Les corpus dans l'enseignement du FLE : exemplification avec le cas des quasi-synonymes

Avlin Pamuksac

Séminaire collectif : Linguistique de corpus: méthodes et analyses Coordinatrice : Corinne Rossari Semestre de printemps 2019

Plan

- Définitions
- Historique de l'utilisation des corpus dans l'enseignement/apprentissage des langues
- Pourquoi utiliser des corpus ?
- Comment utiliser les corpus ?
- Avantages et inconvénients des corpus dans l'enseignement/apprentissage des langues
- Etude de cas : les quasi-synonymes

Définitions

- « Un corpus est une collection de textes (oraux ou écrits) réunis selon un certain nombre de principes et permettant d'observer des faits de langue. » Tyne 2013 : 7
- Data driven-learning: « Selon cette approche, qui prend les données comme point de départ pour l'apprentissage, les apprenants sont amenés à découvrir et à cerner la langue, ainsi que les différents usages, à travers le contact avec les données généralement réunies sous forme de corpus (Chambers 2010). » Tyne 2013: 8

Historique de l'utilisation des corpus dans l'enseignement/apprentissage des langues

« Les corpus nous renseignent aussi sur le fonctionnement de la langue comme système, sur les tendances, les fréquences, etc.: c'est ainsi que des dictionnaires (à commencer par le Dictionnaire fondamental—coir Gougenneim 1958: voir aussi le projet plus récent de constitution du dictionnaire COBUIL pour l'anglais—voir Sinclair 1987), des grammaires (four l'anglais—par ex. Biber et al. 1999), et plus récemment des manuels (pour l'anglais—par ex. McCarthy et al. 2005) ou des guides pour enseignants (pour l'anglais—par ex. O'Keefe et al. 2007; pour le français un ouvrage assez isoir—iraitant davantage de la présentation des données-ressources que du travail sur corpus—Detey et al. 2010), ont pris le parti de présenter la langue telle qu'elle se manifeste à travers les données attestées, en mettant en avant des aspects tels que la fréquence d'utilisation, mais aussi les liens qu'entretiennent les mots avec leur 'voisinage' ou contexte d'utilisation (collocations), la distribution de la grammaire en fonction du registre ou de la situation d'utilisation, etc. » Tyne 2013 : 7-8

Historique de l'utilisation des corpus dans l'enseignement/apprentissage des langues

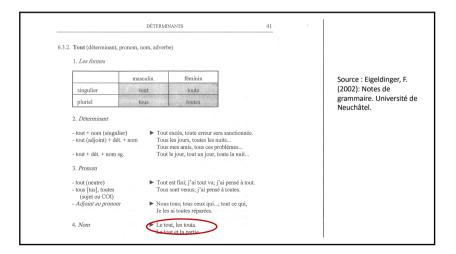
• En anglais: bien avancé sur les corpus: « La langue cible (L2) était l'anglais dans 33 des 39 études, une langue européenne dans toutes les autres. Ceci s'explique en grande partie par la demande pour l'anglais L2 à l'échelle mondiale, ainsi que par la gamme d'outils, de logiciels et de corpus disponibles pour ces langues; le fait que les publications sont souvent en anglais contribue certainement à une plus grande prise de conscience pour cette langue. Quelles que soient les raisons, il semble que le DDL soit une approche qui s'inspire largement du monde anglo-saxon. » Boulton 2008: 40

Historique de l'utilisation des corpus dans l'enseignement/apprentissage des langues

• En français, pas encore beaucoup utilisé car importance accordée au bon usage : « En attendant, le plus grand corpus actuellement disponible au public est Frantext avec ses quelques 200 millions de mots. La composition même de ce corpus (principalement des textes littéraires datant du 16e au 20e siècle) est révélatrice de l'importance accordée au « bon usage » de la langue française. En effet, le manque relatif d'intérêt pour la linguistique de corpus en France peut être attribué à une méfiance envers une approche purement descriptive de l'emploi courant de la langue. » Boulton 2008 : 37

Historique de l'utilisation des corpus dans l'enseignement/apprentissage des langues

 « [...] les corpus doivent encore faire leurs preuves dans les trois usages de base décrits par Fligelstone (1993), à savoir « teaching about » (enseigner la linguistique de corpus), « exploiting to teach » (exploiter les corpus pour enseigner les langues) et « teaching to exploit » (former les apprenants à exploiter les corpus pour apprendre une langue). » Boulton 2008 : 37



