

**TECHNOLOGIES DE LA LANGUE EN EUROPE : MARCHE ET TENDANCES**

**SYNTHESE DE L'ETUDE 2005-2006**

**1. OBJECTIF**

L'étude 2005-2006 « Technologies de la langue en Europe : marché et tendances », réalisée à la demande de la Direction de la Stratégie du Ministère délégué à la Recherche et aux Nouvelles Technologies, dans le cadre du programme Technolangue qualifie et évalue le marché européen des outils linguistiques traitant les données textuelles et vocales, et en identifie les axes de développement.

Elle consiste en une actualisation de l'étude conduite en 2003 sur les 13 pays de l'Union européenne (Allemagne, Autriche, Belgique, Danemark, Espagne, Finlande, France, Grèce, Irlande, Italie, Pays Bas, Royaume Uni, Suède), avec un élargissement aux 12 autres pays membres, dont 10 nouveaux entrants en 2004 (Chypre, Estonie, Hongrie, Lettonie, Lituanie, Luxembourg, Malte, Pologne, Portugal, République tchèque, Slovaquie, Slovénie).

**2. METHODOLOGIE**

La même méthodologie que pour l'étude précédente a été appliquée. A savoir :

- recherche et exploitation de l'information économique et financière disponibles ;
- l'extraction des indicateurs clés du marché ;
- des entretiens complémentaires auprès des principaux acteurs des outils linguistiques ;
- présentation à un groupe d'experts pour ajustement éventuel et enrichissement.

L'étude s'est déroulée, pour l'essentiel, sur l'année 2006 selon 3 phases :

- Phase 1 : identification des acteurs et analyse de l'offre en Europe et en France.
- Phase 2 : analyse du « cœur du marché » des outils linguistiques en France.
- Phase 3 : détermination des tendances et des perspectives à l'horizon 2010.

**3. RESULTATS**

Le marché des technologies de la langue se confirme comme un marché de « niche » représenté par une multitude d'acteurs dont certains sont à la fois éditeurs de logiciels et offreurs de services. La segmentation basée sur les usages et les applications déployées a été reprise afin de confirmer les tendances observées en 2003 et appréhender au mieux les évolutions.

Cette typologie comprend neuf segments :

1. Gestion du contenu	4. E-business	7. Interfaces vocales, multimodales
2. Gestion des documents	5. Ressources linguistiques	8. E-learning
3. Traduction	6. Production et édition	9. Moteurs de recherche intelligents

**3.1. IDENTIFICATION ET EVALUATION DE L'OFFRE**

Le marché des outils linguistiques est étroitement corrélé aux variations des différents éléments structurels et conjoncturels des TIC : évolutions des dépenses en NTIC, et, en Recherche et Développement ; situation du marché des logiciels ; croissance du stockage et de la répartition des données dans les réseaux ; déploiement du secteur des NTIC dans l'économie mondiale.

Selon l'OCDE, les 25 pays de l'union européenne compte 460 millions de personnes en 2005 dont plus de 160 millions d'internautes, et 55% d'entre elles parlerait au moins une seconde langue.

Les dépenses en NTIC en Europe (Administration, GE et PME, particuliers) sont de 625 milliards d'euros en 2005. Pour 2006: prévisions à environ 6% du PIB.

La part des dépenses en R&D en Europe (investissements privés et publics) : 200 milliards d'euros en 2004 soit 1,90 % du PIB). Allemagne, France, et Royaume-Uni en représentent 120 milliards d'euros, soit 60% du total.

La R&D orientée linguistique est estimée à 800 millions de dollars en 2003. Les dépenses en R&D des secteurs producteurs de TIC en France est de 0,16% (contre 0,5% aux Etats-Unis).

La situation du marché des éditeurs de logiciels et des SSII en France: représente 31,5 milliards d'euros en 2005. Le marché français du logiciel se monte à 7,2 milliards d'euros en 2005 avec 2500 entreprises comptabilisant 60 000 personnes en France et 8000 à l'international.

### Evaluation du marché de l'offre en matière de technologies linguistiques :

**En Europe (25 pays)**, sur la période 2000-2005, plus de 50 sociétés européennes spécialisées dans le traitement du langage ont disparu (liquidation ou rachat). Les créations de sociétés ont considérablement augmenté dans les années 90 et 2000.

