévoient un contexte plus large (programme d'apprentissage, macrotâche) dans lequel les microtâches pourraient s'insérer. Regardons de plus près quelques-unes des activités afin d'observer la nature des tâches, des consignes et des données proposées.

Toutes les tâches sont conçues dans un esprit "hors ligne", c'est-à-dire qu'elles peuvent être imprimées sur papier et ne nécessitent pas l'utilisation d'un ordinateur. Ceci limite les possibilités d'exploration de concordances et d'apprentissage incident. Presque toutes les consignes sont assez ouvertes, comme dans l'exemple ci-dessous.

> *Try to spot what the typical cases of 'great' and its collocations are on the basis of the following examples* (Sarka Canova et Jarka Ivanova).

"*Typical cases*" est un terme qui nécessite un contexte. Si la recherche des sens variés de "*great*" était motivée par une macrotâche de production en langue de spécialité, cette microtâche d'observation des caractéristiques de "*great*" serait pertinente. Nous trouvons l'amorce d'une telle recherche de contextualisation dans la consigne suivante.

> *How do we use the prepositions ABOUT and ON? Which one is used more often? Which one tends to be used in academic texts? In which cases is there the occurence of only one of them?* (Kvìta Rychtárová).

Le titre de la microtâche suivante semble annoncer la possibilité d'une contextualisation.

> *POLITICALLY CORRECT EXPRESSIONS: "person" v. "man" etc. Study the following examples and try to describe the thinking behind them* (Dana Slancarova et Petra Tovarkova).

Il s'agit de repérer des formulations qui sont "*politically correct*". Ce terme n'est pas explicité ici. Les apprenants sont invités à repérer comment pensent les scripteurs qui utilisent les mots composés désignant des personnes (*...man, ...person, ...mother*). Ceci semble faire référence au mouvement *top down*. Il s'agit toutefois d'une activité dont les contextes sont multiples et non explicités. Les apprenants doivent d'abord lire 16 petits textes, classés dans trois catégories dont la différence n'est pas expliquée. Ensuite, 21 termes désignant des personnes doivent être utilisés pour compléter des lignes de concordance dont le mot clé principal manque (*gap-filling exercice*). Dans l'activité suivante, nous avons affaire à la situation contraire, un nombre d'occurrences que j'estime insuffisant.

> *Look carefully at the following citations. What difference in the meaning of the adjectives classic/classical can you spot? (...) Now fill in the blanks with appropriate adjectives* (Zuzana Šaffková & Vladislav Smolka).

Cette activité propose un seul contexte pour deux types d'adjectifs qui se terminent soit par *...ic*, soit par *...ical*, comme l'annonce le titre de l'activité. Il pourrait être utile ici que les apprenants aient accès à une concordance en ligne.

Les microtâches que je viens de présenter gagneraient toutes à être intégrées dans un contexte plus large. Ceci permettrait de revoir certaines de leurs caractéristiques, comme la trop grande ou trop petite quantité de données à observer. Les questions de recherche sont formulées dans les consignes. Il s'agit donc des questions des enseignants. Il serait intéressant de partir de questions existantes des apprenants et d'identifier des éléments de réponse dans la concordance. Afin de montrer comment on peut établir le lien entre macrotâche, microtâches et concordances en prenant comme point de départ un élément microsyntaxique, je présente ci-dessous la genèse et l'analyse d'une activité que j'ai développée en février 2009 pour des apprenants en première année d'allemand. Cette activité ne fait pas appel à un corpus réfléchi ; elle s'appuie toutefois sur une page dans le manuel d'allemand qui représente trois textes appartenant au même genre textuel. Des lignes de concordance, faites à l'aide du corpus *deWac* (*Sketch Engine*) sont mises à disposition des apprenants.

4.2.4.2. Je t'invite à mon anniversaire

Le 5 et 9 février 2009, dans le contexte d'un chapitre dont le thème est l'anniversaire, j'ai proposé à onze apprenants de sixième (première année d'allemand) une activité destinée à leur faire découvrir les caractéristiques particulières du verbe "*einladen*" [inviter]. Ce verbe transitif se distingue par sa particule détachable "*ein*" et par la modification de la voyelle de son radical à la 2ème et 3ème personne du singulier. L'activité est basée sur onze lignes de concordance, contenant toutes le mot clé "*laden*" ayant comme contexte de droite la particule "*ein*".

Die Buslinie 137	**lädt**	zu einer Zeitreise ***ein***.
Wir	**laden**	alle Unternehmen herzlich ***ein***.
Die Westküste	**lädt**	zu Touren ***ein***.
"Fête de la Science" : Frankreich	**lädt**	100 deutsche Jugendliche ***ein***.
Ich	**lade**	euch herzlich ***ein***, mit dabei zu sein.
Wir danken für das Interesse und	**laden**	Sie herzlich ***ein*** zum Chat mit Tom.
Da	**lädt**	ihn eine alte Frau ***ein***.
Ich	**lade**	alle ***ein***.
Du	**lädst**	sie auf ein Eis ***ein***.
Ich lade dich ein, du	**lädst**	mich ***ein***.
Ihr	**ladet**	uns ***ein***, wir kommen auch in Schulen.

Tableau 16 – Concordance de "*laden* + *ein*", faite à l'aide de *Sketch Engine*.

Les apprenants étaient invités à observer par groupes de deux ou trois les lignes de concordance présentées ci-dessus. La figure suivante présente les consignes initiales destinées à soutenir l'activité d'observation.

1. Afin de comprendre la conjugaison de "einladen", observe ces lignes de concordance. Que constates-tu ?
2. Quelle personne reconnais-tu ? (par exemple : 1ère personne du singulier, 3ème personne du pluriel, etc.) Note les personnes que tu reconnais dans le cadre à gauche de la concordance. Abrège, par exemple "1ère p.sg." et "3ème p.pl.".
3. Quelle personne manque ? Peux-tu la déduire des autres formes ?
4. Quelle forme n'arrives-tu pas à classer ? As-tu une explication ?
5. Note la conjugaison complète dans ton cahier.

Figure 29 – Consignes initiales pour l'activité "Je t'invite à mon anniversaire".

Analyse de l'exécution de la tâche

Lors de la séance d'observation de lignes de concordance, deux des apprenants, Dorian et Larissa, emploient un métalangage grammatical normé ; leur raisonnement illustre l'habitude de parler de phénomènes grammaticaux. D'autres apprenants, par exemple Thierry, ont été particulièrement silencieux. Thierry m'a dit que sa triade n'avait "rien trouvé", ce qui ne s'est pas avéré exact quand j'ai échangé avec les apprenants faisant partie de ce groupe, à savoir Thierry, Martin et Bernard. Dans trois binômes sur les quatre et dans la triade, un apprenant a particulièrement dominé et parlé davantage que son ou ses partenaires. Le binôme qui a travaillé sans ma présence a établi un vrai dialogue entre apprenants. Chaque fois que je me suis jointe à un des petits groupes, le dialogue a eu tendance à s'établir entre moi et un des apprenants. Sans ma présence, les apprenants n'ont pas eu d'autre choix que de parler l'un avec l'autre, surtout en situation de binôme. Rappelons que Portine (2008) recommande cette constellation qu'il trouve fructueuse et qu'il trouve trop peu employée lors de situations d'apprentissage des langues.

Modification de l'activité

J'ai constaté après coup que j'ai dirigé les apprenants vers une démarche de *bottom up* en évoquant comme objectif la découverte de la conjugaison "étonnante", comme je le dis, de "*einladen*". Les apprenants auraient peut-être observé autre chose que la conjugaison si j'avais été moins directive dans ma consigne. Afin de permettre lors d'une future réalisation de cette activité des observations plus larges, moins focalisées sur la conjugaison, j'ai modifié certaines consignes, et j'en ai rajouté d'autres. Le rajout s'explique par le projet de production qui motive l'activité d'apprentissage. Si l'arrière-plan de l'activité est la production d'un carton d'invitation à une fête, il faudrait faire trouver les lignes qui correspondent à cette situation. Il est également intéressant de faire repérer les compléments d'objet direct qui concernent les personnes que l'on invite : "*Ich lade dich/ihn/euch/alle ein*" et l'évènement auquel on invite quelqu'un : "*zum Geburtstag*". On peut aussi faire repérer les adverbes qui expriment la

politesse dans l'invitation. Une analyse des cooccurrents, donc des mots apparaissant le plus souvent dans un environnement proche d'un mot clé, fait à l'aide de *Cosmas II*, montre que "*einladen*" est fréquemment accompagné de l'adverbe "*herzlich*" [cordialement]. J'ai donné comme devoir à la maison de noter la conjugaison de toutes les personnes au présent de "*einladen*". Lors de leur présentation de la conjugaison, peu d'apprenants ont placé la particule détachable "*ein*" à la fin, par exemple en écrivant "*Ich lade ... ein*." [J'invite…]. Larissa a écrit "**Ich einlade*", Laurent et d'autres ont écrit "**Ich lade*".

L'analyse de l'activité "Je 'invite à mon anniversaire" et de son exécution par un groupe d'apprenants m'a aidée à modifier la première version des consignes de l'activité. Cette première version met en avant la conjugaison et l'ordre des mots. On est donc dans un même dans une approche formaliste qui se concentre sur la morpho-syntaxe. Il s'avère lors du devoir à la maison que le repérage des caractéristiques d'un verbe ne mène pas immédiatement à son emploi conforme aux normes grammaticales. Il conviendrait d'intégrer un mouvement de type *top down*, et il manque une phase d'entraînement. La mise en évidence que le projet "Écrire un carton d'invitation" sous-tend cette activité permet de reformuler les consignes initiales et de les compléter comme suit.

1. Afin de faire connaissance avec le verbe "*einlade*n", observe ces lignes de concordance. Que constates-tu ?
2. Quelles personnes "grammaticales" reconnais-tu ? (par exemple : 1ère personne du singulier, 3ème personne du pluriel, etc.). Note les personnes que tu reconnais dans le cadre à gauche de la concordance. Abrège, par exemple "1ère p.sg." et "3ème p.pl.".
3. Qui est invité ? Cite les mots qui nous renseignent sur cela. Que constates-tu ?
4. Dans quelles lignes parle-t-on de l'évènement auquel on invite quelqu'un ? Que constates-tu ?
5. Quelles lignes peuvent t'aider à écrire un carton d'invitation ?
6. Écris un carton d'invitation.

Figure 30 – Consignes définitives pour l'activité "Je t'invite à mon anniversaire".

La première consigne a perdu le terme "comprendre la conjugaison de..." au profit de la formule plus ouverte "faire connaissance avec le verbe…". La deuxième consigne est reprécisée : il s'agit de repérer des "personnes", certes, mais des "personnes grammaticales". Ce rajout est devenu nécessaire en raison de la nouvelle consigne, la troisième, qui invite à repérer les invités, donc de "vraies personnes". La quatrième consigne a été rajoutée afin de faire repérer l'évènement auquel quelqu'un est invité. Sans l'annoncer explicitement, la troisième et la quatrième consigne peuvent permettre d'observer la façon dont se présentent des noms et pronoms ayant la fonction de COD et de voir quelle préposition est utilisée en

rapport avec l'évènement (*zu* + datif). La cinquième et la sixième consigne sont clairement orientées vers la tâche de production écrite. On est davantage dans le mouvement de *top down* ici, mais sans négliger l'observation des traces matérielles du sens que l'on souhaite faire reconstruire. J'ai testé la nouvelle version des consignes qui ont été données comme devoir à la maison pour le 10 février 2009. Les apprenants m'ont dit qu'ils ont trouvé "difficile" de repérer les formes qui représentent "les invités" et "les évènements". À partir du repérage de ces formes par le groupe-classe, j'ai listé les pronoms personnels qui ont la fonction de COD dans la concordance. Ensuite, les apprenants ont fait une activité de type "exercice à lacunes" lors de laquelle ils devaient trouver des pronoms, des formes de "*einladen*" et des prépositions (certaines incluant un déterminant) repérées dans les cartons d'invitation présentés dans leur manuel d'allemand (*zu, am, um, bei*). Ces prépositions aident à introduire le lieu, l'heure, le jour de la fête et l'évènement. J'ai pu observer les effets positifs d'un tel réaménagement des consignes.

4.3. Pour résumer

Après une première observation des tâches proposées par des chercheurs comme Johns, Stevens, Flowerdew, Kennedy et Miceli, Lamy et Klarskov Mortensen, etc., j'ai fait le choix d'analyser ces tâches à la lumière de Portine (2008). Ceci implique un rapprochement entre macrotâches, microtâches et l'activité langagière qui peut consister en des traitements de type *bottom up* ou de type *top down* ou d'un mélange des deux. Les microtâches ont-elles un arrière-plan qui leur donne du sens ? Font-elles partie d'un projet, d'un scénario qui fournit un contexte social ? Sont-elles orientées vers l'apprentissage individuel ? S'intéressent-elles à l'aspect cognitif de l'apprentissage ? Y a-t-il un compromis possible entre "le tout-culturel" et "le tout-linguistique", entre une approche favorisant la perception holistique de l'objet "langue" et une approche favorisant une perception analytique de l'objet "langue" ? Comment faire en sorte qu'aucun des deux ne prenne trop de poids, mais qu'il ne soit pas non plus négligé ? Le site *Kibbitzers* (Johns, nd) présente des séances de révision de la production écrite. Ce site permet de voir comment un scripteur peut être guidé pour réviser sa production écrite à l'aide de concordances. Kennedy et Miceli (2001) décrivent l'utilisation de concordances comme outils pour des ateliers d'écriture. Mon scénario s'inspire des démarches de Johns et Kennedy et Miceli qui proposent une synthèse entre une approche *top down* et une approche *bottom up*. En effet, le travail est contextualisé, mais on s'intéresse aussi à la langue, parfois à des phénomènes très "petits", comme des terminaisons, parfois à des phénomènes sémantiques ("Le sens du mot est-il approprié dans le contexte local xy ?") ou à des

phénomènes discursifs ("Peut-on dire cela dans un texte représentant le genre textuel xy ?"). Si nous tenons compte du genre textuel, nous pouvons peut-être faire la synthèse entre l'individuel et le social. Selon Rastier (2001), "le genre partage tout à la fois le caractère public de l'action individuelle socialisée et de la norme sociale où elle prend place." Rastier (2004) rappelle que les genres s'inscrivent matériellement dans le texte.

> En d'autres termes, le contenu d'un texte ne se réduit certes pas à une mystérieuse représentation mentale : un texte est fait de deux plans, celui des formes sémantiques et celui des formes expressives, dont le genre notamment norme la mise en corrélation (Rastier, 2004).

Dans le prochain chapitre, je présenterai en détail ma recherche-action. Elle a pour objectif l'observation de l'évolution d'un dispositif qui est conçu en accord avec les principes méthodologiques évoqués dans les chapitres précédents.

5. Une recherche menée au collège

5.1. Types de recherche

En sciences humaines, on peut distinguer entre deux types de recherche, la recherche expérimentale et la recherche qualitative. Le deuxième type de recherche peut prendre la forme d'une recherche-action (cf. 5.1.2.).

5.1.1. Recherche expérimentale

La recherche expérimentale se caractérise par un protocole de recherche hautement standardisé et contrôlé. Ce protocole nécessite l'élaboration d'une hypothèse, la constitution de groupes dits "homogènes" et de groupes témoin qui n'effectuent pas les tâches expérimentales, l'observation de paramètres isolés, la formulation des résultats et leur généralisation (Grosbois, 2007). La recherche expérimentale est souvent exécutée dans des laboratoires. Les tâches et les types d'observation sont construits pour les besoins de la recherche.

5.1.1.1. Types d'observation

La recherche expérimentale vise à observer les faits "existants". Elle chercher à se donner les moyens de créer des relations entre causes et effets, afin d'arriver à une explication de la "réalité" observée.

> Dans le cadre d'une recherche expérimentale, tout doit pouvoir être contrôlé, vérifié, reproduit, recalculé. De plus, l'expérimentation suppose une comparaison de résultats obtenus dans des situations dont on connaît les variables et dont on en fait varier UNE pour en étudier les conséquences sur les autres (Grosbois, 2007).

Ce type d'observation restreint la complexité des phénomènes observés afin de mieux les contrôler. L'expérience doit pouvoir être reproduite (Grosbois, 2007).

5.1.1.2. Postures de chercheur

En recherche quantitative, les chercheurs ont en général une posture explicative. Ils visent à observer leur objet "de l'extérieur", c'est-à-dire, sans participer directement à l'expérience, et en cherchant à limiter le plus possible les biais liés à leur individualité quand ils interprètent les données. Toutefois, l'adhésion à une théorie ou méthodologie peut également représenter un biais (Narcy-Combes, 2005 : 66).

5.1.2. Recherche-action

La recherche-action est une forme de recherche qualitative. La recherche qualitative vise à comprendre l'objet étudié dans toute sa complexité. Le point de départ n'est pas forcément théorique ; la question initiale peut être formulée par des praticiens et reprise par des chercheurs qui font le lien entre cette question et des théories existantes. En particulier dans sa version anglo-saxonne, ce type de recherche a un ancrage idéologique fort (Chateau, 2003 : 60). Il peut, entre autres, viser à modifier une situation ou un comportement considérés comme problématiques.

> Il s'agit de recherches dans lesquelles il y a une action délibérée de transformation de la réalité ; recherches ayant un double objectif : transformer la réalité et produire des connaissances concernant ces transformations (Hugon et Seibel, 1988 : 13).

Les modifications que l'on souhaite obtenir nécessitent l'implication aussi bien des acteurs du terrain (apprenants, enseignants, chefs d'établissement, etc.) que du chercheur.

> [La recherche-action est] un processus interventionniste conçu par ou à tout le moins avec les sujets impliqués et dont l'objectif est la modification par les praticiens de leur relation à leurs postures d'enseignement/apprentissage (Montagne-Macaire, 2007).

Je reviendrai ci-dessous aux postures d'enseignement/apprentissage des enseignants et chercheurs. Les postures, ou – autrement dit – les attitudes des apprenants seront traitées en 6.2.

5.1.2.1. Pourquoi choisir la méthodologie de la recherche-action ?

La recherche-action est considérée comme une des méthodologies qui conviennent à des recherches dans le domaine d'une didactique qui s'intéresse à la pratique (Narcy-Combes, 2005 : 83). Ce type de recherche semble, en effet, convenir à l'observation d'apprenants de langue, en particulier si cet apprentissage est organisé par une institution. Dans ce cadre, les participants sont en général de niveau hétérogène. Il ne paraît défendable ni du point de vue de la pratique, ni du point de vue de l'éthique d'éliminer les participants qui ne cadreraient pas avec un protocole de recherche.[17] Il convient néanmoins de s'interroger sur le bienfondé du choix de cette méthodologie. Un chercheur en didactique des langues qui ne choisit pas la méthodologie de la recherche-action se positionne dans le champ d'une ou de plusieurs des disciplines contributoires de la didactique, ce qui l'oblige à être ou devenir spécialiste du

[17] Cela ne veut pas dire qu'il faut ignorer les différences de niveau qui existent au sein d'un groupe classe.

domaine sollicité. Ses résultats de recherche sont alors moins participatifs et la collaboration avec le terrain est moins égalitaire que lors d'une recherche-action (Narcy-Combes, 2005 : 112). Les informations recueillies par un observateur ou une observatrice participant(e), menant une recherche-action, peuvent fournir des résultats de type compréhensif susceptibles d'aider à réguler des actions comparables futures qui auront lieu dans des contextes similaires.

5.1.2.2. Types d'observation

> Cette démarche qui peut emprunter des outils à la méthodologie qualitative des sciences humaines se distingue par la finesse d'étude des situations humaines, la capacité de travailler sur la complexité et de se placer au cœur de la réalité sociale (Bazin, 2009).

Lors d'une recherche-action, l'observation est, dans une large proportion, directe et transparente, c'est-à-dire que le chercheur est idéalement présent à la plupart des situations observables, et il informe les acteurs impliqués dans la recherche sur ce qu'il fait et comment il le fait. Par exemple, il leur signale qu'il utilisera un dictaphone, dans quel cadre il diffusera les enregistrements, quels documents il compte analyser, etc. Certaines données peuvent être recueillies par quelqu'un d'autre que le chercheur ; il paraît judicieux de faire mener les entretiens oraux par un intervenant extérieur, surtout si le chercheur qui organise la recherche-action est en même temps l'enseignant du groupe sur lequel on enquête. Il est souhaitable de trouver plusieurs points d'entrée pour observer les données. Une triangulation doit permettre de donner de la crédibilité à l'interprétation des données. Le chercheur croise les informations qu'il a obtenues en analysant des produits (dans mon cas, ce sont des textes), des paroles des acteurs (enregistrées ou notées après la situation observée) et des actions qu'il a observées pendant les séances de travail en présentiel ou à distance.

5.1.2.3. Postures de chercheur

Un observateur participant adopte une posture subjective. Celle-ci peut aller de pair avec une attitude volontariste et interventionniste. La posture du chercheur joue un rôle important lors de l'exécution de la recherche. Bien la connaître peut l'aider à prendre conscience de biais éventuels, de visées, affects ou autres obstacles épistémologiques susceptibles de l'empêcher d'aborder les faits de façon suffisamment ouverte (Narcy-Combes, 2005 : 78 ; Demaizière & Narcy-Combes, 2007). Narcy-Combes (2005 : 104) propose la figure suivante qui décrit quatre postures de recherche. Cette figure s'inspire de Juan (1999) et de Chalmers (1987).

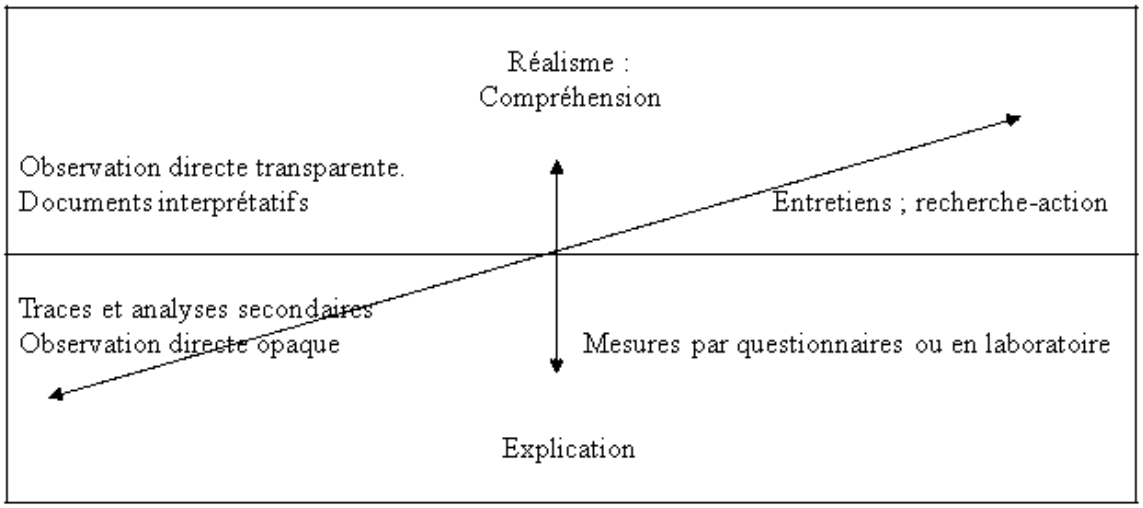


Figure 31 – Approches et horizons de la méthode en sciences humaines.

Selon Missotte (2003 : 248), pendant les années 60, la méthodologie de la recherche-action a été employée dans certains pays du tiers monde, afin de les aider à se développer. La recherche-action, grâce à des processus interactifs qui semblent constitutifs pour cette méthodologie, permet de transformer des pratiques. Elle vise à "permettre à chacun de se décentrer, de considérer d'autres points de vue et d'entrer dans de nouveaux schèmes cognitifs" (Missotte, 2003 : 248). Afin de rendre possible un tel décentrage, tous les acteurs de la recherche-action doivent se sentir libres d'échanger leurs points de vues concernant l'objet étudié. Un tel échange peut être favorisé par un environnement de travail adapté.

> D'où les questions sur les environnements de travail, les contenus de savoir, les modes de régulation à mettre en place pour construire (…) un milieu de travail encourageant les échanges et confrontations cognitivistes de tous (Missotte, 2003 : 248).

5.1.2.4. Ma posture de chercheur

La recherche-action semble une méthodologie adaptée à la plupart des facettes de ma posture de chercheuse.

- Je vise à prendre en compte la complexité de la situation d'enseignement/apprentissage.
- J'ai une attitude volontariste, voire interventionniste.
- Je suis chercheuse et enseignante, ce qui facilite une observation participante.
- Je ne m'intéresse pas seulement à un élément isolé, mais je cherche à comprendre plusieurs facettes de l'objet d'étude, en réduisant peu sa complexité.
- Je pense être prête à me laisser transformer par l'action.

- C'est un choix éthique de ma part d'informer le plus possible les acteurs concernés, en particulier les apprenants, sur ce que je fais lors de la recherche. Je vise à les tenir au courant des éléments que j'observe et les traces que je garde dans le but de les analyser.
- J'essaie d'être réceptive quant au vécu de la recherche par les acteurs impliqués, dont les apprenants, et d'ajuster le projet en fonction des propositions de modification. Si un ajustement ne peut être effectué lors de la phase en cours, je retiens les propositions pour le développement des projets à venir.

Lors de mes recherches, j'ai procédé à une observation directe transparente, c'est-à-dire, une observation effectuée après avoir informé préalablement les acteurs concernés (Narcy-Combes, 2005 : 104). À l'aide d'une intervenante extérieure, j'ai fait faire des entretiens oraux filmés.

5.2. Légitimer une recherche au collège

Une recherche-action est une intervention qui risque de bouleverser les pratiques "établies" et de remettre en question les postures d'enseignement/apprentissage des acteurs impliqués dans la recherche. Pour cette raison, il convient de trouver des points d'entrée dans le contexte institutionnel qui légitiment de façon officielle le projet qu'un chercheur propose lors d'une recherche-action.

5.2.1. S'inspirer de manuels scolaires

Rohrbach (2003), enseignant d'anglais au collège, décrit l'utilisation d'un corpus de textes par un groupe d'apprenants d'anglais, langue étrangère, de niveau troisième. En coopération avec Rohrbach, Braun (2007) a également fait utiliser des corpus pendant des séances de LE par des apprenants d'anglais de classe de troisième. Rohrbach et Braun ont choisi de s'appuyer sur un chapitre du manuel scolaire utilisé par le groupe afin de s'assurer de la "faisabilité" du module dans des conditions d'enseignement/apprentissage "ordinaires". Braun (2007) justifie ce procédé par les obligations qui ont été imposées à la chercheuse et à l'enseignant par l'institution.

> *A crucial prerequisite to take over the class was to cover the main points of the textbook unit as far as possible, in spite of using the corpus materials. This was in line with the overall curricular issues (...), meaning that the case study took place under very realistic conditions* (Braun, 2007 : 310).

Lors de *Brieffreunde.org*, le dernier projet de ma recherche (cf. 5.3.5.), j'ai exploré la possibilité de choisir un chapitre de manuel comme point de départ pour un projet basé sur des corpus et concordances. Toutefois, comme les manuels scolaires sont censés respecter et refléter les programmes officiels, j'estime qu'il convient de se placer à un niveau supérieur et de s'aligner plutôt sur les principes et propositions évoqués dans les documents officiels d'enseignement. Les résultats des recherches doivent laisser la possibilité au chercheur de se prononcer, entre autres, sur la validité didactique de certaines composantes de ces programmes.

5.2.2. Programmes officiels d'enseignement

Les programmes officiels des trois disciplines, français, technologie et allemand, contiennent certains éléments d'enseignement obligatoires ou facultatifs dont le projet tient compte.

5.2.2.1. Programme officiel pour la technologie

Dans les programmes pour l'enseignement de la technologie en classe de troisième, en vigueur avant 2008, la création d'un document multimédia est préconisée.

> Des réalisations assistées par ordinateur élargissent et diversifient la pratique de l'outil informatique pour concevoir, produire et échanger des documents, services et objets. Cette plus grande familiarité avec la technologie de l'information dépasse la pratique machinale des ordinateurs. Elle vise à leur utilisation raisonnée (Bulletin Officiel n° 10, 15 octobre 1998).

5.2.2.2. Programme officiel pour le français

En ce qui concerne le français, les extraits suivants des documents officiels en vigueur avant 2009 soulignent le lien étroit avec certains principes et outils qui caractérisent ma recherche-action. En classe de troisième, il est prévu de faire observer des textes documentaires à l'aide de banques de données (Bulletin officiel de l'Éduction Nationale n° 10, 15/10/1998, hors série).

> Pour conduire les élèves à une plus grande autonomie dans le choix et le maniement des documents, on développe l'usage de dictionnaires, d'usuels et d'ouvrages de références. On leur apprend à consulter les banques de données, notamment informatiques et télématiques (Ministère de l'éducation nationale, 1998).

Les enseignants de lettres sont encouragés à utiliser des blogues comme support d'écriture (Site de l'académie de Versailles). La "pédagogie de l'intertextualité" qui est décrite dans le

document d'accompagnement des programmes officiels fait écho à la méthodologie qui sous-tend *Projekt Prospekt*.

> Le travail de l'écriture prend également en compte le fait que l'élève rédacteur est par ailleurs lecteur d'autres textes. Il y puise des contenus, des formes et des stratégies d'écriture, il imite et transforme. (…) L'écriture, même imaginative et créative, se travaille avec des matériaux préexistants. L'enseignant souligne ainsi l'importance, dans toute production écrite, de la relation avec d'autres écrits mis en mémoire dans la communauté culturelle (Accompagnement du programme de français de 3ème : 174).

5.2.2.3. Programme officiel pour l'allemand et *Projekt Prospekt*

Lors de l'entretien oral en mai 2008, Géraldine, une apprenante, s'exprime au sujet du programme d'enseignement en allemand.

ENQUÊTRICE : Autre proposition ?
GÉRALDINE : Oui, sinon, enfin… (rit), ben, la dernière fois, j'avais dit à Madame Lacroix, parce qu'elle nous avait donné une fiche pour dire ce qu'on aimait bien et ce qu'on n'aime pas dans le projet, et moi, j'avais marqué qu'on n'apprenait pas vraiment la grammaire allemande, des choses comme ça. Donc, elle nous avait donné une fiche sur le passif qui correspondait à ce qu'on faisait en salle info, donc, avec les corpus et tout. Et ce serait bien si elle continuait à nous donner… faire des fiches comme ça.
ENQUÊTRICE : Des fiches pour chaque point grammatical qui correspond…
GÉRALDINE : …au programme par exemple, parce que là, on ne suit pas le programme, donc, par contre, je ne sais pas si ce qu'on fait en salle info, ça rejoint un peu. Là, on n'apprend pas vraiment la grammaire, donc, plus tard…

J'ai réagi à cette remarque de Géraldine en publiant, le 7 mai 2008, l'article suivant sur le blogue *esldeutsch*.

> *Projekt Prospekt* et le programme officiel d'allemand.
> J'ai regardé les films qui ont été faits lors des entretiens. Deux de vos préoccupations semblent s'y dessiner : les programmes officiels pour l'enseignement de l'allemand en fin de collège et particulièrement le programme grammatical. Notre projet s'appuie en effet sur les programmes. Je cite ci-dessous (…) les grandes lignes pour l'allemand :
> - "utiliser des supports d'apprentissage authentiques, attrayants et de qualité (...)" (authentique veut dire, ce qui n'est pas dans les livres d'allemand, mais qui a été produit pour les Allemands, donc les sites Internet que nous utilisons).
> - "avoir recours aux nouvelles technologies (Internet pour la recherche d'informations."

- "favoriser le travail en interdisciplinarité : établir des "ponts" entre la langue étudiée et les autres disciplines."

Le scénario de *Projekt Prospekt* s'appuie sur les programmes officiels pour l'enseignement de l'allemand au collège, palier 2, fin de collège (2007). Le tableau ci-dessous reprend de façon exhaustive la colonne "Compétence grammaticale/Compétences grammaticale et orthographique". J'ai mis entre parenthèses les points de grammaire qui ne font pas explicitement l'objet du projet. Le contenu des colonnes "Exemples de textes/Exemples d'énoncés" et de la colonne "Compétences culturelle et lexicale" sont seulement reprises si elles trouvent un écho dans notre projet. La colonne "Formulations" n'est pas du tout reprise ici. Elle contient des exemples de phrases et segments.

Compréhension de l'écrit			
Projekt Prospekt	**Programme officiel**		
	Exemples de textes	**Compétences culturelle et lexicale**	**Compétence grammaticale**
Instructions Les apprenants réalisent la tâche d'un dépliant et site touristique". Blogue, fiches de travail Production du texte "Commentaire dans le livre d'or d'un musée"	**Instructions** L'élève sera capable de comprendre ce qui est écrit pour réaliser une tâche : consignes ; courrier des lecteurs.	langage de classe	(*brauchen* + *zu*). Énoncé injonctif (l'impératif). Le superlatif de l'adjectif ; coordination avec *denn* [car] ; l'adverbe *trotzdem* [malgré cela].
Textes informatifs Identification d'éléments d'un texte du corpus : titre, informations pratiques, description, etc. Observation de sites touristiques et de dépliants touristiques Identification du passif dans la production du texte "Visite guidée" Observation de noms de musées (souvent des mots composés)	**Textes informatifs** L'élève sera capable de localiser des informations dans un texte assez long pour accomplir une tâche. Pages *web*. Documentaires.		(Préposition *trotz* [malgré]). Connecteurs logiques et temporels. (Génitif saxon). (Quantificateurs (*zahlreiche*, *manche*, *viele*, etc.)). Construction passive. Reconnaissance des mots composés.
Les apprenants ne lisent pas de textes de fiction, ni de biographies, mais il y a des éléments narratifs dans les textes de type "Visite guidée" et "Commentaire dans le livre d'or d'un musée". Il peut y avoir des éléments biographiques ("Visite guidée").	**Textes narratifs** L'élève sera capable de comprendre un enchaînement de faits dans des : textes de fiction ; biographies.		Prétérit. (Plus-que-parfait). Conjonction : *als*, *da*, *obwohl* [quand, lorsque, même si]. Place du verbe dans la subordonnée. Connecteurs temporels (Génitif).
Textes argumentatifs Les apprenants ne lisent pas de tracts, ni de magazines, mais les "Commentaires dans le livre d'or d'un musée" contiennent un schéma argumentatif.	**Textes argumentatifs** L'élève sera capable de reconnaître les grandes lignes d'un schéma argumentatif dans des : tracts ; magazines.		(*nämlich* [il se trouve que]).

Expression écrite			
Projekt Prospekt	**Programme officiel**		
	Exemples d'énoncés	**Compétences culturelles et lexicales**	**Compétences grammaticale et orthographique**
Reformulation Les apprenants reformulent des informations de plusieurs sources pour écrire les trois textes demandés. Ils utilisent leurs "propres mots", mais aussi des outils d'aide à la formulation, comme les concordances et les dictionnaires électroniques.	**Reformulation** L'élève sera capable de transmettre une information en utilisant ses propres mots.		Connecteurs logiques : *weil* [parce que] ; (*daher* [c'est pour cela que]).
Les apprenants n'écrivent pas de notes, ni de messages, mais le directif et le locatif font l'objet des textes de type "Informations pratiques" et le passif est utilisé dans les textes de type "Visite guidée".	**Notes et messages**		Directif. Locatif. Passif présent. (Lien phonie - graphie (St = cht, st = s-t)).
"Commentaire dans le livre d'or d'un musée" "Visite guidée" et "Informations pratiques" "Informations pratiques"	**Correspondances / lettres**	Règles du genre épistolaire	Genre et pluriel des substantifs. Subjonctif 2 des verbes de modalité. Datif (pronoms personnels et groupes nominaux). Prépositions suivies du datif.
Genre textuel qui n'est pas couvert par *Projekt Prospekt*, sauf éventuellement dans le "Commentaire dans le livre d'or d'un musée".	**Compte-rendu**		Pronoms réfléchis au datif. Pronoms réfléchis à l'accusatif. Parfait/Prétérit. Connecteurs temporels ; construction infinitive avec *um…zu* + infinitif [pour… + infinitif]. Prépositions suivies de l'accusatif.
"Visite guidée" Le projet ne propose d'écrire ni chant, ni poème. Les textes dans des corpus peuvent être considérés comme des exemples pour leurs productions écrites.	**Écriture créative** L'élève sera capable d'écrire : un récit ; une description ; un chant ; un poème, d'après un modèle.		(Infinitif substantivé). Déclinaison de l'adjectif épithète. Subordonnée relative. Pronoms indéfinis.

Tableau 17 – *Projekt Prospekt* et le programme officiel pour l'allemand, palier 2.

La comparaison des activités et supports proposés lors de *Projekt Prospekt* et de ceux préconisés par le programme officiel montrent une converge sur de nombreux points, sans que tout ait été traité de façon explicite. Pour prendre un exemple : l'apprentissage du genre et du pluriel des noms n'a pas été l'objet d'un "cours" ou d'une "leçon". Toutefois, j'ai publié un article sur le blogue *esldeutsch* qui informe sur la possibilité de trouver le genre et pluriel des noms à l'aide du traducteur *Reverso*. J'y ai joint un tutoriel concernant l'interprétation de la présentation des informations morphologiques données par ce dictionnaire électronique.

5.2.2.4. Valider certains points du b2i

Projekt Prospekt, projet principal de ma recherche, permet de faire valider certains items du B2i (Brevet informatique et Internet)[18] qui est une des composantes du DNB (Diplôme National du Brevet). Citons comme exemple les items qui concernent la recherche d'informations sur Internet, la gestion de fichiers et dossiers et le courrier électronique. Cette validation peut être effectuée par tous les enseignants, donc en l'occurrence par les enseignantes de technologie, de français et d'allemand.

5.2.3. Droit à l'expérimentation

Depuis 2005, l'article 34 du *Code de l'Éducation* donne officiellement le droit à l'expérimentation au sein de l'enseignement dans une école ou un établissement d'enseignement public du second degré.

> Droit à l'expérimentation
>
> Le projet d'établissement peut prévoir après autorisation des autorités académiques la réalisation d'expérimentations pédagogiques dans les domaines définis par l'article L. 401-1 du code de l'éducation : l'enseignement des disciplines, l'interdisciplinarité, l'organisation pédagogique de la classe, de l'école ou de l'établissement, la coopération avec les partenaires du système éducatif, les échanges ou le jumelage avec des établissements étrangers d'enseignement scolaire. Il s'agit d'encourager les équipes éducatives à exercer leur créativité et leur responsabilité pour proposer des démarches et des pratiques nouvelles de nature à contribuer à la réussite des élèves. Lors de la préparation du volet pédagogique du projet d'établissement, le conseil pédagogique étudiera les propositions d'expérimentation et vérifiera qu'elles sont en cohérence avec le projet global de l'établissement avant de les y inscrire. (…) Des expérimentations, nationales ou académiques, pourront également être proposées aux équipes pédagogiques qui pourront s'y engager volontairement avec l'accord des autorités académiques, dans le cadre de l'article 34 (Bulletin Officiel de l'Éducation Nationale, 2006).

[18] http://www2.educnet.education.fr/formation/certification/b2i/

Une analyse des dossiers déposés dans le cadre du droit à l'expérimentation fait ressortir sept domaines d'intervention (Muller et Teillard, nd). Si j'avais pu déposer un dossier, j'aurais souhaité que mon projet de recherche soit classé dans la rubrique "individualisation des parcours". Le rythme scolaire est aussi affecté par mon projet. En effet, les apprenants ont eu l'occasion de travailler à distance. Cela peut être considéré comme un assouplissement des horaires de travail.

	TYPE 1 expérimentation structurelle	TYPE 2 innovation d'établissement	TYPE 3 expérimentation pédagogique	TYPE 4 Innovation pédagogique	TOTAL	Relevant de Exp. Art 34
aménagement du rythme scolaire	9	1	0	0	**10**	8
individualisation des parcours	6	5	1	0	**12**	7
prise en charge de la difficulté scolaire	8	8	8	0	**24**	6
intégration des enfants à handicaps	2	0	3	0	**5**	3
ouverture internationale	5	1	0	0	**6**	6
détour culturel et artistique	2	1	3	2	**8**	3
enseignements scientifiques	1	1	3	0	**5**	1

Tableau 18 – Types de dossiers déposés dans le cadre du droit à l'expérimentation.

5.2.4. Projet d'établissement

J'ai choisi plusieurs points d'entrée susceptibles d'assurer le lien entre l'institution et le projet de création de trois textes pouvant s'insérer dans un dépliant ou site touristique (*Projekt Prospekt*, cf. 5.3.4.). J'ai fait un dossier pour faire apparaître le projet dans le projet d'établissement 2007/08 du collège. Je l'ai présenté en réunion plénière devant l'équipe administrative et pédagogique du collège. Le principal qui a exercé en 2006/07 dans l'établissement où j'ai mené la recherche n'a pas transmis le dossier à ses supérieurs hiérarchiques. Ceci est en contradiction avec le droit à l'expérimentation. C'est regrettable, car des projets interdisciplinaires sont attendus dans l'académie dont dépend cet établissement. Cela peut être comme déduit de la description des *Itinéraires de découverte* sur le site de l'Académie de Versailles.[19]

[19] Itinéraires de découverte : http://www.lettres.ac-versailles.fr/article.php3?id_article=698

5.3. Scénario pédagogique

5.3.1. Descriptif général du scénario pédagogique

Ma recherche s'adresse à des apprenants d'allemand en classe de troisième du collège. Ils ont commencé l'allemand en classe de sixième et sont donc en quatrième année d'allemand. L'objet de l'enseignement/apprentissage du projet est la révision individualisée de la production écrite. Je propose comme méthode de travail la sélection de segments pertinents dans un *input* écrit (corpus et concordances) qui correspond au types de textes que les apprenants sont censés produire et le réemploi (modifié ou non modifié) des segments lors de la révision. Les références et méthodologies qui sous-tendent ma recherche-action sont les suivantes : la recherche d'informations (Tricot, 2007 ; Rouet *et al.*, 2004), la théorie du repérage de Schmidt (identification de collocations et traitement en LE), l'approche par tâches (R. Ellis, 2003), l'idée que les connaissances se construisent individuellement, mais aussi à l'aide d'autres personnes (Piaget, 1969 ; Bruner, 1983), le *data driven approach* (Johns, 1991) et la conceptualisation (Chini, 2004).

Ma recherche s'est déroulée en trois phases. La première, la phase d'exploration, a eu lieu en 2006/07. Elle a servi à observer la façon dont les treize apprenants du groupe révisaient leurs productions écrites lors du projet *Création de dépliants touristiques*. À partir des données obtenues lors de cette phase, j'ai développé le scénario pour *Projekt Prospekt*, projet principal de la recherche, que j'ai soumis à un groupe de 17 apprenants en 2007/08. En 2008/09, j'ai créé un scénario qui a exploité les éléments de réponse à mes questions de recherche que j'ai obtenus lors des deux premières phases. J'ai proposé à 33 apprenants, répartis en deux groupes classe, un projet d'écriture d'annonces pour trouver un correspondant. Certaines des références qui ont sous-tendu mes actions ont progressivement perdu de leur pertinence. D'autres se sont avérées opératoires au fil de la recherche.

Projekt Prospekt (Création d'un dépliant touristique)	***Projekt Prospekt*** (Création de textes pouvant s'intégrer dans un site et/ou dépliant touristique)	***Brieffreunde.org*** (Écriture d'annonces comme celles que l'on écrit pour trouver des correspondants)
L'approche par tâches (Ellis, 2003) J'ai visé l'exécution d'une version radicale.	**L'approche par tâches** (Ellis, 2003) J'ai proposé une version modérée.	**L'approche par tâches** (Ellis, 2003) J'ai proposé une version modérée.
Formules et production écrite (Siepmann, 2004 ; Gaonac'h, 2006 ; Tutin et Grossmann, 2002 ; Šimunek, 2007). Acquisition de formules à court, moyen et long terme (Henrici, 1995).	**Formules et production écrite** (Siepmann, 2004 ; Gaonac'h, 2006 ; Tutin et Grossmann, 2002 ; Šimunek, 2007Aguado, 2002b).	**Formules et production écrite** (Siepmann, 2004 ; Šimunek, 2007).
Media literacy (Hallet, 2008) Cette compétence a dû être fortement développée auprès de certains apprenants.	***Media literacy*** (Hallet, 2008) Certains apprenants ont eu besoin de développer cette compétence plus que d'autres.	***Media literacy*** (Hallet, 2008) Cette compétence a été bien développée auprès de presque tous les apprenants.
	La linguistique de corpus et l'analyse discursive de textes (Henry et Roseberry, 1996 ; 2001 ; Mazière, 2005 ; Kennedy et Miceli, 2001) Le texte et ses sections ; son exploration au sein d'un corpus et à l'aide d'une concordance.	**La linguistique de corpus et l'analyse discursive de textes** (Henry et Roseberry, 1996 ; 2001 ; Kennedy et Miceli, 2001) Le texte et ses sections ; son exploration au sein d'un corpus et à l'aide d'une concordance.
	Le repérage d'informations (Schmidt, 1990 ; Johns, 1991)	**Le repérage d'informations** (Schmidt, 1990 ; Johns, 1991) .
	Produire un *output* compréhensible (Swain, 1985).	**Produire un *output* compréhensible** (Swain, 1985).
	Conceptualisation (Chini, 2004 ; Audin, 2008 ; Fuchs et Le Goffic, 1992).	**Conceptualisation** (Chini, 2004 ; Audin, 2008 ; Fuchs et Le Goffic, 1992).
		L'apprenant-chercheur (Johns, 1991 ; Bernardini, 2004).
		Recherche d'informations (Rouet *et al.*, 2004).

Tableau 19 – Références pour les trois phases de la recherche-action.

J'ai fait le choix de motiver les activités basées sur des concordances par l'exécution de tâches. Des exemples en sont la création de documents pouvant s'intégrer dans un dépliant touristique et l'écriture d'une annonce comme celles que l'on publie en ligne pour trouver un correspondant. À défaut de pouvoir m'appuyer sur des corpus allemands existants dans les domaines cités ci-dessus, j'ai créé des mini-corpus pédagogiques, en m'inspirant de Braun (2005), à partir de données trouvées sur la Toile pour lesquelles j'ai obtenu les droits d'auteur. J'ai ensuite soumis les données des corpus au concordancier *Text-based concordances (v2)* de Cobb (nd). J'ai publié les corpus et leurs concordances sur un de mes blogues pédagogiques afin que les apprenants y aient accès pendant et en dehors des séances d'allemand.

Dans ce qui suit, je décrirai les phases du processus rédactionnel de Hayes et Flower (1980 ; cf. 3.1.1) et ensuite le lien entre ces phases et le projet *Création d'un dépliant touristique*, rebaptisé ensuite *Projekt Prospekt*.

5.3.2. Phases du processus rédactionnel dans les projets de ma recherche

Planification

Lors de la phase de planification, je privilégie l'apprentissage par *exemplars*, aussi appelés *chunks*, formules, collocations ou segments (cf. 2.3.). Les apprenants sont invités à observer des textes de corpus et des concordances dans le but d'y identifier des sections et segments qui sont pertinents en vue de la préparation de leur production écrite. Ils surlignent un certain nombre de segments, ou ils les associent selon des critères clairement définis (cf. 5.3.7.3.1.).

Formulation

Lors de la phase de formulation, les apprenants rédigent leurs textes en s'inspirant des sections et segments observés. Ils les publient sur le blogue du projet, ou ils me les envoient par courriel.

Révision

Je relève dans les productions provisoires les passages qu'il convient d'améliorer. Je souligne ou surligne ces passages, et je rédige une note dans une colonne à part. J'encourage les apprenants à chercher la solution par eux-mêmes. Une des aides de révision possibles est l'utilisation d'une concordance.

Exécution

La mise en page et l'impression des productions écrites peuvent être réalisées pendant des séances de technologie.

5.3.3. Création de dépliants touristiques (2006/07)

La conception du projet *Création de dépliants touristiques* illustre une version radicale de l'approche par tâches (Ellis 2003). Ce projet a eu pour objectif la production de dépliants touristiques pour un musée ou établissement culturel français existant qui devait nécessairement avoir un site Internet sans version allemande. Les apprenants ont rédigé plusieurs textes en ayant eu le libre choix du sujet et des outils d'aide à la rédaction, à l'exception du traducteur électronique que déconseillé l'utilisation. Selon mon intention, les apprenants devaient créer ces dépliants dans le but de les offrir aux correspondants allemands qu'ils avaient l'occasion d'inviter pendant une semaine en mars 2007, juste après la fin du projet. À l'origine, les prospectus étaient censés préparer des correspondants allemands à des sorties avec les apprenants français. L'objectif de création d'un document papier n'a pas été atteint. Peu d'élèves ont accepté d'accueillir un correspondant en mars, et personne n'a manifesté son envie de faire un prospectus pour les Allemands accueillis par les apprenants des autres groupes. Le groupe classe a d'abord observé des dépliants touristiques allemands. Ensuite, j'ai proposé une initiation à la recherche de documents sur Internet et au dépôt de commentaires sur le blogue pédagogique du projet. Je n'ai pas donné de consignes précises en ce qui concerne les outils d'aide à la rédaction. Les apprenants ont rédigé leurs textes à l'aide de dictionnaires papier ou électronique. Certains ont utilisé des traducteurs. J'ai fini par déconseiller cette pratique, car les textes qui en ont résulté ont été difficiles à comprendre, à annoter et à réviser. Pour soutenir la révision, j'ai proposé aux apprenants des activités personnalisées, basées sur l'observation de lignes de concordance. Ces activités ont aidé certains apprenants à surmonter des difficultés dans le domaine linguistique ou discursif. En raison de l'obstacle qu'a représenté la littératie électronique pour plusieurs apprenants, j'ai réajusté certaines composantes du projet pour la deuxième phase de la recherche-action (cf. 5.4.2.2.).

5.3.4. Projekt Prospekt (2007/08)

Comme le projet de la première phase de la recherche-action, *Projekt Prospekt* avait comme objectif la création de documents touristiques. Les apprenants étaient invités à inventer un établissement culturel (musée, parc d'attraction, etc.) et à rédiger trois textes pouvant s'insérer

dans le site Internet virtuel de cet établissement. J'ai conçu le projet comme projet interdisciplinaire (allemand, français, technologie). Un groupe de sept apprenants sur dix-sept a effectivement eu l'occasion de travailler sur le projet pendant des séances d'allemand, de français et de technologie. J'ai légèrement modifié le scénario pour les dix apprenants dont les enseignants de français et de technologie n'étaient pas impliqués dans le projet. Ces dix apprenants ont eu comme tâche la rédaction de trois textes pouvant compléter les textes d'un site Internet d'un musée existant.

La première tâche, la création d'un texte du genre "Informations pratiques" d'un site Internet touristique, a demandé aux scripteurs peu de travail de planification. En raison du caractère très normé de la langue utilisée dans cette rubrique, les contraintes ont été fortes au niveau de la formulation. À l'aide de deux exemples proposés sur le blogue *Projekt Prospekt*, les apprenants ont observé les caractéristiques de la rubrique "Informations pratiques". Ensuite, ils ont manipulé les données de textes appartenant à ce genre textuel (cf. 5.3.7.3.1.), et ils ont rédigé la rubrique "Informations pratiques" pour leur musée ou établissement culturel.

La deuxième tâche, la création d'un texte du genre "Visite guidée", demande un temps de planification plus long. La formulation d'un texte de ce type est plus libre que pour les textes de type "Informations pratiques". Les apprenants ont d'abord observé quelques textes d'un corpus réfléchi afin de découvrir les sections des textes de type "Visite guidée" et d'observer les caractéristiques linguistiques de ces sections. Ils ont exploré le corpus réfléchi à l'aide d'une concordance afin d'être en mesure de formuler et réviser leur production écrite. La figure suivante est une capture d'écran du blogue *Projekt Prospekt*. L'article présente les informations pratiques d'un parc d'attraction.

Figure 32 – Informations pratiques d'un parc d'attraction.

La troisième tâche, l'écriture d'un "Commentaire dans le livre d'or électronique d'un musée", demande aux apprenants une prise en charge énonciative plus importante que lors des deux premières productions écrites. Les normes linguistiques sont plus libres que pour les deux premières productions. Les scripteurs de commentaires en ligne renoncent parfois aux majuscules des noms, ils font des phrases courtes, certains dédoublent les points d'exclamation, etc.

5.3.5. Brieffreunde.org (2008/09)

Le projet *Brieffreunde.org* [correspondants.org] a été initialement conçu comme un pré-projet, destiné à préparer les apprenants au projet de production d'une critique de film qui devait être le projet principal. Toutefois, *Brieffreunde.org* a pris de l'ampleur, et il est devenu ce que l'on pourrait appeler un "prototype viable". Lors de ce projet, les apprenants ont créé des annonces comparables à celles qui se trouvent sur des sites Internet destinés à mettre en contact des personnes qui cherchent des correspondants. La tâche n'est pas radicalement "réaliste", puisque les apprenants n'ont pas déposé leurs annonces sur le site dont provenaient les textes du corpus "Correspondants". Le point de départ a été une tâche proposée dans le chapitre 3 du manuel *Team Deutsch* (Esterl *et al.*, 2008 : 38). Nous avons rapidement constaté ensemble que les trois annonces proposées comme point de départ ne correspondent pas au profil de beaucoup d'apprenants et que le "corpus réfléchi" était donc trop petit. J'ai ensuite présenté le site *Brieffreunde.org*, et j'ai introduit mon corpus, portant le même nom. Après avoir observé des sections et leurs caractéristiques linguistiques et après s'être familiarisés avec le corpus et sa concordance, les apprenants ont rédigé leur annonce. La salle informatique a moins été au centre de ce projet que lors des phases antérieures de la recherche-action. Le vidéoprojecteur en salle de classe habituelle m'a permis de faire des démonstrations devant la classe entière, ce qui s'était révélé nécessaire lors des projets précédents. Les apprenants ont presque tous rédigé leurs textes à la maison, et ils les ont déposés sur le blogue. Le projet a été bref et intense, ce qui l'a rendu trop chronophage du point de vue guidage. Des concordances sélectionnées, prêtes à l'emploi, et peut-être un logiciel d'assistance à la correction de textes devraient permettre, lors d'une éventuelle reprise du projet, de faire gagner du temps à l'enseignant. J'ai observé une certaine stéréotypie des segments co-présents dans les productions et le corpus. Certains segments sont plus souvent co-présents que d'autres (cf. 6.4.1.1.), sans que cela se justifie par la quantité d'occurrences dans le corpus. Il serait souhaitable d'obtenir des productions plus "impliquées", c'est-à-dire qui correspondent à quelque chose dont les

apprenants veulent parler (exprimer leur opinion concernant les films que nous avons vus ensemble, par exemple).

5.3.6. Outils électroniques proposés aux apprenants

5.3.6.1. Corpus et concordanciers

5.3.6.1.1. Première et deuxième phase (2006/07 et 2007/08)

Pour la première phase, j'ai créé, à l'aide de l'outil *BootCat* (Baroni et Bernardini, 2003 ; cf. 4.1.4.) un corpus spécialisé dans le domaine du tourisme. J'ai toutefois renoncé à choisir ce corpus comme corpus principal du projet, ceci pour deux raisons. Premièrement, il aurait eu besoin d'être manuellement revu et corrigé afin de devenir un corpus "réellement" réfléchi (cf. 1.2.4.). Deuxièmement, j'ai estimé que l'utilisation du concordancier de *Sketch Engine* (Kilgarriff *et al.*, nd) permettant d'interroger ce corpus, était techniquement trop difficile pour les apprenants, compte tenu le temps que nous avions à notre disposition. En plus, l'interface d'interrogation des corpus sur *Sketch Engine* est en anglais, ce qui rajoute pour certains des apprenants des difficultés de compréhension. Toutefois, j'ai utilisé ce corpus afin de créer des activités de révision individuelle ou collective lors de la première phase de la recherche. Il n'est plus accessible à ce jour.

Lors de la deuxième phase de la recherche-action, j'ai principalement travaillé avec des corpus réfléchis, non annotés, que j'ai créés à partir de textes publiés en ligne pour lesquels j'ai obtenu les droits d'auteur. J'ai soumis les textes de ces corpus au concordancier *Text-concordancer* (v2) de Cobb (nd) afin de pouvoir mettre à disposition des apprenants des concordances sous forme de fichiers texte, publiés sur le blogue pédagogique accompagnant *Projekt Prospekt*. L'avantage de ce concordancier est son accessibilité et sa maniabilité pour des non-experts en informatique. Son inconvénient majeur est la gamme restreinte des types d'interrogation (cf. 1.6.4.).

5.3.6.1.2. Troisième phase (2008/09)

Lors de la troisième phase de la recherche-action (2008/09), j'ai utilisé trois types de concordanciers : *Konkordanzer-Corpus Deutsch (v5.1)*, proposé sur le site *Lextutor* de Cobb (nd), le concordancier de la collection d'outils *Antconc3.2.1w* d'Anthony (2007) et le concordancier de la collection d'outils *Lexico3* de Fleury *et al.* (2003).

À la lumière des critères de *Projekt Prospekt, Konkordanzer-Corpus Deutsch (v5.1)* a trois inconvénients : il est basé sur un corpus que l'on peut considérer comme petit, vu qu'il n'est pas spécialisé. On n'a pas accès aux textes sources, et la taille minimale d'une ligne de concordance est de cent caractères ce que je considère comme beaucoup, compte tenu de mon intention de soutenir une lecture ciblée lors du projet. Je l'ai fait utiliser de façon très ponctuelle.

AntConc3.2.1w est une collection d'outils convient aux types d'activités que j'ai proposés lors de ma recherche-action. Je l'ai fait utiliser lors d'un module d'apprentissage indépendant. Les apprenants de la troisième phase de la recherche ont observé la fréquence de mots dans la ballade *L'apprenti sorcier* de Goethe, afin de tester, à l'aide de concordances et de listes de fréquence de mots, mon hypothèse concernant le symbolisme numéral caractérisant ce texte. Les apprenants se sont rapidement appropriés certaines fonctions d'*AntConc3.2.1w* ce qui confirme mon intention de privilégier l'emploi de cette collection d'outils lors de futures recherches.

Je n'ai pas fait utiliser *Lexico3* par les apprenants lors de ma recherche-action. Sa prise en main est moins simple pour des non-experts en informatique que celle de *Text-Concordancer (v2)* de Cobb (nd) et d'*AntConc3.2.1w* (Anthony, 2007 ; cf. 1.6.4.). Je l'ai utilisé pour certaines analyses de mes données.

5.3.6.2. Outils de communication

5.3.6.2.1. Boîte courriel

J'ai offert aux apprenants la possibilité de me contacter en dehors des cours, en utilisant une adresse courriel créée pour le projet. J'ai fait créer les adresses par la "personne-ressource" (cf. 3.7.2.) du collège. Pour la première phase, j'ai proposé une boîte courriel à chaque apprenant. Pour la deuxième phase, j'ai souhaité simplifier la réception des courriels en les regroupant en une seule boîte utilisée par tout le groupe classe. Une deuxième motivation pour ce choix était mon intention de rendre les envois publics. J'ai espéré reproduire ainsi l'effet de visibilité qui est créé quand on envoie un commentaire sur un blogue.

5.3.6.2.2. Blogues pédagogiques

Les blogues permettent de publier des articles et des documents, sans avoir de connaissances techniques très avancées. Ils offrent aux lecteurs la possibilité de déposer des commentaires

concernant les articles. Lors de la recherche-action, j'ai utilisé des blogues non seulement pour publier des articles, mais aussi pour donner aux apprenants l'occasion de communiquer avec moi en dehors des cours et de déposer leurs travaux (productions écrites, activités d'entraînement). En parallèle, il était possible de me contacter par courriel. Selon les phases de la recherche-action, le blogue a été davantage utilisé que la boîte courriel et vice versa. Le blogue permet, davantage que la boîte courriel, de visualiser les travaux des apprenants. Je considère cela comme un avantage. Toutefois, les apprenants ne souhaitent pas toujours que leurs travaux soient publiés, et l'échange par courriel offre une plus grande confidentialité. Le besoin de sécurité semble davantage satisfait lors de ce type de communication. Dans ce qui suit, je présenterai les blogues que j'ai utilisés lors des trois phases de la recherche-action.

Première phase (2006/07)

Blogue Projektprospekt

Le premier blogue de la recherche-action, *projektprospekt*, ne comporte pas de rubriques (appelées "catégories" par l'hébergeur *hautetfort.com*). Les apprenants n'ont pas toujours déposé leurs commentaires à la suite à l'article auquel ils se sont référés. Ceci a causé des problèmes de repérage aussi bien aux apprenants qu'à moi-même : qui a écrit quoi, à quel sujet, et quand ? Dans le questionnaire distribué à mi-chemin du projet, Charles décrit ce problème de repérage comme suit : "Je trouve cela compliqué de s'y retrouver avec toutes les rubriques, boîtes e-mail… ."

Les apprenants ont utilisé le blogue pour déposer leurs textes et pour me poser des questions. L'envoi de productions écrites par courriel a été plutôt rare. J'ai observé un phénomène de collaboration : certains apprenants semblent avoir utilisé les productions de leurs camarades comme *input* (cf. 6.3.1.5.).

Deuxième phase (2007/08)

Blogue Tourismus

Le blogue *Tourismus* comporte les rubriques suivantes :

Das Projekt [Le projet]

Die Aufgaben [Les tâches]

Die Übungen [Les activités d'entraînement]

Technische Infos [Informations techniques]

Websites [Sites Internet]

Suite à un incident de type *spam*, provoqué par un apprenant qui n'a pas fait partie de la recherche-action, j'ai fermé le blogue *Tourismus*, afin de pouvoir passer à un abonnement de mon hébergeur de blogue me permettant de modérer les commentaires.

Blogue Projekt Prospekt

Ce nouveau blogue comporte les mêmes rubriques que le précédent, accompagnées de trois rubriques supplémentaires.

Grammatik [grammaire]

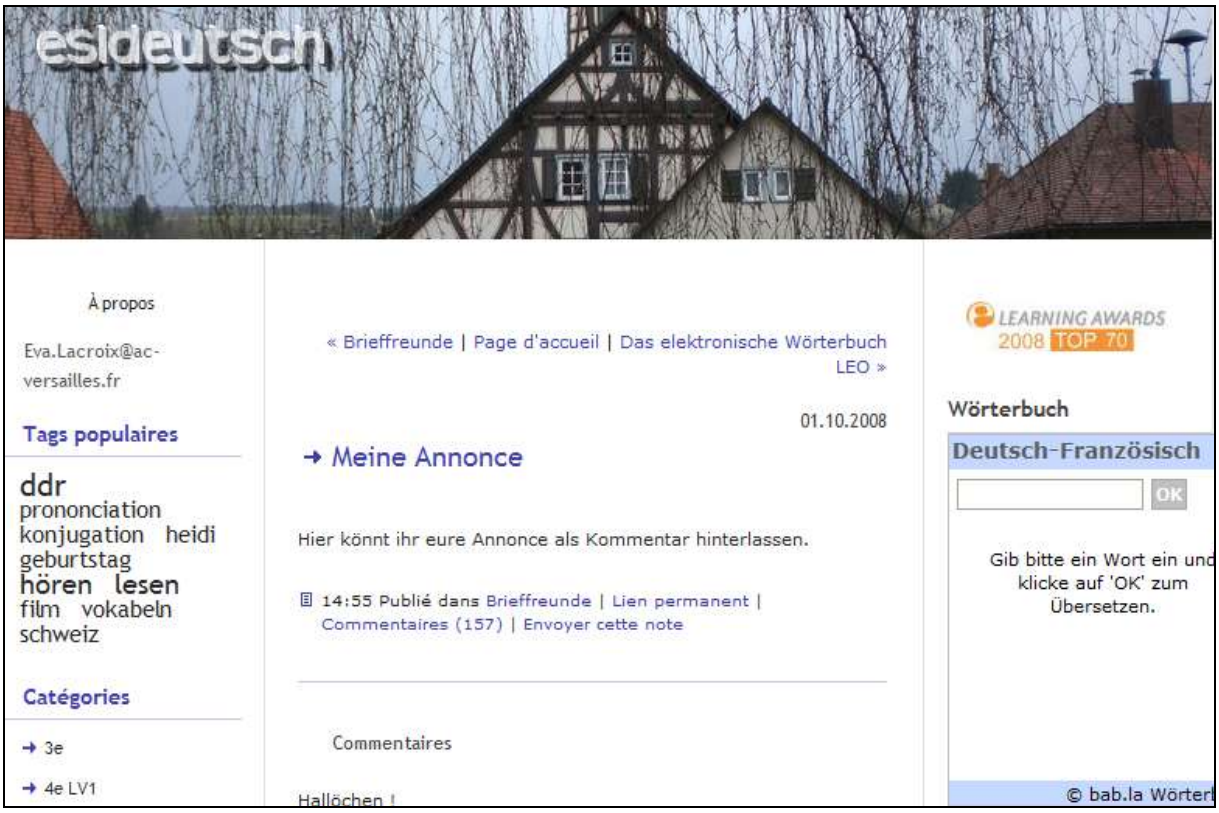
Kommunikation [communication]

Korpora und Konkordanzen [corpus et concordances].