- L'industrie du disque allait s'emparer de touts les styles musicaux et les reproduire à l'échelle industrielle. (www.cmaq.net, rampé 21.10
- . J'espère vraiment pouvoir venir un jour pour savourer touts ces paysages magnifiques et partager la joie de vivre des Mauriciens. (ilem
- Toutes les données, touts les programmes installés depuis que vous l'avez seront perdus. (loiseljp.perso.sfr.fr, rampé 30.11.0002)
- Rien que une bonne coordination entre touts les travailleurs activé, le démênagement peut s'effectuer rapidement avec compétence. (v
- L'atelier est ouvert, sur rendez-vous, à toutes visites et touts projets. (www.delanoestudioglas.ch, rampé 30.11.0002)
- Nous aimons souligner les fêtes importantes de l'année avec les touts petits. (www.poussonspoussettes.com, rampé 30.11.0002)
- Il faut avoir beaucoup volonté et de courage pour endurer touts les efforts physiques et psychologiques. (www.irida.fr, rampé 30.11.000
- Entre-temps est éclaté sûr ce sondage une campagne électoral entre touts les partis. (www.gerardo-mada.de, rampé 30.11.0002)
 Le pilote britannique veut mettre touts les chances de son côté pour signer un nouveau succès dans le cadre de l'Intercontinental Rally
- Comme hébergeur nous voulons mettre à votre disponibilité touts les outils dont vous avez besoin. (www.vevida.ni, rampé 30.11.0002)

Source: Corpus français (http://corpora.uni-leipzig.de/fr?corpusId=fra_mixed_2012), Consulté le 20.04.19

5. Adverbe = tout à fait, entièrement (invariable).

tout + adjectiftout + adverbe

- tout + préposition

Ils n

► Ils sont tout tristes. Ils sont tout heureux.

Ils marchent tout lentement.

Leur maison est tout à côté de la poste.

Attention: Si tout, adverbe, est placé devant un adjectif féminin commençant par une consonne ou un «h» aspiré, il s'accorde pour des raisons d'euphonie:

► Elle est toute triste, toute honteuse. Elles sont toutes tristes, toutes honteuses.

Mais: Elle est tout inquiète.

Elles sont tout aimables, tout heureuses.

Source : Eigeldinger, F. (2002): Notes de grammaire. Université de Neuchâtel.

▲ Voisins de gauche significatifs:

avant (345,541), , (148,177), malgré (143,075), En (139,772), que (101,247), Malgré (78,623),

Source: Corpus français (http://corpora.uni-leipzig.de/fr?corpusId=fra_mixed_2012), Consulté le 20.04.19

Pour apporter des réponses...

3) Placez les adjectifs avant ou après le nom. Expliquez vos choix.

Beau/laid sale/propre héroïque/mauvais

Avec son visage... Avec son vêtement... Avec son geste...

Avec son ... visage. Avec son ... vêtement. Avec son ... geste.





Exemple avec « sale » et « propre »

	Sens 1	Sens 2
Propre	Qui est bien lavé, bien tenu.	Qui appartient exclusivement ou en particulier à (une personne ou une chose).
Sale	pureté [] et donne à la	Qui est digne de mépris ; dont la conduite, les actes sont condamnables.

Source: Trésor de la langue française (http://atilf.atilf.fr/)

Méthodologie du data-driven learning

- Méthode traditionnelle : « rule first, practice second. » Selinger in Johns 1991 : 3
- « [...] il est avant tout synonyme d'autonomie, et l'apprenant est un chercheur, voire même un "détective" (Johns 1997) ou un enquêteur enregistrant et transcrivant les données (Tyne 2009); pour les autres, dans une forme plus "soft", il s'agit plutôt de profiter (directement ou indirectement) de ce qu'offre l'approche sur données, tout en maintenant un cadre plus ou moins 'classique' [...] » Tyne 2013 : 9

Concrètement...

- Deux possibilités :
- 1. Par l'enseignant
- 2. Par l'élève

Par l'enseignant

- a. Lisez les phrases. Soulignez le verbe décider, puis entourez les sujets et les compléments de chaque phrase.
 - 1. Le président a décidé de supprimer le budget.
 - 2. Elle a décidé de faire de sa page personnelle un site officiel.
- 3. Un juge a décidé de prolonger leur garde à vue.
- 4. Il décide de partir à l'étranger.
- 5. Elles décident d'étudier la médecine.
- 6. J'ai décidé d'accepter le poste de ministre d'Etat.

b. Tous les suiets sont ...

