



HAL
open science

Corpus et dictionnaires : prise de conscience linguistique chez des apprenants Lansad

Corinne Landure

► To cite this version:

Corinne Landure. Corpus et dictionnaires : prise de conscience linguistique chez des apprenants Lansad. Bulletin suisse de Linguistique appliquée, 2013, Data-driven approaches to language learning: theories, methods, applications and perspectives, 97 (spécial), pp.59-76. halshs-01246741

HAL Id: halshs-01246741

<https://shs.hal.science/halshs-01246741>

Submitted on 22 Dec 2015

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

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

2013. Corinne Landure. Corpus et dictionnaires : prise de conscience linguistique chez des apprenants LANSAD. In H. Tyne (Ed), *Bulletin Suisse de Linguistique Appliquée*, Numéro Spécial 97, *Data-driven approaches to language learning: theories, methods, applications and perspectives*. p.59-76.

Résumé

Les dictionnaires bilingue et unilingue sont principalement utilisés par les étudiants comme ressources de références et particulièrement appréciés pour leur apparente simplicité d'utilisation et leur efficacité, notamment l'immédiateté des résultats. Pourtant et peut-être pour ces mêmes raisons, ils sont bien souvent sous-exploités par les utilisateurs. Avec la démocratisation de technologies innovantes, la consultation sur corpus offre de nouvelles possibilités et perspectives aux apprenants d'une langue seconde (L2). De nombreuses études témoignent de l'importance du travail sur corpus tant au niveau du développement de l'autonomie que de la conscience linguistique des apprenants. Partant de ces constats, le présent article décrit une étude comparant l'utilisation du dictionnaire bilingue, unilingue et d'un corpus générique pour repérer et comprendre certains faits de langue tels que les noms indénombrables en anglais. Cette expérience s'est déroulée dans un Institut Universitaire de Technologie en France avec la participation d'étudiants en première année de Techniques de Commercialisation. Cette mise en œuvre expérimentale évalue la conscience langagière des étudiants en fonction de l'utilisation de ces trois ressources. Les résultats de cette étude nous engagent à mener une réflexion critique des moyens à développer pour des pratiques effectives d'observation et d'intégration de faits de langues par des apprenants L2.

Abstract

Bilingual and monolingual dictionaries are mainly used by students as reference resources and are generally appreciated for their apparent simplicity of use and their efficiency, in particular concerning the immediacy of the results. Nevertheless, and maybe for these same reasons, they are very often underexploited by the users. With the democratization of innovative technologies, the consultation of corpora offers new possibilities and perspectives to second language learners. Numerous studies have shown the importance of corpus exploitation both at the level of the development of autonomy and of the linguistic awareness of the learners. Based on these reports, the present article describes a comparative study using a bilingual dictionary, a monolingual dictionary and a generic corpus to identify and understand certain items such as uncountable nouns in English. The experiment took place in a University Institute of Technology in France with the cooperation of students in the first year of Commercial Marketing. This work looks at the students' linguistic awareness according to the use of these three resources. The results of the study invite critical reflection on the means to be developed for effective practices of observation and intake of language by ESP learners.

Keywords: Data-driven learning, dictionaries, linguistic awareness, language events, corpus

1. Introduction

Quels que soient les formes et les moyens d'apprentissage utilisés, ceux-ci doivent correspondre aux besoins et aux spécificités des apprenants afin de faciliter leur apprentissage. Pour autant "la décision d'apprendre ne peut être prise que par l'apprenant" (Bogaards 1988: 100). Mais que signifie 'apprendre' pour un apprenant? Le fait de se poser la question n'est pas neutre. La forme de l'apprentissage et les moyens mis à disposition reposent bien souvent sur des choix faits par l'enseignant ou dépend de directives institutionnelles qui peuvent ne pas convenir à tous les étudiants pour différentes raisons. Pour l'apprentissage du lexique par exemple, l'apprenant dispose de manuels spécialisés qui concentrent le vocabulaire à acquérir sous forme de listes

thématiques. Il peut également avoir recours à la lecture de divers ouvrages où les unités lexicales sont diluées dans le texte (Tournier 2004). Pourtant, comme le fait remarquer Cobb (1999), si ces manières d'apprendre présentent des avantages, elles ne sont pas garanties d'un apprentissage efficace. Dans le premier cas, l'apprentissage peut aboutir à des connaissances incomplètes et dans le second, l'acquisition par la lecture s'avère lente. Bien souvent, dans ces deux situations, les étudiants s'aident d'un dictionnaire bilingue, même parfois d'un dictionnaire unilingue, et comme le souligne Galisson (1991: 17) "[les dictionnaires bilingues] se portent plutôt bien, faisant l'objet d'une prédilection des apprenants". Surcouf (2010: sn), quant à lui, remarque:

Qu'il soit plébiscité par les apprenants, et plus ou moins discrédité par les enseignants, le dictionnaire bilingue, en raison même de sa popularité, modèle incontestablement l'apprentissage de la langue étrangère.