Contrairement à la première phase de la recherche-action, les apprenants ont déposé très peu de commentaires sur le blogue *Projekt Prospekt*. Ils ont davantage communiqué par courriel ou sur papier (cf. annexe X). Je n'ai pas observé de phénomène de travail collaboratif basé sur l'emploi de la boîte courriel ou du blogue, et aucun apprenant ne s'est référé à ce qu'un autre apprenant a écrit sur le blogue. J'ai toutefois observé des situations de collaboration pendant certaines séances en présentiel (cf. 6.6.2.).

Troisième phase (2008/09)

Blogue esldeutsch



Lors de la troisième phase de la recherche-action, j'ai utilisé deux blogues différents. Le premier, *esldeutsch*, est celui que je propose à tous les groupes classe à qui j'ai affaire dans l'établissement où j'enseigne. J'ai hébergé le projet *Brieffreunde.org* (écriture d'annonces comme celles que l'on écrit pour trouver un correspondant) sur ce blogue afin de donner aux apprenants un accès aux ressources d'aide lexicale et grammaticale qui y sont proposées.

L'inconvénient de ce choix est l'éventuelle difficulté de repérage d'articles dédiés au projet que certains apprenants peuvent ressentir. En effet, les apprenants sont habitués à consulter la rubrique dédiée à leur niveau d'enseignement (rubrique nommée "3e"). C'est pour cela que j'ai créé une rubrique particulière, nommée *Brieffreunde*, pour le projet.

Blogue Filmblog

Le deuxième blogue proposé lors de la troisième phase de la recherche-action est réservé au projet d'écriture de critiques de film. Il propose des rubriques qui portent les noms de groupes de deux ou trois apprenants. Ainsi, l'endroit où il convient de déposer les productions écrites est spécifié. Les apprenants peuvent consulter de façon ciblée les travaux de leurs pairs, s'ils le souhaitent. L'inconvénient de ce blogue est l'absence de rubriques d'aide linguistique.

5.3.7. Activités préparatoires

5.3.7.1. Préparation aux outils et méthodes

Les nouvelles pratiques de formulation et de révision que les trois enseignantes qui ont participé au projet ont proposées aux apprenants nécessitent une familiarité avec l'outil informatique et avec la réflexion inductive qui dépassent le cadre généralement offert à ce niveau d'enseignement/apprentissage. Pour compenser cela, depuis le début de l'année scolaire 2007/08, j'ai encouragé les apprenants à utiliser en cours d'allemand le blogue *esldeutsch*, mis

en place en 2005. Depuis la classe de quatrième, le groupe qui a participé à la recherche-action en 2007/08 a été exposé à des lignes de concordance sur papier. Je leur ai proposé des activités de conceptualisation grammaticale et de recherche lexicale. "*Find the gloups*" (d'après Lamy et Mortensen, 2006) en est un exemple.

und berichtet davon in seinem	Gloups	.
Schreiben Sie Ihre Gefühle in ein	Gloups	, tauschen Sie sich mit anderen Betroffenen aus.
Er ist ein Idiot !	Gloups	
Heute, liebes	Gloups	möchte ich etwas über Freundschaften schreiben.
So, liebes	Gloups	, das wars für Heute!
was bin ich sauer .... 5. Oktober. Liebes	Gloups	, das war ja ' ne lange Zeit
27. April. Sorry, liebes	Gloups	, aber das waren harte Zeiten die letzten Tage.
So , endlich mal wieder Zeit für mein	Gloups	. Alle kritisieren mich!
Habe keine Zeit, um große Eintragungen im	Gloups	zu machen. Ich muss heute Abend um 18h weg.
habe ich mich dazu entschlossen, mein	Gloups	auf dieser Seite zu beenden.

Figure 33 – Activité de recherche lexicale "*Find the Gloups*".

L'activité de type "*Find the Gloups*" consiste à rechercher le mot clé qui a été enlevé dans une suite de lignes de concordance. Ainsi, lors d'un travail sur le complément du nom, les apprenants ont essayé de tirer des conclusions concernant les terminaisons du déterminant et du nom en observant des lignes de concordance.

5.3.7.2. Séance de présentation de *Projekt Prospekt*

Malgré le fait que mon projet n'a pas été déclaré officiellement "expérimentation au collège" au niveau de l'académie, j'ai pu le mettre en place dans mon établissement d'exercice sans rencontrer des résistances ouvertes, ni de la part de l'administration, ni des acteurs plus directement concernés, à savoir des apprenants et de leurs parents. Afin de m'assurer le soutien des acteurs concernés, j'ai pris soin d'informer auparavant les apprenants et leurs parents par courrier. J'ai organisé une réunion d'information pour les deux classes concernées.

La séance de présentation du projet *Projekt Prospekt* a eu lieu le 14 janvier 2008 dans la salle que j'occupe habituellement pour enseigner. Elle a été encadrée par l'enseignante de

technologie, et par moi-même. L'enseignante de français a eu un empêchement. Deux classes étaient présentes : les vingt-cinq élèves de la classe 3ème 2 que nous avions toutes les trois en commun, comportant sept germanistes, et les neuf germanistes de la classe 3ème 4. Il aurait été préférable que les seize germanistes ait été regroupés en une seule classe pour avoir les mêmes enseignants de français et de technologie et pour travailler dans des conditions comparables (voir les remarques d'Olivia dans l'entretien, cf. annexe I.2.5.). J'ai introduit la réunion par la question suivante, adressée aux apprenants. "Avez-vous déjà publié un texte ? Si oui, où et quoi ?" Une dizaine élèves ont confirmé écrire régulièrement sur un blogue. À la deuxième question, "Que veut dire "savoir rédiger", "avoir une belle plume" ?", j'ai obtenu les réponses suivantes.

Avoir un beau style et un contenu (Estelle, Camille).

Quand on rédige sur un blogue, on réfléchit rapidement et on ne se relit pas (Camille).

À la troisième question, "À quoi bon lire ?", j'ai pu récolter les réponses suivantes.

On lit pour avoir du vocabulaire, pour s'habituer aux belles phrases (Estelle).

Le style de la lecture nous influence (Camille).

Après cet échange, j'ai montré un diaporama illustrant les étapes du projet (cf. blogue *Projekt Prospekt*). Les élèves ont ensuite été invités à réagir à l'oral ou à l'écrit. Des remarques concernant les domaines suivants ont été faites : la durée du projet, le lieu d'exécution, l'organisation du travail (en groupe ou individuel ?), son apport en termes d'apprentissage pour le français, la difficulté d'exécution du projet et son évaluation. Certains apprenants ont demandé s'il faudrait faire un dépliant papier, blogue ou site, et si le site serait privé ou ouvert au public.

Lors de l'échange oral, certains apprenants ont fait repréciser la tâche.

Ah bon, il faut faire un dépliant et dire aux gens de venir et tout ça ?

On peut tout inventer ?

Véronique a écrit que la traduction risque de poser problème lors du projet.

Le projet me semble bien, mais peut-être assez dur au niveau de la traduction du français vers l'allemand.

Estelle, une élève qui n'apprend pas l'allemand, voulait savoir si le "Commentaire dans un livre d'or" serait la seule "rédaction personnelle".

5.3.7.3. Activités d'entraînement

Les activités d'entraînement que je présente ci-dessous se réfèrent presque exclusivement à la deuxième phase de la recherche. Les deux derniers points, 5.3.7.3.2. et 5.3.7.3.3., traitent de développements conçus pour la première, deuxième ou troisième phase de la recherche.

5.3.7.3.1. Repérage d'informations dans un texte appartenant à un genre textuel

Le scénario *Projekt Prospekt* que j'ai proposé lors de la première et deuxième phase de ma recherche-action privilégie la pratique de la langue écrite. Cette pratique s'appuie sur des textes d'un même genre textuel (les "Informations pratiques" d'un dépliant ou site touristique, par exemple), rassemblés dans un "corpus de la langue réelle", selon l'expression de Bailly (1998 : 50), ce qui équivaut, dans ma recherche-action, à un corpus réfléchi (cf. 1.2.4.). Les apprenants travaillent d'abord sur les caractéristiques discursives des textes. Ils repèrent, entre autres, des sections, nommées "séquences" chez Bronckart (1997 : 220). Parmi les sept sections présentées en 3.4.2.2., toutes, à part la séquence dialogale et le script, sont représentées dans au moins un des textes que les apprenants sont invités à produire lors de ma recherche-action. Des activités de manipulation des données ("glisser-déposer") et de repérage (surligner des sections) visent à familiariser les apprenants avec les éléments discursifs des textes.

Observation des caractéristiques linguistiques de sections d'un genre textuel

Les sections ont des caractéristiques linguistiques observables. Les apprenants observent, dans un corpus réfléchi, les faits de langue qui paraissent avoir une importance pour la rédaction d'un texte du même genre textuel. Par exemple, pour écrire des textes du genre "Visite guidée" que l'on peut trouver dans des documents touristiques, les apprenants ont besoin de maîtriser le présent, éventuellement le futur, et éventuellement l'imparfait, par exemple, si la visite dont ils parlent est de nature historique. L'impératif apparaît dans les titres de sections et/ou dans le corps du texte. Les scripteurs peuvent avoir besoin du passif qui apparaît fréquemment dans des textes allemands du type "Visite guidée".

Le tableau suivant juxtapose les sections et les caractéristiques linguistiques d'un texte de type "Visite guidée". La colonne de droite indique les mots clés de la concordance "Visite guidée" qu'il peut être utile d'observer afin de réviser son texte.

Sections de texte	**Caractéristiques linguistiques**	**Mots clés**
Titre Peut contenir… le thème ; le public ; le lieu.	Groupe infinitif Impératif Interrogative	
Accroche donner au lecteur envie de participer à l'activité ; lui permettre de voir si la visite pourrait lui convenir (en posant des questions).	*S'adresser directement au lecteur* : Impératif Interrogative	*du* *ihr* *euch* *wir* *uns* ! ?
Description de l'activité Ce que l'on fait lors de la visite. Ce qui se faisait autrefois (pour des sujets historiques). Informer sur ce qui est prévu, possible ou permis lors de la visite.	*Formulations personnelles* : au présent ; au futur ; à l'imparfait *Formulations impersonnelles* : Le passif Le passif combiné avec un verbe de modalité Le pronom impersonnel "*man*"	 formes du verbe "*werden*" + infinitif formes verbales + infixe **te** (verbes faibles) ou voyelle modifiée (verbes forts) et terminaisons □, st, □, etc. formes du verbe "*werden*" + participe passé *kann* *könnt* *soll* *sollt* (terminaisons □, st, □, en, t, en) *man*

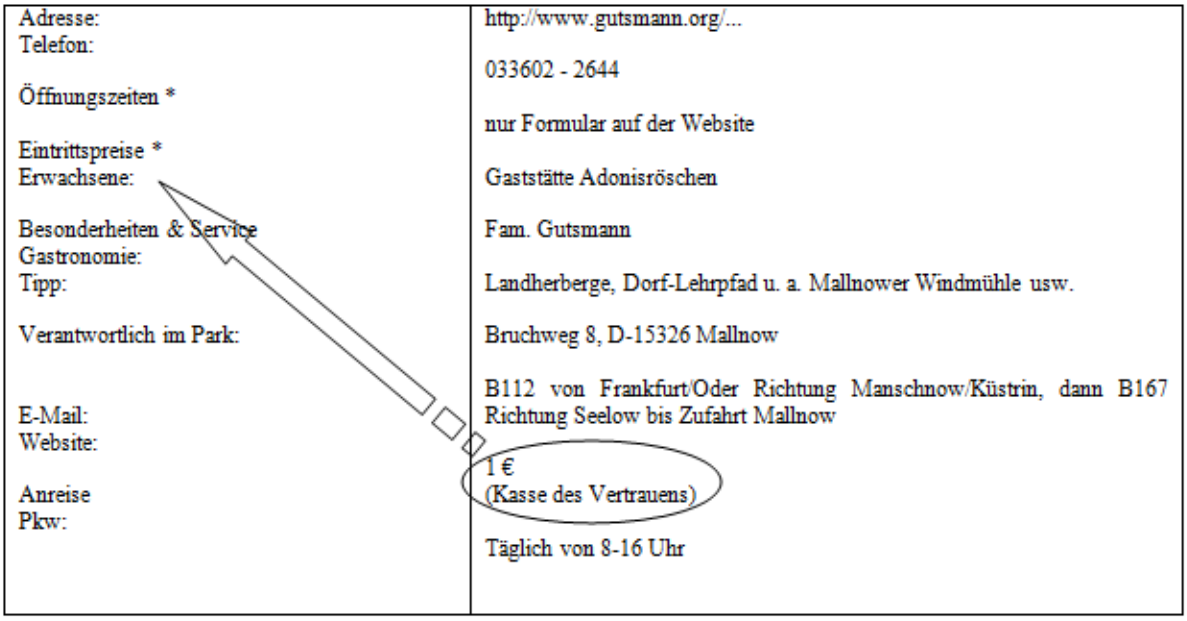
Tableau 20 – Caractéristiques linguistiques des sections d'un texte de type "Visite guidée".

Un exemple concernant les sections des textes du corpus *Brieffreunde*.org peut être consulté dans l'annexe VII.1.

Assembler des informations appartenant aux mêmes sections (Glisser-déposer)

L'activité appelée "glisser-déposer" a été créé afin de préparer les apprenants à la première tâche, la création d'un texte pour la rubrique "Informations pratiques" d'un site touristique. Pour cette activité, le contenu existe déjà. Il ne doit pas être inventé, mais manipulé. Lors de

cette activité, les apprenants assemblent des informations, et ils ont l'occasion de créer un exercice de type "glisser-déposer" pour un camarade. Cette double manipulation du contenu est susceptible d'habituer les apprenants aux caractéristiques linguistiques et discursives du genre textuel "Informations pratiques". L'activité de glisser-déposer est susceptible d'aider les apprenants à prendre conscience de certains segments. Savoir les repérer est un premier pas vers l'amélioration de la compétence de révision. L'exécution de la tâche "Informations pratiques" se trouve ainsi facilitée. Cette activité semble convenir à des scripteurs novices, pour reprendre le terme de Bereiter et Scamardalia (1987), puisqu'elle est peu exigeante au niveau de la planification et de la formulation. Ci-dessous, un exemple d'activité de type "glisser-déposer" illustre comment les apprenants ont fait glisser les segments du cadre de droite vers les éléments correspondants dans le cadre de gauche.



Adresse:	http://www.gutsmann.org/...
Telefon:	033602 - 2644
Öffnungszeiten *	nur Formular auf der Website
Eintrittspreise * Erwachsene:	Gaststätte Adonisröschen
Besonderheiten & Service	Fam. Gutsmann
Gastronomie: Tipp:	Landherberge, Dorf-Lehrpfad u. a. Mallnower Windmühle usw.
Verantwortlich im Park:	Bruchweg 8, D-15326 Mallnow
E-Mail: Website:	B112 von Frankfurt/Oder Richtung Manschnow/Küstrin, dann B167 Richtung Seelow bis Zufahrt Mallnow
Anreise Pkw:	1 € (Kasse des Vertrauens)
	Täglich von 8-16 Uhr

Figure 34 – Activité de type "glisser-déposer".

Selon le planning, les apprenants devaient créer des activités de type "glisser-déposer" pour leurs camarades lors d'une séance de technologie. Cette activité devait les habituer au lexique très contraint des "Informations pratiques". Comme la séance de technologie n'a pas eu lieu, les apprenants qui avaient terminé leur travail avaient l'occasion de créer des activités pendant une séance d'allemand ou à la maison. Pour ce faire, je les ai invités à s'aider d'un tutoriel (cf. annexe V.2).

Repérer et surligner des sections

Pour l'exécution de la deuxième production écrite, la "Visite guidée", j'ai proposé une activité de surlignage de sections.

"Führungen und Workshops für Kinder und Jugendliche"
Aufgabe : Textbausteine identifizieren (identifier des sections de texte)

1. Trouvez le titre (et le sous-titre, s'il existe). Surlignez-le(s) en bleu.
2. Trouvez les passages dans lesquels on l'on s'adresse à quelqu'un (2ème personne). Il peut s'agir entre autres de questions ou bien d'une formulation à l'impératif. Surlignez ces passages en rouge.

3. Trouvez la description de l'activité. Surlignez-la en vert.
4. Trouvez la section "Informations pratiques". Surlignez-la en rose.
5. Trouvez les recommandations qui sont prononcées dans certains des textes. Surlignez-les en jaune.

Voici un exemple :

Spukschlösser und Geheimgänge – 3D-Welten entwerfen mit einem Spieleditor
ab 12 Jahren
19.-20.03., Mi-Do, 2-tägig, 11-17 Uhr
Wie entstehen eigentlich die fantastischen Welten der Computerspiele? In diesem Kurs kannst du selbst geheimnisvolle Gebäude mit Treppen, Türen und Geheimgängen entwerfen – auch Rätsel und Tricks lassen sich einbauen. Deine eigene Welt kannst du auf einer CD mit nach Hause nehmen. Vorkenntnisse sind nicht erforderlich. Räumliches Vorstellungsvermögen und ein geübter Umgang mit dem Computer sind hilfreich.
Achim Däschner, Medienkünstler
Kosten € 60, max. 7 Pers.

Figure 35 – Activité de type "surlignage de sections de texte".

Repérer des avis positifs, négatifs et mitigés

Afin de préparer les apprenants à la dernière production écrite, le "Commentaire dans le livre d'or électronique d'un musée", j'ai fait identifier les avis positifs, négatifs et mitigés. J'ai proposé aux apprenants de marquer les sections identifiées par une binette.

Observez ces commentaires dans le livre d'or électronique d'un musée.

1. Quelles sections reconnaissez-vous ?
2. Quels moyens langagiers sont utilisés (temps, mode, vocabulaire, orthographe, signes typographiques, etc.) ?
3. Rajoutez devant les diverses sections les binettes qui conviennent (flèche = proposition).

Figure 36 – Activité de type "rajout de binettes aux sections identifiées".

5.3.7.3.2. Démonstrations de techniques de recherche d'informations

Lors de la deuxième phase de la recherche-action, j'ai proposé des démonstrations de recherche d'informations au groupe classe et des séances individuelles d'entraînement. Plusieurs activités ont eu pour but l'observation de corpus et concordances, comme par exemple l'activité suivante, destinée à soutenir la formulation du nom du musée ou établissement culturel. J'ai indiqué, sur cette fiche de travail, des combinaisons de mots clés précises, susceptibles de faire trouver des noms de musées dans le corpus *deWac*. J'ai donné des exemples de termes (enfants, nature, arts) permettant de trouver des noms de musées liés à des publics ou thèmes particuliers (musées pour enfants, écomusées, musées artistiques, etc.) sur le site *Deutsches Museumsverzeichnis* (cf. annexe V.1).

Pour guider les apprenants lors de la troisième phase de la recherche-action, j'ai rédigé une feuille de route. Une partie de cette feuille de route contient des consignes concernant la recherche de segments pertinents pour la révision du texte de type "annonce pour trouver un correspondant".

Comment exploiter la concordance pour améliorer ton annonce ?

Trouve un mot clé utile.

Exemple :
Mot clé "*Geburtstag*" [anniversaire] pour corriger l'expression "**Mein Geburtstag ist der 3 mai*".

Repère la formulation qui convient. Pour notre exemple, nous trouvons la ligne de concordance suivante :

und bin 13 Jahre alt. Ich habe am 2. November ***GEBURTSTAG****. Ich wohne in Tondern*
Marque la formulation qui convient et copie-la dans ton document de type *Word*.
und bin 13 Jahre alt. Ich habe am 2. November ***GEBURTSTAG****. Ich wohne in Tondern.*

Adapte la formulation à ce que tu veux dire : *Ich habe am 3. Mai* ***GEBURTSTAG****.*
Enregistre ton document après chaque modification et envoie le résultat final à Mme Lacroix.

Figure 37 – Procédure d'identification et de réemploi d'un segment lors de la révision de texte.

Une partie d'un enregistrement de la séance lors de laquelle j'ai présenté cette feuille de route peut être écoutée sur le blogue *esldeutsch*.

5.3.7.3.3. Observation de lignes de concordance

Lors de la rédaction, des besoins d'apprentissage individuels se sont révélés. J'ai répondu à ces besoins en créant des tâches basées sur des concordances que j'ai mises en ligne sur le blogue

afin de les rendre accessibles à tous. En fonction de ce que j'ai relevé dans les versions provisoires des productions écrites des apprenants, je les ai encouragés à consulter directement les corpus et concordances, souvent en leur conseillant un ou des mots clés dont j'ai testé auparavant la pertinence. L'exemple ci-dessous illustre le guidage proposé à Pia lors de la première phase de la recherche-action. Elle avait à réviser un groupe nominal qui comportait deux erreurs de détermination. La préposition *mit* doit être suivie d'un article défini, et au lieu du génitif saxon (*Ritters*), il convient d'utiliser le cas du génitif (*des Ritters*). La traduction des consignes a été faite pour favoriser la lisibilité de la thèse.

**__mit__ Ritters __Hilfe__*

a*) Kopiere die Konstruktion "mit Ritters Hilfe" und suche sie im Google [garde les guillemets]. Was findest du? Suche eine andere Konstruktion in der Konkordanz.*	a) Copie la construction "*mit Ritters Hilfe*" et cherche-la à l'aide de *Google*. Que trouves-tu ? Cherche une autre construction dans la concordance.
b) "Mit der Hilfe" oder "Mit Hilfe"? Argumentiere.	b) "*Mit der Hilfe*" ou "*Mit Hilfe*" ? Argumente.

muss eine von drei Schriftrollen finden und mit ihrer	Hilfe	eine der drei magischen Waffen zu erwerben,
gegen den Drachen vor dem Turnier mit der	Hilfe	eines sehr starken Spielers vorbereitet hatte.
mit dem sagenhaften Elfenbein - Horn OLIFANT zu	Hilfe	ruft , um Unterstützung im Überfall der
Haar. Sie verliebt sich in ihn , und er vollbringt mit	Hilfe	des Pferdes Taten, durch die er schliesslich
Held heiratet die Königstochter, weil er mit	Hilfe	des wilden Mannes drei Jahre lang die Hasen
dem Meer . - Grossrussisch aus Rjäsan: Pferde mit	Hilfe	des wunderbaren Männleins geheilt,
der Rittoboter als würdiger Ritter erweisen. - Mit	Hilfe	eines Katapultes soll ein böser Drachen verjagt
soll ein böser Drachen verjagt werden. - Mit	Hilfe	von Merlin muss das Burgtor geöffnet und

Figure 38 – Révision guidée à l'aide de lignes de concordance.

En plus du travail individualisé, j'ai proposé des activités collectives. Elles s'appuient sur des lignes de concordance que j'ai sélectionnées en fonction des besoins d'apprentissage linguistiques et discursifs. J'ai mis les activités à disposition des apprenants, soit sur papier, soit en ligne. Le tableau ci-dessous donne un aperçu des activités collectives que j'ai proposées lors de la troisième phase de la recherche-action. J'y présente les problèmes que les apprenants avaient à résoudre. J'indique également le domaine (linguistique, discursif, etc.) concerné par la révision, et je mentionne les mots clés et les corpus employés pour créer les tâches.

Question de travail	**Mots clés**	**Domaines traités**	**Corpus**
Quand faut-il mettre un "e" avant "…*sten*" dans les adverbes au superlatif ?	"**esten*"	Orthographe.	*DeReKo* (Cosmas II)
Comment dit-on "cours" [scolaire] ?	"*Erdkunde*"	Lexique (lexème simple).	*DeReKo (Cosmas II)*
Comment dit-on "J'aime nager" ou "J'adore danser" ?	"*liebe*" "*mag*" "*gern*"	Lexique (combinaison de lexèmes). Syntaxe (place de "gern").	*DeWac*
Comment dit-on "x habite avec y" ?	"*wohnt mit*"	Lexique (combinaison de lexèmes). *Wohnen + Person + Ort* ou *wohnen + Person + zusammen.*	*DeWac*
Comment dire où l'on habite par rapport à un autre endroit ?	"*wohne*"	Lexique et grammaire. Prépositions locales. Observation d'occurrences, dans le but de montrer sur une carte de Berlin où quelqu'un habite.	*DeWac*

Tableau 21 – Activités de révision collectives.

Dans ce qui suit, je présente une activité d'observation du segment "*wohnt mit*" [habite avec] faite à l'aide de *Sketch Engine* (Kilgarriff *et al.*, nd). Le groupe classe a observé les lignes de concordance afin de trouver comment on peut dire en allemand l'équivalent de "x habite avec y". J'ai pris comme point de départ du travail un énoncé proposé par Cathy.

**Er wohnt mit seine Mutter und seine Schwester.* [Il habite avec sa mère et sa sœur.]

Unser Sohn	**wohnt mit**	Familie in der Nähe von Linz/Rhein.
Der sechsjährige Daniel	**wohnt mit**	seinen Eltern in Stuttgart.
Die Mutter von fünf Kindern	**wohnt mit**	ihrer Familie in Milagro.
Die kleine Hexe Piccolina	**wohnt mit**	ihrer schwarzen Katze Tabata im Zauberwald.
Sie	**wohnt mit**	der Mutter im Elternhaus.
Ein Bekannter	**wohnt mit**	seiner Familie in einem Hochhaus.
Sie	**wohnt mit**	einem Franzosen zusammen.
Raymond	**wohnt mit**	seinem Bruder zusammen.
Die sechzehnjährige Lilja	**wohnt mit**	ihrer Mutter in einer Mietskaserne.
Sherman McCoy	**wohnt mit**	seiner Frau und seiner Tochter in Manhattan.
„Gut die Hälfte meiner Freunde	**wohnt mit**	anderen zusammen" , sagt Adela.

Figure 39 - Exemple d'une tâche exécutée collectivement.

Cette tâche a aidé certains apprenants du groupe classe à détecter deux cas de figure différents (*wohnen* + personne + lieu ou *wohnen* + personne + *zusammen*), et, en passant, à se rendre compte de la terminaison de datif appropriée. Suite à cette activité d'observation, Cathy, l'apprenante qui avait écrit "**Er wohnt mit seine Mutter und seine Schwester*", a été en mesure de réviser son texte en fonction de ses observations ou de celles formulées par un de ses camarades.

5.4. Documents analysés

Les documents que j'ai recueillis en vue de les analyser sont les suivants.

– Les scripts des entretiens filmés.

– Les questionnaires écrits.

– Les scripts de certaines séances en présentiel (séquences que j'ai filmées pendant deux séances ; enregistrements audio pour de nombreuses séances).

– Les échanges écrits : courriels et commentaires sur le blogue.

– Les différentes versions des textes produits par les apprenants.

– Mes annotations des productions écrites.

– Mes journaux de bord.

5.4.1. Entretiens

Le 6 mai 2007, une enquêtrice extérieure a mené des entretiens auprès de quinze des dix-sept apprenants qui ont participé à la deuxième phase de la recherche-action. Elle a interrogé cinq groupes de trois apprenants pendant vingt minutes environ. Les entretiens ont été filmés à l'aide d'une caméra fixe, et un dictaphone a tourné en même temps afin d'assurer la qualité sonore des enregistrements. Les questions que l'enquêtrice a posées aux apprenants ont visé la possibilité de mener une enquête compréhensive et semi-dirigée ; elles ont été établies en fonction de critères évoqués chez de Singly (1992 : 27 – 41). Dans l'annexe VIII.1, on peut consulter le questionnaire et les réponses auxquelles l'enquêtrice et moi-même nous sommes attendues.

5.4.2. Questionnaires

5.4.2.1. Première phase

Lors de la première phase de la recherche-action, j'ai récolté, au bout de dix séances de travail, les avis écrits des apprenants concernant ce que j'ai appelé "les points forts et les

points faibles du projet". Les aspects évoqués par les apprenants touchent surtout à la surcharge cognitive, la gestion du temps et l'utilisation de l'ordinateur. Assez peu d'apprenants parlent du contenu du projet et des apports éventuels qu'ils ont perçus. Suite à ces informations, j'ai décidé de proposer, lors de la deuxième phase de la recherche-action, un travail interdisciplinaire (allemand, français, technologie) afin d'alléger la masse de travail en la répartissant dans trois matières.

5.4.2.2. Deuxième phase

Lors de la deuxième phase de la recherche-action, j'ai proposé aux apprenants deux questionnaires écrits. Le premier a été rempli au bout de 12 séances de travail. Ce questionnaire reflète le fait que, au stade où j'ai rédigé les items, j'ai oscillé entre deux postures, celle d'enseignante et celle de chercheuse (cf. 6.8.2.2.). J'ai distribué le deuxième questionnaire lors de la dernière séance de *Projekt Prospekt*. Il contient des questions fermées, semi-ouvertes et ouvertes (cf. annexe VIII.2).

5.4.3. Échanges en présentiel

Selon les phases de la recherche-action, les trois heures hebdomadaires dont j'ai disposé avec les groupes d'allemand ont été partiellement ou entièrement consacrées au projet. J'ai régulièrement débuté les séances en présentiel par un moment d'information et d'échange, en allemand ou en français, selon la situation. Si les séances ont eu lieu en salle informatique, j'ai arrangé l'espace de sorte que tous puissent s'asseoir autour d'un seul bloc de tables. Ainsi, j'ai essayé de donner aux débuts de séance un caractère de réunion de travail conviviale. Les apprenants ont ensuite travaillé de façon individuelle en utilisant chacun un ordinateur. Parfois, je me suis occupée d'un apprenant en particulier lors de ces séances (cf. 6.3.2.2.). Il est aussi arrivé que je donne rendez-vous à quelques apprenants avant l'heure officielle de la séance afin de pouvoir prendre le temps pour eux dans une situation où j'ai estimé ils avaient besoin d'un soutien particulier (cf. 6.8.1.).

5.4.4. Échanges à distance

5.4.4.1. Courriels

Les apprenants ont eu l'occasion de me transmettre par courriel leurs productions, leurs questions et leurs remarques. En ce qui concerne *Projekt Prospekt*, j'ai sauvegardé toutes les productions écrites envoyées par courriel et les traces d'une partie des échanges de type

"questions et remarques". Vers la fin de la recherche, j'ai appris que la boîte courriel qui nous a été mise à disposition par le serveur de l'académie de Versailles allait subir des travaux de maintenance et qu'elle allait fermer dans les deux heures qui suivaient cette information. En l'espace d'une seule matinée, j'ai perdu l'accès aux données de type courriel. La vérification des dates et nombres d'envoi de courriels est ainsi devenue impossible.

5.4.4.2. Articles et commentaires sur le blogue

J'ai rédigé des articles sur les blogues afin de mettre à disposition des apprenants des informations concernant le déroulement du projet, afin de répondre aux questions qu'ils posaient en présentiel ou sur le blogue, et afin de leur donner accès aux corpus, concordances et activités du projet. J'ai diffusé sur le blogue des aides de type linguistique ou technique, et j'ai proposé des liens externes vers des sites susceptibles de soutenir leur travail. J'ai encouragé les apprenants à déposer des commentaires sur le blogue afin de poser des questions, de faire des remarques et de diffuser les diverses versions de leurs productions écrites. Les tableaux ci-dessous comparent les moyens de communication utilisés par les apprenants des trois phases de la recherche-action. Le nombre d'articles et de commentaires produits par les apprenants et par moi-même lors des divers projets est détaillé dans l'annexe X.

Première phase	**Deuxième phase**	**Troisième phase**
Échanges en présentiel. Envoi des textes par *courriel*. Adresses fournies par le collège. J'avais accès aux boîtes courriel des élèves.	Échanges en présentiel. Envoi des textes par *courriel*. Adresses fournies par le collège. J'avais accès aux boîtes courriel des élèves.	Échanges en présentiel. 3e LV2 : la première version des productions écrites sous forme papier, le reste a été déposé sur le blogue. 3e LV1 : toutes les versions ont été déposées sur le blogue. Certains élèves de 3e LV1 ont rendu l'une ou l'autre version sous forme papier.
Blogue projektprospekt (avec mot de passe). Sans rubriques. Transmission des textes principalement par blogue.	Transmission des textes principalement par courriel. Blogue corpus (avec mot de passe : corpus / wanda). Avec rubriques. Transmission des textes occasionnellement par voie de blogue.	Blogue esldeutsch.

Tableau 22 – Comparaison des moyens de communication utilisés par les apprenants.

Le groupe de la première et de la troisième phase de la recherche-action a largement investi ce moyen de communication. Le groupe de la deuxième phase de la recherche-action a davantage communiqué avec moi par courriel que par publication de commentaires.

5.4.4.3. Annotations de productions écrites des apprenants

J'ai annoté les productions écrites de façon plus ou moins détaillée. Parfois, j'ai simplement relevé un segment ou passage à revoir, en le soulignant, l'entourant ou le surlignant. Au fil de l'avancement de *Projekt Prospekt*, j'ai opté pour des annotations plus explicites, allant de l'indication d'un mot clé à rechercher dans le corpus jusqu'à une remarque de nature linguistique ou discursive, voire une consigne précise.

KAMPFSPORT FÜR KINDER IN PARIS	
Eintrittspreise: - Erwachsener : 10E *- Kinder : - (14-18 Jahre) : 8E* *- (6-13 Jahre) : 5E* *Öffnungszeiten : täglich 10-18 Uhr* *Das Museum ist Feiertag geschlossen.*	Pluriel à mettre. Mettre le symbole pour € (Touche alt gr + e) La concordance "Praktische Informationen" à consulter (mot clé "Feiertag") pour corriger ce passage.
Adresse : 15 Avenue des Camps-Elysés *PARIS 75 000* *tel:0139950203* *fax:0139950204* *e-mail: kampfsport@yahoo.fr* *site: www.kampfsport.paris.com*	Orthographe
Magst du Kampfsport ? Komm ins Kampfsportmuseum und du wirst viel Sports kennen,	Pour trouver le mot qui correspond à "différents types de sport", va sur la page Internet http://www.sportschule-ff.de/sport/index_sport.php et cherche l'endroit où l'on présente les sports.
für Kinder ab 6 Jahre.	Place cette information ailleurs dans le texte, ou supprime-la.
Es gibt 13 Kampfsportstanden.	Le pluriel de "Stand" à trouver (voir démo sur le blogue, catégorie "Grammatik").
Du wirst mit Kampfsportlehrerinen diskutieren.	Orthographe. Tu as choisi la forme féminine du métier. Est-ce exprès ?
In das Museum, du wirst mit Lehrerinen Kampfsport zu praktizieren oder,	Autre cas pour "das". Laisse tomber "zu".
für den Ruhen Kampfsporthistorische aufzudecken. *Bis bald!*	Je ne comprends pas. Explique-moi ce que tu voulais dire.

Figure 40 – Annotation du texte "Visite guidée, V3" de Simon.

5.4.5. Journal de bord

J'ai tenu un journal de bord lors de la première et de la deuxième phase de la recherche-action. Celui de la deuxième phase retrace de façon détaillée la plupart des séances de *Projekt Prospekt*. Dans le chapitre suivant, je citerai des parties du journal de la deuxième phase afin de contribuer à une triangulation de l'observation (Demaizière et Narcy-Combes, 2007).

5.5. Méthode d'analyse des données

5.5.1. Relevés faits à l'aide de collections d'outils de statistique textuelle

Je soumets les scripts des entretiens et des séances individuelles et mon journal de bord à une analyse de statistique textuelle. Pour ce faire, j'utilise les collections d'outils *Lexico3* (Fleury, Lamalle, W. Martinez et Salem, 2003) et *AntConc3.2.1w* (Anthony, 2007).

5.5.1.1. Identification des thèmes abordés par les acteurs de la recherche dans les documents recueillis

Suite à la lecture des scripts de séances et entretiens, des questionnaires et de mon journal de bord, j'établirai une liste de thèmes abordés par les acteurs de la recherche quand ils parlent des projets auxquels ils ont participé. J'analyserai les données en recherchant la fréquence du lexique lié à ces thèmes. Je cherche, entre autres, à savoir combien de fois chaque thème est évoqué dans le discours des apprenants et dans mon discours d'enseignante et de chercheuse. La mise en évidence de ces thèmes permettra de rendre compte du lien qui existe, à un degré variable selon les points abordés, entre le cadre théorique et les données récoltées.

5.5.1.2. Fréquence lexicale dans les productions écrites des apprenants

Je fais appel aux fonctions "*Concordance*" et "*Word List*" [liste de mots] *AntConc3.2.1w* (Anthony, 2007) pour comparer le lexique d'un corpus d'apprenants au lexique d'un corpus réfléchi, servant de corpus de référence. Cela me permet d'évaluer des tendances concernant la rédaction des apprenants, comme par exemple la sur-ou sous-utilisation de certains lexèmes et de signes de ponctuation.

5.5.1.3. Localisation de passages co-présents dans les corpus de référence et les corpus d'apprenants

AntConc3.2.1w propose la fonction "*Concordance plot*" (cf. 1.6.2). Cette fonction permet de localiser des passages dans la production d'un apprenant et dans le corpus auquel il a eu accès. Les passages co-présents sont ainsi visuellement mis en évidence. Grâce à cette fonction, je peux évaluer dans quelle mesure les apprenants ont repéré et réutilisé des segments pertinents dans le corpus ou la concordance. Si les passages co-présents varient légèrement – par exemple, ce qui dans le corpus de référence est au singulier est au pluriel dans le texte de l'apprenant –, il est possible d'avancer que l'apprenant a été en mesure d'effectuer une lecture ciblée de l'*input* et de tenir compte de la spécificité de l'environnement du segment dans sa production d'apprenant par rapport à l'environnement de ce segment dans le corpus réfléchi.

5.5.1.4. Fréquence des prises de parole dans les entretiens

Dans le script des entretiens, je relève, à l'aide de *Lexico3* (Fleury *et al.*, 2003), le nombre des prises de parole des apprenants. J'étudie aussi le nombre de fois où l'enquêtrice s'adresse aux apprenants lors de l'entretien, en utilisant leur prénom ou en s'adressant à eux à l'aide d'un pronom personnel de deuxième personne. Je juxtapose ces chiffres au nombre de mots prononcés par les apprenants lors de la prise de parole. Ce relevé servira à nuancer mes observations concernant leur décentration lors de l'exécution de *Projekt Prospekt* (cf. 6.9.).

5.5.2. Relevés faits à l'aide du traitement de texte

5.5.2.1. Comptage du nombre de versions et du nombre de mots des productions d'apprenants

Pour certains relevés, je fais appel à des outils proposés par le traitement de texte *Word 2007*. Je relève, par exemple, le nombre de versions des productions écrites des apprenants. À l'aide de la fonction "Statistiques", je calcule, entre autres, le nombre de mots des productions des apprenants.

5.5.2.2. Comparaisons de versions

Je compare les différentes versions des trois productions écrites pour *Projekt Prospekt* à l'aide de la fonction "Comparer" du menu "Révision" du traitement de texte *Word 2007*. Cette fonction permet de localiser et de compter les insertions, suppressions, déplacements, les modifications concernant la mise en forme et les commentaires ajoutés. Je fais deux lectures

des productions écrites. Lors de la première lecture, je recense le nombre des insertions, suppressions et déplacements de caractères, lexèmes ou segments, sans tenir compte ni des modifications concernant la mise en forme, ni de la ponctuation. Pour ce faire, je décoche les options "mise en forme", "modification de la casse" et "espace". La deuxième lecture est consacrée à l'observation des modifications dans le domaine de la ponctuation dans les productions écrites des apprenants (cf. 6.5.1.2.).

5.5.3. Relevé fait manuellement : évolution de caractéristiques linguistiques des productions écrites

J'observe les modifications dans le domaine de l'adjectif (absence ou présence de terminaisons, place de l'adjectif) et de la syntaxe du groupe verbal et nominal afin de rendre compte du degré de complexification des productions écrites des apprenants en fonction de leur utilisation ou non-utilisation des corpus et concordances.

5.5.4. Analyse des productions écrites

5.5.4.1. Grille d'analyse des productions (d'après la grille du DCL)

J'ai conçu les trois productions écrites du projet principal, *Projekt Prospekt*, à savoir les "Informations pratiques", la "Visite guidée" et le "Commentaire dans un livre d'or électronique", en me basant sur les critères du *DCL* (Diplôme de Compétence en Langue à usage professionnel, 2002). Je n'utiliserai pas ces critères à des fins d'évaluation. Par contre, ils ouvrent des pistes pour l'analyse des productions des apprenants.

Capacités démontrées	**PREMIER DEGRE (A2)**	**DEUXIEME DEGRE (B1-1)**	**TROISIEME DEGRE (B1-2)**
Caractéristiques générales	Repérage de quelques éléments. Transmission intelligible dans des situations de communication parfaitement prévisibles.	Repérage et transmission d'un ensemble d'éléments d'informations en relation avec des situations de communication prévisibles.	Sélection et classement d'éléments d'informations en adéquation avec la tâche délimitée. Autonomie partielle dans des situations de communication qui restent prévisibles.
Compréhension de l'écrit et de l'oral	Repérage de quelques éléments factuels simples	Compréhension de la plupart des éléments explicites marquants.	Recueil d'informations multiples provenant de sources diversifiées.
Production écrite et orale	Reproduction intelligible des éléments repérés.	Restitution des éléments compris dans un format simple.	Réorganisation et réutilisation personnelles des éléments retenus.

Tableau 23 – Extrait de la grille utilisée pour le *DCL*.

5.5.4.2. Grille pour l'identification de traces d'actions vis-à-vis des corpus et concordances (d'après Kennedy et Miceli, 2001)

Quel emploi les apprenants font-ils des corpus et concordances ? Cet emploi, permet-il de trouver des indices d'une attitude d'"apprenant-chercheur" ? Afin de creuser ces questions, il convient d'observer les actions des apprenants sur les concordances et, le cas échéant, des corpus. Je distingue quatre niveaux d'actions.

a) L'apprenant a identifié un segment pertinent dans la concordance. Il l'a l'inséré dans son texte, et il a effectué les modifications nécessitées par le nouveau contexte.
b) L'apprenant a identifié un segment pertinent, mais il n'a pas effectué les modifications nécessitées par le nouveau contexte.
c) L'apprenant a identifié un segment non pertinent dans la concordance.
d) L'apprenant n'a pas identifié de segment dans la concordance.

Je relève le nombre d'actions de type a) à d) par rapport au nombre total de textes du corpus d'apprenants. J'interprète les données en m'inspirant des deux critères de recherche non pertinente suivants, présentés par Kennedy et Miceli (2001 ; cf. 4.2.4.).

e) L'apprenant n'a pas trouvé de segment pertinent parce qu'il a fait une lecture partielle des concordances (par exemple, il a lu de gauche à droite d'une ligne de concordances en commençant par le mot clé, mais il n'a pas lu à gauche du mot clé).
f) L'apprenant a davantage observé certaines caractéristiques du mot clé que d'autres (par exemple, il s'est arrêté aux caractéristiques grammaticales et lexicales et il a ignoré les caractéristiques discursives, ou vice versa.

5.5.5. Évaluation des productions écrites

Le contexte "institution scolaire" motive mon choix de non seulement analyser, mais aussi évaluer, dans une certaine mesure, ce que les apprenants ont produit tout au long de *Projekt Prospekt*. Selon Tardieu (2006), Perrenoud (2004), Schlemminger (2003) et Springer (1999), un certain consensus semble exister concernant les divers types d'évaluation que l'on peut prendre en compte en fonction des types d'activités que l'on propose aux apprenants. Les institutions scolaires distinguent entre des évaluations de type formative, sommative et certificative. L'évaluation formative n'est pas notée. Elle sert à informer les apprenants sur le chemin qu'ils ont déjà parcouru et sur ce qu'il convient de faire pour progresser encore.

L'évaluation sommative sert à faire le point sur ce qu'un apprenant a retenu suite à une séquence de travail. Elle est en général notée. L'évaluation certificative sert à classer les apprenants sur une échelle afin de déterminer s'ils peuvent ou non obtenir une qualification officielle (DNB, baccalauréat, etc.). La premier et le troisième premier type d'évaluation ont joué un certain rôle dans mon scénario pédagogique.

5.5.5.1. Grilles d'évaluation pour des tâches complexes

Ces dernières années, les tâches complexes ont fait massivement leur entrée dans l'enseignement/apprentissage (Laurier 2005, chapitre 6). Comment réguler ce type d'activités, et comment évaluer les performances des apprenants au terme d'une telle tâche ? Les conditions d'exécution de la tâche peuvent déterminer son degré de difficulté : des contraintes de temps, des aides diverses, la difficulté plus ou moins grande au niveau technique peuvent être des facteurs qui allègent ou alourdissent une tâche. Il peut être, par exemple, admis de donner aux apprenants l'occasion de s'appuyer sur des exemples de tâches si l'évaluation porte principalement sur une technique particulière (Laurier 2005 : 112). Ceci justifie, à mon avis, la proposition d'activités de repérage et de réemploi de segments lors de la première phase de *Projekt Prospekt*. Avant l'exécution de la tâche "Informations Pratiques", les apprenants ont l'occasion de se familiariser avec la technique de sélection et de récupération d'informations. Ils ont la possibilité d'apprendre à réinsérer les données récoltées dans un nouveau (co-)texte. La créativité n'est pas un critère décisif lors de cette tâche.

Laurier présente un type d'évaluation certificative qui a été adopté en 2003 par le Ministère de l'Éducation du Québec pour l'épreuve officielle de français langue d'enseignement, à la fin de l'école primaire. La tâche, la production d'un texte argumentatif, est évaluée à l'aide d'une grille à deux volets. Le premier est plutôt qualitatif. Il concerne le niveau discursif ("cohérence de l'argumentation", "organisation stratégique", "continuité et progression"), le deuxième est plutôt quantatif. Il évalue le "respect du code linguistique" ("utilisation des mots", "construction des phrases", "ponctuation", "orthographe"). Les extrémités de l'échelle à cinq niveaux sont formulées comme suit : "compétence insuffisante" et "compétence marquée" (Laurier 2005 : 127). Une telle répartition en deux volets semble convenir également pour *Projekt Prospekt*. Selon le type de la tâche, le volet discursif ou le volet linguistique auront plus ou moins de poids. Une des tâches sera soumise à une évaluation certificative. Il s'agit de la dernière lors de laquelle les apprenants auront à rédiger en temps

limité une production écrite qui rappellera la tâche "Commentaire dans le livre d'or d'un musée" (cf. 5.3.4. et 5.5.5.3.2.).

5.5.5.2. Grilles d'évaluation pour *Projekt Prospekt*

5.5.5.2.1. The Five requirements (Siepmann)

La grille de Siepmann (2003), *The Five Requirements*, propose des critères qui permettent de prendre en compte aussi bien la précision que les éléments discursifs. Siepmann présente cinq domaines qu'il convient d'évaluer dans une production écrite de lycéens. Le premier domaine concerne le contenu qui doit être "nouveau" et "intéressant". Le deuxième évalue la précision grammaticale (prépositions, temps, *patterns* grammaticaux, collocations). Le troisième domaine concerne le style et le lexique qui convient à un style donné. Le quatrième domaine est celui de l'unité thématique. Le cinquième domaine évalue la cohérence du texte, définie ici comme enchaînement logique des énoncés.

1. The requirement of content: Every sentence should offer the reader new and interesting information. *2. The requirement of correctness: Every sentence must be constructed by the grammatical rules (facts: prepositions, tenses; patterns: verb, noun and adjective patterns; choices: tense use, etc.) and the lexical rules (collocation) of the language.* *3. The requirement of style: Every sentence should contain appropriate functional and content vocabulary.* *4. The requirement of unity: Every sentence should be related to the topic sentence of the paragraph.* *5. The requirement of coherence: Every sentence must fit in logically with the sentences surrounding it.*

Tableau 24 – *The Five Requirements* (Siepmann).

Je me suis servie de cette grille pour évaluer la production de départ et la production finale de *Projekt Prospekt* (cf. 5.5.5.3.2.).

5.5.5.2.2. Niveaux de compétence déterminés par le CECR

Selon les programmes officiels de l'Éducation Nationale, les apprenants qui sont en classe de troisième du collège doivent être entraînés en vue de l'obtention du niveau B1, défini dans le *CECR* (*Cadre européen commun de référence pour les langues*, 2001). La conception de *Projekt Prospekt* rappelle les compétences évoquées pour le troisième degré du DCL

(Diplôme de Compétence en Langue) : l'action de repérage et la restitution de données pertinentes deviennent plus complexe que pour les deux premiers degrés ; il s'agit de recueillir des "informations multiples provenant de sources diversifiées", le traitement des données récoltées devient également plus exigeant, comme le rappelle la formulation "Réorganisation et réutilisation personnelle des éléments retenus". Le troisième degré du *DCL* équivaut au niveau B1-2. À la lumière des critères du *DCL*, *Projekt Prospekt* paraît ambitieux. Je reformule les critères pertinents du DCL comme suit.

A1 :
- Utilise entre un et trois documents source ;
- identifie des segments pertinents ;
- présente les segments identifiés dans un ordre qui donne du sens ;
- effectue de légères modifications (horaires, adresse, etc.).

A2 :
- Utilise un corpus de documents ;
- rédige librement quelques segments et/ou phrases ;
- adapte les segments qu'il a repérés dans le corpus à son propre texte
(rajoute des conjonctions, des expressions anaphoriques, modifie les compléments, des terminaisons, etc.) ;
- révise sa production à l'aide de lignes de concordance sélectionnées par l'enseignante.

B1 :
en plus des critères pour A2 :
- complète les parties copiées par des paragraphes rédigés librement ;
- révise sa production à l'aide de lignes de concordance qu'il a sélectionnées lui-même.

Tableau 25 – Critères pour l'évaluation des tâches de *Projekt Prospekt*.

5.5.5.3. Productions évaluées lors de *Projekt Prospekt*

J'ai choisi d'ouvrir et de clore le projet par une production de type "devoir sur table" afin d'obtenir quelques indices concernant l'évolution des caractéristiques de rédaction des apprenants (style parataxique, nativisation, etc.) au fil de *Projekt Prospekt*.

5.5.5.3.1. Production de départ

En février, dans le contexte des épreuves du "Brevet blanc"[20], j'ai fait faire au groupe qui a participé à la deuxième phase de la recherche-action (2007/08) une rédaction en temps limité (une heure trente), non notée, destinée à déterminer d'où "partent" les apprenants. Ils n'ont eu accès à aucun outil d'aide à la rédaction. Suite à une phase de travail préparatoire, j'ai invité les apprenants à faire le portrait de trois personnes et à rédiger une anecdote ayant pour titre "Une histoire de famille". La conception de la tâche a donné aux apprenants l'occasion de produire un texte descriptif et un texte narratif, donc d'utiliser le présent et le passé. Le degré de complexité au niveau du groupe nominal et du groupe verbal n'était pas le même dans la partie descriptive que dans la partie narrative

5.5.5.3.2. Production finale

Lors de l'avant-dernière séance de *Projekt Prospekt*, les apprenants ont rédigé un texte équivalant à la troisième production écrite du projet, le "Commentaire dans un livre d'or éléctronique". Ils ont eu accès à un texte du corpus "Livre d'or" et à 87 lignes de concordance, faites à partir de ce même corpus. Le temps était limité à 55 minutes, et le travail a été noté à l'aide de la grille de Siepmann (2003 ; cf. 5.5.5.2.1.). Cette production permet d'observer la façon dont les apprenants font usage d'un texte de corpus et de lignes de concordance s'ils ne disposent pas d'autres outils d'aide à la rédaction.

Depuis l'année scolaire 2007/08, l'obtention du niveau A2 en une des langues étrangères apprises au collège doit être certifiée en classe de troisième. Le DNB (Diplôme National du Brevet) ne peut être accordé sans que ce niveau ne soit attesté. C'est pour cela que la production finale a fait l'objet d'une certification du niveau A2 dans le domaine de la production écrite.

5.6. Pour résumer

Dans le chapitre 5, j'ai esquissé les méthodologies de recherche, le scénario et les activités proposées aux apprenants, les outils d'aide à la rédaction et les outils de communication, ainsi que les procédés d'analyse de données. Le chapitre 6 sera consacré à l'analyse des données.

[20] Période d'épreuves destinée à préparer les apprenants au passage de l'examen DNB (Diplôme National du Brevet) qui marque la fin de la scolarisation au collège.

6. "Plongés dans le réel" – apprenants et corpus

"Je suis satisfait du projet, car je me sentais plongé dans le réel". C'est en ces termes que Fernand, un des apprenants ayant participé à *Projekt Prospekt*, s'exprime dans le questionnaire distribué à la fin du projet. Il n'est pas le seul à utiliser la métaphore du milieu aquatique pour caractériser les expériences que l'on peut faire en abordant un *input* en langue étrangère. Chez Robinson (2002 : 7), nous trouvons le terme de *input flooding*, et Zanettin (2001) intitule un de ses articles ainsi : "*Swimming in Words: Corpora, Translation, and Language Learning*". En travaillant avec des corpus, on prendrait alors un "bain de mots". La métaphore des "flots de mots" n'a pas uniquement des connotations positives. Certes, une de ses facettes illustre le plaisir que l'on peut éprouver en ayant libre accès à des matériaux linguistiques, dits "authentiques", non présélectionnés par un enseignant ou un concepteur de tâches d'apprentissage. Selon Bernardini (2004 : 16), travailler avec des données de corpus peut mener à des découvertes de type sérendipité, et Johns (1988) s'attend à ce qu'un tel travail aide à apprendre la langue (cf. 2.6.2.). D'un autre côté, la métaphore de l'eau suggère qu'il suffirait d'être exposé à de grandes quantités d'*input* pour apprendre. Deux facteurs peuvent, en effet, favoriser un apprentissage en "immersion" : avoir beaucoup de temps et avoir réellement besoin d'utiliser la langue. Dans le cas de l'apprentissage d'une LE au sein d'une institution, le temps est limité ; de plus, les apprenants ne voient pas toujours l'intérêt de s'exprimer en une langue étrangère qu'ils n'ont d'ailleurs pas forcément choisie d'apprendre (cf. 2.2.1.). Appréhender un nouvel environnement linguistique complexe, tel que les corpus, ne peut pas se faire sans "balisage". Les concordances peuvent être considérées comme des "bouées", destinées à éviter aux apprenants la "noyade" quand ils vont à la rencontre des corpus. Dans un premier temps, la présentation matérielle du texte, obtenue à l'aide de concordanciers, peut être ressentie comme déroutante ; une concordance n'est pas d'emblée une aide sans guidage adapté. L'idéal serait d'accompagner les apprenants sans les empêcher d'explorer le milieu "inconnu". Comment les apprenants abordent-ils les corpus et concordances ? Qu'est-ce qu'ils en font ? Quels types d'aide sont bénéfiques lors de l'exécution d'un scénario de production écrite, basé sur des corpus ? Les attitudes des apprenants face à l'*input* en LE, changent-elles au fil du projet ? L'observation des données de ma recherche vise, entre autres, à répondre à ces questions.

Dans ce qui suit, je présenterai des résultats obtenus principalement lors de la deuxième phase de la recherche. Parfois, les résultats des projets de la première et de la troisième phase complèteront les données du projet principal *Projekt Prospekt*. Lors de ce projet, j'ai travaillé

avec 17 apprenants, ayant entre 14 et 15 ans, et avec deux enseignantes. Par souci d'anonymat, les prénoms des apprenants et les noms des enseignantes ont été modifiés.

6.1. Thèmes dans le discours des acteurs de la recherche

Dans les chapitres précédents, j'ai établi un lien entre l'apprentissage des langues à l'aide de corpus et les thèmes suivants : repérage d'informations, révision, normes linguistiques et discursives, intertextualité, altérité, décentrage, attitudes par rapport à l'apprentissage et guidage. Afin de croiser ces thèmes avec ceux qui ont joué un rôle significatif lors de la deuxième phase, phase principale de la recherche, j'ai analysé le script de la séance de présentation du projet, les entretiens faits en mai 2008, les questionnaires distribués aux apprenants pendant et après l'exécution du projet et mon journal de bord. J'ai utilisé *AntConc3.2.1w* (Anthony, 2007) pour dresser une liste provisoire de thèmes. Ce premier balayage des données et l'analyse des listes des mots les plus fréquents dans les entretiens et mon journal de bord ont fait apparaître les thèmes cités ci-dessus dans des proportions variées. D'autres thèmes, comme la gestion du temps, le rôle de la traduction et du travail collaboratif ont émergé lors de l'analyse plus approfondie des entretiens, des scripts de séances et de mon journal de bord.

Après le paragraphe dédié aux analyses de fréquence de mots dans le discours des personnes qui ont participé à la recherche, je vais présenter sept des thèmes identifiés de façon détaillée, par le biais d'études synthétiques et/ou d'études de cas.

6.1.1. Identification de thèmes récurrents dans les entretiens et dans mon journal de bord

À l'aide de la fonction "*Word List*" [liste de mots classés selon leur fréquence] d'*AntConc3.2.1w* (Anthony, 2007), j'ai comparé la fréquence des mots qui sont utilisés dans le script des entretiens et dans mon journal de bord, en me restreignant aux lemmes des noms et verbes qui y apparaissent dix fois ou plus souvent. J'y ai ajouté le lemme "rep*" afin de pouvoir identifier les formes du verbe "repérer" qui est constitutif pour mes recherches. Les deux documents contiennent de nombreuses occurrences des termes de "texte(s)" et "travail". Aussi bien les entretiens que le journal de bord font ressortir un emploi plus fréquent du terme de "mots" que celui du lemme "formul*". Les termes de "temps", "dictionnaire" et "phrase(s)" sont mieux représentés dans le script des entretiens que dans le journal de bord, tandis que

j'utilise, plus souvent que les apprenants et l'enquêtrice, les termes d'"élève" et de "projet" et le lemme "formul*". Le terme de "tâche" et le lemme "identif*" apparaissent uniquement dans mon journal de bord. La question du repérage est peu visible dans les deux listes de fréquence. J'y ai trouvé respectivement une et quatre occurrences du lemme "rep* qui permet d'identifier les différentes formes verbales du verbe "repérer" et les formes nominales dérivées de ce verbe. La distribution des termes et lemmes que je viens de décrire m'amène à isoler deux thèmes supplémentaires qui se rajoutent à ceux introduits dans les chapitres précédents, à savoir le rôle de la traduction et la gestion du temps. Dans le tableau suivant, les termes pour lesquels la fréquence varie de façon significative sont imprimés en majuscules. Comme le script des entretiens contient 1,27 fois plus de mots que le journal de bord, j'ai multiplié les chiffres obtenus pour ce dernier par 1,27 afin de rendre les données comparables.

Script des entretiens (15466 mots)	**Nombre d'occurrences**	**Journal de bord** (12121 mots)	**Nombre d'occurrences**
texte(s)	87	ÉLÈVE(S)	74 (x 1,27 = 93,98)
travail*	81	texte(s)	59 (x 1,27 = 74,93)
TEMPS	65	travail*	56 (x 1,27 = 71,12)
concordance(s)	61	PROJET	55 (x 1,27 = 69,85)
trad*	48	concordance(s)	42 (x 1,27 = 53,34)
mot(s)	40	TROUVE*[21]	33 (x 1,27 = 41,91)
aide*	39	corpus	38 (x 1,27 = 48,26)
corpus	37	aide*	38 (x 1,27 = 48,26)
DICTIONNAIRE	21	mot(s)	27 (x 1,27 = 34,29)
phrase(s)	20	FORMUL*	25 (x 1,27 = 31,75)
cherch*	20	trad*	25 (x 1,27 = 31,75)
PROJET	19	TEMPS	22 (x 1,27 = 27,94)
TROUV*	16	TÂCHE(S)	21 (x 1,27 = 26,67)
FORMUL*	11	cherch*	19 (x 1,27 = 24,13)
ÉLÈVE	2	phrases	14 (x 1,27 = 17,78)
rep* (repérer)	1	IDENTIF*	13 (x 1,27 = 16,51)
tâche(s)	0	rep* (repérer)	4 (x 1,27 = 5,08)
IDENTIF*	0	DICTIONNAIRE	2 (x 1,27 = 2,54)

Tableau 26 – Noms et lemmes des verbes fréquents dans les entretiens et journal de bord.

6.1.2. Thèmes retenus pour une analyse approfondie

Suite à l'analyse lexicale des données que je viens de décrire, j'ai regroupé les thèmes récurrents que j'ai identifiés dans les rubriques suivantes : les attitudes des apprenants face à l'*input* en LE, leurs actions vis-à-vis de l'*input*, l'évolution des caractéristiques des textes des

[21] J'ai éliminé les occurrences de "trouv*" qui sont synonymes à "estim*".

apprenants, la littératie électronique, le travail collaboratif, l'interdisciplinarité et le rôle du guidage pour l'exploitation de corpus et concordances.

*6.2. Attitudes des apprenants face à l'*input *en LE*

6.2.1. Que veut dire "avoir une attitude d'apprenant-chercheur" ?

Johns et King (1991 : iii) et d'autres chercheurs qui travaillent sur l'utilisation des corpus et concordances pour l'apprentissage des langues, citent comme un des avantages de ces outils la possibilité offerte aux apprenants de développer une attitude d'apprenant-chercheur (cf. 2.6.3.). Que veut dire "avoir une attitude d'apprenant-chercheur" ? Que peuvent bien chercher les apprenants dans les concordances ? Les traces qui renseignent sur cela sont difficiles à détecter dans le contexte de mes recherches. Il paraît, par contre, possible de trouver des indices qui renseignent sur la façon dont les apprenants perçoivent les corpus et comment ils les utilisent. Dans le sous-chapitre 6.2., je m'intéresse au premier de ces deux volets. Le deuxième sera traité en 6.3.

Souvent, les apprenants cherchent à comprendre un texte dans sa totalité (Rouet *et al.*, 2004). Les lecteurs qui ont pour objectif une lecture exhaustive d'un texte ont tendance à prendre comme point de départ de leur lecture l'endroit qui marque visuellement le début du texte, donc celui qui se situe en haut à gauche pour les langues telles que l'allemand et de français.[22] Dans le contexte d'un projet d'écriture basé sur l'exploitation de concordances, donc sur des textes matériellement réorganisés, les apprenants se mettent-ils à lire en modifiant le point de départ et la direction spatiale de leur lecture ? Prennent-ils en compte les informations fournies par les lignes qui se situent à gauche, à droite, en haut et en bas d'un mot clé ? J'ai retenu quelques situations qui donnent des éléments de réponse aux questions évoquées. Il serait intéressant de mener ultérieurement une étude de type quantitatif afin de compléter mes résultats par des données obtenues à l'aide d'une caméra qui filme les mouvements oculaires.

Dans ce qui suit, je présenterai trois regards possibles des apprenants sur l'*input* présenté sous forme de corpus et concordances : le regard exhaustif, le regard synthétique et le regard analytique. Je citerai des situations observées lors de ma recherche qui attestent l'un ou l'autre

[22]Ici, je ne me réfère pas aux textes qui se caractérisent par une mise en forme matérielle qui fait ressortir certains termes davantage que d'autres ou qui résorganise la linéarité des données du texte, comme par exemple les résultats d'une mise en forme matérielle à l'aide du logiciel *wordle*.

type de regard. Ensuite, à la lumière des trois types de regards, j'étudierai les traces d'une attitude d'apprenant-chercheur dans le discours des apprenants en m'intéressant aux champs lexicaux "rechercher", "*input*" et aux co-textes dans lesquels apparaît le terme "concordance" dans le discours des apprenants.

*6.2.2. Regards des apprenants sur l'*input *en LE*

6.2.2.1. Regard exhaustif

Les apprenants qui posent un regard exhaustif sur l'*input* en LE cherchent à saisir le sens du passage ou texte entier. Dans le cas extrême, ils commencent leur lecture en haut à gauche, et ils terminent en bas à droite. Ne pas comprendre un des mots peut être ressenti comme un obstacle pour eux. La figure suivante illustre le sens d'une lecture exhaustive du début d'un texte du corpus "Informations pratiques".

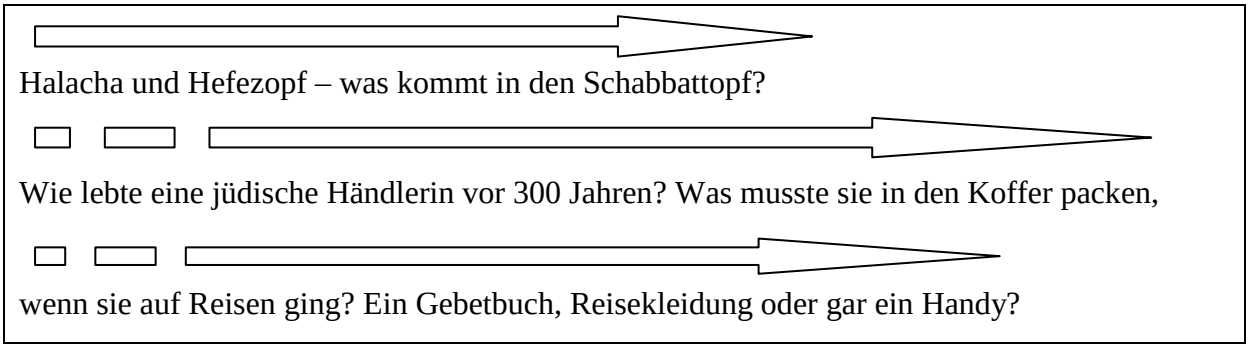


Figure 41 – Sens d'une lecture exhaustive d'un texte de corpus.