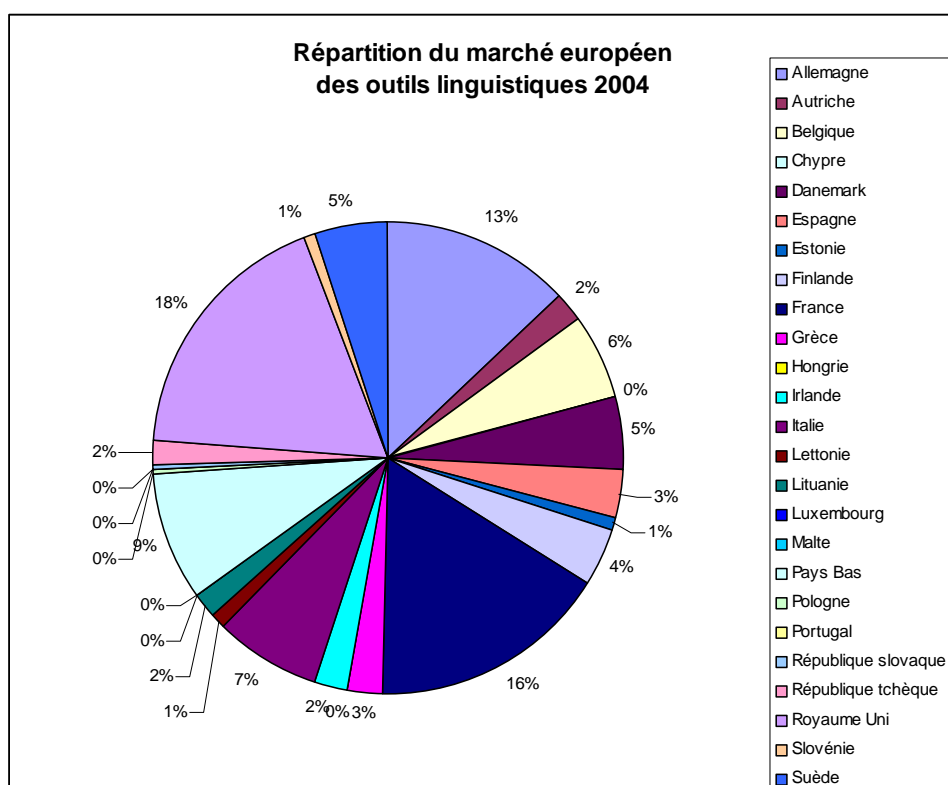
En vingt ans, le nombre de sociétés créées a été multiplié par plus de 10, passant d'une trentaine à plus de 500 sociétés en Europe en 2006.

482 offreurs européens sont présents sur au moins un des 9 segments d'application :

L'offre en Europe se répartit entre le traitement du texte (70%) et le traitement de la voix (30%) ;

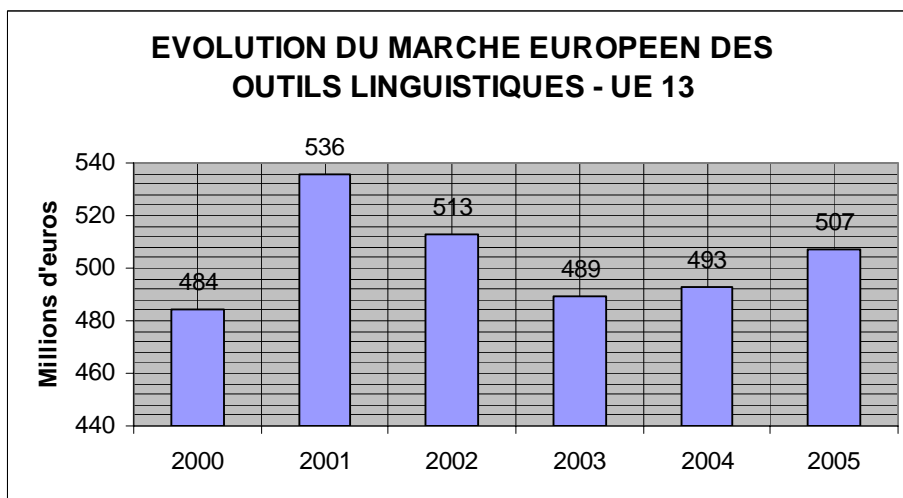
Le marché européen est évalué à 610 millions d'euros en 2004 . Les prévisions sont estimés à 628 millions pour 2005.

