



CARTE ROUTIÈRE TECHNOLOGIQUE (CRT)

**RAPPORT PRÉSENTÉ AU
CENTRE DE RECHERCHE EN TECHNOLOGIES LANGAGIÈRES**

PRÉPARÉ PAR DAVIDE TURCATO

JANVIER 2006

TABLE DES MATIÈRES

1.	SOMMAIRE.....	7
2.	INTRODUCTION ET ANTÉCÉDENTS.....	33
2.1	Mission et vision	36
2.2	Envergure et limites de la carte routière	39
2.2.1	GESTION DU CONTENU	40
2.2.1.1	Définition et vue d'ensemble de la gestion du contenu.....	40
2.2.1.2	Ventilation de la gestion de contenu par technologies, applications et créneaux de marché.....	44
2.2.1.3	Le cycle de vie du contenu numérique.....	50
2.2.2	FORMATION LINGUISTIQUE	58
2.2.2.1	Évolution de la formation linguistique enrichie par les technologies au Canada.....	58
2.2.3	TRAITEMENT DE LA PAROLE.....	70
2.2.3.1	Industrie du traitement de la parole	70
2.2.3.2	Principales technologies de traitement de la parole.....	71
2.2.4	TRADUCTION.....	84
2.2.4.1	Vue d'ensemble du secteur de la traduction	84
2.2.4.2	Catégories principales de ressources et d'outils technologiques de traduction.....	88
2.2.4.3	Le processus de traduction	110
2.3	L'industrie aujourd'hui : ses produits, ses clients, ses fournisseurs et ses processus de fabrication.....	111
2.3.1	GESTION DU CONTENU.....	111
2.3.1.1	Organisations de gestion du contenu actives au Canada.....	111
2.3.1.2	Applications de gestion de contenu : études de cas.....	116
2.3.2	FORMATION LINGUISTIQUE	125
2.3.2.1	Formation linguistique enrichie par les technologies au Canada : Forces et limites	125
2.3.2.2	Relation entre la formation linguistique actuelle et l'industrie canadienne de la langue	135
2.3.2.3	Les chefs de file et les joueurs clés du marché, par région.....	144
2.3.2.4	Structures de soutien à la formation linguistique	160
2.3.3	TRAITEMENT DE LA PAROLE.....	169
2.3.3.1	Principales normes du traitement de la parole.....	169
2.3.3.2	Fournisseurs technologiques principaux.....	177
2.3.3.3	Applications de traitement de la parole.....	197
2.3.3.4	Recherche en traitement de la voix.....	206
2.3.4	TRADUCTION.....	230
2.3.4.1	Produits de technologie traductionnelle	230

2.3.4.2	Fournisseurs de services en matière de traduction et de localisation	250
2.3.4.3	Centres de recherche en technologie du langage.....	250
2.4	Tendances et prévisions du marché	253
2.4.1	GESTION DU CONTENU	253
2.4.1.1	C'est un monde dangereux.....	253
2.4.1.2	La génération de nouvelles recettes	257
2.4.1.3	Consolidation par l'acquisition de petites et de grandes entreprises	260
2.4.2	FORMATION LINGUISTIQUE	261
2.4.2.1	Résumé des tendances et des prévisions du marché au Canada et sur le plan international	261
2.4.3	TRAITEMENT DE LA PAROLE.....	273
2.4.3.1	Le marché de la technologie.....	274
2.4.3.2	Services	275
2.4.3.3	2004 Prévisions de la tendance dans le marché de traitement de la parole.....	281
2.4.4	TRADUCTION.....	282
2.4.4.1	Les mutations dans le marché de la traduction.....	282
2.4.4.2	Données générales sur la tendance à l'intérêt croissant à l'égard de la technologie de la traduction.....	284
2.4.4.3	Facteurs qui affectent le recours à la technologie de la traduction.....	286
2.4.4.4	Prévisions du marché de la traduction	293
2.4.4.5	Taux et salaires	302
3.	BESOINS ET POSSIBILITÉS TECHNIQUES	307
3.1	Produits ciblés	308
3.1.1	GESTION DU CONTENU	308
3.1.1.1	Applications de gestion du contenu	308
3.1.1.2	Ressources	312
3.1.1.3	Outils.....	313
3.1.2	FORMATION LINGUISTIQUE	313
3.1.2.1	Les besoins des étudiants des langues	313
3.1.2.2	Prestation linguistique.....	314
3.1.2.3	Compétences particulières.....	315
3.1.2.4	Secteurs mal desservis	321
3.1.3	TRAITEMENT DE LA PAROLE.....	322
3.1.4	TRADUCTION.....	323
3.1.4.1	Identification des professionnels utilisant les ressources et outils, identification des ressources et outils	324
3.1.4.2	Exploitation des applications de traduction	329
3.2	Notes de performance et caractéristiques de produit	332

3.2.1	GESTION DU CONTENU	332
3.2.1.1	Attributs cruciaux	332
3.2.1.2	Défis de commercialisation.....	332
3.2.1.3	Défis technologiques.....	333
3.2.2	FORMATION LINGUISTIQUE	336
3.2.2.1	Attributs cruciaux	336
3.2.3	TRAITEMENT DE LA PAROLE.....	337
3.2.3.1	Attributs cruciaux	337
3.2.4	TRADUCTION.....	337
3.2.4.1	Attributs cruciaux	337
3.2.4.2	Intégration des outils au processus de traduction	338
3.2.5	SYNOPSIS	347
3.3	Possibilités actuelles de la science et des technologies	350
3.3.1	GESTION DU CONTENU	350
3.3.1.1	État actuel de l'industrie canadienne de la gestion du contenu	350
3.3.2	FORMATION LINGUISTIQUE	355
3.3.2.1	La position du Canada dans un contexte global	356
3.3.3	TRAITEMENT DE LA PAROLE.....	360
3.3.4	TRADUCTION.....	362
3.4	Écarts et obstacles	365
3.4.1	GESTION DU CONTENU	365
3.4.2	FORMATION LINGUISTIQUE	365
3.4.2.1	Modèles novateurs	365
3.4.2.2	Applications de modèles pour les fournisseurs canadiens de formation linguistique	370
3.4.2.3	Les défis du Canada dans la mise en œuvre de ces modèles.....	373
3.4.3	TRAITEMENT DE LA PAROLE.....	378
3.4.4	TRADUCTION.....	379
3.4.4.1	Le défi du processus de traduction	379
3.5	Stratégie et objectifs de développement.....	381
3.5.1	APERÇU	381
3.5.2	GESTION DU CONTENU	382
3.5.2.1	Stratégie de développement.....	382
3.5.3	FORMATION LINGUISTIQUE	384
3.5.3.1	Ouverture des liens de communication – Intégration des éléments disparates.....	384
3.5.3.2	Viser l'or – À la recherche des ressources motrices du développement.....	387
3.5.3.3	Optimisation de la réputation du Canada	389

3.5.3.4	La normalisation par l'image de marque – Le concept d'une marque maîtresse de l'industrie linguistique	390
3.5.3.5	Rôles et responsabilités	392
3.5.4	TRAITEMENT DE LA PAROLE.....	393
3.5.4.1	EuroMAP – Un exemple à imiter	393
3.5.4.2	Recommandations du comité du sous-secteur du traitement de la parole	394
3.5.5	TRADUCTION.....	394
3.5.6	SYNOPSIS	395
3.5.6.1	Synergie.....	395
3.5.6.2	Visibilité.....	397
3.5.6.3	Retombées économiques	398
3.5.6.4	Catégorisation des logiciels	398
3.5.6.5	Conclusion.....	399
4.	STRATÉGIE DE DÉVELOPPEMENT DES TECHNOLOGIES.....	400
4.1	Évaluation et sélection des technologies	401
4.1.1	CRITÈRES D'ÉVALUATION ET DE SÉLECTION	401
4.1.1.1	Critères proposés	401
4.1.1.2	Synopsis	402
4.1.2	ÉVALUATION DES TECHNOLOGIES	403
4.1.2.1	Évaluation générale.....	403
4.1.2.2	Résultats sélectionnés de l'évaluation	416
4.1.3	SÉLECTION DES TECHNOLOGIES	432
4.1.3.1	Examen critique des données d'évaluation.....	432
4.1.3.2	Pollinisation croisée	433
4.1.3.3	Perspectives commerciales.....	436
4.2	Technologies recommandées.....	438
4.2.1	GESTION DU CONTENU	439
4.2.1.1	Élaboration d'un WordNet multilingue	439
4.2.1.2	Centre de soutien de langues croisées.....	441
4.2.2	TRAITEMENT DE LA PAROLE.....	443
4.2.3	TRADUCTION.....	443
4.2.3.1	« Traduction » en temps réel de la parole en parole dans une même langue ..	443
4.3	Plan de mise en œuvre et budget.....	444
5.	CONCLUSIONS.....	449
5.1	Recommandations	450
5.1.1	SOUS-COMITÉ DU TRAITEMENT DE LA PAROLE, ÉTAPE III : PROPOSITIONS DE PLAN D'ACTION, 6 MARS 2005.....	450
5.1.2	PRÉSIDENCES DE COMITÉ, RECOMMANDATIONS PRÉLIMINAIRES DE LA CARTE	

ROUTIÈRE TECHNOLOGIQUE EN MATIÈRE DE LANGUE, 9 NOVEMBRE 2005	451
5.1.2.1 Recommandations.....	451
5.1.2.2 Plan d’action.....	452
5.2 Plan de mise en œuvre des recommandations	454
6. BIBLIOGRAPHIE ET RÉFÉRENCES.....	456
6.1 Gestion du contenu.....	457
6.2 Formation linguistique.....	458
6.3 Traitement de la parole.....	461
6.4 Traduction.....	462
7. ANNEXES	464
7.1 Processus d’élaboration de la carte routière	465
7.1.1 DOCUMENTS INTÉRIMAIRES DE LA CRT	465
7.1.1.1 Biens livrables du 31 mars 2005.....	465
7.1.1.2 Réunion des présidents, 10 mars 2005, Montréal.....	466
7.1.1.3 Carte routière technologique (toile de fond)	466
7.1.1.4 Sous-comité de la gestion du contenu.....	466
7.1.1.5 Sous-comité de la formation linguistique	467
7.1.1.6 Sous-comité du traitement de la parole.....	467
7.1.1.7 Sous-comité de la traduction	467
7.2 Participants, collaborateurs et auteurs.....	469
7.2.1 GESTION DU CONTENU.....	469
7.2.2 FORMATION LINGUISTIQUE	469
7.2.3 TRAITEMENT DE LA PAROLE.....	469
7.2.4 TRADUCTION.....	469
7.2.5 AUTRES PARTICIPANTS, COLLABORATEURS ET AUTEURS.....	470

1. SOMMAIRE

Le présent document trace la Carte routière technologique (CRT) de l'industrie canadienne de la langue. La phase initiale de la Carte routière technologique, contenant les premières recommandations, a été terminée et présentée à Industrie Canada et au CNRC le 31 mars 2003. Une des recommandations du rapport initial portait sur la création d'une association représentant l'industrie canadienne de la langue : l'AILIA. À la suite de ce rapport, le CNRC a organisé, de concert avec Industrie Canada et l'AILIA, une série de tables rondes portant sur les technologies de la langue et l'industrie de la langue au Canada. En conséquence, l'AILIA a décidé de procéder à la création d'une Carte routière technologique pour l'industrie canadienne de la langue. Une Carte routière technologique se compose de trois grandes parties. En premier lieu, elle présente l'envergure de l'industrie actuelle, y compris les tendances et les prévisions de marché. En deuxième lieu, elle pose un regard sur les besoins des utilisateurs des technologies de la langue existantes, ainsi que sur les possibilités qui leur sont offertes. En troisième lieu, elle propose une stratégie de développement technologique. Ce type de stratégie s'attarde à des questions comme l'évaluation et la sélection de technologies, résumant les facteurs décisionnels et offrant un échéancier pour le développement de nouvelles technologies.