Dans les entretiens de Christine, Jean, Léa, Perrine et Véronique, j'ai trouvé des traces qui permettent de parler d'un regard exhaustif sur l'*input* : "tous les mots", "pleins de mots", "on ne comprend pas tout" (cf. 6.2.3., rubrique "tout" et "mots").

6.2.2.2. Regard synthétique

Porter un regard synthétique sur un texte veut dire que l'on voit le lien qui existe entre son sens global et les parties qui le composent. Afin de soutenir ou développer un tel regard, j'ai proposé aux apprenants des activités d'entraînement, comme, par exemple, l'identification et l'observation de sections d'un texte qui caractérise les textes de type "Visite guidée" (cf. 5.3.7.3.1.). L'extrait suivant de mon journal de bord relate la façon dont Vicky a décrit, le 11 avril 2008, donc au bout de six semaines de travail, le changement qu'elle a observé concernant sa façon de lire des textes allemands lors de *Projekt Prospekt*. L'échange verbal entre Vicky et moi a permis de clarifier qu'elle a bien eu l'impression d'avoir appris à porter

un regard synthétique sur les textes des corpus, même si la formule "regarder de façon globale" semble aller dans le sens d'une lecture exhaustive.

> Vicky a dit qu'elle a balayé les textes [du corpus "Visites guidées"] du regard et qu'elle arrive maintenant à les regarder de façon "globale". Je lui ai dit que j'interprétais son regard comme un regard de radiologue, ce qui semblait lui convenir. Je pense qu'elle a fait un pas vers la méthode que j'aimerais faire utiliser : lire pour déconstruire un texte et pour se servir de certaines de ses composantes.

6.2.2.3. Regard analytique

Le regard analytique, ou, dans d'autres termes, la lecture ciblée, peut avoir comme objectif l'identification d'un élément précis qui répond à une question, formulée explicitement ou non. J'ai encouragé les apprenants à procéder à une lecture ciblée des concordances afin de les aider à se familiariser avec le co-texte sémantique ou morpho-syntaxique d'un mot clé sélectionné en fonction de besoins de révision du texte. La figure suivante présente les sens spatial de lecture qui sont possibles quand on prend comme point de départ de la lecture un mot clé dans une concordance. Il s'agit ici du mot clé "*KINDER*" [enfants] dans la concordance faite à partir du corpus "Informations pratiques".

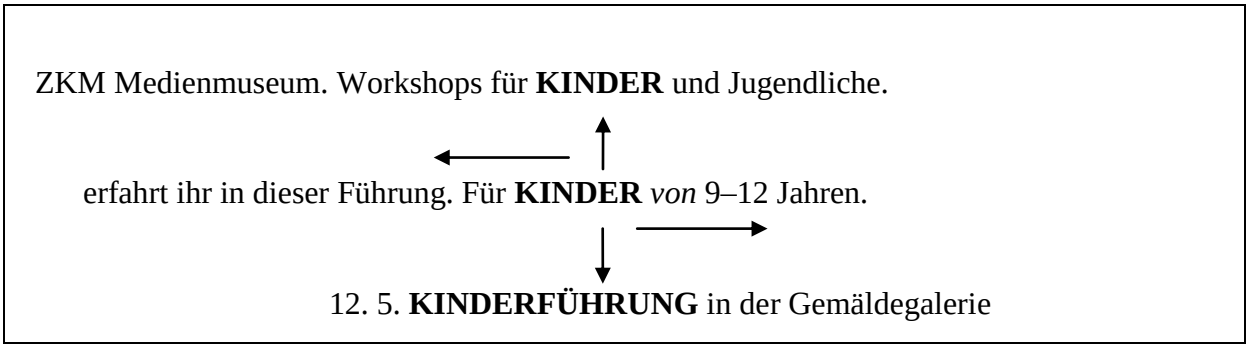


Figure 42 – Sens spatial de la lecture à partir d'un mot clé dans une concordance.

Le 1er avril 2008, je décris dans mon journal de bord deux situations qui illustrent un stade peu avancé de lecture ciblée de concordances.

> Il faut apprendre à lire suffisamment large, pas seulement à droite du mot clé. Par exemple, Géraldine devait revoir le passage "*Reservierung*" en s'inspirant de la ligne suivante.
>
> *Führung nach Anmeldung* ***UND*** *Zooschul-Reservierung bei Horst* Krull [Visite guidée sur inscription et sur réservation du zoo auprès de Horst Krull].
>
> L'idée était de lui faire trouver l'expression "*nach Reservierung*" [sur réservation]. Elle ne l'a trouvée que grâce à un guidage individualisé de ma part. Chloé était censée trouver l'équivalent allemand d'"arrêt de bus". En tapant le mot clé "*Bus*", elle m'a dit n'avoir rien trouvé. Pourtant, le mot

"*Bushaltestelle*" [arrêt de bus] a terminé la série des lignes de concordance comportant ce mot clé.

Pyrmont zu Fuß erreichbar 2 km oder mit dem **BUS** Richtung Stadtzentrum. Bahn bis Gmund per **BUS** weiter bis Bad Wiessee-Abwinkl. bis Nistertal-Bad Marienberg **BUS** zum Wildpark oder ca. 1 km zu Fuß bis Haltestelle Blindham. **BUSHALTESTELLE** direkt am Park [jusqu'à l'arrêt de bus Blindham. **ARRÊT DE BUS** qui donne sur le parc].

Géraldine ne semble pas avoir lu de droite à gauche, et Chloé n'a apparemment pas lu de haut en bas. Cela les a empêchées de repérer un élément pertinent. Je conclus de ces observations que la lecture de lignes de concordance est un nouveau type d'activité qui nécessite la création de nouvelles habitudes concernant le sens de lecture. Cela demande un certain temps d'entraînement collectif et individuel. Les paragraphes suivants recenseront la façon dont les apprenants parlent de l'*input* en LE en général et des concordances en particulier.

*6.2.3. Comment les apprenants désignent-ils l'*input *?*

Dans les entretiens, j'ai détecté des termes qui illustrent la façon dont les apprenants parlent de l'*input* en LE, en particulier des concordances et de leurs composantes. L'utilisation explicite du terme "concordance(s)" fera l'objet d'un paragraphe à part. Je cite également l'enquêtrice une fois, et je comptabilise sa citation. Les termes sont classés selon la complexité décroissante de l'élément désigné, donc en allant du plus macrosyntaxique au plus microsyntaxique (Py, 2002). La première colonne présente le mot clé recherché à l'aide du concordancier. La deuxième colonne contient un choix de citations provenant du script des entretiens. La troisième colonne indique le nombre de fois où ces termes ont été utilisés par les apprenants dans l'entretien pour désigner l'*input* en LE.

Termes	**Citations**	**Fréquence**
tout*	LÉA : On ne comprend pas tout. PERRINE : Il n'y a pas toujours tout ce qu'on cherche, tous les mots. JANINE : Il y avait tout [dans le mini-corpus papier].	5
infos	OLIVIA : Des fois, il y avait trop d'infos à trier. ENQUÊTRICE : C'est-à-dire, vous êtes face à beaucoup d'informations et vous êtes un peu perdus ?	2
reformu-lations	GÉRALDINE : [Dans les concordances,] On a des reformulations complètement allemandes.	1

phrase, phrases	PERRINE : Des fois, ça peut servir quand on ne sait pas comment on formule la phrase… une phrase. FERNAND : Déjà le fait de devoir aller sur l'outil à chaque fois pour…, ensuite trouver.. trouver les phrases qui peuvent m'aider, j'ai trouvé ça un petit peu une perte de temps.	26
expres-sions	CHLOÉ : Sur le dictionnaire, t'as plein de définitions, et on n'est pas du tout sûr que ce soit le bon mot. (…) Et il y a le truc qui nous manque des concordances : il n'y a pas les expressions toutes faites. VÉRONIQUE : Ça revient un peu au même [si on travaille avec des documents trouvés sur Internet] : si on copierait les expressions qui existent sur des sites.	3
construc-tions	FERNAND : Dans le dictionnaire, il n'y a pas les constructions toutes faites et tout ça.	1
bouts	VICKY : En fait, on ne pouvait pas traduire des grands bouts [à l'aide du dictionnaire électronique]. VICKY : C'est bien, les concordances, mais comme c'est des bouts de texte, on ne formule pas tant que ça.	2
mot, mots	CHRISTINE : Il avait, déjà, plein de mots (…) en allemand. PERRINE : il n'y a pas toujours tout ce qu'on cherche, tous les mots. PERRINE : Non, mais, on cherche, et il n'y a pas le mot, des fois, ça sert à rien. VÉRONIQUE : Il n y a pas tous les mots. Je ne vois pas en quoi ça peut m'aider. ALEXA : Quand on cherche par exemple autour d'un mot (…). [mot = mot clé ici]. GÉRALDINE : Et, enfin… on met les mots… on les écrit comme ça, comme on les a, donc comme en français, dans le même ordre, donc, ça veut rien dire. Même s'il y a les mêmes mots, il n'y a pas le même sens.	40

Pour parler de l'*input* présenté sous forme de concordances, les apprenants privilégient les termes de "mot(s)" (40 occurrences) et de "phrase(s)" (26 occurrences). On peut distinguer trois types de regards sur les concordances : certains les perçoivent comme "un tout", d'autres y voient (ou souhaitent y voir) des "mots" et encore d'autres y distinguent "des phrases".

Christine, Léa, Janine, Perrine semblent percevoir les concordances comme un "tout", contenant "pleins de mots". Olivia dit qu'il faut trier "trop d'informations" dans les concordances. Elle exprime ainsi qu'elle trouve l'*input* complexe et qu'elle estime qu'il doit

être structuré à l'aide d'une action de sa part. Olivia remarque toutefois que dans le dictionnaire, " il y a encore plus d'informations à trier" que dans les concordances.

Janine et Perrine formulent (directement ou indirectement) une demande vis-à-vis de l'*input* en LE : il doit contenir "tous les mots" dont elles ont besoin pour rédiger. Selon les expériences dont elles parlent, les concordances ne répondent pas à cette demande.

Chloé, Fernand, Géraldine, Perrine et Vicky disent reconnaître dans les concordances des phrases, de (grands) bouts ou des "formulations" ou "expressions toutes faites". Ils considèrent ces segments comme une aide à la rédaction dont ils souhaitent disposer (Chloé, Vicky et Géraldine) ou qui ne leur convient pas (Fernand et Perrine). Ce groupe d'apprenants parle de la possibilité de lire les concordances de façon ciblée, contrairement à Christine et Léa qui décrivent une lecture de type exhaustive. Les citations de Perrine permettent de la classer aussi bien dans la catégorie de ceux qui décrivent une expérience de lecture globale que de ceux qui parlent d'une lecture ciblée.

6.2.4. Comment les apprenants utilisent-ils le terme de "concordances" ?

Dans ce qui suit, je présente des lignes de concordance, obtenues dans les entretiens pour le mot clé "concordance" à l'aide du concordancier d'*AntConc3.2.1w*. Il ne s'agit pas de dialogues suivis.

ALEXA : Et l'ordre des mots aussi. ENQUÊTRICE : La concordance aide pour l'ordre des mots ?

CAMILLE : Et après, soit on utilise Internet avec les **concordances** quand on a besoin. Ou sinon, le traitement de texte quand on réécrit.

CHLOÉ : ça [être à l'aise avec l'informatique] m'a aidé à me servir des **concordances.** Je trouve que c'est, c'est pas mal comme système

CHLOÉ : C'est plus facile pour les copier-coller, des **concordances** et tout ça parce que ça limite le risque d'erreur.

CHLOÉ : Et il y a le truc qui nous manque des **concordances** : il n'y a pas les expressions toutes faites…

FERNAND : Donc, d'un autre côté, c'est plus pratique, les **concordances** et les corpus.

GEORGES : on travaillait, quoi, avec les **concordances**, ou on faisait le texte.

GEORGES : Oui. De temps en temps, les **concordances,** c'est pas très pratique. On ne comprend pas tout

GÉRALDINE : Et après, on va sur les corpus et **concordances.**

JANINE : Mais non, rien en particulier. Peut-être [expliquer] les **concordances,** parce qu'il y a des trucs où je n'arrive pas trop

OLIVIA : je modifie avec les **concordances** et corpus.

VICKY : sinon, on s'est beaucoup servis des **concordances,** mais ça, c'est après, quand on avait déjà écrit

VICKY : Avec les **concordances,** on est beaucoup plus libre pour écrire vraiment

VICKY : si on devait chercher sur le site ou sur des **concordances,** donc…

VICKY : c'est pas que le côté informatique était indispensable, mais les **concordances,** pour savoir ce qui se dit vraiment

VICKY : Oui. parce que, c'est bien, les **concordances,** mais comme c'est des bouts de texte, on ne formule pas tant que ça.

ENQUÊTRICE : je vois, vous n'avez pas vraiment utilisé les **concordances,** bon.

ENQUÊTRICE : vous traduisiez, et ensuite, vous alliez sur les **concordances** ? OLIVIA : Non, non.

ENQUÊTRICE : Ça t'attire pas, toi, les corpus et **concordances** ? CHRISTINE : Non.

ENQUÊTRICE : Sans utiliser le corpus et concordances, ce qu'a fait Christine.

ENQUÊTRICE : t'es repassée finalement un petit peu par les concordances pour vérifier. CHRISTINE : Oui.

Georges distingue le travail avec concordances du travail d'écriture. Il ne semble pas faire le lien entre les deux. Il dit avoir eu du mal à comprendre l'*input* sous forme de concordances. Janine exprime également son besoin d'explications pour mieux comprendre les corpus et concordances. Christine, Perrine et Fernand se montrent réservés vis-à-vis de ce type de présentation de l'*input*. Fernand admet toutefois que les concordances peuvent être utiles. Alexa, Chloé et Vicky décrivent le potentiel qu'elles attribuent aux concordances : les concordances peuvent, à leur avis, aider au niveau de la syntaxe, des collocations et pour faire moins d'"erreurs", donc au niveau de la précision. Vicky dit faire appel aux concordances après avoir rédigé d'un premier jet. En 6.4., je reviendrai sur le potentiel des concordances évoqué ci-dessus, en m'intéressant particulièrement à l'évolution des apprenants dans les domaines de l'adjectif et de la syntaxe.

6.2.5. Comment les apprenants décrivent-ils la recherche d'informations ?

Dans les scripts de séances et des entretiens, j'ai repéré les termes qui illustrent la présence ou l'émergence d'une attitude de recherche lors de *Projekt Prospekt*. Les verbes suivants font partie du champ lexical "recherche d'informations". Je les ai classés selon le degré décroissant de charge cognitive probablement nécessitée par l'action que le verbe désigne. La deuxième colonne contient des citations provenant du script des entretiens. La troisième colonne indique le nombre de fois que ces termes ont été utilisés par les apprenants dans l'entretien pour désigner une activité liée à la recherche d'informations.

Verbes	**Citations**	**Fréquence**
comprendre	GEORGES : On peut aller sur les ordinateurs, essayer de comprendre.	15, dont 11
	JANINE : C'est mieux avec l'ordinateur. On comprend mieux.	à la forme

	JANINE : Les concordances,] c'est pareil que les corpus…je ne comprends pas.	négative
trier	OLIVIA : [Le dictionnaire papier] est encore plus… il y a encore plus d'informations à trier [que dans les concordances].	2
chercher	OLIVIA : C'est dur de rechercher avec les bons mots clés parce qu'on ne parle pas couramment allemand. Donc, on ne sait pas trop dans quelle phrase ça pourra se trouver. CAMILLE : Je trouvais que c'était par rapport au fait qu'on doit chercher nous-mêmes, on voit nous-mêmes nos erreurs. ENQUÊTRICE : Ce qui vous gêne en fait, c'est d'avoir… d'avoir à chercher, etc., et de pas trouver.	13
repérer	SIMON : En fait, on peut essayer de repérer quelque chose qu'on comprend.	1
lire	SIMON : Ça [la lecture de textes de corpus et de concordances] nous permet de s'habituer, parce que du coup, on lisait beaucoup. On a lu beaucoup de textes.	5
trouver	CAMILLE : Soit on prend dans la concordance, soit on ne trouve pas, on… FERNAND : …on va sur un dictionnaire, il nous le trouve en allemand. OLIVIA : On ne savait pas forcément comment… comment trouver [dans la concordance] ce qui nous intéressait. OLIVIA : On ne sait pas trop dans quelle phrase ça [le bon mot clé] pourra se trouver. Donc, on avait plus de mal à parfois trouver certains mots. Donc, on était obligés d'aller sur le traducteur.	14
savoir	VICKY : En fait, c'est pas que le côté informatique était indispensable, mais les concordances, pour savoir ce qui se dit vraiment.	3
regarder	GÉRALDINE : On pouvait aussi regarder sur le traducteur pour des mots isolés. CAMILLE : Je regarde sur le corpus, je regarde ce qui va et ce que je peux compléter. JANINE : En fait, j'y suis allée une fois [sur la concordance], (…) il y avait plusieurs possibilités. Comme ça, je regardais laquelle me convenait le mieux, et je pouvais la mettre après.	11
voir	CAMILLE : Enfin, [nos phrases ne] sonnaient pas forcément 'plus allemand', mais visuellement, enfin, la façon de former des phrases, parce que comme elles sont déjà faites, avec les verbes, les noms, comment ils sont placés, c'est vrai que c'est plus pratique, ça… quand on voit directement. ENQUÊTRICE : Quand vous avez corrigé vos textes, vous les corrigiez plutôt à la	25, dont 2 à la forme négative

main ou plutôt en version électronique ? GEORGES : Les deux. LÉA : Moi, c'est un peu les deux aussi, à la main et aussi informatique... GEORGES : …pour utiliser les dictionnaires électroniques, les dictionnaires, et puis, pour voir ce voulaient dire quelques mots.

Les citations désignant des actions de recherche d'informations font ressortir que les actions de tri et de repérage sont assez peu représentées : Simon est le seul à utiliser le verbe "repérer", Olivia est la seule à employer le verbe "trier". Ces verbes désignent des actions cognitivement couteuses. Les verbes "regarder" et, en particulier, le verbe "voir" sont bien représentés. Toutefois, quatre sur vingt-six occurrences de "voir" sont liées à l'utilisation du concordancier. Chaque fois, c'est Camille qui parle. Les autres fois, "voir" est mis en relation avec d'autres outils ou lieux : le blogue, la boîte courriel, Internet, le traducteur, le dictionnaire électronique, etc. Le verbe "comprendre" est également bien représenté. Dans onze sur quinze occurrences, le verbe est à la forme négative. Mes observations concernant l'emploi de verbes désignant des actions liées à la recherche d'informations renvoient à la présence d'un obstacle épistémologique important pour un bon nombre des apprenants ayant participé à *Projekt Prospekt*. Les paragraphes suivants s'intéressent aux obstacles liés à l'utilisation des concordances et aux apports perçus par les apprenants.

6.2.6. Obstacles perçus par les apprenants lors de l'utilisation de concordances

6.2.6.1. Trop d'informations

Plusieurs fois, quand j'ai présenté des lignes de concordance à des apprenants, j'ai obtenu des réactions que l'on peut interpréter comme indices de surcharge cognitive. Lors d'une activité d'observation de lignes de concordance pendant la 3ème phase de la recherche-action, Tom, qui s'est par ailleurs montré plutôt enthousiaste vis-à-vis de cette façon de présenter l'*input*, s'est exclamé : "Il y a tellement de mots !" Dans un des entretiens, ce ressenti semble être partagé par Christine et par Olivia.

ENQUÊTRICE : Qu'est-ce qui vous a posé problème pendant que vous travailliez avec les corpus et concordances ? Bon, Christine vient de nous expliquer… (rit) que ce n'était pas forcément facile à utiliser.
CHRISTINE : Ce n'était pas simple. Il y avait plein, déjà, des mots en anglais.
OLIVIA : En allemand.
OLIVIA, GÉRALDINE, CHRISTINE : (rient)

ENQUÊTRICE : Plein de quoi ?
CHRISTINE : Des mots en allemand. Il fallait taper le mot clé et tout, et c'est en allemand.
ENQUÊTRICE : Il faut passer obligatoirement par l'allemand pour le mot clé, c'est ça ?
OLIVIA : Oui. Des fois, il y avait trop d'infos à trier. En plus, on ne savait pas forcément comment… comment trouver… ce qui nous intéressait, parce que c'était très dur de faire le tri parce qu'on avait… si on tapait un mot clé, il y aura pas forcé-, il y aura plein de phrases différentes, et ce sera dur de choisir la bonne qui convient à… dans nos trucs.

Christine relève comme difficulté la quantité importante de mots à laquelle elle est confrontée dans l'*input* en langue étrangère. Olivia précise cette difficulté en disant qu'il y a parfois "trop d'infos à trier". Elle explique que les apprenants ne savent pas forcément comment trouver ce qui les intéresse, parce qu'il y a "plein de phrases différentes" et que c'est "dur de choisir la bonne qui convient" pour réviser leur production écrite. Olivia résume ainsi l'intensité de l'activité cognitive que demande l'emploi des concordances.

6.2.6.2. Manque de certains types d'informations

Les apprenants ont signalé dans les entretiens que les concordances ne comportaient pas toujours ce qu'ils y cherchaient.

ENQUÊTRICE : Et vous, vous avez plutôt choisi de ne pas l'utiliser [le concordancier]. Pourquoi ?
PERRINE : En fait, il n'y a pas toujours tout ce qu'on cherche, tous les mots. Donc, des fois, ça ne servait un peu à rien, quoi.
ENQUÊTRICE : T'avais l'impression de perdre ton temps ?
PERRINE : Non, mais, on cherche, et il n'y a pas le mot, des fois, ça sert à rien. Mais des fois, ça peut servir quand on ne sait pas comment on formule la phrase… une phrase.
ENQUÊTRICE : D'accord.
VÉRONIQUE : Oui, moi, ça ne me sert pas trop, un peu.
ENQUÊTRICE : Pourquoi ?
VÉRONIQUE : Il n'y a pas tous les mots. Je ne vois pas en quoi ça peut m'aider.
ENQUÊTRICE : Tu préfères utiliser le dictionnaire ? Non, le traducteur, toi, c'était le traducteur.
VÉRONIQUE : Oui, parce qu'on rentre directement… même si des fois, ça ne correspond pas vraiment à la façon de parler.

ENQUÊTRICE : C'est plus… c'est plus rapide en tout cas ?
VÉRONIQUE : Oui.
ENQUÊTRICE : C'est ça. Ce qui vous gêne en fait, c'est d'avoir… d'avoir à chercher, etc., et de pas trouver… le fait de perdre un peu de temps, ça vous agace, quoi.
VÉRONIQUE : Mhm.
ENQUÊTRICE : Non, mais vous pouvez dire oui ou non ou pas du tout, c'est moi qui interprète vos propos (rit).
PERRINE : Mais non. Quand on cherche et qu'on ne trouve pas, vu qu'il n'y a pas forcément tous les mots, quoi.

Perrine dit qu'elle n'y a pas trouvé "tous les mots", contrairement à ce qui se passe quand elle utilise le traducteur, comme elle le précise auparavant. Par contre, elle dit que les concordances l'aident pour trouver "des phrases". Véronique avance le gain de temps comme argument en faveur du traducteur, ce qui est en même temps un argument en défaveur de la concordance. Fernand précise également que cela lui prend trop de temps de consulter les concordances.

FERNAND : Déjà le fait de devoir aller sur l'outil à chaque fois pour, ensuite, trouver les phrases qui peuvent m'aider, j'ai trouvé ça un petit peu une perte de temps.

Perrine et Véronique disent qu'il n'y a pas "tous les mots" dans les concordances. En effet, mes corpus sont très petits. Des corpus comme *deWac* ou *DeReKo* auraient peut-être permis de trouver l'un ou l'autre mot que mes corpus ne comportent pas. Toutefois, j'ai souvent proposé aux apprenants des mots clés, testés auparavant pour leur pertinence en vue d'un passage précis à réviser. L'absence de certains mots ne devrait alors pas être l'obstacle principal. La difficulté réside peut-être plutôt en leur identification et/ou compréhension.

6.2.6.3. Mise en forme matérielle comme obstacle

Dans l'extrait du script d'entretien ci-dessous, Camille décrit la présentation du texte du corpus sous forme de lignes de concordance, comportant toutes un mot clé qui apparaît selon l'ordre alphabétique dans le fichier fait à l'aide de *Text-Concordancer (v2)* (Cobb, nd).

ENQUÊTRICE : Est-ce que vous avez eu parfois l'impression que les documents que vous aviez, des concordances ou des corpus, étaient compliquées ou difficiles à comprendre ?

CAMILLE : Au début, c'était – je ne sais plus si c'était concordance ou corpus – il y en a un où c'est par ordre alphabétique, c'est vrai qu'il fallait chercher des mots, et puis il y en a qui sont rangé par… c'était avec des majuscules… Mais maintenant, ça va.

Perrine, Véronique et Fernand semblent ressentir l'effort cognitif nécessité par la consultation de concordances comme un obstacle épistémologique trop important. Une présélection et/ou un "nettoyage" de lignes de concordance auraient certainement pu les aider à aborder ce type d'*input* avec davantage d'espoir d'y trouver ce qu'ils recherchent. Le 6 juin 2008, dans mon journal de bord, je décris la façon dont Vicky semble avoir profité de lignes présélectionnées et nettoyées par moi pour la production finale, un devoir sur table.

> Vicky dit avoir particulièrement apprécié les lignes de concordance "nettoyées" que j'ai proposées pour le devoir sur table du 2 juin et dans lesquelles j'ai enlevé les bouts de phrase que j'ai estimés inutiles. Voici un extrait des lignes d'origine et, en-dessous, les lignes "nettoyées".

*nen ist eine sehr lebendige und anschauliche **AUSSTELLUNG** gelungen in der es neben wissenswerten*
*ist eine sehr lebendige und anschauliche **AUSSTELLUNG*** [est une exposition très vivante et parlante]

*arbeiter sagt Familie Schuh das gabreta-team **BEDANKT** sich recht herzlich für ihr lob und freut*
*das gabreta-team **BEDANKT** sich recht herzlich für ihr lob* [l'équipe gabreta vous remercie très chaleureusement de votre compliment]

*d heute hier im netz... hat mich beides sehr **BEEINDRUCKT** sind eigentlich nur eigene veranstaltu*
*... hat mich beides sehr **BEEINDRUCKT*** [… m'ont impressionné tous les deux]

> Vicky m'a dit que ce "nettoyage" l'a beaucoup aidée pour exprimer ce qu'elle "pense vraiment" ; elle trouve que ce le devoir final est mieux réussi que la tâche de même nature qu'elle a faite auparavant.

Comme Perrine et Véronique ont été absentes lors de ce devoir, je ne puis savoir si le nettoyage de lignes de concordance leur aurait été bénéfique. Fernand en a tiré modérément profit lors de la production finale. Il a toutefois intégré dans son texte certains passages de la concordance qui ne sont pas cohérents par rapport à ce qu'il dit ailleurs dans le texte. Dans l'exemple suivant, une partie de la version V1 du "Commentaire dans un livre d'or", le passage souligné est co-présent dans la production de Fernand et dans une ligne de concordance.

> *Hallo!*
> *Ich heiße Fernand Lelièvre. Im **Oktober** habe ich das Museum besucht. (...)*
> **<u>**Einem** besuch haben wir *schön eingeplant</u> und ich habe endecken ein super technische in die PS3.*

> [Salut ! Je m'appelle Fernand Lelièvre. J'ai visité le musée en octobre. (…) Nous avons l'intention de visiter (le musée) et j'ai découvrir une super technique dans la PS3.]

Fernand ne semble pas avoir entièrement compris le passage co-présent.

6.2.7. Apports des concordances perçus par les apprenants

6.2.7.1. Voir des formulations toutes faites

ENQUÊTRICE : Préféreriez-vous écrire un texte allemand d'une autre façon, c'est-à-dire avec d'autres outils, plutôt que d'utiliser les concordances et le corpus ? Est-ce que vous auriez préféré utiliser autre chose, par exemple un dictionnaire papier classique ?
CAMILLE : Je pense qu'on aurait eu plus de difficultés, peut-être, enfin… C'est sûr qu'au début, c'était lent parce qu'il fallait apprendre à utiliser tous les outils, mais à la fin, je pense que si on avait eu le dictionnaire et à faire nous-mêmes nos phrases, cela aurait été plus compliqué que de voir des formulations toutes faites.
FERNAND : Oui, voilà. Dans le dictionnaire, il n'y a pas les constructions toutes faites et tout ça.

Camille décrit l'apprentissage de l'utilisation des concordances comme une condition pour pouvoir profiter de l'aide qu'elles peuvent fournir. Cette aide consiste en une visualisation de "formulations toutes faites". En effet, Camille est une des premières à avoir pleinement exploité ce potentiel des concordances. Sa première production pour *Projekt Prospekt* illustre cela (cf. annexe II.1.). Fernand rebondit sur ce que vient de dire Camille. Toutefois, pendant *Projekt Prospekt*, il s'appuie rarement sur des lignes de concordance pour revoir ses textes.

6.2.7.2. Voir soi-même ses erreurs

Comme dans le paragraphe précédent, Camille et Fernand mettent en avant que l'utilisation de concordances donne la possibilité d'apprendre à corriger soi-même son texte. Fernand précise de nouveau qu'il peut être avantageux de se pencher sur les concordances, tout en montrant qu'il ne se sent pas prêt à le faire.

ENQUÊTRICE : Quels avantages voyez-vous à travailler avec des corpus et concordances électroniques pour écrire un texte en allemand ? Est-ce que vous avez vu des avantages, des points positifs à utiliser les concordances et les corpus ?

CAMILLE : Ben, moi, en fait, je trouvais que c'était… par rapport au fait qu'on doit chercher nous-mêmes, on voit nous-mêmes nos erreurs. Enfin, pour moi, ça marche comme ça quand je trouve ce que j'ai écrit d'abord une première fois. Après, je regarde sur le corpus, je regarde ce qui va et ce que je peux compléter, donc, je vois moi-même mes erreurs.

ENQUÊTRICE : Tu peux t'autocorriger grâce au corpus ? Tu y arrives ?

CAMILLE : Oui.

ENQUÊTRICE : Oui.

FERNAND : Moi, j'ai plus eu un petit peu de mal à me prendre en main au début, donc…

ENQUÊTRICE : T'a mis du temps à t'habituer.

FERNAND : Oui, voilà, c'est ça.

ENQUÊTRICE : Oui. Et après, tu as vu des points positifs, parce que c'était plutôt une difficulté ?

FERNAND : Oui, c'est ça. Mais… c'est plutôt pratique, dans les textes et tout, pour voir ses erreurs, mais… moi, sans plus, quoi.

Selon les paroles de Camille et de Fernand, l'utilisation de concordances pourrait avoir comme conséquence que les apprenants cessent d'attendre de l'enseignant qu'il leur dise ce qu'il aurait "fallu" écrire. Abandonner cette attente de la part de l'enseignant ne va pas de soi, comme l'illustre l'évolution des quatre premières versions des "Informations pratiques" de Géraldine. Afin de l'aider à trouver une expression qui correspondrait à "ouvert tous les jours", j'ai trois fois répété l'annotation suivante :

> "*Sommer*" [été] : chercher autour de ce mot clé une expression qui ressemble à "*Tag*" [jour].

Géraldine a réussi à trouver le segment "*jeden Tag*" [tous les jours] et le lexème "*täglich*" [quotidiennement]. Les deux peuvent convenir, à condition d'être complétés par "*geöffnet*" [ouvert]. Au lieu de rajouter ce participe passé, dans sa dernière version, Géraldine a intégré le mot clé "*Sommer*" que je lui ai indiqué afin de la guider dans sa recherche de segments pertinents. Elle semble avoir confondu aide et solution.

V1 : *Juli, August: *alle Tagen.* [Juillet, août : tous les jours.]

V2 : *Juli, August: *alle Tagen.* [Juillet, août : tous les jours.]

V3 : *Juli, August: jeden Tag.* [Juillet, août : tous les jours.]

V4 : *Juli, August: Sommer täglich* [Juillet, août : été tous les jours]

6.3. Actions des apprenants vis-à-vis des corpus et concordances

En 5.5.4.2., j'ai présenté quatre niveaux d'actions que les apprenants sont susceptibles d'entreprendre quand on leur propose de réviser leur texte à l'aide de concordances. Les niveaux s'échelonnent entre identification et intégration de lexèmes ou segments dans le texte de l'apprenant et une situation lors de laquelle il n'y a pas de traces d'identification d'éléments dans la concordance. Dans ce qui suit, je présenterai les traces des actions des apprenants vis-à-vis de textes de corpus et vis-à-vis de lignes de concordance lors de la rédaction de la production finale.

6.3.1. Utiliser le corpus comme source pour une écriture intertextuelle

La production finale, de type "Commentaire dans un livre d'or électronique", se prête particulièrement bien à l'observation des actions entreprises par les apprenants par rapport à des lignes de concordance lors d'une activité de rédaction de texte. Je considère ces actions comme éléments d'une d'écriture intertextuelle (cf. 3.4.2.3.2.). La production finale a été rédigée dans des conditions de devoir sur table (cf. 5.5.5.3.2.). J'ai connaissance des documents d'aide à la rédaction auxquels les apprenants ont eu accès pour écrire leur texte. Il s'agit d'un texte du corpus "Livre d'or" et de 87 lignes de concordance, sélectionnées dans ce même corpus et nettoyées par moi afin de faciliter leur utilisation par les apprenants en situation de temps limité (55 minutes). En plus de cela, les apprenants avaient le droit de se servir des mini-corpus papier "Informations pratiques" et "Visite guidée" que je leur ai distribués pendant le projet (contenant respectivement 4672 et 2996 mots). Christine, Corentin et Fernand n'avaient apporté aucun des corpus papier. Alexa n'avait qu'une page des "Informations pratiques" sur elle, et Chloé a disposé du mini-corpus des "Informations pratiques". Georges, Géraldine, Jean, Léa et Vicky avaient apporté le mini-corpus des "Visites guidées". Camille, Janine, Olivia et Simon avaient les deux mini-corpus sur eux. Les apprenants pouvaient se servir de deux grands tableaux de déclinaison affichés dans la salle. Louise, Perrine et Véronique ont été absentes.

Dans ce qui suit, j'observe les lexèmes, segments et passages longs qui sont co-présents dans les productions finales des apprenants et dans les corpus et/ou concordances auxquels ils avaient accès. Je relève le nombre de passages co-présents non modifiés et le nombre de passages co-présents modifiés conformes ou non conformes. Je distingue ici entre "segment" et "passage long co-présent" par leur longueur. Les premiers comportent deux à trois lexèmes.

Les deuxièmes correspondent à un espace qui est délimité par deux signes de ponctuation de type "point", "point d'exclamation" ou point d'interrogation. Je ne différencie pas entre les passages contenant un groupe verbal et les passages sans groupe verbal. Je ne distingue pas non plus les phrases simples, ne contenant qu'une phrase principale, des phrases complexes qui se composent, par exemple, d'une phrase principale et d'une subordonnée. Je définis ici le terme "conforme" dans le sens "qui correspond aux normes linguistiques et discursives, fixées par le genre textuel 'Commentaire dans un livre d'or'". En cas de double emploi d'un passage co-présent, je ne le compte qu'une seule fois.

Prénom	**Lexèmes co-présents**	**Segments co-présents**	**Passages longs co-présents**	**Nombre de passages co-présents non modifiés**	**Nombre de passages co-présents modifiés conformes**	**Nombre de passages co-présents non conformes**	**Nombre de passages co-présents (total)**	**Nombre de mots co-présents**	**Nombre de mots du texte**
Alexa	4	4	0	4	1	3	8	10	115
Camille	2	6	6	8	4	0	12	45	92
Chloé	6	8	1	14	2	1	15	29	103
Christine	4	4	5	7	3	3	12	45	90
Corentin	6	1	0	6	1	0	7	9	79
Fernand	1	4	4	6	1	2	8	33	119
Georges	5	2	2	9	0	0	9	27	84
Géraldine	3	4	4	6	3	3	11	28	119
Janine	3	1	4	8	1	2	9	24	100
Jean	2[23]	5[24]	0	3	1	3	7	16	89
Léa	2	4	2	5	6	0	6	22	84
Olivia	3	7	2	8	2	2	12	44	166
Simon	3	5	4	9	3[25]	1	12	44	107
Vicky	2	7	2	8	1	2	11	35	135
Moyenne	**3,28**	**4,42**	**2,57**	**7,21**	**2,07**	**1,57**	**9,92**	**29,35**	**105,85**

Tableau 27 – Caractéristiques intertextuelles des productions finales.

Ramené à des chiffres qui correspondent à des textes ayant 100 mots (cf. annexe IV.3.), le tableau ci-dessus permet de dire qu'en moyenne, dans les productions finales, 27,72 % de mots sont co-présents dans le texte d'apprenant et les corpus et/ou concordances mis à disposition lors de la rédaction. Les textes contiennent peu de passages non conformes (1,57 pour un texte de 100 mots). La valeur moyenne la plus importante (7,21) concerne le nombre

[23] Un des lexèmes est un segment tronqué.
[24] Un des segments est un passage long tronqué.
[25] Un des segments longs comporte deux modifications.

des passages co-présents non modifiés. Viennent ensuite les valeurs pour le nombre des segments (4,42), des lexèmes (3,28) et des passages longs co-présents (2,57). Ce résultat ne va pas dans le sens de l'observation de Siepmann (2008 : 194) selon laquelle les apprenants réemploient davantage de lexèmes simples, repérés dans un *input* en LE, que de segments qui se composent d'au moins deux lexèmes. J'ai observé la présence de deux fois plus de segments et passages longs (4,42 + 2,57 = 6,99) que de lexèmes simples (3,28).

6.3.1.1. Segments et passages longs co-présents et quantité d'*input* en LE

Les deux tableaux présentés dans ce paragraphe cherchent à mettre en regard la présence de segments ou passages longs co-présents et la quantité d'*input* auquel les apprenants avaient accès lors de la rédaction de leur "Commentaire dans un livre d'or électronique". Les prénoms écrits en majuscules marquent des résultats qui ne sont pas conformes à mes attentes. Les chiffres correspondent à une longueur de texte de 100 mots.

Prénom	**Nombre de segments co-présents**	**Prénom**	**Nombre de mots à disposition sous forme de mini-corpus**
Chloé	7,76	Camille	4672
Camille	6,52	JANINE	4672
Jean	5,61	OLIVIA	4672
Vicky	5,18	Simon	4672
Léa	4,76	GEORGES	2996
Simon	4,67	GÉRALDINE	2996
Christine	4,44	Jean	2996
OLIVIA	4,21	Léa	2996
Alexa	3,47	Vicky	2996
Fernand	3,36	Chloé	1676
GÉRALDINE	3,36	Alexa	362
GEORGES	2,38	Christine	0
Corentin	1,26	Corentin	0
JANINE	1	Fernand	0

Tableau 28 – Segments co-présents et quantité d'*input* accessible aux apprenants.

Pour dix apprenants sur quatorze, le tableau ci-dessus permet de trouver une certaine corrélation entre la quantité d'*input* à disposition et la quantité de segments co-présents dans le texte qu'ils ont produit. Olivia, Géraldine, Georges et surtout Janine semblent avoir peu exploité les mini-corpus auxquels ils avaient accès lors de la rédaction de la dernière version de leur "Commentaire dans un livre d'or électronique".

Prénom	**Nombre de passages longs co-présents**	**Prénom**	**Nombre de mots à disposition sous forme de mini-corpus**
Camille	6,52	Camille	4672
CHRISTINE	5,55	Janine	4672
Janine	4	OLIVIA	4672
Simon	3,73	Simon	4672
FERNAND	3,36	Georges	2996
Géraldine	3,36	Géraldine	2996
Georges	2,38	JEAN	2996
Léa	2,38	Léa	2996
Vicky	1,48	Vicky	2996
OLIVIA	1,2	Chloé	1676
Chloé	0,97	Alexa	362
Alexa	0	CHRISTINE	0
Corentin	0	Corentin	0
JEAN	0	FERNAND	0

Tableau 29 – Passages longs co-présents et quantité d'*input* accessible aux apprenants.

Pour dix apprenants sur quatorze, le tableau ci-dessus montre une corrélation entre le nombre de mots des corpus accessibles aux apprenants et le nombre de passages longs co-présents dans les corpus et leur production écrite. Christine et Fernand, qui n'ont pas eu accès à des mini-corpus, intègrent dans leur production plus de passages longs qu'attendu (respectivement 5,55 et 3,36 passages pour un texte de 100 mots). Ils semblent avoir choisi de façon parfois aléatoire certains passages dans le texte de l'énoncé et les 87 lignes de concordance, car certains des passages co-présents apparaissent comme des "pièces rapportées" dans leurs textes. Voici un extrait de la production de Christine. Le passage co-présent, non cohérent par rapport à ce qui précède, est souligné.

> *Ich bin *gekommen mit *meinen Familie (*Mein *Vatter, Meine Mutter und mein 2 *bruders). Wir haben *gefinden es *interressant, *auffreged. (...) Ich würde gerne mal Ihr Museum besuchen denn ich bien Rentner.*
>
> [Je suis venue avec ma famille (mon père, ma mère et mes deux frères. Nous l'avons trouvé intéressant, captivant. (…) J'aimerais, un jour, visiter votre musée car je suis retraité.]

Jean a eu à disposition le plus grand des deux corpus. Pourtant, il n'y a aucun passage long co-présent dans sa production finale. Comme dans le tableau précédent, il apparaît qu'à proportion, Olivia n'a pas exploité autant que les autres apprenants les données auxquelles elle a accès.

6.3.1.2. Segments co-présents et erreurs de syntaxe

Certains des segments co-présents contiennent des unités qui paraissent "préfabriquées" du point de vue syntaxique. Afin d'observer s'il y a un lien entre l'utilisation de segments co-présents et les caractéristiques syntaxiques de la production écrite des apprenants, je présente ci-dessous un tableau qui juxtapose le nombre de segments co-présents et le nombre d'erreurs de syntaxe dans la production finale. J'y confonds tous les types d'erreurs.

Prénom	**Nombre de segments co-présents (valeur décroissante)**	**Prénom**	**Nombre d'erreurs de syntaxe (valeur croissante)**
CHLOÉ	7,76	Camille	0
Camille	6,52	VICKY	0,74
JEAN	5,61	JANINE	1
VICKY	5,18	Léa	1,19
Léa	4,76	GÉRALDINE	1,68
SIMON	4,67	CHLOÉ	1,94
CHRISTINE	4,44	JEAN	2,24
OLIVIA	4,21	GEORGES	2,38
Alexa	3,47	Alexa	2,5
Fernand	3,36	SIMON	2,8
GÉRALDINE	3,36	CHRISTINE	3,33
GEORGES	2,38	Fernand	3,36
Corentin	1,26	OLIVIA	3,61
JANINE	1	Corentin	5,06
Moyenne	**4,14**	**Moyenne**	**2,27**

Tableau 30 – Segments co-présents et erreurs de syntaxe.

La comparaison des deux valeurs ne permet pas de voir de lien entre une quantité importante de segments co-présents et une quantité faible d'erreurs de syntaxe. Étant donné que Janine, Georges et Géraldine ont utilisé moins de segments co-présents que la moyenne du groupe classe, on pouvait s'attendre à ce qu'ils fassent davantage d'erreurs. Je m'attendais à ce que Simon, Christine et Olivia en fassent moins, vu le nombre de segments co-présents que j'ai trouvés dans leurs textes. Il convient de relativiser ce résultat par étude plus approfondie, portant sur l'évolution de la syntaxe d'un échantillon d'apprenants au fil de *Projekt Prospekt*. Une telle étude sera présentée en 6.4.3. Elle tiendra compte de la complexité des groupes verbaux et nominaux dans les productions afin de rendre compte des changements en matière de syntaxe en relation avec l'exploitation de corpus et concordances.

6.3.1.3. Sur-utilisation de segments co-présents

Olivia et Janine réutilisent plusieurs fois les mêmes lexèmes ou segments co-présents. Olivia emploie dans sa production finale entre deux et trois fois les mêmes lexèmes et segments paratextuels qui introduisent ou closent un texte de type "Commentaire dans un livre d'or électronique". Les lexèmes "*Eintrittspreise*" [prix d'entrée] et "*entrichten*" [s'acquitter de] apparaissent également deux fois. Elle varie essentiellement les noms dans des segments ou des passages longs. Le texte de Janine contient douze passages co-présents. Je n'en retiens que neuf, car elle utilise deux fois le lexème "*Hallo*" [salut] et trois fois "*danke*" [merci]. Un autre segment est utilisé deux fois, chaque fois à l'intérieur d'un passage long co-présent, non modifié.

Ich finde es sehr schön. [Je le trouve très beau].

*Ich finde es *interresant.* [Je le trouve intéressant.]

Les éléments co-présents dans la production de Janine sont essentiellement des lexèmes simples ou des passages longs. Deux des passages longs apparaissent dans le nouveau contexte d'une façon qui n'est pas acceptable du point de vue sémantique. J'ai repéré un seul segment co-présent dans son texte. Janine présente l'adjectif de ce segment sans terminaison, ce qui est en adéquation avec le nouveau contexte.

Concordance : "*eine sehr gute Führung*" [une très bonne visite guidée]

Texte de Janine : "*Das Museum ist sehr gut*" [Le musée est très bien].

L'observation de Siepmann (2008 : 194) concernant la sur-utilisation de certains lexèmes ou segments venant d'un *input* en LE se trouve confirmée par les caractéristiques des productions d'Olivia et de Janine que je viens de décrire.

6.3.1.4. Incorporer l'*input* en LE dans sa production – Jean et le salon du modélisme

Selon Tricot (2007 : 10), apprendre est "la capacité [des organismes vivants] à s'adapter à leur(s) environnement(s)" (cf. 2.2.1.). Un tel processus d'adaptation progressive aux corpus et concordances peut être observé quand on étudie les diverses versions des "Informations pratiques" produites par Jean.

Jean a rédigé quatre versions du texte de type "Informations pratiques". D'une version à l'autre, il est allé de plus en plus vers l'*input* en LE. J'ai observé ce que l'on peut appeler une

"SPA" (Séquence Potentiellement Acquisitionnelle, cf. de Pietro, Matthey et Py, 1989) dans le domaine des mots composés.

V1 : **weltweit *vom modellbau* [mondial du modélisme] / *Messe* [salon]
V2 : *Messe* [salon, foire]
V3 : **Weltweit *Modellbaumusseum* [mondial musée du modélisme] / *Messe *Musseam* [musée du salon]
V4 : **Modellbau Museum* [musée du modélisme] / **Modellbau Messe* [salon du modélisme].

Tableau 31 – Émergence de mots composés lors de la rédaction des "Informations pratiques".

Dans la version V1, Jean propose un segment ayant une structure syntaxique non conforme, proche de celle du segment français "mondial du modélisme". Le lexème "*Messe*" est conforme, mais il ne précise pas la spécialité du salon, à savoir le modélisme. Dans la version V2, Jean renonce complètement au segment correspondant à "mondial du modélisme". Ce segment réapparaît en V3 sous la forme de "**Weltweit *Modellbaumusseum* [*mondial musée du modélisme]. Dans mes annotations de la V2, j'ai proposé à Jean d'observer le nom de musée "*Modellbaumuseum*" [musée du modélisme], afin de l'inciter à s'en inspirer pour trouver un mot composé se terminant en "…*messe*". Jean a repris ce terme (ce qui n'était pas prévu par moi), et il a rajouté "*Messe *Musseam*". Les mots composés de la version V3 correspondent, davantage que les termes choisis auparavant, à un résultat acceptable. Les mots composés que Jean propose dans la version V4 sont presque parfaits. Il suffit d'enlever l'espace entre les deux noms du mot composé pour obtenir le terme conforme aux normes fixées : *Modellbaumuseum, Modellbaumesse* [musée du modélisme, salon du modélisme].

Tout au long du projet, la précision dans le domaine de l'orthographe et des terminaisons et les prépositions spatiales et temporelles ont constitué des difficultés pour lui. Pendant plusieurs versions, Jean s'est, en effet, concentré sur la révision de lexèmes, ce qui caractérise particulièrement les apprenants en situation d'insécurité (cf. 2.9.). Pour sa dernière version, il a changé de procédé. Il a copié un texte intégral du corpus "Informations pratiques", et il a adapté ce texte à son sujet. On peut lui reprocher d'être resté trop proche de l'original ; Jean n'a pas modifié ni le nom de la ville, ni les numéros de téléphone, etc. Néanmoins, il a enlevé les parties superflues, et il a retravaillé plusieurs parties. La version finale est acceptable. La liste suivante indique les types et le nombre de changements effectués par Jean lors de la révision de la version V3 ses "Informations pratiques" : insertions : 10 ; suppressions : 11 ;

déplacements : 1 ; modifications pertinentes : 5 ; modifications non pertinentes : 3 ; modifications neutres : 4. La comparaison des versions V3 et V4 de Jean et le tableau concernant la révision des "Informations pratiques" du groupe classe entier peuvent être consultés en annexe II.3.

Jean n'a pas intégré dans sa dernière version tous les éléments qu'il a rédigés auparavant. Il manque, par exemple, l'information concernant la validité du ticket d'entrée pendant plusieurs jours. Cette information est caractéristique pour un évènement de type salon. Le texte du corpus prend une place importante par rapport à la contribution personnelle. Au moins le temps de l'exécution de la dernière version, Jean se décentre beaucoup, presque au point de perdre sa propre "structure", selon les termes de Piaget (cf. 2.7.1.).

6.3.1.5. Formes de plagiat

Pendant la première et deuxième phase de la recherche, certains apprenants (en particulier Yvan et Jean) ont plagié des passages ou des textes entiers, trouvés sur Internet. Plusieurs apprenants ont fait le choix de se servir de l'*input* d'un de leurs camarades, en le copiant à partir du blogue ou dans la boîte courriel commune où les textes déposés ont été accessibles à tous. Lors de la rédaction de la version V1 de ses "Informations pratiques", Georges a présenté le même texte que celui que Christine venait de rendre. Perrine a participé à la première et à la deuxième phase de la recherche-action. Lors de la première phase de la recherche-action, deux des segments utilisés par elle dans son texte "Commentaire dans un livre d'or" suggèrent qu'elle s'est s'appuyée, comme lors de la formulation de la tâche précédente, sur des éléments qu'elle a pu identifier dans les productions de ses pairs, visibles pour tous sur le blogue du projet. Les éléments co-présents sont soulignés dans le texte ci-dessous.

> *Ich denke dass das Museum besser wäre, wenn es <u>Aktivitäten für Kinder</u> gibt. <u>Danke für dieses shönes Museum</u>!!!*
> Ecrit par : perrine | 26.03.2007

Pour rédiger la version de son texte qu'elle a publiée le 26 mars 2007, Perrine a pu s'inspirer du texte de Charles (publié le 18 mars 2007) et de celui de Jasmine (publié le 21 mars 2007). Elle fait la même erreur de terminaison que Jasmine sur l'adjectif "schönes".

> *Man kann auch Aktivitäten und Führungen für Kindern schaffen.*
> *Ecrit par : Charles | 18.03.2007*

Danke für dieses schönes Museum.
Ecrit par : Jasmine | 21.03.2007

Toujours lors de la première phase de la recherche, Béatrice, une apprenante peu avancée, semble également avoir utilisé les textes d'autres apprenants comme *input*, lors d'une tâche qui a été exécutée en dehors du projet, mais selon les mêmes modalités que les tâches liées à la recherche-action. Béatrice précise d'ailleurs dans le mini-questionnaire du 18 février 2007 que le travail collaboratif lui convient : "J'ai du mal à formuler mes rubriques. J'ai bien aimé l'idée de faire un peu de travail ensemble, en classe, car j'ai du mal chez moi."

Le fait de s'appuyer sur des productions d'autres apprenants rappelle la graduation observée par Šimunek (2007 : 272) lors d'activités de recherche lexicale en petits groupes. Šimunek explique que, lors de sa recherche-action, les apprenants qui avaient besoin de trouver un terme en anglais ont eu tendance à demander d'abord conseil aux partenaires de leur groupe. Ensuite, ils ont demandé de l'aide en allant voir leurs pairs des autres groupes. Si personne n'a réussi à les aider, ils ont eu recours au dictionnaire. Il est possible d'avancer que l'*input* fourni par les pairs est moins menaçant pour des apprenants peu avancés que l'*input* d'un corpus de taille assez importante. Cette démarche peut alléger le coût cognitif d'un travail d'identification de segments. Il convient toutefois de trouver d'autres moyens susceptibles d'alléger l'*input*. Je pense par exemple à une réduction de la taille du contexte d'un mot clé dans une ligne de concordance ou au nettoyage des lignes de concordance (cf. 6.2.6.3.).

6.3.2. Utiliser la concordance comme aide pour la révision de texte

Pour certains apprenants, l'adaptation aux nouvelles méthodes de révision a été aisée, pour d'autres moins, comme l'illustrent les quatre situations décrites dans ce sous-chapitre. Les deux premières permettent de voir comment deux apprenants modifient ou ne modifient pas des lignes de concordance lors de leur intégration dans le nouveau co-texte. La troisième situation décrit un procédé d'exploitation de lignes qui mène à une conceptualisation dans le domaine des mots composés. La quatrième situation présente une amorce de travail collaboratif et une découverte de type sérendipité.

6.3.2.1. Les annonces de Marinette et de Bastien

Les deux situations de ce paragraphe ont été observées lors du projet *Brieffreunde.org* [correspondants.org].

La première version de l'annonce de Marinette contient une phrase entière qui est co-présente dans le corpus *Brieffreunde.org*. L'apprenante a recopié manuellement cette phrase, en remplaçant le mot "*jede*" [chaque] par "*gehe*" qui veut dire "(je) vais". Dans sa deuxième version, elle a identifié dans la concordance le segment "*jede ANTWORT*" [chaque réponse], et elle l'a intégré dans sa production en s'aidant du traitement de texte, toutefois sans changer la casse.

V0 :
*Ich freue mich über *gehe antwort*. [*Va réponse me fera plaisir].
V1 :
Ich freue mich über jede ANTWORT. [Toute RÉPONSE me fera plaisir].

Les deux procédés de reprise d'un segment de Marinette – le recopier manuellement ou le copier à l'aide du traitement de texte – ont entraîné des effets non souhaités. Une troisième version aurait été nécessaire pour "gommer" la trace d'importation du segment, comme Bastien l'a fait.

La première version de l'annonce de Bastien contient le segment "*von 13 bis 16*" [de 13 à 16]. Ce segment apparaît aussi dans la concordance, mais à une différence près : à la place de la préposition "*bis*" [jusqu'à, à], nous trouvons un tiret ("*von 13-16*"). Dans le co-texte de la première version de l'annonce de Bastien, ce segment n'est pas acceptable. Suite à une annotation de ma part, il a intégré dans son texte un segment plus long que le précédent : "*im alter von 9-13*" [entre 9 et 13 ans]. Il a modifié les chiffres indiquant l'âge, et il a gardé le tiret proposé dans la concordance. Il y a une différence supplémentaire entre son segment et celui du corpus : le segment du corpus n'est pas inséré dans une relative, contrairement au sien. Lors de la révision de la version V0, Bastien a aussi modifié "*Schreib*-mir!*" en "*Schreib mir!*" et "*Kussi*" [bisous, salutation plutôt prononcée par des filles] en "*Gruß*" [salutations].
V0 :
*Ich suche *jemand *das *von 13 bis 16 ist. Schreib-mir! Kussi, Bastien* [Je cherche quelqu'un entre 13 et 16 ans. Écris-moi ! Bisou, Bastien].
V1 :
*Ich suche *jemand, der im alter von 13-16 ist. Schreib mir! Gruß, Bastien.* [Je cherche quelqu'un qui a entre 13 et 16 ans. Écris-moi ! Salutations, Bastien].

Bastien a tiré profit des lignes de concordance : il s'est adapté à la présentation d'une tranche d'âge avec tiret, et il l'a enlevé du groupe verbal où un tiret n'est pas conforme.

6.3.2.2. Vicky et la joueuse de clarinette

Lors de la phase préparatoire de *Projekt Prospekt*, le groupe classe s'est entraîné à utiliser le corpus et la concordance faits à partir d'un petit corpus de critiques du film *Jenseits der Stille* [Au-delà du silence] de Caroline Link que nous avons vu ensemble. J'ai encouragé les apprenants à se servir de ce corpus et des lignes de concordance associées lors de la rédaction de leur critique de ce film.

Suite à la rédaction de la version V0 de sa critique, Vicky avait, entre autres, à réviser le passage suivant : "*eine gute *Klarinette Spielerin*" [une bonne joueuse de clarinette]. Je lui ai proposé comme outil de révision la concordance "Critiques de film". Mon objectif a été de l'aider à trouver le terme de *Klarinettistin* [clarinettiste, au féminin], terme qui ne se trouve pas tel quel dans la concordance. Je m'attendais à ce que Vicky repère la forme masculine au datif *Klarinettisten* qui se trouve dans la dernière des huit lignes de concordance pour le lemme *Klarinett**.

ein Pflichtfilm für jeden der KLARINETTE spielt.
und sich in den Klang der KLARINETTE verliebt. Daraufhin wird sie den Eltern
Nun spielt sie KLARINETTE entwickelt Talent und bricht aus.
Auch bin ich durch diesen Film zum KLARINETTE spielen gekommen
besucht ihre Tante in Berlin die auch KLARINETTE spielt
bezaubernde Augenblicke anrührende KLARINETTEN-MUSIK wundervolle Landschaften
Ich mag KLARINETTENMUSIK sehr
Wer Fan von dem KLARINETTISTEN Giora Feidman ist

Figure 43 – Lignes de concordance pour le mot clé "*Klarinett**".

VICKY : *Klarinette* prend bien deux "t" ?

ENSEIGNANTE : Oui, ce n'est pas ça, le problème.

VICKY : Alors, pourquoi… ah…

ENSEIGNANTE : Tu veux dire "joueuse de clarinette".

VICKY : Ah oui.

ENSEIGNANTE : Hein ? Alors, là,…

VICKY : Hein…

ENSEIGNANTE : …tu ne trouveras pas parfaitement là-dedans, mais il y a un terme qui est un petit peu comme ça.

VICKY : "KLARINETTENMUSIK" là ? Ou alors, ici ?

ENSEIGNANTE : Là, tu as… oui, tu peux te calquer sur ce m… Alors, tu ne veux pas dire "musique", mais "joueur", donc ? Fais-le sur le modèle.

VICKY : *Spieler* ?

ENSEIGNANTE : Oui.

VICKY : *Klarinetten…spieler* ?

ENSEIGNANTE : Oui. Comme ça. Tu rajoutes juste un "n" et tu le… Ah oui ! Je trouve ça super intéressant.

VICKY : Je le mets attaché ?

ENSEIGNANTE : Oui. J'ai essayé de faire la démarche, de le trouver, à ta place…

VICKY : Mhm.

ENSEIGNANTE : … et je n'ai pas réussi. Et toi, tu as réussi parce que tu as une autre technique que moi. Tu as pris en fait un mot composé qui existe. Ah, ça, c'était super, ça. Moi, j'ai dû passer par un autre corpus qui existe [pour trouver la forme féminine].

Vicky a repéré dans la concordance les mots composés *KLARINETTEN-MUSIK* et *KLARINETTENMUSIK*, et elle a créé à partir de ces mots le terme de *Klarinettenspielerin*. J'ai jugé nécessaire d'évaluer l'acceptabilité de ce terme en testant sa fréquence. Pour ce faire, j'ai recherché *Klarinettistin** et *Klarinettenspielerin** et leur pendant masculin dans deux concordances et sur la Toile, en utilisant deux moteurs de recherche.

Lemme recherché	***Sketch Engine***	***Cosmas II***	***Google***[26]	***Exalead***
*Klarinettenspielerin** *Klarinettistin**	0 53	8 744	1.133 14.850	105 1.650
*"Klarinettistin**"* est … fois plus fréquent que *"Klarinettenspielerin**"*.	53	93	13,1	15,71
*Klarinettenspieler** *Klarinettist**	22 375	54 3.253[27]	4430 167.100	241 16.933
*"Klarinettist**"* est … fois plus fréquent que *"Klarinettenspieler**"*.	17,04	60,24	37,72	70,26

Tableau 32 – Fréquence de *Klarinettenspieler(in)* et de *Klarinettist(in)* dans des corpus et des moteurs de recherche.

[26] Je n'ai pas effectué de recherche par lemme pour *Google*, ni pour *Exalead*. J'indique la somme des occurrences au singulier et au pluriel nominatif et accusatif. Je ne mentionne le nombre d'occurrences au datif ou génitif.

[27] Chiffre obtenu après déduction du nombre d'occurrences pour *Klarinettistin**.

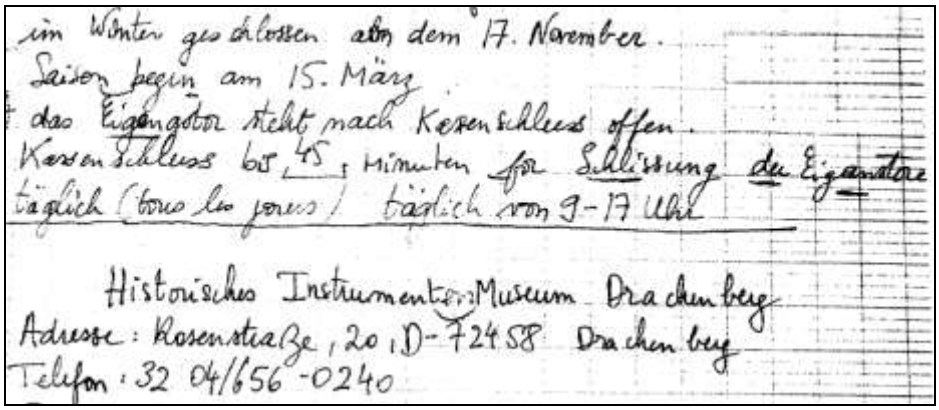
Je constate que les deux très grands corpus *Sketch Engine* et *Cosmas II* affichent une différence de fréquence importante entre *Klarinettist* et *Klarinettenspielerin* : *Klarinettist(in)* est attesté entre 53 et 93 fois plus souvent que *Klarinettenspieler(in)*. *Cosmas II* atteste aussi une différence importante pour la forme masculine. Les moteurs de recherche attestent une différence de fréquence moins importante, sauf *Exalead* qui trouve 70,26 fois plus d'occurrences pour la forme masculine *Klarinettist* que pour *Klarinettenspieler*.

Pourquoi Vicky a-t-elle opté pour le terme de **Klarinette Spielerin* lors de la rédaction de sa première version ? En français, la différence de fréquence entre "clarinettiste" et "joueuse/joueur de clarinette" est également importante : le moteur de recherche *Exalead* affiche 96,46 fois plus d'occurrences pour "clarinettiste" que pour "joueuse/joueur de clarinette". Le choix de **Klarinette Spielerin* n'est donc pas forcément une trace de nativisation, ou en tout cas, Vicky ne semble pas avoir utilisé la langue française comme (seule) référence. D'ailleurs, la première page que le moteur de recherche *Exalead* affiche pour le mot clé *Klarinettenspielerin* renvoie à deux sites ayant trait au film *Jenseits der Stille* qui a constitué la base de travail pour la tâche que Vicky avait à effectuer. En plus, l'interface française d'*Exalead* propose un terme de recherche alternatif qui ressemble à s'y méprendre à la version que Vicky avait à réviser : *klarinetten spielerin*.

Figure 44 – Résultats de recherche pour "*Klarinettenspielerin*" (*Exalead*).

Malgré la fréquence absolue plus importante de "*Klarinettistin*", je considère le terme de "*Klarinettenspielerin*" comme acceptable dans la critique de film de Vicky, d'autant plus que la concordance à laquelle Vicky a eu accès ne comporte pas le terme de "*Klarinettistin*". Elle a su tirer profit de l'outil qui a été mis à sa disposition pour réviser un passage non conforme aux normes fixées.