- des êtres humains □ - des choses, des objets □

c. Choisissez la bonne réponse

- J'ai décidé à ne pas manger.

- J'ai décidé de ne pas manger.

d. Écrivez une phrase avec le verbe décider.

e. Trouvez 3 phrases avec le verbe décider sur Google.

Source: Noghrechi & Kamber 2018: 201-202

Par l'élève

- « Car le propre du 'data-driven learning', dans son état le plus 'pur', est de faire de l'apprenant un chercheur, capable de s'informer sur le fonctionnement de la langue cible à partir des données brutes. » Tyne 2013:10
- Exemple : Landure 2011 : Dictionnaire bilingue vs consultation des corpus

Avantages et inconvénients de l'utilisation des corpus en classe de langue

- Avantages
 « Il se peut en effet que l'intérêt principal des corpus en L2 soit de compléter les outils de référence traditionnels : dictionnaires, grammaires, etc. » Boulton 2008 : 41-42
 - « [...] numerous researchers who claim that engaging in DDL activities has the potential to bring general cognitive benefits which are vital to language learning such as the development of analytical skills, the ability to notice linguistic features, and problem-solving abilities [...] » Kerr 2013: 19
- Inconvénients
- « le DDL nécessite une remise en question des rôles traditionnels : les enseignants peuvent craindre une perte de contrôle, de pouvoir, de statut d'expert; les apprenants peuvent refuser la responsabilité du fils auront pour leur propre apprentissage. » Boulton 2008 : 38
 - « La quasi-totalité des recherches à ce jour vise des apprenants de niveau relativement avancé, en partie sans doute parce que les chercheurs dans des milieux, universitaires y ont plus facilement acces. S'il est d'une dévidence pour certains que le DDL se prête meux aux inveaux avances, les que les partes de la company de la compa

Etude: Les quasi-synonymes

- « One of the commonest types of question asked by the enquiring learner is "What is the difference between ...?" » Johns 1991 : 4
- « The grammar / vocabulary dichotomy is invalid; much language consists of multi-word 'chunks'. » Kerr 2013: 22
- « [...] nous rappellerons une idée majeure de Wittgenstein selon laquelle l'étude de l'usage des mots, réalité manifeste, observable et au besoin quantifiable, est susceptible de nous renseigner sur leur signification – leur nature profonde en quelque sorte. » Blumenthal 2008 : 37

Etude: Les quasi-synonymes

- Dans quel contexte on utilise un verbe plutôt qu'un autre ?
- → Savoir
 - Connaître, être informé

Procédure

- Utiliser Corpus français (http://corpora.uni-leipzig.de/fr?corpusId=fra_mixed_2012) → problème : pas lemmatisé
- Voir les formes les plus fréquentes sur ScienQuest (https://corpora.aiakide.net/scientext20/)

Traits morpho-syntaxiques

Token ု	Fréq. absolue	% des occurrences	‱ des mots du corpus
VINE	992	45.59%	1.77260 ‱
VCONJS	769	35.34%	1.37412 ‱
VCONJP	202	9.28%	0.36095 ‱
Ppr	153	7.03%	0.27340 ‱
<u>PpaMS</u>	57	2.62%	0.10185 ‱
<u>PpaFS</u>	2	0.09%	0.00357 ‱
PpaFP	1	0.05%	0.00179 ‱

Résultats pour « savoir »

▲ Cooccurrences significatives:

•

si (131,746), falt (71,858), faut (69,549), question (62,243), Reste (61,599), comment (45,515), plus (35,972), que (31,090), s'ii (27,075), ce (26,413), Pour (25,555), qu'ii (24,331), faire (23,743), Rédigez (19,808), quel (17,899), quelle (17,057), à (15,953), il (13,829), ou (13,187), d'en (12,241), pour (11,775), pas (11,602), combien (10,885), croit (10,733), pourquoi (10,695), vois (10,208), visitez (10,136), quelles (9,672), le (9,504), sans (9,152), quels (9,337), sur (8,477), réagir (8,463), ne (6,422), quoi (8,163), nous (7,599), . (7,481), ii (7,148), consultez (6,694), En (6,448), cliquez (6,415), répondu (6,342), qui (5,897), tout (5,806), voulez (6,773), commentaire (5,773), qu'elle (5,744), interrogé (5,494), pose (5,369), a (5,306), va (5,149), J'aimerais (5,091), s'ils (4,994), conduite (4,943), attendre (4,834), " (4,803), ... (4,654), rèples (4,603), veut (4,602), je (4,578)