Pour leur part, Schmitt *et al.* (1997) notent que les dictionnaires bilingues sont parmi les premiers outils utilisés par les étudiants quand ils lisent dans une langue étrangère. Le fait est que le dictionnaire bilingue répond de façon satisfaisante à un besoin ponctuel, notamment pour la recherche de la signification d'une unité lexicale, et les étudiants apprécient autant l'apparente facilité de son utilisation que l'immédiateté du résultat ainsi obtenu. Pourtant, comme le font remarquer Schmitt *et al.* (1997) dans leur étude, bien que le dictionnaire soit le premier choix de consultation, il n'est pas cité comme une ressource efficace d'apprentissage. Ce constat corrobore les résultats d'une enquête réalisée auprès de 160 de nos étudiants: 90% utilisent le dictionnaire uniquement pour la recherche d'une traduction ou d'une définition et 72% déclarent n'avoir jamais reçu de formation théorique et pratique concernant l'utilisation d'un dictionnaire bilingue durant leur scolarité (ce taux atteint 89% pour un dictionnaire unilingue). Le dictionnaire serait avant tout utilisé comme ressource de référence plutôt que ressource d'apprentissage pour preuve le faible taux de rétention lors de l'utilisation de dictionnaire à des fins d'apprentissage (Landure 2011). C'est aussi un support 'historique' (ce qu'utilisaient les enseignants eux-mêmes ainsi que leurs enseignants...) transmis de génération en génération. Dans ce cas, quelles ressources l'apprenant a-t-il à disposition pour repérer et traiter efficacement les phénomènes linguistiques?

Cette étude a un double objectif: tout d'abord évaluer la prise de conscience linguistique des apprenants dans leur façon d'appréhender les faits de langue (dans notre étude il s'agit des noms indénumérables) selon qu'ils utilisent un dictionnaire bilingue, unilingue ou un corpus; ensuite déterminer dans quelle mesure une formation préalable à l'utilisation de ces trois ressources influe sur la prise de conscience linguistique. Dans un premier temps, nous définirons ce que sont les indénumérables puis nous présenterons les corpus avant de décrire nos deux expérimentations.

Nous supposons qu'un apprentissage efficace se reconnaît par une connaissance des éléments lexicaux, non pas uniquement en tant qu'unité de sens isolé mais en tant qu'unité de sens aux spécificités propres, régissant l'unité lexicale ainsi que sa dimension morphosyntaxique. Certaines distinctions linguistiques entre langue source et langue seconde (L2) peuvent représenter des écueils lors de l'apprentissage; la traduction littérale des items permet rarement d'en cerner les spécificités linguistiques, même si le mot fait sens dans le contexte. Il en est ainsi des indénumérables en anglais qui, s'ils ne comportent pas de difficultés sémantiques particulières, constituent des obstacles difficilement repérables par l'apprenant, par exemple: *luggage* se traduit par un pluriel en français, ne prend pas de 's' à la fin, ne s'utilise jamais avec un article indéfini et le verbe qui le suit est au singulier. Ces difficultés concernent l'item dans sa morphosyntaxe. Le Priault, auteur de la *Grammaire progressive de l'anglais*, définit ainsi les indénumérables (1996: 8):

La distinction dénombrable/indénombrable a des conséquences grammaticales importantes, en particulier sur l'accord verbal, ainsi que sur le phénomène de détermination (les articles et la quantité).

Il souligne le caractère complexe des indénombrables pour des apprenants français en l'occurrence et conclut "que chaque langue découpe différemment la réalité" (1996: 11).

2. 'Data-driven learning' et apprentissage expérientiel

Dans le cadre de notre travail de recherche, nous procédons à une étude comparative impliquant la consultation de trois ressources pour évaluer la prise de conscience linguistique d'apprenants d'une L2: un dictionnaire bilingue anglais-français, *Le Grand Robert & Collins* (GR&C), un dictionnaire unilingue anglais, *Oxford Advanced Learner's Dictionary of Current English* (OALD), et un corpus générique, le British National Corpus (BNC). Si les deux premières ressources sont familières aux étudiants, l'exploitation de données de corpus est une nouveauté pour nos apprenants. En effet, la forme d'apprentissage sur corpus (ASC) baptisée par Johns ([1990] 1991) '*data driven learning*' est peu connue ou reconnue dans les sphères enseignantes traditionnelles (Boulton 2008b; Lapaire 2009). Pourtant, de nombreuses études recensées par Boulton (2008a) soulignent l'importance d'une telle approche pour améliorer la prise de conscience linguistique des apprenants (cf. O'Sullivan 2007; voir aussi Chambers ce volume). L'ASC offre aux apprenants la possibilité de gérer et de déterminer leur propre apprentissage, d'appréhender et d'organiser de manière inductive leurs connaissances (Boulton 2008a). Aston (2001) souligne également l'importance d'une telle approche pour développer la prise de conscience des apprenants et pour améliorer leur compétence linguistique en découvrant par eux-mêmes des faits de langue qu'ils n'avaient pas forcément remarqués avant cela. Boulton (2010: sn) fait remarquer:

[qu'] avec des outils appropriés, l'approche permet l'étude des usages habituels et probables de façon empirique et non à partir de simples intuitions ou de quelques exemples isolés.

