

50 ANS DE MÉLANGES À DÉMÊLER

Emmanuelle Carette

ATILF-CNRS & Université de Lorraine

Mots-clés

Revue en didactique des langues – évolution thématique – analyse lexicale – cooccurrences lexicales - univers sémantiques

Keywords

Language Teaching and Learning Journal - thematic evolution - lexical analysis - semantic universes

Résumé

L'objectif de cette étude est de rendre compte de cinquante années de publications dans la revue *Mélanges Crapel*, et de décrire les évolutions thématiques. Atteindre cet objectif et garantir une certaine objectivité a nécessité le recours à des outils informatiques permettant de dégager des univers sémantiques, des fréquences de mots lexicaux et de donner une représentation des structures thématiques des publications, sur cinq périodes de dix ans, et sur la totalité de l'échantillon constitué pour l'étude. La structure constante des thèmes sur cinquante ans et les spécificités de chaque décennie sont présentées, ainsi que quelques réflexions personnelles en conclusion.

Abstract

The aim of this study is to report on fifty years of publications in the journal *Mélanges Crapel*, and to describe the thematic developments. In order to achieve this objective and to guarantee a certain objectivity, computer tools were used to identify semantic universes, lexical word frequencies and to give a representation of the thematic structures of the publications, over five ten-year periods, and over the entire sample constituted for the study. The constant structure of the themes over fifty years and the specificities of each decade are presented, as well as some personal considerations in conclusion.

Introduction

69, année didactique : le CRAPEL, Centre de Recherches et d'Applications Pédagogiques en Langues, est né, et en 1970, le premier numéro des *Mélanges*, alors *Pédagogiques*. En 2021, sollicitée pour fêter le cinquantenaire des *Mélanges*, je me suis donné l'objectif de porter un regard sur l'ensemble de la revue, qui décrit cette longue période de travaux en didactique des langues. J'ai cherché à exposer de manière à la fois représentative et le plus simplement possible une évolution des contenus publiés et en extraire une connaissance plus générale intéressant la didactique des langues.

La revue a nécessairement évolué sous les contraintes environnementales, depuis les idéologies de liberté, le développement des domaines de la didactique des langues, du français langue étrangère, jusqu'à la politique d'évaluation des enseignants-chercheurs, en passant par les critères de valeur des publications scientifiques.

Une évolution contextuelle est celle du mode d'évaluation des enseignants-chercheurs, qui dévalorise la publication dans une revue de l'équipe ou du laboratoire d'appartenance de l'auteur. Ce changement de modalité de fonctionnement n'est pas étranger à l'évolution des contenus de la revue, qui ne reflète plus vraiment les travaux de l'équipe.

Conséquence de cette évolution institutionnelle, un aspect probablement influent est l'évolution de la politique de publication. Jusqu'au numéro 27 en 2004, les articles étaient majoritairement ceux des membres du CRAPEL (alors équipe d'accueil) et de ses doctorants. Des contributions de personnes extérieures pouvaient être sollicitées par les membres de l'équipe, mais elles étaient minoritaires. Les numéros ne faisaient pas l'objet d'appels à contribution largement diffusés. Ce fonctionnement a connu une rupture complète avec le numéro 28 en 2006. L'évolution du contexte de la recherche et de la publication scientifiques universitaires a conduit l'équipe à réorganiser la revue, de manière à ce qu'elle corresponde aux critères de « qualité » désormais incontournables. S'en est suivi la création d'un comité éditorial, d'un comité de rédaction, d'un comité scientifique mêlant membres internes et externes, l'explicitation de la sélection par double évaluation anonyme sur

critères établis, ainsi que l'accueil d'articles provenant d'enseignants et/ou de chercheurs du monde entier.

D'autres sources d'évolution sont sans doute à trouver, et il s'agit donc de savoir comment se structurent les publications des années écoulées. La satisfaction de cet objectif m'a menée à des options méthodologiques : comment détecter une évolution du contenu qui ne constitue pas une œuvre artistique mais satisfasse au mieux quelques repères académiques telle que la transparence méthodologique ?

S'agissant d'identifier du « sens », seul un cerveau est capable d'une telle œuvre mais ce seul outil ne permet pas de satisfaire à la transparence et à un critère de reproductibilité de la démarche. C'est pourquoi le choix a été fait de réaliser l'exploration à l'aide d'une méthodologie publique.

La méthodologie adoptée est fondée sur le postulat selon lequel la distribution du lexique dans un corpus et plus précisément l'analyse des cooccurrences lexicales (qui sont des éléments objectifs) donne un accès, au moins pour partie, aux univers sémantiques qui structurent un corpus. J'ai cherché à identifier des variations formelles avec les années et des invariants dans l'emploi d'ensembles de mots lexicaux, postulant qu'ils pourront me servir d'indicateurs d'évolutions thématiques, méthodologiques, voire épistémologiques des recherches publiées dans la revue.

Dans une première partie, j'exposerai la méthodologie adoptée. Ayant grandi professionnellement dans cette équipe, je courais le risque de l'analyse subjective : j'ai donc pris des précautions méthodologiques pour que ma subjectivité soit limitée autant que possible. Dans une deuxième partie, je présenterai les résultats de l'analyse lexicale menée, en dégagant des tendances thématiques spécifiques pour chaque décennie ainsi que des invariants sur l'ensemble des 50 ans. Dans une troisième partie, je discuterai les limites de cette étude. Je conclurai en présentant quelques réflexions, toutes personnelles cette fois, que m'ont inspiré les résultats de cette étude.

1. Méthodologie

Mon objectif est donc de considérer cinquante ans de publications et d'en faire une synthèse.

Mais comment étudier les 371 articles de la revue publiés entre 1970 et 2020 ? La lecture intégrale n'était pas envisageable, faute de temps d'une part, et d'autre part parce qu'une synthèse d'un tel volume d'information n'est pas à la portée des capacités d'un cerveau humain (le mien, en tout cas). Une autre manière de procéder aurait pu être de lire les introductions des numéros, mais tous les numéros n'en comportent pas, ce qui rendait impossible de représenter les contenus de toutes les périodes.

Il fallait donc chercher un procédé de synthèse automatique qui me permette de résumer l'information sans perdre en qualité de représentation. Mais pour faire fonctionner un logiciel d'analyse lexicale, il fallait constituer le corpus.

1.1. Sélection des textes de l'échantillon

Différents procédés auraient été utilisables : j'aurais pu me baser sur une connaissance du contenu de ce qui différencie les textes, et échantillonner les textes sur les éléments qui les distinguent. Mais je n'avais pas cette connaissance. En outre, mes propres thèmes de prédilection auraient très probablement influencé la sélection.

Il m'a paru préférable d'effectuer un échantillonnage aveugle, sur la base du hasard, en espérant que le hasard permette d'obtenir un échantillon représentatif de l'ensemble, ce qui n'est jamais parfaitement le cas. Toutefois, j'ai été assez agréablement surprise du résultat.

J'ai procédé à un tirage au sort aléatoire de cinquante articles (en utilisant un logiciel de tirage au sort de nombres compris entre 1 et 371). J'ai dressé la liste des articles tirés au sort (cf. annexe), et me suis posé la question de savoir combien d'unités temporelles j'allais créer afin de mettre en évidence des évolutions. Connaissant les contraintes institutionnelles évoquées en introduction, j'aurais pu scinder l'analyse en deux époques, avant et après 2006. Mais des changements sont sûrement intervenus tout au long de la vie des *Mélanges*, certains dont je n'ai peut-être rien su, d'autres dont je ne me suis pas aperçue. Après quelques hésitations, j'ai opté pour faire l'analyse en cinq décennies, ainsi que sur la totalité. J'ai alors pu déterminer dans ma liste combien étaient prélevés par décennie.

Le résultat de cette extraction a été évalué en fonction de la connaissance que j'ai de la revue et du domaine de la didactique des langues et de la sociolinguistique. J'ai eu le sentiment que chaque décennie était bien représentée, avec 10 articles en moyenne, et des variations qui coïncident avec le nombre d'articles produits dans chaque décennie. Les thèmes des articles prélevés sont variés et m'ont paru relever de multiples thématiques et d'auteurs variés. Les limites de l'échantillon seront présentées en discussion.