Mission et vision

Dans *Le prochain acte : un nouvel élan pour la dualité linguistique canadienne*, le gouvernement du Canada a formulé un Plan d'action pour les langues officielles, propulsé par la vision d'un pays qui bénéficie des avantages du bilinguisme et surmonte les obstacles parsemant la route qui y mène. Devant ces défis, le Plan d'action demande à l'industrie de la langue de jouer un rôle clé dans l'atteinte des objectifs qu'il énonce. Le Plan d'action aborde trois secteurs prioritaires : 1) l'éducation, concernant l'enseignement de la langue de la minorité et l'enseignement d'une langue seconde; 2) le *développement communautaire*, visant à améliorer l'accès des communautés aux services publics dans les deux langues officielles; et 3) *un service public exemplaire*, cherchant à améliorer l'offre de services fédéraux aux Canadiens dans les deux langues officielles, la participation des Canadiens anglophones et francophones au gouvernement fédéral et l'emploi des deux langues au travail. Le Plan d'action souligne la pertinence mutuelle et le renforcement de la dualité linguistique du Canada et de l'industrie de la langue. La dualité linguistique fait de l'industrie de la langue un facteur vital dans l'atteinte de la vision élaborée par le plan. En même temps, elle est considérée comme un atout pour l'avenir du Canada, et comme un avantage concurrentiel sur la scène internationale. La vision élaborée par le Plan d'action peut être adoptée comme sienne par le secteur des technologies de la langue. La dualité linguistique présente à la fois un avantage concurrentiel et un défi technologique pour les technologies de la langue : le premier inscrit le secteur canadien des technologies de la langue en meilleure position que ses concurrents, en matière de ressources linguistiques, etc. En même temps, en relevant avec succès le défi que représente le marché canadien, le secteur des technologies de la langue a de fortes chances de se positionner en amont de la concurrence et de réussir sur le marché international. À cet égard, la vision de « faire du Canada un chef de file mondial de l'industrie de la langue » adoptée par l'AILIA se concilie avec l'adoption d'une vision dérivée du Plan d'Action : en contribuant à la poursuite de cette vision et en s'attaquant aux défis imposés par la dualité linguistique du Canada, les technologies canadiennes de la langue ont de fortes chances de mettre en œuvre la vision précédente.

Envergure et limites de la carte routière

Lorsque le processus de Carte routière de l'industrie canadienne de la langue a débuté, au milieu de 2003, au terme d'études préliminaires et de discussions, le travail a été distribué entre quatre sous-comités, chargé chacun d'étudier un sous-secteur particulier du domaine des technologies de la langue et de prendre part au comité de coordination de niveau supérieur. Ces sous-secteurs sont : gestion de contenu, formation linguistique, traitement de la parole et traduction.

La *gestion de contenu* est l'interprétation et l'analyse automatiques du texte pour toute application appelée à générer et à transformer du contenu. Les sociétés ou les services de gestion de contenu au sein de sociétés de logiciel élaborent des technologies de traitement intelligent du texte. Bien sûr, le texte se retrouve dans les documents écrits, mais il peut aussi se retrouver en marge ou à l'intérieur d'autres types de contenu, tel que les images, les films et le son, grâce à quoi les technologies de gestion de contenu sont en mesure d'aider les humains à gérer tous les types de contenu électronique. La gestion de contenu et le type de problèmes concrets que celle-ci est appelée à résoudre peuvent être expliqués par la présentation du cycle de vie du contenu numérique. Au cours de son cycle de vie, le contenu numérique est soumis aux activités suivantes : *conversion*, y compris la reconnaissance de caractères optiques, la reconnaissance du texte manuscrit et les applications de conversion texte-parole; *édition*, y compris les générateurs d'idée, les générateurs de texte, les assistants à la rédaction, les outils de langue contrôlée, les vérificateurs d'orthographe et de grammaire et les applications de contrôle de la sécurité; *révision*, y compris les applications de contrôle des versions, les applications de contrôle d'accès et les applications de protection de la vie privée; *gestion*, y compris les extracteurs de métadonnées, les indexeurs automatiques, les moteurs de catégorisation automatique, les agents personnels, les moteurs de recherche intelligents et la gestion de contenu interne; *adaptation*, y compris le résumage, le résumage de langues croisées, les assistants à la traduction, les logiciels de cueillette d'information en langues croisées, les assistants à l'internationalisation et les assistants à la localisation; *distribution*, y compris les applications d'extraction de contenu, les applications de contrôle du routage, les outils de rédaction collaborative et les créateurs de liens automatiques; et *découverte et réutilisation*, y compris les systèmes de dialogue, les systèmes experts, les systèmes de questions et réponses, les logiciels de traduction automatique, les applications de découverte de connaissances, les systèmes d'analyse textuelle, les systèmes de veille stratégique, les systèmes de veille concurrentielle et les systèmes de gestion des relations avec la clientèle. De même, les produits de gestion de contenu peuvent répondre aux besoins de divers secteurs de marché, dont : *gestion sûre de contenu*, y compris les outils de gestion fondés sur les politiques et touchant la protection contre les virus, la protection contre les pourriels, la confidentialité du contenu, le contenu Web, le contenu des courriels et les pièces jointes, ainsi que l'exécution d'applications téléchargeables; *gestion documentaire*, y compris les systèmes et les stratégies de gestion des documents électroniques et papier; *gestion des dossiers*, couvrant toute la gamme des dossiers, des documents papier, objets, lettres et actes, aux dossiers électroniques, tels les fichiers de traitement de texte, les dessins d'ingénierie et les courriels; *gestion des connaissances*, le secteur préoccupé par l'identification et la correspondance d'actifs intellectuels au sein de l'entreprise, générant de nouvelles connaissances au sein de l'entreprise à des fins concurrentielles, etc.; *gestion des relations avec la clientèle*; *systèmes de répondeur pour courriel*; *veille stratégique*, *gestion du courriel* et *archivage du courriel*.

La *formation linguistique* comprend : anglais langue seconde (ALS), français langue seconde (FLS), enseignement et apprentissage des langues secondes, autres langues, langues supplémentaires, langues internationales/ancestrales et langues secondes. L'ALS/FLS est réglementé dans le système d'écoles publiques par les divers ministères de l'éducation; le programme d'enseignement en

ALS/FLS des écoles primaires et secondaires est administré par les commissions scolaires provinciales, qui établissent des lignes directrices en matière de résultats (ou d'objectifs d'apprentissage) pour chaque niveau de langue, ainsi que des critères de performance détaillés en vertu de l'état de la recherche actuelle. Les collèges et universités disposent de plus de marge de manœuvre dans l'établissement de leurs propres lignes directrices en matière d'examen, de classement et d'évaluation de leurs étudiants, dérivée de leurs longues recherches dans le domaine de l'apprentissage d'une langue seconde. La *formation linguistique du secteur des entreprises*, c.-à-d. l'intégration de la formation en anglais ou en français et de l'amélioration des connaissances dans ces deux langues, représente un autre secteur pertinent. Plusieurs organismes et entreprises ont entrepris d'élaborer leurs propres systèmes de formation afin de répondre à la demande d'employés hautement compétents, dictée par la concurrence des secteurs industriels. Une telle formation, bien que faisant rarement appel aux méthodes d'apprentissage d'une langue seconde, s'avère en général efficace, en autant qu'elle suive les normes établies en matière de formation linguistique (une fois de plus, régies par une concurrence saine sur le marché). Les programmes s'inscrivant dans cette catégorie et considérés ici comme programme de formation linguistique sont : cours de mise à niveau des compétences, séminaires et programmes professionnels à distance, offerts sous le vocable « cours de mise à niveau des adultes » et toute formation donnée en anglais ou en français. Le secteur commercial semble de plus en plus laisser les tendances du marché et la concurrence régir ses services. En fait, les organismes d'ALS du Canada ont établi leurs propres systèmes rigoureux de réglementation et d'accréditation de leurs membres (instituts de formation des enseignants et des personnes) et de certification des enseignants. Cela représente un net contraste avec nos plus gros concurrents aux États-Unis, où les normes d'accréditation sont réglementées par l'US Department of Education (Département fédéral de l'éducation), en contribution avec les États. En ce qui a trait à l'*enseignement des langues internationales – ou étrangères –* (autres que l'anglais et le français, les deux langues officielles du Canada), les commissions scolaires sont responsables d'offrir et de soutenir des cours de langues internationales (auparavant connues sous le vocable « langues ancestrales ») aux niveaux primaire et secondaire. Ces cours sont très importants, tant pour les élèves que pour le Canada dans son ensemble : chez les élèves, l'apprentissage d'une langue étrangère (en général la langue du pays d'origine de leurs parents) renforce le sentiment d'identité et l'estime de soi par la reconnaissance de leur identité culturelle, ce qui contribue à leur identité canadienne; pour le Canada, cette reconnaissance de son héritage linguistique et culturel par une personne favorise le civisme et la compréhension mutuelle au sein de la mosaïque canadienne du multiculturalisme.

Le *traitement de la parole* offre des possibilités importantes en matière d'amélioration de l'interaction entre humains et machines et entre humains *utilisant* des machines. Le traitement de la parole englobe le groupe d'applications, d'outils et de techniques et composants logiciels répliquant la capacité de l'humain à entendre, identifier et émettre un langage humain parlé. Le traitement de la parole peut aussi être enrichi par la technologie de traitement du langage naturel, afin de modéliser la capacité humaine à comprendre et à traiter le contenu du langage et permettre la traduction d'une phrase parlée d'une langue à une autre, en dans plusieurs autres applications linguistiques intelligentes. L'industrie du traitement de la parole se compose de la reconnaissance automatique de la parole, de la conversion texte-parole, de la biométrie vocale – leurs applications, plateformes et services. Les technologies de la parole peuvent s'appliquer à une vaste gamme de problèmes commerciaux et administratifs dans le but d'offrir de meilleurs services et de nouvelles solutions efficaces aux entreprises et à leurs clients. Plus particulièrement : la reconnaissance de la parole au téléphone est désormais largement utilisée; les solutions prenant en charge la parole ont commencé à remplacer des systèmes d'applications de centre d'appel; les sociétés de télécommunications déploient des applications de la parole comme hameçon afin d'empêcher leurs clients de changer de

fournisseur; le texte en parole est vastement répandu sur le Web, la biométrie vocale peut contrôler l'accès à des bureaux, des ordinateurs et des réseaux grâce à l'identification et à l'authentification vocale; la fonction de traitement intégré de la parole est offerte dans des outils de plus en plus petits et a prouvé sa rentabilité et sa profitabilité dans des applications telles la téléphonie cellulaire et la télématique (l'électronique de l'automobile); des applications de conversion texte-parole, de texte en parole et de parole en parole ont fait leur apparition dans le domaine de la traduction; la montée de la transcription hors contexte dépendante et indépendante du locuteur, des applications de sous-titrage et de forage sonore (par ex., indexation MPEG7) voient le jour. Le sous-comité de traitement de la parole s'est entendu sur la définition suivante des technologies de base en matière de traitement de la parole : reconnaissance automatique de la parole, dépendante et indépendante du locuteur; synthèse de la parole, y compris la génération de texte en parole et de langage; biométrie vocale, y compris la vérification de l'identité du locuteur (c.-à-d. l'authentification), l'identification du locuteur et la détection de mensonges; classification sonore (silence/bruit/musique/parole), y compris la compréhension du langage naturel et le repérage du sujet; la gestion du dialogue; des outils et applications; et des services et plateformes de reconnaissance de la parole. Le traitement de la parole englobe en outre les catégories suivantes, situées au-delà du mandat actuel du sous-comité : compression de la parole (voix sur IP, MP3); traitement des signaux numériques de faible niveau (annulation de l'écho, suppression du bruit); et production de messages sonores.