Afin d'introduire la notion de prise de conscience linguistique chez des apprenants LANSAD ('LANGues pour Spécialistes d'Autres Disciplines') et d'en identifier les postures, notre étude s'appuie sur le célèbre modèle d'apprentissage expérientiel de Kolb (1984). Bien qu'il existe d'autres cadres de référence traitant ce domaine, le modèle de Kolb, largement employé dans la formation aux adultes, se trouve être le mieux adapté à notre public composé d'étudiants âgés de 19 à 21 ans, en première année de Diplôme Universitaire de Technologie (DUT) en Techniques de Commercialisation. Kolb, bien que ne traitant pas spécifiquement l'apprentissage des langues, considère l'apprentissage expérientiel en quatre étapes: l'expérience concrète, l'observation réfléchie, la conceptualisation abstraite et l'expérimentation. Ces différents stades mettent en opposition d'une part le concret et l'abstrait au niveau de la 'préhension' c'est-à-dire de l'action, et d'autre part le réflexif et l'actif au niveau de la 'transformation', c'est-à-dire de la réflexion (Chevrier & Charbonneau 2000; Wittorski 2001). Les associations de ces formules aboutissent à quatre différents styles d'apprentissage (cf. Boulton 2010) sur lesquels nous ne nous attarderons pas dans cette étude. Notre expérience privilégie plus particulièrement l'étude et l'analyse comparative des procédés et attitudes cognitives de nos apprenants à travers les quatre étapes de l'apprentissage expérientiel telles que définies par Kolb mais transposées et adaptées par nos soins dans le cadre de la consultation de dictionnaires bilingue, unilingue et d'un corpus pour l'appréhension de faits de langue. En rapprochant les quatre étapes de l'apprentissage expérientiel de Kolb (tableau 1 ci-dessous), on

s'aperçoit qu'elles correspondent aux principales notions de l'apprentissage sur corpus, telles que décrites par Johns (1986: 159):

By concentrating and making it easy to compare the contexts within which a particular item occurs, [a concordance] organizes data in a way that encourages and facilitates inference and generalization.

Au mode exploratoire, c'est-à-dire l'expérience concrète où l'apprenant est directement acteur voire chercheur, correspond l'exploration des données sur corpus avec l'utilisation d'un concordancier par l'apprenant. L'observation réfléchie des données se rapporte à l'observation et à la comparaison des données en contexte et co-texte qui aboutissent à une réflexion que l'on peut éventuellement considérer comme étant une des phases essentielles de la prise de conscience linguistique de l'apprenant. Cette réflexion fondée sur l'observation va permettre de mettre en place une conceptualisation abstraite (repères, notions, hypothèses) et ainsi contribuer à une réflexion inductive qui participe également à une autre phase majeure de la prise de conscience langagière. Enfin, le mode vérificateur s'inscrivant dans l'expérimentation conduit à une généralisation des concepts en les confrontant à des applications actives.

Apprentissage expérientiel (Kolb 1984)	Data-driven learning (Johns 1986)
Mode exploratoire expérience concrète	Exploration corpus & 'hands-on' (<i>concordance</i>)
Mode réfléchi observation réfléchie	Observation & comparaison (<i>organize data, compare contexts</i>)
Mode abstraitif conceptualisation abstraite	Induction (<i>inference</i>)
Mode vérificateur expérimentation	Généralisation & application (<i>generalization</i>)

Tableau 1. Comparaison entre apprentissage expérientiel et 'data-driven learning'

3. Méthode

Comme nous l'avons déjà mentionné, l'objectif principal de cette étude est de situer et d'évaluer la prise de conscience linguistique d'apprenants d'une L2 selon qu'ils utilisent un dictionnaire bilingue, unilingue ou un corpus générique dans leur capacité à repérer et à traiter un fait de langue. Pour ce faire, nous avons procédé à deux expériences sur le terrain à une année d'intervalle. La première consiste en la participation de 78 étudiants en première année de DUT de Techniques de Commercialisation qui n'ont reçu, à dessein, qu'une formation théorique brève. Cette présentation d'une quinzaine de minutes, sous forme de diapositives, est assurée par l'enseignante pour chacune des quatre classes composées de vingt étudiants environ et porte sur la manière d'utiliser des dictionnaires et de consulter le corpus. La deuxième expérience concerne 82 nouveaux étudiants en première année de DUT de Techniques de Commercialisation qui ont bénéficié d'une formation préalable de quelques heures sur l'utilisation des dictionnaires bilingue, unilingue et d'un corpus. Le niveau d'anglais de ces deux groupes est moyen et correspond au niveau B1 du Cadre Européen Commun de Référence pour les Langues (CECRL – Conseil de l'Europe 2001), comme l'attestent les résultats à un test informel du TOEIC administré au début du semestre pour chaque année respective. Ces 160 étudiants sont des francophones natifs, à l'exception de quatre Chinois et d'un Marocain, et ont un parcours scolaire très différent: ils sont titulaires soit d'un Bac professionnel, soit d'un Bac technique ou encore général ce qui peut expliquer l'hétérogénéité de leur niveau en anglais. D'origine modeste, ces étudiants n'ont jamais séjourné plus de deux semaines dans un pays anglophone et sont tous domiciliés dans la même région.

Ne pouvant avoir accès à la salle informatique, les étudiants travaillent sur supports papier comportant les consignes qui se rapportent aux tâches à effectuer à partir des définitions d'items extraites du GR&C bilingue pour le premier couple d'items (*information / inquiry*), du dictionnaire unilingue OALD pour le deuxième couple (*advice / suggestion*), et du BNC dans lequel nous avons sélectionné quinze phrases pour l'item *luggage* et quinze autres phrases pour *suitcase*. Ces mots ont été choisis par nos soins pour plusieurs raisons. Tout d'abord, parce qu'ils avaient été vus précédemment en cours dans divers contextes; ils n'étaient donc pas totalement inconnus de nos étudiants au niveau sémantique, ce qui n'est pas négligeable puisque nous voulions mettre l'accent sur la particularité morphosyntaxique d'une unité lexicale composant une paire plutôt que sur son sens. La raison de ce choix s'explique par un constat. Nous avons remarqué lors des cours précédents ces expériences que les étudiants s'attachaient à retenir davantage le sens des mots (cf. Kerr ce volume) au détriment de ses dimensions morphosyntaxiques malgré les efforts et l'insistance de l'enseignante. D'autre part, les mots composant chaque paire ont un sens proche bien qu'ils ne soient pas synonymes mais surtout l'un d'entre eux est indénombrable, ce qui va nous permettre, plus particulièrement dans l'exercice lacunaire, d'évaluer la prise de conscience des étudiants sur ce fait de langue et de la dimension morphosyntaxique de l'unité lexicale. L'impact du sens dans l'exercice lacunaire est très réduit puisque cet exercice a été conçu de telle sorte qu'il ne prend principalement en compte que l'unité lexicale dans sa fonction morphosyntaxique et non sémantique.