	1970-79	1980-89	1990-99	2000-09	2010-20	Total
Nb articles revue	77	67	42	74	111	371
Nb articles Echantillon	11	10	8	7	14	50

Tableau 1. Nombre d'articles de la revue et de l'échantillon

Une grande partie des textes sélectionnés aléatoirement n'était pas utilisable telle quelle. En effet, les textes des numéros les plus anciens sont disponibles dans un format pdf de type image qui n'est pas interrogeable. J'ai donc utilisé une version de Word qui me permette de transformer le texte pdf en txt. J'ai dû contrôler le résultat très imparfait de ce traitement et procéder à des corrections et à un nettoyage de nombreuses coquilles liées aux signes mal identifiés par le logiciel. Pour les textes en anglais, je les ai traduits avec DeepL¹, traduction vérifiée et corrigée par mes soins.

Je disposais ainsi d'un ensemble de textes en format txt, qui a été soumis à l'analyse outillée que je décris maintenant.

1.2. Procédé d'analyse

Pour mener l'analyse, mon idée était de m'appuyer sur des éléments « objectifs » en ce qu'ils seraient reproductibles pour déterminer la distribution du lexique dans cinquante articles. Je suis partie d'un présupposé selon lequel le lexique est porteur d'une grande partie de la signification : ce sont surtout les cooccurrences qui m'intéressent, car elles concernent les relations entre les mots, et

¹ <https://www.deepl.com/fr/translator>

la structure de ces relations. L'objectif était de découvrir comment varient ces structures sur cinq décennies.

Pour ce faire, j'ai eu recours à deux outils informatiques : Alceste (version 2012) pour déterminer les variantes significatives d'ensembles thématiques, et Iramuteq (version 7 alpha 2) pour identifier des invariants thématiques sur l'ensemble de la période.

Je vais donner quelques informations sur ces deux logiciels, afin que le lecteur comprenne les données que je reproduis dans l'exposition des résultats qui suivra. Les lecteurs informés peuvent aller directement à la partie 2, Résultats, ci-dessous.

La méthodologie Alceste (Reinert, 1983, 1986, 1990) a été mise au point en vue de décrire la structure sémantique d'un corpus donné unique. L'objectif de Reinert est d'analyser dans un corpus donné la structure des distributions de trois types d'unités qui sont le mot, l'énoncé et le texte. L'analyse se fait à partir d'un tableau logique à double entrée, comprenant en lignes les énoncés et en colonnes les mots. L'analyse statistique a pour objectif de construire des classes d'énoncés utilisant plus spécifiquement tel ou tel type de vocabulaire.

Pour construire le tableau logique à double entrée, Alceste segmente arbitrairement le texte (définition des énoncés) et met en rapport le segment et les différents mots rencontrés dans le texte. Vu la diversité du lexique total, le résultat de cette mise en rapport est très souvent nul. Le logiciel doit donc résoudre un problème d'analyse de tableaux de données quasi « vides ». Il augmente les quantités statistiques en fusionnant des colonnes correspondant à des formes lexicales. Il opère ainsi diverses formes de réduction (par exemple, lemmatisation : les verbes sont ramenés à leur forme infinitive ; autre exemple, les mots grammaticaux sont exclus). Pour éviter un résultat dépendant du découpage arbitraire, Alceste procède à deux analyses, fondées sur des concaténations différentes de segments. Il ne retient de ces deux résultats que les communautés, réputées indépendantes du découpage.

L'objectif de l'analyse automatique est de répartir en deux classes les énoncés du corpus de telle façon que soit maximisée la différence de fréquence d'occurrence des mots dans les classes. L'analyse est descendante : une fois une classe

d'énoncés constituée, elle est répartie en deux nouvelles classes, et ainsi de suite jusqu'à un seuil fixé par l'utilisateur, comprenant un indice statistique pour évaluer la qualité de représentation du résultat, et un indice pour décider de la stabilité d'un résultat. C'est ainsi que sont obtenues des classes dites « stables ».

Iramuteq (Ratinaud, 2009) procède comme Alceste. Le même corpus a été soumis au traitement d'Iramuteq, avec un objectif différent : repérer un invariant sur les 50 années. Pour repérer les thématiques récurrentes sur l'ensemble du corpus, Iramuteq permet de représenter des associations à l'aide de différents coefficients. Pour l'analyse c'est le coefficient d'association, dit de Jaccard qui a été retenu. Le logiciel produit une représentation spatiale des proximités entre les mots dans les énoncés. Iramuteq a été aussi utilisé pour éditer un tableau de fréquence des mots, différenciés selon la décennie de publication.

2. Résultats par décennies

La présentation des résultats suit un plan identique pour faciliter le repérage des informations. D'abord des éléments chiffrés sur la masse textuelle traitée par Alceste, puis la structuration en classes thématiques qu'il a déterminées. Pour résumer l'information et caractériser les classes, j'ai relevé les mots les plus fréquents et les plus fortement liés à la classe (la statistique est celle du KHI2 et sa représentation standardisée à l'aide du coefficient PHI). Les crochets délimitent les sous-ensembles de la classe.

Exemple : Classe 1 : [texte, écrit, oral, compréhensi*], [document]

Je fais ensuite un bref commentaire sur la relation qui existe entre chaque classe et les numéros de la revue dont sont issues les formes (dans cet exemple : texte, écrit, oral, document). Pour ce faire, j'ai consulté les résultats et identifié les textes d'où proviennent les segments utilisés par Alceste, comme le montre cet exemple :

uce n° 1647 Phi = 0,03

(A) cause de leur (caractere) (unique), les (documents) (authentiques) ne peuvent pas etre (utilises) (bruts) comme modeles a (repetér) (ou) a (imiter), mais ils peuvent etre (exploités) en differe comme materiau d' (observation) preparatoire a l' (expression:)

Commentaire [EC4]: 1979

Enfin, je reproduis un tableau de fréquences des mots spécifiques à la période indiquée. Ce sont des mots très présents dans une période, et absents ou presque

absents dans les autres périodes. J'ai extrait ces spécificités d'un tableau général produit par Iramuteq sur les effectifs relatifs sur cinq décennies.

2.1. Première période : 1970-1979.

2.1.1. Eléments chiffrés :

L'analyse porte sur **11** textes (sur les 77 de la décennie)

Nb de formes contenues dans le corpus : 71216

Nb de formes distinctes : 6920

Nb de formes prises en compte pour l'analyse après réduction : 1341

Nb d'unités textuelles (ou unités de contexte élémentaires) : 1826

Classement des mots du lexique des onze textes selon leur effectif : Les 10 plus fréquents : langue, apprenti*, apprendre, texte, document, étudiant, oral, enseigner, travail*, écrit

2.1.2. Caractérisation des classes

Nb de classes stables retenues par Alceste : 3

Classe 1 : [texte, écrit, oral, compréhensi*], [document]

Les mots de cette classe thématique sont très bien répartis **sur toute la décennie**.

Classe 2 : [programme, étudiant, travail, bibliothèque, sonore], [laboratoire], [heure, semaine]

Les mots de cette classe sont relevés dans trois numéros : **1972, 1974, 1975**.

Classe 3 : [objectif, apprenant], [apprenti*], [apprendre, évaluation, performant]

Les mots de cette classe sont relevés dans quatre numéros : **1971, 1973, 1977, 1979**

Les 3 classes retenues sont représentées dans plusieurs des dix numéros de la période. Elle apparaît en conséquence assez homogène.

2.1.3. Eléments lexicaux spécifiques (Iramuteq):

	1970- 1979	1980- 1989	1990- 1999	2000- 2009	2010- 2020
bibliothèque	2,37	0	0	0	0

spot	3,54	0	0,08	0	0
littéraire	2,71	0,12	0,16	0	0,05
laboratoire	4,58	0,37	0,16	0	0,09
sonore	4,66	0,06	0,73	0,72	0,18
performance	3,37	0,44	0,49	0	0,5
bibliothèque	2,33	0,25	0,33	0,41	0,05
pédagogie	1,96	0,12	0,33	0,62	0,14
cassette	1,87	0,93	0,82	0,21	0
adulte	2,29	0,25	0,33	0,72	0,45
enregistrement	1,66	0,5	0,24	1,13	0,23
lire	2	0,68	1,06	3,4	0,64
test	1,25	0,19	0,82	0,72	0,59
motivation	1,54	0,62	0,82	0,31	0,23
méthode	2,37	1,55	0,82	0,21	1,04

Tableau 2. Liste des mots notablement plus employés dans la période 70-79

N.B. : les nombres du tableau sont à lire « pour mille » : 2,37 signifie, en pourcentage, 0,2% des mots lexicaux sur le nombre total de mots pour la décennie.

