



HAL
open science

La traduction assistée par ordinateur dans le contexte égyptien. Étude d'un système de TAO lié à Internet (français-arabe/arabe-français). Session 1 - Terminologie et Traduction

Dima El Hussein

► To cite this version:

Dima El Hussein. La traduction assistée par ordinateur dans le contexte égyptien. Étude d'un système de TAO lié à Internet (français-arabe/arabe-français). Session 1 - Terminologie et Traduction. Tralogy I. Métiers et technologies de la traduction : quelles convergences pour l'avenir ?, Mar 2011, Paris, France. 15p. <hal-02495540>

HAL Id: hal-02495540

<https://hal.science/hal-02495540v1>

Submitted on 2 Mar 2020

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.



HAL Authorization

TRALOGY

La traduction assistée par ordinateur dans le contexte égyptien

Etude d'un système de TAO lié à Internet (français-arabe/arabe-français)

Dima El Hussein

Faculté des Langues Appliquées, Université Française d'Égypte

TRALOGY I - Session 1
Date d'intervention : 03/03/2011

Le présent article se propose de mener une réflexion sur les outils d'aide à la traduction dans le contexte égyptien où la traduction littéraire prime et où la tâche de traduction des textes spécialisés est souvent assignée aux professionnels et spécialistes du domaine (médecins, informaticiens, juristes, etc.). Il nous paraît opportun d'articuler notre pensée autour de l'appréhension de l'usage de la traduction assistée par ordinateur dans son rapport avec la professionnalisation et la promotion de la profession de traducteur. Nous nous proposons de partir d'une interrogation préalable sur la manière dont il faut envisager les outils d'aide à la traduction. Nous envisageons dans cette étude la présence d'un logiciel de TAO lié à Internet. Il s'agit d'introduire, à côté d'un corpus parallèle qui s'enrichit au fur et à mesure de l'action traduisante, un corpus comparable à la fois permanent et renouvelable. Cette mise à jour se ferait systématiquement pour emboîter le pas aux changements qui s'opèrent dans le domaine technique et dans ses moyens d'expression. Il y aurait ainsi plusieurs degrés de recherche à définir et à effectuer.

lien video : http://webcast.in2p3.fr/videos-session1_la_traduction_assistee_par_ordinateur_dans_le_contexte_egyptien



Introduction

L'intérêt qu'accordent les instituts de recherche à la traduction assistée par ordinateur TAO en Egypte n'est pas récent. Bien que la traduction littérature occupe une place de choix dans la politique de traduction en Egypte depuis bien des années, le Centre National de la traduction (CNT), dépendant du Ministère de la Culture, a appelé à la promotion de la traduction technique. Une nouvelle tendance a émergé il y a quelques années, manifestée par l'intérêt des institutions de recherche publiques et privées à l'application de la nouvelle technologie à la langue arabe. Les projets menés jusqu'à nos jours en sont le témoin. Ce qui nous intéresse dans ce contexte c'est l'émergence de cette tendance grandissante à l'application de la nouvelle technologie à la traduction et son rapport avec la professionnalisation, le monde professionnel et la promotion de la profession de traducteur dans le monde arabe en général. En effet, le profil du traducteur arabe est en train de changer. Bien que ce changement soit relativement lent par rapport au nouveau profil requis, le traducteur arabe manifeste un intérêt particulier à l'application de la nouvelle technologie à la traduction, ce qui lui permettra de s'initier aux «métiers de la traduction». [Gouadec, D. (2003)], [Froeliger, N. (2009)].

Il nous semble important de Jeter la lumière sur la TAO en Egypte. Nous tenons à préciser que l'exemple de la TAO en Egypte est pris comme un indicateur dans le but de mesurer l'état de la recherche, du développement et de l'évolution de l'application des outils de TAO dans le monde arabe. Le grand nombre de recherches qui ont été menées sur la question dans le monde arabe et leurs résultats encouragent de plus en plus la recherche dans ce sens. En effet, le marché de la traduction libérale apprécie les connaissances des logiciels de TAO. En plus, l'usage de la TAO au sein des organisations internationales et la croissance de son intérêt à la fois économique et professionnel montre qu'il est important de mesurer son usage et son développement dans le contexte égyptien.

Les informations contenues dans cet article proviennent d'une enquête téléphonique et d'une série d'entretiens réalisées en décembre-janvier 2011. Il s'agit de deux universités (publique et *ahlia*¹), onze départements de traduction dans les institutions publiques, onze agences de traduction, onze enseignants, sept traducteurs indépendants, cinq sociétés d'informatique et six entreprises. L'enquête téléphonique et les entretiens ont été complétés par des informations supplémentaires tirées d'internet.

1. Traduction technique et TAO

La traduction technique est prise dans cet article dans l'acception de traduction pragmatique «qui a une fonction de communication» [Froelinger, N. (2009)]. Ainsi, nous traiterons des domaines techniques, scientifique, juridique, économique, financier, informatique, etc. Contenant des concepts et des objets liés à un domaine de connaissance «fortement thématique» [Lerat, P. (1995)], les textes spécialisés contiennent des termes relevant d'un domaine de spécialité, ne prêtent pas à confusion et ne sont pas sujets à plusieurs lectures comme dans la traduction littéraire ou la traduction des textes d'analyse. C'est ainsi que la traduction technique a l'avantage de pouvoir être automatisée.

Prise dans sa dimension de communication, la traduction est loin d'être un travail sur la langue. Ainsi, la méthodologie de la traduction insiste sur la compréhension du sens pour assurer sa transmission. La compréhension et la transmission du sens s'opèrent donc au niveau du discours et non au niveau de la langue [Séleskovitch, D. (2001)]. Pour acquérir la connaissance

(1) Il y a deux types d'universités non gouvernementales en Egypte : *ahlia* et privée. Contrairement à l'université privée, l'université *ahlia* est à but non lucratif

du domaine de spécialité qu'il traduit, le traducteur effectue la recherche documentaire et la recherche terminologique et phraséologique. Le lien est étroit entre traduction et recherche terminologique [Ammour, E. (2005)]. Cette recherche s'inscrit dans le cadre de la démarche traduisante. La recherche terminologique et phraséologique s'effectue à travers la lecture de textes parallèles portant sur le même domaine [Plassard, Freddie (2007)]. Ce processus ne se limite pas à l'établissement d'un index terminologique bilingue (ou multilingue) mais il assure la compréhension du domaine de spécialité à travers une méthode de recherche documentaire bien assimilée. [Israël, Fortunato (2001)]. En effet, cette recherche est effectuée par le traducteur dans une optique de traduction.