La *traduction* se voit attribuer une définition à la fois étroite et plus vaste. Dans son sens le plus étroit (c.-à-d. la traduction en elle-même), elle désigne spécifiquement la discipline dans laquelle un message écrit est transféré d'une langue à une autre. Cependant, la traduction s'entend aussi dans un sens plus large, incluant un nombre de disciplines et d'activités diverses – telles l'interprétation, la terminologie et, plus récemment, la localisation – liées d'une façon ou d'une autre à la traduction. Là où il y a lieu dans le présent rapport, les termes « traduction » et « traducteur » désignent exclusivement l'ensemble de ces disciplines et professions connexes. Les professions contribuant au secteur de la traduction peuvent être regroupées parmi les activités suivantes : traduction, révision, interprétation, terminologie, localisation, postédition, traduction audiovisuelle, gestion de la traduction, formation en traduction, recherche en technologies de la langue et rédaction technique et professionnelle. Les ressources développées par les traducteurs ou les chercheurs en technologies de la langue et principalement utilisées par le secteur de la langue comprennent : les banques terminologiques, les ressources de corpus textuels précompilées et les portails de traduction. Les technologies et outils particuliers à l'industrie de la langue et utilisées largement par le secteur de la traduction comprennent : outils de reconnaissance vocale; outils d'analyse et de diagnostic documentaire; systèmes de gestion terminologique; outils de dépouillement terminologique; outils de vérification de la langue contrôlée et de prédiction; outils d'alignement; concordanciers bilingues; systèmes de mémoire de traduction; outils de localisation; systèmes de traduction automatique; technologie de traduction « intergicelle »; poste de travail du traducteur. La majorité des technologies et outils énumérés sont mieux décrits par l'expression « outils de traduction assistée par ordinateur » (TAO) et leur objectif est de soutenir ou d'assister le traducteur dans sa tâche. La traduction automatique constitue l'exception : elle procède à la tâche de traduction en elle-même, bien qu'une intervention humaine puisse être nécessaire afin d'améliorer la qualité de la traduction.

L'industrie aujourd'hui : ses produits, clients, fournisseurs et processus de fabrication

Étant donné la diversité des technologies et des produits du secteur de la gestion de contenu, il est difficile d'établir une liste des entreprises œuvrant dans ce domaine. En outre, ces listes ont une faible longévité. Néanmoins, au moins 37 sociétés se consacrant principalement à la gestion du contenu peuvent être identifiées au Canada. Les termes clés figurant le plus fréquemment en

association avec la description des activités de telles sociétés sont : logiciel antipourriel, gestion des relations avec la clientèle, gestion du courriel – ainsi que les autres activités connexes au courriel, comme la conformité des courriels, l’archivage, le filtrage et les systèmes de répondeur –, correcteurs, gestion documentaire et traitement du langage naturel, en plus du terme général « gestion de contenu ». En ce qui concerne les organismes de recherche œuvrant au Canada dans le secteur de la gestion de contenu, il est possible d’identifier sept centres de recherche, principalement associés aux universités. Pour ce groupe, les termes clés les plus fréquents sont : traitement du langage naturel, extraction d’information, cueillette d’information et reconnaissance de la parole; d’autres termes clés pertinents comprennent : intelligence artificielle, génération automatique du texte, forage de données, regroupement documentaire, systèmes et logiciels intelligents, représentation des connaissances, programmation logique et fonctionnelle, analyse syntaxique, catégorisation du texte et forage de texte. Il est possible d’illustrer l’application des technologies de gestion de contenu aux différentes activités du cycle de vie du contenu numérique par une série d’études de cas dont chacune est associée à une activité particulière du cycle de vie : les technologies de conversion texte-parole pour un meilleur accès et une plus grande productivité, comme cas de conversion du contenu numérique; l’emploi d’outils langagiers pour l’enseignement d’une langue seconde ou la génération automatique de contenu afin de mieux assister les acheteurs, comme cas d’édition; la création d’un Web sémantique pour faciliter la recherche d’éléments, comme cas de gestion; un centre de soutien multilingue, comme cas d’adaptation; et l’amélioration de la biosurveillance à l’aide de la gestion de contenu ou l’amélioration de la veille concurrentielle, comme cas de découverte et de réutilisation.

La *formation linguistique* désigne l’apprentissage linguistique et les services et produits de formation offerts par les secteurs gouvernemental et non gouvernemental, les collèges et universités, de même que par un secteur commercial très important mais fragmenté. La recherche sur l’évolution de la formation linguistique enrichie par les technologies au Canada et sur l’utilisation des technologies en tant qu’outil indique que le Canada est un chef de file dans ses recherches sur le bilinguisme et que l’expertise canadienne repose principalement sur le secteur des universités ouvertes et des réseaux de formation à distance, qui offrent de la formation en anglais, ainsi que des cours connexes en informatique, en enseignement et en commerce. Les sociétés canadiennes investissant dans les infrastructures technologiques visent à répondre à leurs besoins en matière de communication ainsi qu’en matière de transactions commerciales électroniques. La formation de leur effectif a évolué de façon constante, particulièrement suite à l’exemple de réussite de sociétés de premier rang du secteur de la formation électronique; cependant, il n’existe aucune donnée portant sur la dominance du marché par un seul joueur. Le Canada possède aussi une infrastructure de formation en ligne très bien développée, quoique fragmentée et consacrée au développement et à la publication de contenu. Le secteur de la formation linguistique doit capitaliser sur cette infrastructure et créer les partenariats nécessaires avec le reste des parties intéressées (gouvernement, institutions d’enseignement supérieur). L’industrie de la langue est extrêmement fragmentée, avec de petites et moyennes entreprises offrant de la formation linguistique et des services-conseils. Ces sociétés composent un secteur bien positionné dans l’économie canadienne et exportent leurs services à l’étranger, tout en invitant des étudiants étrangers à visiter et à étudier au pays. Les tendances internationales entraîneront probablement des effets positifs sur l’industrie des services, étant donné que de plus en plus de gens cherchent à apprendre et à enseigner l’anglais – la lingua franca d’Internet.

Le monde du traitement de la parole a fait l’objet de grands perfectionnements depuis 1999. Au cours des dernières années, les moteurs de reconnaissance de la parole eux-mêmes se sont améliorés, mais les progrès énormes de la *convivialité des technologies* constituent la plus importante contribution à ce marché. En 1999, les utilisateurs qui investissaient dans l’élaboration d’un système de parole avaient plus de chance d’échouer que de réussir à la première tentative, étant donné que les

applications étaient difficiles à utiliser. Aujourd'hui, les chances de réussite sont aussi importantes que les retombées potentielles. Il existe une foule d'applications de traitement de la parole dont la liste suivante en cite quelques exemples : réponse vocale interactive (centres d'appels et préposés automatiques); portails téléphoniques vocaux; portails Web vocaux; messagerie unifiée (messagerie vocale activée vocalement, etc.); reconnaissance intégrée de la parole (téléphonie cellulaire et ANP); télématique (automobiles); dictée informatisée; contrôle et commande des applications et systèmes d'exploitation informatiques; sécurité (détection de la fraude, vérification de l'identité du client, contrôle d'accès); domotique; contrôle informatique pour les personnes souffrant d'un handicap; assistance-annuaire; applications de composition d'appels sortants; applications de diffusion de nouvelles; transcription automatique, indexation et recherche de documents audiovisuels (MPEG7); formation linguistique assistée par le traitement de la parole; traduction parole en parole; détecteur de mensonge; doublage et postproduction audionumérique. La part du lion du marché du traitement de la parole revient à la réponse vocale interactive (RVI) et aux centres d'appels. Les services financiers, y compris les banques, les sociétés d'assurances et les courtiers, sont les principaux utilisateurs des technologies de reconnaissance de la parole. Parmi les autres segments de marché clés, nous retrouvons : les télécommunications; les soins de la santé; le tourisme; la fabrication et la vente au détail; les jouets et appareils électroniques (parole intégrée); la télématique (traitement de la parole en véhicule); les services publics; le gouvernement; la dictée; et le contrôle et la commande bureautiques. Par exemple, Bell Mobilité, la société sans fil de Bell Canada, a intégré une application canadienne dans ses services. Elle a lancé dernièrement le premier portail vocal au Canada, permettant ainsi à ses abonnés sans fil d'accéder à de l'information Web et à leurs courriels à l'aide de leur téléphone cellulaire numérique ou analogique. Il existe au moins 22 fournisseurs clés canadiens, à savoir des sociétés qui vendent une technologie, des produits, des applications de plateforme ou des services de traitement de la parole. Sur le plan de la recherche, les secteurs principaux de la recherche en matière de traitement de la parole comprennent : classification et segmentation sonore; reconnaissance de la parole en environnement bruyant; reconnaissance d'un vaste vocabulaire; transducteur d'état arithmétique; sous-titrage automatique; transcription automatique; indexation et recherche de documents audiovisuels; extraction d'information (forage sonore); traduction (parole en parole); formation linguistique (traitement du langage naturel et traitement de la parole); « lecture sur les lèvres » pour améliorer l'exactitude de la reconnaissance automatique de la parole; agents conversationnels; biométrie vocale; synthèse de la parole et génération du langage naturel. Il n'existe pas moins de 12 laboratoires de recherche consacrés au traitement de la parole.

Il est possible d'établir la liste des technologies de traduction au Canada à partir d'un bref inventaire des outils offerts sur le marché et des sociétés qui leur sont associées. La liste comprend : les outils de gestion et de dépouillement terminologiques (tels *Dictionary Generator*, *SynchroTerm* et *XTracTerm* de BridgeTerm; *EDITerm* d'EDIT Inc.; *LogiTerm* et *LogiTermWeb* de Terimnotix Inc.); les outils d'alignement (tels *Aladin* de RALI; et *Alignment Robot* de Terminotix Inc.); les concordanciers bilingues (tels *Find* de Beetext; *LogiTerm* et *LogiTermWeb*, déjà mentionnés; et *TransSearch* de RALI); les outils de mémoire de traduction (tels *Fusion Translate* d'Orca; et *LogiTrans* de Terminotix Inc.); les logiciels intermédiaires des technologies de la traduction (tels *BabelTRACK* de BabelFish; *Coordinator's Dashboard* de BridgeTerm; et *Flow* de Beetext); les systèmes de traduction automatique (tels *XLT* de Socatra Inc.); et les postes de travail de traducteur (tels *Lingua Technologies Toolbox* de Lingua Technologies Inc.; *MultiTrans* de MultiCorpora R&D Inc.; et *Promemoria* de BridgeTerm). On dénombre au moins 16 fournisseurs de services de traduction et de localisation au Canada. Les centres de recherche en technologies de la langue sont au nombre de quatorze, au minimum. De ce nombre, tous ne sont pas associés à des universités :

certain, dont ALIS et BabelFish, sont des fournisseurs de services de localisation; d'autres, dont Beetex, BridgeTerm, Lingua Technologies, Socatra et Terminotix, sont des sociétés de production d'outils commerciaux de traduction.

Tendances et prévisions de marché

Le marché de la gestion de contenu est demeuré flou pendant longtemps. Par exemple, les technologies de gestion de contenu permettent le résumage automatique du texte, mais cela ne signifie pas qu'il y ait une demande réelle pour cette fonction. À l'inverse, la cueillette d'information en langues croisées demeure à ce jour une solution imparfaite à un problème concret à résoudre. Quels sont les problèmes que les utilisateurs souhaitent résoudre aujourd'hui et que les technologies de gestion de contenu permettraient de résoudre? Voici une liste de tendances qui ont donné lieu ou qui pourraient donner lieu à des occasions commerciales pour les entreprises de gestion de contenu. Ces tendances peuvent être divisées en trois grandes catégories. La première catégorie concerne la sécurité et la protection dans un monde dangereux et comprend : les logiciels antivirus, surtout en raison des risques provenant d'Internet; le flux constant de pourriels; le secteur de la défense et de la sécurité, surtout aux États-Unis, où la possibilité d'explorer de vastes quantités de renseignements est vitale pour la détection et l'évaluation de menaces potentielles; la nécessité de la conformité aux lois et règlements, en réponse aux mesures sévères exigeant plus de transparence des entreprises, afin de protéger les renseignements confidentiels de leur clientèle; et la protection des biens commerciaux. La deuxième catégorie concerne la génération de nouveaux revenus et comprend : l'explosion du commerce en ligne; les services à la clientèle automatisés, étant donné le flot important de courriels généré par le commerce en ligne, surtout aux États-Unis; la montée du contenu multimédia; la nécessité d'innover en matière d'outils de veille concurrentielle pour les sociétés nord-américaines et européennes souhaitant préserver l'avantage concurrentiel sur leurs rivaux des pays en voie de développement; et l'accès mobile. La troisième et dernière catégorie concerne la consolidation des sociétés de gestion de contenu par l'acquisition de petites et de grandes sociétés. Plutôt que de recourir à une capitalisation par les banques et les fonds d'investissement pour devenir des sociétés ouvertes autosuffisantes grâce aux investissements, aux actionnaires et aux clients, les sociétés de GC peuvent mettre en marché leurs produits en étant achetées par de grandes sociétés à la recherche de nouveaux produits et de nouveaux revenus.