La liste est caractérisée par ce que peut permettre la technologie moderne de l'époque (enregistrement de la performance sur cassettes) pour l'apprentissage de l'oral par des adultes en bibliothèque sonore, opposée au laboratoire de langues, spécifiquement en compréhension. La lecture, et la nature littéraire des textes traditionnellement utilisés, sont également des thèmes fréquents de cette période.

2.2. Deuxième période : 1980-1989

2.2.1. Eléments chiffrés :

10 textes (sur 67)

Nb de formes contenues dans le corpus : 49113

Nb de formes distinctes : 5707

Nb de formes prises en compte pour l'analyse après réduction : 1069

Nb d'unités textuelles (ou unités de contexte élémentaires) : 1231

Classement des mots du lexique des dix textes selon leur effectif : Les 10 plus fréquents : apprenti*, apprenant, objectif, discours, travail, langue, apprendre, action, enseigner, communic*

2.2.2. Caractérisation des classes

Nb de classes stables retenues par Alceste : 7

Classe 1 : [plan], [interac*, acte, parole, participer, interactiv*, stratég*, discours], [partenaire, communic*, usager]

Les mots de cette classe sont typiques des numéros 1981, 1982 et dans une moindre mesure 1988.

Classe 2 : [recherche], [centre, chercheur, théori*, langue, enseignement]

Les mots de cette classe sont typiques du numéro 1989.

Classe 3 : [apprenant, enseigner, travail, groupe], [liberté, année, semaine, heure], [diriger, autodirigé, autonom*, système, apprenti*, école, mine*, anglais]

Les mots de cette classe sont typiques du numéro 1983.

Classe 4 : [étudiant, souhait, connaitre, culture], [spécialité, professeur, grammaire, salle]

Les mots de cette classe sont en très grande majorité sortis du numéro 1985.

Classe 5 : [évaluation, progresser, support], [conseil, arbre], [évolutif, schéma, domaine, entretien, critère, nombre, augmentation]

Les mots de cette classe sont typiques du numéro 1986-87.

Classe 6 : [transcri*, syntagme, cohésion, mot, porteur], [élément, compréhensi*, instruction, francophone, difficulté, reprise, percevoir, audit*]

Les mots de cette classe sont typiques du numéro 1982.

Classe 7 : [français, article, auteur, passif, verbe, malade], [passage, section, descriptif, séquence, assertion], [marque, page, intertitre, typographique, paragraphe, thème, unité]

Les mots de cette classe sont typiques du numéro 1980.

On constate beaucoup moins de continuité thématique entre les numéros : 5 des 7 classes retenues sont représentées dans un seul numéro de la décennie, qui en compte neuf. La classe 4 est très fortement majoritaire dans un numéro, seule la classe 1 est représentée principalement dans deux numéros, et dans un troisième à un moindre degré.

2.2.3. Eléments lexicaux spécifiques (Iramuteq):

	1970- 1979	1980- 1989	1990- 1999	2000- 2009	2010- 2020
communicatif	0,96	2,8	0,49	0,1	0,32
auteur	0,83	2,92	0,08	0,31	1,73
acte	1,12	5,72	1,14	0,41	0,95
séquence	0,25	2,61	0,08	1,34	1,5
énoncé	0,92	2,8	0,16	0,52	1,82
parole	1,17	2,98	0,16	1,55	0,73

Tableau 3. Liste des mots notablement plus employés dans la période 80-89

Ce qui frappe d'emblée est que cette liste est très marquée par l'avènement de l'approche communicative en didactique des langues et l'analyse de discours. L'influence de l'analyse de discours sur l'objet d'apprentissage apparaît clairement dans les numéros de la période.

2.3. Troisième période : 1990-1999

2.3.1. *Éléments chiffrés :*

8 textes sur 42

Nb de formes contenues dans le corpus : 35900

Nb de formes distinctes : 4622

Nb de formes prises en compte pour l'analyse après réduction : 862

Nb d'unités textuelles (ou unités de contexte élémentaires) : 926

Classement des mots du lexique des huit textes selon leur effectif : Les 10 plus fréquents : apprendre, langue, apprenti*, travail, élève, formation, apprenants, français, action, document

2.3.2. *Caractérisation des classes*

Nb de classes stables retenues par Alceste : 4

Classe 1 : [deuxième, élève, classe], [année, minute, travail, groupe], [soutien, autogérer, ressource, centre]

Les mots de cette classe sont typiques du numéro **22**

Classe 2 : [formation, éducation, polit*, niveau], [département, discours, didactique, scientifique, universit*, étudiant], [académi*, discipline, arabe, étude, problem*, supérieur, enseignement]

Les mots de cette classe sont typiques des numéros **1990 et 21**.

Classe 3 : [partenaire, bois, Xylolingua, grec], [moduler, présent*, utilisation, expressif, compréhensi*]

Les mots de cette classe sont typiques des numéros **23 et 24**.

Classe 4 : [agir, définir, évaluer], [apprendre, apprenti*, objectif, acquisition], [critère, décision, savoir, savoir-faire, techn*]

Les mots de cette classe sont assez bien répartis sur la période, puisqu'on les retrouve liés aux numéros **1990, 22, 23, 24**.

Cette période pourrait sembler plus homogène que la précédente, puisque seule la classe 1 est typique d'un seul numéro de la décennie, qui en compte cinq. Toutefois, quatre classes thématiques pour cinq numéros semble un nombre assez élevé au regard des trois classes bien réparties sur plusieurs des dix numéros de la première décennie.

2.3.3. Eléments lexicaux spécifiques (Iramuteq):

	1970- 1979	1980- 1989	1990- 1999	2000- 2009	2010- 2020
module	0,62	0,06	4,08	0	0,54
utilisateur	0,42	0	3,91	0,41	1
soutien	0	0,37	2,36	0,93	0,27
vidéo	0,17	0,12	1,88	0,52	0,86
grammaire	1,08	0,56	1,88	0,21	0,36
image	0,37	0,37	1,88	0,41	0,86
support	0,42	0,93	2,2	0,31	1,23
représentation	0,29	0,68	2,36	1,44	0,82

Tableau 4. Liste des mots notablement plus employés dans la période 90-99

Cette liste est marquée par l'utilisation des supports vidéo d'apprentissage de l'oral selon les besoins de l'utilisateur, dans des modules visant à faire évoluer les représentations grâce au soutien à l'apprentissage.

2.4. Quatrième période : 2000-2009

2.4.1. *Éléments chiffrés :*

7 textes sur 74

Nb de formes contenues dans le corpus : 33336

Nb de formes distinctes : 5265

Nb de formes prises en compte pour l'analyse après réduction : 895

Nb d'unités textuelles (ou unités de contexte élémentaires) : 830

Classement des mots du lexique des sept textes selon leur effectif : Les 10 plus fréquents : langue, étudiant, action, enfant, enseigner, accompagner, document, apprendre, apprenti*, autonom*

2.4.2. *Caractérisation des classes*

Nb de classes stables retenues par Alceste : 9.

Classe 1 : [taux, résultat, niveau, difficulté, facile], [suisse, émission, sonore], [document, oralisé, spontané, oral, compréhensi*, information, détailler, exercice, type]

Les mots de cette classe sont typiques du numéro **31a**

Classe 2 : [école, progrès, envie, livre, parler, lire, atelier, plaisir], [action, parent, langage, facilitateur, animation, an], [enfant, dialogue, adulte, réaliser, janvier, mars, juin, enregistrer, entretien, syntax*, complexité, introduction]

Les mots de cette classe sont typiques du numéro **29**.

Classe 3 : [direction, projet, autodirig*, apprenti*, formation], [innover, élève, enseigner, autonom*, représentation], [condition, initiation, phase, Holec, enseignement, produit]

Les mots de cette classe sont typiques du numéro **25**.

Classe 4 : [découverte, travaux, fonctionnement, observer, phénomène, variation], [étude, observation, exploiter, utilisation, corpus, travail], [locuteur, étudiant, réflexion, constitution, donnée, enquête, cours, sociolinguistique]

Les mots de cette classe sont typiques du numéro **31b**.