Afin de pouvoir déterminer les attitudes et comportements des étudiants lors de leurs diverses consultations, nous leur avons demandé de procéder à quatre tâches. La première activité consiste à consulter un couple de mots dans une ressource de référence et à relever le temps passé pour chaque consultation, grâce à une horloge numérique provenant d'un site Internet et projetée au tableau (<http://www.horlogeparlante.com>). La deuxième activité concerne la rédaction d'une phrase dans la langue cible intégrant l'item consulté afin de savoir si l'étudiant a repéré dans la ressource consultée le fait de langue et la dimension morphosyntaxique du mot recherché. La troisième porte sur la traduction dans la langue source de leur production qui n'a d'autre intention que d'évaluer les interférences possibles avec le français et ainsi de confirmer ou d'infirmer si l'apprenant en a saisi et assimilé les spécificités dans la langue cible mais également comment il les traduit dans la langue source. Enfin, la quatrième activité consiste à remplir un exercice à trous pour chaque couple d'items; cet exercice a été conçu par nos soins à partir de lignes de concordance provenant du BNC et de telle sorte que les choix à opérer par les apprenants portent davantage sur une réflexion ayant trait à la dimension morphosyntaxique de l'unité lexicale plutôt que sur sa fonction sémantique. Chacune des trois activités (consultation, rédaction et traduction) va permettre d'apprécier les cheminements et comportements des apprenants dans le traitement des données provenant des trois ressources dans la phase active d'exploration puis d'observation et enfin d'induction aboutissant à une production écrite. Ainsi, notre étude va s'attacher à évaluer l'item dans sa spécificité propre en tant que fait de langue, sa traduction dans le contexte et ses répercussions sur le co-texte selon la ressource consultée. La phase vérificatoire consiste en un exercice à trous pour évaluer la prise en compte du co-texte par les apprenants lors du remplissage et ces résultats seront à rapprocher de ceux provenant de l'activité de production.

4. Résultats

4.1 Première étude comparative de la prise de conscience d'un fait de langue selon la ressource consultée

Le tableau 2 (voir ci-dessous) nous indique que tous les étudiants ont consulté les différentes ressources puisqu'ils ont tous indiqué sur une grille le temps passé pour la consultation de

chaque item. Nous remarquons qu'ils ont passé plus de temps sur les quinze lignes de concordance provenant du corpus, deux à trois fois moins de temps pour le dictionnaire unilingue et encore moitié moins pour le dictionnaire bilingue. Il est intéressant de noter la différence de participation dans la phase productive; elle s'élève à 76% après consultation sur corpus, arrive à 95% après consultation du dictionnaire unilingue et atteint les 100% après consultation du GR&C bilingue. Concernant ces activités, aucune limite de temps n'a été fixée, et tous les étudiants ont déclaré avoir terminé ces trois tâches au bout d'une heure. Ceux qui ne sont pas parvenus à tout faire (24%) n'ont pas évoqué un manque de temps mais plutôt une incompréhension ou une incapacité à exécuter l'activité demandée. Ils mentionnent un problème de traduction du mot, d'incompréhension de l'ensemble des occurrences ou d'un manque d'inspiration.

Ressources & items	Consultation			Rédaction	Traduction
	nombre %	temps moyen par apprenant	écart-type	nombre%	nombre%
Dictionnaire bilingue					
<i>information</i>	100%	34 sec	07 sec	100%	100%
<i>inquiry</i>	100%	30 sec	08 sec	100%	100%
Dictionnaire unilingue					
<i>advice</i>	100%	58 sec	18 sec	96%	94%
<i>suggestion</i>	100%	50 sec	16 sec	94%	93%
Corpus					
<i>luggage</i>	100%	3 mn	1.58 mn	76%	75%
<i>suitcase</i>	100%	2 mn	1.30 mn	76%	76%

Tableau 2. Activités du 1^{er} groupe de 78 étudiants sans formation

Le tableau 3 renseigne quant aux indénumbrables consultés respectivement dans les différentes ressources. L'unité lexicale *information* a été traitée complètement par les étudiants lors des activités de consultation, de rédaction et de traduction. L'item dans le contexte n'a pas posé de problème puisqu'en plus d'être un mot transparent, les étudiants ont pu s'assurer du sens dans le dictionnaire bilingue.