Pour le secteur de la formation linguistique, il est prévu qu'en 2005 90 % de toute la formation au Canada se fera par Internet. Un autre facteur pertinent réside dans le fait que la formation en ligne attire les investisseurs. La formation en ligne permet aux sociétés de diversifier leurs sources de revenus et de capitaliser sur le potentiel commercial d'Internet. En retour, cette stratégie attire les investisseurs et entraîne la croissance du secteur. À long terme, les industries de la langue devraient tirer parti de ce développement de la formation en ligne. Étant donné que l'infrastructure sera en place et que les coûts de l'équipement auront tendance à diminuer, on anticipe que la concurrence obligera les sociétés de formation en ligne à former des partenariats avec l'industrie de la langue et à attirer de nouveaux marchés. La grande diversité des langues au Canada peut aussi représenter un actif puissant pour les industries de la langue émergentes. Malgré le fait que la langue prédominante d'Internet soit l'anglais, il existe une grande demande pour d'autres langues, au Canada et au niveau international. Au niveau international, on s'attend à ce que la formation sur le Web ou en ligne devienne la méthode d'enseignement et de formation privilégiée par tous les secteurs de l'industrie et celle qui connaîtra la plus forte croissance. Ces prévisions sont justifiées en partie par l'expansion des technologies. Les possibilités de formation en ligne se multiplient à la vitesse d'Internet. La bande passante représentait une limite importante à la formation en ligne, mais elle ne sera plus un frein bientôt grâce aux nouvelles technologies. La croissance de l'industrie est également motivée par

les énormes économies que la formation en ligne offre aux entreprises. CDI, important fournisseur d'analyses industrielles, prévoit que les revenus commerciaux provenant de la formation en ligne atteindront 23,7 milliards de dollars en 2006, par rapport à 234 millions de dollars en 1997. Il semble juste d'affirmer que les entreprises sont de plus en plus conscientes des avantages de la formation en ligne.

La convergence des technologies de l'information, de la téléphonie et des appareils électroniques personnels, doublée du besoin d'une interface utilisateur de rechange, stimule la demande pour des systèmes de traitement de la parole. Des investissements continus en recherche et développement, combinés à une infrastructure existante qui réduit les coûts de pénétration du marché, propulseront cette industrie vers une croissance impressionnante. Si certains segments des industries des télécommunications américaine et canadienne ont dû faire face à des défis économiques importants, les dépenses globales aux États-Unis ont augmenté de 4,9 pourcent en 2003. Les augmentations à deux chiffres des services sans fil, des services de soutien d'équipement et des services spécialisés comme les communications unifiées, la conférence audiovisuelle et l'accès Internet haute vitesse font contrepoids aux diminutions des dépenses d'équipement liées aux revenus des services locaux et interurbains. Au cours de la période de 2004 à 2007, la croissance de l'industrie des télécommunications américaine se fera à un taux annuel composé prévu de 9,2 % pour atteindre un milliard de dollars. Le traitement de la parole fera partie intégrante de cette croissance. La mise en marché des technologies de traitement de la parole a connu un cycle complet : s'adressant d'abord aux utilisateurs « innovateurs », puis aux « utilisateurs du début », ces technologies ont maintenant rejoint le grand public. Sur le marché des technologies de première génération, les utilisateurs innovateurs avaient des besoins très précis (par ex., appels à frais renversés 0+, assistance-annuaire 411) ou souhaitaient innover (par ex., Reuter a utilisé très tôt la technologie de traitement du langage naturel afin de recueillir de l'information pour son service de nouvelles). Les utilisateurs du début ont exploité la maturité croissante du traitement de la parole afin de générer des revenus et d'améliorer leurs produits (par ex., automatisation des centres d'appel, portails vocaux et télématiques). Le traitement de la parole est aujourd'hui connu du grand public, étant déployé auprès de la majorité des utilisateurs précoces. La reconnaissance de la parole au téléphone est désormais largement utilisée; les solutions prenant en charge la parole ont commencé à remplacer des systèmes d'applications de centres d'appel; les sociétés de télécommunications déploient des applications de la parole comme hameçon pour empêcher les clients de changer de fournisseur; le texte en parole est grandement répandu sur le Web; le traitement intégré de la parole est offert sur des appareils de plus en plus petits et a prouvé sa rentabilité. De plus, nous assistons à la montée de la transcription hors contexte dépendante ou indépendante du locuteur et à l'émergence d'applications de forage sonore (par ex., indexation MPEG7). L'Allied Business Intelligence (ABI) prévoit que le marché international de la reconnaissance de la parole connaîtra une croissance de 991 millions de dollars en 2003, comparativement à 677 millions de dollars en 2002. Les faits démontrent que les revenus de 2003 se chiffraient plutôt aux environs de 800 millions de dollars, selon Datamonitor (janvier 2004). À long terme, on prévoit que le marché de la reconnaissance de la parole atteindra 5,3 milliards de dollars en 2008 (ABI, 2 novembre 2003). On estime que le marché canadien du traitement occupe entre 10 et 20 % du marché mondial.

La croissance annuelle anticipée du marché mondial de la traduction est de l'ordre de 15 à 20 %. Si on se fie aux prévisions d'ABI à l'effet que l'effectif de traduction croîtra à un taux annuel qui varie entre 3 et 6 %, les traducteurs devraient de plus en plus se tourner vers les technologies afin d'accroître leur productivité et répondre ainsi à la demande croissante en matière de traduction. Le créneau de la traduction technique, en particulier, devrait connaître une grande croissance. On évalue qu'entre 8 et 10 % des documents techniques sont traduits. Mais dans l'avenir, cette proportion va

augmenter pour se chiffrer aux alentours de 25 à 30 %. Les secteurs où les besoins de traduction se font le plus sentir sont les services commerciaux, la finance, les hautes technologies, l'aéronautique et les soins de la santé. Étant donné que les documents techniques et spécialisés sont particulièrement propices à la traduction assistée par la technologie, la demande toujours plus grande pour ce type de traduction entraînera sûrement une utilisation plus grande des outils de traduction. La traduction se transforme en activité économique d'importance grandissante, devenant un facteur contribuant à l'augmentation des ventes, des revenus et des emplois pour les entreprises. En parallèle, la qualité de la traduction devient une mesure de capacité concurrentielle. Étant donné que les technologies de la traduction ne font pas qu'augmenter la productivité mais qu'elles permettent aussi une augmentation de la qualité (par ex., en favorisant la cohérence), les outils de traduction seront probablement de plus en plus recherchés. Les prévisions pour le marché mondial intégral de la traduction ont été calculées avec l'ajout des autres données présentées par ABI (2002) pour les marchés suivants : traduction humaine, localisation de logiciels, localisation de sites Web et traduction automatique. Ces données combinées indiquent qu'en 2001, le marché intégral de la traduction représentait environ 8,8 milliards de dollars américains. Les prévisions modérées font état d'une progression qui atteindra 13,6 milliards de dollars américains en 2007, tandis que les prévisions optimistes parlent plutôt de 18,1 milliards de dollars américains en 2007. En bref, on s'attend à ce que la croissance des revenus de la traduction se poursuive (au moins jusqu'en 2007). La traduction humaine continuera de détenir une avance considérable sur la traduction automatique (probablement en raison de la faible qualité des textes issus de la traduction automatique). Cependant, la traduction automatique dispose d'un potentiel de croissance exponentiel, surtout dans la mesure où il est possible de convaincre les utilisateurs d'accepter une qualité moindre pour certains types de traduction et de réaliser que la traduction automatique et la traduction humaine sont chacune pertinentes à différentes applications. La localisation est le créneau qui connaît la croissance la plus rapide et celui dont la dépendance envers les technologies est la plus grande. En ce moment, la localisation de logiciels détient une avance considérable sur la localisation de sites Web, bien que cette dernière continue de croître rapidement, réduisant ainsi l'écart qui les sépare.

Produits ciblés

Au cours des prochaines années, l'industrie canadienne de la gestion de contenu devrait tenter de créer des produits qui, soit répondront à des besoins importants du marché, soit viendront en aide aux sociétés et aux centres de recherche en matière de gestion de contenu pour les aider à mettre au point de meilleures applications de gestion de contenu. Ces produits peuvent être divisés en applications, ressources et outils. Les applications ciblées comprennent les systèmes de cueillette d'information de langues croisées, les systèmes de dialogue; les logiciels intelligents de conversion texte-parole, les foires aux questions multilingues, la cueillette d'information non structurée, l'analyse textuelle et les générateurs de texte. Les ressources ciblées comprennent : les terminologies multilingues, les ontologies, les corpus parallèles annotés et les corpus écrits. Les outils ciblés comprennent les outils d'acquisition et de maintenance des ontologies, les outils d'acquisition et de maintenance des terminologies et les moteurs de recherche de corpus alignés.

Un ensemble de solutions de technologies de formation linguistique, représentant le plus vaste potentiel pour les fournisseurs canadiens de formation linguistique et cadrant bien dans le contexte canadien, peut être davantage subdivisé en applications (ou plateformes), ressources et outils. Les applications ou plateformes pertinentes comprennent les applications en ligne pour des écoles de formation linguistique entièrement personnalisée, les laboratoires de langue numériques virtuels, les ressources de perfectionnement professionnel en ligne destinées aux enseignants et gestionnaires, les logiciels d'analyse de la parole et de rétroaction sur appareils portatifs, les systèmes d'évaluation en

ligne conformes aux normes nationales, les applications de formation linguistique en ligne par abonnement pour les appareils portatifs, les applications de texte en parole, les applications de création d'évaluations, les cours TESL en ligne, les outils de création de didacticiels et de systèmes de gestion de la formation, les applications de développement des compétences conformes aux normes nationales, les systèmes de gestion et d'évaluation de l'avancement conformes aux normes nationales. Les ressources pertinentes comprennent les sites Web offrant des plans de cours aux enseignants ou des activités aux élèves, les sites Web par abonnement offrant aux enseignants le matériel d'apprentissage conforme à un programme d'enseignement national, les médias importants offrant de l'information et des ressources au milieu de l'éducation, les réseaux Internet mondiaux de jumelage d'élèves, les répertoires sur Internet des cours à l'étranger offerts aux élèves d'ici et d'ailleurs, les bibliothèques de documentation en ligne avec liens vocabulaires individuels ou modélisation de la parole, les référentiels en ligne sur l'industrie de la formation linguistique, avec services-conseils, les ressources pédagogiques normalisées visant à aider les enseignants à utiliser la technologie, les dictionnaires multilingues sur téléphone cellulaire ou ANP pour la traduction et la formation linguistique, les sites de ressources linguistiques par abonnement destinés aux enseignants d'ALS, les dictionnaires en ligne, les dictionnaires interactifs et les sites Web d'accompagnement aux manuels, les sites d'enrichissement de texte et le soutien didactique en ligne.