▲ Voisins de gauche significatifs: fait (238,847), à (155,890), faut (142,903), de (68,873), En (55,566), pour (46,475), en (43,362), d'en (32,454), croit (25,363), sans (15,565), faire (13,804), ∧ (11,902), voudrais (8,374), voulu (8,207), tout (8,155), doit (7,644), 'alimerais (6,803), doivent (6,493), devez (5,734), le (5,717), voulez (4,755), 'alimerais (4,639), veut (4,424), laissé (4,389), voulais (4,201), ren (4,119), faut-il (3,927), Pour (3,717), veux (3,398), pas (3,257), Tout (3,189), veulent (3,170), allez (2,665), du (2,429), voulant (2,086), crois (1,976), croient (1,916), l'alimerai (1,892), J'alimerai (1,882), tera (1,823), Comment (1,779), voulons (1,602), à (1,594), son (1,568), notre (1,503), fall-il (1,500), faudrait (1,459), Allez (1,417), aussi (1,283), croit-on (1,279), leur (1,243), souhalterais (1,219), iont (1,217), Faut (1,213), den (1,207), trop (1,203), blen (1,200), devons (1,166), voudrait (1,111), voulaient (1,105)







Résultats pour « sait »

Voisins de droite significatifs:

que (144,459), pas (138,255), qu'il (35,291), jamais (18,157), comment (16,826), rien (13,504), bien (12,883), aussi (12,296), ce (11,263), déjà (10,841), toujours (8,878), pertinemment (8,783), très (8,735), combien (8,222), qu'elle (7,825), plus (7,332), maintenant (7,318), où (6,645), qu'on (6,404), faire (6,223), si (6,123), parfaitement (4,992), exactement (4,404), désormais (4,267), aujourd'hui (3,565), quelque (3,191), depuis (3,084), quelle (3,049), pourquoi (3,009), trop (2,795), qu'ils (2,584), qu'un (2,478), qu'une (2,383), qu'en (2,301), tout (2,286), quel (2,227), quil (2,144), toutefois (2,105), vraiment (2,044), d'où (2,005), cependant (1,773), lire (1,655), mieux (1,641), quoi (1,606), pourtant (1,558), seulement (1,536), jouer (1,485), encore (1,470), quand (1,463), également (1,406), qu'il (1,384), s'il (1,346), peu (1,300), tous (1,295), s'adapter (1,288), d'ores (1,226), même (1,206), donc (1,164), se (1,150), qu'à (1,139)

Conclusion pour « savoir – sait »

- Construction type:
 - sujet humain VERBE (adverbe de précision du savoir : bien, pertinemment ou du temps du savoir : jamais, aussi, déjà, toujours), subordonnée (introduite par que, ce que, si, où...) ou interrogation indirecte (introduite par comment, poruquoi...)
- Phrases tirées du corpus :
 - L'étudiante sait sans doute que Steve Jobs a pour habitude de répondre parfois personnellement aux messages qu'on lui envoie.
 - On ne sait toujours pas s'il existe une infinité de nombres unitaires parfaits pairs.
 - Claude Bergeaud sait qu'il peut compter sur son guerrier préféré pour guider sa

Problème

- Construction sait + *infinitif* ne ressort pas dans les collocations
- · Que ce soit les scènes de bataille rythmées, des images de poésie pure, des seconds rôles fantaisistes, ou les séquences d'émotion, Miyazz (i sait mettre en scène son immense talent de conteur. (www.ecrannoir.fr.txt, rampé 22.10.2007) Peu d'informations toutefois, on sait que cet e-book offrirait un format d'écran un peu moins.
- grand que le Beader de Sony. (www.lesnumeriques.com. rampé 24.03.2009)
- On sait le discours stéréotypé qui sort toujours du complexe français. (www.marianne2.fr, rampé
- Isait encaisser (www2.canoe.com, rampé 06.05.2006)

30 11 0002)