Ressources utilisées	Items consultés	0 réponse	Contexte approprié	Traduction appropriée	Forme de l'item appr	Co-texte approprié
Dic biling	<i>information</i>	0%	100%	86%	81%	83%
Dic uniling	<i>advice</i>	4%	76%	28%	95%	56%
Corpus	<i>luggage</i>	24%	88%	54%	88%	76%

Tableau 3. Analyse des résultats des activités du 1^{er} groupe

La traduction a été correcte (ou acceptable) dans 86% des cas, ce qui peut paraître peu mais qui s'explique par le fait que les étudiants l'ont traduit souvent par un singulier en français alors qu'il s'agissait d'un pluriel dans le contexte. La forme de l'unité lexicale est correcte à hauteur de 81%, la différence provient essentiellement de l'ajout d'un 's'. On relève peu d'erreurs au niveau du co-texte, celles-ci proviennent des quantifieurs (15%) ou de l'accord de la forme verbale (2%). Pour l'item *advice* consulté dans le OALD, 4% des participants n'ont pas exécuté les tâches demandées. 76% l'ont intégré dans un contexte approprié et seulement 28% ont donné une traduction adéquate; on relève un important nombre de traductions au singulier à la place d'un pluriel dans le contexte, beaucoup d'approximations ('opinion', 'avis'), de nombreux contre-

sens ('ordre', 'ordonnance', 'notice', 'information') ainsi qu'un certain nombre de changements de classe grammaticale, par exemple: *to advice* (*sic*). La forme de l'item a été plutôt bien respectée, on ne relève que 5% de 's' à la fin de *advice*. Le co-texte a posé un certain nombre de problèmes, on note entre autres la présence du déterminant *an* et la confusion entre le verbe et le substantif (*I / you advice* (*sic*) et *advice*). Concernant l'item *luggage* que les étudiants devaient consulter dans les quinze lignes de concordance provenant du BNC, on relève une défection de presque un quart des étudiants concernant les activités de rédaction et de traduction. Cela pourrait s'expliquer par notre choix intentionnel d'avoir fait peu de formation sur l'exploitation de corpus, par le manque de motivation des apprenants ou de leur incompréhension face cette nouvelle ressource. A ce stade, nous ne pouvons qu'émettre des hypothèses.

Le mot *luggage* dans le contexte (88%) n'a pas posé trop de problèmes aux étudiants qui ont plus souffert au niveau de la traduction. La plupart l'ont traduit par 'valise', 'objet', 'affaires', d'autres ont pensé qu'il pouvait s'agir d'une marque de voiture, d'une commode. 12% ont ajouté un 's' à l'item et 76% ont respecté le co-texte, les erreurs se situant au niveau des quantifieurs (57%) et de l'accord de la forme verbale (43%).

Ces résultats confirment en partie l'efficacité du dictionnaire bilingue pour la traduction de l'item et son sens dans le contexte. En revanche, la forme de l'item a été moins prise en compte au niveau de sa spécificité linguistique par rapport aux résultats obtenus avec les autres ressources. Les étudiants se focalisent certainement plus sur le sens du mot et sont moins attentifs à ces spécificités. Pourtant, les répercussions du fait de langue sur le co-texte ont été bien traitées; afin d'être en mesure de confirmer ou d'infirmer ces résultats, nous envisageons de les rapprocher de ceux provenant de l'exercice à trous portant sur le co-texte. Si le dictionnaire unilingue a généré de gros problèmes de traduction, nous notons un résultat mitigé pour le co-texte que les étudiants ont, semble-t-il, ignoré alors que la forme de l'item a été dans l'ensemble bien traitée. Mise à part la traduction, les résultats provenant de la consultation sur corpus sont plutôt satisfaisants, notamment au niveau de la forme de l'item en tant que fait de langue et des répercussions sur son co-texte. En revanche, on note une déstabilisation conséquente des étudiants puisqu'un quart d'entre eux n'ont pas participé aux activités de production ce que nous allons tenter de comprendre et d'expliquer dans la partie suivante.

La particularité de l'exercice à trous conçu à partir du BNC, est de confirmer ou d'infirmer les résultats obtenus lors de la phase de production écrite. Cet exercice concerne plus particulièrement les répercussions du fait de langue dans le co-texte. Nous estimons que s'il y a bien prise de conscience linguistique, les manières de procéder de l'apprenant se révéleront dans la façon de remplir les blancs en fonction du co-texte. Cette prise de conscience a pu commencer lors de la phase d'exploration, se poursuivre lors de l'observation pour aboutir à une réflexion inductive. Les résultats de ce cheminement vers une prise de conscience linguistique se traduisent dans la manière d'aborder et d'appréhender une activité, dans notre cas dans la façon de traiter un fait de langue dans son co-texte.

L'exercice lacunaire comporte 24 phrases extraites du BNC. Les huit premières doivent être complétées avec *information* et / ou *inquiry*, les huit suivantes avec *advice* et / ou *suggestion* et les huit dernières avec *luggage* et / ou *suitcase* dans leur forme singulier ou pluriel. Les étudiants disposent de 45 minutes pour cette activité et aucun document n'est autorisé.

Le tableau 4 présente les ressources utilisées pour chaque couple d'items. On note que les résultats provenant de la consultation des dictionnaires pour l'appréhension d'un fait de langue sur son co-texte atteignent respectivement 41% pour le dictionnaire bilingue et 40% pour l'OALD alors que la consultation sur corpus approche les 34%. Pourtant lors de l'activité de production, les résultats du co-texte suite à la consultation du dictionnaire bilingue atteignaient le double (83%), 56% après consultation du dictionnaire unilingue et 76% après la consultation sur corpus.