Il est de coutume en Egypte de confier la traduction d'un texte spécialisé à un spécialiste du domaine. Un texte juridique ou scientifique, par exemple, est mis entre les mains d'un juriste ou d'un scientifique pour assurer une meilleure qualité de traduction. Il n'est pas lieu ici de relancer le débat sur ce sujet, quoique ce soit intéressant. Mais, ce qui nous intéresse, dans ce contexte, c'est de mener une réflexion sur la profession de traducteur dans le contexte de la nouvelle technologie et sur les moyens qui permettent au traducteur arabe d'améliorer sa performance en utilisant les outils mis à sa disposition.

Le présent article développe des points déjà traités dans différentes études menées dans le domaine de la TAO notamment en ce qui concerne la langue arabe. Il se divise en trois parties. Dans la première partie, l'état actuel de la TAO en Egypte sera examiné à trois niveaux : enseignement supérieur, monde professionnel et recherche. Dans la deuxième partie, l'idée d'un logiciel de TAO lié à Internet sera exposée. Dans la troisième partie, des pistes de réflexions pour l'avenir seront évoquées.

2. Etat actuel de la TAO en Egypte

2.1 TA et TAO

Pour jeter la lumière sur le statu quo de la TAO en Egypte et proposer des pistes de réflexions, il nous paraît opportun de distinguer d'abord la traduction automatique TA de la traduction assistée par ordinateur TAO [Bouillon, P. et Clas, A. (1994)]. La TA que nous n'aborderons pas dans notre article est une traduction purement automatique qui n'implique pas l'intervention du traducteur professionnel. Mais, ce dernier est amené à réviser la traduction fournie par l'ordinateur. Quant à la TAO, elle nécessite l'intervention du traducteur qui crée la base textuelle du logiciel et alimente en permanence sa mémoire. Lorsqu'il utilise la TAO, le traducteur alimente la mémoire de traduction avec des textes sources et leurs traductions. Ainsi, les phrases précédemment traduites par le traducteur lui seront proposées par la mémoire du logiciel lors de la traduction d'un nouveau texte. Les textes peuvent être ainsi réutilisés. Il est donc important de distinguer la traduction automatique de la traduction faite par un humain assisté par la machine :

«On distingue TA (traduction automatique) et AT (aide au traducteur). En TA, la machine traduit, et les humains qui aident le système à le faire, en répondant à des questions de désambiguïsation, sont supposés ne connaître qu'une des langues en question. En AT, c'est un traducteur (professionnel ou non) qui fait le travail, même s'il est aidé par le système, qui lui propose des « suggestions de traduction », soit à partir d'une mémoire de traductions (MT), soit à partir des résultats d'un ou plusieurs systèmes de TA. On distingue ainsi classiquement la TA-R (pour réviseurs), la TA-A (pour auteurs), et la TA-W (pour veilleurs, ou *watchers* en anglais). [Boitet, C. (2007)]

2.2 Domaine de l'enseignement supérieur

Initier les apprentis traducteurs à la TAO commence à l'université en offrant une formation professionnalisante qui répond aux besoins du marché [Arrouart, C. (2003)]. Formés à la TAO, les

futurs traducteurs, pourront ainsi intégrer le monde du travail de la traduction qui demande, de plus en plus, des connaissances en matière d'outils technologiques en rapport avec la traduction.

En effet, une nouvelle tendance se dessine dans l'univers de l'enseignement supérieur égyptien où se manifeste un intérêt croissant aux outils technologiques et à leurs applications à la traduction. L'Université Française d'Égypte (UFE) et l'Université Ain Chams (deux statuts différents)² s'intéressent à étudier l'intégration de la TAO dans leurs cursus universitaires. A la faculté des Langues Appliquées de l'UFE, l'application des outils technologiques est déjà intégrée au cursus d'informatique. D'où le projet de construction d'une base de données terminologiques par les étudiants et l'utilisation des extracteurs et concordanciers dans les cours de terminologie française et dans la préparation du mémoire terminologique bilingue (français-arabe). Le département de traduction spécialisée compte intégrer à partir de l'année prochaine 2011-2012 l'enseignement de la TAO à son cursus. Les apprentis traducteurs qui participent cette année au projet de traduction (du français vers l'arabe) du livre «le marketing international» [Karsaklian, E. (2007)], auront à alimenter la mémoire de traduction avec la traduction qu'ils auront terminée à la fin du semestre prochain³.

En ce qui concerne les enseignants, ils sont divisés sur le sujet de l'enseignement de la TAO. Sept enseignants sur quatre sont favorables à l'idée. Ils évoquent l'avantage du temps épargné et de l'effort investi. Ils soulignent la nécessité d'emboîter le pas aux nouvelles technologies appliquées à la traduction et de contribuer à la formation du nouveau profil du traducteur.

Ceux qui sont défavorables à l'enseignement de la TAO, parlent surtout de qualité médiocre, de perte de temps (traduction révisée ou presque refaite par le traducteur). Ils craignent que les outils d'aide à la traduction n'altèrent la méthodologie de la traduction et ne font en sorte que les apprentis traducteurs dépendent entièrement de la machine.

Il est intéressant de faire remarquer que les programmes à des fins d'enseignement lancés par SDL Trados (<http://www.translationzone.com/en/community/academic-portal/>), Multitrans (<http://www.multicorpora.com/solutions/academic/>) et Wordfast (<http://www.wordfast.net/index.php?whichpage=buy&lang=engb>) sont concurrentiels à ce sujet surtout que certaines sociétés offrent un programme académique à titre gratuit. Ces programmes assurent une formation aussi bien pour les étudiants que pour les professeurs.

2.3 Domaine professionnel

Les outils d'aide à la traduction sont utilisés dans une large mesure dans les organisations internationales et leurs bureaux régionaux. En ce qui concerne les onze institutions publiques que nous avons contactées, aucune n'utilise la TAO pour des raisons de limites budgétaires. Les responsables trouvent que l'achat d'un outil d'aide à la traduction n'est pas rentable. En effet, ces institutions font appel à des traducteurs indépendants ou à des traducteurs travaillant à mi-temps, même si elles sont dotées d'une unité ou d'un département de traduction.

En ce qui concerne les entreprises (locales ou multinationales), elles ont une politique commune qui consiste à recruter une personne bilingue qui, à côté de ses compétences dans son domaine de spécialité, s'occupe de la traduction.

Sur les onze agences de traduction contactées quatre utilisent des outils d'aide à la traduction. Certaines agences offrent une formation payante à la TAO, notamment à SDL Trados qui est très répandu en Égypte. En plus, les agences de traduction ont de plus en plus tendance à rechercher le profil du traducteur qui maîtrise la TAO. Les cinq sociétés d'informatique qui offrent le service de traduction - localisation utilisent les logiciels de TAO.