- Et manifestement, elle sait de quoi elle parle puisque diplômée de psychologie clinique et de psychopathologie, elle anime des ateliers d'écriture visant la décharge, par l'écrit, du trop-plein de souffrance. (www.jose-corti.fr, rampé 30.11.0002)
- Le cimentier calt aussi adopter une attitude sociale. (www.letemps.ch, rampé 10.11.2005)
- Chaque employe sait ce qu'il devra faire pendant la journée, chaque armoire est clairement étiquetée et les bureaux n'ont pas de téléphone, (www.largeur.com, rampé 18.06.2010)
- L'acheteur ne sait pas à l'avance quelles figurines contient le paquet. (fr.wikipedia.org, rampé 04.01.2010)
- On sait aujourd'hui qu'ils passaient l'hiver autour des baies, se nourrissant de mollusques, de
- crustacés et de poissons. (Karankawa, rampé 22,10,2007) Räikkönen ne sait nas vraiment à quel résultat sattendre ce week-end (www.f1-action.net rampé
- Harmonie des goûts et couleurs, Claude sait y faire. () www.standdegilamont.ch, rampé
- Pas de doute, Care sait choisir les vrais bons produits et les accommoder avec justesse. (www.tbx.be, rampé 18.10:2011)

Résultats pour « connaître »

▲ Cooccurrences significatives:

faire (60,950), mieux (44,056), les (11,609), pour (11,081), fait (10,461), Pour (8,804), de (8,058), vous (7,715), afin (6,816), permet (5,770), et (5,170), apprendre (4,804), nous (4,790), se (4,357), votre (4,276), devrait (4,261), public (3,789), attendre (3,686), leur (3,579), comprendre (3,314), site (3,190), faut (3,167), leurs (2,966), cliquez (2,965), il (2,820), vérité (2,753), ses (2,694), qui (2,634), succès (2,579), permettra (2,550), consulter (2,537), consultez (2,486), notre (2,432), le (2,374), va (2,363), son (2,339), à (2,281), vos (2,268), d'en (2,220), Mieux (2,168), fera (2,164), auprès (2,131), exacte (2,117), savoir (2,044), bien (2,042), avant (2,009), faudra (1,974), nos (1,964), détails (1,869), fit (1,808), permis (1,799), sans (1,776), intentions (1,770), permettent (1,707), appris (1,702), permettre (1,696), d'apprendre (1,640), sa (1,615), sort (1,613), Reste

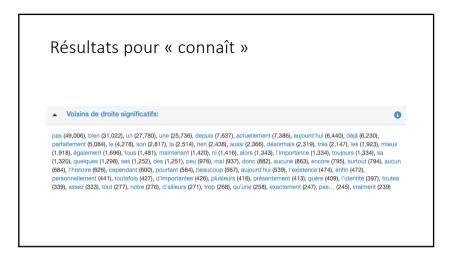










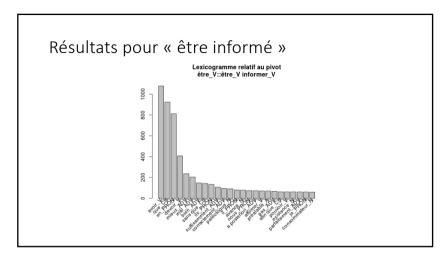


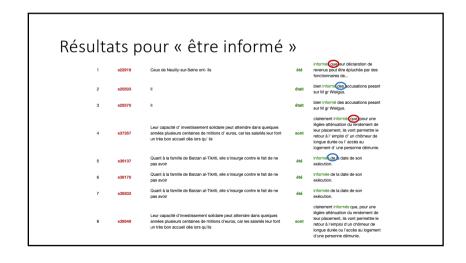
Conclusions pour « connaître – connaît »

- Construction type
 - Sujet collectif géographique (pays, ville, région...) VERBE adverbe de temps (depuis, actuellement, aujourd'hui, déjà...), groupe nominal (catégorie sémantique « abstraite » : succès, difficulté, essor, croissance...)
- Phrases tirées du corpus
 - Mais, à cette époque, Collioure connaît un tournant essentiel, qui contribuera grandement à son développement touristique.
 - Nous avons reprogrammé certaines destinations mais le reste de la région connaît une croissance très saine.
 - Le marché marocain **connaît** cependant une saturation avec une position monopolistique de cinq acteurs majeurs qui le dominent.