Ressources	items	approprié au co-texte 1 ^{er} groupe
Dictionnaire bilingue	<i>information and / or inquiry</i>	41%
Dictionnaire unilingue	<i>advice and / or suggestion</i>	40%
Corpus	<i>luggage and / or suitcase</i>	34%

Tableau 4. Résultats de l'exercice lacunaire pour le co-texte

Notre hypothèse de départ supposait qu'un apprentissage efficace allait se reconnaître par une connaissance des items non pas uniquement en tant qu'unité de sens isolé mais en tant qu'unité de sens aux spécificités propres, régissant l'unité lexicale ainsi que son entourage proche. Notre étude partant de cette hypothèse consistait à comparer trois ressources de référence pour évaluer la prise de conscience linguistique d'apprenants d'une L2. Les résultats sont surprenants voire décevants si nous nous fondons sur des études antérieures prêtant à l'ASC des qualités notamment pour l'amélioration de la prise de conscience linguistique et le développement de l'autonomie des apprenants. En parlant de l'autonomie, Holec (1979: 3) souligne que:

Cette capacité n'est pas innée, elle doit s'acquérir, soit de manière 'naturelle', soit (c'est le cas le plus fréquent) par un apprentissage formel, c'est-à-dire systématique et réfléchi.

Et si justement, comme l'autonomie, la prise de conscience linguistique n'était pas innée et nécessitait une formation? C'est ce que nous avons voulu vérifier avec la deuxième expérience se passant dans les mêmes conditions à un an d'intervalle. Les 82 participants sont des étudiants en première année de DUT en Techniques de Commercialisation. La seule différence réside dans une formation de deux heures (au lieu de 15 minutes pour la première expérience) dont ils ont bénéficié avant cette expérience pour l'utilisation des dictionnaires et de la consultation de corpus, et ciblant, bien entendu, d'autres phénomènes langagiers que les indénombrables.

4.2. Deuxième étude comparative de la prise de conscience d'un fait de langue selon la ressource consultée

Si nous comparons les résultats du tableau 5 aux résultats obtenus une année auparavant (tableau 2) avec des étudiants n'ayant bénéficié d'aucune formation préalable, on remarque que le temps moyen par étudiant pour les consultations du dictionnaire bilingue et unilingue varie entre 1 et 8 secondes supplémentaires, en revanche le temps passé pour la consultation sur corpus a triplé. On note également une augmentation sensible de participation aux activités de production après la consultation sur corpus.

ressources		consultation		rédaction	traduction
& items	nombre %	temps moyen / apprenant	écart-type	nombre%	nombre%
Dictionnaire bilingue					
<i>information</i>	100%	35 sec	05 sec	100%	100%
<i>inquiry</i>	100%	32 sec	06 sec	100%	100%
Dictionnaire unilingue					
<i>advice</i>	100%	65 sec	13 sec	100%	100%
<i>suggestion</i>	100%	58 sec	15 sec	98%	98%
Corpus					
<i>luggage</i>	100%	8 mn	1.02 mn	91%	89%
<i>suitcase</i>	100%	7 mn	0.50 mn	95%	95%

Tableau 5. Activités du 2^e groupe de 82 étudiants avec formation

Ressources utilisées	Items consultés	0 réponse	Contexte approprié	Traduction appropriée	Forme de l'item appr	co-texte approprié
Dic biling	<i>information</i>	0%	100%	96%	91%	85%
Dic uniling	<i>advice</i>	1%	72%	30%	90%	66%
Corpus	<i>luggage</i>	8%	91%	65%	95%	82%

Tableau 6. Analyse des résultats des activités du 2^e groupe

Le tableau 6 montre une augmentation des résultats dans toutes les rubriques, ce que nous pourrions imputer à la formation dont les étudiants ont bénéficié. En revanche, il est intéressant de comparer ces résultats avec ceux de l'exercice à trous (tableau 7). On remarque à nouveau une différence non négligeable entre l'activité de production et l'exercice à trous au niveau des résultats du co-texte qui pourrait s'expliquer par la teneur de l'exercice lui-même, trop complexe peut-être, ou au contraire par une minimalisation de la production écrite qui intègre l'item dans une phrase trop simple ou qui s'inspire d'une phrase trouvée dans les ressources consultées. En revanche, le tableau 7 indique une augmentation sensible de résultats appropriés au co-texte et cela quelles que soient les ressources utilisées, avec un résultat presque doublé pour l'utilisation de corpus.

Ressources	items	approprié au co-texte (1 ^{er} groupe)	approprié au co-texte (2 ^e groupe)
Dictionnaire bilingue	<i>information and / or inquiry</i>	41%	52%
Dictionnaire unilingue	<i>advice and / or suggestion</i>	40%	48%
Corpus	<i>luggage and / or suitcase</i>	34%	63%

Tableau 7. Etude comparative de l'exercice à trous

A ce stade, peut-on avancer que la formation pour l'utilisation de ressources a augmenté la capacité de prise de conscience linguistique d'apprenants d'une L2? Peut-on avancer que la consultation de corpus améliore efficacement la prise de conscience et donne les moyens à l'apprenant de développer sa sensibilité linguistique en fonction de ses besoins? Les résultats montrent que la formation préalable portant sur l'utilisation des corpus est plus efficace pour une meilleure prise de conscience linguistique, alors que la formation préalable portant sur l'utilisation des dictionnaires bilingue et unilingue n'a pas vraiment amélioré la prise de conscience linguistique des apprenants.