(2) L'Université Française d'Égypte est une université *ahlia* et l'Université Ain Chams est une université publique

(3) Ce projet s'étend sur une année universitaire et constitue une simulation d'un projet de traduction de la vie professionnelle : signature de contrat, désignation d'un chef de projet, formation d'équipes de traducteurs, de réviseurs et de réviseurs linguistiques

En ce qui concerne les traducteurs indépendants, il est curieux de noter qu'un certain nombre de jeunes traducteurs s'intéressent à l'achat des mémoires de traduction par souci de professionnalisme et pour répondre au nouveau besoin du marché, d'autant plus que la maîtrise de la TAO figure dans les pré-requis du profil du traducteur recherché. Sur les sept traducteurs indépendants contactés, quatre utilisent la TAO. Ceux qui travaillent auprès de l'ONU, disposent en général d'une mémoire de traduction étant donné qu'ils avaient déjà expérimenté l'usage de ces outils au sein de l'organisation et ne peuvent pas s'en passer.

Il y a une sensibilisation grandissante vis-à-vis de l'utilisation de la TAO ces dernières années à travers les annonces de formations gratuites en ligne proposées par certaines organisations et associations arabes de traduction. Cette sensibilisation met en premier lieu l'accent sur la différence entre traduction automatique et traduction humaine assistée par ordinateur pour qu'il n'y ait pas de confusion.

2.4 Domaine de la recherche

Le domaine de la traduction assistée par ordinateur et, par extension, celui de la traduction automatique, attire l'attention de nombreux chercheurs dans le monde arabe. Présentés dans des colloques nationaux ou internationaux ou dans le cadre de projets, ces recherches s'orientent vers la définition des difficultés rencontrées dans le traitement automatique de la langue arabe et proposent des solutions pour son développement. Trois projets financés par la commission européenne sont à mentionner dans ce sens. Le projet NEMLAR (Network for Euro-Mediterranean Language Resources) (www.medar.info/The_Nemlar_Project/Publications/annual-report2003.pdf), le projet MEDAR (Mediterranean Arabic language : www.medar.info) et le projet DINAR (Dictionnaire Informatisé de la Langue Arabe : <http://silat.univ-lyon2.fr/>). Trois entreprises égyptiennes⁴ ont participé au projet MEDAR dont la société Sakhr (entreprise koweïtienne implantée en Egypte <http://www.sakhr.com/HomePage.aspx>) qui développe les logiciels de traitement automatique de la langue arabe, de la traduction automatique et des outils d'aide à la traduction.

L'Institut des recherches en électronique au Caire accorde une importance aux recherches menées dans le domaine du traitement automatique de la langue arabe et à la traduction assistée par ordinateur. Ces recherches attestent d'une tendance positive vers le développement des logiciels de TAO de langue arabe [Hamada, S. (2010)]. Une étude intéressante est en cours pour concevoir un système de TAO à un prix abordable pour le traducteur égyptien [Hamada, S. (2010)].

Le rôle de l'Organisation Mondiale de la Santé n'est pas à ignorer. Les séminaires de l'Organisation portant sur l'arabisation de la médecine au bureau régional du Caire attestent d'un 'intérêt croissant vis-à-vis de l'usage de la TAO dans la traduction des documents scientifiques de l'organisation [EL Jazaeri, H. (2005)].

En ce qui concerne les sujets des thèses et des mémoires soutenus dans les universités égyptiennes ou ceux qui sont en cours, ils s'inscrivent dans cette nouvelle tendance qui s'intéresse à l'application de la nouvelle technologie à la langue arabe et à la traduction.

2.5 Problèmes constatés

Après ce tour d'horizon, il ressort que grâce aux recherches menées, des progrès ont été réalisés au niveau de l'analyse morphologique de la langue arabe, des logiciels de mesure de fréquence et des outils d'alignement pour la langue arabe [Guidère, M. (2005)]. Certes, il y a un intérêt manifesté et un progrès réalisé. Toutefois, les initiatives au niveau du monde arabe ne sont pas satisfaisantes. Il reste beaucoup à faire pour promouvoir la recherche dans ce domaine au niveau des institutions aussi bien publiques que privées.

(4) Engineering company for the Development of Computer systems (RDI), Human Language Technologies group (IBM) et Sakhr softwar company

En effet, les raisons de limites budgétaires ont été évoquées par toutes les parties contactées. Mais, si dans le domaine de l'enseignement certains enseignants s'opposent fermement à l'enseignement de la TAO, c'est parce qu'ils ne préfèrent pas se lancer dans un domaine qu'ils n'ont pas expérimenté et qu'ils ne maîtrisent pas encore. Certes, cette opposition est due à la confusion qu'ont encore certains enseignants entre TA et TAO. Quant aux entreprises, il est clair que le concept de profession de traducteur et de traductologie n'est pas encore bien implanté dans le milieu professionnel.

Nous remarquons que les institutions publiques et les agences de traduction ont majoritairement recours aux traducteurs indépendants qui, en expérimentant la TAO, expriment certains besoins dont l'alimentation permanente de la base de données terminologiques du logiciel et le besoin d'effectuer une recherche terminologique et phraséologique pour certains domaines. Les traducteurs indépendants affirment que les bases de données ne contiennent pas les nouveaux termes relatifs aux dernières recherches et découvertes.

2.6 Solution envisagée

Il nous semble opportun, à l'heure actuelle, d'œuvrer à élargir le recours à la TAO dans le monde arabe et ce en imitant la démarche du traducteur. Ce dernier entreprend une recherche documentaire, terminologique et phraséologique pour acquérir la connaissance thématique et notionnelle du domaine de spécialité. L'extraction phraséologique et terminologique se fait à partir d'un corpus comparable spécialisé que le traducteur se charge de confectionner. Cette recherche qui se fait dans une optique de traduction est complétée par le recours aux spécialistes. Ces derniers orientent le traducteur vers des ressources, lui donnent des précisions sur les concepts et le renseignent sur l'usage réel des termes dans le milieu professionnel. La solution envisagée dans ce contexte est de lier le logiciel de TAO à Internet. Le logiciel entreprendra la démarche du traducteur. Il confectionnera le corpus comparable spécialisé sur le web et aura accès aux répertoires de sites prédéfinis par les spécialistes. Ce qui permettra l'alimentation de la base de données terminologiques.