### Sur la base des chiffres d'affaires communiqués, le marché européen des outils linguistiques en 2004 est estimé à environ 610 millions d'euros, ainsi répartis:

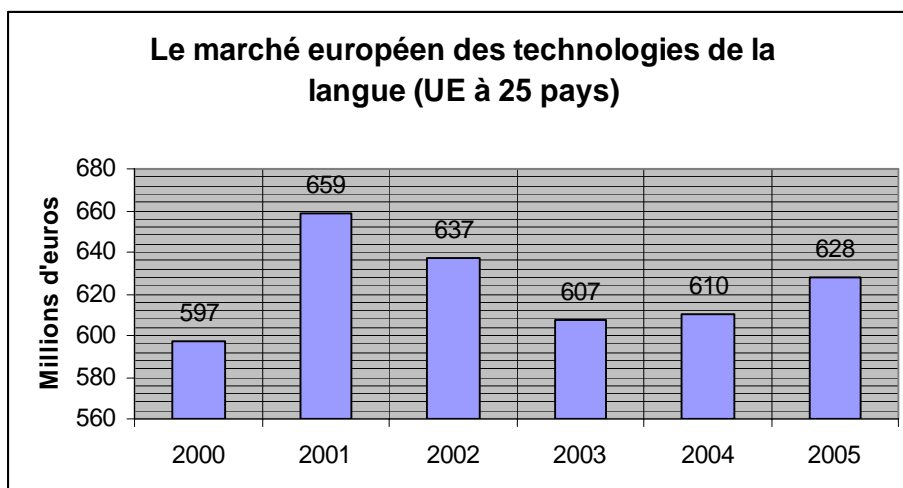


Les pays leaders du marché des outils linguistiques sont donc le Royaume-Uni, la France, et l'Allemagne représentant à eux trois 47% du marché européen. Le Royaume-Uni est leader sur le marché des outils linguistiques. Cette position peut être expliquée par la place prédominante du traitement vocal dans ce pays.

Dans la zone européenne à 13 pays, 394 offreurs européens sont présents sur au moins un des 9 segments d'application. Selon nos estimations actualisées, le marché européen des outils linguistiques (Europe 13 pays) confirme le retour à la croissance amorcée en 2004 après l'éclatement de la bulle internet.

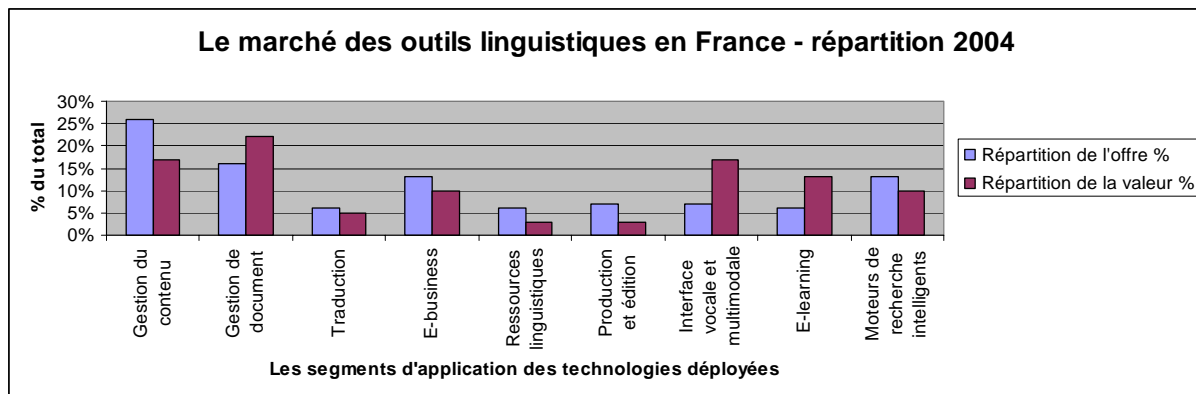


Dans la Zone européenne à 25 pays 482 sociétés ont été identifiées, et seulement 252 ont communiqué leur CA.. L'évolution du marché (zone 25 pays) suit la même tendance que celle constatée dans la zone de l'Union européenne à 13 pays.