1. Discussion

Afin de répondre à ces deux questions, nous nous appuyons sur les travaux de Wittorski (2001) qui, s'inspirant des recherches de Foerster (1960), rapporte qu'un individu se fabrique une représentation de la situation en fonction de ses expériences antérieures, et que par conséquent les modèles d'action mobilisés dépendront de ses expériences personnelles. Puis, Wittorski se fonde sur les travaux de Serre (1992) pour illustrer les processus qui vont guider l'action et que nous avons représentés schématiquement (tableau 8):

Modèle d'action par rapport aux expériences passées	Hypothèses émises en fonction de la situation	Action engagée suivant les hypothèses	Vérification de l'(in)efficacité de l'action
---	---	---------------------------------------	--

Tableau 8. Schéma inspiré des travaux de Serre pour les processus de mobilisation

Ces processus s'appliquent à notre deuxième expérience pour laquelle un modèle d'action ou des modèles d'actions ont été mis en place par les apprenants et qui découlent directement des expériences passées; dans notre cas, il s'agit de la formation dont ils ont bénéficié dans le cadre de la consultation des ressources. Des hypothèses ont été alors émises en fonction de la situation; en ce qui nous concerne, des hypothèses ont été exprimées par rapport à des situations de consultation intéressant diverses ressources de référence. Ces dernières ont contribué à engager une action en fonction des expériences antérieures et postulats émis en amont; dans notre cas, cela consiste en la production écrite faisant suite aux consultations des trois ressources. Enfin, l'exercice lacunaire a permis de vérifier l'efficacité ou l'inefficacité de la consultation des trois ressources utilisées pour la compréhension des faits de langue et ses répercussions sur le co-texte. Les résultats de l'exercice lacunaire montrent qu'une formation préalable ciblant l'utilisation de dictionnaires bilingue et unilingue n'a que peu d'impact sur la prise de conscience alors qu'une formation impliquant l'utilisation des corpus obtient une meilleure prise de conscience linguistique. Les apprenants peuvent ainsi reproduire des modèles d'action à mettre en œuvre selon les situations rencontrées, émettre des hypothèses par rapport aux données et procéder à une réflexion inductive surtout dans le cas de la consultation de corpus. Effectivement, on peut avancer que la formation a plus profité en termes d'efficacité à la consultation de corpus, peut-être du fait de sa nouveauté et de la manière de l'aborder. Les corpus offrent aux apprenants la possibilité d'exploiter des données, d'observer le co-texte et le contexte et permet d'améliorer la sensibilité linguistique en fonction des besoins. Alors que les résultats moyens provenant de l'utilisation des dictionnaires accentuent leur efficacité médiocre dans la prise de conscience linguistique, Wittorski, citant les travaux d'Argyris *et al.* (1987), évoque l'inefficacité de l'action engagée en l'expliquant par la non-capacité de l'acteur à remettre en cause sa perception de la situation et à reproduire de manière récurrente des modèles d'action inefficaces. Deux solutions d'apprentissage sont alors proposées pour y remédier: la première consiste à adopter une stratégie comportementale différente pouvant être initiée par une autre personne mais sans remettre en question le cadre de la situation que les auteurs ont appelé "*single loop learning*"; l'autre alternative intervient au niveau du cadre de perception qui lui permet d'adopter de nouveaux modèles d'action reposant sur le questionnement et l'analyse de la situation. Ce deuxième choix, "*double loop learning*" serait un véritable changement en termes d'apprentissage, ce que Schön (1983) appelle "*reshaping the situation*". C'est effectivement ce que nous avons observé dans le cadre de notre expérience. Les dictionnaires sont des ressources très appréciées mais très mal utilisées et il est quelquefois plus difficile de changer de mauvaises habitudes, fossilisées par des années de pratique que d'apprendre à en utiliser de nouvelles efficacement. De plus, Surcouf (2010: sn), en évoquant les dictionnaires, déclare:

[qu'] en traduisant, l'apprenant se prive, entre autres, de la possibilité de construire des réseaux d'associations morphologiques, lexicaux et sémantiques en L2. Ainsi, en vient-il en partie à inhiber, à son insu, l'un des procédés essentiels de l'apprentissage de toute langue: l'organisation des savoirs en réseaux permettant un traitement regroupé des informations et une rétention en mémoire plus structurée, plus aisée et donc plus efficace.

2. Conclusion

Cette étude semble démontrer que la prise de conscience linguistique peut être favorisée par la consultation de corpus qui permet de repérer et d'appréhender des phénomènes linguistiques de manière efficace. L'ASC s'apparente dans ses grandes lignes à l'apprentissage expérientiel (Kolb 1984) avec lequel il partage des notions essentielles. Nous pensons qu'une prise de conscience linguistique adopte la même ligne que l'apprentissage expérientiel et partage des attitudes et des comportements semblables. Pour autant, la prise de conscience linguistique n'est pas innée et nécessite une formation afin de pouvoir aborder différentes situations en se créant des modèles d'action (Serre 1992). Un aspect majeur de la prise de conscience s'inscrit dans la manière dont nous appréhendons l'apprentissage d'une L2. L'explication réside en partie dans les modalités cognitives de construction et de réalisation associées à l'apprentissage ou au savoir apprendre expérientiel. L'apprenant organise ses connaissances par rapport à la représentation de la tâche et également par rapport aux ressources utilisées, il adopte consciemment ou pas un cheminement qui a des répercussions sur sa façon d'apprendre. Il modéliserait mentalement un parcours d'apprentissage qui agira sur sa façon d'appréhender ses connaissances acquises ou en cours d'acquisition. Les résultats linguistiques de la deuxième expérience indiquent que l'exploitation de données sur corpus augmente sensiblement la capacité de prise de conscience linguistique d'apprenants d'une L2 à repérer et à traiter des faits de langue au niveau de l'item et du co-texte. Cependant, cette prise de conscience n'est efficace que si elle est précédée d'une formation linguistique. La comparaison des résultats avec d'autres ressources de référence telles que le dictionnaire bilingue et unilingue accentuent la propension à mener une réflexion inductive lors de la consultation de corpus et comme le soulignent Chevrier et Charbonneau (2000: sn) "le savoir ainsi constitué est réutilisable pour vivre de nouvelles expériences".