3. TAO lié à Internet

3.1 Corpus comparable sur le web

Internet est une arme à double tranchants. Grâce au web et au flux de l'information, des ressources en différentes langues sont accessibles. Toutefois, certains liens sur internet ne sont pas crédibles. On y trouve parfois des documents de moindre qualité, mal rédigés et comportant des informations erronées. Nul doute que le web constitue une source importante pour la confection d'un corpus comparable [Kilgarriff, A. et Grefenstette, G. (2003)], [Cavaglià, G. Kilgarriff. (2001)], [Déjean, H. Gaussière, E. (2002)]. Le problème des mots clés de recherche et la proposition d'un outil de recherche pour la collecte rapide d'un corpus comme celui de WebBootCaT ont été étudiés [Baroni, M. *et al.* (2006)]. Citons dans ce contexte le projet TTC de la commission européenne «Extraction terminologique, outils de traduction et corpus comparables», qui traite de l'extraction et de la gestion des termes des corpus comparables spécialisés constitués à partir du web dans le but de développer les outils d'aide à la traduction. [Blancafort, H. *et al.* (2010)]. Il n'est pas inutile de rappeler ici l'importance du traitement automatique des équivalences des corpus comparables [Morin, E. *et al.* (2004)], [L'Homme, M. (2004)].

A la différence des corpus alignés, les corpus comparables ne constituent pas une traduction. Les critères régissant le choix des textes qui composent un corpus comparable concernent «le même niveau de langue, la même tranche chronologique ou la même variété régionale [...]

Pour le terminographe, la parenté se définit en fonction des thématiques abordées dans les textes à savoir le domaine de spécialité ou la subdivision d'un domaine générique» [L'Homme, M. (2004)]. Il est inutile de rappeler les avantages des corpus comparables qui ont été attestés sur le plan pratique et théorique [Maniez, F. (2001)]. L'accès direct à l'usage réel des termes dans leurs domaines de spécialité est assuré et le traducteur effectue sa traduction sans être imprégné par la langue du texte de départ [Déjean, H., Gaussier, E. (2002)].

Pour pouvoir bien profiter d'Internet et pour que la confection des corpus comparables soient à l'abri des sources incroyables, l'évaluation du document web s'avère être une étape importante. Austermühl a soulevé les points essentiels relatifs à l'évaluation dont la vérification de la signature du site et des coordonnées du concepteur, l'identification du concepteur du site (organisme, université, gouvernement, etc.), la précision de la date de la dernière mise à jour, l'existence d'une bibliographie, de liens vers d'autres sites et le renvoi d'autres sites au site consulté [Austermühl, F. (2001)].

3.2 Corpus arabe sur le web : tour d'horizon

Si on compare, au niveau des ressources, la collecte des corpus français ou anglais à celle des corpus arabes sur le web, nous nous rendons compte que les ressources accessibles sur Internet en français ou en anglais sont plus abondantes que les ressources en arabe. Plusieurs recherches ont mis l'accent sur l'importance de la constitution d'un corpus arabe en faisant allusion aux corpus existants dans d'autres pays comme l'American National corpus, le British National Corpus et le corpus français FRANTEXT [Abbès, R. et Dichy, J. (2008)]. D'autres recherches ont fait état des difficultés relatives à la constitution d'un corpus arabe sur le web et ont proposé des solutions que nous mentionnerons dans cette partie [Abbès, R. et Dichy, J. (2008)]. Nous allons plutôt nous focaliser sur le corpus arabe dont la confection sur le web représente jusqu'à nos jours certaines difficultés.

En effet, il existe des corpus arabes importants accessibles sur Internet. Ces corpus ont vu le jour grâce aux projets qui ont participé au développement des différents outils de traitement automatique de la langue arabe. Les corpus arabes confectionnés dans le cadre des projets NEMLAR (http://catalog.elra.info/product_info.php?products_id=873), MEDAR et DINAR sont accessibles sur le site de l'Agence pour l'évaluation et la distribution des ressources linguistiques ELDA (http://www.elda.org/medar_knowledge_base/monolingual_corpora.php) qui exécute les plans et les stratégies de l'association européenne des ressources linguistiques ERLA. Nous trouvons aussi dans le catalogue d'ERLA les corpus de deux journaux libanais. Le corpus du journal Al Hayat (http://catalog.elra.info/product_info.php?products_id=632&language=fr) regroupe 42 591 articles confectionnés dans le cadre d'un projet de recherche de l'Université d'Essex au Royaume Uni, en collaboration avec Open University et le corpus du journal Al Nahar (http://catalog.elra.info/product_info.php?products_id=767) est constitué d'articles publiés entre 1995 et 2000. Il y a aussi le corpus du site d'information Al Jazeera «Aljazeera News Arabic Dataset» qui regroupe 1500 documents et le corpus de la magazine «Al-jazirah Magazine Arabic Dataset» qui contient 4470 articles publiés entre 2001 et 2005 [Said, D. *et al.* (2009)].

Les universités ne manquent pas de manifester leur intérêt à diffuser en ligne les corpus arabes comme l'Université de Leeds (<http://smlc09.leeds.ac.uk/query-ar.html>) qui offre également un accès aux liens des corpus arabes payants et gratuits (<http://www.comp.leeds.ac.uk/eric/latifa/ArabicCorporWebConc.htm>), l'Université d'Essex (<http://privatewww.essex.ac.uk/~melhaj/easc.htm>), et l'université de Brigham Young aux Etats-Unis (<http://arabiccorpus.byu.edu/>) [Anizi, M. et Dichy, J. (2009)].

4. Confection d'un corpus arabe du web

4.1 Difficultés

La recherche des ressources arabes sur internet constitue une difficulté surtout lorsqu'il s'agit des domaines spécialisés. Ainsi, la collecte de textes pour confectionner un corpus spécialisé en arabe avec une similitude de style et de contenu est souvent difficile [Nikkhou & Choukri, 2004]. Ce point est très important car «pour être comparables dans ce contexte les textes doivent présenter un nombre élevé de similitudes» [L'Homme, M.C. (2004)].

Dans le cadre des projets précités, les problèmes qui se posent lors du traitement automatique de la langue arabe ont été identifiés. Ces difficultés se résument à la variation morphologique, à la variété orthographique des mots (malgré leur graphie identique, ils sont différents de point de vue sémantique), à l'absence de signes de vocalisation, à l'absence du signe diacritique *la chadda*⁵, à l'absence de la dernière voyelle, à l'omission du signe notant la *hamza*⁶ et à la translittération des mots étrangers en arabe. [Abbès, R. *et al.* (2005)], [Abbès, R. et Dichy, J. (2008)], [Ali, N. (1994)]. Dans ce sens, des solutions ont été proposées en 1989 par l'équipe SAMIA⁷ (Synthèse et Analyse Morphologiques Informatisées de l'Arabe) en ce qui concerne le traitement automatique de la morphologie de l'arabe et l'analyse du mot graphique non vocalisé [Dichy, J. et Hassoun, M.O (1989)]. Plus tard, des propositions ont été avancées pour surmonter le problème de l'absence (totale ou partielle) des signes de vocalisation qui gêne la détection du sens du mot : «ce type d'ambiguïté pose encore de réels problèmes à la machine.» [Abbès, R. et Dichy, J. (2008)].

