

Exploiter des concordances pour la rédaction d'annonces allemandes

Eva Schaeffer-Lacroix, maître de conférences à
IUFM de Paris, école interne de Paris 4.

Résumé

L'exploitation de lignes de concordance peut soutenir l'enseignement-apprentissage de la production écrite. J'ai étudié la façon dont des adolescents en tirent profit lors d'un projet de production d'annonces pour trouver un correspondant. Dans cet article, je présenterai d'abord l'arrière-plan méthodologique. Je décrirai ensuite le scénario d'apprentissage *Brieffreunde.org* que j'ai développé et testé en classe de troisième. Le projet a permis aux apprenants de réviser leurs premiers jets en semi-autonomie et de produire des textes qui respectent suffisamment les normes linguistiques et discursives du genre textuel "annonce pour trouver un correspondant".

Mots clés : allemand, production écrite, annonces, tâches constructivistes, genre textuel, corpus, lignes de concordance

Keywords : German, writing tasks, penpal ads, constructivism, text genre, corpora, concordance lines

1. Introduction

Comment aider à rédiger en langue étrangère sans imposer des solutions préconçues ? Comment faire en sorte que la réflexion des apprenants concernant la pertinence de leur façon d'écrire ait des chances de se mettre en marche ? Le recours à des corpus réfléchis, donc des banques de données dédiées à un genre textuel en particulier, semble offrir des possibilités intéressantes pour l'enseignement-apprentissage de la production écrite. J'ai étudié la façon dont des adolescents français en troisième ou quatrième année d'allemand perçoivent et exploitent des corpus lors d'un projet de production de textes qui pourraient être publiés sur des sites Internet donnant l'occasion de trouver un correspondant ou une correspondante. Dans cet article, je définirai d'abord les termes "corpus", "concordance" et "collocation". Je présenterai ensuite l'arrière-plan méthodologique applicable à des tâches constructivistes impliquant l'exploitation de corpus et de lignes de concordance. Je décrirai les composantes du scénario d'apprentissage *Brieffreunde.org*, un projet de rédaction d'annonces en allemand, langue étrangère, que j'ai développé et testé lors d'une recherche-action en classe de troisième du collège. Je terminerai par la présentation de quelques résultats de cette recherche qui illustreront la façon dont les apprenants ont révisé leur texte en semi-autonomie.

2. Définitions

2.1. Corpus

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. Un des objectifs des corpus peut être la recherche d'informations dans un contexte d'apprentissage d'une langue étrangère. Les énoncés des textes du corpus

ont un statut de norme linguistique, norme spécifique au genre textuel qui y est représenté, le cas échéant.

2.2. Concordance

Une concordance est une mise en forme particulière des données d'un corpus, obtenue à l'aide d'un logiciel appelé "concordancier". Ce logiciel réorganise matériellement le texte. Il donne l'occasion de rechercher toutes les occurrences d'un mot clé ou d'une combinaison de mots clés dans le corpus. Le résultat de la requête apparaît sous forme de liste des occurrences. Le nombre de caractères des lignes d'une telle liste dépend des options de recherche définies préalablement.

Ich	suche	nette Brieffreundinnen
und	suche	nette Kontakte aus der ganzen Welt
Hallo Leute,	Suche	nette Leute zum Schreiben
Hallo	suche	neue Leute zum Briefe schreiben.
Hallo^^ Ich	suche	selbstverständlich nette Briefkontakte .

Tableau 1 – Lignes de concordance, obtenues avec Sketch Engine.

La figure 1 présente un choix de cinq lignes obtenues pour la requête suivante dans le corpus *Brieffreunde.org* : j'ai recherché tous les noms communs qui se trouvent à une distance entre un à cinq mots à droite du mot clé "suche". Cela a permis d'afficher plusieurs variantes pour le terme "correspondant(e)" : Brieffreundinnen [correspondantes], "Kontakte" [des contacts], "Leute" [des gens] et "Briefkontakte" [des contacts épistolaires]. Les lignes de concordance aident donc à observer les mots clés dans un contexte plus ou moins étroit. Les applications possibles du concordancier sont, entre autres, l'étude de texte (linguistique, discursive, etc.) et l'enseignement-apprentissage de la production écrite en langue étrangère.