Les produits ciblés du secteur du traitement de la parole sont mieux subdivisés entre technologies et applications. En retour, les technologies pertinentes dans le champ d'application du traitement de la parole peuvent être regroupées en catégories comme suit : *reconnaissance de la parole*, y compris la reconnaissance intégrée de la parole, la reconnaissance de la parole dépendante du locuteur, la reconnaissance de la parole indépendante du locuteur, la reconnaissance de la parole de vaste vocabulaire, la reconnaissance de la parole en environnement bruyant, l'extraction d'information (forage sonore), le repérage du sujet, l'analyse de maîtrise de la parole et la « lecture sur les lèvres » afin d'augmenter le niveau d'exactitude de la reconnaissance automatique de la parole; *traduction de la voix en texte*, y compris la synthèse de la parole et la concaténation de la parole; *biométrie vocale*, y compris l'identification du locuteur, la vérification de l'identité du locuteur et la détection de mensonge; *traitement de signaux*, y compris la classification sonore (silence, bruit, musique ou parole) et la compression de la parole; et *traitement du langage naturel*, y compris la génération du langage naturel, la génération du langage et les agents conversationnels. Les applications ciblées, selon les cinq catégories de technologie ci-dessus, dans une certaine mesure, comprennent : les centres d'appels et portails Web, la réponse vocale automatique, l'assistance-annuaire, la composition d'appels sortants, la parole en véhicule, les tests d'aptitude, la dictée informatisée, les commandes et contrôle informatiques, la diffusion de nouvelles, la transcription, la messagerie unifiée, la reconnaissance intégrée de la parole, la sécurité, la domotique, la technologie d'assistance, l'indexation et la recherche audiovisuelles, la traduction de parole en parole, le doublage, la postproduction audiovisuelle et le sous-titrage.

Les produits de traduction ciblés peuvent également être répartis entre les applications et les technologies. Cela étant, les applications et les technologies peuvent elles-mêmes être subdivisées en trois catégories principales : *traduction automatique*, *mémoire de traduction* et *flux des travaux*. Les applications de traduction automatique peuvent être classées plus avant comme suit : installation client, Web, moteurs, forage de données et condensé textuel. Les applications de mémoire de traduction peuvent être classées plus avant comme suit : installation client, Web, recherche, alignement, optimisation de l'utilisation (par fichiers, société ou industrie), et conversion de fichiers. Les systèmes de flux de travaux sont classés plus avant comme suit : installation client pour la traduction, Web pour la traduction; et adaptés pour la traduction. En parallèle, les technologies de traduction automatique peuvent être caractérisées par les éléments suivants : unilinguisme,

multilinguisme, moteurs, entrée contrôlée, entrée non structurée, extraction d'information (forage de données), bidirectionnelles, installation sur place, intégrations (logiciel-outil), Web et identification des erreurs de la traduction humaine. Les applications de mémoire de traduction se caractérisent par les éléments suivants : multilinguisme, interface multilingue, Web, installation sur place, intégrations (logiciel-outil), alignement de phrases, optimisation de l'utilisation du contenu (au niveau des fichiers, de la société ou de l'industrie). Les technologies de flux de travaux se caractérisent par les éléments suivants : particulières à la traduction, systèmes adaptés, particulières à une société, installation sur place et Web. En outre, deux classes de technologies supplémentaires sont ciblées : *terminologie* et *ressourcement*. Les technologies terminologiques se caractérisent comme suit : paires de langues, interface multilingue, Web, installation sur place, forage de données et dépouillement terminologique et technologies particulières à l'industrie. Enfin, les technologies de ressourcement se caractérisent comme suit : fournisseurs multilingues (FM), fournisseurs unilingues (FU) et traduction directe en ligne.

Notes de performance et caractéristiques de produit

Le tableau suivant présente une vision statistique des attributs cruciaux identifiés comme étant les plus pertinents par chaque sous-comité de la CRT afin d'évaluer et de sélectionner les produits ciblés. Afin de mettre en évidence les éléments communs entre les secteurs, on a indiqué en gris les attributs communs à tous les sous-secteurs ou à au moins trois sous-secteurs. Les attributs originaux sont indiqués comme chaque sous-comité les a présentés, même dans les cas où le regroupement de divers attributs aurait été plus simple. Dans ces cas particuliers, pour mettre les attributs en gris, il a fallu traiter les attributs différents comme un seul. C'est le cas notamment des attributs « pollinisation croisée » et « applicabilité au sous-secteur x », aussi séparés par une ligne pointillée plutôt que par une ligne solide, afin de souligner leur communauté d'éléments. Il convient également de noter qu'en ce qui concerne la gestion de contenu et la formation linguistique, les attributs choisis s'appliquent également à toutes les sous-classes du produit (c.-à-d.. applications, ressources et outils).

ATTRIBUTS CRUCIAUX	SOUS-SECTEUR					
	Gestion du contenu	Formation linguistique	Traitement de la parole		Traduction	
			Technologie	Applications	Technologie	Applications
Mise en œuvre	✓					
Marché intérieur	✓	✓	✓	✓	✓	✓
Marché international	✓	✓	✓	✓	✓	✓
Facilité d'utilisation		✓		✓		✓
Facilité d'intégration			✓		✓	
Élaboration du modèle			✓			
Mise au point de la grammaire			✓			
Coût de développement	✓	✓		✓		✓
Perfectionnements					✓	
RCI utilisateur final	✓	✓	✓	✓	✓	✓
Conformité aux normes	✓	✓	✓	✓	✓	✓
Durée de vie	✓	✓	✓	✓	✓	✓
Pollinisation croisée	✓	✓				
Applicabilité au sous-secteur x			✓		✓	
Applications et outils			✓		✓	
Forces actuelles	✓					
Spécificité de la FL		✓				
Modèle de revenu		✓		✓		✓
Maturité		✓	✓		✓	

Légende : | : commun aux quatre sous-secteurs; | : commun à trois sous-secteurs.

Possibilités actuelles de la science et des technologies

Il est possible d'évaluer l'état actuel de l'industrie canadienne de la gestion de contenu en tenant compte de quatre grandes catégories de facteurs incontournables en matière d'avantage concurrentiel : *état des facteurs*, incluant les ressources humaines, l'argent, les connaissances et les ressources linguistiques; *état de la demande*; *industries connexes et de soutien*; et *stratégie, structure et rivalité des sociétés*. Les deux premiers facteurs – état des facteurs et état de la demande – comportent des lacunes importantes. En ce qui concerne le quatrième facteur décisionnel, il vaut la peine d'en souligner certains aspects : les petites et moyennes entreprises de gestion de contenu ont souvent de la difficulté à maîtriser les aspects du développement de produit non connexes à la gestion de contenu; il semble que plus la concurrence est forte dans une industrie nationale, plus cette industrie s'avère concurrentielle à l'échelle internationale; certaines sociétés de gestion de contenu ont commencé à offrir des services linguistiques afin d'équilibrer leur budget car elles sont incapables d'atteindre la rentabilité simplement par la vente d'outils automatisés; certaines sociétés collaborent parfois ensemble lorsqu'elles ne sont pas directement rivales; enfin, la plupart des sociétés commercialisent des produits plutôt que des technologies OEM.

Les organismes canadiens ont créé certains produits novateurs qui ont connu un succès dans le secteur de la formation linguistique. Cependant, ils doivent faire face à des obstacles structurels, surtout attribuables à l'absence d'un corps d'enseignement fédéral au gouvernement qui permettrait d'en intégrer les diverses composantes. Bien qu'il existe certaines entreprises solides en matière de formation linguistique au pays, peu d'entre elles peuvent être reconnues mondialement. Les enseignants de langues, dans leur ensemble, sont réticents à adopter la technologie en tant que support efficace, et certains même considèrent les dépenses en ressources, nécessaires au développement ou à l'utilisation de la technologie, comme un obstacle important. Il existe ici un manque de connaissance généralisé des développements en cours au sein des autres créneaux de l'industrie de la formation linguistique, ainsi qu'au sein des autres sous-secteurs de l'industrie du langage, alors que dans d'autres pays, on a adopté une structure de communication plus ouverte. Bien qu'un certain nombre d'organismes particuliers à ce segment au Canada établissent et respectent un ensemble de normes, il n'existe aucune norme générale permettant de les unifier. Même dans les pays où la situation est similaire, on a réussi à établir des groupes de coordination qui offrent l'apparence de normes nationales, facilitant ainsi la création de produits technologiques soutenus par les normes. Enfin, jusqu'à tout récemment, il n'y avait que très peu de recherche pour l'application des technologies à la formation linguistique. Encore aujourd'hui, les résultats de cette recherche ne sont même pas largement diffusés dans l'industrie. L'importance de la commercialisation de la recherche s'avère plus grande pour les autres marchés.

En ce qui concerne le secteur du traitement de la parole, nous présentons ici les principales forces de l'industrie canadienne de la parole : les technologies de la parole ont atteint leur maturité, et leur consolidation en est la preuve; l'automatisation des centres d'appel représente aujourd'hui le rendement du capital investi en matière de parole; le Canada dispose des meilleurs chercheurs et développeurs d'applications de la parole; et il existe une synergie entre l'industrie de la parole et la recherche, grâce à l'AILIA et le présent processus de carte routière. À l'inverse, les principales faiblesses de l'industrie canadienne de la parole sont les suivantes : une absence de connaissance de l'utilisateur final à propos des avantages des technologies de la parole; une absence d'identification des perspectives d'occasions commerciales et de soutien à la commercialisation; il doit y avoir une adhésion plus grande du gouvernement en ligne face au traitement de la parole, comme c'est le cas en Nouvelle-Zélande, en Australie et en Europe; et enfin, les budgets de recherche et de développement de l'industrie diminuent.

Il est possible d'obtenir un aperçu de l'état actuel du secteur canadien de la traduction à partir du tableau statistique suivant, qui illustre sous forme de matrice les possibilités techniques des fournisseurs canadiens des technologies de traduction.

FOURNISSEUR CANADIEN DE TECHNOLOGIES DE LA TRADUCTION	POSSIBILITÉS TECHNIQUES																			
	Programme préexistant	Apportez votre programme	Reconnaissance vocale	Analyse documentaire	Prétraduction	Gestion terminologique	Consultation terminologique automatique	Dépouillement terminologique unilingue	Dépouillement terminologique multilingue	Langue contrôlée / prédiction	Préalignement automatique	Alignement interactif par transfert optique	Recherche plein texte	Concordancier bilingue	Mémoire de traduction	Localisation	Traduction automatique	Intergiciel de traduction	Interconnectivité	Client-serveur
TransSearch (RALI-Terminotix)	O													O						O
Dict. Generator (BridgeTerm)		O					O	O		O							I		O	
EDITerm (J. Chandioux)		O			C	O		C												
LogiTerm (Terminotix)		O			O	O	O			O		O	O	O	I				O	
LogiTermWeb (Terminotix)		O			O	O	O			O		O	O	O	I				O	O
SynchroTerm (BridgeTerm)		O				O		O				O							O	
XTracTerm (Bridgewater)		O					O													
Aladin (RALI/Terminotix)		O								O										
Alignment Robot (Terminotix)		O								O									O	
Find (Beetext)		O												O					O	
LogiTrans (Terminotix)		O		O	O			I		I		I	I	O					O	
LogiTransWeb (Terminotix)		O		O	O			I		I		I	I	O					O	O
Babel-TRACK (BabelFish)																			O	O
Flow (Beetext)																			O	O
XLT (Socatra)																	O			
Lingua Technologies Toolbox		O				O		O		O			O							
MultiTrans (MultiCorpora)		O	I	O	O	O		O			O	O	O	O			I		O	O
Promemoria (BridgeTerm)		O						O				C		O			I			

Légende :

O = Oui (c.-à-d. l'outil offre cette fonctionnalité)

C = Compagnon (c.-à-d. ne fait pas partie du produit principal, mais s'obtient par l'ajout de modules)

I = Interconnectivité (c.-à-d. ne fait pas partie du produit, mais s'obtient par l'interconnectivité avec un autre produit)

Écarts et obstacles

Comme il a été mentionné plus haut, il existe des lacunes au sein de l'industrie canadienne de la gestion de contenu en ce qui concerne l'état des facteurs, une des catégories décisives en matière

d'avantage concurrentiel. Pour ce qui est du facteur *ressources humaines* : la gestion de contenu n'est ni aussi connue ni aussi soutenue que les autres secteurs d'enseignement, ce qui explique pourquoi si peu d'étudiants choisissent de poursuivre des études supérieures dans ce domaine; étant donné le manque de programmes d'enseignement spécialisés dans ce domaine, les sociétés peinent à trouver des candidats présentant un profil linguistique et un profil informatique; enfin, plusieurs employés formés dans la gestion de contenu ont été mis à pied lors de l'éclatement récent de la bulle des hautes technologies et plusieurs d'entre eux quittent le Canada et le domaine de la gestion du contenu. Pour ce qui est du facteur *argent*, les centres de recherche manquent de fonds, ce qui mine leur capacité à mener des activités de R-D et à former le personnel futur pour ce secteur. Au niveau des entreprises, ce manque de fonds les empêche de développer et de mettre en marché de nouvelles technologies et de nouveaux produits. En ce qui concerne le facteur *connaissances*, la pénurie de ressources de gestion de contenu avec lesquelles les ressources humaines doivent travailler limite leur capacité à créer une nouvelle propriété intellectuelle dans le secteur. En outre, plusieurs raisons rendent difficile le transfert de connaissances du milieu universitaire vers l'entreprise. La situation n'est pas non plus idéale en ce qui concerne le facteur *ressources linguistiques*, surtout en ce qui concerne les ressources en français, et ce, pour plusieurs raisons : les ressources humaines en gestion de contenu ne connaissent souvent pas les ressources disponibles, ni ne savent où se les procurer; souvent, elles refusent de partager les outils qu'elles ont mis au point; il arrive parfois que ces ressources n'existent tout simplement pas (par ex., un *WordNet* en français); et parfois, surtout dans le cas de petites entreprises, ces ressources sont trop onéreuses. Enfin, l'état de la *demande* au Canada peut représenter un obstacle supplémentaire au développement de l'industrie, étant donné, par exemple, les politiques d'approvisionnement gouvernementales qui ne favorisent pas les PME mais plutôt les grands intégrateurs de systèmes et les sociétés américaines, et les utilisateurs qui ignorent souvent tout de la gestion de contenu.