BIBLIOGRAPHIE

Aston, G. (2001). *Learning with corpora*. Houston: Athelstan.

Argyris, C., Putman, R. & Mac Lain Smith, D. (1987). *Action science*. San Francisco: Jossey Bass Publishers.

Bogaards, P. (1988). *Aptitude et affectivité dans l'apprentissage des langues étrangères*. Paris: Crédif-Didier.

Boulton, A., Château, A., Pereiro, M. & Azzam-Hannachi, R. (2008b). Learning to learn languages with ICT – but how? *CALL-EJ*. <http://www.tell.is.ritsumei.ac.jp/callejonline/journal/9-2/boulton.html> (10.02.2011).

Boulton, A. (2008a). Esprit de corpus: promouvoir l'exploitation de corpus en apprentissage des langues. *Texte et corpus*, 3, 37-46.

Boulton, A. (2008b). Looking for empirical evidence of DDL at lower levels. In: B. Lewandowska-Tomaszczyk (ed.), *Corpus linguistics, computer tools, and applications – state of the art* (pp.581-598). Francfort: Peter Lang.

Boulton, A. (2010). Consultation de corpus et styles d'apprentissage. *Cahiers de l'APLIUT*, 29, 1, 98-115. <http://apliut.revues.org/967> (24.02.2012).

BYU-BNC. *British national corpus*. <http://corpus.byu.edu/bnc/x.asp> (10.02.10).

Chevrier, J. & Charbonneau, B. (2000). Le savoir-apprendre expérientiel dans le contexte du modèle de David Kolb. *Revue des sciences de l'éducation*, 26, 2, 287-324. <http://id.erudit.org/iderudit/000124ar> (15.09.11).

Cobb, T. (1999). Breadth and depth of vocabulary acquisition with hands-on concordancing. *Computer Assisted Language Learning*, 12, 345-360.

Conseil de l'Europe (2001). *Cadre européen commun de référence pour les langues: apprendre, enseigner, évaluer*. Paris: Didier.

Foerster Von, H. (1960). On self-organizing systems and their environments. In: M. Yovits & S. Cameron (eds.), *Self-organizing systems*. Londres: Pergamon Press.

Galisson, R. (1991). *De la langue à la culture par les mots*. Paris: Clé international.
Grand Robert & Collins (2008). *Dictionnaire anglais-français*. Paris: Le Robert.

Holec, H. (1979). *Autonomie et apprentissage des langues étrangères*. Strasbourg: Conseil de l'Europe. Repris dans: B. André (éd.) (1989). *Autonomie et enseignement / apprentissage des langues étrangères*. Paris: Didier/Hatier].

Johns, T. (1986). Micro-concord: a language-learner's research tool. *System*, 14, 2, 151-162.

Johns, T. ([1990] 1991). From printout to handout: grammar and vocabulary teaching in the context of data-driven learning. *CALL Austria*, 10, 14-34. [Version remaniée: T. Johns & P. King (eds.), *Classroom concordancing*. *English Language Research Journal*, 4, 27-45.

Kolb, D. (1984). *Experiential learning: experience as the source of learning and development*. Englewood Cliffs: Prentice Hall.

Landure, C. (2011). Data-driven learning: apprendre et enseigner à contre-courant. *Mélanges CRAPEL*, 32, 162-178. http://www.atilf.fr/IMG/pdf/163-178_landure.pdf. (15.05.2013)

Lapaire, J. (2009). De la grammaire anglaise balado diffusée au portfolio grammatical électronique: vers de nouvelles pratiques adaptées à l'ère du numérique. *ALSIC*, 12. <http://alsic.revues.org/index1314.html> (08/09/2011).

Le Priault, H. (1996). *Grammaire progressive de l'anglais*. Paris: Belin.

O'Sullivan, Í. (2007). Enhancing a process-oriented approach to literacy and language learning: the role of corpus consultation literacy. *ReCALL*, 19, 3, 269-286.

Oxford advanced learner's dictionary (1974). Oxford: Oxford University Press.

Schmitt, N. & McCarthy, M. (1997). *Vocabulary: description, acquisition and pedagogy*. Cambridge: Cambridge University Press.

Schön, D. (1983). *The reflective practitioner: how professionals think in action*. New York: Basic books.

Serre, F. (1992). La science en action: le rapport entre la science et la pratique professionnelle. In: R. Tessier & Y. Tellier (éds). *Changement planifié et développement des organisations: méthodes d'intervention, consultation et formation* (pp. 395-422). Québec: Presses de l'Université du Québec.

Surcouf, C. (2010). Le dictionnaire bilingue peut-il s'intégrer profitablement dans une stratégie d'apprentissage d'une langue étrangère? *Cahiers de l'APLIUT*, 29, 1, 116-136.

Test of English for international communication. <http://www.toeic.eu/> (02.02.2010).

Tournier, J. (2004). *Précis de lexicologie anglaise*. Paris: Ellipses édition.

Wittorski, R. (2001). Contribution de l'apprentissage expérientiel et de la science action à la pratique professionnelle. In: M.-P. Mackiewicz (éd.), *Praticien et chercheur: parcours dans le champ social* (pp. 107-119). Paris: L'Harmattan.