4.2 Engins de recherche

Dans le cadre du projet MEDAR, les problèmes qui se posent lors de la recherche de documents arabes sur internet ont été soulevés au 2^{ème} colloque international sur les ressources et le traitement automatique de la langue arabe, tenu au Caire en 2009. En effet, le mode de fonctionnement des engins de recherche sur Internet dans leur rapport avec la morphologie de la langue arabe, la variation terminologique d'une région arabe à une autre et le manque de sites en langue arabe pour certains domaines de spécialité posent des difficultés au niveau de la recherche de ressources en arabe [Guidère, M. (2005)].

L'engin de recherche communément utilisé pour effectuer une recherche de ressources en arabe est Google (www.google.com.eg). Les résultats obtenus suite à l'interrogation de Google ne sont pas entièrement satisfaisantes en ce qui concerne la recherche des ressources spécialisées. Ces résultats posent les mêmes problèmes déjà mentionnés (ambiguïté due à la morphologie, à l'absence des signes de vocalisation, à la variété orthographique etc.) [Abd El Salam, H. (2009)]. Ainsi, un mot comme *kawkab* (planète) recherché dans le moteur google donne des résultats qui nécessitent une sélection rigoureuse en fonction des différents contextes. En l'absence d'un analyseur morphologique, la recherche exige la précision des différentes variations morphologiques du mot [Anizi, M. et Dichy, J. (2009)]. En ce qui concerne la variation orthographique, elle est montrée à travers l'exemple du mot graphique : علم «Hors contexte, ce mot peut recevoir sept vocalisations / 'alam (drapeau) عِلْم ; / 'ilm (science) عِلْم ; / 'alima (il a su) ; عِلْم / 'ulima (il a été su) عِلْم ; / 'allama (il a enseigné) عِلْم ; / 'allim (enseigne - impératif) عِلْم ; عِلْم / 'ullima (il a été enseigné)». [Alrahabi, M. et Dichy, J. (2009)]

Les deux engins de recherche arabes www.ayna.com et www.albawaba.com sont très limités et ne donnent pas de résultats satisfaisants en matière de recherche dans les domaines spécialisés. Par contre, le logiciel de recherche Aldaal (www.arabtext.ws/EN_Daal.htm) développé par la

(5) La *chadda*, placée au dessus de la lettre, double sa prononciation

(6) La *hamza* est une lettre qui s'écrit parfois comme un diacritique

(7) Laboratoire d'Informatique documentaire de Lyon 1 et Centre de Recherches en Terminologie et Traduction de Lyon 2

société jordanienne arabtext (www.arabtext.ws) apporte des solutions aux problèmes soulevés ci-dessus surtout qu'il contient un analyseur morphologique pour l'arabe. Le logiciel permet aussi d'effectuer une recherche en langue anglaise [Hattab, M. *et al.* (2009)]. Mais, il n'est pas en accès libre. La nécessité de développer des engins de recherche arabes accessibles sur Internet est un besoin de plus en plus pressant [EL Jazaeri, H. (2005)].

4.3 Variation terminologique

Au niveau de l'écrit, l'arabe est une langue commune à tout le monde arabe. Nous n'aborderons pas dans notre article la question du dialecte, mais nous soulèverons l'idée de la variation terminologique. Dans certains pays, des termes sont plus d'usage que d'autres et les termes étrangers sont empruntés. La variété terminologique concerne le *Machreq* (Jordanie, Liban, Palestine, Syrie), le *Maghreb* (les pays du Maghreb) et l'Égypte. Citons comme exemple le pluriel d'usage du terme capital رأس مال en arabe. Au Liban, le pluriel du terme capital est رساميل (rassamil) comme on le lit souvent dans le quotidien Dar Al Hayat, tandis qu'en Égypte, le pluriel d'usage est أموال رؤوس (rou'ouss amwal). Pour campagne de sensibilisation, on dit au Maghreb حملة تحسيس tandis qu'en Égypte et au Machreq on dit توعية حملة. Le domaine de droit abonde en exemples. Citons celui de la Cour d'appel. En Algérie, on utilise قضاء مجلس tandis qu'en Égypte, on dit محكمة الاستئناف.

Avec la révolution égyptienne qui a éclaté le 25 janvier 2011 et a conduit à l'éviction du régime du président Hosni Mubarak le 11 février 2011, le vocable égyptien «baltagui», utilisé dans le registre dialectal, est devenu courant dans la presse arabe (et étrangère aussi) est dorénavant lié aux événements de la révolution égyptienne. Le mot «baltagui» signifie voyou, sans foi ni loi, armé de couteau ou de fouet. Le «baltagui» peut être un homme ou une femme «baltagueyya « (fortes-à-bras). Dans le contexte égyptien, ce vocable a été toujours lié aux élections. Ces «baltaguis» sont payés par le régime au pouvoir pour perturber le cours des élections. Leur rôle va jusqu'à l'orientation de l'avenir politique du pays. [Al Korachi, A. (2010)]. Lors des affrontements meurtriers au Caire entre partisans et opposants du président Mubarak, les «baltaguis», armés de couteaux, de fouets et de bâtons ont attaqué les manifestants à la place Tahrir avec des chevaux et des dromadaires. Dorénavant, le vocable «baltagui» renvoie aux «mains payés», aux «hommes de main du régime» [Al Gamal, Samar (2011)] et à une réalité politique qui s'inscrit dans l'histoire de la révolution égyptienne.

4.4 Variété orthographique

L'écriture de la *hamza* diffère d'une région arabe à une autre. En Égypte, dans le vocable responsable مسئول la hamza s'écrit sur la ligne tandis qu'au Liban et en Syrie, pour le même vocable, la hamza s'écrit sur la lettre و (waw) مسؤول. Une autre différence se situe au niveau de la transcription arabe de la lettre g. Cette dernière est transcrite par la lettre غ (gh) au Liban et en Syrie et par la lettre ع chez les Égyptiens. Ainsi, nous avons deux formes orthographiques différentes pour le vocable google en arabe : غوجل (Liban, Syrie) et جوجل (Égypte). Une autre forme orthographique différente est celle de la lettre ي qui s'écrit de deux manières différentes. Au Liban et en Syrie la lettre ي à la fin de la phrase s'écrit avec deux points (ي) tandis qu'en Égypte, la lettre s'écrit sans les deux points [Abbès, R. et Dichy, J. (2008)]. Ainsi, nous avons deux orthographes différentes pour la lettre ي dans le vocable «révolutionnaire», par exemple, ثوري (Liban, Syrie) et ثوري (Égypte).