**En France**, 109 sociétés ont pu être identifiées dans le périmètre de l'ingénierie linguistique : le chiffre d'affaire dégagé représente environ 78,8 millions d'euros, soit environ 16% du marché européen et la France est le deuxième pays leader après le Royaume-Uni.

En se basant sur les leaders du marché en France, nous pouvons établir la prédominance du traitement du texte sur le marché français : 77% des sociétés leaders du marché français sont positionnées sur le texte, qui est encore aujourd'hui le moteur de la valeur du marché des outils linguistiques en France, cependant les revenus générés par la voix sont en expansion importante.



En 2005, environ 36% (40% en 2003) des acteurs identifiés sont positionnés sur un seul segment d'application. Les sociétés les plus présentes sur le marché des outils linguistiques sont celles traitant la gestion du contenu (26%), la gestion des documents (16%), les moteurs de recherche intelligents (13%), puis les spécialistes du e-business (13%), les sociétés spécialisées dans les interfaces vocales et multimodales (7%), et les sociétés spécialisées dans l'édition et la production (7%).

Les sociétés les plus présentes sur le marché français des outils linguistiques en terme de potentiel de CA sont les sociétés de GED (22%), de gestion de contenu (17%), de reconnaissance et de synthèse vocale (17%), de e-learning (13%), de e-business (11%), et les moteurs de recherche intelligents (10%), de traduction automatique (4%), d'édition et de production (4%) et de ressources linguistiques (3%). En terme de potentiel de CA, les sociétés de GED (27%) et de gestion de contenu (21%) représentaient près de 50% du marché en 2002.

Le marché français reste concentré. Les 10 premières sociétés de l'échantillon français du marché des outils linguistiques représentent 70% du CA, les 10 suivantes seulement 20%. Les 20 dernières font moins de 5% du CA. Le premier acteur, principalement présent sur les segments 7 et 8, présente un chiffre d'affaires très supérieur à ceux des autres sociétés de l'échantillon. La majorité des chiffres d'affaires importants du secteur proviennent souvent d'acteurs qui déploient des activités annexes à l'ingénierie linguistique.

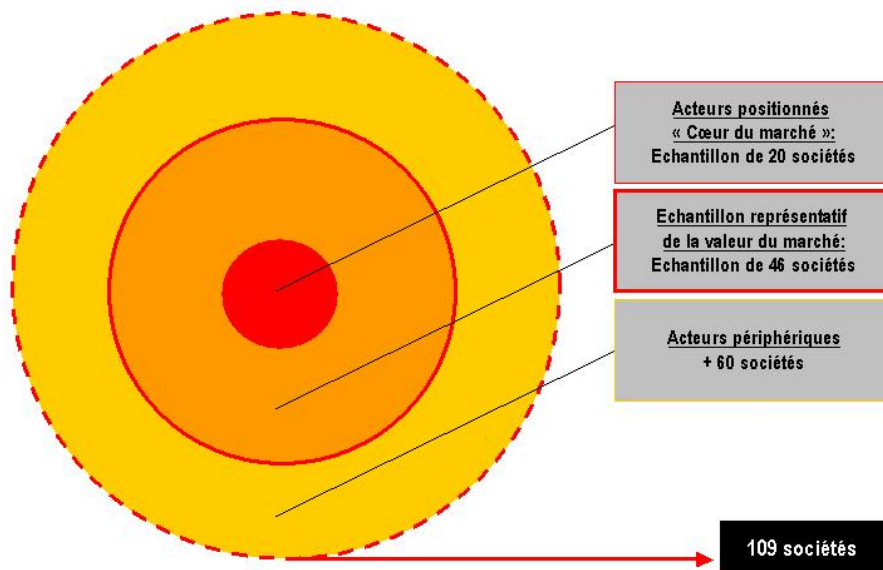
**Remarque :**

Sur le marché des outils linguistiques, le chiffre d'affaires moyen des sociétés les plus petites est de 300 000 euros. Il est de 5 millions d'euros pour les sociétés les plus importantes principalement positionnées sur le secteur des technologies de la langue.