La figure ci-dessous illustre que le travail de révision à l'aide de concordances a eu des conséquences sur une des productions ultérieures de Vicky, la version V2 de ses "Informations pratiques". Lors de la séance individuelle de révision, elle m'a adressé la question suivante : "Je le mets attaché [les deux noms du mot composé] ?" Dans l'extrait de texte ci-dessous, elle dessine deux fois un lien d'"attachement" sous un groupe de deux noms dont elle fait ainsi un mot composé conforme : *Saison beginn* > *Saisonbeginn* [ouverture de la saison] ; *Instrument Museum* > *Instrumentenmuseum* [musées des instruments].


im Winter geschlossen am dem 17. November.
Saison begin am 15. März
das Eigengstor steht nach Kassenschluss offen.
Kassenschluss bis 45 Minuten für Schlissung der Eigenstore
täglich (tous les jours) täglich von 9-17 Uhr

Historisches Instrumentenmuseum Drachenberg
Adresse: Rosenstraße, 20, D-72458 Drachenberg
Telefon: 32 04/656-0240

Figure 45 – Traces d'entraînement avec des lignes de concordance.

6.3.2.3. La découverte de Virgile – révision collaborative

Dans l'entretien, Camille cite un des procédés de recherche que l'on peut appliquer à des concordances, faites à l'aide de *Text-Concordancer (v2)* de Cobb (nd) proposant peu d'options d'interrogation. Ce procédé consiste à parcourir un fichier texte, en faisant des recherches à l'aide de la présentation des mots clés selon l'ordre alphabétique (cf. 6.2.6.3.). Illustrons cela par la description d'une séance de révision de passages d'un texte à l'aide d'une concordance, afin de voir à quels résultats peut mener cette façon de consulter une concordance.

La séance a eu lieu le 5 décembre 2008, donc lors de la troisième phase de la recherche-action. J'ai proposé à Marinette, Juliette et Virgile de réviser les passages soulignés dans la critique de film de Chloé, une apprenante qui a participé à la deuxième phase de la recherche.

Dans ce qui suit, je reproduis la feuille de travail que j'ai distribuée aux apprenants. Ensuite, je présente un extrait du script du compte-rendu oral que j'ai enregistré tout de suite après la séance.

> *Ich mag den Film. Die Taubheit ist ein selten Thema. Aber ich mag die musik des Films. Ich habe gefunden, dass Laras Onkel sympatisch ist aber ich habe Laras Tante nicht gemögt.* (Chloé)
>
> Regardez ensemble ce texte. Qu'est-ce qui "ne va pas", à votre avis, dans les passages soulignés ? Qu'est-ce qui devrait être modifié ? Formulez vos hypothèses à ce sujet et testez-les à l'aide de la concordance, éventuellement de LEO ou du corpus. N'ayez pas recours au traducteur.
>
> Rédigez toutes vos observations, questions, hypothèses, même si vous n'êtes pas d'accord l'un avec l'autre. Indiquez les outils (concordance, LEO, corpus) que vous avez utilisés pour tester vos hypothèses. Rédigez la solution pour laquelle vous avez opté.

Figure 46 – Activité de révision collaborative.

Le texte à réviser contient quatre types d'erreur : une erreur d'orthographe (**sympatisch*), une erreur de terminaison dans un groupe nominal complexe (*ein *selten Thema*), une erreur discursive (*Ich mag... *Aber ich mag…* [J'aime… *mais j'aime…]) et une erreur de syntaxe et/ou de temps du verbe principal (*Ich habe nicht…* *gemögt [Je n'ai pas aimé…]). Les apprenants ont rapidement trouvé l'erreur de type discursif, sans consulter la concordance. Ils se sont surtout concentrés sur les erreurs de terminaison et d'orthographe qu'ils ont révisées à l'aide de la concordance.

> [La séance] était très très vivant[e], très intéressant[e], on a vraiment pu voir que l'intérêt des concordances s'est manifesté, que les élèves ont réussi à accepter que ça peut apporter quelque chose de prendre la peine d'y aller. Et ce qui est particulièrement intéressant, c'est qu'il y a des résultats inattendus, par exemple, une fois qu'on a trouvé quel était le problème avec "*ein selten Thema*", où il manquait en fait la terminaison de l'adjectif, Virgile a observé une formulation qui était écrite juste en dessous de quelque chose qu'on a observé, il a dit : "Oui, et maintenant, pourquoi là, il y a '…er' ?" Donc, Marinette lui a expliqué que c'était en raison du nom masculin qui suivait et en raison de l'absence de terminaison sur le déterminant "*ein*". Cela a donné des résultats supplémentaires. (…) Virgile a cherché avec "contrôle F – rechercher dans ce texte", ce qui n'a pas toujours aidé, comme par exemple pour "*ist*" [est], suivi d'un adjectif. On a obtenu un résultat quand on a observé le mot clé *IST*, en profitant du classement alphabétique des mots clés de la concordance. Ça donne quand même pour eux des idées sur comment il faut procéder après quand ils ont à chercher. J'ai vu vraiment plusieurs fois des effets comme "ah, oui, ah, c'est bien".

Ce script témoigne d'un changement concernant le sens de la lecture : Virgile lit de haut en bas. Cela semble l'amener à s'interroger sur la terminaison d'un adjectif, et Marinette lui explique l'information transmise par cette terminaison. À part la découverte de Virgile, de type sérendipité, et l'amorce d'un travail collaboratif, le bilan de la séance d'observation de lignes de concordance n'est pas entièrement positif : la révision de "**sympatisch*" prend plus de temps qu'avec un dictionnaire, et la recherche par ordre alphabétique dans une concordance faite à l'aide de *Text-Concordancer (v2)* est également chronophage. Comme l'illustre la figure suivante, ce concordancier ne permet pas aux apprenants de réduire le co-texte des mots clés à un minimum suffisant.

Die Soldaten die am Ostfront waren schienen SYMPATHIE mit Sophies Bruder zu haben. Der böse Ri
mmer noch empfehlenswert weil es einfach ein SYMPATHISCHER unterhaltsamer Familienfilm ist. 24
arn gefilmt worden. Die Charaktere sind sehr SYMPATISCH und das Ende finde ich auch gelungen mit
Dafür sorgen wahrscheinlich die frechen und SYMPATISCHEN Hauptdarsteller das Tempo der Handlung

Figure 47 - Lignes de concordance proposées par *Text-Concordancer (v2)* pour **sympatisch*.

Les concordanciers faisant partie des collections d'outils *AntConc3.2.1w* et *Lexico3* comportent, par rapport à *Text-Concordancer (v2)*, des avantages indéniables en matière de recherche ciblée d'informations, notamment la possibilité de réduire ou d'élargir la taille du co-texte du mot clé.

KWIC
aren, schienen Sympathie mit Sophi
tere sind sehr sympatisch und das
ie frechen und sympatischen Hauptd
es einfach ein sympathischer, unte
arrogant und unsympathisch rüber!!

Figure 48 – Lignes de concordance proposées par *AntConc3.2.1.w* pour *symp**.

Dans le corpus *Kundenrezensionen* [critiques de film], *AntConc3.2.1w* permet d'identifier une seule ligne pour le mot clé **sympatisch*. Si on ne tape que le début du mot, *symp*, cinq lignes s'affichent, c'est-à-dire une de plus que *Text-Concordancer (v2)*. En effet, *AntConc3.2.1w* affiche aussi l'antonyme *unsympathisch* [antipathique]. Trois des cinq lignes contiennent un résultat pertinent.

6.4. Évolution des productions écrites des apprenants

Dans ce qui suit, j'essayerai d'établir un lien entre l'utilisation des corpus et concordances et l'évolution des productions des apprenants. J'ai choisi d'examiner en 6.4.1. le domaine des adjectifs dans toutes les productions du groupe classe concernant les textes de type "Visite guidée" et "Commentaire dans un livre d'or électronique". En 6.4.2., je propose une étude des caractéristiques syntaxiques des textes les plus longs produits par Alexa, Géraldine et Vicky.

6.4.1. Concordances et évolution des adjectifs

6.4.1.1. Pourquoi et comment relever les adjectifs dans les corpus ?

Inspirée de l'observation de Siepmann (2008 : 194) concernant la sur-utilisation de certains lexèmes et collocations par les apprenants (cf. 2.9.), j'ai étudié la liste de fréquence de tous les mots des corpus contenant les textes des apprenants et des corpus de référence de *Projekt Prospekt*. Le premier résultat est le suivant : les corpus d'apprenants contiennent certains adjectifs qui sont utilisés à une fréquence plus élevée que dans les corpus réfléchis. Les lemmes des trois adjectifs les plus fréquents dans le corpus d'apprenants "Visites guidées" sont *gut** [bon, bien], *schön** [beau] et *interessant** [intéressant].

Type de corpus	***gut***	***schön***	***interessant***
Corpus d'apprenants "Visite guidée" (CA)	77,51	64,03	33,7
Corpus réfléchi "Visites guidées" (CR)	22	16	11
Proportion CR : CA	1 : 3,52	1 : 4	1 : 3,06

Tableau 33 – Comparaison de fréquence des adjectifs *gut*, *schön* et *interessant*.

Ces trois adjectifs apparaissent entre trois et quatre fois plus souvent dans le corpus d'apprenants que dans le corpus de référence. C'est pour cela que j'ai choisi d'observer, à l'aide de la collection d'outils *AntConc3.2.1w*, tous les adjectifs dans les "Productions de départ", les "Visites guidées" et les "Commentaires dans le livre d'or d'un musée". J'ai écarté les textes de type "Informations pratiques", car la proportion d'adjectifs y est moins élevée. Pour faciliter l'identification des adjectifs, j'ai utilisé l'outil en ligne *TreeTagger output visualisation module* (Koller, nd). Cet outil permet d'annoter un corpus d'une taille maximale de 1000 mots. Grâce à ce logiciel, j'ai pu annoter un échantillon de 875 mots du corpus de référence "Commentaires dans un livre d'or" et un échantillon de 875 mots du corpus d'apprenants, contenant la dernière version des commentaires de chaque apprenant.

Je constate que les adjectifs apparaissent plus souvent sous forme non fléchie dans les corpus d'apprenants que dans les corpus de référence. Dans les deux échantillons des "Visites guidées", j'ai repéré les fréquences suivantes. Les chiffres correspondent à des textes contenant 100 mots.

Type de corpus	**Formes fléchies**	**Formes non fléchies**	**Total**
Corpus réfléchi "Visite guidée" (CR)	5,71	1,6	7,31
Corpus d'apprenants "Visite guidée" (CA)	2,17	6,85	9,02
Proportion CR : CA	2,63 : 1	1 : 4,28	1 : 1,23

Tableau 34 – Formes adjectivales fléchies et non fléchies dans les corpus "Visites guidées".

Le tableau 34 atteste une sur-utilisation de formes non fléchies et une sous-utilisation des formes fléchies par les apprenants, par rapport aux chiffres du corpus de référence.

Si l'on retrace l'évolution du nombre des formes fléchies et non fléchies des adjectifs dans les corpus d'apprenants, on peut avoir des pistes pour savoir comment les apprenants exploitent les concordances faites à partir des corpus réfléchis. Grâce à une telle étude, il sera peut-être possible de trouver des éléments de réponse à la question de savoir si les concordances sont une aide pour les scripteurs qui rédigent en LE. Selon Py (2002), lors de l'apprentissage de LE, la prise en compte d'éléments macrosyntaxiques contribue à l'évolution de la microsyntaxe (cf. 2.4.2.), donc, par exemple, à l'évolution du marquage de l'adjectif. J'essayerai d'évaluer la pertinence de ce point de vue en l'appliquant au domaine de la révision des adjectifs à l'aide de corpus et concordances. Je considère les segments des textes des corpus, ayant une certaine cohésion discursive, comme éléments macrosyntaxiques.

6.4.1.2. Remarques concernant l'adjectif allemand

Les adjectifs allemands ayant la fonction d'épithète ont une terminaison qui indique le genre, le nombre et le cas du nom spécifié. Ils se placent systématiquement devant ce nom. Les adjectifs, ayant la fonction d'attribut du sujet ou d'adverbe, n'ont pas de terminaison. Comme détaillé en 6.4.1., les formes fléchies sont environ deux à trois fois plus fréquentes dans mes corpus réfléchis que dans mes corpus d'apprenants. Je m'attends donc à ce que la variation de la fréquence des formes fléchies ou non fléchies soit un indicateur pour l'évolution du style d'écriture des apprenants. Springer (1999) caractérise les productions (certes orales) des apprenants peu avancés comme parataxiques. Ce style comporte des structures attributives simples de type "x est y" ou "x s'appelle y", etc. J'ai identifié, dans les productions de départ de certains apprenants (Géraldine, Fernand, Olivia, Véronique, Perrine, Janine) un style de

rédaction qui a les caractéristiques nommées ci-dessus. Il est particulièrement visible dans la production de départ de Janine.

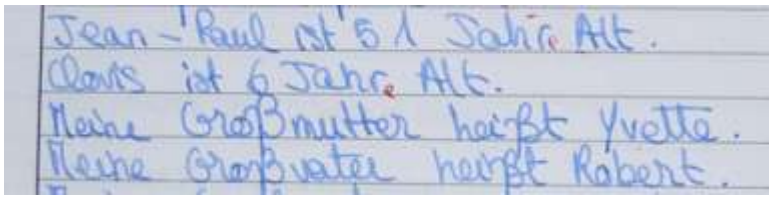


Figure 49 – Style parataxique chez Janine.

6.4.1.3. Comment je procède pour étudier les adjectifs

Dans ce qui suit, je présente des relevés concernant les adjectifs à forme fléchie, non fléchie, et ceux portant une erreur de marque ou de syntaxe. J'y ai ajouté le nombre d'adjectifs faisant partie d'un segment qui est co-présent dans les corpus de référence et les corpus d'apprenants.[28] Je cherche à savoir s'il existe un lien entre l'utilisation des corpus et concordances et l'émergence d'un style d'écriture complexe. Pour les besoins de mon étude, je définis une production "complexe" comme une production qui se caractérise par d'autres mises en relation des notions que ceux de type "attribut du sujet".

Mon étude se base sur 73 textes qui ont été produits par 13 apprenants. Après avoir présenté les résultats individuels, je regrouperai les chiffres afin de présenter les grandes lignes en matière d'évolution de l'emploi de l'adjectif. En raison du nombre peu important des productions qu'ils ont rendues, je ne présente par les études des textes de Louise, Corentin, Perrine et Véronique. Leur travail peut être consulté dans l'annexe III, sauf celui de Louise qui n'a rendu aucun texte. Je ne comptabiliser pas les adjectifs qui prennent la forme d'un participe passé, sauf s'ils ont la fonction d'épithète. Je classe les adjectifs fléchis qui font partie d'un segment co-présent et qui portent une marque dans la colonne "adjectifs fléchis ayant une fonction d'épithète", même s'ils n'ont pas cette fonction. Il peut s'agir d'expressions comme *am besten* [le mieux]. Ce segment peut avoir la fonction d'adverbe ou d'attribut du sujet. Les graphiques se basent sur des chiffres correspondant à de textes ramenés à 100 mots (cf. annexe IV.1.).

J'observe les points suivants dans les textes des apprenants :

- les adjectifs (qui devraient être) fléchis, ayant la fonction d'épithète ;

[28] Dans la production de départ, les segments ne sont pas co-présents, puisque cette production a été faite sans faire appel à un corpus. Néanmoins, j'ai repéré les groupes de mots qui relèvent du phénomène collocatif, afin de pouvoir comparer les données.

- les adjectifs non fléchis, ayant la fonction d'attribut du sujet ;
- les adjectifs non fléchis, ayant la fonction d'adverbe ;
- le nombre de terminaisons manquantes (on y retrouve ceux qui ont la fonction d'adverbe) ;
- le nombre de terminaisons superflues ;
- le nombre de marques erronées (l'apprenant a mis une terminaison là où il en faut une, mais il s'est trompé de marque) ;
- le nombre d'erreurs de syntaxe sur l'adjectif ;
- le nombre d'adjectifs faisant partie d'un segment co-présent (défini comme segment qui se trouve dans le corpus réfléchi et dans le texte de l'apprenant et qui a des caractéristiques collocatives, dans un sens large).

La plupart des apprenants ont rédigé la première et la dernière production dans des conditions de devoir sur table, donc en temps limité. La première production a été faite sans aide à la rédaction. Pour la rédaction de la dernière production, les apprenants avaient accès à un texte du corpus "Commentaire dans un livre d'or", à 87 lignes de concordance faites à l'aide de ce même corpus, et ils avaient le droit d'utiliser les mini-corpus "Visites guidées" et "Commentaires dans un livre d'or", sous forme papier.

Les tableaux qui introduisent chaque étude individuelle comprennent tous les points étudiés. Les graphiques se concentrent sur quatre des points, à savoir le nombre des adjectifs ayant la fonction d'épithète, d'attribut du sujet ou d'adverbe et les adjectifs faisant partie d'un segment co-présent.
J'utilise les abréviations suivantes : pdd = production de départ ; vg = "Visite guidée" ; cl= "Commentaire dans un livre d'or". Les études sont classées par ordre alphabétique.

6.4.1.4. Les adjectifs dans les productions d'Alexa

Type de production écrite	**Adjectifs (qui devraient être) fléchis +fonction épithète**	**Adjectifs non fléchis + fonction attribut du sujet**	**Adjectifs non fléchis + fonction adverbe**	**Nombre de terminaisons manquantes**	**Nombre de terminaisons superflues**	**Nombre de marques erronées**	**Nombre d'erreurs de syntaxe (adjectif)**	**Nombre d'adjectifs faisant partie d'un segment co-présent**	**Nombre total de mots**
vgV1	3	0	0	0	0	0	0	1	57
vgV2	4	0	0	0	0	0	0	3	65
clV0	7	9	0	0	1	3	0	4	115

Tableau 35 – Adjectifs dans les productions d'Alexa.

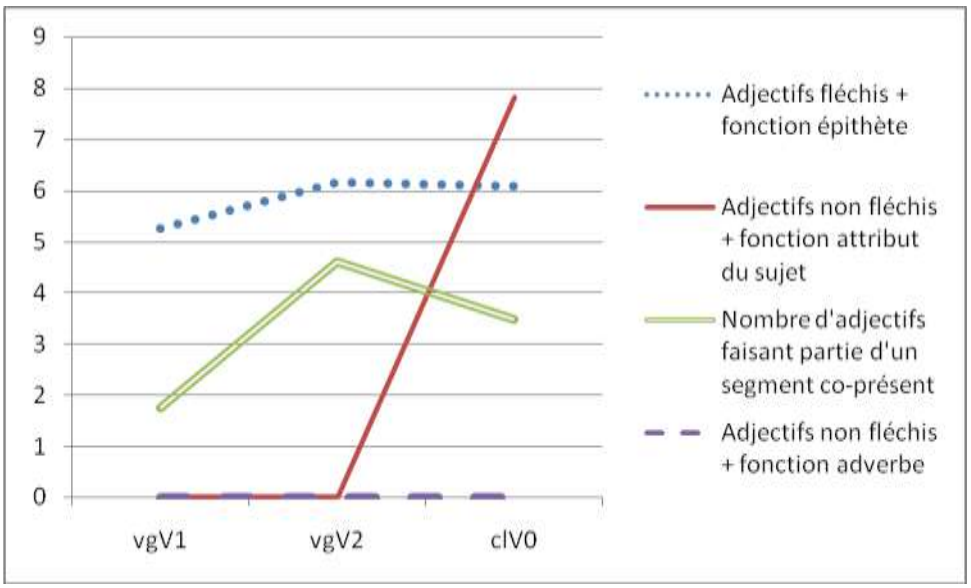


Figure 50 – Adjectifs dans les productions d'Alexa.

Alexa n'a pas été présente lors de la production de départ. Elle a rendu peu de versions de ses textes. Toutefois, dès sa première production, il y a des traces de segments co-présents qui sont presque toujours bien intégrés dans le nouveau co-texte. Le devoir sur table fait remonter le score des adjectifs non fléchis, ayant la fonction d'attribut du sujet, dans son texte. La tendance à utiliser une grande proportion de tels types d'adjectifs dans des situations où les apprenants disposent de peu d'outils d'aide à la rédaction se confirmera pour presque tous les apprenants.

6.4.1.5. Les adjectifs dans les productions de Camille

Type de production écrite	Adjectifs (qui devraient être) fléchis +fonction épithète	Adjectifs non fléchis + fonction attribut du sujet	Adjectifs non fléchis + fonction adverbe	Nombre de terminaisons manquantes	Nombre de terminaisons superflues	Nombre de marques erronées	Nombre d'erreurs de syntaxe (adjectif)	Nombre d'adjectifs faisant partie d'un segment co-présent	Nombre total de mots
pdd	0	5	0	0	0	0	0	5	187
vgV1	1[29]	1	0	0	0	0	0	1	63
vgV2	1[30]	1	0	0	0	0	0	1	54
clV1	1	2	0	0	0	0	0	1	35
clV2	2	2	0	0	0	0	0	2	74
clV3	1	8	3	0	0	0	0	4	92

Tableau 36 - Adjectifs dans les productions de Camille.

[29] "für *alle die Familie".
[30] "für *alle Familie"

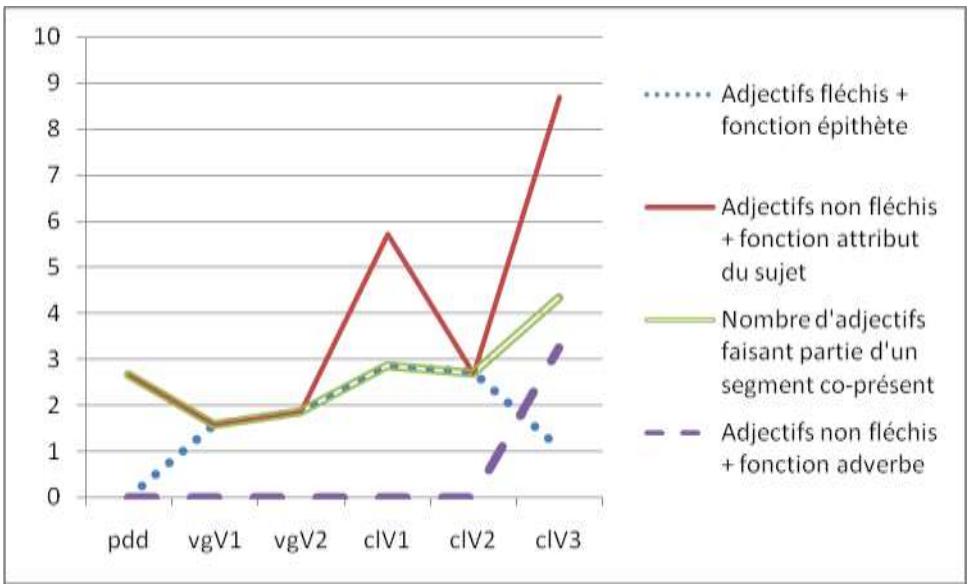


Figure 51 - Adjectifs dans les productions de Camille.

La courbe retraçant le nombre d'adjectifs fléchis monte régulièrement pour redescendre à l'occasion du devoir sur table. La courbe des adjectifs faisant partie d'un segment co-présent monte en parallèle, mais elle continue à monter même lors de la production finale. Comme chez les autres apprenants, le graphique des adjectifs employés par Camille illustre sa tendance à privilégier l'emploi d'adjectifs non fléchis lors d'une situation d'évaluation qui se caractérise par la limitation du temps et par la limitation des outils d'aide à la rédaction. Toutefois, cette tendance est compensée par l'augmentation de segments co-présents contenant un adjectif ayant la fonction d'adverbe. Lors de la dernière production, elle a utilisé trois adjectifs de ce type, dont deux qui font partie d'un segment co-présent et un qui fait partie d'un segment construit par elle, ayant la même structure que les deux autres segments. Camille a su tirer profit des corpus et concordances et complexifier ses textes au fil du projet.

6.4.1.6. Les adjectifs dans la production de Chloé

Type de production écrite	Adjectifs (qui devraient être) fléchis +fonction épithète	Adjectifs non fléchis + fonction attribut du sujet	Adjectifs non fléchis + fonction adverbe	Nombre de terminaisons manquantes	Nombre de terminaisons superflues	Nombre de marques erronées	Nombre d'erreurs de syntaxe (adjectif)	Nombre d'adjectifs faisant partie d'un segment co-présent	Nombre total de mots
pdd	0	10	0	0	0	0	0	3	165
vgV1	1	0	0	0	0	1	0	0	52
vgV2	2	0	0	0	0	0	0	1[31]	63
clV0	2	1	0	0	0	0	0	1	46
clV1	2	3	1	0	0	0	0	1	56
clV2	2	4	1	0	0	0	0	2	63
clV4	2	6	1	0	0	0	0	3	103

Tableau 37 - Adjectifs dans les productions de Chloé.

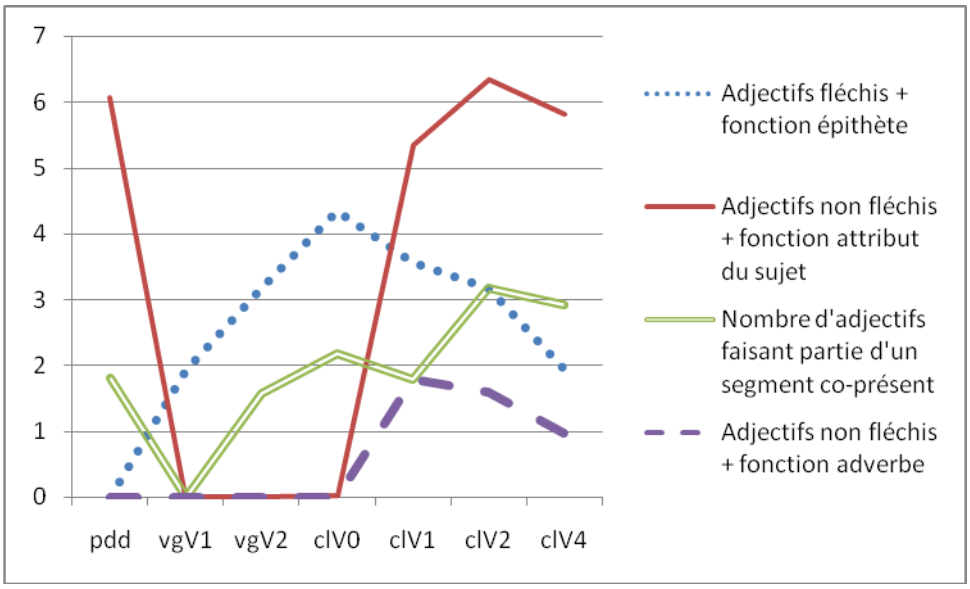


Figure 52 - Adjectifs dans les productions de Chloé.

Chloé s'est mise rapidement à utiliser des corpus et concordances (cf. annexe I.2.1.). Dès la deuxième version de son texte "Visite guidée", j'ai repéré des passages contenant un adjectif qui existent aussi bien dans ses productions que dans le corpus réfléchi, donc des segments co-présents. Chloé a réussi à adapter certains de ces passages au nouveau co-texte, en remplaçant un nom ou une terminaison, par exemple. Comme chez Camille, l'apparition d'adverbes adjectivaux va ici de pair avec une régression du nombre d'adjectifs fléchis. Les adjectifs non fléchis réapparaissent déjà lors de la version V1 du "Commentaire dans un livre d'or", donc plus tôt que chez les autres apprenants, et ils sont également présents dans la

[31] *Es*können eigenes Zeichenmaterial mitgebracht werden*, partiellement co-présent à *Es können eigene Instrumente mitgebracht werden.* [Il est possible d'apporter son propre matériel/ses propres instruments.]

production finale. Toutefois, nous trouvons dans ce dernier texte également des adjectifs fléchis et des adjectifs faisant partie d'un segment co-présent, quoique à proportion moindre par rapport aux productions précédentes.

6.4.1.7. Les adjectifs dans les productions de Christine

Type de production écrite	Adjectifs (qui devraient être) fléchis +fonction épithète	Adjectifs non fléchis + fonction attribut du sujet	Adjectifs non fléchis + fonction adverbe	Nombre de terminaisons manquantes	Nombre de terminaisons superflues	Nombre de marques erronées	Nombre d'erreurs de syntaxe (adjectif)	Nombre d'adjectifs faisant partie d'un segment co-présent	Nombre total de mots
pdd	2	3	1	1	0	0	1	0	98
vgV1	1	2	1	0	0	1	0	2	90
vgV3	1	2	0	0	0	0	0	2	92
clV1	0	4	0	0	0	0	0	0	36
clV2	0	10	3	0	0	0	0	4	90

Tableau 38 - Adjectifs dans les productions de Christine.

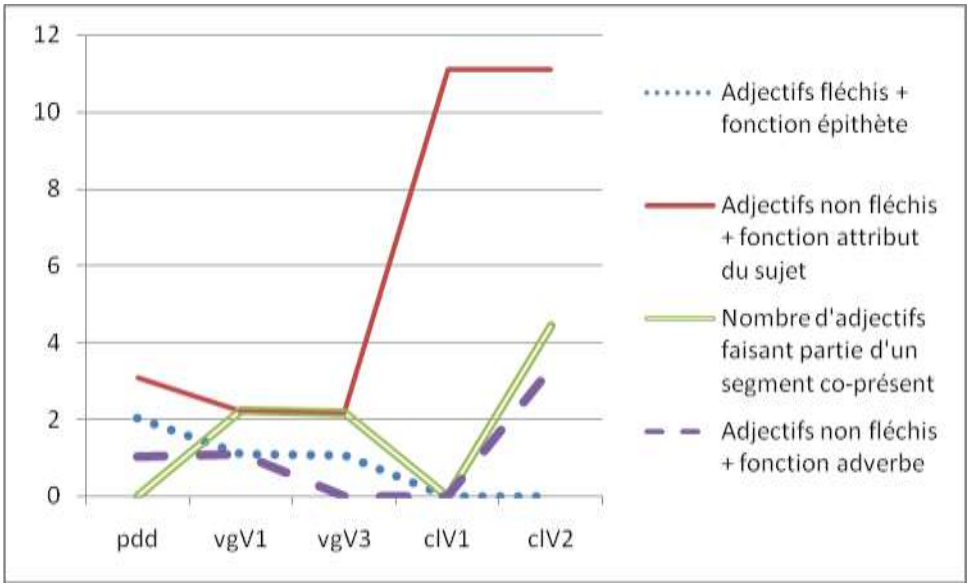


Figure 53 – Adjectifs dans les productions de Christine.

Les productions de Christine contiennent peu de formes fléchies. La courbe des adjectifs faisant partie d'un segment co-présent et celle des adjectifs ayant la fonction d'adverbe se confondent presque lors des deux dernières productions, et elles sont en augmentation. Pour la version V1 du "Commentaire dans un livre d'or", le graphique atteint un "point zéro", à part pour les adjectifs non fléchis, ayant la fonction d'attribut. Une légère complexification peut être observée dans le domaine de l'emploi des adjectifs adverbaux.

6.4.1.8. Les adjectifs dans les productions de Fernand

Type de production écrite	Adjectifs (qui devraient être) fléchis + fonction épithète	Adjectifs non fléchis + fonction attribut du sujet	Adjectifs non fléchis + fonction adverbe	Nombre de terminaisons manquantes	Nombre de terminaisons superflues	Nombre de marques erronées	Nombre d'erreurs de syntaxe (adjectif)	Nombre d'adjectifs faisant partie d'un segment co-présent	Nombre total de mots
pdd	1	10	0	0	0	1	0	0	122
vgV1	0	0	0	0	0	0	0	0	15
vgV2	4	0	1	2	0	0	0	0[32]	32
vgV3	4	0	1	2	0	0	0	0	53
vgV4	4	0	2	1	0	1	0	3	88
vgV5	4	0	1	0	0	1	0	2	82
clV1	4	4	0	1	0	1	0	3	119

Tableau 39 - Adjectifs dans les productions de Fernand.

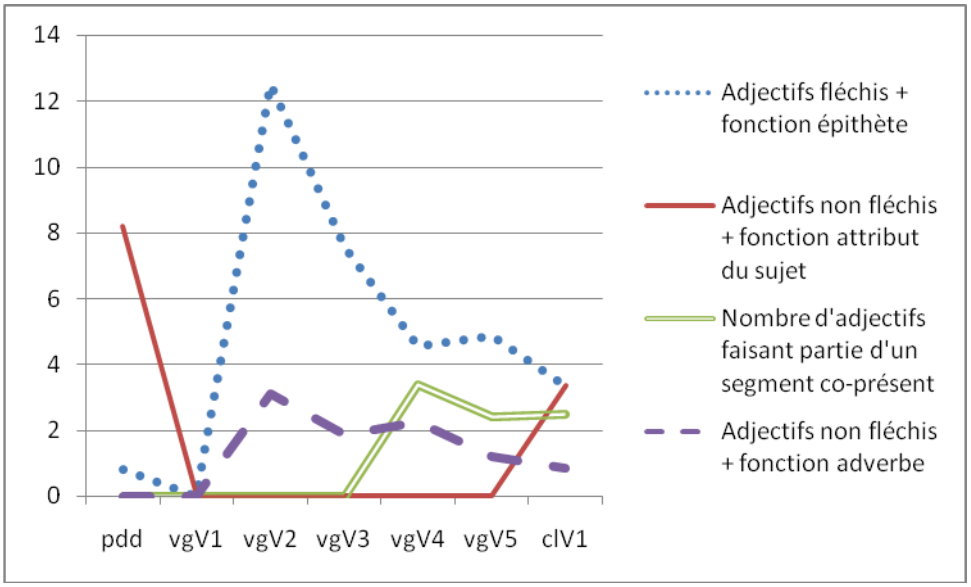


Figure 54 - Adjectifs dans les productions de Fernand.

La courbe des adjectifs (qui devraient être) fléchis présente un pic lors de la version V2 de la "Visite guidée". Ensuite, elle retombe jusqu'à un niveau légèrement supérieur au niveau de départ. Cela s'explique par les faits suivants : le pic d'adjectifs fléchis, qui caractérise la version V2 de la "Visite guidée", concerne un texte qui ne contient que 32 mots. Le chiffre obtenu est donc peu représentatif. Fernand a étoffé son texte à partir de cette version, sans toutefois modifier le nombre d'adjectifs fléchis, qui reste stable jusqu'à la version V5 de sa "Visite guidée". D'ailleurs, dans les versions V2 et V3, deux des quatre adjectifs qui devraient porter une marque n'en portent pas. Dans la version V4, trois sur les quatre adjectifs épithètes

[32] J'ai essayé de faire trouver à F un segment contenant "Generation".

portent une marque. Une des trois formes est erronée. Dans la version V5, les quatre adjectifs épithètes portent une marque, dont une erronée. On peut constater une évolution, toutefois lente et instable, dans le domaine des adjectifs fléchis et des adjectifs adverbaux.

6.4.1.9. Les adjectifs dans les productions de Georges

Type de production écrite	Adjectifs (qui devraient être) fléchis +fonction épithète	Adjectifs non fléchis + fonction attribut du sujet	Adjectifs non fléchis + fonction adverbe	Nombre de terminaisons manquantes	Nombre de terminaisons superflues	Nombre de marques erronées	Nombre d'erreurs de syntaxe (adjectif)	Nombre d'adjectifs faisant partie d'un segment co-présent	Nombre total de mots
pdd	1	7	0	0	0	0	0	0	62
vgV1	1	0	0	1	0	0	0	0	51
vgV3	1	0	0	1	0	0	0	0	81
vgV5	1	0	0	0	0	0	0	0	78
clV1	1	6	0	0	0	0	0	1	84

Tableau 40 - Adjectifs dans les productions de Georges.

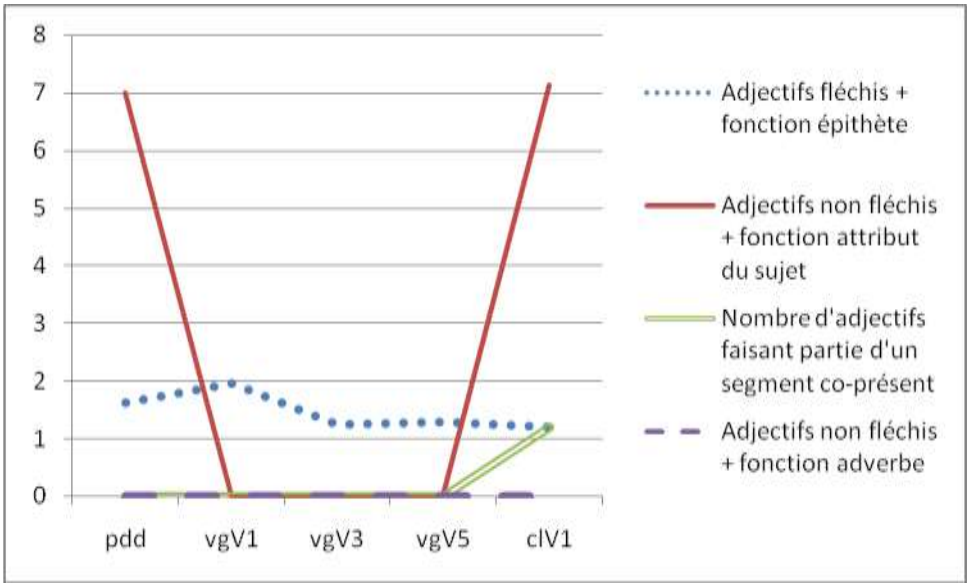


Figure 55 - Adjectifs dans les productions de Georges.

Georges a souvent travaillé à l'aide du traducteur. Ses productions ne comportent aucun adjectif ayant la fonction d'adverbe. Quand il n'a pas eu accès aux outils d'aide à la rédaction électroniques, il a fait un emploi assez important d'adjectifs non fléchis, ayant la fonction d'attribut du sujet. Dans les productions pdd, vgV1, vgV3 et cgV5, Georges a fait un emploi modéré d'adjectifs épithètes qui devraient être fléchis. Il ne les a pas pourvus de marque. La dernière production, le devoir sur table, contient un adjectif fléchi, portant une marque exacte. Une évolution d'utilisation conforme des adjectifs semble s'amorcer dans ses travaux.

6.4.1.10. Les adjectifs dans les productions de Géraldine

Type de production écrite	Adjectifs (qui devraient être) fléchis +fonction épithète	Adjectifs non fléchis + fonction attribut du sujet	Adjectifs non fléchis + fonction adverbe	Nombre de terminaisons manquantes	Nombre de terminaisons superflues	Nombre de marques erronées	Nombre d'erreurs de syntaxe (adjectif)	Nombre d'adjectifs faisant partie d'un segment co-présent	Nombre total de mots
pdd	4	22	1	2	4	2	0	4	166
vgV1	1	1	0	1	0	0	0	0	90
vgV2	3	1	0	1	0	0	0	0	95
vgV3	3	1	0	1	0	0	0	0	92
clV1	6	3	0	0	0	1	0	2	95
clV2	5	3	0	0	0	0	0	2	95
clV3	5	6	1	1	0	1	0	2[33]	119

Tableau 41 - Adjectifs dans les productions de Géraldine.

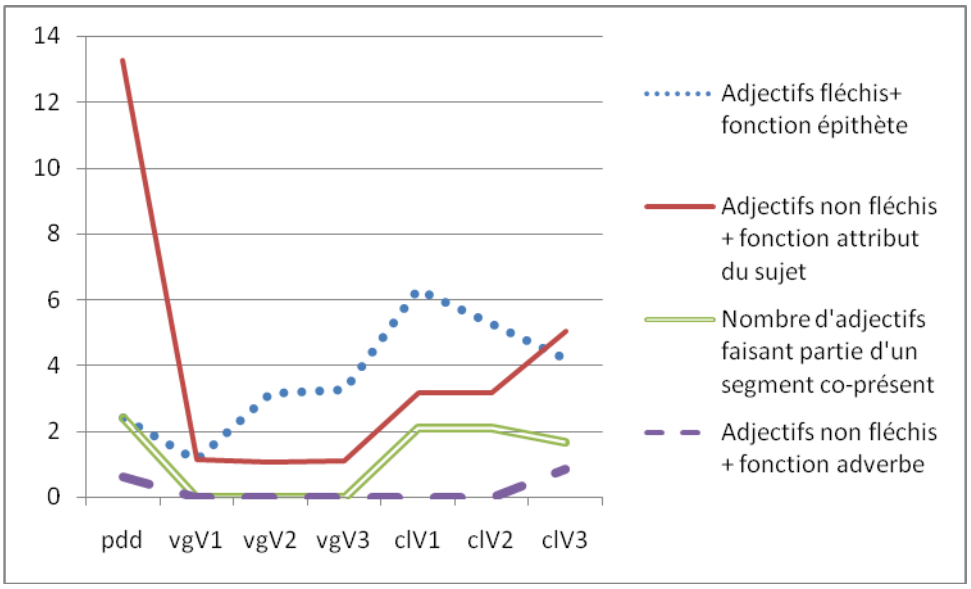


Figure 56 – Adjectifs dans les productions de Géraldine.

Dans sa production de départ, Géraldine a utilisé un nombre important d'adjectifs non fléchis, ayant la fonction d'attribut du sujet (13,25 pour 100 mots). Au cours de *Projekt Prospekt*, elle a considérablement réduit l'emploi de telles formes, au profit d'une augmentation constante de formes adjectivales fléchies. Cette augmentation est accompagnée par une courbe montante, légèrement inférieure, désignant l'existence d'adjectifs faisant partie d'un segment co-présent. Lors de la deuxième version de son commentaire dans le livre d'or, le score des adjectifs fléchis chute légèrement. La chute continue dans la production finale, le devoir sur table. Une courbe quasi parallèle à celle des adjectifs fléchis illustre la chute des adjectifs faisant partie

[33] Il y a en tout trois segments contenant un adjectif, mais *bis nächste Woche* [à la semaine prochaine] et *dankeschön* [merci bien] n'apparaissent pas dans le corpus.

d'un segment co-présent. Dans la dernière production, Géraldine fait de nouveau un emploi assez important d'adjectifs non fléchis (5,04 pour 100 mots). Par contre, pour la première fois, elle utilise des adjectifs non fléchis, ayant la fonction d'adverbe. Globalement, il est donc possible d'avancer que ses productions se sont complexifiées au fil de *Projekt Prospekt*.

6.4.1.11. Les adjectifs dans les productions de Janine

Type de production écrite	Adjectifs (qui devraient être) fléchis +fonction épithète	Adjectifs non fléchis + fonction attribut du sujet	Adjectifs non fléchis + fonction adverbe	Nombre de terminaisons manquantes	Nombre de terminaisons superflues	Nombre de marques erronées	Nombre d'erreurs de syntaxe (adjectif)	Nombre d'adjectifs faisant partie d'un segment co-présent	Nombre total de mots
pdd	0	5	0	0	0	0	0	5	96
vgV3	2	0	0	1	0	0	1	0	33
vgV7	2	1	0	0	0	0	0	0	73
clV0	0	0	0	0	0	0	0	0	24
clV1	1	8	0	0	0	0	0	3	100

Tableau 42 - Adjectifs dans les productions de Janine.

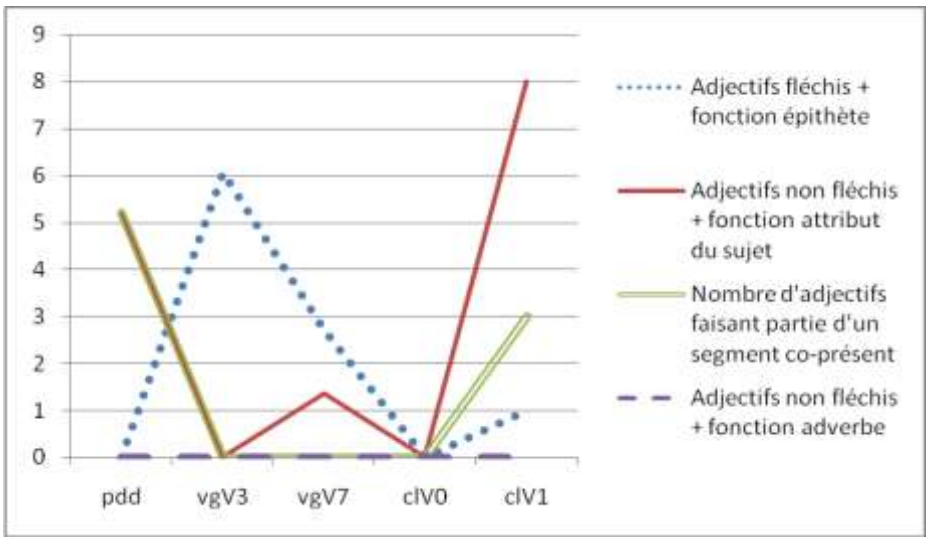


Figure 57 - Adjectifs dans les productions de Janine.

Dans la production de départ de Janine, les courbes des adjectifs non fléchis et des adjectifs faisant partie d'un segment co-présent se confondent. Il convient de relativiser cela en précisant que le seul adjectif utilisé par elle dans cette première production est le mot *alt* [vieux], employé cinq fois dans une expression pour dire l'âge. Le pic qu'affiche le graphique pour la production "Visite guidée, V3" pour l'utilisation de formes (qui devraient être) fléchies est également à lire avec prudence. En effet, dans ce texte très court, d'une longueur de 33 mots, se trouvent deux adjectifs épithètes. L'un est fléchi, mais il se trouve après le nom qu'il

spécifie, ce qui est une erreur de syntaxe en allemand. La toute dernière production contient un adjectif fléchi qui fait partie d'un segment co-présent. Il paraît possible de parler d'une légère évolution au niveau de la microsyntaxe grâce à la prise en compte d'éléments macrosyntaxiques. Toutefois, dans les textes que j'ai observés, Janine n'a pas utilisé d'adjectif ayant la fonction d'adverbe, et les adjectifs ayant la fonction d'attribut du sujet restent très présents. Ces deux caractéristiques sont des signes d'un style d'écriture peu complexe.

6.4.1.12. Les adjectifs dans les productions de Jean

Type de production écrite	Adjectifs (qui devraient être) fléchis +fonction épithète	Adjectifs non fléchis + fonction attribut du sujet	Adjectifs non fléchis + fonction adverbe	Nombre de terminaisons manquantes	Nombre de terminaisons superflues	Nombre de marques erronées	Nombre d'erreurs de syntaxe (adjectif)	Nombre d'adjectifs faisant partie d'un segment co-présent	Nombre total de mots
pdd	1	1	0	0	0	0	0	3	88
vgV1	0	0	0	0	0	0	0	0	24
clV0	0	2	0	0	0	0	0	0	24
clV1	2	7	0	2	1	0	0	2[34]	89

Tableau 43 - Adjectifs dans les productions de Jean.

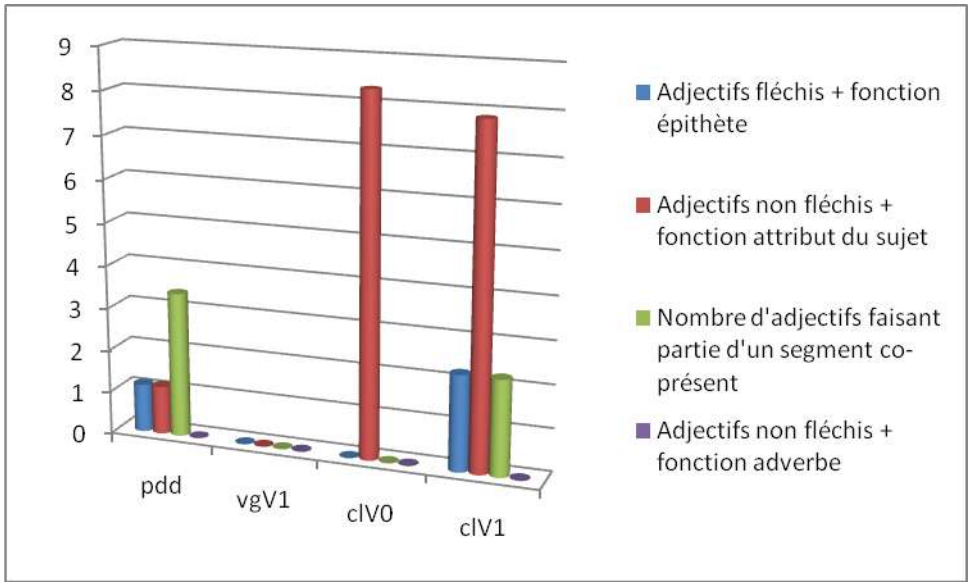


Figure 58 - Adjectifs dans les productions de Jean (1).

Ce premier graphique concernant les adjectifs utilisés par Jean illustre le parallélisme entre le nombre des adjectifs (qui devraient être) fléchis et les adjectifs non fléchis, attribut du sujet, pour la première production. Je constate en plus un parallélisme entre le nombre des adjectifs fléchis et les adjectifs faisant partie d'un segment co-présent lors de la dernière production.

[34] dont un segment incomplète : **besten* au lieu de *am besten*.

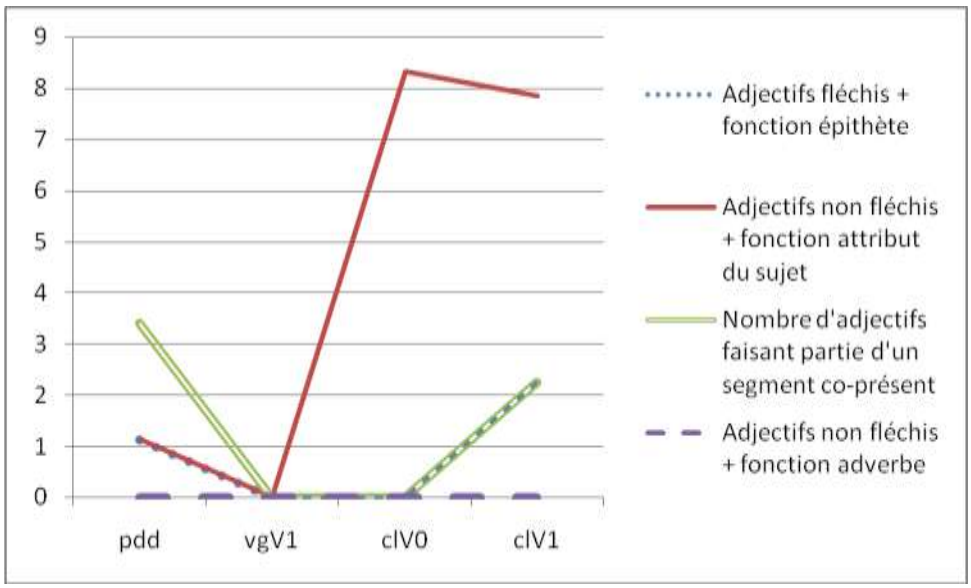


Figure 59 - Adjectifs dans les productions de Jean (2).

Selon ce deuxième graphique, dans les deux productions intermédiaires, Jean a uniquement utilisé des adjectifs ayant la fonction d'adverbe. Ses deux versions du texte "Commentaire dans un livre d'or" contiennent 8,33 adjectifs non fléchis (pour 100 mots). Dans sa toute dernière production, le devoir sur table, il y a presque autant d'adjectifs non fléchis (7,86 pour 100 mots), mais des adjectifs ayant la fonction d'épithète apparaissent, en parallèle à la présence de segments co-présents contenant des adjectifs. Dans cette dernière production, les adjectives épithètes ne comportent toutefois pas de marque, malgré le fait que Jean a eu accès à une petite concordance dans laquelle se trouvent des segments co-présents, contenant un adjectif qui porte une marque. Le nombre des adjectifs attribut du sujet reste élevé. On ne peut pas réellement parler d'une complexification des productions, mais une légère évolution vers l'emploi d'adjectifs épithètes, parallèle à celle de l'emploi de segments co-présents, peut être constatée.

6.4.1.13. Les adjectifs dans les productions de Léa

Type de production écrite	Adjectifs (qui devraient être) fléchis +fonction épithète	Adjectifs non fléchis + fonction attribut du sujet	Adjectifs non fléchis + fonction adverbe	Nombre de terminaisons manquantes	Nombre de terminaisons superflues	Nombre de marques erronées	Nombre d'erreurs de syntaxe (adjectif)	Nombre d'adjectifs faisant partie d'un segment co-présent	Nombre total de mots
pdd	0	4	0	0	0	0	0	0	78
vgV1[35]	6	1	0	0	0	1	0	1	97
vgV2	2	2	0	1	0	0	0	1	82
clV0	3	2	0	0	0	0	0	0[36]	54
clV1	4	0	0	0	0	0	0	1	77
clV2	2	3	2	2	0	0	0	1	84

Tableau 44 - Adjectifs dans les productions de Léa.

Le style des textes vgV1, vgV2, clV0 et clV1 de Léa permet d'envisager qu'elle n'est pas l'unique auteur de ces productions. Je complèterai donc le graphique (1) par deux autres graphiques, afin de pouvoir interpréter les caractéristiques d'apparence contradictoire du premier graphique.

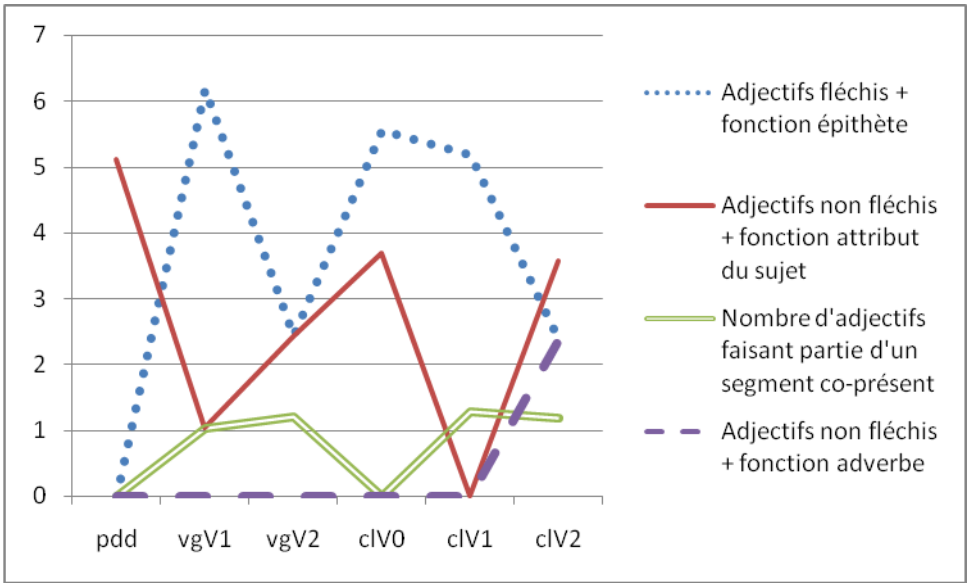


Figure 60 - Adjectifs dans les productions de Léa (1).

Le premier graphique paraît peu cohérent. Il comporte plusieurs pics, et aucune évolution ne se dessine clairement, à part celle concernant l'emploi des adjectifs ayant la fonction d'adverbe qui croît lors de la dernière production. Si je dissocie ce dont je sais que Léa l'a produit et ce qu'un autre auteur aurait pu écrire, j'obtiens le deuxième graphique qui pourrait décrire la

[35] Je ne suis pas sure que ce soit Léa qui ait rédigé ce texte.
[36] Il y a un segment : *nächstes Mal* [la prochaine fois], mais elle n'existe pas dans le corpus réfléchi. Aussi ici, j'ai un doute concernant l'auteur.

façon dont rédige un apprenant avancé et le troisième graphique qui peut correspondre à l'évolution d'un apprenant peu avancé.

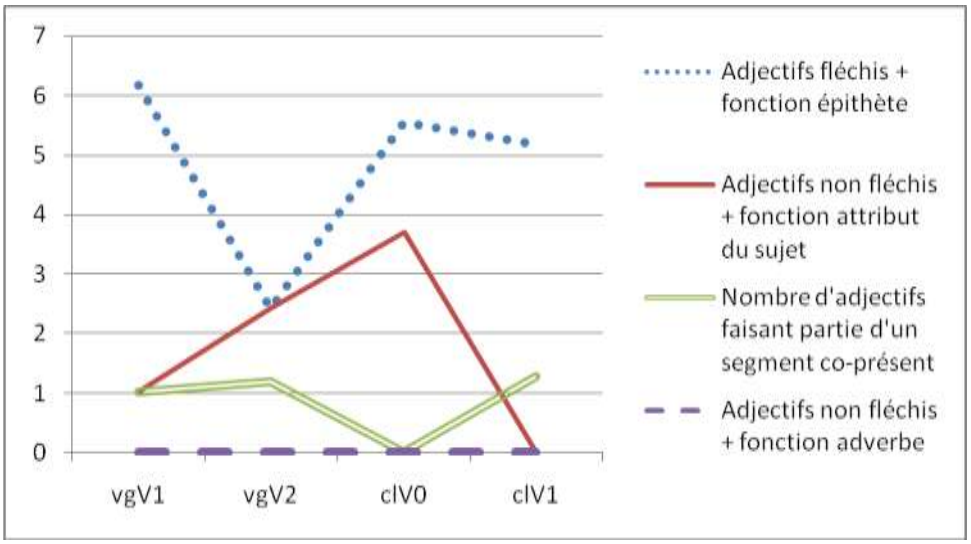


Figure 61 - Adjectifs dans les productions de Léa (2).

Le graphique ci-dessus montre, en alternance, deux situations distinctes. La première et la dernière production de ce graphique se caractérisent par une utilisation relativement fréquente et maîtrisée des adjectifs fléchis et une utilisation rare d'adjectifs non fléchis. Entre les deux, il y a des productions qui se caractérisent par la présence de moins d'adjectifs fléchis et davantage d'adjectifs non fléchis. Cela pourrait s'expliquer par le fait que Léa n'a peut-être pas toujours travaillé seule à la maison.

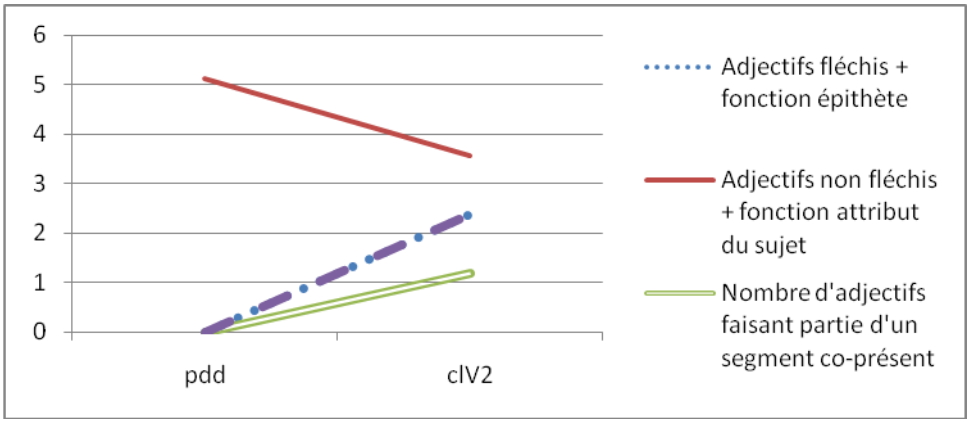


Figure 62 - Adjectifs dans les productions de Léa (3).

Le troisième graphique montre un mouvement qui part d'une sur-utilisation d'adjectifs non fléchis et d'une sous-utilisation des trois autres types d'adjectifs observés pour elle, vers un état contraire, donc un emploi plus équilibré entre adjectifs fléchis et non fléchis, évolution secondée par la présence d'un adjectif faisant partie d'un segment co-présent. Les deux adjectifs épithètes dans son dernier texte ne portent pas de marque. L'emploi de deux adjectifs

adverbiaux, inexistants dans les productions précédentes, est un signe de complexification des textes.

6.4.1.14. Les adjectifs dans les productions d'Olivia

Type de production écrite	Adjectifs (qui devraient être) fléchis +fonction épithète	Adjectifs non fléchis + fonction attribut du sujet	Adjectifs non fléchis + fonction adverbe	Nombre de terminaisons manquantes	Nombre de terminaisons superflues	Nombre de marques erronées	Nombre d'erreurs de syntaxe (adjectif)	Nombre d'adjectifs faisant partie d'un segment co-présent	Nombre total de mots
pdd	8	9	0	5	0	1	0	0	103
vgV1	0	0	0	0	0	0	0	0	57
vgV3	1	0	0	0	0	1	0	0	54
vgV4	1	0	0	0	0	0	0	0	48
vgV5	1	0	0	0	0	0	0	0	125
clV1	4	14	0	1	0	0	0	2	166

Tableau 45 - Adjectifs dans les productions d'Olivia.

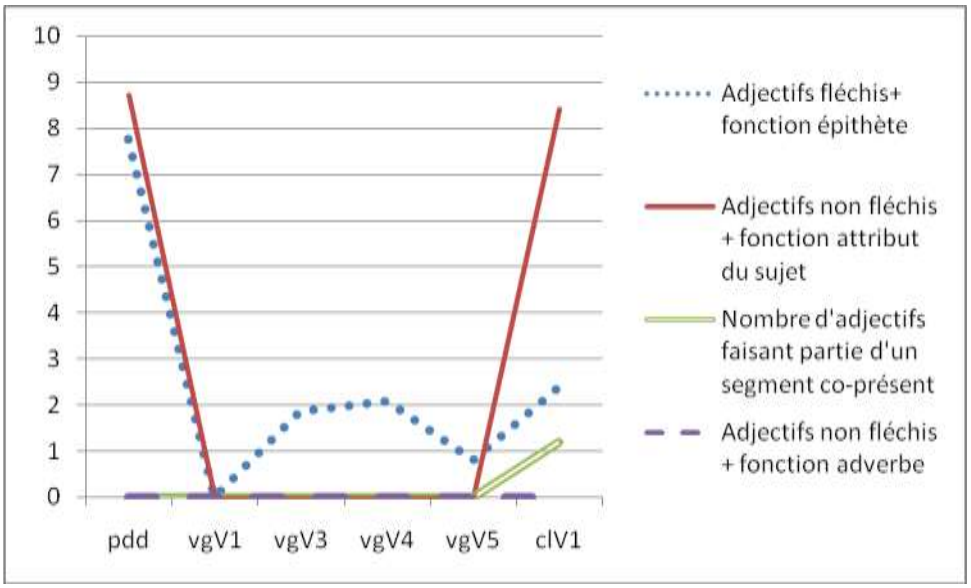


Figure 63 - Adjectifs dans les productions d'Olivia.

Les différentes versions de la "Visite guidée" d'Olivia contiennent peu d'adjectifs. Le nombre d'adjectifs, attribut du sujet (d'ailleurs employés de façon répétitive), augmente dans la dernière production, pour atteindre presque le niveau initial. Les textes ne contiennent pas d'adjectif adverbial. La courbe des segments co-présents suggère que, dans le domaine des adjectifs, Olivia ne s'est apparemment pas appuyée sur les corpus et concordances, à part pour la production finale. Dans cette production, elle a utilisé deux segments contenant un adjectif fléchi, portant la marque adéquate.

6.4.1.15. Les adjectifs dans les productions de Simon

Type de production écrite	Adjectifs (qui devraient être) fléchis +fonction épithète	Adjectifs non fléchis + fonction attribut du sujet	Adjectifs non fléchis + fonction adverbe	Nombre de terminaisons manquantes	Nombre de terminaisons superflues	Nombre de marques erronées	Nombre d'erreurs de syntaxe (adjectif)	Nombre d'adjectifs faisant partie d'un segment co-présent	Nombre total de mots
pdd	0	4	1	0	0	0	0	0	70
vgV1	0	0	0	0	0	0	0	0	32
vgV3	1	2	0	0	0	0	0	2	93
clV0	0	3	1	0	0	0	0	0	28
clV1	2	5	2	0	0	0	0	2	99

Tableau 46 - Adjectifs dans les productions de Simon.

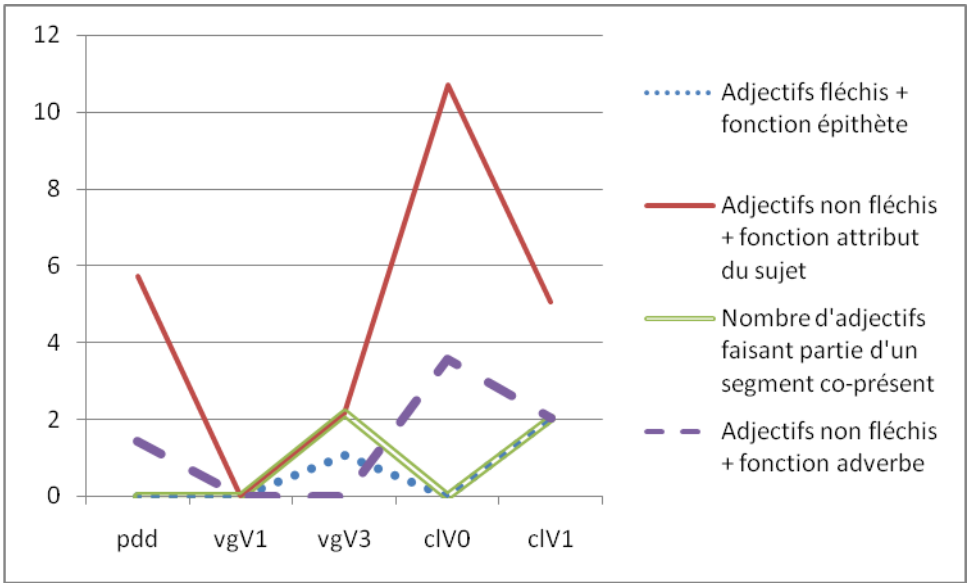


Figure 64 - Adjectifs dans les productions de Simon.