4.5 Automatiser la démarche de pré-traduction

Nous envisageons dans cette étude la présence d'un logiciel de TAO lié à Internet. Abordée de point de vue méthodologique, cette approche a pour objectif la conception d'un outil d'aide à la traduction pour automatiser la démarche de pré-traduction et l'adapter aux besoins du traducteur.

La mémoire du logiciel de traduction serait non seulement alimentée par un corpus parallèle indexé (arabe-français et français-arabe) mais aussi par des sources d'Internet pouvant être utiles au traducteur dans sa recherche terminologique et phraséologique bilingue et répertoriées en fonction du domaine de spécialité défini par le traducteur. Il s'agit donc d'introduire, à côté d'un corpus parallèle qui s'enrichit au fur et à mesure de l'action traduisante, un corpus comparable à la fois permanent et renouvelable.

Ces sources seraient déjà mémorisées dans le logiciel. Le traducteur aurait la possibilité d'alimenter la mémoire en ajoutant d'autres sources à celles qui lui seraient fournies par le logiciel. Ces sources seront répertoriées en fonction des deux langues française et arabe. Elles orienteraient le logiciel qui, après définition du domaine de spécialité, irait chercher sur Internet les documents qui comporteront les mots clés précisés par le traducteur. Nous rappelons qu'il s'agit de la constitution d'un corpus comparable spécialisé qui viendrait en aide au traducteur lorsque la base de données terminologique du logiciel ne comporte pas les éléments utiles à la traduction. En effet, les nouveaux termes relatifs aux dernières découvertes dans les domaines technique et scientifique ne sont pas disponibles en arabe dans les bases de données terminologiques [Besançon, R. *et al.* (2009)]. Le traducteur pourrait alimenter en permanence la base de données. La recherche terminologique et phraséologique des textes spécialisés assure la mise à jour des bases de données surtout que la langue est dans une évolution continue comme l'a déjà souligné Karl-Heinz Brinkmann :

«le langage technique, en particulier, dans plusieurs domaines évolue à un rythme vertigineux non seulement en ce qui concerne le vocabulaire mais aussi les autres aspects (par exemple, la phraséologie, la syntaxe, la sémantique)» [Brinkmann, K. (1979)]

4.6 Méthode de recherche

Le traducteur choisirait dans le menu déroulant le domaine de spécialité du texte à traduire. Un menu de sous-domaines s'afficherait sur l'écran. Après la sélection des sous-domaines, une liste de pays arabes s'afficherait. Le traducteur pourrait, s'il le juge nécessaire, sélectionner le pays (Egypte, Syrie, etc.) pour ne pas tomber sur des sites qui représentent une variation terminologique dont il n'aura pas besoin. Ainsi, il aura la possibilité d'affiner sa recherche terminologique et phraséologique. Cette question doit être prise en compte dans la constitution du répertoire des sites. Ensuite, vient l'étape de la définition des mots clés de recherche que le logiciel ira récupérer sur les sites correspondants aux sous-domaines choisis et déjà enregistrés dans le logiciels. Pour chaque site enregistré, il y aurait un descriptif succinct précisant le type de ressource : revue, rapport, procès verbal, etc. et le concepteur du site, notamment le pays d'origine (pour l'arabe littéral : Egypte, Syrie, Tunisie, etc. Pour le français : France, Canada, Suisse, etc.). La recherche par mots clés devra respecter l'occurrence des termes et des phrases contenus dans le document à traduire et ce, pour garantir des résultats satisfaisants.

Le logiciel de TAO lié à Internet comporterait également les liens de ces corpus existants sur la Toile. Après la recherche des corpus, le logiciel les téléchargerait sur l'ordinateur. Pour extraire la terminologie et la phraséologie, le traducteur utilisera les extracteurs et les concordanciers.

Il n'est pas de doute que la participation des spécialistes à la mise en place de ce système revêt d'une grande importance. Maîtrisant leur domaine de spécialité, les spécialistes seront les mieux qualifiés pour proposer une panoplie de sites à intégrer au logiciel. Un spécialiste en électronique ou un scientifique, par exemple, indiquera les sites qu'il a l'habitude de consulter pour rechercher des informations ou diffuser les derniers résultats et découvertes : «on trouve à présent énormément de textes de chercheurs en sciences se rapportant aux dernières recherches effectuées dans les différents domaines» [Sanchez, X. et Bonneville, C. (2006)]. Il s'agit évidemment, dans le cas de la langue arabe, des spécialistes de tous les pays du monde arabe pour avoir des sites qui représentent les différentes variations terminologiques.

4.7 Avantages du système proposé

Le logiciel de TAO lié à Internet serait alimenté par des ressources consultables en ligne. Comme nous l'avons signalé, la recherche du corpus spécialisé s'opère par mots clés définis par le traducteur. La possibilité pour un traducteur d'avoir accès à des sites spécialisés où il pourrait consulter des corpus spécialisés lui permet de suivre les mises à jour des nouveautés du domaine, d'une part, et de la langue, d'autre part.

Nous estimons que le traducteur pourra profiter des avantages de ce système pour améliorer sa performance :

Gain de temps. Le traducteur épargnera du temps au niveau de la recherche et de la collecte de documents sur Internet puisque le logiciel ira chercher dans les sites enregistrés en fonction des paramètres sélectionnés. En plus, il ne perdra pas son temps à vérifier l'authenticité du site.

Crédibilité des sources. Le site consulté sera crédible étant donné qu'il a été recensé par des spécialistes du domaine qui ont l'habitude de le consulter ou d'y publier leurs recherches récentes. Revenant aux critères de l'évaluation d'un document tiré du web établis par Austermühl, nous nous rendons à l'évidence que la vérification de la signature du site ou des coordonnées du concepteur sont sans doute assurées. En effet, la liste de sites établie par les spécialistes devrait répondre aux critères établis par Austermühl.

Qualité des documents. Les documents spécialisés rédigés par des spécialistes ou du moins vérifiés ou consultés par des spécialistes seront directement accessibles⁸. L'un des avantages de la définition du domaine de spécialité pour la collecte de textes spécialisés est d'éviter la polysémie comme dans l'exemple du vocable *kawkab* (planète) cité plus haut [Anizi, M. et Dichy, J. (2009)]. Puisqu'on parle de domaine, on parle de terme et non de mot. Dans ce cas, la polysémie n'a pas lieu d'être car le terme, dans un domaine donné, ne peut pas avoir plusieurs sens en fonction du contexte comme dans le cas du mot : «les termes, objets de la recherche terminologique, se trouvent-ils étroitement liés à des domaines spécialisés» [Durieux, C. (1997)]. Le résultat devrait être, méthodologiquement parlant, satisfaisant puisque le traducteur tomberait directement sur la terminologie et la phraséologie du domaine défini. L'avantage à mentionner ici est sans doute celui de l'exclusion de la source wikipedia de la constitution des corpus vu que les textes qui la constituent sont rédigés par des volontaires et pourraient contenir des informations erronées [Le Genre, B. (2008)].