2.3. Collocation

Une collocation est une expression qui 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. Une collocation forme une unité de sens. Elle est repérable dans un corpus grâce à sa fréquence significative. Pour donner un exemple, la formule "(ich) freue mich (auf/über...)" [je me réjouis (à l'avance/de...)] apparaît cinq fois dans le petit corpus *Brieffreunde.org*. Je considère cette formule comme une collocation parce qu'elle apparaît fréquemment en fin d'annonces ou de lettres, dans un contexte où on lirait en français des formules qui n'apparaissent pas comme une traduction littérale de "sich freuen", comme, par exemple, "j'ai hâte de te lire!" ou "Écris-moi vite !".

3. Arrière-plan méthodologique

Depuis les années 80, un certain nombre d'études décrit le potentiel des corpus pour enseigner et apprendre une langue étrangère. Dans ce contexte, plusieurs méthodologies sont évoquées. 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)*. Le tableau suivant distingue entre méthodologies ayant une dominante de type *top down* ou une dominante de type *bottom up*. Le terme de *top down* désigne des opérations cognitives lors desquelles le traitement de l'information est dirigé par des concepts. C'est le domaine des représentations des apprenants concernant les caractéristiques de la langue. Le terme de *bottom up* s'applique à des opérations lors desquelles le traitement de l'information est dirigé par les données, donc par du texte qui peut prendre la forme d'un corpus ou de lignes de concordance.

Concept	Constructivisme	
Méthodologie	Plutôt orienté <i>bottom up</i> : <i>noticing</i> [repérage linguistique]	Plutôt orienté <i>top down</i> : approche par tâches apprentissage incident
	Va-et-vient entre des mouvements <i>bottom up</i> et <i>top down</i> : <i>Data Driven Learning</i> apprentissage par la découverte grammaire pédagogique conceptualisation	

Tableau 2 - Méthodologies associées à l'utilisation de corpus

Le tableau 2 regroupe 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. D'autres méthodologies, comme l'apprentissage par tâches, incitent les apprenants à créer un produit. Dans le contexte d'un travail basé sur des corpus, la tâche consiste à appréhender le corpus dans le but de rédiger un texte qui pourrait y être intégré. La méthodologie appelée *Data driven learning* considère les corpus comme éléments constitutifs du travail d'apprentissage. Selon les adeptes de ce courant, les concordances donnent l'occasion d'interroger, de manipuler et de modifier du texte produit par des scripteurs experts. 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 susceptibles de soutenir une attitude réflexive aussi bien de l'enseignant que des apprenants.

Je défends l'idée que l'apprentissage est favorisé par l'alternance des mouvements *top down* et *bottom up* au sein d'un scénario pédagogique. Si les tâches d'apprentissage du scénario incluent une composante d'observation de la langue dans un contexte de sens, elles peu-

vent aider à coordonner les deux mouvements. Dans ce qui suit, je vais présenter le scénario pédagogique *Brieffreunde.org* qui propose la réalisation d'une tâche constructiviste, basée sur l'exploitation de lignes de concordance.

4. Description du projet *Brieffreunde.org*

4.1. Dispositif

Le projet *Brieffreunde.org* a été proposé à 33 collégiens inscrits soit en troisième, soit en quatrième année d'allemand. Il a eu comme objet la rédaction d'une annonce comme celles que l'on trouve sur des sites Internet qui mettent en contact des personnes qui cherchent un correspondant. Une dizaine de séances ont été partiellement dédiées au projet. Le travail a été effectué en salle de classe ordinaire, dotée d'un chariot multimédia avec accès à Internet. Deux séances ont eu lieu en salle informatique, et les apprenants ont été encouragés à travailler à distance, donc à la maison ou au CDI.

4.2. Ressources et outils

J'ai créé le corpus *Brieffreunde.org* à partir d'annonces publiées sur le site Internet portant le même nom (Ingelmann, nd). J'ai obtenu le droit d'utiliser ces données pour la recherche. À l'aide de l'outil *Text Concordancer* (Cobb, nd), j'ai produit un fichier *Word* présentant les données du corpus sous forme de lignes de concordance. Le corpus et la concordance ont été mis à disposition sur le blogue yyy (Schaeffer-Lacroix, nd) comme l'illustre l'extrait d'article de blogue suivant.