Classe 5 : [apprenants, langue, maternel*, texte, objectif, lecture], [romane, presse, séance, connaissance, stratég*], [lecteur, diversité, sens, formuler, hypoth*, exclusif, processus, comprendre, déchiffrer, sensib*]

Les mots de cette classe sont typiques du numéro **27**.

Classe 6 : [femme, identité, sexuel*], [catégori*, service, interac*, usager, prestataire, sentiment, négatif, positif, attente], [libraire, bibliothécaire, respect, financier, politesse, apprécier, comportement, échange, bonjour, contact, mauvais, agressif, sourire, absent]

Les mots de cette classe sont typiques du numéro **30**.

Classe 7 : [conversation, magasin, Brésil, vendeur], [Mexique, banque, tourisme, client], [attendre, Gabon, énerver, commerce, Chypre, Iran]

Les mots de cette classe sont typiques du numéro **30**.

Classe 8 : [cognitiv*, social, relation], [autoformation, accompagnateur], [paradigme, médiation, terme, technolog*], [accompagner, notion, transformation]

Les mots de cette classe sont typiques du numéro **28**.

Classe 9 : [situe, ligne, conception, cas], [tâche, fonction, apparaître, étape]

Les mots de cette classe sont typiques du numéro **28**.

Il semble qu'il y ait plus de diversité thématique. La période apparaît très hétérogène, comme le montre le nombre élevé de classes retenues, qui correspondent toutes à un seul numéro de la décennie, qui en compte sept.

2.4.3. *Éléments lexicaux spécifiques (Iramuteq):*

	1970-1979	1980-1989	1990-1999	2000-2009	2010-2020
accompagnement	0,04	0	0	8,87	0,73
client	0,25	0,12	0,08	6,39	0,32
enfant	0,33	0,12	0,41	10,31	2,36
livre	0,58	0,25	0,16	2,78	0,14
service	0,29	1,31	0,41	5,26	0,32
langage	0,42	0,31	0,82	3,71	0,73
réussite	0,54	0,06	0,24	1,55	0,86

transcription	0,83	1,24	0,82	4,54	0,54
contact	0,67	0,75	0,41	3,09	0,54
questionnaire	0,29	0,44	0,82	2,17	0,77
comportement	0,79	0,68	0,57	2,37	0,36
relation	2,29	1,06	0,98	5,26	1,45
maternel	1,33	0,31	0,9	2,47	0,68

Tableau 5. Liste des mots notablement plus employés dans la période 2000-09

La lecture de la liste ne me permet pas de faire une interprétation d'une thématique dominante pour la période. L'hétérogénéité thématique se retrouve également dans la liste des mots spécifiques. On peut déceler au moins trois thèmes, avec la « naissance » de la notion d'accompagnement, désignée dans d'autres décennies par « conseil » ou « soutien ». Le public de « client » est à relier à « service », « enfant » à « langage » et « livre ».

2.5. Cinquième période : 2010-2020

2.5.1. *Éléments chiffrés :*

14 textes sur 111

Nb de formes contenues dans le corpus : 75061

Nb de formes distinctes : 8356

Nb de formes prises en compte pour l'analyse après réduction : 1346

Nb d'unités textuelles (ou unités de contexte élémentaires) : 1909

Classement des mots du lexique des 14 textes selon leur effectif : Les 10 plus fréquents : langue, enseigner, action, étudiant, enseignement, apprenti*, recherche, élève, genre, compétent

2.5.2. *Caractérisation des classes*

Nb de classes stables retenues par Alceste : 7

Classe 1 : [input, transfert, enfant, langue de scolarisation] [communauté familiale, diaspora, enquête, parler, maternel*].

Les mots de cette classe sont typiques des numéros 39 et 35.

Classe 2 : [primaire, milieu, inclusif], [facteur, réussite, condition] [orthographe, élève, enseignement, apprenti*, lexical, explicite], [dictée, mot, phonème, régularité], [règle, expliquer]

Les mots de cette classe sont typiques du numéro 39.

Classe 3 : [entrevue, universit*, littérat*, profession] [stratég*, pédagog*, scripteur, produire, genre, écrit, initial, formation]

Les mots de cette classe sont typiques du numéro 37.

Classe 4 : [autonom*, projet, disposition, étudiant] [site, apprenants, ressource, utilisation, corpus, concordancier]

Les mots de cette classe sont typiques des numéros 32, 37, (38), 40.

Classe 5 : [société, curriculum, actionnelle, cultur*] [ethnographi*, professionnel*]

Les mots de cette classe sont typiques des numéros 34 et 40.

Classe 6 : [norme, social, rapport, sexe, femme, homme], [énonc*, ironie, domestique], [humour, valeur, rire, support, inégal*, masculin, personnage, féminin]

Les mots de cette classe sont typiques du numéro 36.

Classe 7 : [interculturel, changement, questionner], [dialogue, dynam*, diariste, altérit*, critiquer, point de vue] [(branche avec des fréquences plus faibles) vie, quotidien, collectif, interprétation, sens, chercheur, agir, construction]

Les mots de cette classe sont typiques du numéro 41.

Les classes 1 et 5 sont liées à deux numéros. 4 classes sont représentées dans un seul numéro. Seule une classe est représentée dans 4 numéros de la décennie, sur un total de quatorze.

La période apparaît donc assez hétérogène. Toutefois, l'utilisation de ressources par des apprenants/étudiants semble un thème plus commun que les autres.

2.5.3. Eléments lexicaux spécifiques (Iramuteq):

	1970- 1979	1980- 1989	1990- 1999	2000- .2009	2010- 2020
enquêtés	0	0	0	0	2,32
kurde	0	0	0	0	4,04
fleuron	0	0	0	0	2,32
orthographe	0,04	0	0	0,1	3,22
site	0	0,06	0,33	0,1	3,27
humour	0,12	0,25	0,08	0,1	4,63

transfert	0,37	0	0,08	0,1	1,86
genre	0,75	0,19	0,08	0,82	7,45
origine	0,33	0,25	0,08	0,1	2,86
expert	0,08	0,06	0,33	0,21	2,27
critique	0,25	0,31	0,08	0,21	2,63
femme	0,08	0,12	0,08	1,65	2,5
formateurs	0,33	0,19	0,24	0,1	2,09
universitaire	0,17	0,37	0,57	0,41	3,72
dispositif	0,62	0,06	0,08	1,86	2,91
valeur	0,71	1,62	0,24	1,13	4,45
perspective	0,17	0,19	0,33	1,03	1,68
chercheur	0,46	1,18	0	1,34	2,59
langagières	0,33	0,31	0,24	0,52	1,86
projet	0,25	0,99	1,3	1,96	3,95
france	0,33	0,56	0,65	1,24	2,36
culturel	0,46	0,99	1,3	1,13	3,59
scolaire	0,54	0,5	0,57	0,31	1,95
modèle	1,04	1,37	0,73	0,82	3,81
expliquer	0,46	0,99	0,9	1,55	2,91

Tableau 6. Liste des mots notablement plus employés dans la période 2010-20

La lecture de la liste ne me permet pas de faire une interprétation d'une thématique dominante pour la période.

Les 3 premiers mots « enquêtés », « kurde » et « FLEURON », sont très employés dans les textes de la dernière décennie et complètement absents des textes d'autres décennies. Est-ce un effet du tirage au sort ? Sans doute. Un autre article de la période porte sur le Maroc ; s'il avait été tiré au sort, « marocain » remplacerait probablement « kurde ». Ce qui importe est de constater la spécificité de la période dont on peine à interpréter le sens. On remarque que « chercheur » et « universitaire » sont plus fréquents en cette période récente, de même que « projet », qui correspond peut-être à l'évolution du financement des recherches par appels à projets.

3. Résultats sur l'ensemble des 50 années

3.1. Éléments chiffrés :

50 textes sur 371.