Des modèles innovateurs existent dans le secteur de la formation linguistique. Les organismes de média de masse sont des partenaires de distribution et de ressources importants en matière de diffusion du contenu de formation linguistique. Avec l'intérêt accru du monde d'aujourd'hui pour la mobilité des solutions, le contenu de formation linguistique fait son apparition, combinant une pédagogie de formation linguistique simple à une technologie et des interfaces utilisateur conviviaux. Les organismes disposant déjà de contenu peuvent bénéficier des marques et produits établis afin de créer de nouvelles perspectives et sources de revenus. La valeur de la normalisation apparaît particulièrement évidente en ce qui a trait à l'évaluation, où une évaluation frontale intégrée et axée sur des normes régionales ouvre la porte à un cheminement complet au cœur de ce système régional. En adaptant leur produit à une norme généralement acceptée, les développeurs de technologie en formation linguistique peuvent s'assurer un accès à un marché très vaste. Des partenariats ont été formés entre des organismes possédant des forces très diversifiées, par l'adhésion au modèle *financier-fournisseur de contenu-développeur de plateforme-distributeur*. Les produits technologiques de formation linguistique répondant à un besoin particulier en matière de ressources humaines peuvent connaître un grand succès sur le marché du travail, surtout s'ils adhèrent aux normes de cette industrie. Comment ces modèles peuvent-ils être appliqués? Il existe de nombreuses possibilités de partenariat entre les organismes, fondées sur les quatre avantages principaux suivants : i) les plus petits fournisseurs peuvent mettre à profit leur contenu, repérer les besoins en formation linguistique et approcher les chefs de file de cette industrie en tant que partenaires de ressources en développement d'outils destinés aux ressources humaines; ii) il existe un besoin en matière de contenu de formation linguistique personnalisé au sein des organismes de diffusion de vaste contenu; iii) les produits de formation linguistique actuels n'arrivent pas à concilier les perfectionnements technologiques et les perfectionnements pédagogiques; une plus grande coopération au sein de ce

sous-secteur permettrait de réduire cet écart; et iv) il a été démontré que les partenariats entraînent une augmentation notable des réseaux et de l'accès à l'information, ce qui ouvre la voie à de nouvelles perspectives et sources de revenus. En l'absence d'un plan de normalisation national – qui représenterait un engagement pertinent mais majeur pour le Canada –, il est impossible de mettre en œuvre un plan de facto qui puisse offrir l'apparence et la fonctionnalité de la normalisation et renforcer ainsi la crédibilité et le potentiel de commercialisation des produits canadiens de la technologie du langage. Les fournisseurs canadiens ont tendance à sous-estimer la valeur du contenu ou, du moins, ses applications technologiques. Les organismes canadiens devraient prendre du recul pour mieux cerner les éléments clés de leurs systèmes. Ils pourraient ainsi les utiliser dans la technologie ou les présenter à des partenaires potentiels en matière de ressources. La création de modèles novateurs soulève quatre enjeux : i) bien qu'il existe un certain nombre de sources de financement potentiel, la structure actuelle de l'industrie ne permet pas de trouver facilement l'information relative à ces sources; ii) la complexité et la diversité de l'industrie de la formation linguistique rendent très difficile l'utilisation de normes comme base de développement d'un produit technologique qui plaise au marché le plus vaste possible; iii) même lorsqu'il est possible d'identifier des normes, un dilemme se pose, pour les organismes, entre la conformité à une norme globale promettant un accès à un marché plus vaste mais offrant moins de potentiel pour un créneau, et la conformité à une norme plus locale offrant plus de perspectives mais limitant la taille du marché potentiel; et iv) contrairement aux pays en concurrence avec le Canada, l'absence de compétence fédérale en matière d'éducation et de formation n'encourage pas la création d'un environnement propice à l'innovation dont le développement technologique a besoin.

En ce qui concerne le traitement de la parole, si les grandes sociétés n'envisagent pas objectivement la reconnaissance automatique de la parole comme une solution possible, c'est tout simplement parce qu'elles ne connaissent pas ou très peu les données suivantes : à savoir si la reconnaissance automatique de la parole fonctionne réellement; les types de solutions particulières de reconnaissance automatique de la parole et les différences entre elles; où et en quoi la reconnaissance automatique de la parole peut s'avérer utile pour l'entreprise; où en quoi la reconnaissance automatique de la parole peut ne pas s'avérer utile pour l'entreprise; la pertinence des solutions de reconnaissance automatique de la parole dans le paysage global du commerce et des technologies de l'information; les avantages potentiels de la reconnaissance automatique de la parole; et la valeur concrète d'un déploiement efficace de la reconnaissance automatique de la parole dans le monde réel.

Le sous-comité de la traduction a cerné deux secteurs dans lesquels les besoins de traduction dans leur ensemble sont cruciaux. Le premier secteur a trait aux systèmes de gestion de la traduction. Les étapes centrales du processus de traduction sont celles qui sont les plus touchées : i) les étapes de prétraduction, telles la création d'un glossaire; ii) le processus de traduction en soi; iii) les étapes de révision et de correction d'épreuves; iv) les étapes d'édition et d'assurance de la qualité linguistique; et v) l'étape de mise au point finale, notamment la réimportation des données de test, etc. Les facteurs cruciaux suivants ont été identifiés : la création d'une trousse de traduction pour l'étape de prétraduction; la distribution et le suivi, la résolution de problèmes, la rétroaction à 360 degrés en matière de qualité et les mesures d'évaluation. Le second secteur crucial est la création de systèmes de flux de travaux de bout-en-bout conviviaux, dont les facteurs suivants s'avèrent essentiels : coût, convivialité et facilité de mise en œuvre.

Stratégie et objectifs de développement

Pour amorcer la discussion relative à une stratégie de développement, il faut commencer par analyser l'état actuel de l'industrie de la langue, qui est décrit dans le Plan d'action pour les langues officielles, et le plan dressé à cet effet par le gouvernement. Selon le Plan d'action, les industries de

la langue font face à quatre grands enjeux : la fragmentation de l'industrie, le manque de visibilité, la relève insuffisante et des investissements en recherche-développement insuffisants. Par conséquent, le Plan d'action fait valoir que, afin de relever les deux premiers défis, il faut avant tout resserrer les liens entre les industries de la langue et accroître leur visibilité. Ces questions seront réévaluées au terme de la discussion portant sur les éléments d'une stratégie de développement, soumis par les divers sous-comités.

Le sous-comité de gestion de contenu suggère une liste de cinq éléments que le Canada devra prendre en considération pour conserver et renforcer sa position dans le domaine de la gestion de contenu. Cette liste correspond largement à la liste des facteurs décisifs en matière d'avantage concurrentiel. En ce qui concerne les *ressources humaines*, les activités suivantes sont proposées : créer des programmes de synergie entre les grandes sociétés, les PME, les laboratoires de R-D et le milieu universitaire, notamment des incitatifs fiscaux; endosser la création de programmes d'études supérieures dans le domaine de la gestion de contenu; créer des centres d'expertise universitaires et y inviter les employés des PME et des grandes sociétés à des programmes de formation; offrir des bourses aux étudiants qui décident de poursuivre leurs études dans le domaine de la gestion de contenu et aux professeurs qui travaillent à la recherche dans ce secteur; et offrir des incitatifs fiscaux aux PME et aux grandes sociétés œuvrant dans le secteur de la gestion de contenu afin qu'elles offrent de la formation aux nouveaux diplômés. Les activités concernant *l'argent* comprennent : offrir aux PME un meilleur accès au capital; offrir des incitatifs aux sociétés d'investissement en capital de risque et aux banques afin qu'elles financent les PME du secteur de la gestion de contenu; mettre sur pied un fonds d'investissement en capital de risque AILIA pour l'industrie de la langue et les PME – particulièrement celles établies par les R-D gouvernementale et universitaire; créer un programme particulier assurant le financement d'initiatives de recherche solides dans le secteur de la gestion de contenu; et s'assurer que les centres de recherche participant à des projets de recherche internationaux ont accès aux sources de financement nationales. Les activités concernant *l'argent* comprennent : offrir un soutien financier aux organismes sans but lucratif pour encourager la création et le partage de ressources ou logiciels langagiers précommerciaux utiles aux centres de recherche et aux sociétés du secteur de la gestion de contenu. Les éléments suivants se rapportent au facteur *demande* : adopter des normes, telles XML, propices à l'échange d'information dans l'emploi des technologies de gestion du contenu et garantissant aux citoyens canadiens un haut niveau de service à la clientèle; s'assurer que les organismes gouvernementaux allouent une partie de leur budget technologique à l'acquisition de produits et de services de gestion de contenu; créer un répertoire de sociétés et de centres de recherche du domaine de la gestion de contenu, mis à jour par les sociétés et les organismes de recherche eux-mêmes; et financer les expositions technologiques démontrant les capacités réelles des outils de gestion de contenu pour les organismes gouvernementaux. Enfin, les activités suivantes concernent la *stratégie* : suivant en cela le conseil du CDI, encourager les sociétés à offrir des systèmes de gestion de contenu permettant « de régler la crise de complexité par la simplification des processus d'affaires et l'unification de l'accès au contenu », « d'offrir des fonctions de gestion de dossiers et d'archivage » et « d'unifier l'accès des sociétés au contenu structuré et non structuré et d'offrir des outils permettant de les gérer en même temps »; convoquer un forum réunissant les grandes sociétés, les PME, le milieu universitaire et les laboratoires de recherche gouvernementaux afin d'établir un plan d'action pour l'industrie canadienne de la gestion de contenu; établir des cibles annuelles qui serviront de mesures d'évaluation pour l'industrie canadienne de la gestion de contenu; et organiser des conférences portant sur des thèmes particuliers de la gestion de contenu afin d'encourager l'échange d'information.