Wie schreibt man Annoncen, um Brieffreunde zu finden?

Hier ist eine Internetseite zu diesem Thema:

<http://www.brieffreunde.org/>

Hier ist ein Textkorporus: [Korpus Brieffreunde.rtf](#)

Hier ist die Konkordanz: [Konkordanz Brieffreunde.rtf](#)

[Comment rédiger des annonces pour trouver un correspondant?

Voici une page Internet à ce sujet :
<http://www.brieffreunde.org/>

Voici un corpus de textes : [Korpus Brieffreunde.rtf](#)

Voici la concordance : [Konkordanz Brieffreunde.rtf](#)]

Toutes les versions des productions des apprentis-scripteurs, ainsi que mes propositions de révision, ont été publiées sur ce même blogue.

4.3. Sections d'une annonce

Avant la phase de rédaction, le groupe a observé les sections d'une annonce dans le corpus *Brieffreunde.org*. Ce découpage s'inspire des travaux de Bronckart (1997 : 220) sur les séquences d'un texte.

Sections d'une annonce	Exemples dans le corpus <i>Brieffreunde.org</i>
Dire bonjour	<i>Hallo</i> <i>Hallöchen</i> <i>Hej</i> <i>hallo liebe aktive briefschreiber</i>
Se présenter	<i>Ich heiße</i> <i>Meine Hobbys sind</i>
Dire qui l'on cherche	<i>aktive briefschreiber</i> <i>ich suche nette Leute</i>
Moyen de communication	<i>Ich suche selbstverständlich nette Briefkontakte.</i> <i>Ich suche auch Mailfreundschaften.</i>
Inciter à répondre	<i>Bitte meldet euch.</i> <i>Also einfach drauf los schreiben und ab die Post.</i>
Dire au revoir	<i>Hoffentlich bis ganz bald, euer Daniel</i> <i>Gruß Thomas</i> <i>kussi Kimi</i>

Tableau 3 – Sections d'une annonce

4.4. Rédaction et révision

Suite à l'observation des caractéristiques linguistiques et discursives des annonces, les apprenants ont rédigé le premier jet de leur texte. Ils l'ont ensuite publié sur le blogue, sous forme de commentaire posté à la suite de l'article "Meine Annonce" [mon annonce].

Meine Annonce

Hier könnt ihr eure Annonce als Kommentar hinterlassen.

Publié dans Wanda: Brieffreunde.org | Lien permanent | Commentaires (157)

[Mon annonce

Vous pouvez déposer votre annonce ci-dessous, sous forme de commentaire.]

J'ai répondu à chacun individuellement, également sous forme de commentaire publié sur le blogue. Pour ce faire, j'ai copié les parties dont j'ai estimé qu'elles nécessitaient une révision, et je les ai mis entre guillemets. Parfois, j'ai donné des indications supplémentaires susceptibles de guider les apprentis-scripteurs lors de la phase de révision. J'ai, par exemple, conseillé d'observer le contexte d'un mot clé en particulier, ou j'ai renvoyé les apprenants à des explications grammaticales ou lexicales publiées sur le blogue ou ailleurs. Parfois, j'ai proposé le recours à des outils supplémentaires, comme des dictionnaires en ligne. J'ai également répondu aux questions posées dans les commentaires publiés en ligne. En réaction aux caractéristiques des premiers jets, j'ai publié des informations de nature linguistique ou discursive sur le blogue. Après un va-et-vient pouvant aller jusqu'à trois échanges, j'ai clos le projet, en considérant la dernière version publiée comme définitive. L'échange suivant entre Rose¹ et moi exemplifie cette démarche de révision guidée.

¹ Les prénoms sont modifiés pour les besoins de cet article.

HallÖchen !

Ich HEIBE Rose und ich bin 14. Ich reite gern. Meine Hobbys sind Musik und sport Ich liebe es,Guitare zu spielen. Ich wohne in st-leu. Ich suche eine nette Brieffreund. Ich schreibe gern Mails. Ich freu mich auf jede Antwort.