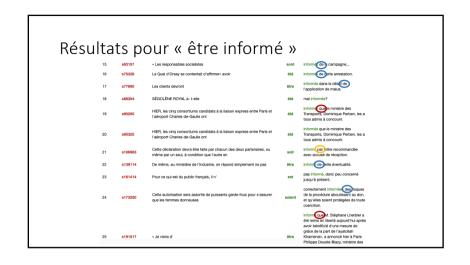


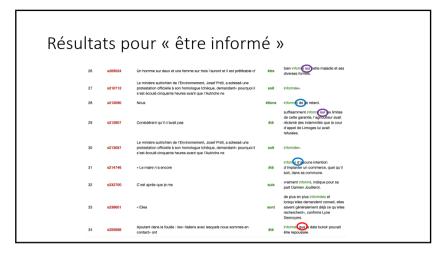


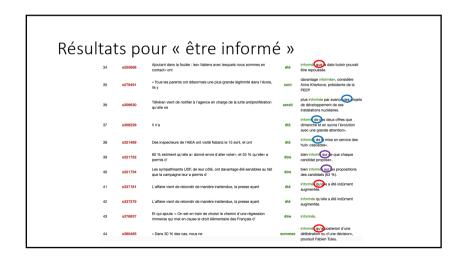












Conclusions pour « être informé »

- Construction type
 - Sujet humain VERBE (adverbe du degré d'information inséré entre « être » et « informé » : bien, clairement...) préposition (de, par, sur) avec des compléments plutôt abstraits (accusation, déclaration, limite...) ou subordonnée (introduite par que)
- Phrases crées sur ce modèle :
 - Les étudiants seront informés des résultats par e-mail.
 - Le président a été informé que la deuxième tour a été touchée par un avion.

Récapitulatif

Savoir

- Sujet humain
 Adverbe de la
- précision du savoir (bien, pertinemment) ou du temps du savoir (jamais, aussi, déjà)
- Complément subordonnée (introduite par que, ce que...) ou interrogation indirecte (introduite par comment, si...)

Connaître

- Sujet collectif (géographique)
- Adverbe de temps (depuis, actuellement, déjà...)
- Complément groupe nominal (catégorie sémantique abstraite: succès, difficulté, essor)

être informé

- Sujet humain
- Possibilité d'ajouter un adverbe concernant le degré de l'information
- Complément introduit par une préposition (le plus souvent de) avec un complément de type abstrait ou une subordonnée (introduite par que)

Comment l'enseigner?

- Souligner les mots qui ont le même sens ou un sens proche
- Avec des exercices d'observation (à l'aide du contexte mieux comprendre le sens du verbe)
- Comment se construit ce verbe ?
- Introduisez le bon verbe (parmi une liste ou non)

Conclusion

- Corpus peuvent être utiles pour l'enseignement d'une langue à condition de bien savoir les utiliser (authenticité, nombre de données...)
- Certains corpus n'offrent pas les outils nécessaires, il faut savoir les combiner avec d'autres
- Pour être encore plus précise, pouvoir télécharger les données pour faire d'autres calculs

Bibliographie

- Blumenthal, P. (2008). Combinatoire des prépositions : approche quantitative. Langue française, 157, 37-51.
- Boulton, A. (2008). Esprit de corpus : promouvoir l'exploitation de corpus en apprentissage des langues. Texte et corpus, 3, 37-46.
- Flament-Boistrancourt, D. (2004). Quelle contribution un corpus natif/non natifs est-il susceptible d'apporter à la linguistique et à l'enseignement du français L2: le cas de LANCOM? Pré-textes franco-danois, 4, 45-69.
- Johns, T. (1991). Should you be persuaded two samples of data-driven learning materials. ELR Journal, 4, 1-16.
- Kerr, B. (2013). Grammatical description and classroom application. Theory and practice in data-driven learning. Bulletin suisse de linguistique appliquée, 97, 17-39.
- Landure, C. (2011). Data-driven learning: apprendre et enseigner à contre-courant. Mélanges CRAPEL, 32, 163-178.
- Leray, M. & Tyne, H. (2016). Homophonie et maîtrise du français écrit: apport de l'apprentissage sur corpus. Linguistik Online, 78 (4/16), 131-148.
- Noghrechi, H. & Kamber, A. (2018). Réviser les manuels pédagogiques et les référentiels à la lumière des corpus écrits: le cas des locutions à verbe support pour les apprenants de français langue étrangère des niveaux A1 et A2. Synergies France, 12, 189-205.
- Tyne, H. (2013). Corpus et apprentissage-enseignement des langues. Bulletin suisse de linguistique appliquée, 97, 7-15.