Selon le sous-comité de formation linguistique, une banque de ressources doit être créée afin de diffuser les perspectives de financement et afin d'aider les organismes à mieux comprendre le

processus permettant d'y accéder. Cette initiative devrait s'ajouter à l'élaboration d'un groupe industriel permettant aux fournisseurs de ressources de se consacrer davantage aux perspectives des technologies de formation linguistique. L'industrie doit développer un portail de communication qui permettra à toutes les parties intéressées de communiquer entre elles à propos de perspectives de collaboration et d'informer l'industrie des nouveaux développements. L'industrie canadienne de la formation linguistique doit faire des efforts afin de tirer avantage de la réputation mondiale du Canada. Il serait possible d'utiliser la prochaine tenue des Jeux olympiques d'hiver à Vancouver-Whistler pour faire valoir les solutions canadiennes de la technologie de la langue. L'industrie canadienne de la langue doit élaborer une méthode de normalisation rapide par le développement d'une marque maîtresse, qui offrirait une façon simplifiée de comprendre, de comparer et d'intégrer tous les sous-secteurs et segments de l'industrie. En matière de rôles et de responsabilités, les exploitants privés doivent reconnaître la valeur d'une approche Équipe Canada, s'élever au-dessus des disputes internes de l'industrie et choisir un dirigeant organisationnel qui soit à l'avant-garde des efforts visant à une communication, une normalisation et une intégration de plus grande envergure. Le gouvernement doit continuer à raffermir son engagement envers l'innovation dans l'industrie de la langue, et communiquer l'importance de l'industrie à tous ses services et secteurs de compétence. La communauté universitaire doit piloter la question de la normalisation, tout en demeurant assez souple pour poursuivre les progrès en faveur de l'intégration et en soutenant la création d'une structure de communication entre les créneaux de l'industrie. Enfin, au niveau associatif, l'AILIA doit se consacrer à l'augmentation de ses ressources, afin d'accélérer les efforts nécessaires à la diffusion de l'information, à la communication entre les créneaux, à la facilitation et à la promotion.

Le sous-comité du traitement de la parole nomme comme modèle à suivre EuroMap – organisme européen dont la tâche consiste à « conscientiser, établir des ponts et offrir des services propres à la commercialisation afin d'encourager le marché à tirer parti des résultats des activités européennes et nationales de recherche et de développement de la technologie du langage humain ». En outre, les recommandations suivantes sont soumises : offrir un aperçu de la technologie de la parole et de sa terminologie; créer un inventaire des fournisseurs, chercheurs et adaptateurs canadiens de la technologie du traitement de la parole; créer un répertoire et un inventaire des données sonores; offrir un répertoire des « personnalités » canadiennes de la technologie du langage humain; offrir un portail d'information sur la recherche canadienne en matière de technologie de la parole; établir des forums afin d'établir des alliances stratégiques et d'interagir avec les autres joueurs de la technologie de la langue; créer un portail d'information portant sur le financement commercial et de recherche; fournir des renseignements et des activités à jour en matière de technologie du langage humain, au Canada et dans le monde; offrir des liens vers d'autres sites utiles; et tenter de normaliser les méthodologies et mesures d'évaluation de la technologie de la parole.

Le sous-comité de la traduction fournit des recommandations préliminaires afin de cerner les lacunes de l'industrie canadienne de la traduction comme mesure prioritaire. Plus particulièrement, l'accent porte sur les lacunes en matière de fonctions et produits développés au Canada et en matière de services connexes technologiques offerts au Canada.

Trois concepts essentiels permettent de résumer les recommandations proposées à différents paliers du processus de la carte routière : *synergie, visibilité et retombées économiques*.

La question de la *synergie* a été soulevée de façon quasi unanime, de diverses façons interreliées, par la plupart des parties ayant pris part au processus de la carte routière. Tel que mentionné précédemment, la question de la fragmentation de l'industrie de la langue et le besoin correspondant de renforcer les liens de l'industrie de la langue ont été soulignés dans le Plan d'action pour les langues officielles. Dans le même ordre d'idées, toutes les stratégies de développement mises de

l'avant par les sous-comités ont en commun le même enjeu. En général, la portée du consensus sur le besoin de synergie et le degré d'importance accordé à cette question, tant en matière de stratégie de soutien au développement qu'en matière de développement technologique, font de la question de la synergie un facteur primordial de la Carte routière technologique. En effet, la question de la synergie est sous-jacente à de nombreux attributs technologiques identifiés comme étant cruciaux, tels la *pollinisation croisée*, l'*interopérabilité* et les *normes*. Par exemple, tous les sous-comités considèrent la pollinisation croisée comme un élément critique. Cependant, ce terme est utilisé de trois façons interreliées, qu'il peut être utile de distinguer : dans son sens premier, la pollinisation croisée s'entend d'un chevauchement de la portée de deux secteurs différents; dans son deuxième sens, la pollinisation croisée résulte du fait qu'une technologie donnée soit aussi pertinente à divers secteurs; et dans son troisième sens, la pollinisation croisée est liée à la distinction importante entre technologies et applications, telle que présentée par certains sous-comités, et largement acceptée dans les discussions entourant le processus de la Carte routière technologique. Ce dernier sens – asymétrique de par sa nature, contrairement aux deux premiers – indique qu'une pollinisation croisée est présente lorsqu'une technologie d'un secteur peut s'avérer pertinente à une application d'un autre secteur. Si nous considérons la pollinisation croisée comme un attribut crucial essentiel, notons qu'en plus de favoriser la synergie, la pollinisation croisée constitue également un excellent moyen de ne pas restreindre le développement technologique à une seule avenue. En effet, l'élaboration de technologies utiles à différents secteurs garantit l'accès simultané à plusieurs avenues technologiques. Les deux autres notions interreliées, l'interopérabilité et les normes, traitent aussi de l'importance de contourner la fragmentation.

Le deuxième enjeu que soulève le Plan d'action pour les langues officielles et, par conséquent, la deuxième priorité stratégique, est la *visibilité*. Il s'agit d'un problème dont la pertinence fait l'objet d'un vaste consensus parmi les sous-secteurs de l'industrie de la langue. Le sous-comité de la formation linguistique, par exemple, perçoit le manque de reconnaissance de l'industrie canadienne de la langue comme un problème de taille, et propose, comme solution possible, la création d'une image de marque. Dans la même veine, le sous-comité de traitement de la parole soutient que si les grandes sociétés n'envisagent pas objectivement la reconnaissance de la parole comme une solution possible, c'est tout simplement parce qu'elles ne connaissent pas ou que très peu l'existence de cet outil. Le mandat de l'AILIA comprend aussi l'augmentation de la visibilité de l'industrie. Dans le contexte actuel, le problème est plus présent au niveau de la stratégie de soutien au développement, vu l'importance d'appuyer les technologies ayant atteint un degré de maturité satisfaisant, qu'il ne l'est au niveau du développement technologique.

Enfin, les retombées économiques des nouvelles technologies sont mises en évidence par les attributs cruciaux sélectionnés par tous les sous-comités. Outre l'évaluation des *possibilités des marchés intérieur et international*, les sous-comités ont choisi à l'unanimité deux autres attributs importants : *coût de développement* et *rendement du capital investi pour l'utilisateur final*.

Une autre question clé issue du processus de la carte routière concerne la catégorisation des logiciels. Certains sous-comités, soit ceux du traitement de la parole et de la traduction, font une distinction profonde entre les applications et les composants technologiques. Le fait que certains sous-comités plutôt que d'autres aient soulevé cette distinction explique en partie la présence de disparités dans la série d'attributs cruciaux fournis par les différents sous-comités, qui ne sont pas entièrement comparables. Toutefois, plutôt que de considérer ces disparités comme une faiblesse, il faudrait les voir comme un indicateur de la nature particulière de chaque sous-secteur. On pourrait, notamment, faire une distinction fondamentale entre certains secteurs, comme celui du traitement de la parole, qui sont des fournisseurs de technologies, et d'autres secteurs, comme celui de la formation linguistique,

qui sont des bénéficiaires de ces technologies. Une telle distinction peut être mise à profit en faveur de la pollinisation croisée asymétrique. Autrement dit, au lieu de sélectionner les technologies à l'aide de critères qui placeraient les sous-secteurs sur un même pied d'égalité et qui les forceraient se faire concurrence, nous pourrions établir au préalable un critère de distinction, entre fournisseurs et bénéficiaires de technologies, et utiliser ce critère dans un but de sélection ou d'élimination, ce qui mènerait à l'évaluation d'autres scénarios possibles. Il serait ensuite pertinent d'examiner plus en détail les scénarios qui répondent d'emblée aux exigences de pollinisation croisée asymétrique (par ex., proposer l'utilisation du traitement de la parole dans le cadre de la formation linguistique) en fonction de leur utilisation optimale dans d'autres formes de pollinisation croisée et en fonction d'autres attributs cruciaux. Enfin, l'inclusion des ressources en tant que catégorie particulière parmi les produits ciblés par certains sous-comités peut être considérée comme la sélection implicite de ressources linguistiques dans leurs propositions en matière de stratégie de développement. Comme le souligne le Plan d'action pour les langues officielles, la dualité linguistique offre au Canada un avantage concurrentiel en matière de ressources de développement, étant donné qu'elle encourage la création de ressources langagières qui sont particulières à une langue sans être particulières à une nation. Les possibilités de pollinisation croisée sont tout aussi importantes, sinon davantage, pour les ressources langagières que pour les technologies et les applications. Dans le cas actuel, nous pourrions envisager de répartir le travail entre les sous-secteurs, à titre de complément aux mesures déjà évoquées pour les technologies et les applications.

En somme, le scénario qui cadrerait le mieux avec les suggestions de stratégie de développement mises en évidence par les différents sous-comités est celui qui permettrait une distinction fondamentale entre le rôle du développement technologique et celui de la stratégie de soutien au développement technologique. En vertu de ce scénario, le développement technologique pourrait contrer la fragmentation de l'industrie de la langue, tandis que la stratégie de soutien au développement technologique s'attaquerait au problème de visibilité de l'industrie de la langue, possiblement par la création d'une image de marque. En ce qui a trait au développement technologique, il faudrait établir une distinction entre les technologies, les applications et les ressources. Nous pourrions ainsi envisager de répartir, toujours selon le même modèle, le travail et les responsabilités des sous-secteurs. En outre, nous pourrions élaborer une stratégie de développement technologique en deux étapes, en convenant au préalable de la répartition des tâches entre les sous-secteurs et du classement séparé des fournisseurs de technologies, des bénéficiaires de technologies, des concepteurs de ressources et autres. Les applications visées seraient celles des secteurs bénéficiaires de technologies, et les technologies visées, celles des secteurs fournisseurs de technologies, cela afin d'assurer une pollinisation croisée verticale optimale. Ce projet de pollinisation croisée pourrait servir de critère à l'évaluation détaillée des solutions de développement technologique possible. Les paramètres d'évaluation et de sélection des demandes de candidatures, des technologies et des ressources, s'inscriraient au projet préliminaire et seraient surtout axés sur les attributs cruciaux ayant fait l'objet d'un consensus entre les sous-comités, tels que la pollinisation croisée horizontale (pour ce qui est des technologies et de la portée sectorielle des applications), les retombées économiques et la promotion de la normalisation.

Évaluation et sélection des technologies

Chaque sous-comité a fourni une évaluation quantitative des produits ciblés, selon les attributs cruciaux identifiés précédemment. Ces données, illustrées plus avant, représentent l'information centrale sur laquelle devrait reposer la sélection des technologies. Cela ne signifie pas pour autant que l'information contenue aux présentes est de nature exhaustive, mais qu'au contraire, il serait possible et envisageable d'y intégrer de nouvelles données. Nous recommandons toutefois d'intégrer

les nouvelles données de manière à les concilier au cadre de travail fourni dans la présente section.

En outre, divers critères d'évaluation et de sélection de la technologie ont été proposés par les parties intéressées dans le cadre du processus de la carte routière. En résumé, les critères proposés peuvent se diviser en trois groupes : i) les critères correspondant aux attributs utilisés dans les résultats de l'évaluation, c.-à-d. les attributs cruciaux identifiés par les sous-comités, tels que les possibilités de marché et la pollinisation croisée; ii) les critères ne correspondant pas exactement aux attributs utilisés dans l'évaluation, mais pour lesquels certains attributs, seuls ou combinés, représentent une approximation du critère à l'étude. Par exemple, le critère « faisabilité » peut correspondre à un attribut « coût de développement », considéré comme un bon indicateur du premier critère; iii) les critères ne correspondant à aucun attribut, comme « parties potentielles » et « qualité des ressources humaines en cause ». Les deux premières catégories semblent englober la majorité des critères indiqués ci-dessus. Nous pouvons donc dire, qu'avec un brin d'ingéniosité, il est possible d'établir une certaine correspondance entre la plupart des séries de critères et les combinaisons d'attributs correspondantes.