Travail accompli en temps réel. Le logiciel travaillerait en temps réel. Contrairement au corpus enregistré sur l'ordinateur qui exige une mise à jour, le corpus confectionné en temps réel et adapté aux besoins du traducteur a l'avantage d'être mis à jour.

Sélection et adaptation de la variation terminologique. L'affichage de l'origine du site permet au traducteur de sélectionner le pays pour affiner sa recherche terminologique et phraséologique. La recherche ponctuelle par pays répond au besoin du traducteur. Elle lui permet également d'adapter le choix terminologique et phraséologique au pays destinataire. Il est à noter que le traducteur arabe est, habituellement, au courant de ces variations terminologiques.

Alimentation de la base de données terminologique. Le traducteur va ajouter à la base de données du logiciel de TAO les terminologies et phraséologies qu'il a extraites. C'est aussi un moyen de rentabiliser le temps investi dans la recherche.

(8) La possibilité d'avoir accès à des sites créés par des spécialistes et qui nécessitent une inscription payante est à envisager.

4.8 Limites du système proposé

Il nous semble que des difficultés se poseront lors du recensement des sites par les spécialistes. Dans certains domaines très pointus, la recherche est rédigée exclusivement en anglais ou en français (à l'exclusion de la Syrie où ce problème ne se posera pas dans le domaine de la médecine vu la politique d'arabisation appliquée dans le domaine de l'enseignement). D'où la rareté et même l'absence de ressources techniques en arabe notamment dans le domaine de la technologie de l'Information [Besançon R. *et al.* (2009)]. En Egypte, par exemple, les thèses et les recherches portant sur la médecine ou sur des domaines pointus comme l'électronique sont soutenues, enseignées, rédigées et publiées en anglais. Toutefois, nous estimons que ce point ne constitue pas un vrai problème vu la rareté de la demande en traduction vers l'arabe dans ces domaines ne serait-ce que pour la vulgarisation. Notons également que la rareté de ressources des domaines pointus en arabe peut être due au progrès assez lent de la production scientifique dans le monde arabe [Daoud, M. *et al.* (2009)].

5. Perspectives

En effet, les logiciels ne sont qu'une simple aide pour la traduction des textes spécialisés. Ils ne peuvent pas remplacer la réflexion et l'analyse du traducteur. Il n'est pas de doute que sans méthodologie, les résultats ne seront pas satisfaisants.

Au terme de cette étude, nous pourrions tirer des conclusions assez optimistes de point de vue méthodologique. L'imitation de la démarche du traducteur est un élément clé à développer en encourageant la coopération entre les traducteurs professionnels et les ingénieurs pour parvenir à l'automatisation de la démarche traductive. Nous avons tenu à jeter la lumière sur les domaines porteurs de l'application de la TAO en Egypte. L'intérêt manifesté à la TAO par les traducteurs indépendants, les professionnels et les enseignants est encourageant et le désir d'acquérir des compétences à la fois traductionnelles et technologiques leur permet de se positionner vis-à-vis du nouveau profil du traducteur défini par Gouadec [Gouadec, D. (2007)].

Nous estimons que l'apprentissage des logiciels de la TAO et de leurs fonctionnalités devient un besoin pressant non seulement pour les apprentis traducteurs et les traducteurs professionnels mais aussi pour les enseignants. En effet, les étudiants ne doivent utiliser les outils d'aide à la traduction qu'une fois les réflexes méthodologiques acquis et après un entraînement sérieux à la recherche et la gestion documentaire. En tout cas, pour éviter toute appréhension chez les enseignants, il faudrait idéalement qu'ils puissent s'essayer à la mémoire de traduction eux-mêmes et la pratiquer un minimum pour en connaître les avantages et les limites.

Il nous paraît opportun de mener une réflexion sur les moyens de sensibiliser les institutions publiques vis-à-vis de l'utilisation des outils d'aide à la traduction. L'organisation de journées consacrées à la TAO regroupant les professionnels de la traduction, les enseignants, les institutions publiques et privées ainsi que les bureaux régionaux des organisations internationales permettra d'échanger des expertises et définir la demande potentielle du marché.

Bibliographie

Abbès, Ramzi (2002), « AraFreq : un outil pour le calcul de fréquences de mots arabes », in A. Braham, ed. *Colloque international sur le traitement automatique de l'arabe*, Avril 18-20, Université de la Manouba, Tunis.

Abbès, Ramzi *et al.* (2005), Morpho-lexical ambiguities in the recognition of written Arabic word-forms, evidence from the DIINAR.1 lexical resource, ACIDCA-ICMI'05 – Tozeur, 5-7 novembre.

Abbès, Ramzi et Dichy, Joseph (2008), « Extraction automatique de fréquences lexicales en arabe et analyse d'un corpus journalistique avec le logiciel AraConc et la base de connaissances DIINAR.1 » in Heiden, Serge et Bénédicte Pincemain, *Actes des JADT 2008, 9es journées internationales d'analyse statistique des données textuelles*, Lyon 12-14-mars, Presses Universitaires de Lyon, 2 vol. , pp. 31-44.

Al Jazaeri, Hussein Abdel Razak (2005), 3^{ème} séminaire scientifique sur l'arabisation de la médecine 6-7 juin, Organisation mondiale de la Santé, bureau régional, le Caire.

Ali, Nabil (1994), *les Arabes et l'ère de l'information*, Koweït, Conseil National de la culture, des arts et de la littérature, coll. Monde de de la connaissance ('alam al ma'rifa) n° 184.

Alrahabi, Motasem et Dichy, Joseph (2009), « Levée d'ambigüité par la méthode d'exploration contextuelle : la séquencec'alif-nûn (ان) en arabe », in Ghenima, Malek, Ouksel, Aris et Sidhom, Sahbi eds., *Systèmes d'Information et Intelligence Economique*, 2^{ème} Conférence Internationale (SIIE),Tunis, Hammamet, 12-14 février, IHE éditions, pp. 573-585.

Ammour, Elisabeth (2005), « La recherche documentaire dans l'enseignement de la traduction » in Israël, Fortunato et Lederer, Mariane eds, *Théorie Interprétative de la Traduction. t.III : De la formation à la pratique*, Paris-Caen, Lettres modernes minard, pp. 85-94.