Nb de formes contenues dans le corpus : 256640

Nb de formes distinctes : 15333

Nb de formes prises en compte pour l'analyse après réduction : 1358

Nb d'unités textuelles (ou unités de contexte élémentaires) : 6522

Classement des mots du lexique du corpus total selon leur effectif : Les 10 plus fréquents : langue, apprenti*, apprendre, enseigner, travail, apprenant, étudiant, action, objectif, enseignement

3.2. Caractérisation des classes

Nb de classes stables retenues par Alceste : 4

Classe 1 : [interact*, social*}

Classe 2 : [texte, lecture], [authent*, document], [compréhensi*,oral, écrit, expressif]

Classe 3 : [apprendre, apprenti*, autonom*, évaluation]

Classe 4 : [travail, groupe], [cours, étudiant, année, heure, semaine]

Les classes 2, 3 et 4 reflètent le versant didactique des publications, la classe 1 est plus orientée sur la sociolinguistique. Cette proportion n'est pas étonnante si l'on observe un résumé des classes thématiques sur les cinq périodes :

1970-79	1980-89	1990-99	2000-2009	2010-2019	50 ans
3 classes	7 classes	4 classes	9 classes	7 classes	4 classes
laboratoire, bibliothèque, étudiant, travail, heure, semaine	interaction, stratégie, discours, communicat*	élève, classe, année, minute, groupe, soutien, autodirig*, centre, ressource	niveau difficulté, émission, sonore, document, oral, compréhens*, exercice	acquisition, langue, seconde, communauté familiale, scolarisation	interact*, social*
oral, écrit, document, texte, compréhension	recherche, théori*, langue, enseignement	formation, éducation, niveau, discours, didactique, étudiant, enseignement, supérieur	école, progrès, livre, plaisir, action, parent, langage, enfant, adulte, entretien, complexité, syntax*	primaire, réussite, orthographe, élève, apprenti*, enseignement, dictée, phonème, règle	texte, lecture, document, authent*, compréhens*, oral, écrit
apprenti, apprendre, apprenant, objectif, évaluation	apprenant, enseigner, année, semaine, autonom*, autodirig*	Xylolingua, moduler, utilisation, expressif, compréhens*	autodir*, apprenti*, formation, innover, autonom*, Holec, enseignement	universit*, littérat*, pédagog*, scripteur, texte, genre, formation	apprendre, apprenti*, autonom*, évaluation
	étudiant, culture, professeur, grammaire	agir, évaluer, apprendre, apprenti*, objectif, acquisition, critère, décision	découverte, observer, variation, exploiter, utilisation, corpus, étudiant, enquête, sociolinguist*	autonom*, étudiant, ressource, utilisation	travail, groupe, cours, étudiant, année, heure, semaine
	évaluation, progresser, conseil, entretien, critère		apprenants, langue, maternel*, texte, romane, stratég*, processus, comprendre	curriculum, cultur*, professionnel*	
	transcri*, syntagme, cohésion, compréhens*, difficulté, audit*		femme, identité, interac*, prestataire, usager, sentiment, attente, comportement, échange	norme, social*, sexe, femme, homme, énonc*, humour, inégal*	
	article, passif, verbe, descriptif, page, titre, typographie		conversation, magasin, vendeur, tourisme, client	interculturel*, changement, dialogue, altérit*, critiquer, point de vue	
			cognitiv*, social, relation, accompagnateur, médiation, transformation		
			ligne, conception, cas, tâche, étape		

Tableau 7. Résumé des classes thématiques du corpus

Avec Iramuteq, une analyse fondée sur l'intensité des cooccurrences des termes qui apparaissent au moins 100 fois dans l'ensemble du corpus des 5 décennies fait apparaître une organisation thématique schématiquement résumée ainsi :

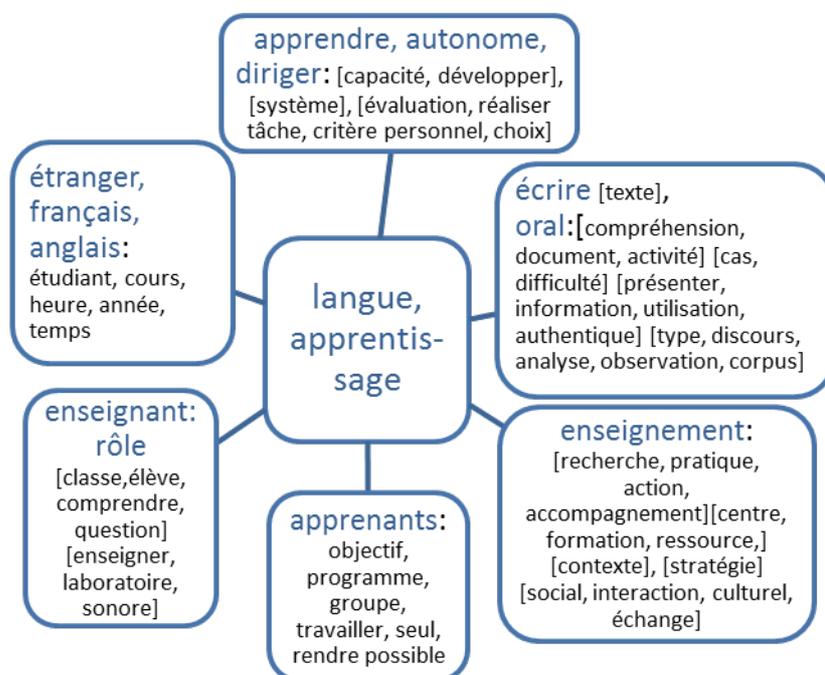


Schéma 1. Organisation thématique de la revue sur cinquante années

3.3. Eléments lexicaux constants sur les 5 décennies (Iramuteq):

Le tableau ci-dessous présente la proportion pour mille de l'occurrence d'un mot par rapport à l'ensemble des occurrences de toutes les formes du corpus qui apparaissent au moins dix fois. Je ne reproduis ici que les valeurs les plus élevées (les données initiales portent sur 545 formes). Dans le tableau, sont soulignées par un fond jaune les fréquences supérieures à la moyenne des valeurs. Celle-ci se situe autour de 5‰, ou 0,5%, correspondant à un mot pour 200 mots lexicaux. En bleu clair ce sont les valeurs les plus faibles qui sont mises en évidence, autour et inférieures à un pour mille mots lexicaux.

	1970-1979	1980-1989	1990-1999	2000-2009	2010-2020
langue	19,94	10,57	23,65	14,95	18,39
apprentissage	19,65	15,1	21,12	8,35	9,95
apprendre	11,07	10,69	31,23	8,66	4,81
apprenants	7,49	14,17	9,46	6,39	4,27
étudiant	8,99	4,29	7,5	11,65	10,31
enseignant	8,03	6,53	4,73	9,9	12,67
enseignement	7,28	3,73	5,63	2,47	10,31
enseigner	6,91	0,19	1,14	0,31	3,22
acquisition	1,91	1,12	2,28	1,34	1,04
compétence	3,16	3,36	2,45	1,86	6,58
classe	2,37	2,11	1,96	1,03	6,54
cours	6,45	5,66	4,81	4,54	4,18
groupe	4,91	5,03	6,36	1,96	3,22
situation	6,37	3,54	4,65	3,92	5,22
travail	5,91	9,57	9,13	6,7	5,77
travailler	2,54	1,74	2,69	2,78	1,27
pratique	1,46	1,99	3,18	3,71	5,4
écrire	8,16	3,42	5,87	3,71	6,49
activité	1,96	4,91	7,58	3,92	6,36
mot	1,17	2,61	4,24	1,65	5,81
français	1,17	4,04	8,73	6,6	5,63
anglais	6,66	5,28	4,73	0,52	1,63
étranger	2,33	1,99	4,4	4,95	1,32
linguistique	4,7	3,42	3,59	1,44	4,5
didactique	1,42	2,61	1,63	1,86	4,31
formation	2,91	2,3	9,3	6,7	4,13
pédagogique	4,12	1,8	2,36	2,27	4,13
comprendre	2,83	2,98	2,94	6,08	3,81
compréhension	6,78	3,54	7,09	6,39	1,36
texte	9,32	4,79	3,51	6,08	2,86
document	9,24	2,36	8,32	9,18	3,68
difficulté	2,37	2,73	2,61	4,64	2
oral	8,66	4,29	3,67	5,05	2,5
niveau	6,99	3,42	2,12	3,2	4,13
temps	3,91	3,98	3,83	5,67	2,59
centre	0,54	4,41	4,65	1,24	0,95
autonomie	3,25	2,05	1,96	5,36	1,41
besoin	4,95	3,42	3,67	1,24	1,41
objectif	7,45	13,55	6,85	2,89	3,59
évaluation	4,04	4,1	2,94	2,47	2,13
contexte	0,67	1,8	0,73	1,44	4,54
interaction	2,29	1,99	0,57	5,16	4,09
discours	1,96	11,75	2,85	1,34	1,91

Tableau 8 : fréquence de 43 formes fréquentes dans le corpus total (exprimée en %)

On remarque que les premiers mots (langue, apprentissage, apprendre, apprenants, étudiant, enseignant, enseignement) pèsent beaucoup et sont réguliers. Ce n'est pas une surprise dans la mesure où ils sont apparus régulièrement dans les dix mots les plus fréquents de chacune des périodes étudiées.