Chez Simon, l'évolution de la courbe des adjectifs fléchis est assez similaire à celle des adjectifs faisant partie d'un segment co-présent. Pour les deux dernières productions, les deux courbes se confondent. Le score des adjectifs non fléchis, ayant la fonction d'attribut du sujet, chute lors de la dernière production, pour atteindre toutefois le niveau de la production de départ. Le pic signalant les adjectifs non fléchis, attribut du sujet, concerne une production courte de 28 mots. Je le considère donc comme peu significatif. La courbe retraçant l'emploi des adjectifs, ayant la fonction d'adverbe, a une forme de zigzag montant, avec une chute lors de la dernière production. Une évolution, toutefois modeste, vers une complexification du style d'écriture peut être constatée.

6.4.1.16. Les adjectifs dans les productions de Vicky

Type de production écrite	Adjectifs (qui devraient être) fléchis + fonction épithète	Adjectifs non fléchis + fonction attribut du sujet	Adjectifs non fléchis + fonction adverbe	Nombre de terminaisons manquantes	Nombre de terminaisons superflues	Nombre de marques erronées	Nombre d'erreurs de syntaxe (adjectif)	Nombre d'adjectifs faisant partie d'un segment co-présent	Nombre total de mots
pdd	3	8	1	0	0	1	0	0	136
vgV1	5	0	0	0	0	0	0	1	84
vgV3	6	0	0	0	0	0	0	1	81
clV0[37]	7	0	0	1	0	2	0	5	136
clV1[38]	8	6	2	1	0	1	0	5	135

Tableau 47 - Adjectifs dans les productions de Vicky.

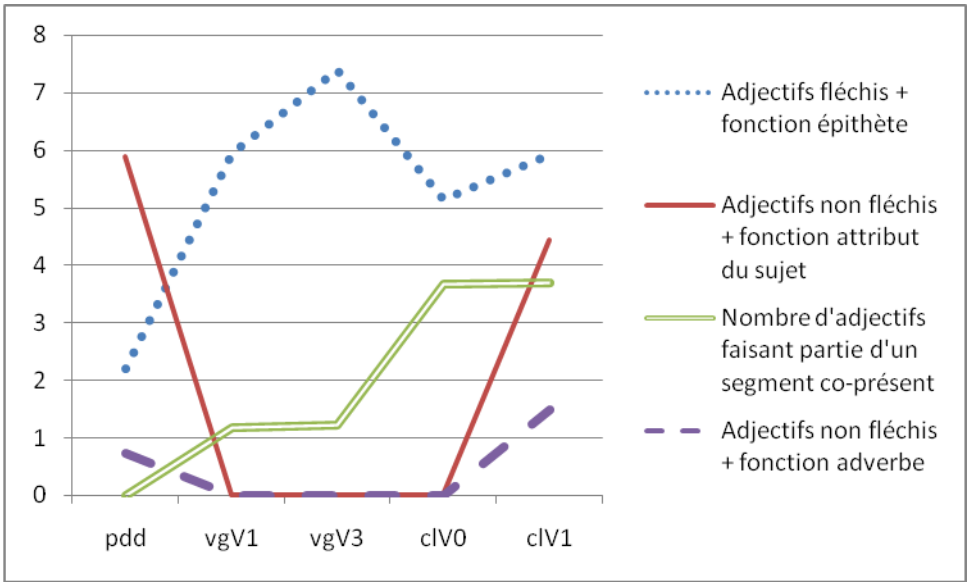


Figure 65 - Adjectifs dans les productions de Vicky.

Entre la production de départ et la production finale, Vicky a évolué au niveau de l'emploi des adjectifs : le nombre de formes fléchies qu'elle emploie augmente de 3 à 8, et le nombre de formes non fléchies, attributs du sujet, chute de 8 à 0, pour remonter à 5 lors de la dernière production, donc le devoir sur table. Le nombre d'adjectifs faisant partie de segments co-présents augmente de 0 à 5, et le nombre des adjectifs ayant la fonction d'adverbe, au point zéro pendant les trois productions intermédiaires, monte lors de la dernière production pour dépasser le niveau initial. Vicky a complété deux segments co-présents supplémentaires par un adjectif fléchi, portant la marque exacte. Quand Vicky a rédigé dans une situation autre

[37] Rédigé pendant les séances de cours et à la maison
[38] Devoir sur table et en même temps la dernière version du "Commentaire dans un livre d'or".

que celle d'un devoir sur table, elle a produit des textes qui ne comportent aucun adjectif non fléchi. Dans sa production vgV3, j'ai repéré six adjectifs fléchis, dont un qui fait partie d'un segment co-présent. Les deux productions suivantes illustrent que le nombre d'adjectifs fléchis augmente, aussi bien que le nombre d'adjectifs faisant partie d'un segment co-présent. La production finale, faite en temps limité, confirme cette tendance. De plus, Vicky y utilise des adjectifs adverbiaux. On peut parler d'une nette évolution vers une complexification du style d'écriture.

6.4.2. Les adjectifs dans les productions du groupe classe

Les études individuelles des adjectifs dans les productions de type "Visite guidée" et "Commentaire dans un livre d'or" permettent d'avancer que pour quatre apprenants sur treize (Camille, Chloé, Géraldine et Vicky), l'utilisation de corpus et concordances comme outil d'aide à la rédaction mène clairement à une complexification de leur style d'écriture. Pour huit autres apprenants (Alexa, Christine, Fernand, Georges, Janine, Jean, Léa, Olivia et Simon), l'évolution de la complexification est plus modeste, mais néanmoins visible.

Dans ce qui suit, je regroupe les résultats obtenus pour le groupe classe entier. N'y figurent pas les productions ni de Louise, pour les raisons déjà évoquées, ni d'Alexa, Corentin, Perrine et Véronique, qui n'ont pas produit soit la production de départ, soit la production finale.

6.4.2.1. Observations générales

Le tableau ci-dessous retrace l'évolution de 12 apprenants sur 17 dans le domaine des adjectifs. J'ai choisi un échantillon de quatre textes : la production de départ, la première et la dernière version du texte "Visite guidée" et la production finale, qui est en même temps la dernière – et pour certains la première – version du texte "Commentaire dans un livre d'or". Les chiffres se réfèrent à des textes contenant 100 mots. Selon ce tableau, en matière d'emploi d'adjectifs, les moyennes des apprenants évoluent, sans s'en approcher de très près, vers les moyennes des scripteurs experts que j'ai observées dans un échantillon de 875 mots du corpus de référence "Visite guidée" (cf. 6.4.). Les adjectifs attributs sont particulièrement nombreux dans la première et de la dernière production.

Type de production écrite	Adjectifs (qui devraient être) fléchis + fonction épithète	Adjectifs non fléchis + fonction attribut du sujet	Adjectifs non fléchis + fonction adverbe	Nombre d'adjectifs faisant partie d'un segment co-présent
pdd	1,49	6	0,31	1,29
vg 1ère version	2,15	0,49	0,08	0,5
vg Version finale	2,53	0,92	0,46	0,95
cl Version finale	2,31	6,64	0,98	1,33
		3,51	**0,45**	
Moyenne	**2,12**	**3,96**		**1,01**
Corpus de référence "Visite guidée"	**5,71**	**1,6** (toutes formes non fléchies confondues)		

Tableau 48 - Adjectifs dans les productions du groupe classe (1).

Le tableau suivant ne présente que les productions intermédiaires, donc celles pour lesquelles les apprenants avaient libre accès à tous types d'outils d'aide à la rédaction, dont des corpus et concordances de taille assez "conséquente". Les moyennes obtenues s'approchent, davantage que dans le tableau précédent, des moyennes du corpus de référence.

Type de production écrite	Adjectifs (qui devraient être) fléchis + fonction épithète	Adjectifs non fléchis + fonction attribut du sujet	Adjectifs non fléchis + fonction adverbe	Nombre d'adjectifs faisant partie d'un segment co-présent
vg 1ère version	2,15	0,49	0,08	0,5
vg Version finale	2,53	0,92	0,46	0,95
Moyenne	**2,34**	**0,48**		**0,72**
CR "Visite guidée"	**5,71**	**1,6**		

Tableau 49 - Adjectifs dans les productions du groupe classe (2).

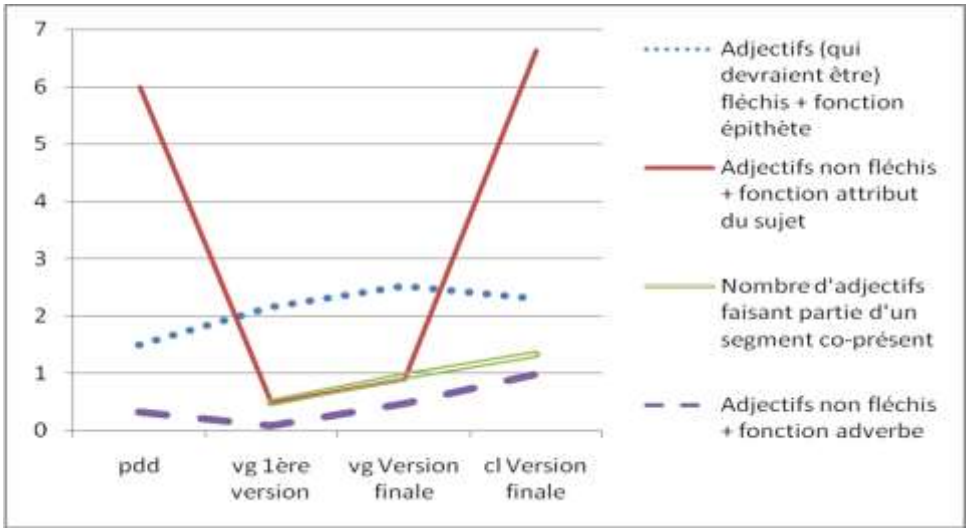


Figure 66 – Évolution des adjectifs lors de *Projekt Prospekt*.

Le graphique ci-dessus affiche le pic d'adjectifs attribut du sujet, non fléchis, caractérisant la plupart des productions qui ont été soumises à une restriction en termes de temps et d'aides à la rédaction. *Projekt Prospekt* a mené à une légère augmentation des formes adjectivales fléchies. La courbe des segments co-présents, contenant des adjectifs, monte de façon parallèle à la celle des formes fléchies, et, de façon encore plus nette, en parallèle à celle des adjectifs ayant la fonction d'adverbe. La courbe des adjectifs, faisant partie d'un segment co-présent, et celle des adjectifs-adverbes continue à évoluer même lors des productions finales. En ce qui concerne le devoir sur table, donc dans une situation où les apprenants avaient accès à un seul texte du corpus et à 187 lignes de concordance, le score des adjectifs fléchis chute légèrement.

Dans la tranche contenue entre la première et la dernière version des "Visites guidées" se dessine l'évolution-type de présence de formes fléchies, non fléchies et de segments co-présents contenant un adjectif. Le tracé des courbes concernant les formes fléchies et les segments co-présents et le tracé concernant l'évolution de l'emploi des adjectifs, ayant la fonction d'adverbe, et les résultats individuels sont des éléments qui permettent de dire que, dans certaines conditions, l'emploi des outils d'aide à la rédaction, proposés lors de *Projekt Prospekt*, mène à une complexification de la production écrite.

6.4.2.2. Segments répétés, segments uniques

Une analyse tenant comparant les formes et segments co-présents répétés plusieurs fois à ceux qui n'apparaissent qu'une seule fois dans un texte permet d'affiner le résultat positif esquissé ci-dessus et de confirmer le potentiel des concordances pour une complexification de la production écrite. Le tableau suivant présente des chiffres pour la production de départ. L'indication (R) pointe les zones de répétition d'un même adjectif. Cette production a été réalisée après quelques séances de préparation au lexique de la famille et du quotidien. Les apprenants n'avaient accès à aucun outil d'aide à la rédaction. Olivia et Véronique ont été convoquées à une épreuve d'anglais pendant l'heure du devoir sur table en allemand. Elles ont donc rédigé leur production de départ à la maison ce qui explique pourquoi je ne puis dire avec certitude si elles ont rédigé sans aide.

Prénoms	Adjectifs fléchis	Adjectifs non fléchis	Adjectifs faisant partie de segments co-présents	Nombre de mots du texte
Fernand	1	10	0	122
Camille	0	5	5 (R)	187
Chloé	0	10	3 (R)	165
Christine	2	3	0	98
Georges	1	7	0	62
Géraldine	4	22	4 (R)	166
Janine	0	5 (R)	5 (R)	96
Jean	1	2	3 (R)	88
Léa	0	4	0	78
Olivia	8	9	0	103
Perrine	0	10 (R)	0	94
Simon	0	4	0	70
Véronique	0	13 (R)	2 (R)	158
Vicky	3	8	0	136

Tableau 50 – Adjectifs utilisés de façon répétitive dans les productions de départ.

Dans les productions de départ, sept apprenants ont répété plusieurs fois un même adjectif. J'ai détecté dans ces productions de rares segments de type collocatif, contenant un adjectif. Il s'agit principalement de la collocation qui désigne l'âge de quelqu'un : *ist … Jahre alt* [a … ans]. Les autres passages contenant des adjectifs ne sont pas très cohésifs. Certains éléments de segments ont une certaine tendance à apparaître fréquemment ensemble : Jean écrit "*Ich habe schlechte *note*" [J'ai de mauvaises notes] ; chez Christine, j'ai trouvé "*spielt gut Klavier*" [joue bien du piano], et Olivia utilise le segment *hat *schwarz Haare* [a des cheveux noirs].

Dans les productions finales des apprenants, les nombres de segments co-présents uniques, contenant un adjectif, sont les suivants : Camille : 4 ; Chloé : 3 ; Christine : 4 ; Fernand : 1 ; Georges : 1 ; Géraldine : 3 ; Janine : 3 ; Jean : 2 ; Léa : 1 ; Olivia : 4 ; Simon : 2 ; Vicky : 5. Cela fait en moyenne 2,75 segments. Camille a utilisé trois fois le même adverbe adjectival *total*, toutefois en variant l'adjectif attribut spécifié par *total*. Les autres apprenants n'ont répété aucun de ces segments co-présents. Cette observation est un élément supplémentaire qui plaide en faveur de l'idée que l'emploi de concordances lors d'un projet d'écriture en LE peut être avantageux pour des apprenants de niveau A2/B1.

6.4.2.3. Évolution des types d'erreurs dans le domaine de l'adjectif

J'ai recensé les moyennes des nombres d'erreurs du groupe classe lors de *Projekt Prospekt* dans le domaine de l'adjectif. Je définis une erreur comme une forme qui ne correspond pas aux normes attestées pour un genre discursif déterminé. Les chiffres concernant les textes de type "Visite guidée" équivalent à la somme des chiffres pour chaque version, divisée par le nombre des versions produites par les apprenants. J'ai opté pour ce procédé en raison du nombre variable de versions produites par les apprenants. J'ai fait de même pour les textes de type "Commentaire dans un livre d'or", en comptabilisant à part la toute dernière version.

Type de production écrite	**Nombre de terminaisons manquantes**	**Nombre de terminaisons superflues**	**Nombre de marques erronées**	**Nombre d'erreurs de syntaxe**
pdd	0,58 (0,2 sans Olivia)	0	0,46	0,08
vg (toutes versions confondues)	0,65	0	0,29	0,12
cl (toutes versions, sauf version finale)	0	0	0,16	0
cl (version finale)	0,94	0,28	0,27	0,09

Tableau 51 – Erreurs du groupe classe dans le domaine de l'adjectif.

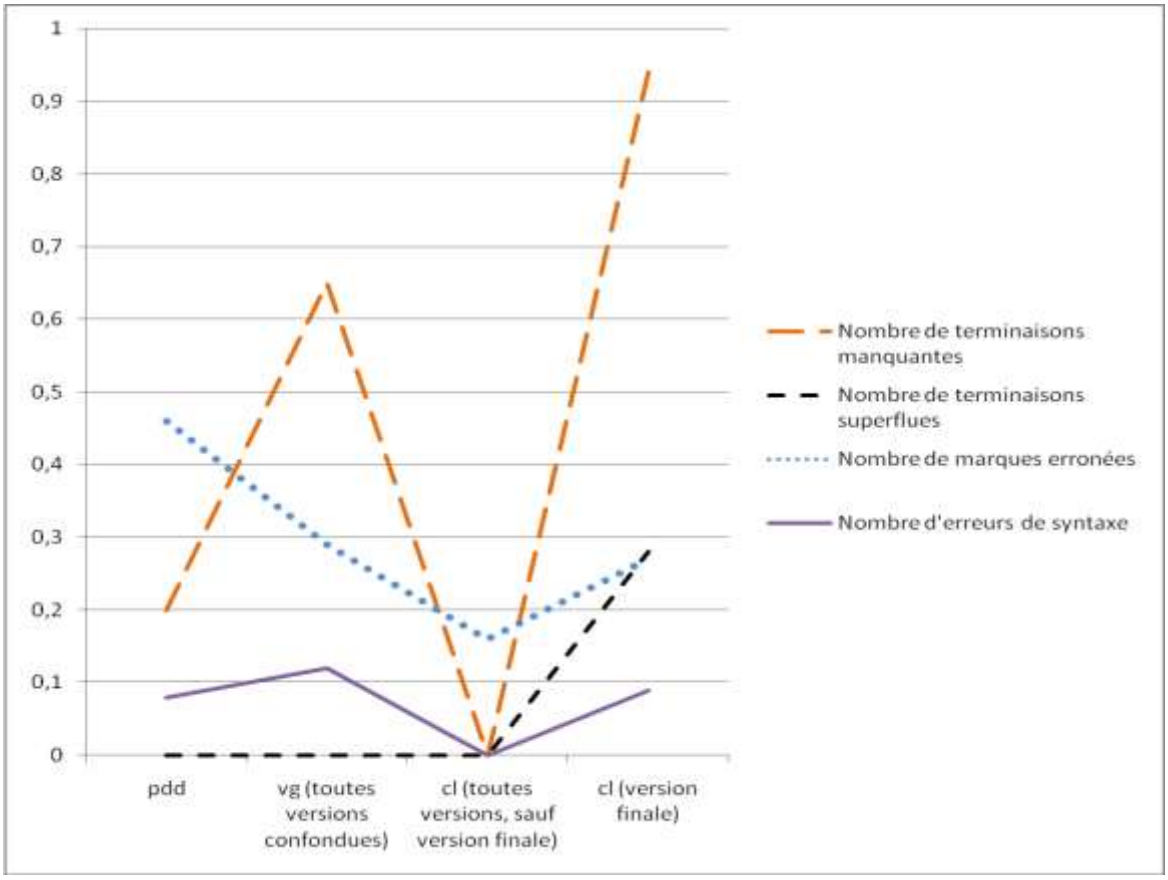


Figure 67 – Évolution des types d'erreurs dans le domaine de l'adjectif (groupe classe).

Le graphique ci-dessus confirme mes observations précédentes concernant l'emploi intensif d'adjectifs sans marque, surtout dans des situations d'évaluation pendant lesquelles le temps et les aides à la rédaction sont limités. Le troisième type de production, présenté dans le graphique, ne comporte ni erreur de syntaxe, ni terminaison superflue. Il n'y a aucune marque absente. Le taux de marques erronées est faible. Il chute d'ailleurs régulièrement, pour remonter lors de la dernière production, le devoir sur table. C'est seulement lors de la production finale que des marques apparaissent qui sont apposées à des adjectifs qui ne devraient pas en porter une. Je considère cela comme l'émergence de conceptualisation dans le domaine du marquage de l'adjectif qui mène, peut-être temporairement, à un phénomène de sur-généralisation des références. Ce résultat semble tout à fait encourageant : Diehl *et al.* (2000 : 214) ont observé que la flexion de l'adjectif est maîtrisé par de rares scripteurs à partir de la classe de troisième/seconde (appelées en Suisse 9ème et 10ème classe).

> *Wer also kann die Adjektive tatsächlich richtig flektieren? Die Muttersprachlerinnen und Muttersprachler – und wenige sehr fortgeschrittene L2-Erwerbende, in der Regel aber offensichtlich erst nach der Schulzeit, und zwar dann, wenn sie sich weiter um die deutsche Sprache bemühen. Die Beherrschung der Flexion komplexer Nominalgruppen – tatsächlich ein unermesslicher Erfolg.*
>
> [Qui sait alors fléchir les adjectifs correctement ? Les locutrices et locuteurs natifs – et quelques très rares apprenants de L2, mais en général apparemment seulement après le temps de scolarité obligatoire, à condition qu'ils continuent à entretenir la langue allemande. Arriver à dominer la flexion de groupes nominaux complexes – effectivement un succès immense.]

À la lumière de cette citation, les résultats obtenus lors de la séance collaborative de révision avec Juliette, Marinette et Virgile sont très encourageants (cf. 6.3.2.3.).

Tout au long de *Projekt Prospekt*, le nombre d'erreurs de syntaxe adjectivale est négligeable. L'observation d'Alexa que les concordances sont une aide dans le domaine de l'ordre des mots (cf. annexe I.2.1.) semble être confirmée par les chiffres. Le paragraphe suivant sera consacré à l'étude des caractéristiques des productions écrites dans le domaine de l'ordre des mots, afin de voir dans quelle mesure l'utilisation des concordances va de pair avec l'obtention de production écrites conformes du point de vue de la syntaxe.

6.4.3. Concordances et évolution de la syntaxe

Diehl *et al.* (2000 : 322) ont observé que la syntaxe et les prépositions s'acquièrent davantage à l'aide de la mémorisation de *chunks*, ou, comme je dirais, de segments, qu'à l'aide de références grammaticales. Cela semble être particulièrement le cas à des stades

d'apprentissage peu avancés (cf. 2.4.3.). Ces résultats trouvent un écho dans les remarques suivantes de Chloé et d'Alexa qui ont participé à la deuxième phase de la recherche-action.

ENQUÊTRICE : Quels avantages voyez-vous à travailler avec des corpus et des concordances électroniques pour écrire un texte en allemand ?
CHLOÉ : Honnêtement, c'est beaucoup plus naturel comme style d'écriture. Par exemple, si on écrit un texte en allemand comme ça sur papier, ça… ça… déjà, il y a des fautes de grammaire peut-être et il y a des trucs qui ne se disent pas en allemand et qui se disent en français. Ça va être surtout à ce niveau-là, je pense.
ALEXA : Et l'ordre des mots aussi.
ENQUÊTRICE : La concordance aide pour l'ordre des mots ?
ALEXA : Quand on cherche par exemple autour d'un mot, on l'a plusieurs fois en ligne.

Certains des apprenants ont eu tendance à rédiger dans un style parataxique qui ne comporte pas un grand risque de se tromper dans le domaine de l'ordre des mots. Ainsi, Jean, dont le texte "Visite guidée" ne comporte pas de constructions complexes, n'a pas fait d'erreurs de syntaxe, contrairement à Olivia dont les différentes versions des "Visites guidées" contiennent une grande proportion de dépendantes. Dans ce qui suit, j'observerai l'évolution des caractéristiques syntaxiques dans les productions écrites de Géraldine et Vicky, les deux apprenantes du groupe classe dont tous les textes dépassent 80 mots. En dessous de ce seuil, les résultats risquent d'être peu probants. Au-dessus de ce seuil, je n'ai pas assez de données. Je fais précéder cette étude par l'observation des textes d'Alexa, en raison de la remarque qu'elle a faite, dans les entretiens, concernant le lien entre concordances et syntaxe. Alexa n'a pas fait la production de départ. Elle a produit trois textes dont un seul dépasse 80 mots. J'analyse – le cas échéant – la production de départ, et un choix pertinent des versions des trois productions rédigées pour *Projekt Prospekt*. J'introduis chaque étude par l'avant-dernière production des apprenantes, donc la dernière qui a été exécutée en ayant libre accès à des outils d'aide. Je comptabilise comme "erreurs de syntaxe" uniquement la place de l'élément observé dans la phrase. Je ne tiens pas compte de la forme (flexion, orthographe). Je définis ici "construction complexe" comme construction contenant une subordonnée (dont les groupes infinitifs, introduits par "*zu*"), un groupe verbal complexe avec auxiliaire (*werden* + infinitif, verbe modal + infinitif, formes de passé composé, etc.) et un groupe nominal complexe contenant davantage d'éléments qu'un article défini ou indéfini. Pour recenser les "constructions conformes", je comptabilise le titre du texte, ensuite toutes les unités contenant

un groupe verbal et les unités séparées par des signes de ponctuation, à l'exception des groupes nominaux sans détermination (par exemple, *Erwachsene : 8 €* [Adultes : 8 €) ou les groupes prépositionnels peu complexes, comme "*Bis bald*" [à bientôt]. Je comptabilise chacun des segments répétés plusieurs fois, le cas échéant. Je recense aussi le nombre de segments co-présents qui font partie d'un groupe verbal ou groupe nominal complexe et qui ont été intégrés de façon conforme dans le nouveau co-texte.

6.4.3.1. La syntaxe dans les productions d'Alexa

Hinter den Kulissen

Wir schlagen den Besuch der Werkstätten der Führungen vor. Eine Begegnung mit Mia der Fotograf ist vergesehen am 30.06.08 in La Cité des Découvertes um 14 Uhr. Ihr könnt unterschiedliche den Kleidungstückenwerkstätten besuchen. Die Kinder können einige Kostüm-Kreation anprobieren, um sich zu vergnügen !
Praktische Infos: Jeden Tag geöffnet, von 10 Uhr bis 12 Uhr, danach von 14 Uhr bis 16 Uhr.

Figure 68 – "Visite guidée" d'Alexa (V2).

Je présente ci-dessous les chiffres qui correspondent aux nombres réels de mots des textes. Les graphiques ont été faits à partir des chiffres pour des textes ayant 100 mots (cf. annexe IV.2.).

Type de texte	**Erreur groupe verbal**	**Erreur groupe nominal**	**Constructions complexes**	**Nombre de constructions conformes + un GV ou + un GN complexe**	**Nombre de segments co-présents conformes & élément syntaxique**	**Nombre de segments co-présents non conformes**	**Nombre de mots**
ipV1	1	0	0	2	1	1	79
ipV2	1	0	0	3	1	1	82
vgV1	1	0	5	4	0	0	57
vgV2	0	0	5	6	1	0	65
clV0	4	0	6	14	1	1	115

Tableau 52 – Syntaxe dans les productions d'Alexa.

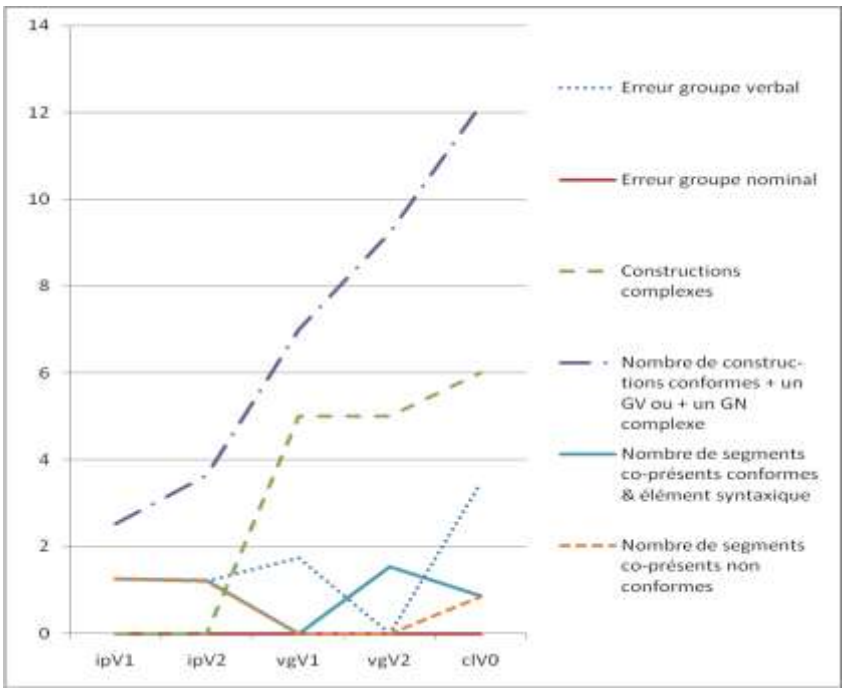


Figure 69 – Syntaxe dans les productions d'Alexa.

Les productions se complexifient et comportent de plus en plus de constructions qui sont conformes du point de vue syntaxique. Cette évolution ne semble pas vraiment avoir de lien avec l'emploi de lignes de concordance, qui reste modeste dans les textes d'Alexa. Lors de la dernière production, pour laquelle j'ai proposé comme aides un texte du corpus de référence et 87 lignes de concordance, les erreurs de syntaxe augmentent de façon significative. Cela peut s'expliquer par le fait qu'Alexa avait un niveau déjà assez avancé en syntaxe au début du projet. Elle ne s'appuie que rarement sur les lignes de concordance pour s'aider pour la syntaxe, peut-être parce que d'autres outils auxquels elle a eu accès pendant les séances lui ont suffi pour être en mesure de produire des textes conformes du point de vue syntaxique.

6.4.3.2. La syntaxe dans les productions de Géraldine

Name: Bartholéméo Grascochon (grascochon_rose@free.fr
69, rue du Porkita Rose
75017 Palot sur Seine)
IP: 168.143.112.8
Hallo Chichoune Accroforest!
Ich bin zu deiner« Strecke für mutige Kinder » gekommen: Ich bin sehr zufrieden! Ich habe alle starke Gefühle gehabt, den ich haben wollte (über ist den Tigern die Strecke sehr hoch...). Außerdem ist der Eintrittspreis nicht sehr teuer für das, was wir machen in Chichoune Accroforest. Aber die Strecke müsste größer sein und müsste während die Ferien geöffnet sein. Ich spreche über die Strecke mit meine Freunde.
Viel Spaß weiterhin an eurem Park! Viele Grüß!

Figure 70 - "Commentaire dans un livre d'or" de Géraldine (V2).

Type de texte	Erreur groupe verbal	Erreur groupe nominal	Constructions complexes	Nombre de constructions conformes + un GV ou + un GN complexe	Nombre de segments co-présents conformes & élément syntaxique	Nombre de segments co-présents non conformes	Nombre de mots
pdd	6	0	3	38			166
ipV1	1[39]	0	0	3	1	0	95
ipV6	0	0	0	3	1	0	101
vgV1	4	0	5	6	0	0	90
vgV2	1	0	5	9	0	0	95
vgV3	2	0	5	8	0	0	95
vgV4	0	0	5	11	0	0	92
clV1	3	0	5	7	1	0	95
clV2	1	0	5	9	1	0	95
clV3	3	0	9	14	2	1	119

Tableau 53 – Syntaxe dans les productions de Géraldine.

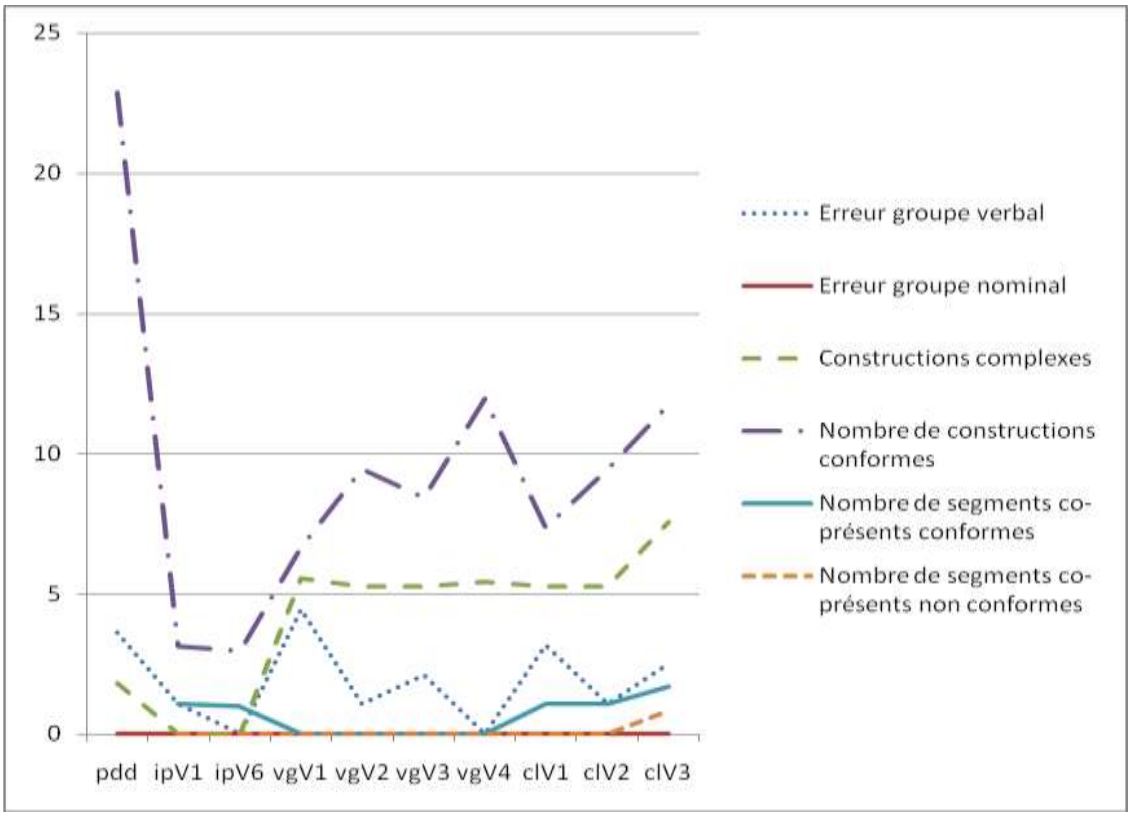


Figure 71 – Syntaxe dans les productions de Géraldine.

Géraldine a débuté par une production qui comporte une large proportion de constructions conformes et qui présente peu d'erreurs syntaxiques. Toutefois, les constructions sont peu complexes. Le nombre des constructions conformes évolue ensuite en zigzagant, presque en parallèle à l'évolution des erreurs syntaxiques autour du groupe verbal. La courbe des constructions conformes continue à monter lors de la dernière production, tout comme celle

[39] Adverbe supprimé par la suite.

des erreurs syntaxiques du groupe verbal, mais dans une proportion moindre pour cette dernière qui atteint un niveau relativement bas par rapport aux situations précédentes. La courbe des constructions complexes stagne jusqu'à l'avant-dernière production. Lors de la dernière production, elle monte de façon significative. Cela ne peut être attribué que de façon très modeste à l'emploi de segments co-présents, car Géraldine n'en a pas fait un usage significatif. Néanmoins, sa toute dernière production contient trois passages co-présents modifiés et bien intégrés dans le co-texte. Cela mène, entre autres, à la production d'un groupe nominal complexe, contenant deux adjectifs : "*eine wirklich gute Strecke*" [un parcours vraiment bien]. À la fin de *Projekt Prospekt*, Géraldine ne maîtrise pas encore parfaitement la syntaxe du groupe verbal, mais un certain progrès se dessine.

6.4.3.3. La syntaxe dans les productions de Vicky

Bert Spielmann (http//...) schrieb am 14.03.2006 um 7:43 Uhr
Hallo,

Die Ausstellung vielen Instrumente interessierten mich sehr. Ich war da und mir hat es gefallen. Ich komme nächste Monat wieder mit meine Familie, und ich hoffe, wenn ich dieses Museum besuche, mehr Zeit haben. She habt keine Audioführer in deusch, english, französich und japanisch (haben). She müsstet Internationalität Besucher angeboten. Ich habe (Es ist ?) ein besonderes Anliegen, durch die Führung aus allen zusammenführen und nicht zu trennen. Mein Kompliment für die interessanten Internet-Seiten, die ich gerne in Anspruch genommen habe.

Mit freundlichen Grüßen,

B. Spielmann

Antworte

Sehr geehrter Herr Spielmann,

vielen Dank für Ihren Eintrag im Gästebuch, und für Ihren Kompliment, der mich natürlich interessiert. Wir werden Audioführer in deutsch, english und französich haben.

A. Schühler schrieb am 18.03.2007

Ein sehr gutes Museum!!!

Figure 72 – "Commentaire dans un livre d'or" de Vicky (V0).

Type de texte	Erreur groupe verbal	Erreur groupe nominal	Constructions complexes	Nombre de constructions conformes	Nombre de segments co-présents conformes	Nombre de segments co-présents non conformes	Nombre de mots
pdd	6	0	6	17			136
ipV0	0	1	3	6	0	1	61
ipV1	0	0	2	8	2	0	92
ipV2	0	0	2	8	2	0	82
vgV0	0	0	8	14	4	0	84
vgV1	0	0	8	14	4	0	81
vgV2	0	0	8	14	4	0	81
clV0	0	0	6	18	9	0	134
clV1	1	0	6	17	3	1	135

Tableau 54 – Syntaxe dans les productions de Vicky.

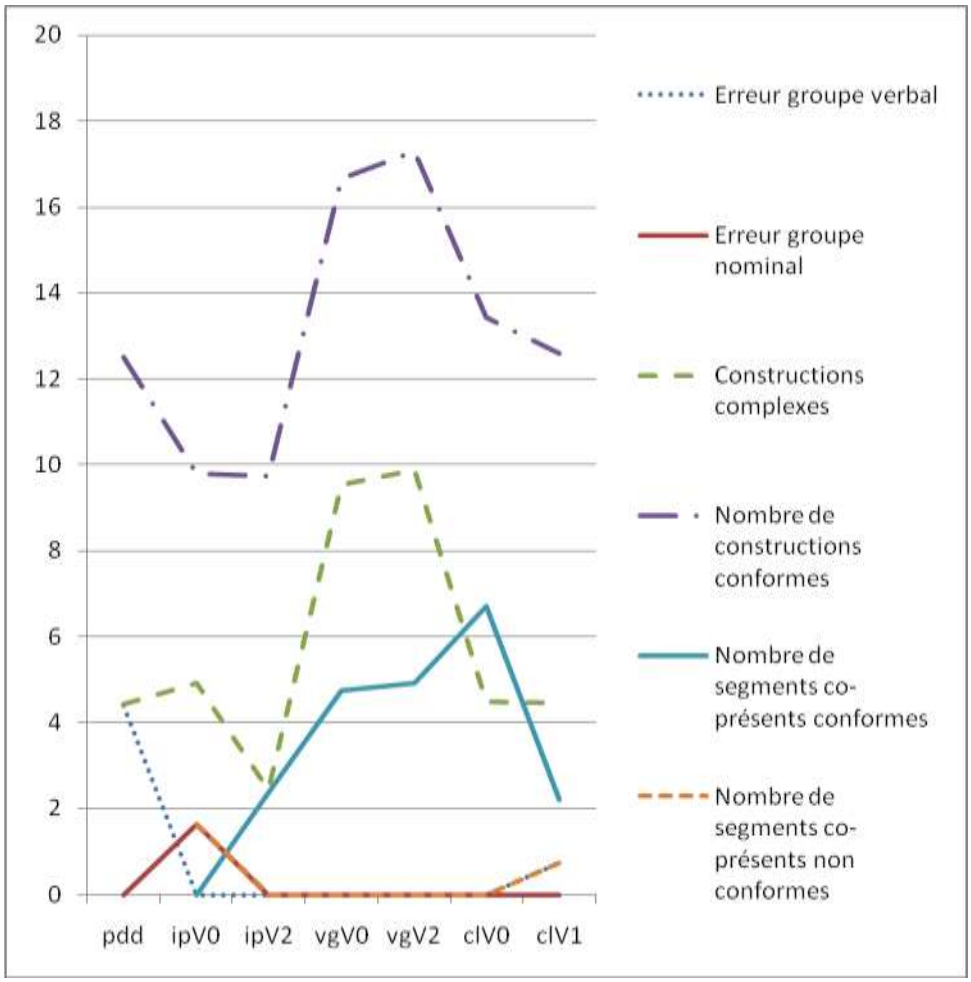


Figure 73 – Syntaxe dans les productions de Vicky.

Dans les textes de Vicky se trouvent de nombreux passages co-présents, parfois des phrases entières. Certains segments co-présents comportent des variantes plus ou moins importantes par rapport aux segments équivalents dans les corpus de référence. Vicky a presque toujours intégré les segments de façon pertinente dans le nouveau co-texte. Le texte de départ comporte six erreurs syntaxiques autour du groupe verbal. Dès la première production, les

erreurs de cette nature disparaissent. Vicky fait une seule erreur syntaxique dans la production finale. La courbe des segments co-présents conformes, des constructions complexes et des constructions conformes montent régulièrement. Les courbes des constructions complexes et celle des constructions conformes baissent à partir de la version V0 du "Commentaire dans un livre d'or". Celle des segments co-présents redescend à l'occasion de la production finale lors de laquelle Vicky a eu accès à des aides à la rédaction très restreintes. Il est possible de voir un lien entre la complexification et la précision des textes de Vicky et l'utilisation de corpus et concordances. Ces outils d'aide à la rédaction semblent l'avoir aidée à progresser dans le domaine de la syntaxe du groupe verbal.

6.4.3.4. Remarques finales concernant la syntaxe

Je constate que les données de ma recherche ne contiennent pas beaucoup d'indices qui permettent d'illustrer que les progrès d'Alexa et de Géraldine dans le domaine de la syntaxe sont à attribuer à la consultation de corpus et concordances. Pourtant, l'évolution des caractéristiques syntaxiques des textes de Vicky donne des arguments en faveur du point de vue que l'emploi intense de concordances peut être bénéfique dans ce domaine. On pourrait objecter que la reprise de passages longs dans une concordance ne signifie encore en rien la maîtrise de la syntaxe. Vicky ne s'est toutefois pas toujours contentée de simplement repérer des passages pertinents. Elle a su, après une phase d'entraînement, en déconstruire et reconstruire certains éléments afin de les intégrer dans son texte à elle. Il s'agit d'une démarche que l'on peut caractériser d'intertextuelle (cf. 3.4.2.3.2.). Selon Narcy-Combes (2005b), la production, vue comme un ensemble comprenant compréhension et expression, déclenche l'apprentissage. Il paraît donc possible que Vicky ait progressé en syntaxe, entre autres à l'aide du repérage de lignes de concordance et de leur réemploi lors de la rédaction de ses textes. Une étude portant sur la syntaxe dans des productions plus complexes pourrait apporter un éclairage supplémentaire concernant l'utilité des concordances dans ce domaine.

6.5. Le corpus comme espace des normes

Comme précisé dans l'introduction, aborder des textes de corpus est susceptible d'aider l'apprenti-scripteur à "se décentrer de sa propre production et à l'évaluer efficacement" (Marin et Legros, 2008 : 120). Si je retiens ce point de vue, il convient de chercher des indices qui renseignent sur la façon dont les apprenants ont pris de la distance par rapport à leurs productions. Ont-ils vu la différence, au niveau linguistique et discursif, entre ce qu'ils ont produit et ce que les scripteurs experts ont écrit ? Les trois types de textes de *Projekt Prospekt*

appartiennent à des genres textuels différents. Chacun de ces genres implique un degré particulier de précision linguistique et de prise en charge énonciative du discours (cf. 4.1.5.). Dans quelle mesure les apprenants prennent-ils en compte les normes qui se matérialisent dans les corpus (Rastier, 2004 et 2008) quand ils rédigent les trois textes ? Quelles traces renvoient à des obstacles qui ont empêché les apprenants de faire un pas vers l'altérité, représentée dans les textes des corpus ? Catach (1980) rappelle que les signes de ponctuation ont un caractère normatif très marqué : "Cet aspect du langage est, tout comme l'orthographe, particulièrement codifié, socialisé, pour ne pas dire aliéné." Si l'on va à l'encontre des conventions dans ce domaine, on affiche soit les limites de sa littératie en production écrite, soit sa volonté de détourner les conventions, par exemple, pour créer un effet de style. Les signes de ponctuation peuvent, entre autres, permettre d'organiser le texte du point de vue syntaxique, ou bien, ils renforcent les caractéristiques sémantiques de certaines unités du texte (Catach, 1980). Donnons un exemple dans le domaine de la syntaxe allemande : les virgules marquent, entre autres, la limite entre propositions principales et subordonnées. Contrairement à la langue française, elles ne séparent pas les compléments circonstanciels du reste de l'énoncé.

Proposition principale :		**Proposition subordonnée :**
<u>In dieser Führung</u> soll gezeigt werden	**,**	*dass das schwarze Gold seinen Namen verdient.*
complément circonstanciel		
[Pendant cette visite, il est prévu de montrer		que l'or noir mérite son nom.]

Dans un des passages longs de sa production finale, Olivia a rajouté une virgule après le complément *In dieser Führung*, tel que l'on le ferait en français. Il est intéressant de remarquer que, par ailleurs, la syntaxe de son énoncé est conforme aux normes allemandes : le verbe conjugué se trouve en deuxième position après un complément circonstanciel.

Passage dans la concordance :

Im September habe ich das Museum besucht. [En septembre, j'ai visité le musée.]

Passage dans le texte d'Olivia :

*Im märz**, habe ich *der Park besucht.* [En mars, j'ai visité le parc.]

Dans le paragraphe 6.5.2., je reviendrai sur la façon dont Olivia "résiste" initialement aux normes des textes de corpus dans le domaine de la ponctuation. Avant de faire cela, je

présente l'emploi que font les apprenants du point d'exclamation qui peut jouer un rôle pour l'organisation discursive du texte.

6.5.1. Normes discursives : points d'exclamation dans les textes de type "Visite guidée"

Dans ce qui suit, je relèverai les points d'exclamation dans le corpus de référence "Visite guidée". Il représente les normes que l'on peut observer auprès de scripteurs experts. Je l'appelle "corpus de référence", en m'appuyant sur la définition de Habert, Salem et Nazarenko (1997 : 144) : "qui contient la globalité des données dont on dispose" (cf. 1.2.4.). Je comparerai les résultats avec ceux obtenus grâce à l'étude du corpus que j'ai constitué à partir des textes de type "Visite guidée", produits par les apprenants. Je nomme ce deuxième corpus "corpus d'apprenants" (cf. 1.2.4.). J'utilise pour les deux corpus les abréviations "CR" (corpus de référence) et "CA" (corpus d'apprenants).

6.5.1.1. Points d'exclamation dans le corpus de référence

Corpus "Visite guidée", 3344 *tokens* ; 10 occurrences de "!".
Le corpus contient uniquement des points d'exclamation simples, donc non répétés.

1 Hier erwartet Dich ein Filmerlebnis der besonderen Art!	[Un évènement filmique tout à fait particulier t'attend ici !
2 Dann ist dieser Workshop genau richtig für Dich!	Alors, cet atelier est exactement ce qu'il te faut !
3 So ein Wetter!	Quel temps !
4 Alle Attraktionen regen zum Mitmachen an!	Toutes les attractions incitent à participer !
5 Handschuhe aus Metall und der Rest eines Kettenhemdes warten auf euch!	Des gants en métal et le reste d'une armure vous attendent !
6 Hier ist Kreativität gefragt!	Ici, c'est la créativité qui prime !
7 Bitte Brotzeit mitbringen!	Merci d'apporter un encas !
8 Ferne Klänge, ganz nah!	Sons lointains, tout proches !
9 So nah warst Du Deinem Star noch nie!	Tu n'as jamais été aussi proche de ta star !
10 Discovery Centre ist für die kleineren Kids reserviert!	Le centre *Discovery* est réservé aux plus jeunes !]

La prise de contact avec le lecteur se fait par un pronom personnel (lignes 1, 2, 5, 9 et 10), par une formulation infinitive à sens injonctif (ligne 7) ou à l'aide d'une formulation

impersonnelle, comme dans "*Hier ist Kreativität gefragt!"* [C'est la créativité qui prime !] et dans " *Discovery Centre ist für die kleineren Kids reserviert!*" [Le *centre Discovery* est réservé aux plus jeunes !]. Il n'y a pas d'occurrences contenant une forme d'impératif. Certains des énoncés ne se termineraient pas forcément par un point d'exclamation en français (lignes 4, 5, 7 et 10). Les occurrences équivalentes en français, exprimant une charge émotionnelle forte, nécessitent le point d'exclamation davantage que les occurrences à sens injonctif.

Citations dans les "Visites guidées" du corpus de référence, contenant un ou plusieurs points d'exclamation	**Énoncés dans lesquels on s'adresse au lecteur en utilisant un pronom personnel**	**Énoncés exprimant une charge émotionnelle forte**	**Énoncés à sens injonctif**
Hier erwartet Dich ein Filmerlebnis der besonderen Art!	x	x	
Dann ist dieser Workshop genau richtig für Dich!	x		
So ein Wetter!		x	
Alle Attraktionen regen zum Mitmachen an!		x	
Handschuhe aus Metall und der Rest eines Kettenhemdes warten auf euch!	x		
Hier ist Kreativität gefragt!		x	
Bitte Brotzeit mitbringen!			x
Ferne Klänge, ganz nah!		x	
So nah warst Du Deinem Star noch nie!		x	
Discovery Centre ist für die kleineren Kids reserviert!			x
Total : 11	**3**	**6**	**2**

Tableau 55 – Points d'exclamation dans le corpus de référence "Visite guidée".

6.5.1.2. Points d'exclamation dans le corpus d'apprenants

Corpus d'apprenants "Visite guidée" (dernière version), 963 *tokens* ; 24 occurrences de "!", "!!" ou "!!!".

Dans le corpus d'apprenants regroupant les dernières versions des "Visites guidées", il y a une proportion plus grande de points d'exclamation que dans le corpus de référence "Visite guidée" (2,38 % *versus* 0,29 %). Quatre apprenants ont choisi de doubler ou de tripler les points d'exclamation, phénomène inexistant dans le corpus de référence.

17 x "!" (Fernand : 2 x, Olivia : 1 x, Géraldine : 1 x, Jean : 1 x, Alexa : 1 x, Léa : 1 x, Véronique : 1 x) ; 6 x "!!" (Jean : 1 x, Géraldine : 3 x, Olivia : 2 x) ; 1 x "!!!" (Léa : 1 x).

Comme les scripteurs des textes du corpus de référence, les apprenants emploient les points d'exclamation dans des situations désignant une injonction, une charge émotionnelle forte ou une prise de contact directe avec le lecteur. Les passages en majuscules marquent un emploi du point d'exclamation que j'estime non motivé.

Citations dans les "Visites guidées " des apprenants, contenant un ou plusieurs points d'exclamation	**On s'adresse au lecteur**	**Charge émotionnelle forte**	**Sens injonctif**
HALLO KINDER !!!	x		
Bringt Kieselsteine mit ! (Léa)			x
Hallo !	x		
Komm in den Playmobil FunPark !			x
Du kannst viel lachen, deine Playmobil anfertigen !	x		
Komm schnell !! (Olivia)			x
Bist du mutig? (...) So komm zur Strecke für mutig Kinder !!			x
Löwen, Bären werden von uns und noch viel sehen...! Um		x	
12h00, werden wir ein Picknick über den Tigern essen!!		x	
Bis bald!! (Géraldine)	x		
FUR KINDER BIS 10!		?	
REALISIRE EIN MODELL !!			x
Male ein Modell! (Jean)			x
Komm in mein Museum !			x
Kommt uns zu zehen ! (Fernand)			x
Zeichne deinen Manga!			x
Willst du deinen ganzen Manga selber machen? (...) Dann komm ins Mangamuseum! Du hast Gelegenheit, deine Figuren zu zeichnen, deine Story zu schreiben und dein Manga zu verwirklichen! (Chloé)	x		x
Speziell für Kinder!		x	
DIE LEGO STEHEN ZUR VERFÜGUNG! (Perrine)		?	
Die welt der Musik" die du liebst ! (Véronique)		x	
Bis bald! (Simon)	x		
Die Kinder können einige Kostüm-Kreation anprobieren, um sich zu vergnügen ! (Alexa)		x	
Total CA : 23	**6**	**7**	**10**
Total CR : 10	**4 (x 2,3 = 9,2)**	**4 (x 2,3 = 9,2)**	**3 (x 2,3 = 6,9)**

Tableau 56 – Points d'exclamation dans la production finale des apprenants.

La comparaison de l'emploi que font les scripteurs experts et les scripteurs-apprenants des points d'exclamation fait ressortir que les apprenants s'adressent moins souvent que les scripteurs experts aux lecteurs en utilisant un pronom personnel. Par contre, ils utilisent

davantage d'énoncés à sens injonctif et davantage d'énoncés à haute charge émotionnelle. Les injonctions sont exclusivement exprimées par une forme à l'impératif, contrairement au corpus de référence. Dans le corpus d'apprenants, l'expression de la charge émotionnelle forte est encore renforcée par le dédoublement ou le triplement des points d'exclamation, parfois dans des situations où un seul point d'exclamation aurait suffi, comme dans cette formule de prise de contact de Léa : "*Hallo Kinder!!!"* [Salut les enfants !!!]. Parfois, même l'emploi d'un seul point d'exclamation paraît non justifié, comme dans cette citation du texte de Jean : "**fur Kinder bis 10!*" [pour enfants jusqu'à 10 ans !] ou "*sie *lego stehen zur Verfügung!* " [Les légos sont à votre disposition !], informations qui, *a priori*, n'ont rien d'injonctif, ni de surprenant et dans lesquelles les points d'exclamation ne sont donc pas justifiés du point de vue discursif. Dans la version V1 de son texte "Visite guidée", Olivia a utilisé un nombre élevé de points d'exclamations et de points d'interrogation.

> *Hallo !! Komm in *der Playmobil Funpark !!! Was kannst du machen ??? Du kannst viel lachen, deine playmobil *herzustellen ! Du kannst ein *Riese playmobil *zu *besuchen ... Du kannst die *Häre , die *Kleide , die Schuhe *zu wählen!!! Komm schnell !!! Erwarten *dir *uns !!*
>
> [Salut !! Viens dans le Playmobil Funpark !!! Que peux-tu faire ??? Tu peux beaucoup rigoler, fabriquer tes playmobil! Tu peux visiter un playmobil géant. Tu peux choisir les cheveux, vêtements, chaussures !!! Viens vite!!! Nous t'attendons !!!]

Dans le texte ci-dessus, elle a utilisé des signes de ponctuation qui conviennent davantage à un "Commentaire dans un livre d'or électronique" qu'à un texte de type "Visite guidée". Pour la rédaction de ce dernier, elle ne semble pas se décentrer de ses habitudes de rédaction. Ces habitudes peuvent être observées dans certains des courriels qu'elle m'a envoyés pour *Projekt Prospekt*. Le courriel qu'Olivia m'a fait parvenir en mars 2008 contient le signe typographique :) qui équivaut au message "sourire". Un triple point d'exclamation clôt le courriel.

> Je pense revenir demain pas de souci , le medecin me dira si je peux aujourd'hui.. J ' ai complété mon texte .Merci :) *bis bald* !!! [à bientôt].

Nous retrouvons dans ce texte les habitudes de rédaction des adolescents de l'âge d'Olivia, comme l'illustre la "parenthèse" que je propose ci-dessous. Il s'agit d'une capture d'écran de deux messages postés par des adolescentes sur le blogue d'une colonie de vacances.

Figure 74 – Points d'exclamation dans des messages d'adolescents sur un blogue.

En matière de signes de ponctuation, le corpus de référence "Visite guidée" ne contient que des points, points d'exclamation, et points d'interrogations simples. Le corpus de référence "Livre d'or" contient une occurrence de triples points, deux occurrences de points d'interrogation dédoublés, une occurrence de points d'interrogation triples et une occurrence de sept points d'interrogation. Du point de vue discursif, le texte d'Olivia n'est donc pas conforme au genre "Visite guidée", mais il le serait s'il s'agissait d'un texte appartenant au genre textuel "Commentaire dans un livre d'or électronique". Dans le paragraphe suivant, je décris plus en détail la façon dont Olivia aborde un *input* en LE lors d'une activité de révision de texte.

6.5.2. Normes linguistiques : Olivia et la date

Olivia est une apprenant dotée d'une forte personnalité. Elle a un an d'avance et elle réussit en général bien sa scolarité. Elle m'a dit le 25 mars 2008 qu'elle préfère travailler "comme ça, sans consulter le corpus". L'extrait suivant d'un de ses textes illustre qu'elle a rédigé le texte allemand en remplaçant le texte français et en laissant les parties françaises qu'elle avait encore à travailler (soulignées ci-dessous par moi). J'ai surligné les passages que j'ai annotés en vue de leur révision par Olivia.

Informations importantes :	***(Wichtige) Informationen :***
Chaque après midi les 22, 23 et 24 mars 2008, viens fêter Pâques au Playmobil FunPark :	*Die 22 , 23 , 24 marz Nachmittags , komm fêter paques in Playmobil funpark*
Tous les jours du 19 avril au 4 mai 2008 :	***Täglich von 19 april bis 4 mai 2008 :***
Viens participer à la grande course de chevaux PLAYMOBIL !	
Qui sera la gagnant?...	
Suspense garanti et surprises Playmobil pour tous les enfants participants !	*für alle kinder*
FunPark tarifs d'entrée par personne:	***Eintrittspreise :***
À partir de 3 ans: 1,50 €	*von 3 jahre : 1.50 euros.Behinderte : 1.00*
Pour les personnes handicapées: 1,00 €	*euros. Kleinkinder:Eintritt frei .*
Pour les enfants de moins de 3 ans: gratuit	*Eintritt frei* à la boutique seule .
Accès à la boutique seule: gratuit	

Tableau 57 – La version V1 des "Informations pratiques" d'Olivia.

Certains passages co-présents de sa production comportent des imprécisions qui montrent qu'Olivia n'a pas copié à l'aide du traitement de texte des formules dans le corpus ou la concordance numérisé. Le 31 mars 2008, elle a corrigé à la main son texte annoté. Elle a réussi à revoir plusieurs passages, mais autour de la date et autour du mot clé "*geschlossen*", les normes fixées pour le genre textuel "Informations pratiques" ne sont pas parfaitement respectées. En plus, sa production n'est pas empreinte du style elliptique qui caractérise certains passages dans les textes du corpus. L'exemple ci-dessous illustre cela.

Formulation d'Olivia :

Dienstag 1 september ist Playmobil Funpark geschlossen [*Mardi 1 septembre Playmobil Funpark est fermé]

Ma proposition de formulation alternative :

am Dienstag 1. September geschlossen [fermé mardi 1er septembre]

Le 16 avril 2008, Olivia m'a envoyé un courriel concernant la version V3 de ses "Informations pratiques". Dans ce courriel, on peut lire des traces de déstabilisation et peut-être le début d'une décentration.

> Je crois , je dis bien je crois , que ça commence à être ça ... peut etre que par miracle , j' ai même tout bon .... Je m' attaque quand même au second sujet !! Courage !!!!!!!!! à Jeudi !!!

Le désarroi et en même temps l'envie de réussir que j'ai ressentis dans ces lignes m'ont poussée à prendre un temps particulier avec Olivia. Le 17 avril, je lui ai proposé de revoir son texte avec moi pendant que les autres faisaient un travail d'observation de lignes de concordance sur papier. Voici une partie du script de cette séance. Le script complet peut être consulté en annexe (I.3.1.).

ESL : Ça, ce n'est pas une date allemande [en ponctuant les deux derniers mots].

OLIVIA : Je ne comprends pas, pourtant j'ai cherché. Le problème avec les extraits que vous avez donnés, il n'y a pas les mots… enfin… dans *Word*… regardez.

ESL : Oui, oui, vas-y, raconte, Olivia. Reviens sur le corpus, et tu tu m'expliqueras.

OLIVIA : Là-dedans, j'ai du mal à me retrouver. (…) C'est que j'ai vu que ce n'était pas vraiment rédigé dans la concordance ou dans le corpus. Des fois, c'était écrit genre… au lieu d'écrire le mois de mars ou le 1er mars, il y avait écrit un zéro trois.

ESL : D'accord, tu voulais un peu imiter ça en fait, en version abrégée. Alors qu'est-ce que tu peux faire pour trouver une date là ?

OLIVIA : Mais je cherche par le mois.

ESL : Oui, voilà, c'est une bonne idée. (…)

OLIVIA : Il y a *ein Mai, ein Mai* [un mai] (rit).

ESL : *Erster Mai* [premier mai]. Tu vois comment ils font ? Tout bêtement, qu'est-ce qu'ils font ?

OLIVIA : Il mettent le jour et le mois.

ESL : Voilà, copie-le, voilà, avec le copier-coller, tu vois... c'est ça qu'il faudrait peut-être un petit peu faire plus pour être moins sujette aux erreurs d'orthographe (…).

OLIVIA : On peut mettre "*Osterfest*", et après, on met…

ESL : (la coupe)… voilà, et pas une phrase, c'est tout à fait bien. (…) Tu te débrouilles pour après essayer d'intégrer tout ça parce que c'est pas encore super, regarde, la date par rapport à ce que l'on avait vu…

OLIVIA : C'est pas ça avant *März* ?

ESL : Essaie de le faire, Olivia, après, on le regardera, hein. (Va voir les autres élèves)

OLIVIA : Et la date, je mets quoi ?

ESL : Tu mets la date là.

OLIVIA : N'importe quelle date ?

ESL : Les dates qui t'intéressent, c'est celles-ci là. *Am*. (…) Tu vois, c'est pas encore conforme. La date c'était quoi ? Tu mets 22 virgule. Tu vois que je tiens à mon point là.

OLIVIA : C'est pas une date, c'est les… plusieurs jours différents.

ESL : Ah… oui mais c'est le 22, 23 – il faut un point chaque fois pour dire que c'est une date, tu vois.

OLIVIA : Ah bon.

ESL : Tu mets 22 point virgule, 23 point virgule… etc.

Ce script permet de se rendre compte qu'Olivia "ne voit pas" les points qui caractérisent la date allemande. Elle confond les adjectifs numéraux ordinaux avec les adjectifs numéraux cardinaux : elle prononce "*ein Mai*" [un mai] au lieu de "*erster Mai*" [premier mai], et elle considère une suite de dates comme une énumération de chiffres cardinaux : "[22, 23] C'est pas une date, c'est les… plusieurs jours différents." Un travail de conceptualisation aurait été nécessaire afin de lui faire prendre conscience de cette différence. Jusqu'à la fin de la séance individuelle, j'insiste pour qu'Olivia revienne sur l'*input* en LE afin de comparer ce qu'elle a écrit et ce qui se trouve dans la concordance. Je finis par être très explicite : "il faut un point chaque fois pour dire que c'est une date". Toutefois, sa remarque finale "Ah bon", dite sur un ton de surprise, peut être considérée comme indice d'une prise de conscience concernant la présentation de la date en allemand (cf. Portine, 1998 : "expériences-déclics que Vygotski appelait 'ah !-ah ! expériences'"). Il est légitime de se demander à quel point la précision doit être exigée dans des productions d'apprenants de niveau intermédiaire. Insister sur un point, est-ce bien nécessaire ? J'estime que le genre textuel "Informations pratiques" nécessite une précision dans le domaine de la date, contrairement au genre textuel "Commentaire dans un livre d'or" qui se caractérise par une orthographe empreinte de nouvelles normes (cf. 3.5.). De plus, j'ai encouragé les apprenants à copier électroniquement certains segments, justement afin de les soulager au niveau de la précision orthographique. J'ai répété cette proposition lors de la séance individuelle. Toutefois, Olivia n'a pas été d'emblée prête à changer ses habitudes de rédaction. Les autres productions d'Olivia ont également montré qu'elle met du temps à développer une attitude lui permettant de se décentrer.

6.6. Littératie électronique

Travailler avec des corpus et concordances nécessite certaines savoirs faire électroniques, aussi appelés "littératie électronique" (cf. 3.6.3.). Les paragraphes suivants s'intéressent à la

maîtrise des outils bureautiques des apprenants et à la formation par les pairs dans le domaine de la littératie électronique.

6.6.1. Maîtrise des outils bureautiques

Au début de la deuxième phase de la recherche, seulement cinq des dix-sept apprenants possédaient suffisamment de compétences en informatique pour suivre le projet sans formation supplémentaire. Ils avaient, entre autres, besoin de bien connaître le traitement de texte, en l'occurrence celui de *Word* et de *Open Office*, notamment la fonction "rechercher" du menu "édition". Ils devaient savoir envoyer des courriels dans une boîte courriel commune, qui se présente autrement que celle qu'ils ont à la maison. Il était important qu'ils sachent retrouver des fichiers dans la boîte courriel et qu'ils arrivent à en joindre lors d'un envoi.
Le 20 mai 2008, Alexa n'a pas remarqué, dans le courriel que je lui ai envoyé, le fichier joint qui contenait son texte annoté par moi. Le 22 mai, elle n'a pas réussi à trouver son message dans la boîte courriel commune qui se présente comme ci-dessous. Il aurait suffi qu'Alexa clique sur la page précédente pour trouver les autres messages.