Cependant, il faut se rappeler que certains attributs, n'étant pas uniformément utilisés par tous les sous-comités, présentent des lacunes. Il convient de souligner que quelques-uns des critères n'ayant aucun attribut correspondant, tels que « participants éventuels », « qualité des ressources humaines » et divisions « besoins comblés par le projet », ne pourraient faire l'objet d'une représentation quantitative. Par conséquent, si nous choisissons d'utiliser ces critères, il faudra sans doute délaissé les paramètres quantitatifs et utiliser des évaluations plus subjectives. Enfin, le critère important proposé quant à la « classification de la *technologie* à l'intérieur de ses *composantes* », qui ne fait pas actuellement l'objet de données d'évaluation, n'est pas de nature quantitative en soi, mais pourrait servir de base à une évaluation quantitative à partir de groupes de logiciels connexes plutôt que d'éléments seuls.

Comme nous l'avons déjà mentionné, les données d'évaluation fournies par les différents sous-comités présentent des disparités. Les disparités se situent partiellement au niveau de l'utilisation d'attributs entièrement différents. Jusqu'à un certain point, de telles disparités doivent être acceptées comme indicatrices de la nature polyvalente des divers sous-secteurs. L'utilisation d'échelles d'évaluation différentes pour un même attribut crucial ou pour des attributs analogues est un autre exemple de disparité. Dans ce cas, un effort de normalisation a été effectué afin d'exprimer les évaluations analogues dans un même format. Plusieurs classements comparatifs en matière de pollinisation croisée peuvent être fournis pour des attributs cruciaux seuls ou combinés. Par exemple, les tableaux suivants illustrent les produits premiers au classement de chacune des catégories pertinentes (applications, ressources, technologies et outils) en vertu d'une note combinée correspondant à la somme des notes individuelles pour chaque attribut partagé entre tous les sous-secteurs. L'échelle de notes des attributs seuls sera généralement de quatre points.

<i>Sous-secteur</i>	<i>Application</i>							
		<i>Marché intérieur</i>	<i>Marché international</i>	<i>Coût de développement</i>	<i>RCI utilisateur final</i>	<i>Conformité aux normes</i>	<i>Durée de vie</i>	<i>Résultats combinés</i>
TP	Réponse vocale interactive	3	4	3	4	4	4	22
TP	Machine à dicter de table	3	3	4	4	4	4	22
TP	Cmd/ctrl de table	3	3	3	4	4	4	21
TP	Centres d'appels, portails Web	3	3	2	4	4	4	20
FL	Cours TESL en ligne	3	3	3	4	3	3	19
FL	Outils de création de didacticiels et systèmes de Gestion de la formation	4	4	1	2	4	4	19
TP	Transcription	3	3	3	4	2	4	19
TP	Messagerie unifiée	3	3	2	3	4	4	19
TP	RP intégrée	3	4	2	4	3	3	19
Tr	Mémoire de traduction / install. client	3	3	2	3	4	4	19

En plus des attributs utilisés ci-dessus, le tableau suivant contient également la *pollinisation croisée*.

<i>Sous-secteur</i>	<i>Ressource</i>								
		<i>Marché intérieur</i>	<i>Marché international</i>	<i>Coût de développement</i>	<i>RCI utilisateur final</i>	<i>Conformité aux normes</i>	<i>Durée de vie</i>	<i>Pollinisation croisée</i>	<i>Résultats combinés</i>
FL	Site Web par abonnement offrant aux enseignants du matériel d'apprentissage conforme au programme d'enseignement national	3	3	2	2	4	4	2	20
FL	Accès aux médias pour diffuser de l'information relative à la formation	3	3	3	2	3	3	2	19
FL	Dictionnaires multilingues sur cellulaire ou ANP pour la traduction et la formation linguistique	3	4	2	2	1	3	4	19

<i>Sous-secteur</i>	<i>Ressource</i>								
		<i>Marché intérieur</i>	<i>Marché international</i>	<i>Coût de développement</i>	<i>RCI utilisateur final</i>	<i>Conformité aux normes</i>	<i>Durée de vie</i>	<i>Pollinisation croisée</i>	<i>Résultats combinés</i>
FL	Ressources pédagogiques normalisées pour aider les enseignants à utiliser la technologie	3	3	1	2	4	4	1	18
FL	Sites de ressources par abonnement s'adressant expressément aux enseignants ALS	2	3	3	2	3	3	2	18
FL	Dictionnaires interactifs	2	2	4	2	3	3	2	18
GC	Corpus parallèles	1	1	3	2	2	4	4	17
GC	Corpus écrits	1	1	3	2	2	4	4	17
FL	Réseaux Internet mondial pour encourager le partenariat entre les élèves	2	3	3	2	1	3	3	17
FL	Répertoires sur Internet des cours à l'étranger pour les élèves d'ici et d'autres pays	3	3	3	2	1	3	2	17

Le tableau suivi comprend les mêmes attributs que ci-dessus, sauf le *coût de développement*.

<i>Sous-secteur</i>	<i>Technologie</i>								
		<i>Marché intérieur</i>	<i>Marché international</i>	<i>RCI utilisateur final</i>	<i>Conformité aux normes</i>	<i>Durée de vie</i>	<i>Pollinisation croisée</i>	<i>Résultats combinés</i>	
TP	Traitement des signaux / compression de la parole	4	4	4	4	4	4	4	24
Tr	Traduction automatique / en ligne	3	4	5	4	3	4	4	23
TP	Reconnaissance de la parole / selon le locuteur	3	3	4	4	4	4	4	22
TP	Conversion de la parole en texte / synthèse de la parole	3	3	4	4	4	4	4	22
TP	Conversion parole-texte / chaîne concat.	3	3	4	4	4	4	4	22

<i>Sous-secteur</i>	<i>Technologie</i>							
		<i>Marché intérieur</i>	<i>Marché international</i>	<i>RCI utilisateur final</i>	<i>Conformité aux normes</i>	<i>Durée de vie</i>	<i>Pollinisation croisée</i>	<i>Résultats combinés</i>
Tr	Terminologie / en direct	3	3	5	3	4	4	22
Tr	Mémoire de traduction / intégration	3	3	3.5	4	4	4	21.5
TP	Reconnaissance de la parole / intégrée	3	3	3	4	4	4	21
TP	Reconnaissance de la parole / indép. du locuteur	3	3	3	4	4	4	21
Tr	Mémoire de traduction / en ligne	4	4	5	1	3	4	21

Le tableau suivant comprend les mêmes attributs que ci-dessus.

<i>Sous-secteur</i>	<i>Outil</i>							
		<i>Marché intérieur</i>	<i>Marché international</i>	<i>RCI utilisateur final</i>	<i>Conformité aux normes</i>	<i>Durée de vie</i>	<i>Pollinisation croisée</i>	<i>Résultats combinés</i>
FL	Outils d'apprentissage en ligne installés sur dispositifs de poche	3	4	3	3	3	3	19
FL	Logiciels de scénarios-maquettes avec jeux de rôle pour enfants	3	3	1	4	3	3	17
FL	Cours d'anglais en ligne par abonnement	3	4	2	3	3	2	17
FL	Systèmes d'apprentissage des langues avec outils multimédias (vidéo, enregistrement, essais interactifs)	3	4	2	3	3	2	17
FL	Milieux immersifs pour la formation linguistique liée à une fonction particulière	2	3	3	4	2	3	17
FL	Formation linguistique par télévision/radio avec utilisation de manuels didactiques	3	4	2	3	4	1	17

Sous-secteur	Outil							
		Marché intérieur	Marché international	RCI utilisateur final	Conformité aux normes	Durée de vie	Pollinisation croisée	Résultats combinés
FL	Outils d'exercices pratiques en ligne conformes aux normes d'évaluation	3	4	2	4	2	2	17
FL	Évaluation technohabilité pour la validation des offres d'emploi	2	2	3	4	4	2	17
FL	Logiciels d'apprentissage des langues en ligne ou sur CD	3	3	2	3	3	2	16
FL	Logiciels de formation linguistique par le jeu, sur CD	3	3	1	4	2	3	16

Les données d'évaluation échantillonnées ci-dessus fournissent l'information de fond à utiliser pour classer les technologies par ordre de priorité. Cependant, au lieu d'utiliser ces données pour la *génération* algorithmique d'un ensemble de technologies dont il faut établir l'ordre de priorité, il semble plus raisonnable et réaliste de les utiliser pour *justifier* l'ordre de priorité défini. Autrement dit, il serait logique d'imposer comme critère qu'une proposition qui favorise la technologie A par rapport à la technologie B justifie sa démarche en démontrant des résultats d'évaluation supérieurs pour la technologie A par rapport à ceux de la technologie B. Ce critère constituerait un paramètre explicite mesurable pour évaluer et sélectionner les technologies qui ont été mises de l'avant à la suite d'un processus non spécifié. Dans cet esprit, il est suggéré d'entreprendre une étape préliminaire permettant l'atteinte d'un consensus au sujet des lignes directrices en matière de stratégie générale. Ces résultats permettraient de filtrer, aux fins d'évaluation et de sélection, les différents scénarios proposés pour les technologies. Comme il a été mentionné plus haut, un point préliminaire semble faire l'objet d'un consensus, à savoir la distinction entre les applications, les technologies et les ressources. Il a par ailleurs été suggéré d'appliquer la même distinction afin d'établir au préalable la répartition des travaux entre les différents sous-secteurs, de façon à délimiter les rôles des fournisseurs de technologies, des bénéficiaires des technologies, des concepteurs de ressources et autres. Un tel consensus dès le départ aurait pour effet de réduire l'espace de recherche des technologies à évaluer en éliminant d'emblée les scénarios inappropriés. Il permettrait aussi, durant le processus de sélection, de tirer parti des disparités d'évaluation entre les sous-secteurs, plutôt que de sentir limité par celles-ci. En outre, les attributs communs à tous les sous-secteurs peuvent constituer un ensemble approprié d'attributs cruciaux à privilégier dans l'évaluation des technologies. Les données et les mesures d'évaluation peuvent être enrichies, intégrées et raffinées de plusieurs façons : de nouveaux attributs peuvent être ajoutés et les écarts entre les attributs existants peuvent être éliminés, de telle sorte que le nombre d'attributs communs à tous les sous-secteurs soit élargi; une note combinée peut être raffinée en lui affectant un poids relatif aux attributs cruciaux comprenant la note combinée, de telle sorte que les attributs les plus pertinents sont favorisés par rapport aux attributs moins pertinents; et une note combinée peut être établie non seulement relativement à divers attributs du même élément technologique, mais, en outre, relativement à un

ensemble d'éléments technologiques possédant des liens entre eux, ajoutant un niveau de complexité. D'autres extensions pourraient facilement être envisagées. Quoi qu'il en soit, à ce stade, il est possible de conclure que : i) quelles que soient les extensions des données d'évaluation, celles-ci devraient trouver suffisamment de preuves à l'appui dans le vaste bassin d'information offert par la carte routière technologique; et ii) quels que soient les critères d'évaluation adoptés, ceux-ci devraient être formulés de façon explicite, afin d'être facilement vérifiables et applicables dans leur ensemble. Ainsi, après plusieurs itérations, on pourrait souhaiter que les parties intéressées en arrivent à une entente commune, que les suggestions provenant de sous-secteurs spécifiques soient corroborées par un processus de validation de pollinisation croisée avec consensus, et que le processus d'établissement de la carte routière aboutisse à un document exhaustif et complet.

Technologies recommandées

Il est suggéré que les spécifications des projets proposés jusqu'à présent par les sous-comités, par exemple, la *Construction d'un WordNet multilingue* et le *Centre de soutien vocal* de langues croisées proposés par le sous-comité de la gestion du contenu, ou la « *Traduction* » *en temps réel de la voix en parole dans une même langue*, proposée par le sous-comité de la traduction, doivent être comparés aux données d'évaluation disponibles afin de tester la pertinence du projet relativement aux attributs cruciaux identifiés précédemment. Cet exercice vise à proposer des façons de comparer de nouvelles propositions et de les classer par priorité et, inversement, de présenter des méthodes afin d'étendre la portée des données d'évaluation et d'utiliser ces données comme banc d'essai pour valider les propositions de projet.