Arrouart, Catherine (2003), « Les mémoires de traduction et la formation universitaire : quelques pistes de réflexion », *Meta : journal des traducteurs*, vol 48, n° 3, pp. 476-479.

Aubin, Marie-Christine (1995), *Splendeurs et misères de la traductique*, Cahiers Franco-canadiens de l'Ouest, vol. 7, n° 2, 1995, pp. 211-226.

Austermühl, Frank (2001), *Electronic Tools for Translators*, Manchester, St.Jerome publishing.

Baroni, Marco and Kilgarriff, Adam *et al.* (2006), *WebBootCaT: a web tool for instant corpora*, *Proc. Euralex*. Torino, Italy.

Besançon, Romaric *et al.* (2009), « Arabic, English and French: three languages in a filtering systems evaluation project », 2nd International Conference on Arabic Language Resources and Tools 22-23 April, Cairo, Egypt.

Blancafort, Helena and Daille, Béatrice *et al.* (2010), *TTC: Terminology extraction, translation tools and comparable corpora*. In *Proc. EURALEX 2010*, Leeuwarden, 5-6 July.

Boitet, Christian (2007), « Corpus pour la TA : types, tailles et problèmes associés, selon leur usage et le type de système », *Revue française de linguistique appliquée/ 1*, vol. XII, pp. 25-38.

Cavaglià, Gabriela and Kilgarriff, Adam (2001), *Corpora from the web, 4th CLUK colloquium*, Sheffield.

Clas, André et Bouillon, Pierrette (dir.) (1995), *TA-TAO : recherches de pointe et applications immédiates*, Actes des Troisièmes Journées scientifiques du réseau thématique de recherche « Lexicologie, Terminologie, Traduction », coll. « Universités francophones », Beyrouth, AUPELF-UREF et FMA.

Daoud, Mohammad *et al.* (2009), « Collaborative Construction of Arabic Lexical Resources », 2nd International Conference on Arabic Language Resources and Tools 22-23 April, Cairo, Egypt.

Déjean, Hervé et Gaussière, Eric (2002), « Une nouvelle approche à l'extraction de lexiques bilingues à partir de corpus comparables », *Lexicometrica*, pp. 1-22

Durieux, Christine (1997), « La recherche terminologique en traduction : pour une approche hypertextuelle », *Meta : journal des traducteurs*, vol. 42, n° 4, 1997, pp. 677-684.

Froeliger, Nicolas, (2009), « A quoi bon enseigner la traduction technique » in *La traduction et ses métiers, aspects théoriques et pratiques*, colette Laplace *ed.*, Paris, Lettres modernes minard, pp. 199-210.

Froeliger, Nicolas, (2009), *De l'intérêt des métiers de la traduction aujourd'hui*, publié en ligne sur le site <http://alerteducation.eu/Declic>, mars 2009.

Gouadec, Daniel (2003), « Le bagage spécifique du localiseur/localisateur, le vrai «nouveau profil» requis », *Meta : journal des traducteurs*, vol. 48, n° 4, 2003, pp. 526-545.

Guidère, Mathieu (2002), « Alignement d'un corpus français-arabe et extraction d'un lexique bilingue » (Alignment of a French-Arabic Corpus), in *La Lettre d'information d'ELRA* (European Language Resources Association), vol. 7, n° 1, p. 6.

Guidère, Mathieu (2005), *La Traduction arabe, méthodes et applications. De la traduction à la traductique*, Paris, Ellipses.

Hamada, Salwa (2006), « Some approach for a Machine Translation Evaluation System (MTES) », *The 18th Saudi National Computer Conference (NCC18)*, Riadh.

Hamada, Salwa (2010), *Le traitement automatique de la langue arabe. Problèmes et solutions*, Le Caire, Dar Gharib.

Hattab, Mamoun *et al.* (2009), *2nd International Conference on Arabic Language Resources and Tools 22-23 Avril, Cairo, Egypt.*

Israël, Fortunato (1998), « Terminologie et traduction : le traducteur a-t-il besoin d'une formation terminologique ? » (Table ronde), *Quelle formation pour le traducteur de l'an 2000 ? Actes du colloque International tenu à L'ESIT les 6,7 et 8 juin, Paris, Didier érudition*, pp. 191-218.

Karsaklian, Eliane (2007), *Le Marketing international*, Paris, Eyrolles.

Kilgarriff, Adam and Grefenstette, Gregory (2003), « Introduction to the Special Issue on Web as Corpus », *Computational Linguistics, September, vol. 29, n° 3*, pp. 333-347

Le Gendre, Bertrand (2008), « Faut-il brûler Wikipédia ? » *Le Monde*, 16 Mars.

L'Homme, Marie-Claude (2008), *Initiation à la traductique*, 2^{ème} édition, Montréal, Lingua tech.

Lerat, P. (1995), *Les langues de spécialités*, coll. « Linguistiques Nouvelles », Paris, PUF.

Maniez, François (2001), « Extraction d'une phraséologie bilingue en langue de spécialité : corpus parallèles et corpus comparables », *Meta : journal des traducteurs*, vol. 46, n° 3, pp. 552-563.

Morin, Emmanuel *et al.* (2004), *Extraction de terminologie bilingues à partir de corpus comparables*, TALN, Fès, 19-21 avril.

Nikkhou, Mahtab and Choukri, Khalid (2004), *Report on Survey on Arabic Language Resources and Tools in the Mediterranean Countries*, ELDA, NEMLAR, Center for Sprogteknologi, University of Copenhagen, Denmark.

Ouerhani, Béchir (2008), *Les problèmes linguistiques de la traduction automatique des prédicats nominaux entre l'arabe et le français*, *Meta : journal des traducteurs*, vol. 53, n° 2, pp. 407-419.

Plassard, Freddie (2007), *Lire pour traduire*, Paris, Presses Sorbonne Nouvelle.

Said, Dina *et al.* (2009), *A Study of Text Preprocessing Tools for Arabic Text Categorization*, 2nd International Conference on Arabic Language Resources and Tools 22-23 April 2009, Cairo, Egypte.

Séléskovitch, Danica et Lederer, Marianne (2001), *Interpréter pour traduire*, 2^{ème} éd., Paris, Didier érudition.

Sanchez, xavier et Bonneville, Christinne. (2006), *Enseignement des langues : de la langue de spécialité à la traduction spécialisée sur corpus électroniques*, Association EPI, publié sur le site : www.epi.asso.fr/revue/articles/a0611b.htm. Page consultée le 14/01/2011

Zughoul, Muhammad et Abu-Alshaar, Awatef (2005), « English/Arabic/English Machine Translation: A Historical Perspective », *Meta: Translators' Journal*, vol. 50, n° 3, 2005, pp. 1022-1041