60	60	21.05.2008	Eva Lacroix	Fernand Führung V1
61	61	21.05.2008	Eva Lacroix	Vicky Führung V1
62	62	21.05.2008	Eva Lacroix	Vicky Infos pratiques V3

Tableau 58 – Interface de la boîte courriel pour *Projekt Prospekt*.

Au bout de trois mois d'utilisation de la boîte courriel commune, Jean et Georges ont modifié l'adresse Thomas@clg-st-leu-landowska.fr, en remplaçant "Thomas" par leur prénom : "jean@clg-st-leu-landowska.fr " et "Georges@clg-st-leu-landowska.fr" (22/05/2008). Cela a provoqué une perte de l'envoi de leur document.

Jean, Janine, Georges et Olivia ont eu du mal à prendre l'habitude d'écrire l'objet de l'envoi, suivi ou précédé de leur nom. Olivia a parfois écrit, dans la ligne de l'objet, des formules de prise de contact, comme "*Hallo!*" [Salut !] ou "*Hilfe!*" [Au secours !], sans s'identifier par son prénom ou nom. Dans l'entretien, Chloé précise qu'elle a rapidement adopté les concordances parce qu'elle est habituée à travailler avec l'ordinateur.

CHLOÉ : Moi, j'ai tout fait sur informatique parce que je suis à l'aise, et honnêtement, ça m'a aidé à me servir des concordances. Je trouve que c'est… c'est pas mal comme système.

Pour d'autres apprenants, le travail avec l'ordinateur a été moins aisé. Le 20 mai 2008, Jean m'a fait savoir qu'il préférait écrire à la main parce que l'envoi de courriels entre deux cours lui prenait trop de temps. La saisie de texte a été un obstacle non négligeable non seulement pour Jean, mais aussi pour Janine, Perrine et Vicky, comme l'illustrent les extraits d'entretiens suivants.

ENQUÊTRICE : Quand vous avez transmis les textes, c'était par e-mail ?
JANINE : Moi pas.
PERRINE : Surtout pas. Moi, je faisais tout à la main.

Vicky dit avoir préféré la rédaction sur papier. Elle remarque toutefois que la numérisation des textes l'aide à réduire les erreurs d'orthographe.

VICKY : Moi, personnellement, j'ai fait une partie [du travail de rédaction] chez moi. (…)
ENQUÊTRICE : Quand tu dis "chez toi", c'était sur l'ordinateur aussi ?
VICKY : Non, non. C'était sur papier, parce que j'ai des problèmes… avec l'informatique. (…) En fait, quand je travaillais à la maison, je l'écrivais vraiment, je le rédigeais, mais je faisais beaucoup de fautes d'orthographe. Donc, le taper de temps en temps, ça me permettait d'être corrigée. Donc, je l'imprimais assez souvent.
ENQUÊTRICE : Et après, quand tu corrigeais, tu corrigeais sur papier, à la main, et ensuite sur l'ordinateur ?
VICKY : Hein non, des fois, je le donnais directement, elle [l'enseignante] le corrigeait, mais surtout ce que j'ai envie de faire, c'est de faire quelque chose de complet et ensuite de le taper chez moi, de le mettre sur une clé USB et de l'imprimer au collège.

Vicky a rédigé deux versions de ses "Informations pratiques" à la main (V0 et V1), et elle a tapé uniquement la version finale à l'ordinateur. Entre les versions V0/V1 et la version V2, il y a plusieurs modifications qui sont dues à l'utilisation du clavier. La lettre "ß" est présente dans "*Rosenstraße*" et "*Ermäßigung*" en version 1, mais en V2, Vicky la remplace par la combinaison de lettres "ss", donc par la variante de la lettre "ß" utilisée dans des formulaires officiels ou en Suisse. En V2, le numéro de téléphone perd la barre oblique et le sigle "€" disparaît pour être remplacé par "*Euros*". Vicky a fait d'autres modifications lors de la numérisation du document. Le mot "*Eintritt*spreise" perd le "s" de liaison lors du passage à la version électronique. En V2, le code postal est plus court qu'en V1. Ce n'est pas "non

conforme", mais il n'était pas nécessaire d'effectuer cette modification. Entre les versions V1 et V2, il y a une perte de dix mots. En fait, Vicky n'a pas tapé le tout dernier paragraphe qu'elle avait rajouté en V1 et qui n'existait pas encore en V0. Passer au traitement de texte a fait perdre à Vicky une orthographe et une ponctuation conformes au profit d'une orthographe et d'une ponctuation qui ne pose pas problème au niveau du clavier, mais qui s'éloigne, pour certaines des modifications, des conventions discursives du genre textuel "Informations pratiques".

6.6.2. Jouer le rôle de tuteur pour ses pairs

Pendant le projet *Brieffreunde.org*, plusieurs apprenants ont choisi ou accepté de s'occuper de leurs camarades qui avaient besoin d'aide technique. Je note dans mon journal de bord les observations suivantes.

> Corentin, le nouvel élève, dit être habitué à utiliser l'ordinateur. (…) Louise et Christine ont choisi de s'occuper de lui pour lui montrer où il faut enregistrer le travail, au lieu de faire leur travail. J'ai officiellement accepté cela pour Louise. De toute façon, il me manquait un ordinateur, et Louise a déclaré qu'elle ne savait pas quoi faire puisque "vous ne m'avez pas dit comment je dois traduire mon texte en allemand" (26/03/2008).
>
> Janine s'est fait expliquer par moi l'utilisation du traducteur *Reverso* pour trouver le pluriel du nom. Ensuite, elle a très bien expliqué la procédure à trois de ses camarades (27/05/2008).

J'ai régulièrement encouragé les apprenants les plus avancés de s'occuper d'un de leurs camarades quand ils avaient fini leur travail. Le 26 mai 2008, j'ai distribué au groupe classe une fiche de travail qui a fixé les rôles de tuteur/tutoré et les tâches pour certains des apprenants. Il s'agit exclusivement de tâches concernant la littératie électronique.

> Christine, à l'aide de Chloé : envoyer son texte sous forme de commentaire sur le blogue (attention : indiquer une adresse e-mail, même fictive). Faire vérifier l'envoi en faisant appel à Mme Lacroix.
>
> Georges, à l'aide de Léa, Jean à l'aide de Chloé : envoyer un mail à Mme Lacroix + fichier attaché (leur dernière version de la "Visite guidée", si possible). Faire vérifier l'envoi en faisant appel à Mme Lacroix.
>
> Alexa à l'aide de Camille, Janine à l'aide de Géraldine : faire une liste de tous ceux qui ont reçu un mail de ma part depuis le 22 mai. Noter le nombre des mails envoyés à ces élèves. Présenter cette liste devant la classe.

Cette forme tutorat est très "réglementée" et ne laisse pas beaucoup de place pour des échanges concernant le sens du texte. Dans l'entretien du 6 mai 2008, Georges et Simon

expriment le souhait de travailler en binôme pour s'aider mutuellement à comprendre l'*input* en LE et pour réviser leurs textes.

GEORGES : Il serait bien aussi que ceux qui sont en avance aident les autres qui sont un peu en retard.
ENQUÊTRICE : Tu aimerais bien travailler en binôme ?
GEORGES : Oui. Ce serait plus rapide, plus pratique.
ENQUÊTRICE : Mais tu penses qu'ils t'apporteraient quoi, les gens qui sont en avance ? Tu attends quoi d'eux ?
GEORGES : Mais qu'ils nous aident à traduire et à comprendre ce qu'on marque parce que si on ne comprend pas, ça n'a aucun sens. (…) C'est beaucoup plus pratique de travailler en binôme.
ENQUÊTRICE : Tu supposes que la personne qui travaille avec toi, ton binôme, verrait par exemple tes erreurs que tu pourrais faire ?
SIMON (sic) : Oui. Pas forcément quelqu'un de niveau plus élevé que le sien, mais c'est mieux de le faire à deux parce qu'il y a des erreurs que l'un relèverait et pas l'autre.

Cathy a participé à la troisième phase de la recherche-action. Plusieurs situations en présentiel m'ont fait comprendre qu'elle avait des difficultés dans le domaine de l'utilisation de l'ordinateur et que cela la freinait pour exécuter le travail basé sur l'emploi de concordances. Régulièrement, Nathalie, une de ses camarades, s'est spontanément occupée d'elle pour la soutenir. En octobre 2008, j'ai filmé une séance du projet *Brieffreunde.org* [correspondants.org], destinée à observer la façon dont Cathy se sert des outils électroniques que j'ai proposés au groupe classe pour rédiger une annonce comme celles que l'on met en ligne pour trouver un correspondant. Les autres membres du groupe classe étaient présents. Ils ont été invités à participer à la formation de leur camarade dans le domaine de l'utilisation de l'ordinateur. L'utilisation d'un vidéoprojecteur a permis de visualiser les actions de Cathy. Je lui ai annoncé qu'elle était censée réviser certains passages de son annonce. J'ai cité ces passages dans le commentaire que j'ai laissé pour elle sur le blogue *esldeutsch*. Cathy avait besoin de repérer la version la plus récente de son texte qu'elle avait publié auparavant sur le blogue et de trouver mon commentaire concernant ce texte. Les notes suivantes que j'ai prises après la séance filmée par moi permettent d'illustrer la façon dont l'apprenante a procédé pour trouver l'un de ces deux documents.

> Nous sommes en salle de classe habituelle. Cathy se trouve devant le vidéoprojecteur. Elle parcourt d'abord la page qui affiche tous les commentaires de l'article "*Meine Annonce*" [Mon annonce]. Il y en a une bonne centaine. Je lui conseille d'utiliser une autre technique pour "faire plus vite". Lucie lui propose d'utiliser un raccourci, la touche "F3", que je découvre par la même occasion. Cette touche correspond à la fonction "rechercher dans cette page…".

L'observation du procédé de Cathy renvoie au sujet des indices métatextuels, donc des moyens de signalement de la localisation d'un contenu dans un document.

6.6.3. Indices métatextuels

Les corpus et concordances peuvent être classés dans la catégorie des dispositifs complexes. L'utilisation de tels dispositifs nécessite un savoir-faire dans le domaine des moyens de signalement concernant le contenu (lecture de menus et sous-menus, etc.) : tables des matières ; index ; menus ; en-têtes, etc. Ces moyens sont appelés "indices métatextuels". Ils sont censés aider le lecteur à localiser une information donnée dans le document complexe (Coutelet, 2003). Leur utilisation peut poser problème à des enfants de moins de 13 ans, mais aussi à certains adultes (Rouet *et al.*, 2004). Le concordancier que j'ai utilisé pendant la phase principale de ma recherche, *Text-Concordancer (v2)* de Cobb (nd), classe les mots clés de façon alphabétique. Cette présentation est un premier indice métatextuel qui peut être exploité par les apprenants lors de leurs recherches. Les signes de ponctuation (points, virgules, etc.) fournissent également de tels indices. Ces deux types d'indices se sont avérés insuffisants lors de ma recherche. La collection d'outils *Antconc3.2.1.w* (Anthony, 2007) et la collection d'outils de statistique textuelle *Lexico3* (Fleury *et al.*, 2003 ; cf. 1.6.4) proposent, entre autres, des options de classement alphabétique des contextes de gauche ou de droite du mot clé, la réduction ou l'élargissement du co-texte du mot clé et la localisation visuelle d'un lexème ou segment dans le corpus. Je prévois l'exploitation du potentiel prometteur d'*AntConc3.2.1.w* et de *Lexico3* lors de recherches futures.

6.6.4. Remarques finales concernant la littératie électronique

En 6.6.1., j'ai mis en évidence les obstacles qui ont alourdi le travail basé sur l'utilisation de corpus lors des deux premières phases de ma recherche. Il ne s'agit pas de uniquement de difficultés dans le domaine de la compréhension de l'allemand ou concernant l'adaptation d'un segment identifié au nouveau co-texte. La première étape à franchir est la maîtrise suffisante de certains outils bureautiques. La deuxième étape est le repérage, par l'apprenant, de son texte parmi de nombreux textes envoyés par courriel ou publiés sur le blogue. La troisième

étape consiste à exploiter les aides à la rédaction et à la révision, offertes par les corpus et concordances. Un bilan de départ en connaissances informatiques et une coopération suffisamment bonne avec l'enseignante de technologie auraient peut-être permis de réduire certaines des difficultés des apprenants.

Le paragraphe suivant décrit les raisons pour lesquelles la coopération interdisciplinaire entre allemand, technologie et français que *Projekt Prospekt* a visée n'a pas donné entièrement satisfaction.

6.7. Interdisciplinarité

À la fin de la deuxième phase de la recherche, deux apprenantes se prononcent en faveur d'un travail interdisciplinaire.

> Je suis satisfaite du projet, car d'une part faire des projets interdisciplinaires me plaît beaucoup (Camille, questionnaire final).
>
> Alexa a trouvé le projet "compliqué", et malgré les difficultés d'organisation, elle a apprécié le côté interdisciplinaire du projet (journal de bord, 4 juin 2008).

Pourtant, le travail interdisciplinaire (français, technologie, allemand) a eu peu de retombés visibles, comme le précisent Chloé et Vicky.

CHLOÉ : Moi, ce qui me dérange un peu, c'est le lien entre les trois matières, techno, allemand et français, parce que en français, bon, on a eu un ou deux cours dessus, alors qu'en allemand, on passe depuis le début à tous les cours, et en techno, elle nous en a parlé une demi-heure.
VICKY : Oui, en techno, une fois qu'on sait taper…
CHLOÉ : Oui, moi, je ne sais pas les autres, mais taper un texte, je peux le faire chez moi. Je n'ai pas besoin d'un cours de techno pour taper mon texte, quoi.
ENQUÊTRICE : Vous regrettez qu'il y ait un manque de coordination entre les matières ? Vous ne voyez pas très bien le lien entre les trois matières ? D'accord. Et ce serait donc une de vos propositions, qu'il y ait une meilleure communication entre les trois matières par rapport à votre projet.

Le travail interdisciplinaire a, en effet, été peu conforme à mes attentes. J'ai pu obtenir une seule réunion – la toute première – qui a réuni les trois enseignantes impliquées dans le projet.

Lors des réunions suivantes, j'ai eu l'occasion de voir soit l'enseignante de français, soit l'enseignante de technologie. L'enseignante de français a été absente lors de la séance de présentation du projet devant le groupe classe. Par contre, elle a assisté à une des séances d'allemand, dédiée à l'observation des genres textuels. L'enseignante de technologie n'a pas réellement suivi le planning (cf. annexe IX.1.), et elle n'a pas rendu le questionnaire final, contrairement à l'enseignante de français. La citation ci-dessous présente une partie de ce que la collègue de français a envoyé comme réponse au questionnaire final.

> La principale motivation a bien sûr été la possibilité d'un travail interdisciplinaire, tel que je l'avais déjà expérimenté lors des *Itinéraires de Découverte*. Je souhaitais travailler différemment avec les élèves, leur donner une vision plus concrète des cours de français et leur permettre de mieux saisir les liens entre les différentes matières. Il aurait été souhaitable d'avoir des heures de concertation entre les trois matières et plus de possibilités d'intervention des trois professeurs concernés dans un cours commun. Il serait intéressant, si l'expérience se renouvelait, qu'une heure fût prévue dans l'emploi du temps des élèves pour permettre de les réunir périodiquement avec les enseignants concernés et de faire le point. Je regrette aussi de ne pas avoir pu travailler avec le professeur de technologie et que les problèmes de temps n'aient pas permis aux élèves qui n'étaient pas à l'aise avec l'outil informatique de mener le projet à terme.

La recherche-action va de pair avec une prise de risque : changer les habitudes des acteurs concernés – enseignants, apprenants, familles, etc. – n'est pas anodin (Missotte, 2003 : 249). En France, dans le milieu de l'enseignement secondaire, travailler en équipe semble être une pratique qu'il convient de cultiver particulièrement afin qu'elle puisse davantage être banalisée.

> Cette démarche [de recherche-action] ne va pas de soi tant l'idée d'une intelligence collective est absente de la formation et de la culture professionnelle (Missotte, 2003 : 251).

J'en conclus qu'il est important d'intégrer des projets de recherche dans le projet d'établissement, afin de pouvoir obtenir des fonds permettant de former les enseignants impliqués dans le projet, et afin d'organiser des moments de concertation. L'enseignante de français a utilisé les documents que j'ai mis à sa disposition, mais j'aurais aimé qu'elle les critique et qu'elle en produise éventuellement elle-même. La formation pourrait alors porter non seulement sur les outils, mais aussi sur la collaboration et la part contributive de tous les acteurs.

6.8. Guidage

Quel effet le guidage que j'ai offert a-t-il eu sur les apprenants ? Mes annotations et les échanges écrits et oraux, leur ont-ils permis d'exploiter l'*input* en LE pour la rédaction ? Mes interventions, les ont-elles aidés à développer une attitude d'apprenant-chercheur ? Une étude de cas et l'analyse de mes annotations de textes d'apprenants visent à donner des éléments de réponse à ces questions.

6.8.1. Guidage en présentiel – le plongeon de Fernand

Suite à une période lors de laquelle Fernand a semblé stagner dans son travail, je l'ai invité à participer à une séance de travail individuelle. Elle a eu pour objet l'entraînement au repérage de segments pertinents afin d'être en mesure de réviser sa production écrite. Lors de cette séance, mes interventions sont nombreuses et parfois très directives, ce qui prolonge les effets de la remarque "musclée" que j'ai fait parvenir à Fernand par courriel deux jours plus tôt. Je lui ai interdit l'utilisation de certains outils d'aide à la rédaction (cf. la gradation des types d'intervention en 6.8.2.1.) : "PAS DE TRADUCTEUR, MAIS LA CONCORDANCE et le CORPUS à utiliser." Précisons que Fernand a utilisé le traducteur pour rédiger la première tâche et qu'il a dit, lors de l'entretien, qu'il considérait l'utilisation de concordances comme "une perte de temps" (cf. 6.2.6.2.). L'extrait du script de la séance, enregistrée à l'aide d'un dictaphone, permet de suivre le cheminement de Fernand dans les eaux, pour lui parfois "troubles", des concordances.

6.8.1.1. Guidage et repérage

ENSEIGNANTE : Alors, comment on pourrait chercher ici si on ne connaît pas le mot "visiter" ? On veut chercher "visiter un musée", on ne connaît pas le mot "visiter".
FERNAND : On met "musée".
ENSEIGNANTE : Voilà. (Sur un ton d'encouragement) On va essayer ? Vas-y.
FERNAND : ++++++
ENSEIGNANTE : Alors, est-ce que t'as "*Museum*" là ? C'est pas dit, mais bon. On peut déjà voir. Alors, là… est-ce qu'il y a un verbe autour de ce mot ? Ah oui, je vois, moi je vois. Je suis sure que tu vas trouver ça.
FERNAND : Donc, là, ce serait "visiter"…
ENSEIGNANTE : (le coupe) Ton projet, c'est de dire "Nous allons visiter le musée" ou "Tu peux visiter le musée librement". Alors, je te jure qu'il y a quelque chose qui va.

FERNAND : +++++++
ENSEIGNANTE : Même deux choses + pour moi.

Comme la question de recherche a été posée par moi et qu'elle n'est pas le fruit d'une réflexion de Fernand, il est peu étonnant que Fernand l'oublie en cours de route.

ENSEIGNANTE : Dis-moi au hasard, où est-ce que tu sens… où est-ce que tu le sens bien ? Quelque chose qui parle de… quelle ligne-là ?
FERNAND : (désigne un mot du doigt). Celle-là.
ENSEIGNANTE : Ça va être "*kunsthistorisches*" [adjectif ayant le sens de "relatif à l'histoire de l'art"]. Tu vois un verbe dedans ? Parce qu'on cherche un verbe.
FERNAND : Ah, ce sera un verbe ?
ENSEIGNANTE : Un verbe, oui, pour dire "visiter", c'est un verbe.

Le dialogue précédent permet d'avancer qu'un travail de conceptualisation serait nécessaire pour aider Fernand à repérer des éléments pertinents dans la concordance. Il semble avoir cherché sans trop savoir de quelle nature doivent être les éléments recherchés (cf. 2.5.2. concernant les attentes de l'apprenant quand il aborde l'*input*).

ENSEIGNANTE : Là, il y a deux choses là qui moi, personnellement me plaisent pour ton texte. Mais bon, j'ai un faible pour ça, je trouve vraiment que c'est ce que tu voulais dire.

Je constate que je me suis trop accaparée le travail de Fernand. À la fin de la séance, j'ai exprimé ma satisfaction concernant les segments que j'ai identifiés à la place de Fernand. Toutefois, il a fini par repérer le verbe "*gehen*" qui fait partie d'un des deux segments pertinents. J'ai proposé à Fernand de "prendre ou de laisser" les segments qui me paraissent pertinents, et il a semblé plutôt "consentant".

ENSEIGNANTE : Tu le mets dans ton texte, si tu veux. Tu n'es pas obligé.
FERNAND : Oui. On verra bien.

6.8.1.2. Rôle de la langue maternelle pour la compréhension

Certains passages du script de la séance avec Fernand montrent le lien étroit que j'ai établi entre la langue maternelle et la compréhension lors de la séance. Je l'ai invité plusieurs fois à

formuler en français ce qu'il voulait dire. Au début de la séance, je lui ai fait préciser en français le sens qu'il a attribué à l'énoncé avec lequel il a débuté sa rédaction : "*Du bist frei*" [Tu es libre].

ENSEIGNANTE : Alors, est-ce que "*du bist frei*", tu l'as vu quelque part ?
FERNAND : Non, j'ai vu le mot "*frei*", et j'ai mis…
ENSEIGNANTE : Alors, il faudrait mettre un verbe du genre "visiter", non ? Ou "Tu es libre dans le musée", c'est tout ce que tu veux dire ?
FERNAND : En fait, j'aimerais bien dire "Vous êtes libres de vous déplacer pendant ce temps dans le musée".

À part l'invitation à une reformulation en français, je n'ai pas vraiment donné l'occasion à Fernand de formuler sa propre question de recherche. Il a été censé trouver ce que j'avais trouvé avant lui. J'ai essayé de lui faire trouver le sens d'un mot allemand en lui proposant de compléter un segment français par l'équivalent français du mot inconnu, "*durch*" [à travers].

ENSEIGNANTE : Qu'est-ce que ça peut être, "*durch das Museum*", "*das Museum*", ça veut dire ? [à travers le musée, le musée]
FERNAND : Le musée.
ENSEIGNANTE : Nous allons… Qu'est-ce que ça peut être, "*durch das Museum*" ? "Nous allons blabla le musée", qu'est-ce qui peut manquer à la place de blabla ?
FERNAND : Découvrant.
ENSEIGNANTE : Non, en fait, c'est une préposition.

La proposition de Fernand m'a surprise : "Nous allons découvrant le musée." Je me serais plutôt attendu à "Nous allons découvrir le musée". Il est vrai que le sens littéral "Nous allons à travers le musée" n'était pas forcément facile à trouver à l'aide de la technique de la "lacune à combler". J'ai néanmoins encouragé Fernand à poursuivre dans cette voie, donc d'aller vers le texte et d'essayer de le comprendre, au risque de se tromper.

ENSEIGNANTE : Essaie de le prendre comme ça, de… de… de prendre les choses qui sont là, de piocher à droite et à gauche, ça, je le comprends, je vais creuser un peu, est-ce que je le comprends ?

Fernand a semblé percevoir l'*input* en LE comme une masse de mots dans laquelle il avait du mal à trouver des repères. Il n'a pas systématiquement pris le mot clé comme point de départ, et il n'a pas lu en direction de gauche, mais seulement vers la droite à partir de son point de départ. La séquence suivante illustre qu'il a peiné à trouver le début et la fin d'un segment pertinent. Cette difficulté a été renforcée par la perte de signes de ponctuation, provoquées par une gestion imparfaite du texte allemand par l'outil *Text-Concordancer (v2)* de Cobb (nd).

ENSEIGNANTE : Ça commence où à ton avis ? C'est un peu perturbant, comme ils ne mettent pas les points et tout ça…
FERNAND : Oui.
ENSEIGNANTE : …mais tu commencerais où là, comme ça, spontanément, la phrase-là ? Est-ce que c'est ça, le premier mot, où est-ce que tu…
FERNAND : Peut-être là.
ENSEIGNANTE : Oui, tu peux laisser tomber l'horaire, on n'a pas trop besoin. Qu'est-ce qu'il y a après ? Lis-le à haute voix, juste pour voir ce que ça te fait.
FERNAND : "*Gemeinsam gehen wir durch das Museum entdecken*"… [Ensemble, nous traversons le musée découvrir…]
ENSEIGNANTE : (le coupe) Voilà. Alors, là il manque une virgule, tu ne peux pas le savoir. Donc, ça veut dire… "*Gemeinsam gehen wir durch das Museum*." [Ensemble, nous traversons le musée.]
ENSEIGNANTE : (…) je sens bien que pour toi c'est difficile d'isoler le début et la fin de l'élément intéressant, hein.

La séance que je viens de décrire est un condensé des obstacles que l'on peut rencontrer en invitant les apprenants à utiliser les corpus et concordances. Tous les apprenants n'ont pas peiné autant que Fernand. Néanmoins, je constate que certaines attitudes de ma part ont contribué aux difficultés des apprenants, tandis que d'autres les ont aidés. Dans ce qui suit, je présenterai une étude de mes annotations des productions écrites en cherchant à mettre particulièrement en valeur ce qui était bénéfique au développement d'une attitude d'apprenant-chercheur auprès des apprenants.

6.8.2. Guidage à distance

6.8.2.1. Annotations des productions écrites

J'ai annoté les diverses versions des productions des apprenants lors de *Projekt Prospekt* et de *Brieffreunde.org*. Les annotations ont eu pour objectif de donner aux apprenants l'occasion de réviser leurs productions en ayant le moins possible recours à mon aide directe. Afin de leur en donner les moyens, je leur ai donné accès à des corpus et concordances, faits pour les besoins des projets, et à d'autres outils d'aide à la rédaction, comme le dictionnaire en ligne *LEO*, parfois un traducteur en ligne, une leçon ou un article sur le blogue, etc. Dans le tableau ci-dessous, j'ai recensé les types d'annotations dont j'ai pourvu toutes les productions du groupe classe. J'ai choisi une présentation qui indique un degré croissant d'intervention directe de ma part. Le terme d'"intervention" correspond ici l'acte de souligner ou surligner un passage que l'apprenant doit réviser. Dans certains items, j'utilise le terme de "nature de l'erreur". Il indique des situations dans lesquelles j'utilise un métalangage (en général français) pour nommer ce qui n'est pas conforme. Dans l'exemple ci-dessous, choisi dans le "Commentaire dans le livre d'or, V1" de Chloé, la première annotation indique la nature de l'erreur, la deuxième ne l'indique pas.

Texte de Chloé :	**Mes annotations :**
wir haben museum besucht	Rajoute un déterminant.
und haben neue mangas entdecken.	Regarde la ligne de concordance ci-dessous et corrige.

Type d'annotation	Total des interventions	Type d'intervention		
		faible	moyen	fort
Je souligne ou surligne un passage.	65	x		
Je souligne ou surligne un passage et j'invite l'apprenant à regarder dans le corpus ou la concordance.	13	x		
J'invite l'apprenant à observer une partie précise du corpus (début ou fin des sections).	4		x	
Je nomme des lexèmes ou segments français dont l'apprenant est censé repérer l'équivalent allemand dans la concordance.	12		x	
J'indique à l'apprenant un mot clé précis dont il est censé observer le co-texte.	104		x	
J'invite l'apprenant à rajouter quelque chose en précisant la nature de l'erreur ou l'équivalent français du rajout.	12			x
J'indique un mot clé ou un passage précis d'un corpus et j'indique la nature de l'erreur.	19			x
Je souligne ou surligne un passage, et je présélectionne des lignes de concordance.	16			x
J'indique la nature de l'erreur et je présente des lignes de concordance présélectionnées.	5			x
J'utilise un métalangage français pour indiquer la nature de l'erreur ou le procédé de révision que je conseille.	186			x
J'invite l'apprenant à chercher un lexème ou segment à l'aide d'un dictionnaire électronique, d'un traducteur, d'un moteur de recherche ou d'une leçon.	13			x
J'invite l'apprenant à supprimer, déplacer ou remplacer un passage précis.	35			x
J'invite l'apprenant à remplacer un passage précis par un autre passage précis que je nomme.	20			x
J'impose ou j'interdis l'utilisation d'un outil particulier d'aide à la rédaction.	1			x
Total des interventions	**505**	**1 annotation tous les 7,3 mots** (pour 3687 mots produits par le groupe classe)		
Total des interventions à degré faible	**78**	**= 15,44 %**		
Total des interventions à degré moyen	**120**	**= 23,76 %**		
Total des interventions à degré fort	**307**	**= 60,79 %**		

Tableau 59 – Types d'intervention lors de l'annotation des textes de *Projekt Prospekt*.

Le tableau ci-dessus illustre le degré d'intervention très élevé pendant *Projekt Prospekt*. Tous les 7,3 mots, nous trouvons une annotation. 60,79 % des annotations peuvent être classées

dans la rubrique "intervention forte". Dans 186 annotations sur 505 (ce qui équivaut à 35,71 % des annotations), j'ai utilisé un métalangage français, souvent très normé.

Devant ce constat, qui va à l'encontre de mes intentions, j'ai entrepris une étude équivalente de mes annotations des productions publiées pour le projet *Brieffreunde.org* sur le blogue *esldeutsch*. Ce projet a été proposé à 33 apprenants, faisant partie de deux groupes différents (3ème LV1 et 3ème LV2), pendant la troisième phase de la recherche-action. Le graphique ci-dessous compare certains des résultats concernant les annotations obtenus pour *Projekt Prospekt* et *Brieffreunde.org*. Le tableau complet pour *Brieffreunde.org* peut être consulté dans les annexes (cf. VII.2.). Afin de pouvoir comparer les chiffres concernant les deux projets, j'ai multiplié les chiffres de *Projekt Prospekt* par 1,0412, ce qui permet de compenser la différence de taille des deux corpus d'apprenants (3687 mots *versus* 3839 mots).

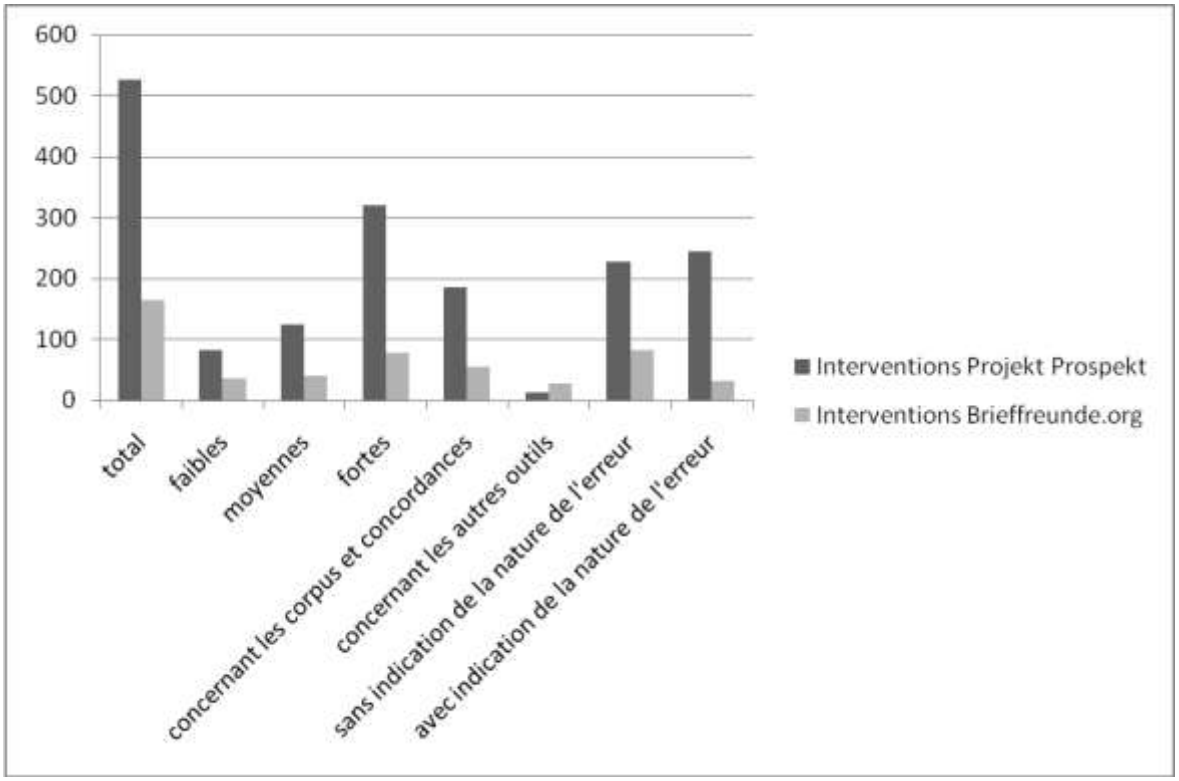


Figure 75 – Types d'intervention dans les annotations pour *Projekt Prospekt* et de *Brieffreunde.org*.

Le graphique ci-dessus compare, entre autres, le nombre d'annotations qui encouragent les apprenants à utiliser les corpus et concordances et le nombre d'annotations qui renvoient à d'autres outils d'aide à la rédaction. J'y indique également le nombre d'annotations qui mentionnent ou qui ne mentionnent pas la nature de l'erreur. Je constate qu'il y a un grand écart concernant le nombre total des annotations : tandis que pour *Projekt Prospekt*, j'ai relevé 525,8 passages à réviser, dans *Brieffreunde.org*, ce chiffre s'élève à 164. Une annotation équivaut en moyenne à une séquence de 7,3 mots pour le premier projet et à une séquence

moyenne de 23,66 mots pour le suivant. Lors du projet de la troisième phase de la recherche, en proportion, j'encourage, davantage que pour le projet précédent, à utiliser d'autres outils d'aide que le corpus ou la concordance. Cela s'explique par le fait que le corpus que j'ai créé pour *Brieffreunde.org* ne contient que 643 mots, ce qui s'est avéré insuffisant dans certaines situations de recherche d'informations. Les corpus de référence que j'ai proposés lors de *Projekt Prospekt* contiennent entre 2996 et 8980 mots. C'est pour cela qu'il a, moins souvent que pour *Brieffreunde.org*, été nécessaire de suggérer des recherches à l'aide d'autres concordances ou outils. Lors du deuxième projet, j'ai moins souvent indiqué la nature de l'erreur que lors du premier projet.

	Projekt Prospekt	***Brieffreunde.org***
Nombre total des annotations	525,8	164
Interventions faibles	15,44 %	21,95 %
Interventions moyennes	23,76 %	24,52 %
Interventions fortes	60,79 %	47,56 %

Tableau 60 – Degrés d'intervention lors de *Projekt Prospekt* et de *Brieffreunde.org*.

À première vue, le tableau ci-dessus montre relativement peu de variations. Toutefois, la somme de l'augmentation des interventions faibles et de la baisse des interventions fortes équivaut à 19,74 %, ce qui peut être considéré comme une évolution nette vers moins d'interventions directes de ma part.

Mon observation des annotations des productions fait ressortir une différence fondamentale. Dans les annotations rédigées pour *Projekt Prospekt*, je m'adresse rarement aux apprenants en allemand. J'ai recensé des annotations contenant des énoncés rédigés en allemand dans seulement neuf productions, et ces textes ont été produits par six apprenants sur seize. Aucune production n'est entièrement annotée en allemand. Les annotations formulées pour *Brieffreunde.org* contiennent une grande proportion d'énoncés allemands, comme l'illustre l'exemple suivant.

> *Eve (V0), danke für deinen Text. Korrigiere noch diese Sätze ("..."): "St Leu la Foret Stadtrand von Paris" (Ich liebe es zu tanzen) "und theater" Bis Montag!* [Eve (V0), merci pour ton texte. Corrige encore ces phrases (…). À lundi. Mme Lacroix].

Cette différence d'utilisation de la langue étrangère s'explique par plusieurs facteurs. Pour *Brieffreunde.org*, j'ai explicitement pris la décision de rédiger le plus possible en allemand. Pourtant, même au sein des annotations de ce projet, il y a des apprenants qui reçoivent

davantage de remarques formulées en français que d'autres. Il y a premièrement une certaine tendance à la lassitude : rédiger en français va plus vite que de rédiger en allemand, compte tenu du fait que je m'adresse aux apprenants dans un allemand qui doit être adapté à leur niveau. Deuxièmement, plus un apprenant utilise le traducteur comme aide à la rédaction, moins il suffit de relever certains lexèmes ou segments que l'ont peut observer dans la concordance. La production de Ronald (2008/09) et mes annotations en sont un bon exemple.

> Guten Tag heiße ich Ronald.Ich bin 15 Jahre alt, ich wohne in Frankreich, hat heiliger Saint-leu la Foret. Ich spiel gehn hat den Computer, den wii, und mächtig die andern Spiele. Ich suche eine Person, das habe, hat pres dasselbe Jahre alt, sogar Freizeit, und gouts sogar weinig.
>
> Ronald (V0), le traducteur fait des dégâts... Essaie de rester simple, et si tu ne trouves pas de solution par toi-même, puise dans la concordance et travaille avec LEO. Je te marque les passages à revoir :
> "Guten Tag" (pas un début d'annonce ; cherche dans le corpus).
> "heiße ich Ronald" (ordre des mots à revoir)
> "hat heiliger Saint-leu la Foret." (le traducteur traduit même les villes... ce n'est pas la peine).
> "Ich spiel gehn hat den Computer" (regarde ce qu'Axel a mis).
> "den wii" (c'est féminin).
> "mächtig" (laisse tomber ce mot.
> "Ich suche eine Person" (cherche le mot clé "SUCHE" dans la concordance pour remplacer "Person").
> "das habe, hat pres dasselbe Jahre alt" (cherche une tranche d'âge qui te va dans la concordance et copie la ligne).
> "sogar Freizeit" (le traducteur se trompe royalement ici. Cherche des lignes avec "Hobby" dans la concordance. Pour remplacer "sogar", cherche "les mêmes" avec LEO).
> "und gouts sogar weinig." (laisse tomber cela).
> Frau Lacroix
> Ecrit par : Frau Lacroix an Ronald (V0) | 07.10.2008

Un dernier facteur qui semble conditionner mon utilisation du français ou de l'allemand dans les annotations est la complexité de la tâche. Le genre textuel "Annonce pour trouver un correspondant" ne demande pas autant de préparation linguistique et discursive que les genres textuels proposés pour *Projekt Prospekt*. Les textes produits pour *Brieffreunde.org* sont en moyenne plus courts que ceux rédigés pour *Projekt Prospekt* : les annonces contiennent entre 10 et 69 mots, et la moyenne est de 32 mots. Pour *Projekt Prospekt*, les textes vont de 21 à 166 mots, et la moyenne est de 76,81 mots par production. La longueur et complexité des textes va de pair avec la nécessité de rédiger des annotations plus élaborées.

Une autre évolution que je souhaite relever est liée à l'utilisation de l'outil de publication des productions et annotations. Pour diffuser les documents concernant *Projekt Prospekt*, nous avons utilisé une boîte courriel, et rarement le blogue. Les apprenants qui m'ont envoyé des

courriels ont été accueillis par moi par des formules de prise de contact en allemand, et j'ai quelques traces qui montrent que me suis adressée à certains apprenants (Camille, Géraldine, Olivia, Alexa, Christine, Simon) uniquement en allemand dans certains courriels. Toutefois, les annotations des textes ont été formulées dans un fichier à part, et parfois, le ton paraît plutôt impersonnel, comme l'illustre l'exemple suivant : "à modifier en regardant dans la concordance ci-dessous" (annotation pour Fernand, VG V3). Pour *Brieffreunde.org*, les apprenants ont publié leurs productions sous forme de commentaires sur le blogue *esldeutsch*, et j'ai répondu sous forme de commentaire, parfois sous forme d'article complet. Ce choix de moyen de communication a entraîné une forme d'échange plus "direct" entre les apprenants et moi.

	Impératif	**Pronom ou adjectif possessif à la 2ème personne**	**Formulation impersonnelle**	**Compliment, encouragement ou remerciement**	**Formule d'ouverture ou de clôture**	**Total**
Nombre d'occurrences dans mes 67 commentaires	135	25	27	59	15	**261**
Pourcentage annotations	51,72 %	9,57 %	10,34 %	22,6 %	5,74 %	

Tableau 61 – Prise en compte du lecteur de mes annotations dans *Brieffreunde.org*.

J'ai 15 fois utilisé des formules d'ouverture et de clôture, et j'ai cité et annoté les passages à revoir en m'adressant directement aux apprenants, à la deuxième personne ou sous forme d'impératif. La fréquence de publication lors de *Brieffreunde.org* (157 commentaires publiés entre le 1er et le 14 octobre, par 33 apprenants et par moi-même) et la qualité linguistique et discursive des productions finales m'encourage à poursuivre dans la voie de développement de projets pour lesquels *Brieffreunde.org* peut représenter un prototype.

6.8.2.2. Formulation d'items dans les questionnaires distribués aux apprenants

Suite à l'étude des questionnaires que j'ai distribués aux apprenants lors de *Projekt Prospekt*, j'ai pu observer un changement concernant ma pratique de communication écrite. J'ai distribué le premier questionnaire au bout de douze séances, le deuxième à la fin du projet, donc au bout de trente-deux séances. Le premier questionnaire est conçu d'une façon qui traduit, à cette période de mes recherches, mon ancrage théorique et ma difficulté de trouver des termes susceptibles de permettre aux apprenants de comprendre les objectifs et le déroulement du

projet. Dans ce questionnaire, je m'adresse aux apprenants à la 1ère ou à la 3ème personne. Ce questionnaire ressemble davantage à un contrôle de connaissances qu'à un échange de type horizontal, tel que le défend l'esprit de la recherche-action. Comme l'illustre le début du questionnaire ci-dessous, il se compose d'énoncés comportant des lacunes que les apprenants sont censés remplir.

Le Projet Nous faisons un projet dans le domaine du ……. . En allemand, les élèves produisent trois textes appartenant chacun à un genre spécifique de ce domaine : …….d'un dépliant ou site touristique ; …….proposée dans un dépliant ou site touristique ; ……..d'un musée ou parc d'attraction.

Figure 76 – Questionnaire intermédiaire pour *Projekt Prospekt*.

Le premier questionnaire ne comporte pas de paragraphe introductif. Il offre aux apprenants l'occasion de s'exprimer librement en quelques lignes à la fin du document.

> Remarques concernant le projet
> Merci de noter ici les avantages et inconvénients du projet à ce stade tels que vous les ressentez. Ces informations ne donneront pas de note.

Le deuxième questionnaire, contient des questions semi-ouvertes. Ce type de questions donne aux apprenants l'occasion de s'exprimer plus librement que dans le questionnaire précédent. Un paragraphe introductif explique l'objectif.

> Tu as participé au *Projekt Prospekt*. Merci de répondre aux questions suivantes, en toute franchise, pour permettre d'ajuster les composantes du projet qui ont besoin d'être revues. Tu peux écrire des remarques dans les lignes qui ne proposent pas d'option afin de nuancer tes propos.

Ensuite, chaque paragraphe comporte une partie de type questionnaire à choix multiples et une partie qui permet aux apprenants de rajouter des remarques.

Quel type de commentaire as-tu préféré de la part du professeur ?

❑ Correction directe ("Écris x à la place de y").

❑ Explication grammaticale ou lexicale.

❑ Renvoi à une concordance.

❑ Renvoi à un corpus.

❑ Renvoi aux articles du blogue.

❑ Renvoi à des dictionnaires ou traducteurs.

Mes remarques :

Figure 77 – Questionnaire final pour "Projekt Prospekt".

L'analyse de ce questionnaire peut être consultée en annexe VIII.2.2.

6.9. Traces de décentration au fil de Projekt Prospekt

Après avoir décrit mon évolution en matière de guidage, je trouve utile de revenir sur l'évolution des attitudes des apprenants face à l'*input* en LE, et plus généralement, face à l'apprentissage de l'allemand. En 6.5., j'ai cité le verbe "se décentrer", proposé par Marin et Legros (2008) pour exprimer la prise de distance vis-à-vis la propre production écrite. Dans ce qui suit, je reprendrai le thème de la décentration, en le plaçant dans un contexte d'apprentissage plus général. Sur le site de la fondation Piaget, la notion de décentration est définie comme suit.

> Intervenant aux côtés d'autres mécanismes tels que l'abstraction réfléchissante, la décentration permet au sujet d'acquérir les constances perceptives, de construire l'objet permanent, l'espace sensori-moteur dans lequel il situe les objets les uns par rapport aux autres (corps propre y compris), puis, quelques années plus tard, d'acquérir les différents invariants de la pensée concrète (la substance, le poids, le volume, le nombre, les classes éléments d'une classification logique, etc.). De manière la plus générale, la décentration est le mécanisme qui permet au sujet individuel ou collectif d'échapper à toute forme de subjectivité déformante, parce que "égocentrée" ou "sociocentrée", pour atteindre des formes variées d'objectivité dans le rapport au monde ou à autrui (Définition collective. Ducret, 2009).

Selon cette citation, la décentration permet aux êtres humains d'élargir le regard qu'ils portent sur les "objets" qu'ils rencontrent et d'en percevoir davantage de facettes qu'auparavant. La décentration peut alors être considérée comme un des moyens susceptibles de mener à l'apprentissage de LE. Dans ce qui suit, j'examinerai le discours de certains des apprenants au sujet de l'énonciateur, donc de la personne qui transmet le message, dans des textes du projet.

Cette étude prépare à la suivante dans laquelle je m'intéresse aux noms de scripteurs que les apprenants ont choisis lors de la rédaction de leur "Commentaire dans un livre d'or électronique".

6.9.1. "De vraies personnes qui parlent" ?

Dans le questionnaire final distribué aux apprenants à la fin de *Projekt Prospekt* (cf. annexe VIII.2.1.), les apprenants ont exprimé les raisons qui leur ont fait préférer la production d'un type de texte plutôt que celle d'un autre. Les arguments peuvent être classés dans les catégories suivantes.

Le type de production…

… nécessite (ou pas) d'avoir de l'imagination ;

… permet de dire/formuler ce qu'on pense ;

… permet de dire son avis ;

… est plus simple que les autres types de production ;

… permet de trouver du plaisir ;

… permet d'apprendre quelque chose.

Le tableau suivant présente les avis de cinq apprenants concernant les types de production qu'ils ont préférés. J'ai sélectionné les avis qui permettent de détecter des indices d'une modification de l'attitude grâce à l'exécution d'une des tâches. Les indices de changement, déstabilisation ou décentration sont marqués en gras. Les indices qui permettent de penser qu'il n'y a pas eu de décentration sont marqués en italique.

Prénom	**Avis concernant la production du "Commentaire dans un livre d'or"**	**Avis concernant la production de la "Visite guidée"**	**Décentration**	
			forte	**faible**
Camille		Même si **j'ai eu plus de mal à l'écrire au départ (je m'accrochais trop au français),** c'est celui dans lequel je trouve, **on a appris le plus de choses**.	x	
Fernand		J'avais l'impression de *me proposer de faire une visite guidée à moi-même dans mon propre musée*.		x

Léa	Ça m'a appris comment dire en allemand *notre avis concernant un musée qu'on a visité*.			x
Simon	*On peut donner son avis, s'exprimer*			x
Vicky	Parce qu'on pouvait *écrire des choses comme on les pensait* puisque **c'était censé être de vraies personnes qui parlaient** → style + direct.		x	x

Tableau 62 – Indices de décentration des apprenants dans le questionnaire final.

Fernand, Léa et Simon font correspondre le "moi" du scripteur et leur "moi" psychologique. Ils ne semblent pas chercher à prendre de la distance en rédigeant.

> J'avais l'impression de me proposer de faire une visite guidée à moi-même dans mon propre musée (Fernand).

D'ailleurs, comme je l'illustrerai dans le paragraphe suivant, Fernand et Simon choisissent leur vrai nom pour désigner le scripteur du "Commentaire dans un livre d'or électronique".

Vicky s'exprime d'une façon qui peut paraître ambivalente. D'un côté, elle dit apprécier le fait de pouvoir écrire "comme on le pensait". Le "moi" n'est pas très loin du pronom "on" qu'elle utilise. D'un autre côté, une mise à distance de sa propre personne peut être vue dans la formulation "c'était censé être de vraies personnes qui parlaient". Le va-et-vient entre son moi et le moi du scripteur peut être observé entre les noms de "*Schühler*" [élève] et "*Spielmann*" [homme qui joue, ménestrel] qu'elle a choisis pour sa toute dernière rédaction.

Camille décrit un changement qu'elle a observé concernant sa façon de formuler pendant l'exécution de *Projekt Prospekt*. Au début, elle dit qu'elle s'est trop "accrochée au français" et qu'à la fin, elle avait appris "quelque chose". En effet, la première version de la critique de film que Camille a rédigée juste avant "*Projekt Prospekt*", contient des lexèmes et segments français.

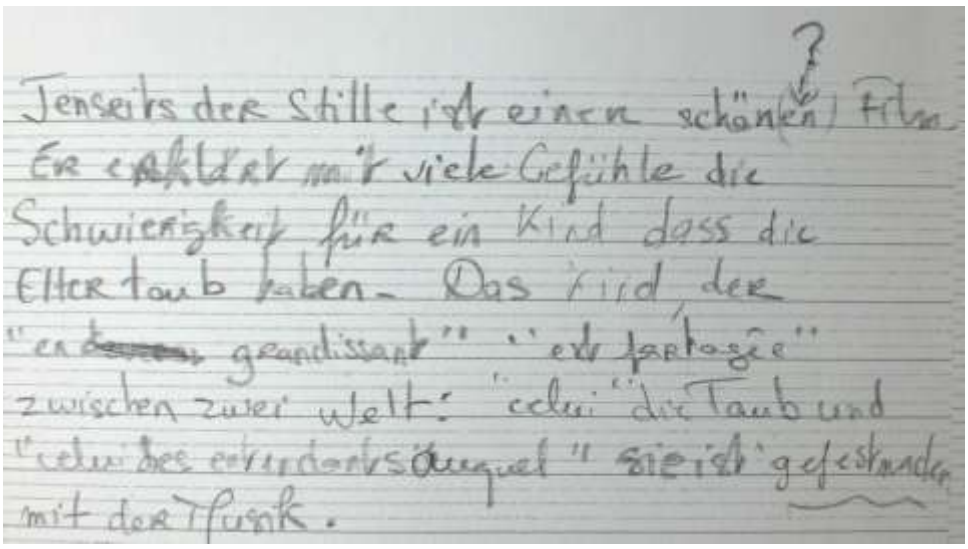

Jenseits der Stille ist einen schönen Film.
Er erklärt mit viele Gefühle die
Schwierigkeit für ein Kind dass die
Eltern taub haben. Das Kind, der
"en grandissant" "est partagée"
zwischen zwei Welt: "celui" die Taub und
"celui des entendants" "auquel" sie ist gestanden
mit der Musik.

Figure 78 – "Critique de film" de Camille.

Camille ne s'est pas appuyée sur le corpus "Critiques de film" pour écrire ce texte. La version suivante de ce texte porte encore des traces de nativisation (cf. 2.2.2.3.). L'avant-dernier texte que Camille a rédigé pour *Projekt Prospekt*, à l'aide de corpus et concordances, entre autres, est pleinement acceptable du point de vue discursif et linguistique.

> *Sonia Derrick (-Keine URL-) schrieb am 29.05.2008 um 12:30 Uhr:*
>
> *Hallo!*
> *Wir erlebten gestern einen tollen Tag auf Sablones Wegen! Und das Museum ist sehr schön!!!!!!!! Aber nach dem Spaziergang waren wir todmüde! Könntet ihr den Kindern ein Vesper geben? Die Führung war sehr gut! Danke!*
>
> *antwort : danke für das lob und die kritischen anmerkungen. wir wollen gerne in die Führung ein Vesper integrieren.*
>
> [Sonia Derrick (-Keine URL-) écrivit le 29/05/2008 à 12:30 : Salut ! Hier, nous avons passé une journée superbe dans les sentes de Sablone ! Et le musée était très beau !!!!!!! Mais après la promenade, nous avons été morts de fatigue ! Pourriez-vous donner un goûter aux enfants ? La visite a été très bien ! Merci !
>
> réponse : merci pour le compliment et les remarques critiques. nous sommes prêts à proposer un goûter lors de la visite].

Je terminerai ce chapitre par la présentation d'une étude sur des noms de scripteurs qui est censée donner un éclairage supplémentaire sur la façon dont *Projekt Prospekt* a "affecté", pour ne pas dire "transformé", les apprenants.

6.9.2. Jeux d'identité

Le tableau ci-dessous recense les prénoms et/ou noms de scripteurs, choisis par quatre des seize apprenants lors des deux dernières versions des "Commentaires dans un livre d'or", rédigées pour *Projekt Prospekt* (cf. annexe VI pour le tableau regroupant les résultats du groupe classe complet). Je classe les prénoms et/ou noms dans les catégories "décentration forte, moyenne ou faible". Un nom à consonance autre que française correspond ici à une décentration forte, un nom mi-français correspond à une décentration à degré moyen, et un nom français est classé dans la catégorie de la décentration faible. Je mets une croix distincte pour chacun des noms qui apparaissent dans les textes.

Prénom	**Commentaire dans le livre d'or (dernière version)**	**Commentaire dans le livre d'or (avant-dernière version)**	**Décentration**		
			forte	**moyenne**	**faible**
Camille	"Anna Derrick"	"Sonia Derrick"	x x		
Georges	"Marian prauf"	pas de nom		x nom français ? prénom féminin	x
Olivia	Olivia a indiqué l'initiale de son vrai prénom et son vrai nom complet pour nommer le "visiteur" du musée. Elle a choisi un nom anglais pour le "webmaster" du livre d'or : "Lenny Barred".	"Lenny Barred" (nom du visiteur qui écrit le commentaire).	x anglais		x propre nom et prénom
Total pour le groupe classe			**10 fort** (dont 9 allemands, 2 anglais)	**5 moyen**	**11 faible**

Tableau 63 – Noms de scripteurs de "Commentaires dans un livre d'or".

Dix noms, sur les vingt-cinq noms sélectionnés par les quatorze apprenants, renvoient à une attitude de décentration de l'auteur. Huit noms ont une consonance allemande plus ou moins marquée. Ils ont été proposés par six apprenants différents, à savoir par Camille (une fois), Géraldine (une fois), Janine (une fois), Jean (une fois), Léa (deux fois) et Vicky (trois fois). Trois apprenants ont choisi des prénoms et/ou noms clairement germaniques (Camille, Jean et

Vicky). Trois autres ont gardé leur propre prénom et/ou nom de famille (Fernand, Olivia et Simon), ce que je considère comme un indice de valeur "zéro" pour une décentration. Deux apprenants ont choisi des noms à consonance anglaise (Corentin et Olivia). Dans l'avant-dernière version de son texte, Géraldine a utilisé un mot espagnol au sein de coordonnées complètes, écrites en français. Léa a utilisé deux prénoms qui pourraient être aussi bien français qu'allemand. Georges a tronqué un prénom féminin qui pourrait, lui aussi, être aussi bien allemand que français et qui rappelle la figure allégorique de la France ("Marian"). Le nom de famille qu'il a choisi, "prauf", ne paraît, a priori, ni allemand, ni français, et il commence par une minuscule. Prononcé à la française, nous obtenons /prof/. Vicky a appelé un des visiteurs "A. Schühler". Prononcé à l'allemande, nous obtenons "*Schüler*", l'équivalent du mot allemand pour "élève". Le vrai prénom de Vicky commence aussi par la lettre "A". Pour un de ses noms de famille, Géraldine a choisi "Fransis", ce qui ressemble à "Française". Janine semble avoir joué avec les consonances de son vrai prénom et nom. Elle a choisi comme prénom le pseudonyme du prénom que je lui ai attribué pour garantir l'anonymat de sa personne lors de la recherche. Georges et Jean ont fait de même dans un de leurs textes. Ce pseudonyme s'inspire d'ailleurs de leur vrai nom de famille. Jean et Georges utilisent une fois un prénom féminin, Vicky utilise deux fois un prénom masculin. Pour donner un nom au scripteur de son commentaire, Jean a opté pour "Brigitte Hochleiner", un segment qui apparaît dans le premier texte du mini-corpus "Visite guidée" auquel Jean a eu accès lors de la rédaction de son texte. "Brigitte" est un prénom féminin franco-allemand. Comme pour la dernière version de ses "Informations pratiques", Jean s'est appuyé sur un texte du corpus de référence, sans le modifier (mise à part l'omission de la lettre "t" du nom "Hochleitner"). Vicky est la seule apprenante qui semble avoir recherché à attribuer à un des noms qu'elle a choisis un sens qui corresponde au type de musée qu'elle a choisi. Il s'agit du "musée des instruments de musique". Le nom "*Spielmann*" qu'elle a utilisé peut être traduit par "homme qui joue" ou "ménestrel". Ce nom apparaît aussi dans des chansons populaires allemandes.

À l'occasion du choix du nom de scripteurs de commentaires, les apprenants se sont mis à jouer avec leur identité nationale ou sexuelle ou avec leur statut d'"élève". Certains ont fait preuve d'humour ou de dérision ("grascochon", "Jean bonneau"). J'ai aussi détecté des clins d'œil aux références "culturelles" françaises ("Marian") et allemandes ("Derrick").

Le choix des prénoms et noms fictifs pour les "Commentaires dans un livre d'or" semble illustrer plusieurs modes d'adaptation des apprenants au nouvel environnement linguistique

que représentent les corpus. Comme esquissé en 2.7.1., l'apprentissage est un processus lors duquel on s'efforce à "conserver sa structure en lui assimilant les aliments tirés du milieu extérieur" et en même temps à "accommoder cette structure aux diverses particularités de ce milieu" (Piaget, 1969 : 208). Olivia a gardé son vrai nom pour désigner le scripteur dans son dernier "Commentaire dans un livre d'or", et le webmestre de ce commentaire porte un nom anglais. Il n'est pas inintéressant de noter qu'elle a particulièrement "résistée" aux méthodes de travail que j'ai proposées. Elle a, par contre, deux fois manifesté sa forte envie de communiquer lors de l'entretien, donc en français et en son nom propre. Le 31 janvier 2008, j'ai noté dans mon journal de bord qu'"Olivia a insisté sur son envie de s'exprimer face à l'enquêtrice". Lors de l'entretien du 6 mai 2008, parmi les 15 apprenants enquêtés, c'est elle qui a le plus souvent pris la parole (74 tours de parole, la moyenne étant de 39,5 ; cf. annexe XI.2.). Un parallélisme semblable peut être démontré pour Fernand. J'ai observé la situation contraire avec Jean, grand silencieux lors de l'entretien (18 tours de parole, avec une moyenne de 4,944 mots par tour de parole). Cet apprenant a choisi d'un patronyme masculin français pour le scripteur de son avant-dernier "Commentaire dans un livre d'or". Lors de la dernière production, il bascule vers un patronyme féminin allemand, identifié dans un texte de corpus. La situation intermédiaire, donc un choix "franco-allemand", existe également : Géraldine a choisi le prénom allemand "Martha" et un nom qui paraît être la forme germanisée du mot "Française". Les choix de noms de scripteurs semblent illustrer la façon dont les apprenants ont tenté de construire des passerelles entre le "terrain français" et le "terrain allemand" afin de pouvoir traverser les flots (linguistiques ? culturels ? conceptuels ?) qui relient les deux, sans risquer de s'y perdre.

6.9.3. Écrire en allemand, une "dimension nouvelle de la parole" ?

> Là se trouve de la façon la plus apparente la contradiction fondamentale qui caractérise (...) la façon dont on apprend à écrire à l'école : écrire est enseigné comme une habitude motrice et non comme une activité culturelle complexe. Pour cette raison (...), la nécessité d'une pratique vivante de l'écrit s'impose, comparable à la nécessité d'une pratique vivante de l'arithmétique. Ceci signifie que l'écrit doit faire sens pour l'enfant, il doit être suscité par un besoin propre, inclus dans une tâche vivante essentielle pour l'enfant. Nous serons alors persuadés que l'écrit se développe chez l'enfant non pas comme un ensemble d'habitudes de la main et des doigts mais comme une dimension authentiquement nouvelle et complexe de la parole (Vygotski, 1931 : 145, cité dans Brossard, 2004 : 100-101).

Un des résultats annexes de *Projekt Prospekt* est la "banalisation" de l'utilisation de la langue allemande, donc son utilisation pour des situations d'échange verbal lors desquelles on

s'exprime en son propre nom. Pendant *Projekt Prospekt*, je me suis régulièrement adressée aux apprenants en allemand. J'ai débuté presque toutes les séances en présentiel par quelques remarques dans cette langue, et je l'ai utilisée de façon privilégiée lors de la rédaction de courriels et d'articles sur les blogues. Les apprenants ont en général choisi de s'adresser à moi en français, à part quelques formules de prise de contact, comme "*Hallo!*" [Salut!] ou "*Bis bald*!" [à bientôt]. Le 4 mai 2008, Olivia m'a envoyé un courriel qui contient plusieurs formules allemandes. Elles sont marquées en majuscules dans le texte.

Date: Sun, 04 May 2008 19:43:21 +0200

De: Thomas Thomas <Thomas.Thomas@clg-landowska-st-leu.ac-versailles.fr>

À: Eva Lacroix <prall@clg-landowska-st-leu.ac-versailles.fr>

Objet: *HILFE !!!!!!!* [au secours]

> bonjour , c' est olivia , je reviens d' Allemagne , je suis allée avec l' Ecole de Musique de Saint Leu (Camille et Vicky étaient là ...) et là je viens de voir que je dois corriger mes infos pratiques ,mais je l' ai déjà fait et je suppose que vous m' avez répondu sur la feuille mais je suis désolée , (*ES TUT MIR SO LEID*) , et je viens de me rappeler que vous l' aviez disrtibuée donc s" il vous plait , *BITTE* , est ce que vous pourriez m' aider , je n' étais pas du tout là pendant les vacances , et là où j' étais , je n' avais pas internet (avant de partir en Allemagne) donc je n' ai pas vu bref , merci d' avance , je vous redis que je suis vraiment désolée et j' éspère que vous verrez mon mail avant demain : *TSCHÜSS* !!!!!!

Olivia a choisi de rédiger son courriel majoritairement en français. Elle y a joint quatre passages en allemand, dont un qu'elle met entre parenthèses, en tant que traduction d'un énoncé français. Le 30 mai 2008, je relate dans mon journal de bord qu'une autre apprenante, Géraldine, a répondu à un de mes courriels en rédigeant tout son texte en allemand. Ces exemples permettent d'entrevoir qu'un scénario pédagogique encourageant les apprenants à s'approcher de près de l'*input* en LE est susceptible avoir des conséquences sur l'acceptation de la langue étrangère comme langue d'échange dans des situations dites "ordinaires".

6.10. Tâches constructivistes ?

Bruner (1983 : 266 ; cf. 2.7.2.) décrit les conditions qu'une tâche d'apprentissage[40] doit remplir afin d'être conforme à des théories constructivistes. Voici un rappel de quelques-unes de ces conditions. La tâche doit être suffisamment complexe pour permettre le développement

[40] Bruner utilise le terme de "tâche" dans le sens d'activité d'apprentissage. J'applique les conditions qu'il énumère à des activités qui se rapprochent en certains points de la définition de "tâche sociale" (cf. 2.8.1.2.).

de l'apprenant et la "modification de sa conduite". Elle doit contenir une structure sous-jacente répétitive afin que l'expérience acquise puisse être appliquée à une activité ultérieure, et elle ne doit pas être totalement au-delà des possibilités des apprenants. Il serait favorable qu'elle soit "excitante et provocante".

6.10.1. Complexité, répétitivité et niveau des tâches

Les tâches, ont-elles été suffisamment répétitives et complexes, sans dépasser les possibilités des apprenants ? J'ai veillé à les construire selon le principe d'une graduation qui a donné aux apprenants l'occasion de se familiariser d'abord avec les outils et de produire, au fil du projet, du texte en ayant une liberté grandissante au niveau de la formulation.

Première production écrite

Pour préparer la rédaction des "Informations pratiques", les apprenants ont d'abord manipulé les outils et les données avant de passer au stade de la production. La formulation était censée être délibérément peu "créative". La reprise de segments dans l'*input* était motivée par le caractère très conventionnel des termes du genre textuel "Informations pratiques". Les constructions linguistiques n'étaient pas très complexes : les textes sont constitués de listes et de constructions infinitives, davantage que de constructions de type subordonnées, par exemple.

Deuxième production écrite

La "Visite guidée" a permis une plus grande créativité dans le domaine de la formulation. Les contraintes collocatives étaient moins fortes que pour le premier genre textuel. Toutefois, les sections du texte ont posé un cadre qui a conditionné les formes linguistiques utilisées par les scripteurs. Les constructions linguistiques étaient plus complexes que lors de la première tâche.