Si l'on veut parler en pourcentage, on peut dire que le mot « langue » est très présent tout au long des cinquante années, avec des valeurs comprises entre 1% et 2,3% des mots lexicaux : on le lit en moyenne entre une et deux fois tous les cent mots lexicaux.

Concernant le thème de l'apprentissage, si j'additionne les nombres de la forme « apprendre » et celle d'« apprentissage », dans l'ordre des cinq périodes j'obtiens environ 3%, 2,5%, 5,2%, 1,6% et 1,4%. Ce thème apparaît de manière constante et est très présent dans les articles de la revue, avec un pic lors de la décennie 1990-99, où l'on peut lire l'une des deux formes cinq fois tous les cent mots lexicaux en moyenne.

On constate également que lorsque apprentissage, apprendre et apprenants baissent notablement en fréquence sur la période 2010-20 relativement aux précédentes décennies, les mots « enseignement, enseignant et enseigner » deviennent plus fréquents. Il semble que l'enseignement supérieur soit plus souvent le contexte des études, puisque « étudiant » est en augmentation dans les deux dernières décennies. Evoquer les cours est une régularité sur les cinquante années. En revanche, comme la fréquence relative des termes besoins, objectifs, évaluation semble baisser sur les deux dernières décennies, en même temps que apprendre et apprentissage, on peut supposer qu'ils persistent toutefois, mais plus souvent dans des contextes d'enseignement. Cette évolution coïncide avec le fait que, depuis 2006, les auteurs viennent majoritairement d'autres laboratoires ou équipes de recherche que celle du CRAPEL/ATILF : leurs recherches sont moins influencées par l'intérêt fort de l'équipe pour l'apprentissage, autodirigé en particulier.

La langue d'apprentissage la plus étudiée change aussi selon les époques : on voit clairement que l'anglais domine lors des deux premières décennies, alors que le français est beaucoup plus présent lors des trois décennies suivantes. Cette évolution s'explique sans doute par l'essor du français langue étrangère comme domaine de recherche et de formation.

Certains mots renvoient à des objets d'étude historiquement marqués : ainsi « discours » a une fréquence élevée entre 1980 et 1989, de même que « parole » et « acte ». « Interaction » est plus fréquent dans les deux dernières décennies. On le trouve toutefois dans le numéro 1978, où Gremmo, Holec et Riley en donnent une définition pour l'analyse de la communication orale. Le terme est très utilisé dans la décennie suivante, on trouve ainsi « type d'interaction, acte interactif, une interaction de type argumentatif, une seule interaction ». En 1990-99, le terme apparaît beaucoup moins, et il revient en 2000-2009 avec interaction de service (dans le n°30), et interaction sociale (n°28). En 2010-2020, le terme est bien distribué dans plusieurs numéros : interactions de classe (34), interactions sociales (40), interactions à la scolarité, entre natifs et non-natifs (37), les interactions en famille (39).

Discussion.

Le tirage au sort m'a évité de « préférer » certains thèmes ou certains auteurs : la sélection a bien été aléatoire, et le corpus pour l'analyse est indépendant de l'auteur.

Le choix d'analyser les résultats décennie par décennie est un compromis entre simplicité et représentativité. Pourquoi pas tous les 7 ans ? tous les 9 ans ? Or, pour avoir un résultat entier de division en périodes, 50 est divisible par 2, 5, 10 ou 25. Il s'agissait de rendre compte de 50 années d'existence de la revue : diviser en 2 périodes de 25 ans me paraissait insuffisamment détaillé ; en 10 périodes de 5 ans trop détaillé, a fortiori en 25 périodes de 2 ans. Donc il restait la solution de faire 5 périodes de 10 ans. Il est possible aussi que l'existence même du mot décennie m'ait aidée dans mon choix. Il a l'avantage de la lisibilité. A posteriori, je m'aperçois que le découpage par décennies a mis en évidence des spécificités pour les trois premières décennies, mais a peut-être moins de pertinence concernant les deux dernières décennies.

J'ai voulu m'assurer de la représentativité de mon échantillon. Pour cela, j'ai procédé à un calcul consistant à chiffrer la part que constituent les articles d'une période sur le total des 371 (proportion revue-décennie/ nb total), et à faire de même sur les 50 articles de l'échantillon (proportion échantillon-décennie/nb total):

	1970-79	1980-89	1990-99	2000-09	2010-20	Total
Nb articles revue	77	67	42	74	111	371
Nb articles Echantillon	11	10	8	7	14	50
% revue- décennie /nb total	20.8 %	18.1%	11.3%	19.9%	29.9%	100%
% échantillon- décennie /nb total	22.0%	20.0%	16.0%	14.0%	28.0%	100%

L'écart observé entre la population totale et l'échantillon est un événement aléatoire dont la probabilité d'occurrence est de 77%. On ne trouve donc pas de différences significatives entre la distribution des articles selon les décennies dans la revue et la distribution des articles dans l'échantillon.

Si j'avais étudié une autre extraction aléatoire, le résultat serait-il le même ? La réponse est négative, sans doute. Mais les différences entre les deux analyses seraient-elles significatives ? Non, sans doute, les différences seraient minimes. Elles affecteraient la liste des mots spécifiques à une période et totalement absents des autres, comme l'exemple de « kurde ». Mais les mots relevant du thème central de l'étude (la langue de scolarisation ou l'acquisition lexicale par exemple), qu'elle porte sur des apprenants kurdes ou marocains, resteraient constants.

La combinaison du tirage au sort et du groupement en périodes de 10 ans conduit probablement à un résultat différent de celui que l'on obtiendrait sur la base d'une autre sélection et d'une autre division temporelle. Mais ce qui me paraît important est d'avoir obtenu une sélection d'un article par numéro, parfois deux lorsque le nombre d'articles était plus élevé que la moyenne et augmentait donc les probabilités d'un deuxième tirage dans le même numéro. J'ai fait l'expérience d'une variation dans la sélection : sur la période 2000-2009 (prise au hasard), la présence ou l'absence d'un article fait varier le résultat de l'analyse thématique, très peu pour ce qui concerne les mots les plus fréquents sur les cinq décennies, mais surtout pour ce qui concerne le nombre de classes retenues dans la période. J'en conclus qu'une autre sélection ferait varier quelques mots spécifiques d'une période, mais probablement pas les invariants.

Conclusion

L'ensemble de cette étude indique une évolution des centres d'intérêt : on parle de moins en moins d'apprentissage et de plus en plus d'enseignement. L'enseignement reste une thématique de recherche plus répandue. Logiquement, le mot « étudiants », est plus fréquent qu'« apprenants ». Le public immigrant est très peu présent : 19 occurrences en 2010-20, par comparaison avec étudiant : 227.

De la même manière, les thèmes comme l'oral, la compréhension, l'autonomie, sont moins fréquents dans la revue depuis 2010. Le fait que la revue ne reflète plus les travaux de l'équipe est probablement lié à cette évolution. D'une certaine manière, la revue a gagné en diversité, mais a perdu son originalité dans le paysage de la didactique des langues. La voix du CRAPEL dont parlait Henri Holec pour le trentième anniversaire des *Mélanges Crapel*, « attendez-vous à ce que la voix du CRAPEL se fasse encore entendre pour bousculer un certain nombre d'idées reçues, car c'est la raison d'être des chercheurs crapéliens » (Holec, 2000, p.12), ne passe plus par ce support.