Kussi, Rose

Écrit par : Rose V1

Rose (V1),

korrigiere noch diese Wörter:

„Guitare“ (mot clé „spielen“ dans la concordance)

„eine nette Brieffreund.“ (mot clé „nette“ dans la concordance).

Korrigiere "HEIßE" ("h...").

Frau Lacroix

Écrit par : Frau Lacroix an Rose (V1)

Hallöchen !

Ich heiße Rose und ich bin 14. Ich reite gern. Meine Hobbys sind Musik und sport. Ich liebe es,Gitarre zu spielen .Ich wohne in st-leu. Ich suche einen netten brieffreund. Ich schreibe gern Mails. Ich freu mich auf jede Antwort.

Kussi, Rose ^^

Écrit par : Rose V2

Rose, Virgile und Constantin (V2),

eure Annonce ist gut geworden. Ihr seid jetzt fertig.

Bis Morgen.

Frau Lacroix

Écrit par : Frau Lacroix an Vincent, Rose und Corentin (V2)

Dans les dernières lignes de l'exemple ci-dessus, je complimente trois des apprentis-scripteurs, dont Rose, pour le travail accompli, et je leur signale clairement la fin de la tâche : "votre annonce est bien comme ça. Vous avez terminé maintenant".

4.5. Aspects linguistiques et discursifs de la révision

Les segments repérés dans les lignes de concordance ont permis aux apprentis-scripteurs de réviser des aspects linguistiques et discursifs des premiers jets. Dans le domaine des aspects linguistiques qui ont été abordés par les apprenants, il est possible de citer les terminaisons et la place des adjectifs et les mots composés, ainsi que des collocations, donc des expressions polylexicales semi-figées ou figées, comme "aus aller Welt" [du monde entier]. Cette expression se trouve, entre autres, dans la phrase du corpus "Ich freue mich auf Briefe aus aller Welt" [Des lettres du monde entier me feraient plaisir]. Au niveau discursif, le corpus se caractérise par des signes typographiques et par la multiplication de lettres ou de signes de ponctuation : "Biiiiittttte meldet euch :)" [S'iiiiil vous plaît, faites-moi signe :)]. Certaines des productions des apprenants comportent ce genre de caractéristiques, comme l'illustre le paragraphe suivant.

4.6. Écriture intertextuelle

Selon Bronckart (2000), tout texte s'inscrit dans des textes du même genre qui sont "déjà là". Cette vision des choses a motivé mon choix de comparer les productions des apprenants au corpus afin de déterminer la co-présence de mots dans les deux types de textes. Cette comparaison renseigne sur les aspects intertextuels de l'écriture des apprentis-scripteurs.

Les textes de Rose (4.4.) montrent des traces de récupération de formules trouvées dans les lignes de concordance : le mot clé "HEIßE" [m'appelle], écrit en majuscules dans la concordance, reste en majuscules dans la version V1, et le segment "Ich freu mich auf jede Antwort." [Toutes les réponses me feront plaisir.], ainsi que la formule finale "Kussi" [bisou] co-existent dans le texte de Rose et dans le corpus. Je constate que dans la deuxième version publiée (V2), le prénom

en fin de texte est suivi par l'émoicône "^^" qui veut dire "sourire avec les yeux fermés".² De plus, Rose remplace "Brieffreund" par "brieffreund" dans la version finale. Je considère ces deux actions comme un signe d'appropriation du discours par Rose. Le genre textuel "commentaire dans un livre d'or électronique" se caractérise, en effet, par la présence d'émoicônes, ainsi que par certaines particularités orthographiques, comme l'abandon occasionnel ou systématique des majuscules des noms.

L'exemple de Cathy ci-dessous présente en détail un processus de révision. Les parties co-présentes sont marquées en gras. Les parties supprimées par rapport à la version précédente sont barrées.

(V0) **Hej** Mein Name ist Cathy! **ich suche auf diesem** interessante **Briefschreiber**. ich bin 15 jahre und ich will eine Brieffreunde, die 15 jahre ist. ich mache volleyball und rocktanze.