Troisième production écrite

Le "Commentaire dans un livre d'or électronique" a permis une liberté encore plus grande que la "Visite guidée" dans le domaine de la formulation. Les sections et les normes linguistiques sont moins standardisées que pour les genres textuels des productions précédentes. Les apprenants avaient l'occasion d'utiliser plusieurs temps et modes. La complexité des constructions a été modulable : elle dépendait des choix faits par le scripteur.

La partie "Informations pratiques" est un élément répétitif dans mon scénario. Elle pouvait trouver sa place dans les trois types de textes.

Certaines composantes des tâches demandaient des aptitudes linguistiques et cognitives ou un savoir-faire dans le domaine de la littératie électronique. Les apprenants n'en disposaient pas tous au moment de l'exécution du travail. Cela a eu comme conséquence qu'ils n'ont rendu aucun texte (Louise), qu'ils ont utilisé de façon trop intense le traducteur électronique (Georges, Véronique), ou qu'ils ont travaillé à un rythme particulièrement lent (Fernand). La présélection de lignes de concordance, leur nettoyage préalable, l'utilisation d'un corpus parallèle et le travail collaboratif sont des solutions susceptibles de soutenir les apprenants ayant un niveau peu avancé dans un des domaines évoqués.

6.10.2. Tâches "excitantes et provocantes" ?

À certains moments de la recherche-action, Olivia, Perrine, Véronique, Marinette et Cathy ont exprimé qu'elles n'avaient pas "envie" d'utiliser les concordances. D'autres apprenants, comme Chloé et Bastien, ont clairement exprimé leur enthousiasme vis-à-vis de cette aide à la rédaction. Le plaisir que l'on éprouve en abordant un corpus est lié à la possibilité de pouvoir en tirer profit, comme le montre l'attitude positive de Marinette pendant une séance de révision en petit groupe (cf. 6.3.2.3.). Cette possibilité dépend, entre autres, du type de concordancier qui est utilisé.

Les apprenants se sont particulièrement impliqués lors de certaines activités préparatoires lors desquelles ils n'avaient pas à rédiger, comme le surlignage de sections de textes ou l'attribution de binettes exprimant des avis de visiteurs d'un musée (cf. 5.3.7.3.1.).

Certains participants se sont montrés attirés par le travail collaboratif. Cette forme de travail n'a pas été constitutive pour ma recherche. Toutefois, j'ai favorisé les initiatives qui allaient dans ce sens – Géraldine et Olivia ont coproduit leurs critiques de film –, et j'ai parfois organisé la collaboration quand certains apprenants avaient besoin d'être particulièrement soutenus (cf. 6.6.2.).

Alexa et Camille ont décrit l'aspect interdisciplinaire du projet comme motivant, malgré le fait qu'il n'y a pas eu beaucoup de collaboration entre les trois collègues (cf. 6.7.).

6.10.3. Pour résumer

Les paragraphes précédents ont traité des effets du scénario pédagogique sur l'enseignante que je suis et sur les apprenants. Il apparaît que *Projekt Prospekt* a été un projet très complexe qui a nécessité, et pour certains, entraîné la modification d'habitudes de travail et de communication. Un nouveau mode de relation entre les apprenants et moi-même a pu s'établir, laissant aux apprenants suffisamment de liberté d'action pour qu'ils soient en mesure de (co-)construire leur apprentissage. *Brieffreunde.org* a permis de tester les idées principales de *Projekt Prospekt* dans une situation d'enseignement/apprentissage que l'on peut appeler "ordinaire". Il serait intéressant de construire un projet ayant des caractéristiques de complexité se situant entre *Brieffreunde.org* et *Projekt Prospekt*, donc entre le niveau A2 et B1 du *Cadre Européen Commun de Référence pour les langues* (2001), afin de pouvoir travailler dans des conditions acceptables en termes d'effort cognitif et de temps à investir. L'intégration d'un volet de travail conceptualisant pourrait contribuer à rendre les effets de tels projets plus "durables".

7. Où nous a menés l'exploration des corpus ?

Après avoir décrit l'arrière-plan théorique et méthodologique de la recherche et après avoir analysé les données, il est temps pour moi de constater où nous a menés la "route sinueuse" que l'on accepte de suivre quand on explore des corpus. Murison-Bowie (1993 : 46) a choisi l'équivalent anglais de cette collocation pour décrire où peut mener la recherche d'informations dans une concordance : on part d'une question qui est ensuite relayée par une autre. On finit par arriver à un endroit que l'on ne pouvait pas prévoir.

> *Whether one opts for putting up a case, or for knocking one down, any search using [a concordancer] is given a clearer focus if one starts out with a problem in mind, and some, however provisional, answer to it. (...). It will frequently be the case that you will want to formulate another question, which will start you off down a winding road to who knows where.*

Dans la séance de présentation du projet de la deuxième phase de la recherche, j'ai expliqué aux apprenants que grâce aux corpus, ils auraient accès à un choix de formulations qui sont possibles dans une situation donnée : "Là, vous voyez quand même qu'on a assez rapidement un éventail des choses qu'on peut dire." J'ai ainsi fait allusion à ce que Culioli (2002 : 143) appelle "cette capacité que nous avons de faire proliférer" (cf. 2.1.2.1.). J'ai expliqué aux apprenants que les corpus leur permettraient d'observer les normes attestées pour un genre textuel donné et de faire moins d'"erreurs" de langue. Les concordances les aideraient à gagner du temps lors de la rédaction. Les outils de la linguistique de corpus seraient aussi un moyen pour les rendre indépendants dans leur travail de révision. Pendant la séance de présentation, j'ai décrit au groupe une situation de l'année précédente lors de laquelle j'ai incité une apprenante à modifier elle-même un de ses passages non conformes aux normes fixées par le genre textuel. Je me cite : "J'aurais pu lui dire directement [à l'apprenant] : 'Il fallait le dire ça comme ça', mais ce qui m'intéressait, c'est qu'elle ait les techniques en main pour que, dans d'autres situations, elle le trouve sans moi." L'analyse de mes données atteste, en effet, de l'émergence d'une telle capacité auprès de certains des apprenants.

7.1. Avoir un autre regard sur l'input

Les corpus et concordances peuvent permettre d'observer des normes linguistiques et discursives de genres textuels ce qui contribue au développement d'un regard synthétique sur les données en LE.

L'exécution des activités d'entraînement concernant les sections des textes représentant un genre textuel (cf. 5.3.7.3.1.) a aidé certains apprenants à développer un regard plus approfondi sur l'*input* et à mieux comprendre le lien entre forme et sens (cf. 6.3.2.2. et 6.5.2.). D'autres apprenants ont gardé des habitudes de lecture exhaustive (cf. 6.2.2.1.). Ils ont donc eu du mal à voir l'intérêt des concordances qui isolent certains éléments dans une grande quantité d'*input*.

7.2. Disposer d'une aide à la formulation

Les corpus donnent accès, davantage que les dictionnaires, à un éventail de "formulations" pertinentes qui sont attestés pour un genre textuel donné.

Ce que j'ai appelé "formulations" lors de la séance de présentation peut être observé, entre autres, dans le domaine des collocations. Afin d'illustrer les limites des dictionnaires et de souligner l'apport des outils de la linguistique de corpus, je présenterai dans ce qui suit quelques observations faites lors de la simulation globale *Schweizer Dorfgemeinschaft* [Communauté de villageois suisses]. Ce projet a été proposé lors de la troisième phase de la recherche. Il n'a pas été basé sur l'emploi de concordances. Lors de l'exécution du projet, les apprenants ont fréquemment signalé leur difficulté à choisir entre plusieurs traductions proposées dans un dictionnaire, comme l'illustrent les deux exemples suivants : "*emeritiert*" [à la retraite, mais seulement en tant que professeur d'université], au lieu de "*in Rente*" [à la retraite] ou "*vielköpfig*" [polycéphal] au lieu de "*zahlreich*" [nombreux], etc. Un apprenant m'a demandé quel verbe conviendrait pour dire "un film va sortir", un autre comment on dit "donner une fête". Ces questions concernent les caractéristiques collocatives, un domaine encore peu traité par les dictionnaires. Pour les mots clés "Fest" [fête] et "Party" [boum], le dictionnaire *Frequency Dictionary of German* de Jones, Tschirner *et al.* (2006) indique les contextualisations suivantes : "*Fest : Weihnachten ist das Fest der Familie*." [Fête : Noël est la fête de la famille] et "*Party : die Party beginnt gegen 8 Uhr*." [La boum commence à 8 heures]. Pour le mot clé "Film", le dictionnaire propose l'énoncé "*Ich habe mir den Film auf DVD gekauft*." [Je me suis acheté le film sur DVD]. On est loin de la quantité d'occurrences collocationnelles potentiellement mise à disposition par les corpus. En particulier dans le domaine des segments, il s'avère donc utile de faire travailler les apprenants avec un corpus réfléchi et avec un concordancier facile d'emploi, comme par exemple *AntConc3.2.1w* d'Anthony (2007).

7.3. Arriver à écrire des textes peu nativisés

Les corpus et concordances peuvent permettre de faire moins d'"erreurs" dans certains domaines linguistiques et discursifs.

Alexa, Camille, Chloé, Fernand et Vicky disent dans l'entretien que les concordances les aident à faire moins d'erreurs. Dans l'entretien (annexe I.2.1.), Chloé précise en quoi les concordances l'aident lors de la rédaction. De sa dernière phrase, je conclus qu'elles la soutiennent dans le domaine de la dénativisation.

> On tape à l'ordinateur directement. C'est plus facile pour les copier-coller, des concordances et tout ça parce que ça limite le risque d'erreur. Si on écrit un texte en allemand comme ça sur papier [sans utiliser les concordances], ça… ça… déjà, il y a des fautes de grammaire peut-être et il y a des trucs qui ne se disent pas en allemand et qui se disent en français.

Camille et Vicky ont également tiré profit du corpus et des concordances en rédigeant des textes allemands de moins en moins nativisés. Camille a travaillé à distance pendant les vacances, avant le début des séances en présentiel, et elle a réussi à exploiter parfaitement les outils proposés en un temps "record". En raison du degré relativement faible de sa littératie électronique au début du projet, Vicky a eu besoin de plus de temps avant d'être à l'aise avec les nouveaux outils. Toutefois, les productions qu'elle obtient sont de grande "qualité" linguistique et discursive.

7.4. Progresser en langue étrangère

Les corpus et concordances peuvent permettre de progresser dans le domaine de la flexion des adjectifs, de la ponctuation et des mots composés.

Le paragraphe précédent s'est davantage prononcé sur les qualités du produit que sur l'apprentissage de la langue étrangère. Mon analyse des données a permis de retracer certaines évolutions du système linguistique des apprenants. Mes études concernant les adjectifs et la syntaxe (6.4.) font ressortir que pour le type de textes que le scénario pédagogique a prévu – des textes d'une longueur de 60 à 120 mots environ –, une évolution peut être davantage attestée pour la flexion des adjectifs que pour la syntaxe. Olivia a progressé dans le domaine de la ponctuation (cf. 6.5.2.). Certains apprenants ont inséré des segments repérés dans l'*input* sans que cela fasse sens dans le nouveau contexte. J'en conclus qu'ils n'ont pas traité les données repérées et que, par conséquent, il n'y a pas eu apprentissage de la LE. Le repérage

doit aller de pair avec un travail de réflexion métalinguistique, menant à la conceptualisation des phénomènes linguistiques observés. L'analyse des travaux de Jean (cf. 6.3.1.4.) et de Vicky (6.3.2.2.) permet de dire qu'une telle réflexion peut mener à une conceptualisation dans le domaine des mots composés.

L'étude présentée en 6.5.1. illustre les apports du scénario en termes de prise en compte de normes discursives concernant la ponctuation. Des recherches complémentaires sont nécessaires dans ce domaine qui n'a pas été prioritaire lors de ma recherche.

Le rôle du temps investi

Lors de la séance de présentation, j'ai dit que les corpus et concordances permettent de "gagner du temps" lors de la rédaction. Il convient de relativiser cette "promesse". L'exploitation des corpus nécessite un travail préparatoire permettant d'habituer les apprenants aux outils et méthodes. Le temps de préparation peut paraître long. Notons que selon Demaizière (2007), lors d'une formation, il faut parfois accepter d'investir du temps afin d'en gagner ultérieurement.

> On aide l'apprenant pour lui permettre de progresser. Cela signifie-t-il le faire "aller plus vite" et donc faire pour lui un certain nombre d'actions ? Comment gérer le temps qu'il peut être nécessaire de sembler perdre (ou laisser perdre) par rapport aux objectifs de la formation ? On sait que des enseignants ou tuteurs sont souvent inquiets de ce qui leur semble une progression trop lente, des pertes de temps ou un "papillonnage" entre ressources au moment où les apprenants découvrent un support ou une modalité de travail qui leur est inhabituelle, ceux-ci semblant passer trop de temps à trouver un mode de travail adéquat à l'environnement. (…) Aider c'est parfois laisser passer le temps qui permettra d'en gagner plus tard.

Le temps s'est d'ailleurs révélé un mot clé fréquent dans mon étude présentée en 6.1.1. Il est un des paramètres que Portine (1998) présente comme décisif pour le développement d'une attitude "autonome" de l'apprenant.

> L'espace de formation (…) est l'un des trois paramètres contextuels de la mise en place de procédures d'autonomisation de l'apprenant. Le deuxième paramètre est le tuteur (…). Le troisième paramètre est le temps. Dans la constitution de l'autonomie, le temps joue un rôle important car il doit être géré comme une composante du processus.

7.5. Réviser de façon plus indépendante

Les corpus et concordances peuvent contribuer à rendre les apprenants plus indépendants dans leur travail de révision.

Le fait d'employer des corpus et concordances n'a pas d'incidence directe sur l'"autonomie" des apprenants, pour reprendre le terme que j'ai utilisé lors de la séance de présentation. Portine (1998) définit ce terme comme suit.

> L'autonomie, c'est construire un projet d'action et gérer la réalisation de ce projet au sein d'une structure qui définit les contraintes globales et apporte une aide lorsqu'elle est nécessaire.

Mon scénario pédagogique n'a pas favorisé ce type d'autonomie. Par contre, il a visé à rendre les apprenants co-responsables de leur travail de révision : je les ai amenés à réfléchir à des solutions en s'appuyant sur des données attestées. Ce type d'indépendance peut être considéré comme une des caractéristiques d'un apprenant-chercheur. L'accompagnement offert aux apprenants lors de l'utilisation des outils joue un rôle crucial pour le développement d'une attitude d'apprenant-chercheur. Au fil du projet, j'ai développé une façon d'accompagner les apprenants susceptible de les aider à aller vers les corpus et concordances en ayant conscience de ce qu'ils y cherchaient, au lieu de les pousser à répliquer ma démarche de recherche. Dans le questionnaire final, Vicky décrit son cheminement comme suit.

> Je suis satisfaite du projet, car on a travaillé en autonomie. Au début, j'étais angoissée car on se dirigeait dans des directions différentes tous. Mais à la fin, j'ai développé mes propres méthodes d'écriture en adaptant ce que je savais faire ou pas.

L'exploitation directe, donc "indépendante", des corpus et concordances par des apprenants de niveau intermédiaire dépend aussi du type de ressources et d'outils d'interrogation dont on dispose. J'ai trouvé que pour l'allemand, dans ces domaines, le choix est moins large que pour l'anglais et que l'emploi d'outils conçus pour cette langue n'est pas forcément une solution acceptable. Des recherches complémentaires et des développements d'outils et de ressources sont nécessaires. Ces recherches pourront se baser, entre autres, sur l'utilisation des concordanciers *AntConc3.2.1w* (Anthony, 2007) et *Lexico3* (Fleury, Lamalle, W. Martinez et Salem, 2003) et sur l'emploi de la version franco-allemande de *Linguee* (Fink et Frahling, 2008) dont j'attends le développement futur. Dans un courriel que j'ai reçu le 13 août 2009 de Laura Alvarez, collaboratrice de *Linguee.com*, une telle évolution est évoquée.

> *Wir freuen uns, dass Linguee sich rasant verbreitet - mittlerweile erreichen uns täglich über 170.000 Suchanfragen. (...) In den nächsten Monaten*

stehen weitere Entwicklungen an. Dazu gehört der Ausbau unseres Community-Bereiches und die Einführung weiterer Sprachpaare.

[Nous sommes heureux du fait que *Linguee* se propage très rapidement – nous comptabilisons quotidiennement au-delà de 170.000 requêtes. (…) Dans les mois à venir, nous prévoyons des développements supplémentaires. L'élargissement de notre espace collaboratif et l'introduction de nouvelles combinaisons de langue en feront partie].

8. Destinations de rêve

8.1. Dynamiser les scénarios basés sur des corpus

Les projets de ma recherche ont privilégié le travail individuel. Lors de la première et de la troisième phase, les apprenants ont publié leurs productions sur des blogues. Cela a eu des effets positifs sur l'investissement et le rythme de travail en particulier lors du projet *Brieffreunde.org* (cf. 5.3.5. et 6.8.2.1.). Un de mes projets qui n'a pas été basé sur des corpus illustre particulièrement bien la dynamique qui peut être créée si l'on donne aux apprenants une liberté assez grande dans le domaine de la publication en ligne. Dans ce qui suit, je décrirai ce projet afin d'illustrer de quels éléments supplémentaires j'aimerais doter mes scénarios futurs, basés sur des corpus.

Comme déjà évoqué au chapitre 7, en juin 2009, j'ai proposé aux apprenants ayant participé à *Brieffreunde.org* la simulation globale *Schweizer Dorfgemeinschaft* [Communauté de village en Suisse]. Ce groupe d'apprenants en classe de troisième a été initié aux blogues et à l'exploitation de lignes de concordance dès la classe de cinquième. Le blogue créé pour le projet reflète la fréquence et la qualité d'échanges que l'on peut obtenir si l'on donne aux apprenants le statut d'auteur afin qu'ils puissent publier des articles et des documents multimédias de façon directe, sans être obligés de me les soumettre auparavant par courriel ou par un autre moyen de transmission électronique. La capture d'écran suivante illustre la façon dont certains se sont emparés des fonctionnalités du blogue. Une apprenante invite ses camarades à laisser des commentaires : "*Stell ein Kommentar für dich einzutragen*" [Mets un commentaire afin de t'inscrire]. Suite à cette invitation, trois commentaires ont été déposés.

Figure 79 – Échanger avec des pairs sur un blogue.

Kommentare	**Commentaires**
Danke !	Merci !
Ich will kommen	Je veux venir.
*Bis Morgen ^^*	À demain ^^
Ecrit par : Kekes | 18.06.2009	
Oh ! Das ist sehr gut für dich !	Oh ! C'est très bien pour toi !
Du gehst gut Geschäft gemacht !	Tu vas faire de bonnes affaires !
Bis Montag :)	À lundi :)
Ecrit par : Chaplin & Einstein | 18.06.2009	
Ja !!! Das ist der Anlass um zu gehen !	Oui !!! C'est l'occasion de venir !
Ich komme, ich komme, ich komme !	Je viendrai, je viendrai, je viendrai !
Ecrit par : Klaryssa | 18.06.2009	

En l'espace de six séances, les 33 apprenants ont rédigé 77 articles et 107 commentaires.

L'échange par courriel que j'ai eu en décembre 2008 avec Théo, un membre de ce même groupe, ayant des résultats variables en allemand, témoigne également de l'énergie qui peut

être mobilisée si un apprenant a l'occasion de présenter ses travaux à d'autres, soit à distance, soit en présentiel. Cette énergie semble encore être décuplée quand il se sent reconnu comme expert dans un domaine donné. Le 12 décembre 2008, j'ai noté dans mon journal de bord les remarques suivantes.

> Théo m'a envoyé un diaporama qu'il a fait sur l'histoire de l'Allemagne, à partir d'un document qu'il a publié sur *Wikipedia*. Il m'a proposé de le publier sur le blogue *esldeutsch*. Quand il m'a dit qu'il est auteur sur *Wikipedia*, j'ai montré beaucoup d'enthousiasme. Théo m'a ensuite envoyé encore deux diaporamas. (…) Je lui ai proposé de préparer un exposé pour le sujet d'histoire que nous allons aborder à partir du mois de janvier. Ceci l'a rendu très actif (…). Il est allé jusqu'à imaginer l'organisation de tout un cours, avec gestion de la classe (faire deux groupes, organiser un vote, etc.), et avec un programme plus qu'ambitieux au niveau du contenu.

Il serait intéressant d'exploiter la dynamique que l'on peut obtenir par la reconnaissance des apprenants, appuyée par la publication en ligne, pour des projets de plus grande envergure, impliquant des corpus et concordances.

8.2. Réfléchir à la langue étrangère

Les projets de ma recherche ont pris comme point de départ les genres textuels et l'observation des caractéristiques linguistiques des sections de texte. Le mouvement premier sur l'objet "texte" a donc été de type *top down* : j'ai invité les apprenants à observer d'abord un texte de façon holistique et à essayer ensuite de reconnaître ses parties, afin de voir comment ils pouvaient construire eux-mêmes des objets ressemblant à ceux qu'ils avaient observés. Comme je l'ai illustré en 4.2.4.2., la concentration sur des phénomènes linguistiques précis, avec prise en compte de l'alternance entre les mouvements de type *bottom up* et de type *top down*, pourrait contribuer à faire évoluer le système linguistique des apprenants.

Demaizière et Berthoud (1998) ont observé que si l'on ne donne pas de repères explicatifs grammaticaux aux apprenants, ils s'en donnent eux-mêmes, sans régulation. Elles précisent que les procédés suivants peuvent permettre de guider l'auto-construction de références : les enseignants désignent des marques retenues comme pertinentes. Ils donnent ensuite la parole aux apprenants afin qu'ils puissent expliquer ce qu'ils ont observé concernant ces marques.

Dans mon scénario, j'ai mis ces propositions en pratique en créant des activités d'observation de lignes de concordance (cf. 5.3.7.3.3.) et d'échange verbal entre binômes d'apprenants (4.2.4.2.) lors de la troisième phase de la recherche. Début mars 2009, j'ai proposé à deux

groupes classe une activité qui s'appuie également sur une mise en forme matérielle de l'*input* et qui permet d'illustrer les possibilités d'un échange métalinguistique autour des formes verbales dans un *input*. Afin de favoriser ce type d'échange, j'ai présenté aux apprenants un graphique fait à l'aide du logiciel *Wordle* (Feinberg, 2009) qui était censé les préparer à la lecture d'un texte assez long, *Die geteilte Stadt* [La ville divisée], suivi de *Der Brief* [La lettre], deux textes tirés de Kordon (1988) qui apparaissent, légèrement modifiés, dans le manuel d'allemand *Sag' mal* (Kahn *et al.*, 1993).

Figure 80 – Repérage de catégories linguistiques dans un graphique *Wordle*.

J'ai invité les apprenants à identifier, entre autres, les formes verbales qui apparaissent dans ce graphique. Certains apprenants ont classé le lexème "*geteilte*" [divisée] dans la catégorie "verbe". Ce mot fait partie de la classe des adjectifs dérivés d'un participe passé. L'erreur d'identification peut s'expliquer par les indices "*ge*..", préfixe des participes passés allemands, et "*teil*...", radical du verbe "*teilen*" [diviser]. L'indice "...*e*" est une terminaison qui peut caractériser des adjectifs, mais qui n'apparaît jamais à la fin des participes passés. Cet indice justifie le classement de "*geteilte*" dans la catégorie "adjectif". Un deuxième indice qui plaide en faveur d'un emploi adjectival de ce participe passé est son apparition dans le titre du texte ("*Die geteilte Stadt*") qui peut être lu au-dessus du graphique *Wordle* (Feinberg, 2009) dans sa version publiée en ligne.

L'exemple que je viens de décrire ne mentionne qu'une partie du travail de conceptualisation que l'on peut mettre en place avec des apprenants (cf. 2.1.2. et 2.8.2.). Il paraît judicieux

d'inclure une réflexion sur le rapport entre formes et valeurs, en s'inspirant, entre autres, d'Audin (2008) qui a développé une méthodologie donnant des pistes d'action passionnantes dans ce domaine. Mon exemple ci-dessus concernant l'observation de formes permet d'entrevoir que partir de l'observation d'éléments microsyntaxiques peut, en effet, ouvrir un dialogue métalinguistique entre apprenants et enseignants qui "en apprend énormément", pour reprendre une formulation de Culioli (2002 : 131) concernant le travail sur corpus (cf. 2.1.2.). Une prochaine étape de mes recherches poursuivra cette voie qui s'annonce prometteuse. Les concordances peuvent être un partenaire privilégié dans de telles entreprises.

Conclusions

Ma recherche a consisté à mettre en place un scénario pédagogique qui a fait appel à des corpus afin de soutenir l'apprentissage de la production écrite en LE. L'analyse de mes données a fait ressortir que les corpus et les outils d'interrogation ne représentent pas un kit d'apprentissage "prêt à l'emploi". Leur emploi nécessite un investissement en termes de temps et d'activité cognitive. Pourtant - ou justement en raison de cela -, le potentiel des corpus pour l'enseignement/apprentissage de LE est réel. Les productions peu nativisées qu'ont produites les apprenants ayant réussi à se familiariser avec les outils et les méthodes de travail en témoignent tout particulièrement. J'ai pu constater que l'exploitation des corpus n'est réservée ni aux adultes, ni aux apprenants ayant un niveau avancé en LE : les apprenants de niveau intermédiaire en tirent profit, à condition d'avoir accès à des ressources, des outils et un guidage adéquats.

Stevens (1990 : 8) énumère certains aspects concernant l'exploitation de corpus qui se trouvent confirmés par l'analyse de mes données : inciter les apprenants à se servir d'outils de linguistique de corpus et à manipuler les données de l'*input* de façon indépendante va de pair avec la remise en question des modes d'enseignement/apprentissage et des relations pédagogiques établis. Je me joins à Stevens pour souligner qu'il est indispensable de proposer aux enseignants de langue des actions de formation, afin de les aider à s'approprier les outils de la linguistique de corpus et les méthodes d'enseignement/apprentissage qui sont en accord avec leur utilisation. Lüdeling et Walter (2009 : 8) conseillent de mettre en place des ateliers destinés à cultiver auprès des apprenants d'allemand une littératie dans le domaine des corpus (*corpus literacy*). De telles actions peuvent tout à fait s'adresser non seulement à des apprenants de langue étrangère, mais aussi à des enseignants, en particulier dans le cadre de la formation initiale.

Les chantiers dans le domaine de l'enseignement/apprentissage de l'allemand à l'aide de corpus sont, en effet, nombreux. Ils ne demandent qu'à être investis par des chercheurs qui défendent le point de vue que l'accès direct aux données est un des éléments qui donnent aux utilisateurs l'occasion de cultiver une pensée indépendante.

Références générales

Tous les documents en ligne étaient accessibles le 19/08/2009.

Pour obtenir les coordonnées de mes blogues dont je n'ai pas indiqué l'adresse Internet, merci de me contacter en m'envoyant un courriel (evalacroix@free.fr).

Adam, J.-M. (1992). *Les textes, types et prototypes. Récit, description, argumentation, explication et dialogue*. Paris : Éditions Nathan ; Armand Colin, 2005[2].

Adam, J.-M. & Viprey, J.-M. (2008). "Appel à communication, n° 9 – 2009". *Corpus*, n° 7. Disponible sur : http://corpus.revues.org/document382.html

Aguado, K. (2002a). *Imitation als Erwerbsstrategie. Interaktive und kognitive Dimensionen des Fremdsprachenerwerbs*. Habilitation à diriger des recherches. Disponible en ligne : http://karin.aguado.de/publ/downloads/ImitationAlsErwerbsstragie.pdf

Aguado, K. (2002b). "Formelhafte Sequenzen und ihre Funktionen für den L2-Erwerb". *Zeitschrift für Angewandte Linguistik*, vol. 37. Frankfurt am Main, Berlin, Bern, Bruxelles, New York, Oxford, Wien : Peter Lang. pp. 27-49.

Alamargot, D., Lambert, É., & Chanquoy, L. (2005). "La production écrite et ses relations avec la mémoire". *A.N.A.E. (Approche Neuropsychologique des Acquisitions de l'Enfant)*, vol. 17. pp. 41-46.

Anthony, L. (2002). *AntConc1.0*. http://www.antlab.sci.waseda.ac.jp/antconc_index.html

Anthony, L. (2006). "Developing a Freeware, Multiplatform Corpus Analysis Toolkit for the Technical Writing Classroom". *IEEE Transactions on Professional Communication*, vol. 49, n° 3. pp. 275-286.

Anthony, L. (2007). *AntConc3.2.1w*. http://www.antlab.sci.waseda.ac.jp/antconc_index.html

Aston, G. (2002). "The learner as corpus designer". In Kettemann, B. & Marko, G. (dir.). *Teaching and Learning by Doing Corpus Analysis*. Amsterdam, New York : Éditions Rodopi B.V. pp. 9-26.

Axtell, T. (nd). *Concordancing in the CEGEPs*. http://pages.infinit.net/jaguar3/lounge/concord/default.htm

Audin, L. (2008). "Langues étrangères à l'école primaire : recherches INRP 1998-2003". In Chini, D. & Goutéraux, P. (dir.). *Psycholinguistique et didactique des langues étrangères. Travaux du GEPED en hommage à Danielle Bailly*. Paris : Ophrys. pp. 143-155.

Bailly, D. (1998). *Les mots de la didactique. Le cas de l'anglais. Lexique*. Gap, Paris : Ophrys.

Bange, P. (2003)."L'usage de la règle dans l'enseignement et l'apprentissage de langue étrangère". *Marges linguistiques*, septembre 2003. Saint-Chamas : M.L.M.S. éditeur.

Barbier, M.-L. (2003). "Écrire en L2 : bilan et perspective de recherches". *Arob@se 7*, vol. 1, n° 2. pp. 6-21. http://www.univ-rouen.fr/arobase/v7/barbier.pdf

Barbier, M.-L. (2004). "Écrire en langue seconde, quelles spécificités ?". In Piolat, A. (dir.). *Écriture. Approches en sciences cognitives*. Aix-en-Provence : Publications de l'Université de Provence. pp. 181-232.

Baron, G.-L. & Dané, É. (2007). "Pédagogie et ressources numériques en ligne : quelques réflexions". *EpiNet : Revue électronique de l'EPI (Enseignement Public & Informatique)*, septembre 2007. http://edutice.archives-ouvertes.fr/docs/00/27/78/25/HTML/a0709c.htm

Baroni, M. (2006). *DeWac.* Corpus interrogeable sur le site de *Sketch Engine* (Kilgarriff *et al.).*

Baroni, M. & Bernardini, S. (2003). *BootCaT toolkit 0.1.2. (Bootstrapping Corpora and Terms from the Web).* Outil de création de corpus. http://sslmit.unibo.it/~baroni/bootcat.html

Baroni, M. & Bernardini, S. (2004). "BootCaT: Bootstrapping Corpora and Terms from the Web". *Proceedings of LREC* 2004. Disponible en ligne : http://people.sslmit.unibo.it/~baroni/publications/lrec2004/bootcat_lrec_2004.pdf

Baroni, M. & Kilgarriff, A. (2006). "Large linguistically-processed Web corpora for multiple languages". In *Conference Companion of EACL 2006 (11th Conference of the European Chapter of the Association for Computational Linguistics)*. East Stroudsburg PA : ACL. pp. 87-90. Disponible sur : http://clic.cimec.unitn.it/marco/publications/eacl2006/dewac_eacl.pdf

Bazin, H. (2009). *Fondamentaux à toute recherche-action*. http://recherche-action.fr/pages/manual12/principes-generaux/fondamentaux-a-toute-recherche-action.php

Bereiter, C. & Scamardalia, M. (1983). "Does learning to write have to be so difficult?". In Freedman, A., Pringle, I. & Yalden, J. (dir.). *Learning to write: First language, second language.* New York : Longman.

Bereiter, C. & Scamardalia, M. (1987). "Knowledge telling and knowledge transforming in written composition". In Rosenberg, S. (dir.). *Reading, Writing and Language learning: Advances in applied* psycholinguistics, vol. 2. Cambrige : Cambridge University Press. pp. 142-175.

Bernardini, S. (2000). "Systematising serendipity: Proposals for concordancing large corpora with language learners". In Burnard, L. & McEnery, T. (dir.). *Rethinking language pedagogy from a corpus perspective*. Frankfurt am Main, Berlin, Bern, Bruxelles, New York, Oxford, Wien : Peter Lang. pp. 225-234.

Bernardini, S. (2004)."Corpora in the classroom. An overview and some reflections on future developements". In Sinclair, J. (dir.). *How to Use Corpora in Language Teaching*. Amsterdam, Philadelphia : John Benjamins Publishing Company. pp. 15-36.

Bernardy, J.-P., Haldenwang, N., Lozachmeur, F., Pavan, C. & Suissa-Le Scanff, P. (2009). *Gute Fahrt! Niveau A1/A2*. Paris : Nathan.

Berry, D. C. & Dienes, Z. (1993) : *Implicit learning : theoretical and empirical issues*. Hove, Hillsdale : Lawrence Erlbaum Associates.

BNC (British National Corpus). Corpus interrogeable sur : http://www.natcorp.ox.ac.uk/

Bommier-Pincemin, B. (1999). *Diffusion ciblée automatique d'informations : conception et mise en oeuvre d'une linguistique textuelle pour la caractérisation des destinataires et des* documents. Thèse de Doctorat en Linguistique. Paris : Université de Paris 4.

Bouchard, R. (2000). "Écriture en L2, registres, préconstructions, des questions pour la seule didactique des langues et/ou pour la RAL?". *ELA (Études de Linguistique Appliquée)* 120. Paris : Klincksieck. pp. 475-486.

Bouchard, R., Simon, J.-P. & Vourzay, M.-H. (2004). "L'intertextualité au service de l'interdiscursivité, étude de la production d'une synthèse de documents : à la recherche du "lieu commun". In *Actes du XXIVe colloque d'Albi, Langages et signification, l'intertextualité*, 2003. CALS/CPST. Toulouse : Presses de l'Université de Toulouse le Mirail. pp. 93-103.

Bouillaguet, A. (1996). *L'écriture imitative. Pastiche, parodie, collage*. Paris : Éditions Nathan.

Boulton, A. (2007). "Esprit de corpus : Promouvoir l'exploitation de corpus en apprentissage des langues". In *Revue électronique Texte et corpus*, n° 3, août 2008, *Actes des Journées de la linguistique de Corpus 2007*. p. 37-46. Disponible sur : http://hal.archives-ouvertes.fr/docs/00/32/69/88/PDF/2007_boulton_JFLS_esprit.pdf

Braun, S. (2005)."From pedagogically relevant corpora to authentic language learning contents". *ReCALL*, vol. 17, n° 1. Cambridge : Cambridge University Press. pp. 47-64.

Braun, S. (2007). "Integrating corpus work into secondary education: From data driven learning to needs-driven corpora". *ReCALL*, vol. 19, n°3. Camb ridge : Cambridge University Press. pp. 307-328.

Braun, S. (nd). *Corpora4Learning.net*. http://www.corpora4learning.net/

Bronckart, J.-P. (1997). *Activité langagière, textes et discours. Pour un interactionisme socio-discursif*. Paris : Delachaux et Niestlé.

Bronckart, J.-P. (2000). "Possibles de l'activité langagière et impossibles de 'la langue'". *Texto!*, vol. 18-19, décembre 2000. pp. 153-172. http://www.revue-texto.net/Parutions/Semiotiques/SEM_n18-19_7.pdf

Brossard, M. (2004). *Vygotski : lectures et perspectives de recherches en éducation. Suivi d'un inédit de L. S. Vygotski*. Traduit par Olga Anokhina et Michel Brossard. Lille : Presses Universitaires Septentrion.

Brown Corpus (nd). Corpus interrogeable sur : http://www.ldc.upenn.edu/cgi-bin/ldc/textcorpus?doc=yes&corpus=BROWN

Bruner, J. S. (1958). "Les processus de préparation à la perception". In Bruner, J. S., Bresson, F., Morf, A. & Piaget, J. (dir.). *Logique et perception*. *Études d'épistémologie génétique*, vol. 6. Paris : PUF (Presses Universitaires de France).

Bruner, J. (1983). "Le rôle de l'interaction en tutelle dans la résolution de problème". In Bruner, J. *Le développement de l'enfant. Savoir faire, savoir dire.* Vendôme : PUF (Presses Universitaires de France), 2002[7]. pp. 261-280.

Bruner, J. S. (1991). *Car la culture donne forme à l'esprit. De la révolution cognitive à la psychologie culturelle.* Paris : Éditions Eshel.

Bubenhofer, N. (2006). *Einführung in die Korpuslinguistik: Praktische Grundlagen und Werkzeuge.* Zurich. http://www.bubenhofer.com/korpuslinguistik

Cadre européen commun de réference pour les langues. Apprendre, enseigner, évaluer (2001). Division des politiques linguistiques, Strasbourg. Paris : Didier.

Catach, N. (1980). "La ponctuation". *Langue française*, n° 45. pp. 16-27. Disponible en ligne : http://www.persee.fr/web/revues/home/prescript/article/lfr_0023-8368_1980_num_45_1_5260

Chalmers, A. F. (1987). *Qu'est-ce que la science ?* Paris : La Découverte.

Chambers, A. (2005)."Integrating Corpus Consultation in Language Studies." *Language Learning & Technology*, 2005, vol. 9, n° 2. pp. 111-125. http://llt.msu.edu/vol9num2/chambers/default.html

Chanquoy, L. & Alamargot, D. (2003). "Mise en place et développement des traitements rédactionnels : le rôle de la mémoire de travail". *Le Langage et l'Homme*, vol. 38, n° 2, décembre 2003. pp. 171-190. Disponible en ligne : http://cogprints.org/3438/1/12_Chanquoy.pdf

Chanquoy, L., Tricot, A. & Sweller, J. (2007). *La charge cognitive*. Paris : Armand Colin.

Charlirelle (1977). *Behind the Words*. *Anglais sixième. Livret d'accompagnement.* Paris : Hatier.

Chateau, A. (2003). *Anglais de spécialité, la communication orale à objectifs spécifiques, une approche expérimentale du guidage*. Thése d'anglais. La Rochelle : Université de La Rochelle.

Chemnitz English-German Translation Corpus. Site de l'université de Chemnitz. http://ell.phil.tu-chemnitz.de/search/

Chemnitz Internet Grammar. Site de l'université de Chemnitz. http://ell.phil.tu-chemnitz.de/cing/frontend/index.html

Cheng, W., Warren, M., & Xun-feng, X. (2003). "The language learner as language researcher: Putting corpus linguistics on the timetable". *System*, vol. 31, n° 2. pp. 173-186.

Chini, D. (2001). "Détournement didactique de la locutionnalité : les blocs lexicalisés". In Tollis, F. (dir.). *La locution et la périphrase du lexique à la grammaire. Actes des Journées d'étude sur la locution*, organisées à l'Université de Pau les 16 et 17 octobre 1998 par le Groupe d'Approches du Langage de Pau. Paris, Budapest, Torino : L'Harmattan. pp. 37-55.

Chini, D. (2004). "Entre didactique et psycholinguistique, l'activité en classe de langue : quel rôle pour la problématique énonciative ?" *Note de synthèse, présentée le 14 septembre 2004 à l'université de Pau et des Pays de l'Adour en vue de l'obtention de l'habilitation à diriger des recherches.* Pau : Université de de Pau et des Pays de l'Adour. Disponible en ligne : http://acedle.u-strasbg.fr/IMG/pdf/DChiniHDR.pdf

Chiss, J.-L., Maingueneau, D. & Filliolet, J. (2001). *Introduction à la linguistique française : Notions fondamentales, Phonétique, Lexique* (Mise à jour de Linguistique française : *Notions fondamentales, Phonétique, Lexique.* 1993). Paris : Hachette.

Ciezielska-Ciupek, M. (2001). "Teaching with the internet and corpus materials: preparation of ELT materials using the internet and corpus resources". In Lewandowska-Tomaszczyk, B. (dir.). *Practical applications in language corpora. Lodz studies in language* 7 *(Palc 2001*). Frankfurt am Main, Berlin, Bern, Bruxelles, New York, Oxford, Wien : Peter Lang. pp. 521-531.

Cobb, T. (1997). "Is there any measurable learning from hands-on concordancing?". *System*, vol. 25, n° 3. pp. 301-315. Disponible en ligne : http://www.er.uqam.ca/nobel/r21270/cv/Hands_on.html

Cobb, T. (2006). "Constructivism, applied linguistics, & language education". In Brown, K. (dir.). *Encyclopedia of language and linguistics*. *Foundations of linguistics*, vol. 3[2]. Oxford : Elsevier. pp. 85-88.

Cobb, T. (nd). *Konkordanzer-Corpus Deutsch (v5.1).* http://www.lextutor.ca/concordancers/concord_g.html

Cobb, T. (nd). *Text-Concordancer (v2).* http://www.lextutor.ca/concordancers/text_concord/

Cobb, T. (nd). *The Compleat lexical tutor*. http://www.lextutor.ca/

Cobb, T. & Horst, M. (2000). "Vocabulary sizes of some City University students". *City University (HK) Journal of Language Studies*, vol.1. pp. 59-68.

Codex Manesse – die große Heidelberger Liederhandschrift. *Vollständiges Faksimile des Codex Palatinus Germanicus 848 der Universitätsbibliothek Heidelberg* (1975-1981). Frankfurt a.M. : Insel-Verlag. Disponible sur : http://digi.ub.uni-heidelberg.de/diglit/cpg848

Cordier-Gauthier, C. & Dion, C. (2003). "La révision de l'écrit en français langue seconde : médiation humain, médiation informatique." *Apprentissage des langues et systèmes d'information et de communication (ALSIC)*, vol. 6, n° 1. pp. 29-43. http://alsic.u-strasbg.fr/Num10/cordier/alsic_n10-rec5.htm

COSMAS II (nd). Outil d'exploration de *DeReKo* (Deutsches ReferenzKorpus), proposé par l'IDS (Institut für Deutsche Sprache), Mannheim. https://cosmas2.ids-mannheim.de/cosmas2-web/

Crinon, J. & Legros, D. (2000). "De l'ordinateur outil d'écriture à l'écriture outil". *Repères* 22. pp. 161-175.

Crinon, J. & Legros, D. (2002). *Psychologie des apprentissages et multimédia*. Paris : Armand Colin.

Crinon, J., Legros, D. & Marin, B. (2002a)."Quels outils pour écrire en lisant des récits ?". *Médialog*, n° 44. pp. 54-59.

Crinon, J., Legros, D. & Marin, B. (2002b). "Écrire et réécrire au cycle 3 : l'effet des mots clés sur la réécriture, avec et sans assistance informatique". In Plane, S. (dir.). *L'écriture et*

son apprentissage à l'école élémentaire. Repères, n° 26-27 (2002/03). Paris : INRP (Institut National de Recherche Pédagogique). pp. 187-202.

Crinon, J. & Vigne, H. (2002). *Écrire en lisant des récits de vie*. Cédérom pour le cycle 3 et le collège. Créteil : CRDP.

Cross, J. (2002). "'Noticing' in SLA: Is it a valid concept?". TESL-EJ, vol. 6, n° 3 A-2. pp. 1-9. http://tesl-ej.org/ej23/a2.html

Culioli, A. (1985). *Notes du séminaire de D.E.A. 1983 – 1984*. Poitiers. Disponible sur : *http://www.scribd.com/doc/15735/Culioli-Notes-de-seminaire-1983*

Culioli, A. (1990). *Pour une linguistique de l'énonciation.* Tome 1. Paris, Gap : Ophrys.

Culioli, A. (2002). *Variations sur la linguistique. Entretiens avec Frédéric Fau*. Paris : Klincksieck.

Das Digitale Wörterbuch der deutschen Sprache (DWDS). Berlin-Brandenburgische Akademie der Wissenschaft, Berlin. http://www.dwds.de/

David, J. & Plane, S. (1996). *L'apprentissage de l'écriture de l'école au collège*. Paris : PUF (Presses Universitaires de France).

Defays, J.-M. (3002). *Didactique du français : Langue maternelle, langue étrangère et langue seconde : vers un nouveau partage ?* Cortil-Wodon : É.M.E. (Éditions Modulaires Européennes) & InterCommunications s.p.r.l.

De Pietro, J.-F., Matthey, M. & Py, B. (1989). "Acquisition et contrat didactique : les séquences potentiellement acquisitionnelles de la conversation exolingue". In Weil, D. & Fugier, H. (dir.). *Actes du troisième colloque régional de linguistique*. Strasbourg : Université des sciences humaines et Université Louis Pasteur. pp. 99-124.

Demaizière, F. (2004)."Ressources et guidage. Définition d'une co-construction". In Develotte, C. & Pothier, M. (dir.). *NEQ (Notions en questions). La notion de ressource à l'heure du numérique.* Lyon : École Normale Supérieure de Lettres et Sciences Humaines. pp. 81-103.

Demaizière, F. (2007). "Didactique des langues et TIC : les aides à l'apprentissage". *Apprentissage des langues et systèmes d'information et de communication (ALSIC)*, vol. 10, n° 1. pp. 5-21. http://alsic.u-strasbg.fr/v10/demaiziere/alsic_v10_01-rec1.htm

Demaizière, F. & Berthoud, A.-C. (1998). "Des formes linguistiques pour communiquer : vers de nouveaux rapports entre linguistique et didactique". In Springer, C. (dir.). *Les linguistiques appliquées et les sciences du langage. Actes du deuxième colloque de linguistique appliquée Cofdela*. Strasbourg : Université Strasbourg 2. pp. 164-178. Disponible sur : http://didatic.net/article.php3?id_article=22

Demaizière, F. & Narcy-Combes, J.-P. (2005). "Méthodologie de la recherche didactique : nativisation, tâches et TIC". *Apprentissage des langues et systèmes d'information et de communication (ALSIC)*, vol. 8, n° 1. pp. 45-64. http://alsic.u-strasbg.fr/v08/demaiziere/alsic_v08_14-rec8.htm .

Demaizière, F. & Narcy-Combes, J.-P. (2007). "Du positionnement épistémologique aux données de terrain". *Les Cahiers de l'Acedle*, n° 4, 2007. http://acedle.u-strasbg.fr/IMG/pdf/Demaiziere-Narcy_cah4.pdf .

DeReKo (Deutsches ReferenzKorpus). IDS (Institut für Deutsche Sprache), Mannheim. Corpus interrogeable à l'aide du concordancier *COSMAS II*. http://www.ids-mannheim.de/kl/projekte/dereko_I/

De Singly, F. (1992). *L'enquête et ses méthodes : le questionnaire*. Paris : Nathan, 2004.

Deutschland-Portal (nd). Site de l'Ambassade d'Allemagne. http://www.deutschland.de/home.php

Develotte, C. & Potolia, A. (2005). "Les bases de données pédagogiques sur Internet : le cas du programme Éducasup". *Apprentissage des langues et systèmes d'information et de communication (ALSIC)*, vol. 8, n° 1. http://alsic.revues.org/index370.html

Diehl, E., Christen, H., Leuenberger, S., Pelvat, I. & Studer, T. (2000). *Grammatikunterricht – Alles für der Katz ? Untersuchungen zum Zweitsprachenerwerb Deutsch.* Tübingen : Niemeyer. Disponible en ligne : http://doc.rero.ch/lm.php?url=1000,10,3,20071125233235-GY/digs_complete_postprint_v1.pdf

Diniz, L. (2005). "Comparative review: TEXTSTAT 2.5, ANTCONC 3.0, and COMPLEAT LEXICAL TUTOR 4.0". *Language Learning & Technology*, vol. 9, n° 3. pp. 22-27. http://llt.msu.edu/vol9num3/pdf/review2.pdf

Divsic, D., Rott, S. & Horst, M. (2008). *Braun Corpus.* Corpus interrogable à l'aide du concordancier *Konkordanzer-Corpus Deutsch (v5.1).* http://www.lextutor.ca/concordancers/braun_info.html

Dodd, B. (1997). "Exploiting a corpus of written German for advanced language learning". In Wichmann, A., Fligelstone, S., McEnery, T. & Knowles, G. (dir.). *Teaching and Language Corpora.* London : Longman. pp. 131-145.

Doquet-Lacoste, C. (2004). "Genèse des textes, genèse des scripteurs : regards sur la langue dans l'écriture". In Piolat, A. (dir.). *Écriture. Approches en sciences cognitives*. Aix-en-Provence : Publications de l'Université de Provence. pp. 233-250.

Dortmunder Chat-Korpus. Technische Universität Dortmund. http://www.chatkorpus.tu-dortmund.de/

Doughty, C. & Long, M. (2003). "Optimal psycholinguistic environments for distance foreign language learning". *Language Learning & Technology*, vol. 7, n° 3, pp. 50-80. http://llt.msu.edu/vol7num3/doughty/default.html

Ducret, J.-J. (2009). "Notion 'décentration'". *Fondation Piaget*. http://www.fondationjeanpiaget.ch/fjp/site/presentation/index_notion.php?PRESMODE=1&NOTIONID=277

Duffner, R. & Näf, A. (2006a). "Korpuslinguistik im Zeitalter der Textdatenbanken". *Linguistik online* 28, 3/06. pp. 1-4. http://www.linguistik-online.de/28_06/index.html

Duffner, R. & Näf, A. (2006b). "Digitale Textdatenbanken im Vergleich". *Linguistik Online* 28, 3/06. pp. 7-23. http://www.linguistik-online.de/28_06/duffnerNaef.html

Ellis, N. (2001). "Constructions, chunking, and connectionism: The emergence of second language structure". In Doughty, C. & Long, M. (dir.). *Handbook of Second Language Acquisition*. Oxford : Blackwell. pp. 63-103.

Ellis, N. (2008). "Phraseology. The periphery and the heart of the language." In Granger, Sylviane & Meunier, F. (dir.). *Phraseology in Language Learning and Teaching*. Amsterdam, Philadelphia : John Benjamins Publishing Company. pp. 1-13.

Ellis, R. (2003). *Task-based Language Learning and Teaching*. Oxford : Oxford University Press.

Ellis, R. & Fotos, S. (1999). *Learning a Second Language Through Interaction.* Amsterdam, Philadelphia : John Benjamins Publishing Company.

Ericsson, K. & Simon, H. (1984). *Protocol Analysis: Verbal Reports as Data.* Cambridge, Massachusetts : MIT Press.

Esterl, U., Körner, E. & Einhorn, A., Kubicka, A. & Jenkins-Krumm, E.-M. (2008). *Team Deutsch, première année*. Stuttgart : Ernst Klett Sprachen GmbH ; Paris : Éditions Maison des Langues.

Falko. Corpus d'apprenants. Institut für Deutsche Sprache und Linguistik, université Humboldt, Berlin. http://www.linguistik.hu-berlin.de/institut/professuren/korpuslinguistik/forschung/falko

Fayol, M. (1996). "La production du langage écrit". In David, J. & Plane, S. (1996). pp. 9-36.

Feinberg, J. (2009). *Wordle*. http://www.wordle.net/

Fillmore, C. (1979). "On fluency". In Fillmore, C., Kempler, D. & Wang, W. (dir.). *Individual differences in language ability and language behavior*. New York : Academic Press. pp. 85-101.

Fink, L. & Frahling, G. (2008). *Linguee.com. The Web as a Dictionary.* http://www.linguee.com/PHP/about.php

Firth, J. R. (1951). "Modes of Meaning". In Palmer, F. R. (dir.). *Papers in Linguistics*. Londres : Oxford University Press. pp. 190-215.

Fleury, S., Lamalle, C., Martinez, W. & Salem, A. (2003). *Lexico3. Outils de statistique textuelle*. Paris : Université de Paris 3. http://www.cavi.univ-paris3.fr/Ilpga/ilpga/tal/lexicoWWW/ Le lien n'est plus actif le 19/08/2009.

Flowerdew, J. (1996). "Concordancing in Language Learning". In Pennington, M. (dir.). *The Power of CALL*. Houston XT : Athelstan. pp. 97-113. Disponible en ligne : http://sunzi1.lib.hku.hk/hkjo/view/10/1000064.pdf

Foster, P. (2001). "Rules and routines: a consideration of their role in the task-based language production of native and non-native speakers". In Bygate, M., Skehan, P. & Swain, M. (dir.). *Researching pedagogic tasks. Second language learning, teaching and testing*. Harlow : Longman.

Free (nd). Traducteur en ligne. http://services.portail.free.fr/traducteur/

Fuchs, C. & Le Goffic, P (1992). *Les linguistiques contemporaines : repères théoriques*. Paris : Hachette supérieur.

Gaonac'h, D. (1991). *Théories d'apprentissage et acquisition d'une langue étrangère*. Paris : Les Éditions Didier.

Gaonac'h, D. (2006). *L'apprentissage précoce d'une langue étrangère : le point de vue de la psycholinguistique.* Paris : Hachette Éducation.

Gaskell, D. & Cobb, T. (2004). "Can learners use concordance feedback for writing errors?" *System*, vol. 32, n° 3. pp. 301-319. Disponible en ligne : http://www.er.uqam.ca/nobel/r21270/cv/conc_fb.htm

Gläser, R. (1986). Phraseologie der englischen Sprache. Tübingen : Niemeyer.

Goodwin-Jones, B. (2001)."Emerging Technologies. Tools and Trends in Corpora Use for Teaching and Learning." *Language Learning & Technology*, vol. 5, n° 3. pp. 7–12. http://llt.msu.edu/vol5num3/emerging/

Grabe, W. & Kaplan, R. B. (1996). *Theory & Practice of Writing*. Harlow, Essex : Addison Wesley Longman Limited, (1997[2]).

Granger, S., Hung, J., & Petch-Tyson, S. (dir.) (2002). *Computer Learner Corpora, Second Language Acquisition and Foreign Language Teaching*. Amsterdam, Philadelphia : John Benjamins Publishing Company.

Granger, S. & Meunier, F. (2008). "Phraseology in language learning and teaching. Where to from here?" In Granger, S. & Meunier, F. (dir.). *Phraseology in Language Learning and Teaching*. Amsterdam, Philadelphia : John Benjamins Publishing Company. pp. 247-252.

Grosbois, M. (2007). "Didactique des langues et recherche expérimentale". *Les Cahiers de l'Acedle*, n° 4. pp. 65-83. http://acedle.org/IMG/pdf/Grosbois_cah4.pdf

Habert, B. (2000). "Des corpus représentatifs : de quoi, pour quoi, comment ?" In Bilger, M. (dir.). *Linguistique sur corpus. Études et réflexions. Actes de la journée Linguistique et corpus*. Pergpignan : Presses Universitaires de Perpignan. pp. 11-58. Disponible en ligne : http://atala.biomath.jussieu.fr/je/010428/Habert/Perpignan00/Perpignan00.html

Habert, B. & Nazarenko, A. & Salem, A. (1997). *Les linguistiques de corpus*. Paris : Armand Colin.

Hadley, G. (2002). "Sensing the Winds of Change: An Introduction to Data-Driven Learning". *RELC Journal*, vol. 33, n° 2. pp. 99-124. Disponible en ligne : http://www.nuis.ac.jp/~hadley/publication/windofchange/windsofchange.htm

Halberstadt, W., Carle, B., Faure-Paschal, N., Hoock-Douilly, M. & Wegener, A. (2008). *Spontan*. 4ème, 6ème. Paris : Éditions Didier.

Hallet, W. (2008). "Schreiben lernen mit dem Hypertext? Hypertextualität und generische Kohärenz in der Schreiberziehung". *Zeitschrift für Interkulturellen Fremdsprachenunterricht*, vol. 13/1. http://zif.spz.tu-darmstadt.de/jg-13-1/beitrag/Hallet1.htm

Hayes, J. R. (1996). "A new framework for understanding cognition and affect in writing". In Levy, C. M. & Ransdell, S. (dir.). *The Science of Writing*. Hillsdale, NJ : Lawrence Erlbaum Associates.

Hayes, J. R., & Flower, L. S. (1980). "Identifying the organization of writing processes". In Gregg, L. W. & Steinberg, E. R. (dir.). *Cognitive processes in writing*. Hillsdale, NJ : Lawrence Erlbaum Associates. pp. 31-50.

Henrici, G. (1995). *Spracherwerb durch Interaktion? Eine Einführung in die Fremdsprachenerwerbsspezifische Diskursanalyse.* Bausteine Deutsch als Fremdsprache, vol. 5. Hohengehren : Schneider.

Henry, A. & Roseberry, R. L. (1996). "A Corpus-based Investigation of the Language and Linguistic Patterns of One Genre and the Implications for Language Teaching". *Research in the Teaching of English*, n° 30. pp. 472-489.

Henry, A. & Roseberry, R. L. (2001). "Using a small corpus to obtain data for teaching a genre". In Ghadessy, M. & Henry, A. & Roseberry, R. L. (dir.). *Small Corpus Studies and ELT*. Amsterdam, Philadelphia : John Benjamins Publishing Compagny. pp. 93-133.

Hugon, M.A. & Seibel, C. (dir.) (1988). *Recherches impliquées. Recherches action : Le cas de l'éducation*. Bruxelles : De Boeck.

Hulstijn, J. (2007). "Psycholinguistic perspectives on second language acquisition". In Cummins, J. & Davison, C . (dir). *The international handbook on English language teaching*. Norwell, MA : Springer. pp. 701-713. Disponible en ligne : http://home.medewerker.uva.nl/j.h.hulstijn/bestanden/Hulstijn_Psycholinguistic_Perspectives_2007.pdf

Institut für Kommunikationswissenschaften. Université de Bonn. http://www.ikp.uni-bonn.de/forschung/computerlinguistik/computerlinguistik

Jauss, H. R. (1978). *Pour une esthétique de la réception*. Paris : Gallimard.

Jenny, L. (2003). "Les genres littéraires". *Méthodes et problèmes.* Genève : Département de français moderne. http://www.unige.ch/lettres/framo/enseignements/methodes/genres/

Johns, T. (1988). "Whence and Whither Classroom Concordancing?". In Bongaerts, T., de Haan, P., Lobbe, S. & Wekker, H. (dir.). *Computer Applications in Language Learning*. Dordrecht : Foris. pp. 9-27.

Johns, T. (1991). "Should you be persuaded - two samples of data-driven learning materials". In Johns, T. & King, P. (dir.) (1991). pp.1-13.

Johns, T. (2000). *Ustí nad Labem DDL Workshop: 21st – 25th March 2000. Participant's work.* Disponible sur : http://www.eisu2.bham.ac.uk/johnstf/unl_ddl.htm

Johns, T. (2002). "Data-driven Learning: the Perpetual Challenge". In Kettemann, B. & Marko, G. (dir.). *Teaching and Learning by Doing Corpus Analysis*. Amsterdam, New York : Éditions Rodopi B.V. pp. 107-117.

Johns, T. (nd). *Kibbitzers. English for International Students Unit*. http://www.eisu2.bham.ac.uk/Webmaterials/kibbitzers/indexes/indexdisc.htm

Johns, T. & King, P. (dir.) (1991). *Classroom concordancing*. *English language research journal* 4. Birmingham : Birmingham University.

Jordan, G. (2004). *Theory Construction in Second Language Acquisition*. Amsterdam, Philadelphia : John Benjamins Publishing Company.

Juan, S. (1999). *Méthodes de recherche en sciences sociohumaines. Exploration critique des techniques.* Paris : PUF (Presses Universitaires de France)

Kahn, H., Guénebaut, E., Houlot, D. & Moine, M. (1993). "Die geteilte Stadt" ; "Der Brief". In *Sag' mal...*, classe de 3ème. Paris : Armand Colin. pp. 28-31. D'après Kordon, K. (1988) : *Die Flaschenpost*. Ravensburg : Otto Maier.

Kaltenböck, G. & Mehlmauer-Larcher, B. (2005). "Computer corpora and the language classroom: on the potential and limitations of computer corpora in language teaching". *ReCALL*, vol. 17, n° 1. Cambridge : Cambridge University Press. pp. 65-84.

Kaur, J. & Hegelheimer, V. (2005). "ESL Students' Use of Concordance in the Transfer of Academic Word Knowledge: An Exploratory Study". *Computer assisted language learning*, vol.18. pp. 287 -310.

Kennedy, C. & Miceli, T. (2001). "An evaluation of intermediate student's approaches to corpus investigation." *Language Learning & Technology*, vol. 5/3, septembre 2001. pp. 77-90. http://llt.msu.edu/vol5num3/kennedy/

Kilgarriff, A., Rychly, P. & Pomikalek J. (nd). *Sketch Engine*. Collection de corpus. http://www.sketchengine.co.uk/

Kucera, H. & Francis, W. N. (1967). *Computational analysis of present-day American English*. Providence : Brown University Press.

Labelle, M. (2001). "30 ans de psycholinguistique". *Revue québécoise de linguistique. 1971-2001 : Trente ans de linguistique*, vol. 30, n° 1. Montréal : Université du Québec. pp. 155-176. Disponible en ligne : http://www.erudit.org/revue/rql/2001/v30/n1/000516ar.pdf

Lamy, M.-N. & Klarskov Mortensen, H. J. (2006). *ICT4LT (Information and Communications Technology for Language Teachers), module 2.4 : Using concordance programs in the Modern Foreign Languages classroom*. http://www.ict4lt.org/en/en_mod2-4.htm

Laurier, M. D. (2005[3]). *Les principes de la mesure et de l'évaluation des apprentissages*. Montréal (Québec) : Gaëtan Morin.

Lebart, L. & Salem, A. (1994). *Statistique Textuelle*. Paris : Dunod. Disponible en ligne : http://ses.telecom-paristech.fr/lebart/ST.html

Lee, D. (nd). *David Lee's Bookmarks for Corpus-based Linguists.* http://personal.cityu.edu.hk/~davidlee/devotedtocorpora/CBLLinks.htm

Leech, G. (2003). "Corpus Processing". In Frawley, W. J. (dir.) *International Encyclopedia of Linguistics*. Oxford : Oxford University Press. Disponible en ligne : http://www.oxford-linguistics.com/entry?entry=t202.e0254

Leki, I., Cumming, A. H. & Silva, T. (2008). *A Synthesis of Research on Second Language Writing in English: 1980-2005*. London : Routledge (Taylor & Francis).

Lemnitzer, L. & Zinsmeister, H. (2006). *Korpuslinguistik – Eine Einführung*. Tübingen : Gunter Narr Verlag.

LEO (nd). Un service en ligne de l'entreprise LEO GmbH. http://dict.leo.org/frde?lang=de&lp=frde&search=

LIMAS (Linguistik und Maschinelle Sprachbearbeitung). Université de Bonn. Voir *DeReKo*. http://www.korpora.org/Limas/

Linard, M. (1996). *Des machines et des hommes*. Paris : L'Harmattan.

Logan, G. (1988). "Towards an instance theory of automatisation". *Psychological Review*, n° 95. pp. 492-527.

Lüdeling, A. & Walter, M. (2009). "Korpuslinguistik für Deutsch als Fremdsprache – Sprachvermittlung und Spracherwerbsforschung". Version élargie de "Korpuslinguistik" (à paraître). Fandrych, C., Hufeisen, B., Krumm, H.-J. & Riemer, C. (dir.). *Deutsch als Fremd- und Zweitsprache. Ein internationales Handbuch. 2. überarbeitete Auflage.* (= HSK, vol. 19).

Berlin, New York : Mouton de Gruyter. Disponible sur : http://www.linguistik.hu-berlin.de/institut/professuren/korpuslinguistik/mitarbeiter-innen/anke/pdf/LuedelingWalterDaF.pdf

Maingueneau, D. (2008[2]). *L'énonciation en linguistique française*. Paris : Hachette.

Mangenot, F. (1996). *Gammes d'écriture*. Logiciel Windows. Paris : CNDP. Documentation consultable sur : http://www.cndp.fr/lettres/gamecrit/

Mangenot, F. (1997). "Aide à la rédaction à travers un dialogue 'ordinateur/apprenant'". *Revue de l'association Enseignement Public et Informatique* n° 87, septembre 1997. Paris : EPI. Disponible en ligne : http://www.epi.asso.fr/revue/87/b87p149.htm

Mangenot, F. (1998)."Classification des apports d'Internet à l'apprentissage des langues". *Apprentissage des langues et systèmes d'information et de communication (ALSIC)*, vol. 1, n° 2. pp. 133-146. http://alsic.u-strasbg.fr/Num2/mangenot/alsic_n02-pra1.htm

Mangenot, F. (2000). "Contexte et conditions pour une réelle production d'écrits en ALAO". *Apprentissage des langues et systèmes d'information et de communication (ALSIC)*, vol. 3, n° 2, décembre 2000. pp. 187-206. http://alsic.u-strasbg.fr/Num6/mangenot/alsic_n06-rec1.htm

Marin, B. & Legros, D. (2008). *Psycholinguistique cognitive : lecture, compréhension et production de texte*. Bruxelles : De Boeck.

Martinez, P. (2004[4]). *La didactique des langues étrangères*. Édition mise à jour. Paris : PUF (Presses Universitaires de France), 1996.