Il y a des mots émergents à partir de 2000 : c'est le cas de corpus. Il ne s'agit pas de corpus pour la recherche. Dans un article du n°37, il s'agit d'un support de l'action de l'apprenant : analyse de/ travail sur corpus. Le mot « donnée » dans le numéro 31b concerne aussi le travail de l'apprenant, qui analyse des données langagières. En revanche, il s'agit de « récolte de données » pour la recherche dans les numéros 39 et 35, expression qui mériterait d'ailleurs d'être révisée, car en utilisant recueil ou récolte des données, on masque le fait, important sur le plan heuristique, qu'avant tout on les produit.

Le travail sur données langagières par l'apprenant est un thème récurrent dans les *Mélanges*. Dans les années 1970-80, les enseignants-chercheurs du CRAPEL cherchaient à permettre ou faciliter le travail individuel, sur objectifs individuels et supports choisis par l'apprenant (étudiant) parmi une sélection effectuée par les enseignants. La diversité des supports était moins accessible par l'apprenant en dehors de l'université et sa bibliothèque sonore. Les documents étaient décrits par des mots clés thématiques, un résumé, des indications de débit, longueur, d'accents. On signalait l'existence d'un document écrit pouvant servir à comprendre le document sonore (cf. Riley & Zoppis, 1976). Ensuite, la fréquentation

du document permettait à chacun de découvrir des aspects diversifiés ; pour les uns, une compréhension globale, pour d'autres, très détaillée, pour d'autres encore un repérage d'utilisation de règles ou de mots grammaticaux, ou une découverte de « manières de dire » à noter et réutiliser en production.

L'utilisation de corpus pour l'apprentissage, via un site, est une évolution technologique de ce qui précède qui me paraît avoir modifié le rapport au document : comme une commande sur Internet, le produit est « dématérialisé ». On le cherche avec sa tête, à l'exclusion du toucher, du déplacement dans des étagères, etc. La fourniture de ressources est, comme précédemment, conçue et mise en œuvre par des enseignants et/ou des chercheurs. Comme précédemment, la nature et les contenus du catalogue dépendent des objectifs du site d'apprentissage. Quand ils visent uniquement l'utilisation individuelle, les ressources sont assorties d'aides de différentes sortes : sous-titres, script, résumé, longueur. S'ils visent l'apprendre à apprendre, des conseils de travail peuvent être ajoutés aux ressources, ainsi que toutes sortes d'outils permettant à des apprenants de se mettre en contact avec des pairs ou avec des conseillers. L'autonomisation a été abondamment documentée, mais il reste à mon avis une question cruciale : à quelles acquisitions langagières conduisent la fréquentation de corpus, de documents bruts, etc. ?

L'acquisition de connaissances et de compétences en langues par des apprenants est le but de l'action des pédagogues et des recherches en didactique des langues. Dans la revue, les descriptions de systèmes, de dispositifs d'enseignement et d'apprentissage, les observations de l'apprentissage par les trajectoires, les activités d'apprenants, sont nombreuses. Mais évalue-t-on leur effet sur l'acquisition de langue? On étudie l'enseignement, l'apprentissage, phénomènes visibles, mais l'évaluation se cantonne souvent à la satisfaction des utilisateurs, pas à celle du résultat acquis.

En effet, cette étude m'a permis de constater la faible présence du thème de l'acquisition, même si le mot apparaît de manière constante. Régulièrement évoqué, le mot est fréquent entre 1 et 2 pour mille quelle que soit la décennie, mais l'acquisition ne fait pas l'objet de recherche. En 1970-79, on parle d'acquisition de compétence communicative (1977), acquisition de savoirs et savoir-faire (1979), en tant qu'objectifs de l'apprentissage ou de l'enseignement, sans jamais mesurer les

acquis. En 1980-89 : acquisition de la compétence langagière, communicative ou d'apprentissage, de techniques de lecture, de formes linguistiques. En 1990-99 : objectif d'acquisition, acquisition d'une nouvelle culture d'apprentissage. Dans la période 2000-2009 : le mot est surtout présent, évidemment, dans le numéro spécial « acquisition » (n°29) : on trouve alors mentionnés les acquis (évaluation des), (des élèves). Chez l'enfant : acquisition de structures langagières, de la lecture. En 2010-20 : dans le numéro 32, Belan et Narcy parlent d'« acquisitions langagières », et d'« analyse des productions des étudiants pour mesurer les acquisitions ». Mais cela ne fait pas l'objet de l'article. Dans le n°35, on trouve une étude des acquisitions en langue de scolarisation et en langue d'origine chez des enfants. Dans le n°37-1, l'évaluation des acquis dont il est question est diagnostique, et l'évaluation du modèle d'enseignement dont il est question n'est que souhaitée.

Il en va de même du mot « progrès », qui est une autre manière de considérer l'acquisition de savoirs et/ou de savoir-faire. En 1983, Moulden écrit que l'enseignant qui fait travailler un apprenant en autonomie doit « voir s'il garde des traces de ses activités et contrôle ses progrès ». Un article de Gremmo et Holec en 86-87 est entièrement consacré à l'observation des progrès d'un apprenant dans son acquisition de compétence d'apprentissage. En 1978, on fait « noter les priorités et les progrès réalisés ». Les auteurs Gremmo, Holec et Riley proposent une question pour le futur fort intéressante : « Cette méthode a-t-elle permis un apprentissage plus efficace que celles employées précédemment ? ». Aucune étude, à ma connaissance, ne répondra à cette question. Il est question, de loin en loin, de progrès et de progression (n°22, 24, 29, 39), mais sans mesure.

Finalement, il semble qu'alors que les enseignants-chercheurs mettent beaucoup d'énergie et de temps à mettre au point des systèmes ou dispositifs d'apprentissage sophistiqués, le résultat de l'apprentissage, qui est l'objectif visé par ce déploiement d'énergie, n'est pas mesuré.

Comment cela se fait-il ?

Faut-il voir dans ce constat des difficultés liées au modèle de recherche en didactique des langues ?

Dans ce domaine, on revendique l'utilité sociale : une partie d'ingénierie consiste donc à partir d'une situation déterminée comme problématique et nécessitant amélioration, un « terrain », et à créer des solutions pour apporter des réponses. Mais il est bien fréquent que ces interventions se limitent à cette dimension ingénierique, sans permettre de répondre à des questions de recherche qui, en général, n'ont pas précédé l'intervention. Celle-ci se déroule en temps limité, et la recherche demande du temps, ainsi que bien construire, tester un protocole d'observation, analyser des résultats et évaluer l'ensemble. La pression qui pèse sur les enseignants-chercheurs, résumée parfaitement par cette belle paire anglaise, *publish or perish*, est une contrainte antinomique aux progrès de la connaissance scientifique. Les publications sur des recherches en cours relèvent alors souvent davantage de l'anecdote que de l'apport d'information sur les résultats d'une méthodologie de recherche, d'une méthodologie d'enseignement/apprentissage sur l'acquisition.

Dans cet extrait du numéro 1982, Hélène Cabut décrit la difficulté que je viens d'évoquer :

Il est certes bien difficile d'évaluer la progression des apprenants dans leur compréhension à l'intérieur d'un texte. Dans ce cas précis, il est en effet pratiquement impossible de mesurer « objectivement » les difficultés relatives que présentent les différentes reprises, tant est grand, comme nous l'avons vu, le nombre de facteurs qui entrent en ligne de compte. Même si l'on réussit à établir avec une certaine précision, par l'analyse du discours, quelles reprises ont le plus de chances d'être perçues par les apprenants et pour quelles raisons, et lesquelles poseront le plus de problèmes, la série d'exercices proposée était trop limitée, trop partielle, pour que l'on puisse en tirer des conclusions certaines quant à la progression des apprenants.