Tchüss Cathy

(Nombre de mots co-présents : 7/31)

(V2) **Hej** Mein name ist Cathy!=p
ich wohne zwar nicht direkt in paris.ich suche auf diesem interessante Briefschreiber eine nette brieffreundin. ich bin 15 jahre alt und ich will eine ~~Brieffreunde~~ Brieffreudin, die ~~15 jahre~~ von 15 jahre von 14 ist. ich mache volley ball und ~~rocktanze~~ tanze rock.Ich mag musik horen. **Schreibt mir!**

~~Tchüss~~ **kussi** ,Cathy:)

(Insertions : 6. Suppressions : 5. Nombre de mots co-présents : 16/46)

² Selon la définition dans l'article "Émoicône" de Wikipedia. <http://fr.wikipedia.org/wiki/%C3%89motic%C3%B4ne>

(V3) **Hej** Mein name ist Cathy!=p
ich wohne zwar nicht direkt in paris.ich suche eine
nette brieffreundin. ich bin 15 jahre alt und ich ~~will ei-~~
~~ne Brieffreundin, die von 15~~ **suche Brieffreunde im al-**
ter von 14 ist. -15.Ich mache volley ball und tanze rock.
 Ich mag musik ~~hören.Schreibt~~ hÖren. Schreib mir!
kussi ,Cathy:)

(Insertions : 3. Suppressions : 4. Nombre de mots co-
 présents : 23/46)

Cathy a deux fois révisé son texte. La première version (nommée "V0") comporte 7 mots coprésents sur un total de 31 mots. Il s'y trouve une collocation qui paraît tronquée : "auf diesem" [par ce] au lieu de "auf diesem Wege" [par cette voie]. Dans la version V2, j'ai identifié cinq passages qui ressemblent à des éléments présents dans le corpus. Trois des passages ont un caractère collocatif. Le premier passage nécessite la suppression de "zwar", un adverbe à caractère anaphorique qui n'a pas de référent dans le texte de Cathy, mais qui en a un dans le passage équivalent du corpus. Au fil des révisions, tout comme Rose dans l'exemple présenté en 4.4., Cathy semble s'approprier le discours représenté dans le corpus : le nombre éléments coprésents augmente pour atteindre la proportion de 50 %.³ L'emploi de la formule finale "kussi" [bisou], d'émoicônes et la modification de "Schreibt mir!" [Écrivez-moi !] en "Schreib mir!" [Écris-moi!] contribuent à ce processus.

5. Pour finir

Le déroulement positif de *Brieffreunde.org* et les productions obtenues permettent d'avancer qu'il est pertinent de proposer un travail basé sur des lignes de concordance dans le milieu de l'enseignement

³ Je considère cette proportion comme excessive. Mon étude globale des productions suggère que la co-présence de 30 % de mots représente un bon rapport entre corpus et propre texte.

secondaire. J'ai toutefois trouvé le guidage chronophage, et de plus, j'ai observé dans l'ensemble des productions une certaine stéréotypie au niveau des collocations. Afin de remédier à cela, il paraît prometteur de proposer le même type de projet avec un corpus plus grand et avec un outil d'exploration plus performant, permettant aux apprenants de faire des requêtes plus variées, en étant moins guidés. Les outils proposés sur le site *Sketch Engine* offrent des possibilités intéressantes à ce sujet : il est possible de télécharger son propre corpus et d'y (faire) appliquer de nombreux types de requêtes à l'aide d'un concordancier facile d'emploi.

6. Références bibliographiques

Tous les liens étaient actifs en novembre 2010.

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

7. Logiciels et sites Internet

Tous les liens étaient actifs en novembre 2010.

Cobb, T. (nd). *Text-Concordancer (v2)*.

<http://www.lex tutor.ca/concordance>

Ingelmann, C. (nd). *Brief freunde.org*. <http://www.brief freunde.org/>

Kilgarriff, A., Rychly, P. & Pomikalek, J. (nd). *Sketch Engine*.

<http://www.sketchengine.co.uk/>

Schaeffer-Lacroix, E. (nd). *Projekt Prospekt*. Identifiant : corpus. Mot de passe : wanda. <http://corpus.hautetfort.com/>