Matthey, M. & Véronique, D. (2004). "Trois approches de l'acquisition des langues étrangères : enjeux et perspectives". *AILE* (*Acquisition et Interaction en Langue Étrangère*), vol. 21. Paris : Association Aile Encrages, université Paris VIII. pp. 203-223.

Mayaffre, D. (2005). "Rôle et place des corpus en linguistique : réflexions introductives". *Texto!*, vol. 10, n° 4. http://www.revue-texto.net/1996-2007/Corpus/Publications/Mayaffre_Corpus.html

Mazière, F. (2005). *L'analyse du discours. Histoire et pratiques*. Paris : PUF (Presses Universitaires de France).

Meyer, C. F. (2002). *English Corpus Linguistics: An Introduction*. Cambrige : Cambridge University Press.

Ministère de la jeunesse, de l'éducation nationale et de la recherche (1998**).** "Accompagnement du programme de français de 3ème". *Bulletin Officiel de l'Éducation Nationale*, n° 10. Disponible sur : ftp://trf.education.gouv.fr/pub/edutel/bo/1998/hs10/hs10vol2.pdf

Ministère de la jeunesse, de l'éducation nationale et de la recherche (2002). "Diplôme de Compétence en Langue". *Bulletin Officiel de l'Éducation Nationale*, n° 22. Disponible sur : http://fleasso.club.fr/diplomes/2002-DCL.htm

Ministère de la jeunesse, de l'éducation nationale et de la recherche (2006). "Droit à l'expérimentation". *Bulletin Officiel de l'Éducation Nationale*, n° 13. *Circulaire n° 2006-051 du 27/03 2006.* Disponible sur : http://www.education.gouv.fr/bo/2006/13/MENE0600903C.htm

Ministère de la jeunesse, de l'éducation nationale et de la recherche (2007). "Programmes de l'enseignement de langues vivantes étrangères au collège. Allemand". *Bulletin Officiel de l'Éducation Nationale, hors série*, n° 7. Disponible sur : ftp://trf.education.gouv.fr/pub/edutel/bo/2007/hs7/hs7_allemand-vol3.pdf

Ministère de la jeunesse, de l'éducation nationale et de la recherche (2008). *Programmes du collège. Programmes de l'enseignement de français*. *Bulletin officiel spécial*, n° 6. Disponible sur : http://media.education.gouv.fr/file/special_6/21/8/programme_francais_general_33218.pdf

Minuth, C. & Schlemminger, G. (2000). "Freinet-Pädagogik im Fremdsprachenunterricht: Lernreflexion über Sprache und Spracherwerb". In Schlemminger, G., Brysch, T. & Schewe, M. L. (dir.). *Pädagogische Konzepte für einen ganzheitlichen DaF-Unterricht*. Berlin : Cornelsen Verlag. pp. 127-140.

Missotte, P. (2003). "La recherche-action pour un développement endogène". In Missotte, P. & Mesnier, P.-M. (dir.). *La recherche-action, Une autre manière de chercher, se former, transformer. Recherche-action en pratiques sociales*. Paris : L'Harmattan. pp. 248-251.

Montagne-Macaire, D. (2007). "Didactique des langues et recherche-action". *Les Cahiers de l'Acedle*, n° 4. http://acedle.org/IMG/pdf/Macaire-D_cah4.pdf

Muller, F. & Teillard, F. (nd). *Expérimenter*. http://www.ac-paris.fr/portail/jcms/piapp1_70440/evaluer-l-experimentation-et-experimenter-dans-l-evaluation-c-est-possible

Mukherjee, J. (2002): *Korpuslinguistik und Englischunterricht: Eine Einführung.* Frankfurt am Main, Berlin, Bern, Bruxelles, New York, Oxford, Wien : Peter Lang.

Murison-Bowie, S. (1993). *MicroConcord Manual*. Oxford : Oxford University Press.

Narcy-Combes, J.-P. (2005). *Didactique des langues et TIC : Vers une recherche-action responsable*. Paris : Éditions Ophrys.

Narcy-Combes, J.-P. et Walski, J. (2004)."Le concept de la tâche soumis au crible de nouvelles questions". In *Les cahiers de l'APLIUT (Association des Professeurs de Langue en IUT)*, vol. 33, n° 1, février 2004. *Actes du XXVième Congrès de l'APLIUT (Association des Professeurs de Langue en IUT)*, IUT Paul Sabatier, Toulouse III, 5-6-7 juin 2003. pp. 27-44.

Narita, M., Kurokawa, K. & Utsuro, T. (2003). "Case study on the development of a computer-based support tool for assisting Japanese software engineers with their English writing needs". *Professional Communication, IEEE Transactions on Publication*, vol. 46/3. pp. 194-209.

NEGRA Korpus Version 2. Université de Saarbrücken. http://www.coli.uni-saarland.de/projects/sfb378/negra-corpus/negra-corpus.html

Nicol-Benoît, W. (2004a). *L'approche par tâches dans l'apprentissage de l'anglais de spécialité – Opérationnalisation contrôlée dans l'enseignement supérieur*. Thèse d'anglais. Nantes : Université de Nantes.

Nicol-Benoît, W. (2004b) : "Comment l'approche par tâches permet de tout intégrer ou presque". Communication RANACLES, Université de Lille III, 25-27 novembre 2004. Disponible en ligne : http://crl.univ-lille3.fr/ranacles2004/commu/nicol-benoit.htm

Norman, D. A., Rumelhart, D. E. & *the LNR Research Group* (1975). *Explorations in cognition*. San Francisco : Freeman.

O'Keeffe, A. & McCarthy, M. & Carter, R. (2007). *From Corpus to Classroom. Language Use and Language Teaching*. Cambridge : Cambridge University Press.

Olive, T. & Piolat, A. (2003)."Activation des processus rédactionnels et qualité des textes". In *Le Langage et l'Homme*, vol. 38, n° 2. pp. 191-206. Disponible en ligne : http://cogprints.org/3605/1/OlivePiolatLH2003.pdf

Partington, A. (1998). *Patterns and Meanings. Using Corpora for English Research and Teaching*. Amsterdam, Philadelphia : John Benjamins Publishing Company.

Partington, A. (2006). "Aims, Tools and Practices of Corpus Linguistics". *IntUne Papers. Integrated and United: A quest for Citizenship in an 'ever closer Europe'*. http://www.intune.it/file_download/23

Perrenoud, P. (1994). *Métier d'élève et sens du travail scolaire*. Issy-les-Moulineaux : ESF Éditeur, 2005[6].

Perrenoud, P. (1996). "Le travail sur l'habitus dans la formation des enseignants. Analyse des pratiques et prise de conscience." In Paquay, L. & Altet, M. & Charlier, É. & Perrenoud, P. (dir.). *Former des enseignants professionnels. Quelles stratégies ? Quelles compétences ?* Paris, Bruxelles : De Boeck & Larcier S.A. pp. 182-207.

Perrenoud, P. (2004). "Évaluer des compétences". In *L'Éducateur*, numéro spécial *"La note en pleine évaluation"*, mars 2004. pp. 8-11. Disponible sur : http://www.unige.ch/fapse/SSE/teachers/perrenoud/php_main/php_2004/2004_01.html

Petzold, D. (nd). *Zoo-Datenbank für Deutschland*. http://www.zoo-infos.de/

Piaget, J. (1969). *Psychologie et pédagogie.* Paris : Éditions Denoël, 1993[2].

Pincemin, B. (2006). "Concordances et concordanciers – De l'art du bon KWAC". Soumission à *Documents numériques et interprétation – Corpus en Lettres et Sciences sociales*, Albi, 10-14 juillet 2006. Disponible en ligne : http://www.revue-texto.net/Parutions/Livres-E/Albi-2006/pincemin.pdf

Plane, S. (1996). "Écriture, réécriture et traitement de texte." In David, J. & Plane, S. (1996). pp. 37-77.

Plane, S. (2006). "Singularités et constantes de la production d'écrit. L'écriture comme traitement de contraintes". In Lafont-Terranova, J. & Colin, D. (dir.). *Didactique de l'écrit. La construction des savoirs et le sujet écrivant*. Presses Universitaires de Namur. pp. 33-54. Disponible en ligne : http://www.vjf.cnrs.fr/umr8606/FichExt/splane-pdf/Plane2004b.pdf

Poiarkova, E. (2006). *Le rôle de la traduction dans la formation des futurs professeurs (exemple de l'Université pédagogique de St Petersbourg)*. Thèse de doctorat. A ix en Provence : Université de Provence.

Portine, H. (1998). L'autonomie de l'apprenant en questions. *Apprentissage des langues et systèmes d'information et de communication (ALSIC)*, vol. 1, n° 1. pp 73-77. http://toiltheque.org/Alsic_volume_1-7/Num1/portine/alsic_n01-poi1.pdf

Portine, H. (2008). Activités langagières, énonciation et cognition. La centration sur les apprentissages. In Narcy-Combes, M.-F. & Toffoli, D. (dir.). *Les Cahiers de l'Acedle*, vol. 5, n° 1. pp. 233 à 254. http://acedle.u-strasbg.fr/IMG/pdf/Portine_Cah5-1.pdf

Projekt Gutenberg (nd). http://www.gutenberg.org/browse/languages/de

Py, B. (2002). "Acquisition d'une langue seconde, organisation macrosyntaxique et émergence d'une microsyntace". *Marges linguistiques*, novembre 2002. Saint-Chamas : M.L.M.S. éditeur. pp. 48-55. Disponible sur : http://www.revue-texto.net/1996-2007/marges/marges/Documents%20Site%200/00_ml112002/00_ml112002.pdf

Randall, M. (2007). *Memory, Psychology, and Second Language Learning*. Amsterdam, Philadelphia : John Benjamins Publishing Company.

Rastier, F. (2001). "Eléments de théorie des genres". *Texto !*, juin 2001. http://www.revue-texto.net/Inedits/Rastier/Rastier_Elements.html

Rastier, F. (2004). "Enjeux épistémologiques de la linguistique de corpus". *Texto !*, juin 2004. Rubrique Dits et inédits. http://www.revue-texto.net/1996-2007/Inedits/Rastier/Rastier_Enjeux.html

Rastier, F. (2008). "Conditions d'une linguistique des normes". In Steuckardt, A. & Siouffi, F. (dir.). *Les linguistes et la norme – Aspects normatifs du discours linguistique.* Frankfurt am Main, Berlin, Bern, Bruxelles, New York, Oxford, Wien : Peter Lang. pp. 3-20. *Texto!*, juillet 2008, vol. 18, n° 3. http://www.revue-texto.net/docannexe/file/1612/rastier_normes.pdf

Reverso (nd). Traducteur en ligne. http://www.reverso.net/text_translation.asp?lang=FR

Rézeau, J. (2007). "L'apport du concordancier à l'analyse et à la remédiation des erreurs des apprenants dans les forums de discussion en ligne". *Apprentissage des langues et systèmes d'information et de communication (ALSIC)*, vol. 10, n° 2. pp. 27-43. http://alsic.u-strasbg.fr/v10/rezeau/alsic_v10_04-pra1.htm

Robert, S. (2002). "Modèles linguistiques de production". In Fayol, M. (dir.). *Traité des Sciences Cognitives*, vol. *Production du langage*. Paris : Hermès. pp. 66-86.

Robinson, P. (2002). *Individual Differences and Instructed Language Learning.* Amsterdam, Philadelphia : John Benjamins Publishing Company.

Rohrbach, J.-M. (2003). "'Don't miss out on Göttingen's nightlife': Genreproduktion im Englischunterricht". *Praxis des Neusprachlichen Unterrichts*, 50. pp. 381-389.

Rouet, J.-F., Coutelet, B. & Dinet, J. (2004). "La recherche d'informations dans les documents complexes : Processus cognitifs, apprentissage et développement". *Conférence invitée à la Journée d'Etudes sur le Traitement Cognitif des Systèmes d'Informations Complexes* (JETCSIC), Genève, 17-18 juin 2004. Disponible en ligne : http://tecfa.unige.ch/tecfa/jetcsic2004/papers/jfrjetcsic2004.pdf

Rumelhart, D. E. (1977). "Toward an interactive model of reading". Dornic, S. (dir.). *Attention and performance VI*. Hillsdale, NJ : Lawrence Erlbaum Associates.

Ruthven-Stuart, P. (nd). *How to use concordances in teaching English.* http://www.nsknet.or.jp/~peterr-s/concordancing/usingconcs.html

Schaeffer-Lacroix, E. (nd). *Esldeutsch*. http://www.didaktik.fr

Schaeffer-Lacroix, E. (nd). *Filmblog* (me contacter : evalacroix@free.fr).

Schaeffer-Lacroix, E. (nd). *Projekt Prospekt*. Identifiant : corpus. Mot de passe : wanda. http://corpus.hautetfort.com/

Schaeffer-Lacroix, E. (nd). *Projektprospekt* (me contacter : evalacroix@free.fr).

Schaeffer-Lacroix, E. (nd). *Tourismus* (me contacter : evalacroix@free.fr).

Schenk, K. (1998). "Die allmähliche Verfertigung der Texte beim Schreiben. Vorschläge zur Strukturierung von Schreibsitzungen". *Beiträge zur Fremdsprachenvermittlung* 34 (1998). Landau, Konstanz : Institut für fremdsprachliche Philologien der Universität Koblenz-Landau, Universität Konstanz. pp. 98-117.

Schiller, A., Thielen, C., Teufel, S. & Stöckert, C. (nd). *STTS (Stuttgart/Tübinger Tagset).* http://www.ifi.uzh.ch/arvo/cl/tagger/UIS-STTS-Diffs.html

Schlobinski, P. (nd). *Schlobi's Linguistic Corner.* http://www.linguistic-corner.uni-hannover.de/_sprachkorpora.html

Schmidt, R. (1990). "The Role of Consciousness in Second Language Learning". *Applied Linguistics*, 1990, 11/2. Oxford : Oxford University Press. pp. 129-158.

Schmidt, R. & Frota, S. N. (1986). "Developing basic conversational ability in a second language: A case study of an adult learner of Portuguese". In Day, R. R. (dir.). *Talking to Learn: Conversation in Second Language Acquisition*. Rowley, MA : Newbury House.

Scripertexte. Voir Crinon, J. & Vigne, H. (2002).

Selva, T. & Verlinde S. (2002). "L'utilisation d'un dictionniare électronique : une étude de cas". In *Actes du dixième congrès EURALEX'2002 (European Association for Lexicography)*, Copenhague. pp. 773-783.

Sharoff, S. (nd). *French web corpus.* Corpus interrogeable sur le site de *Sketch Engine* (Kilgarriff *et al.).*

Shetzer, H. & Warschauer, M. (2000). "An electronic literacy approach to network-based language teaching". In Warschauer, M. (dir.). *Network-based Language Teaching: Concepts and Practice*. Cambridge : CUP. pp. 171-185.

Siepmann, D. (1999). "Linguistique de corpus et grammaire : une nouvelle donne pour l'enseignement du français langue étrangère". *französisch heute* 3. pp. 309-327.

Siepmann, D. (2003). "Fremdsprachliches Schreiben auf der gymnasialen Oberstufe: acht Thesen zur bestehenden Praxis und Vorschläge für einen Teillehrplan". *Neusprachliche Mitteilungen aus Wissenschaft und Praxis 1*. München : Oldenbourg Schulbuchverlag. pp. 24-32. Disponible en ligne : http://www.dirk-siepmann.de/Publications/schreiben.htm

Siepmann, D. (2004). "Kollokationen und Fremdsprachenlernen: Imitation und Kreation, Figur und Hintergrund". *Praxis Fremdsprachenunterricht*, vol. 2. Berlin : Cornelsen Verlag. pp. 107-113.

Siepmann, D. (2008). "Phraseology in learners' dictionaries. What, where and how?" In Granger, S. & Meunier, F. (dir.). *Phraseology in Language Learning and Teaching*. Amsterdam, Philadelphia : John Benjamins Publishing Company. pp. 185-202.

Šimunek, R. (2007). *Formelhafte Sprache und Internetprojekte. Zum Sprachgebrauch und -erwerb fortgeschrittener Fremdsprachenlerner*. Tübingen : Gunter Narr Verlag.

Sinclair, J. (1996). *EAGLES. Preliminary Recommendations on Corpus Typology*. Pisa : Consiglio Nazionale delle Ricerche. Istituto di Linguistica Computazionale. Disponible en ligne : http://www.ilc.cnr.it/EAGLES96/corpustyp/corpustyp.html

Sinclair, J. (2004). *Trust the text. Language, Corpus and Discourse*. London : Routledge.

Sinclair, J. (2005). "Corpus and Text – Basic Principles". In Wynne, M. (2005). *Developing Linguistic Corpora: a Guide to Good Practice*. Oxford : Oxbow Books. pp. 1-16. Disponible en ligne : http://ahds.ac.uk/linguistic-corpora/

Site de l'académie de Rouen (nd). http://www.ac-rouen.fr/pedagogie/plan_tice/2001/pers_ress.htm

Site de l'académie de Versailles (nd). http://www.lettres.ac-versailles.fr/

Springer, C. (1999). "Évaluation de la compétence et problématique de l'acquisition en L2 : Préliminaires pour une définition de profils prototypiques de compétence en L2". In *Actes du XIe Colloque International FOCAL. Acquisition d'une langue étrangère : perspectives et recherches*. Paris, 19-21 avril 1999. Disponible en ligne : http://www.revue-texto.net/1996-2007/marges/marges/Documents%20Site%206/doc0087_springer_c/doc0087.pdf

Springer, C. (2009). "La dimension sociale dans le CECR : pistes pour scénariser, évaluer et valoriser l'apprentissage collaboratif". In Rosen, E. (dir.) *La perspective actionnelle et l'approche par les tâches en classe des langues. Le français dans le monde. Numéro spécial 45 "Recherche et applications"*. Paris : Clé International. pp. 25-34.

Stevens, V. (1988). "Studying vocabulary using concordances on microcomputers". *TESOL Newsletter* 22, 3/27. Disponible en ligne : http://www.vancestevens.com/papers/tesol/newsletter/1988online.htm

Stevens, V. (1991). "Concordance-based vocabulary exercises: A viable alternative to gap-fillers". In Johns, T. & King, P. (dir.). pp. 47-61.

Stevens, V. (1995). "Concordancing with Language Learners: Why? When? What?". *CAELL Journal (Computer Assisted English. Language Learning)*, vol. 6, n° 2. pp. 2-10.

St. John, E. (2001). "A case for using parallel corpus and concordancer for beginners of a foreign langage". *Language Learning & Technology*, vol. 5, n° 3, septembre 2001. pp. 185-203. http://llt.msu.edu/vol5num3/stjohn/default.html

St. John, E. & Cash, D. (1995). "German language learning via email: a case study". *ReCALL* vol. 7, n° 2. pp. 47-51.

Sun Y.-C. & Wang L.-Y. (2003). "Concordancers in the EFL classroom: cognitive approaches and collocation difficulty". *CALL (Computer-Assisted Language Learning)*, vol. 16/1. pp. 83-94.

Swain, M. (1985). "Communicative Competence: Some roles of comprehensible *input* and comprehensible output in its development". Madden, C. & Gass, S. M. (dir.). *Input in Second Language Acquisition.* Rowley, MA : Newbury House. pp. 235-253.

Swain, M. (1995). "Three functions of output in second language learning". In Cook, G. & Seidlhofer, B. (dir.). *Principles and practice in the study of language*. Oxford : Oxford University Press. pp. 125-144.

Swain, M. & Lapkin, S. (1995). "Problems in output and the cognitive processes they generate: A step towards second language learning". *Applied Linguistics* 16. pp. 371-391.

Teubert, W. (2005). "My version of corpus linguistics". *International Journal of Corpus Linguistics*, vol. 10, n° 1, 2005. Amsterdam, Philadelphia : John Benjamins Publishing Company. pp. 1-13. Disponible en ligne : http://www.benjamins.com/jbp/series/IJCL/10-1/art/0001a.pdf

Teubert, W. (2006). "Korpuslinguistik, Hermeneutik und die soziale Konstruktion der Wirklichkeit". *Linguistik Online* 28, 3/06. pp. 41-60. http://www.linguistik-online.de/28_06/teubert.html

Tiedemann, J. (2009). *Opus. An open source parallel corpus*. http://urd.let.rug.nl/tiedeman/OPUS/

TIGERCorpus (2007). Institut für Phonetik/Phonologie, Universität des Saarlandes & Institut für Germanistik, Potsdam & Institut für Maschinelle Sprachverarbeitung, Stuttgart. http://www.ims.uni-stuttgart.de/projekte/TIGER/TIGERCorpus/

TreeTagger. Institut für maschinelle Datenverarbeitung, Stuttgart. http://www.ims.uni-stuttgart.de/ims-home.html.en

Trésor de la langue française informatisé (nd). Universités de Nancy 1 et Nancy 2. http://atilf.atilf.fr/tlf.htm

Tribble, C. (1997). "Improvising Corpora for ELT: Quick and Dirty Ways of Developing Corpora for Language Teaching". In B. Lewandowska-Tomaszczyk and J. Melia (dir.). *PALC'97* (*Proceedings of the First International conference on Practical Applications in Language Corpora)*. Lodz : Lodz University Press. Disponible en ligne : http://www.ctribble.co.uk/text/Palc.htm

Tricot, A. (2007). *Apprentissages et documents numériques*. Paris : Belin.

Tschirner, E., Buchwald, I., Goldhahn, A., Ittner, A. & Jones, R. L. (2006). *A Frequency Dictionary of German: Core Vocabulary for Learners*. London : Routledge.

Tutin, A. & Grossmann, F. (2002). "Collocations régulières et irrégulières : esquisse de typologie du phénomène collocatif". *Revue Française de Linguistique Appliquée, Lexique : recherches actuelles*, vol. 7, juin 2002. pp. 7-25.

Tutin, A. (2005). Le dictionnaire de collocations est-il indispensable ? *Revue Française de Linguistique Appliquée. Dictionnaires : nouvelles approches, nouveaux modèles*, vol. 10/2, décembre 2005. Disponible en ligne : http://w3.u-grenoble3.fr/tutin/Publis/tutin_RFLA.pdf

Véronis, J. & et Khouri, L. (1995). "Étiquetage grammatical multilingue : le projet MULTEXT". *Traitement automatique des langues (T.A.L)*, vol. 36, n° 1-2. pp. 233-248.

Vygotski, L. (1934). *Pensée et langage*. Paris : La Dispute/SNÉDIT, 1997[3].

WaCky (*Web-as-Corpus kool ynitiative*). Site donnant accès à des corpus regroupés dans *Sketch Engine* (Kilgarriff *et al.*). http://wacky.sslmit.unibo.it/doku.php?id=corpora

Weigle, S. C. (2002). *Assessing Writing*. Cambrigde : Cambridge University Press.

Wortschatz-Portal. Université de Leipzig. http://wortschatz.uni-leipzig.de/

Zanettin, F. (2001). "Swimming in Words: Corpora, Translation, and Language Learning". In Aston, G. (dir.). *Learning with Corpora*. Houston XT : Athelstan.

Zimmermann, R. (2000). "L2 writing: subprocesses, a model of formulating and empirical findings". *Learning and Instruction* 10. Amsterdam, Eslevler, Oxford : Pergamon. pp. 73-99.

Références "Corpus et concordances"

Anthony, L. (2002). *AntConc1.0.* http://www.antlab.sci.waseda.ac.jp/antconc_index.html

Anthony, L. (2006). "Developing a Freeware, Multiplatform Corpus Analysis Toolkit for the Technical Writing Classroom". *IEEE Transactions on Professional Communication*, vol. 49, n° 3. pp. 275-286.

Anthony, L. (2007). *AntConc3.2.1w*. http://www.antlab.sci.waseda.ac.jp/antconc_index.html

Aston, G. (2002). "The learner as corpus designer". In KETTEMANN, Bernhard & MARKO, Georg (dir.). *Teaching and Learning by Doing Corpus Analysis*. Amsterdam, New York : Éditions Rodopi B.V. pp. 9-26.

Axtell, T. (nd). *Concordancing in the CEGEPs*. http://pages.infinit.net/jaguar3/lounge/concord/default.htm

Baroni, M. (2006). *DeWac.* Corpus interrogeable sur le site de *Sketch Engine* (Kilgarriff *et al.).*

Baroni, M. & Bernardini, S. (2003). *BootCaT toolkit 0.1.2. (Bootstrapping Corpora and Terms from the Web).* Outil de création de corpus. http://sslmit.unibo.it/~baroni/bootcat.html

Baroni, M. & Bernardini, S. (2004). "BootCaT: Bootstrapping Corpora and Terms from the Web". *Proceedings of LREC* 2004. Disponible en ligne : http://people.sslmit.unibo.it/~baroni/publications/lrec2004/bootcat_lrec_2004.pdf

Baroni, M. & Kilgarriff, A. (2006). "Large linguistically-processed Web corpora for multiple languages". *Conference Companion of EACL* 2006 (11th Conference of the European Chapter of the Association for Computational Linguistics). East Stroudsburg PA : ACL. pp. 87-90.

Bernardini, S. (2000). "Systematising serendipity: Proposals for concordancing large corpora with language learners". In Burnard, L. & McEnery, T. (dir.). *Rethinking language pedagogy from a corpus perspective*. Frankfurt am Main, Berlin, Bern, Bruxelles, New York, Oxford, Wien : Peter Lang. pp. 225-234.

Bernardini, S. (2004). "Corpora in the classroom. An overview and some reflections on future developements". In Sinclair, J. (dir.). *How to Use Corpora in Language Teaching*. Amsterdam : John Benjamins. pp. 15-36.

BNC (British National Corpus). Corpus interrogeable sur : http://www.natcorp.ox.ac.uk/

Boulton, A. (2007). "Esprit de corpus : Promouvoir l'exploitation de corpus en apprentissage des langues". In *Texte et corpus*, n° 3, août 2008. *Actes des Journées de la linguistique de Corpus 2007*. Lorient : Université de Bretagne-Sud. pp. 37-46. Disponible sur : http://hal.archives-ouvertes.fr/docs/00/32/69/88/PDF/2007_boulton_JFLS_esprit.pdf

Braun, S. (2005). "From pedagogically relevant corpora to authentic language learning contents". *ReCALL*, vol. 17, n° 1. Cambridge : Cambridge University Press. pp. 47-64.

Braun, S. (2007). "Integrating corpus work into secondary education: From data driven learning to needs-driven corpora". *ReCALL*, vol. 19, n° 3. Cambridge : Cambridge University Press. pp. 307-328.

Braun, S. (nd). *Corpora4Learning.net*. http://www.corpora4learning.net/

Brown Corpus (nd). Corpus interrogeable sur : http://www.ldc.upenn.edu/cgi-bin/ldc/textcorpus?doc=yes&corpus=BROWN

Bubenhofer, N. (2006). *Einführung in die Korpuslinguistik: Praktische Grundlagen und Werkzeuge.* Zurich. http://www.bubenhofer.com/korpuslinguistik

Chambers, A. (2005). "Integrating Corpus Consultation in Language Studies." *Language Learning & Technology*, 2005, vol. 9, n° 2. pp. 111-125. http://llt.msu.edu/vol9num2/chambers/default.html

Chemnitz English-German Translation Corpus. Site de l'université de Chemnitz. http://ell.phil.tu-chemnitz.de/search/

Ciezielska-Ciupek, M. (2001). "Teaching with the internet and corpus materials: preparation of ELT materials using the internet and corpus resources". In Lewandowska-Tomaszczyk, B. (dir.). *Practical applications in language corpora. Lodz studies in language 7 (Palc 2001*). Frankfurt am Main, Berlin, Bern, Bruxelles, New York, Oxford, Wien : Peter Lang. pp. 521-531.

Cobb, T. (1997). "Is there any measurable learning from hands-on concordancing?". *System*, vol. 25, n° 3. pp. 301-315. Disponible en ligne : http://www.er.uqam.ca/nobel/r21270/cv/Hands_on.html

Cobb, T. (nd). *Konkordanzer-Corpus Deutsch (v5.1).* http://www.lextutor.ca/concordancers/concord_g.html

Cobb, T. (nd). *Text-Concordancer (v2).* http://www.lextutor.ca/concordancers/text_concord/

Cobb, T. (nd). *The Compleat Lexical Tutor*. http://www.lextutor.ca/

COSMAS II (nd). Outil d'exploration de *DeReKo* (Deutsches ReferenzKorpus), proposé par l'IDS (Institut für Deutsche Sprache), Mannheim. https://cosmas2.ids-mannheim.de/cosmas2-web/

Das Digitale Wörterbuch der deutschen Sprache (DWDS). Berlin-Brandenburgische Akademie der Wissenschaft, Berlin. http://www.dwds.de/

DeReKo (Deutsches ReferenzKorpus). IDS (Institut für Deutsche Sprache), Mannheim. Corpus interrogeable à l'aide du concordancier *COSMAS II*. http://www.ids-mannheim.de/kl/projekte/dereko_I/

Diniz, L. (2005). "Comparative review: TEXTSTAT 2.5, ANTCONC 3.0, and COMPLEAT LEXICAL TUTOR 4.0". *Language Learning & Technology*, vol. 9, n° 3. pp. 22-27. http://llt.msu.edu/vol9num3/pdf/review2.pdf

Divsic, D., Rott, S. & Horst, M. (2008). *Braun Corpus.* Corpus interrogable à l'aide du concordancier *Konkordanzer-Corpus Deutsch (v5.1).* http://www.lextutor.ca/concordancers/braun_info.html

Dodd, B. (1997). "Exploiting a corpus of written German for advanced language learning". In Wichmann, A., Fligelstone, S., McEnery, T. & Knowles, G. (dir.). *Teaching and Language Corpora.* London : Longman. pp. 131-145.

Duffner, R. & Näf, A. (2006a). "Korpuslinguistik im Zeitalter der Textdatenbanken". *Linguistik online* 28, 3/06. pp. 1-4. http://www.linguistik-online.de/28_06/index.html

Duffner, R. & Näf, A. (2006b). "Digitale Textdatenbanken im Vergleich". *Linguistik Online* 28, 3/06. pp. 7-23. http://www.linguistik-online.de/28_06/duffnerNaef.html

Falko. Corpus d'apprenants. Institut für Deutsche Sprache und Linguistik, université Humboldt, Berlin. http://www.linguistik.hu-berlin.de/institut/professuren/korpuslinguistik/forschung/falko

Fink, L. & Frahling, G. (2008). *Linguee.com. The Web as a Dictionary.* http://www.linguee.com/PHP/about.php

Fleury, S., Lamalle, C., Martinez, W. & Salem, A. (2003). *Lexico3. Outils de statistique textuelle*. Paris : Université de Paris 3. http://www.cavi.univ-paris3.fr/Ilpga/ilpga/tal/lexicoWWW/

Flowerdew, John (1996). "Concordancing in Language Learning". In Pennington, M. (dir.). *The Power of CALL*. Houston XT : Athelstan. pp. 97-113. Disponible en ligne : http://sunzi1.lib.hku.hk/hkjo/view/10/1000064.pdf

Gaskell, D. & Cobb, T. (2004). "Can learners use concordance feedback for writing errors?" *System*, vol. 32, n° 3, septembre 2004. pp. 301-319. Disponible en ligne : http://www.er.uqam.ca/nobel/r21270/cv/conc_fb.htm

Goodwin-Jones, B. (2001). "Emerging Technologies. Tools and Trends in Corpora Use for Teaching and Learning." In *Language Learning & Technology*, vol. 5, n° 3. pp. 7-12. http://llt.msu.edu/vol5num3/emerging/

Granger, S. & Hung, J. & Petch-Tyson, S. (dir.) (2002). *Computer Learner Corpora, Second Language Acquisition and Foreign Language Teaching*. Amsterdam, Philadelphia : John Benjamins Publishing Company.

Habert, B. (2000). "Des corpus représentatifs : de quoi, pour quoi, comment ?" In Bilger, M. (dir.). *Linguistique sur corpus. Études et réflexions*. *Actes de la journée Linguistique et corpus*. Pergpignan : Presses Universitaires de Perpignan. pp. 11-58. Disponible en ligne : http://atala.biomath.jussieu.fr/je/010428/Habert/Perpignan00/Perpignan00.html

Habert, B. & Nazarenko, A. & Salem, A. (1997). *Les linguistiques de corpus*. Paris : Armand Colin.

Henry, A. & Roseberry, R. L. (1996). "A Corpus-based Investigation of the Language and Linguistic Patterns of One Genre and the Implications for Language Teaching". *Research in the Teaching of English*, 1996, n° 30. pp. 472-489.

Henry, A. & Roseberry, R. L. (2001). "Using a small corpus to obtain data for teaching a genre". In Ghadessy, M. & Henry, A. & Roseberry, R. L. (dir.). *Small Corpus Studies and ELT*. Amsterdam/Philadelphia : John Benjamins Publishing Compagny. pp. 93-133.

Institut für Kommunikationswissenschaften. Université de Bonn. http://www.ikp.uni-bonn.de/forschung/computerlinguistik/computerlinguistik

Johns, T. (1988). "Whence and Whither Classroom Concordancing?". In Bongaerts, T., de Haan, P., Lobbe, S. & Wekker, H. (dir.). *Computer Applications in Language Learning*. Dordrecht : Foris. pp. 9-27.

Johns, T. (1991). "Should you be persuaded - two samples of data-driven learning materials". In Johns, T. & King, P. (dir.) (1991). pp.1-13.

Johns, T. (2000). *Ustí nad Labem DDL Workshop: 21st – 25th March 2000. Participant's work.* Disponible sur : http://www.eisu2.bham.ac.uk/johnstf/unl_ddl.htm

Johns, T. (2002). "Data-driven Learning: the Perpetual Challenge". In Kettemann, B. & Marko, G. (dir.). *Teaching and Learning by Doing Corpus Analysis*. Amsterdam, New York : Éditions Rodopi B.V. pp. 107-117.

Johns, T. (nd). *Kibbitzers. English for International Students Unit*. http://www.eisu2.bham.ac.uk/Webmaterials/kibbitzers/indexes/indexdisc.htm

Johns, T. & King, P. (dir.) (1991). *Classroom concordancing*. *English language research journal* 4. Birmingham : Birmingham University.

Kaltenböck, G. & Mehlmauer-Larcher, B. (2005)."Computer corpora and the language classroom : on the potential and limitations of computer corpora in language teaching". *ReCALL*, vol. 17, n° 1. Cambridge : Cambridge University Press. pp. 65-84.

Kaur, J. & Hegelheimer, V. (2005). "ESL Students' Use of Concordance in the Transfer of Academic Word Knowledge: An Exploratory Study". *Computer assisted language learning*, vol. 18. pp. 287-310.

Kennedy, C. & Miceli, T. (2001)."An evaluation of intermediate student's approaches to corpus investigation." *Language Learning & Technology*, vol. 5, n° 3, septembre 2001. pp. 77-90. http://llt.msu.edu/vol5num3/kennedy/

Kilgarriff, A., Rychly, P. & Pomikalek J. (nd). *Sketch Engine*. Collection de corpus. http://www.sketchengine.co.uk/

Kucera, H. & Francis, W. N. (1967). *Computational analysis of present-day American English.* Providence : Brown University Press.

Lamy, M.-N. & Klarskov Mortensen, H. J. (2006). *ICT4LT (Information and Communications Technology for Language Teachers), module 2.4 : Using concordance programs in the Modern Foreign Languages classroom*. http://www.ict4lt.org/en/en_mod2-4.htm

Lebart, L. & Salem, A. (1994). *Statistique Textuelle*. Paris : Dunod. Disponible en ligne : http://ses.telecom-paristech.fr/lebart/ST.html

Lee, D. (nd). *David Lee's Bookmarks for Corpus-based Linguists.* http://personal.cityu.edu.hk/~davidlee/devotedtocorpora/CBLLinks.htm

Leech, G. (2003). "Corpus Processing". In Frawley, W. J. (dir.). *International Encyclopedia of Linguistics*. Oxford University Press. Disponible en ligne : http://www.oxford-linguistics.com/entry?entry=t202.e0254

Lemnitzer, L. & Zinsmeister, H. (2006). *Korpuslinguistik – Eine Einführung*. Tübingen : Gunter Narr Verlag.

LIMAS (Linguistik und Maschinelle Sprachbearbeitung). Université de Bonn. Voir *DeReKo*. http://www.korpora.org/Limas/

Lüdeling, A. & Walter, M. (2009). "Korpuslinguistik für Deutsch als Fremdsprache – Sprachvermittlung und Spracherwerbsforschung". Version élargie de "Korpuslinguistik" (à paraître). Fandrych, C., Hufeisen, B., Krumm, H.-J. & Riemer, C. (dir.). *Deutsch als Fremd- und Zweitsprache. Ein internationales Handbuch. 2. überarbeitete Auflage.* (= HSK, vol. 19). Berlin, New York : Mouton de Gruyter. Disponible sur : http://www.linguistik.hu-berlin.de/institut/professuren/korpuslinguistik/mitarbeiter-innen/anke/pdf/LuedelingWalterDaF.pdf

Mayaffre, D. (2005). "Rôle et place des corpus en linguistique : réflexions introductives". *Texto!*, vol. 10, n° 4. http://www.revue-texto.net/1996-2007/Corpus/Publications/Mayaffre_Corpus.html

Meyer, C. F. (2002). *English Corpus Linguistics: An Introduction*. Cambrige : Cambridge University Press.

Mukherjee, J. (2002). *Korpuslinguistik und Englischunterricht: Eine Einführung*. Frankfurt am Main, Berlin, Bern, Bruxelles, New York, Oxford, Wien : Peter Lang.

Murison-Bowie, S. (1993). *MicroConcord Manual*. Oxford : Oxford University Press.

NEGRA Korpus Version 2. Université de Saarbrücken. http://www.coli.uni-saarland.de/projects/sfb378/negra-corpus/negra-corpus.html

O'Keeffe, A. & McCarthy, M. & Carter, R. (2007). *From Corpus to Classroom. Language Use and Language Teaching.* Cambridge : Cambridge University Press.

Partington, A. (1998). *Patterns and Meanings. Using Corpora for English Research and Teaching.* Amsterdam/Philadelphia : John Benjamins Publishing Company.

Partington, A. (2006). "Aims, Tools and Practices of Corpus Linguistics". *IntUne Papers. Integrated and United: A quest for Citizenship in an 'ever closer Europe'.* http://209.85.129.132/search?q=cache:I1Gw-k_ecIEJ:www.intune.it/file_download/23+Partington+Alan+Aims+tools&cd=1&hl=fr&ct=clnk&gl=de

Petzold, D. (nd). *Zoo-Datenbank für Deutschland*. http://www.zoo-infos.de/

Pincemin, B. (2006). "Concordances et concordanciers – De l'art du bon KWAC. Soumission à *Documents numériques et interprétation – Corpus en Lettres et Sciences sociales*, Albi, 10-14 juillet 2006.

Rastier, F. (2004). "Enjeux épistémologiques de la linguistique de corpus". *Texto !*, juin 2004. http://www.revue-texto.net/1996-2007/Inedits/Rastier/Rastier_Enjeux.html

Rézau, J. (2007). "L'apport du concordancier à l'analyse et à la remédiation des erreurs des apprenants dans les forums de discussion en ligne". *Apprentissage des langues et systèmes d'information et de communication (ALSIC)*, vol. 10, n° 2. pp. 27-43. http://alsic.u-strasbg.fr/v10/rezeau/alsic_v10_04-pra1.htm

Ruthven-Stuart, P. (nd). *How to use concordances in teaching English.* http://www.nsknet.or.jp/~peterr-s/concordancing/usingconcs.html

Schiller, A., Thielen, C., Teufel, S. & Stöckert, C. (nd). *STTS (Stuttgart/Tübinger Tagset).* http://www.ifi.uzh.ch/arvo/cl/tagger/UIS-STTS-Diffs.html

Schlobinski, P. (nd). *Schlobi's Linguistic Corner*. http://www.linguistic-corner.uni-hannover.de/_sprachkorpora.html

Sharoff, S. (nd). *French web corpus*. Corpus interrogeable sur le site de *Sketch Engine* (Kilgarriff *et al.).*

Siepmann, D. (1999). "Linguistique de corpus et grammaire : une nouvelle donne pour l'enseignement du français langue étrangère". *französisch heute* 3/1999. Seelze, Velber : KallmeyerscheVerlagsbuchhandlung GmbH. pp. 309-327.

Sinclair, J. (1996). *EAGLES. Preliminary Recommendations on Corpus Typology*. Pisa : Consiglio Nazionale delle Ricerche. Istituto di Linguistica Computazionale. Disponible en ligne : http://www.ilc.cnr.it/EAGLES96/corpustyp/corpustyp.html

Sinclair, J. (2004). *Trust the text. Language, Corpus and Discourse.* London : Routledge.

Sinclair, J. (2005). "Corpus and Text — Basic Principles". In Wynne, Martin (2005). *Developing Linguistic Corpora: a Guide to Good Practice*. Oxford : Oxbow Books. pp. 1-16. Disponible en ligne : http://ahds.ac.uk/linguistic-corpora/

Stevens, V. (1988). "Studying vocabulary using concordances on microcomputers*"*. *TESOL Newsletter* 22, 3/27. http://www.vancestevens.com/papers/tesol/newsletter/1988online.htm

Stevens, V. (1991). "Concordance-based vocabulary exercises: A viable alternative to gap-fillers". In Johns, T. & King, P. (dir.). pp. 47-61.

Stevens, V. (1995). "Concordancing with Language Learners: Why? When? What?". *CAELL Journal (Computer Assisted English. Language Learning),* vol. 6, n° 2. pp. 2-10.

St. John, E. (2001). "A case for using parallel corpus and concordancer for beginners of a foreign langage". *Language Learning & Technology*, vol. 5, n° 3, septembre 2001. pp. 185-203. http://llt.msu.edu/vol5num3/stjohn/default.html

Sun, Y.-C. & Wang, L.-Y. (2003). "Concordancers in the EFL classroom: cognitive approaches and collocation difficulty". *CALL*, vol. 16/1. pp. 83-94.

Teubert, W. (2005). "My version of corpus linguistics". *International Journal of Corpus Linguistics*, vol. 10, n° 1, 2005. pp. 1-13.

Teubert, W. (2006). "Korpuslinguistik, Hermeneutik und die soziale Konstruktion der Wirklichkeit". *Linguistik Online* 28, 3/06. pp. 41-60. http://www.linguistik-online.de/28_06/teubert.html

Tiedemann, J. (2009). *Opus*. *An open source parallel corpus*. http://urd.let.rug.nl/tiedeman/OPUS/

TIGERCorpus. Institut für Phonetik/Phonologie, Universität des Saarlandes & Institut für Germanistik, Potsdam & Institut für Maschinelle Sprachverarbeitung, Stuttgart. http://www.ims.uni-stuttgart.de/projekte/TIGER/TIGERCorpus/

TreeTagger. Institut für maschinelle Datenverarbeitung, Stuttgart. http://www.ims.uni-stuttgart.de/ims-home.html.en

Tribble, C. (1997). "Improvising Corpora for ELT: Quick and Dirty Ways of Developing Corpora for Language Teaching". In Lewandowska-Tomaszczyk, B. and Melia, P. J. (dir.). *Proceedings of the First International conference on Practical Applications in Language Corpora (Palc 1997).* Lodz : Lodz University Press.

WaCky (*Web-as-Corpus kool ynitiative*). Site donnant accès à des corpus regroupés dans *Sketch Engine* (Kilgarriff *et al.*). http://wacky.sslmit.unibo.it/doku.php?id=corpora

Wortschatz-Portal. Université de Leipzig. http://wortschatz.uni-leipzig.de/

Zanettin, F. (2001). "Swimming in Words: Corpora, Translation, and Language Learning". In Aston, G. (dir.). *Learning with Corpora*. Houston XT : Athelstan.

Liste des figures

Figure 1 – Évolution de l'utilisation du terme "*Korpus*" entre 1900 et 1990. ……... 19
Figure 2 – Position des corpus par rapport à d'autres ensembles publiant du texte. …. 23
Figure 3 – Ma définition de "corpus". …… 27
Figure 4 – Types d'annotation d'après Partington (2006). …….. 30
Figure 5 – Fonction en ligne "rechercher dans cette page…". …….. 35
Figure 6 – Système de recherche d'informations sur le site du journal *Le Monde*. …… 35
Figure 7 – Localisation du mot clé "werden" dans deux corpus de référence. …….. 36
Figure 8 – Ma définition de "concordance". …….. 38
Figure 9 – Mot clé "man", recherché dans le corpus "Visite guidée". ……... 38
Figure 10 – Segment "Schule geht", recherché dans le corpus *DeWac* (Baroni, 2006). 39
Figure 11 – Cluster pour "*Eintritt* [entrée]" dans le corpus "Visites guidées". ……. 39
Figure 12 – Lignes de concordance sans le mot clé "Museum". …… 42
Figure 13 – L'évolution de la taille des corpus de langue allemande dans *DeReKo*. … 47
Figure 14 – Options d'exploration pour *French Web Corpus* et pour *DeWac*. …….. 48
Figure 15 – *Input*, *intake* et *output*. …… 54
Figure 16 – Traitement local des données. ……. 57
Figure 17 – Ma définition de "collocation". …….. 64
Figure 18 – Illustration du modèle spiralaire de Gläser. …… 65
Figure 19 – Modèle d'apprentissage par exploitation de la boucle phonologique. …… 68
Figure 20 – Repérage d'un élément dans l'*input*. …… 74
Figure 21 – Structure du modèle d'écriture de Hayes et Flower. ……... 93
Figure 22 – Modèle de production écrite de Fayol. …… 95
Figure 23 – Modélisation du processus de formulation en LE par Zimmermann. …… 97
Figure 24 – De l'hypertexte au texte (Hallet). …… 110
Figure 25 – Mot clé "personne ressource/personnes ressources" dans *French Web Corpus*. ……... 115
Figure 26 – Interface de *English–German Translation Corpus*. …… 125
Figure 27 – Perception holistique, synthétique et analytique de l'objet "langue". ……. 134
Figure 28 – Mouvements *top down* et *bottom up* et corpus. ……. 135
Figure 29 – Consignes initiales pour l'activité "Je t'invite à mon anniversaire". ……... 145
Figure 30 – Consignes définitives pour l'activité "Je t'invite à mon anniversaire". …... 146
Figure 31 – Approches et horizons de la méthode en sciences humaines. ……. 152
Figure 32 – "Informations pratiques" d'un parc d'attraction. ……. 165
Figure 33 – Activité de recherche lexicale "*Find the Gloups*". ……. 173
Figure 34 – Activité de type "glisser–déposer". ……. 177
Figure 35 – Activité de type "surlignage de sections de texte". ……. 178
Figure 36 – Activité de type "rajout de binettes aux sections identifiées". …… 178
Figure 37 – Procédure d'identification et de réemploi d'un segment. …… 179
Figure 38 – Révision guidée à l'aide de lignes de concordance. …… 180
Figure 39 – Exemple d'une tâche exécutée collectivement. ……... 181
Figure 40 – Annotation du texte "Visite guidée, V3" de Simon. ……... 185
Figure 41 – Sens d'une lecture exhaustive d'un texte de corpus. …… 198
Figure 42 – Sens spatial de la lecture à partir d'un mot clé dans une concordance. ….. 199
Figure 43 – Lignes de concordance pour le mot clé "Klarinett*". ……. 221
Figure 44 – Résultats de recherche pour "Klarinettenspielerin" (*Exalead*). …….. 223
Figure 45 – Traces d'entraînement avec des lignes de concordance. …… 224
Figure 46 – Activité de révision collaborative. …….. 225
Figure 47 – Lignes de concordance proposées par Text–Concordancer (v2), pour **sympatisch*. …….. 226

Figure 48 – Lignes de concordance proposées par AntConc pour *symp**. …… 226
Figure 49 – Style parataxique chez Janine. …… 229
Figure 50 – Adjectifs dans les productions d'Alexa. …….. 231
Figure 51 – Adjectifs dans les productions de Camille. …… 232
Figure 52 – Adjectifs dans les productions de Chloé. …… 233
Figure 53 – Adjectifs dans les productions de Christine. …….. 234
Figure 54 – Adjectifs dans les productions de Fernand. …… 235
Figure 55 – Adjectifs dans les productions de Georges. …… 236
Figure 56 – Adjectifs dans les productions de Géraldine. ……. 237
Figure 57 – Adjectifs dans les productions de Janine. ……... 238
Figure 58 – Adjectifs dans les productions de Jean (1). ……. 239
Figure 59 – Adjectifs dans les productions de Jean (2). ……. 240
Figure 60 – Adjectifs dans les productions de Léa (1). …….. 241
Figure 61 – Adjectifs dans les productions de Léa (2). …….. 242
Figure 62 – Adjectifs dans les productions de Léa (3). …….. 242
Figure 63 – Adjectifs dans les productions d'Olivia. …….. 243
Figure 64 – Adjectifs dans les productions de Simon. ……... 244
Figure 65 – Adjectifs dans les productions de Vicky. …… 245
Figure 66 – Évolution des adjectifs lors de *Projekt Prospekt*. ……... 247
Figure 67 – Évolution des types d'erreurs dans le domaine de l'adjectif. ……... 250
Figure 68 – "Visite guidée" d'Alexa (V2). ……. 253
Figure 69 – Syntaxe dans les productions d'Alexa. …… 254
Figure 70 – "Commentaire dans un livre d'or" de Géraldine (V2). …… 254
Figure 71 – Syntaxe dans les productions de Géraldine. …… 255
Figure 72 – "Commentaire dans un livre d'or" de Vicky (V0). …….. 256
Figure 73 – Syntaxe dans les productions de Vicky. …….. 257
Figure 74 – Points d'exclamation dans des messages d'adolescents sur un blogue. …... 264
Figure 75 – Types d'intervention dans les annotations pour *Projekt Prospekt* et de *Brieffreunde.org*. …… 281
Figure 76 – Questionnaire intermédiaire pour *Projekt Prospekt*. …….. 285
Figure 77 – Questionnaire final pour "Projekt Prospekt". …….. 286
Figure 78 – "Critique de film" de Camille. …… 289
Figure 79 – Échanger avec des pairs sur un blogue. …….. 304
Figure 80 – Repérage de catégories linguistiques dans un graphique *Wordle*. …….. 306

Liste des tableaux

Tableau 1 – Archives en ligne en langue allemande. ……... 23
Tableau 2 – Hiérarchie entre corpus, document et texte dans *DeReKo*. ……. 24
Tableau 3 – *The "anatomy" and "physiology" of language* (Firth). ……. 30
Tableau 4 – Postures vis–à–vis des corpus et leurs implications théoriques et méthodologiques. …….. 32
Tableau 5 – Comparaison de six concordanciers. ……. 40
Tableau 6 – Corpus proposés par *Sketch Engine*. ……. 43
Tableau 7 – Unités polylexicales et courants/domaines de recherche. ……. 60
Tableau 8 – Éléments pour un précis du bon usage du guidage. …….. 83
Tableau 9 – Caractéristiques d'une tâche. ……. 86
Tableau 10 – Tâches traditionnelles, tâches nouvelles (Perrenoud). …… 87
Tableau 11 – Produit et processus dans des modèles d'écriture. …….. 92
Tableau 12 – *Input*, outils et phases de rédaction. …… 119
Tableau 13 – Genres textuels et prise en charge énonciative. ……. 128
Tableau 14 – Nombre d'occurrences pertinentes pour "*Eintritt*" [entrée] dans quatre corpus. ……... 129
Tableau 15 – Méthodologies associées à l'utilisation de corpus et concordances pour l'enseignement/apprentissage d'une langue étrangère. …….. 131
Tableau 16 – Concordance de "laden + ein", faite à l'aide de Sketch Engine. ……. 144
Tableau 17 – *Projekt Prospekt* et le programme officiel pour l'allemand, palier 2. ……. 158
Tableau 18 – Types de dossiers déposés dans le cadre du droit à l'expérimentation. ….. 160
Tableau 19 – Références pour les trois phases de la recherche–action. ……... 162
Tableau 20 – Caractéristiques linguistiques des sections d'un texte de type "Visite guidée". ……. 176
Tableau 21 – Activités de révision collectives. …… 181
Tableau 22 – Comparaison des moyens de communication utilisés par les apprenants.. 184
Tableau 23 – Extrait de la grille utilisée pour le *DCL*. ……. 188
Tableau 24 – *The Five Requirements* (Siepmann). ……... 191
Tableau 25 – Critères pour l'évaluation des tâches de *Projekt Prospekt*. ……. 192
Tableau 26 – Noms et lemmes des verbes fréquents dans les entretiens et journal de bord. …….. 196
Tableau 27 – Caractéristiques intertextuelles des productions finales. …… 212
Tableau 28 – Segments co–présents et quantité d'*input* accessible aux apprenants. …… 213
Tableau 29 – Passages longs co–présents et quantité d'*input* accessible aux apprenants. 214
Tableau 30 – Segments co–présents et erreurs de syntaxe. …….. 215
Tableau 31 – Émergence de mots composés lors de la rédaction des "Informations pratiques". ……. 217
Tableau 32 – Fréquence de *Klarinettenspieler(in)* et de *Klarinettist(in)* dans des corpus et des moteurs de recherche. . ……..……. 222
Tableau 33 – Comparaison de fréquence des adjectifs *gut*, *schön* et *interessant*. …… 227
Tableau 34 – Formes adjectivales fléchies et non fléchies dans les corpus "Visites guidées". ……..…….. 228
Tableau 35 – Adjectifs dans les productions d'Alexa. …….. 230
Tableau 36 – Adjectifs dans les productions de Camille. ……. 231
Tableau 37 – Adjectifs dans les productions de Chloé. …… 233
Tableau 38 – Adjectifs dans les productions de Christine. ……... 234
Tableau 39 – Adjectifs dans les productions de Fernand. …… 235
Tableau 40 – Adjectifs dans les productions de Georges. …… 236
Tableau 41 – Adjectifs dans les productions de Géraldine. …….. 237

Tableau 42 – Adjectifs dans les productions de Janine. ……... 238
Tableau 43 – Adjectifs dans les productions de Jean. …….. 239
Tableau 44 – Adjectifs dans les productions de Léa. ……... 241
Tableau 45 – Adjectifs dans les productions d'Olivia. ……. 243
Tableau 46 – Adjectifs dans les productions de Simon. ……... 244
Tableau 47 – Adjectifs dans les productions de Vicky. …… 245
Tableau 48 – Adjectifs dans les productions du groupe classe (1). …….. 247
Tableau 49 – Adjectifs dans les productions du groupe classe (2). …….. 247
Tableau 50 – Adjectifs utilisés de façon répétitive dans les productions de départ. …… 249
Tableau 51 – Erreurs du groupe classe dans le domaine de l'adjectif. …….. 250
Tableau 52 – Syntaxe dans les productions d'Alexa. …… 253
Tableau 53 – Syntaxe dans les productions de Géraldine. ……... 255
Tableau 54 – Syntaxe dans les productions de Vicky. ……. 257
Tableau 55 – Points d'exclamation dans le corpus de référence "Visite guidée". …… 261
Tableau 56 – Points d'exclamation dans la production finale des apprenants. ……. 262
Tableau 57 – La version V1 des "Informations pratiques" d'Olivia. …… 265
Tableau 58 – Interface de la boîte courriel pour *Projekt Prospekt*. …….. 268
Tableau 59 – Types d'intervention lors de l'annotation des textes de *Projekt Prospekt*. .. 280
Tableau 60 – Degrés d'intervention lors de *Projekt Prospekt* et de *Brieffreunde.org*. …. 282
Tableau 61 – Prise en compte du lecteur de mes annotations dans *Brieffreunde.org*. …. 284
Tableau 62 – Indices de décentration des apprenants dans le questionnaire final. ……... 288
Tableau 63 – Noms de scripteurs de "Commentaires dans un livre d'or". …… 290

Index des notions

Allemand, 13, 19, 23, 40, 42, 43, 44, 45, 47, 48, 67, 70, 85, 98, 104, 107, 112, 113, 114, 119, 122, 123, 124, 125, 126, 144, 147, 154, 155, 156, 158, 159, 161, 163, 165, 172, 174, 177, 181, 183, 197, 199, 201, 204, 205, 206, 209, 228, 239, 248, 252, 264, 267, 270, 272, 273, 274, 277, 278, 280, 282, 283, 284, 285, 286, 288, 290, 291, 292, 293, 299, 304, 306, 326, 347

Annotation, 29, 30, 31, 33, 43, 182, 185, 210, 217, 220, 275, 278, 279, 280, 281, 282, 283, 284, 345, 347

Apprenant-chercheur, 15, 76, 78, 131, 136, 142, 162, 189, 197, 198, 275, 278

Bloc lexicalisé, 61

Blogue, 119, 120, 155, 159, 163, 164, 165, 166, 167, 168, 169, 170, 171, 172, 174, 179, 182, 184, 185, 205, 218, 263, 264, 270, 271, 272, 279, 281, 283, 286, 303, 304, 345

Chunk, 39, 59, 69, 70, 71, 163, 251

Collaboration, 46, 80, 124, 151, 170, 274, 295

Collocation, 39, 49, 59, 60, 61, 62, 63, 64, 65, 67, 68, 69, 70, 71, 89, 108, 113, 138, 140, 141, 143, 161, 163, 191, 203, 227, 249, 333, 334, 343, 345

Concordance, 3, 4, 15, 18, 36, 37, 38, 39, 40, 41, 42, 52, 56, 59, 65, 69, 70, 72, 76, 77, 78, 89, 90, 105, 108, 113, 115, 117, 119, 120, 121, 122, 126, 127, 130, 131, 132, 133, 134, 135, 136, 137, 138, 139, 140, 141, 142, 143, 144, 145, 146, 147, 154, 158, 161, 162, 163, 164, 165, 166, 167, 168, 171, 173, 175, 179, 180, 181, 184, 185, 187, 188, 189, 192, 193, 194, 196, 197, 198, 199, 200, 201, 202, 203, 204, 205, 206, 207, 208, 209, 210, 211, 212, 214, 215, 216, 219, 220, 221, 222, 224, 225, 226, 227, 228, 229, 230, 232, 233, 240, 243, 246, 247, 248, 249, 251, 252, 254, 258, 259, 265, 266, 267, 268, 271, 272, 273, 275, 276, 278, 279, 280, 282, 283, 284, 286, 289, 295, 297, 298, 299, 301, 303, 315, 320, 324, 330, 333, 336, 338, 339, 341, 342, 343, 345, 347

Concordancier, 15, 34, 35, 36, 38, 40, 41, 42, 46, 102, 108, 119, 125, 127, 130, 132, 134, 136, 163, 167, 194, 200, 202, 205, 206, 226, 272, 298, 317, 318, 328, 330, 338, 342, 347

Courriel, 46, 107, 163, 168, 169, 170, 171, 183, 184, 205, 218, 263, 265, 268, 272, 275, 283, 293, 303, 304, 347

Décentration, 187, 265, 286, 287, 288, 290, 318, 347

Distance, 52, 89, 95, 101, 151, 160, 183, 258, 279, 286, 288, 299, 318

Erreur, 34, 44, 70, 74, 81, 82, 95, 112, 113, 180, 202, 203, 204, 209, 210, 215, 218, 225, 229, 230, 231, 233, 234, 235, 236, 237, 238, 239, 241, 243, 244, 245, 250, 251, 252, 254, 255, 257, 266, 269, 271, 279, 280, 281, 297, 298, 299, 306, 330, 342, 345, 347

Formule, 55, 60, 63, 65, 71, 108, 112, 146

Genre textuel, 14, 27, 28, 50, 56, 63, 64, 71, 91, 98, 102, 103, 104, 117, 118, 119, 123, 124, 126, 129, 144, 148, 165, 175, 177, 212, 264, 265, 267, 270, 283, 294, 297, 298

Guidage, 2, 49, 55, 77, 79, 81, 83, 84, 88, 90, 101, 119, 120, 121, 135, 139, 140, 166, 180, 194, 195, 197, 199, 275, 286, 314, 316, 347

Input, 2, 15, 49, 54, 55, 56, 57, 58, 59, 60, 61, 65, 69, 70, 71, 72, 73, 74, 75, 76, 78, 89, 92, 98, 100, 101, 104, 105, 106, 107, 109, 119, 120, 131, 134, 135, 161, 170, 187, 194, 196, 197, 198, 200, 201, 202, 203, 205, 206, 208, 213, 214, 216, 218, 219, 264, 267, 271, 275, 276, 278, 286, 293, 294, 298, 299, 333, 345, 347

Intake, 49, 54, 55, 57, 58, 74, 345

Intertexte, 2, 98, 104, 105, 106, 131, 154, 195, 312

Lecture, 2, 39, 72, 92, 98, 104, 109, 114, 116, 123, 132, 136, 137, 168, 174, 186, 187, 188, 189, 197, 198, 199, 200, 202, 204, 226, 272, 298, 306, 325, 345

Lexème, 39, 42, 66, 90, 124, 125, 181, 186, 188, 210, 211, 212, 213, 216, 217, 227, 272, 280, 283, 288

Littératie électronique, 90, 110, 123, 164, 197, 267, 270, 272, 273, 299

Nativisation, 58, 118, 192, 223, 289, 317

Observation, 2, 18, 27, 33, 51, 53, 58, 67, 70, 77, 83, 88, 90, 91, 98, 100, 102, 107, 117, 120, 123, 130, 134, 137, 138, 139, 141, 142, 143, 144, 145, 147, 148, 149, 150, 151, 152, 153, 164, 179, 181, 182, 186, 187, 188, 194, 198, 200, 205, 211, 213, 216, 225, 226, 227, 249, 251, 252, 266, 270, 272, 274, 282, 305, 307

Output, 15, 49, 54, 55, 57, 58, 78, 131, 139, 162, 227, 333, 345
Planification, 13, 93, 95, 96, 126, 163, 165, 177
Production écrite, 48, 54, 56, 63, 64, 90, 91, 92, 93, 94, 95, 96, 98, 99, 100, 101, 102, 105, 106, 107, 109, 110, 111, 115, 117, 118, 120, 122, 137, 139, 140, 147, 155, 161, 163, 165, 177, 178, 191, 193, 194, 206, 214, 215, 230, 231, 233, 234, 235, 236, 237, 238, 239, 241, 243, 244, 245, 247, 248, 250, 251, 259, 275, 286, 345
Recherche d'informations, 15, 27, 28, 33, 35, 71, 72, 104, 109, 123, 155, 159, 161, 179, 203, 205, 282, 330, 345
Recherche-action, 2, 16, 65, 81, 89, 90, 95, 98, 106, 108, 111, 112, 113, 114, 118, 120, 123, 126, 127, 130, 134, 141, 148, 149, 150, 151, 152, 153, 154, 161, 162, 164, 166, 167, 168, 169, 170, 171, 172, 173, 175, 179, 180, 182, 183, 184, 186, 193, 195, 205, 218, 219, 224, 252, 271, 274, 281, 285, 295, 311, 326, 327, 347
Repérage, 56, 70, 73, 74, 78, 107, 133, 137, 146, 147, 161, 162, 169, 172, 175, 190, 192, 195, 196, 205, 258, 272, 275, 299
Révision, 31, 55, 56, 73, 93, 94, 95, 96, 98, 102, 105, 117, 118, 119, 120, 139, 141, 147, 161, 163, 164, 167, 172, 177, 179, 180, 181, 195, 199, 217, 219, 220, 221, 224, 225, 226, 228, 251, 264, 273, 280, 281, 297, 301, 315, 345, 347
Scénario pédagogique, 63, 98, 118, 119, 120, 161, 190, 293, 296, 308
Scripteur, 14, 57, 92, 93, 94, 95, 96, 98, 99, 101, 104, 105, 109, 118, 147, 258, 288, 291, 292, 294
Segment, 20, 29, 36, 39, 42, 58, 62, 64, 68, 70, 71, 72, 89, 102, 105, 106, 112, 117, 118, 124, 134, 135, 161, 163, 166, 177, 179, 181, 185, 187, 188, 189, 190, 192, 202, 210, 211, 212, 213, 215, 216, 217, 218, 219, 220, 228, 229, 230, 231, 232, 233, 234, 235, 236, 237, 238, 239, 240, 241, 242, 243, 244, 245, 246, 247, 248, 249, 251, 253, 255, 256, 257, 267, 272, 275, 276, 277, 278, 280, 283, 288, 291, 294, 298, 345
Tâche, 31, 33, 48, 49, 54, 55, 56, 59, 65, 72, 73, 74, 81, 82, 83, 84, 85, 86, 87, 89, 91, 93, 95, 96, 98, 105, 107, 111, 115, 117, 118, 119, 120, 121, 122, 123, 124, 130, 131, 132, 133, 134, 135, 136, 137, 138, 142, 143, 145, 147, 149, 157, 161, 162, 164, 165, 166, 170, 174, 176, 179, 180, 181, 182, 188, 190, 192, 193, 194, 196, 208, 218, 219, 223, 270, 275, 283, 287, 292, 293, 294, 295, 317, 327, 332, 345, 347