Philip Riley exprime très clairement en 1989 que :

« pour l'enseignant comme pour le chercheur, "il n'y a rien de plus pratique qu'une bonne théorie". Ce n'est que sur la base d'une théorie générale que l'on peut identifier et analyser - conceptualiser - des problèmes spécifiques. Et seule une théorie nous permet de générer des hypothèses, de concevoir des expériences et d'évaluer les résultats. »

Il ajoute que :

« la recherche didactique (...) doit donc être considérée non pas comme l'application de théories provenant d'autres disciplines telles que la psychologie, la sociologie, la philosophie

ou même la linguistique, mais comme le développement d'une théorie (et d'un ensemble de connaissances) de l'enseignement et de l'apprentissage des langues. »

Aujourd'hui, j'ai le sentiment à la lecture d'une multiplicité d'articles que le lien entre les connaissances générales issues d'interventions antérieures et les interventions ponctuelles n'est pas établi, ou plutôt, n'est pas biunivoque. Il me semble nécessaire de faire le point, d'en faire une analyse afin de les grouper et en tirer soit des connaissances d'un niveau généralisable, soit des directions de recherche précises auxquelles un ensemble d'expériences pourraient participer. C'est peut-être un objectif inatteignable, dans la mesure où, pour répondre aux exigences quantitatives de publication, les chercheurs limitent leur recherche à un aspect original de chaque expérience et en rendent compte de telle façon qu'il n'est pas possible de reproduire la recherche.

La revue *Mélanges Crapel* avait une identité bien identifiée jusqu'au début du 21^{ème} siècle. Elle s'est diversifiée, avec pour conséquence une perte de fil conducteur. Pour contribuer plus efficacement à la construction de connaissances, pourrait-elle orienter un certain nombre de numéros spéciaux autour d'un but à moyen ou long terme ?

Annexe. Liste des textes de l'échantillon utilisé pour l'étude, par ordre chronologique

1. 1970 : Châlon, Y. Pour une pédagogie sauvage
2. 1970 : Bouillon, C. L'entraînement à l'expression écrite : compte-rendu d'une expérience
3. 1971 : Bouillon, C. Du laboratoire de langues à la bibliothèque sonore : l'individualisation de l'apprentissage en langues vivantes
4. 1972 : Boulanger, C., Heddesheimer, C., Holec, H., Muller M-S, Roussel, F., Towell, R., Zoppis, C. Le discours oral
5. 1973 : Cembalo, S.M., Holec, H. Les langues aux adultes : pour une pédagogie de l'autonomie
6. 1974 : Harding, E, Legras, M. La bibliothèque sonore et ses implications pédagogiques
7. 1975 : Abe, D., Henner-Stanchina, C., Smith, P. New approaches to autonomy : two experiments in self-directed learning
8. 1976 : Abe, D., Duda, R., Henner-Stanchina, C. Un spécialiste - quel spécialiste ?
9. 1977 : Henner-Stanchina, C, Holec, H. Evaluation in an autonomous learning scheme
10. 1978 : Gremmo, M-J, Holec, H., Riley, P. Taking the initiative : some pedagogical applications of discourse analysis
11. 1979 : Abe, D., Carton, F., Cembalo, S.M., Régent, O. Didactique et authentique : du document à la pédagogie
12. 1980 : Régent, O. Approche comparative des discours de spécialité pour l'entraînement à l'anglais écrit
13. 1981 : Carton, F. Le discours oral en élaboration : directions de recherche
14. 1982 : Cabut, H. Français langue seconde : les apprenants face au discours didactique oral
15. 1983 : Moulden, H. Apprentissage auto-dirigé : compte-rendu d'expériences 1978-1983
16. 1984 : Carton, F. Systèmes « autonomisants » d'apprentissage de langues
17. 1985 : Trompette, C. Quand les apprenants seront au centre

18. 1986-87 : Duda, R. Le discours écrit incohérent : les enseignements pour une pédagogie de l'expression ?
19. 1986-87 Gremmo, M-J, Holec, H. Evolution de l'autonomie de l'apprenant : le cas de l'apprenant D
20. 1988 : Duda, R., Carton, F. Production orale : comment mettre en place des stratégies d'enseignement-apprentissage
21. 1989 : Riley, P. « There is nothing as practical as a good theory » : research teaching and learning functions of language centres
22. 1990 : Bailly, S., Petitdemange, G. Construire son apprentissage : compte-rendu d'une expérience d'auto-apprentissage avec soutien en anglais pour étudiants en architecture
23. 1990 : Holec, H. Qu'est-ce qu'apprendre à apprendre ?
24. n°21 : Esch, E., King : Monsieur Dumas, qui c'est ? Documents authentiques et appropriation du rôle de destinataire par les apprenants
25. n°22 : Gremmo, M-J. Former les apprenants à apprendre : les leçons d'une expérience
26. n°22 : Trebbi, T. Apprentissage uto-dirigé et enseignement secondaire : un centre de ressources au collège
27. N°23 : Duda, R. Video Im Deutch Unterricht : un outil télévisuel au service de l'allemand
28. N°23 : Régent, O. Conception d'un produit multimedia : XYLOLINGUA
29. N°24 : Debaisieux, J-M, Régent, O. Un outil multimédia pour apprendre à apprendre les langues étrangères
30. N°25 : Carette, E. Introduction de l'autonomie dans le système éducatif français : des réponses, des questions
31. N°27 : Debaisieux, J-M. Une expérience de lecture en langues romanes
32. N°28 : Barbot, M-J. Rôle de l'enseignant formateur : l'accompagnement en question
33. N°29 : Bosseau, A. Des « facilitateurs du langage » auprès de jeunes enfants en activité périscolaire
34. N°30 : Bailly, S. Prestataires français, usagers étrangers : quels points de rencontre dans l'interaction de service ?
35. N°31 : Tyne, H. Corpus oraux par et pour l'apprenant

- 36.N°31 : Parpette, C., Carette, E., Duda, R., Carton, F., Mangiante, J-M. Table ronde : faut-il aménager les documents authentiques en vue de l'apprentissage ?
- 37.N°32 : Belan, S., Narcy-Combes, M-F. Accompagnement et innovation : quels outils pour un suivi efficace dans un dispositif hybride ?
- 38.N°33 : Duda, R. Au secours ! les 3 « c » K arrivent
- 39.N°34 : Moore, D., Sabatier, C. Cultures éducatives partagées : ethnographie de la salle de classe, postures de recherche et formation des enseignants
- 40.N°35 : Lambelet, A., Desgrippes, M., Decandio, F., Pestana, C. Acquis dans une langue, transféré dans l'autre ?
- 41.N°36 : Ranchon, G. La subversion humoristique des normes socio-sexuelles dans trois activités de FLE
- 42.N°37 : André, V. FLEURON : origines, démarches et perspectives
- 43.N°37-1 : Messier, G., Lafontaine, L. Un portrait des stratégies pédagogiques employées par des formateurs universitaires pour enseigner les genres écrits produits par des étudiants en formation à l'enseignement au Québec
- 44.N°38-1 : Huver, E. Peut-on (encore) penser à partir du CECR ? Perspectives critiques sur la version amplifiée
- 45.N°38-2 : Privas-Bréauté, V. Le jeu et les techniques dramatiques en classe de langue : un levier pour la co-construction des compétences professionnelles
- 46.N°39 : Moreau, A., Dumais, C., Nolin, R., Villeneuve-Lapointe, M., Stanké, B. Enseigner l'orthographe lexicale et l'apprentissage sans erreur: expérience pilote d'un modèle d'enseignement en classe ordinaire du primaire par un enseignant expert
- 47.N°39-1 : Akin, S. La famille et la transmission de langue en contexte diasporique : le cas du kurde
- 48.N°40-1 : Carette, E., Tassinari, G. Introduction
- 49.N°40-2 : Gadbois, M., Vergues, M. Langue, culture et identité : une approche structurée et innovante pour l'intégration de la culture en classe de langue seconde
- 50.N°41-1 : Alaoui, D. Le journal interculturel critique : du questionnement au changement

Bibliographie

- Ratinaud, P. (2009). IRAMUTEQ : Interface de R pour les Analyses Multidimensionnelles de Textes et de Questionnaires. [<http://www.iramuteq.org>.]
- Reinert, M. (1983). Un programme de classification descendante hiérarchique ; application à l'analyse lexicale par contexte. *Cahiers de l'Analyse des Données*, 8, 183-198. [http://www.numdam.org/item?id=CAD_1983__8_2_187_0]
- Reinert, M (1986). Un logiciel d'analyse lexicale (ALCESTE). *Cahiers de l'Analyse des Données*, 4, 471-484. [http://www.numdam.org/item/CAD_1986__11_4_471_0/]
- Reinert, M. (1990). Une méthode de classification des énoncés d'un corpus présenté à l'aide d'une application. *Cahiers de l'Analyse des Données*, 15(1), 21-36 [http://www.numdam.org/item/CAD_1990__15_1_21_0/]