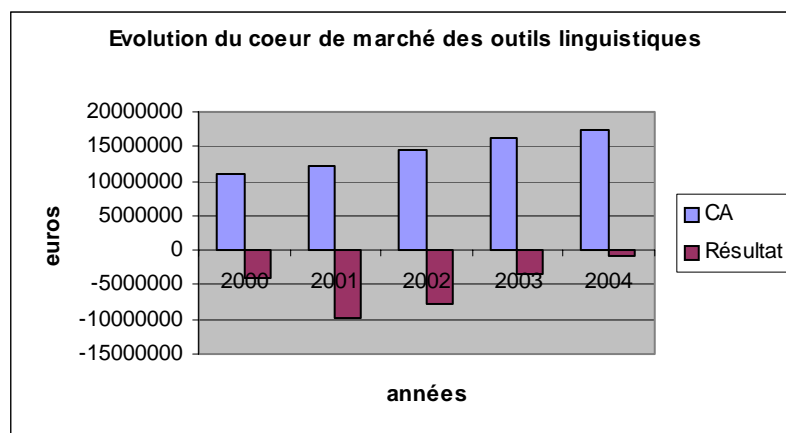
### 3.2. ANALYSE DU CŒUR DE MARCHÉ EN FRANCE

Concernant les 109 acteurs identifiés en France, ils sont situés à différents niveaux sur le marché des outils linguistiques : certains sont historiquement positionnés dans le cœur du marché de l'ingénierie linguistique, d'autres sont placés en périphérie du marché des technologies de la langue.

Le marché des technologies de la langue en France : cœur de marché et périphérie



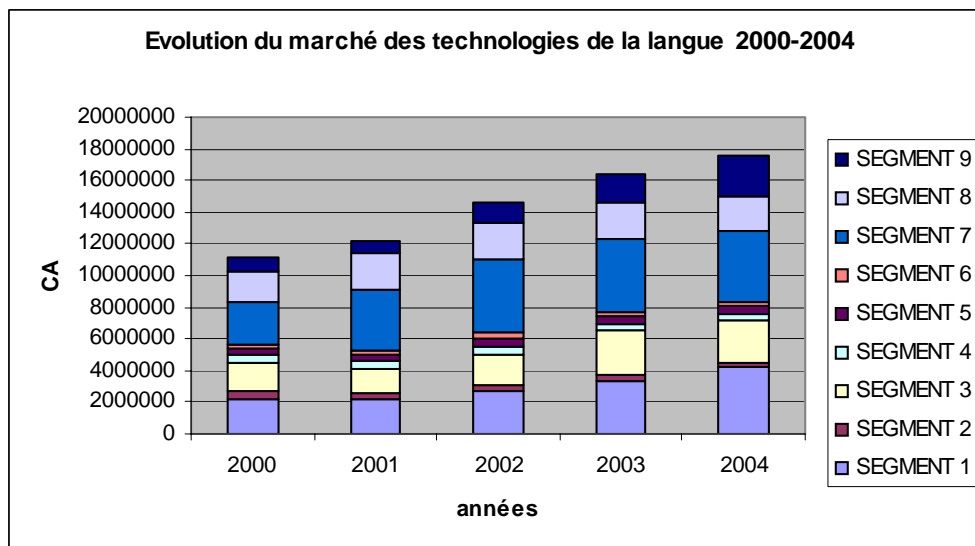
L'analyse des tendances du cœur du marché des technologies de la langue a été conduite à partir d'un échantillon représentatif de 20 sociétés pionnières de l'ingénierie linguistique en France et positionnées dans le noyau dur du marché. La valeur dégagée par ces 20 acteurs positionnés dans le noyau dur du marché a été de 17,5 millions d'euros en 2004 en augmentation de 6,93% par rapport à 2003.



La rentabilité de l'échantillon reste négative sur la période mais les chiffres dégagés concernant le résultat s'améliorent considérablement pour un retour à l'équilibre attendu en 2005 après 5 années de baisse consécutives.

L'évolution de la valeur dégagée par chaque segment illustre la domination de 4 segments qui représentent près de 90% du CA sur la période 2000-2004 : il s'agit du segment 1 « Gestion de contenu », du segment 3 « Traduction automatique », du segment 7 « Applications vocales et multimodales », du segment 8 « E-learning » et du segment 9 « Moteurs de recherche intelligents ».

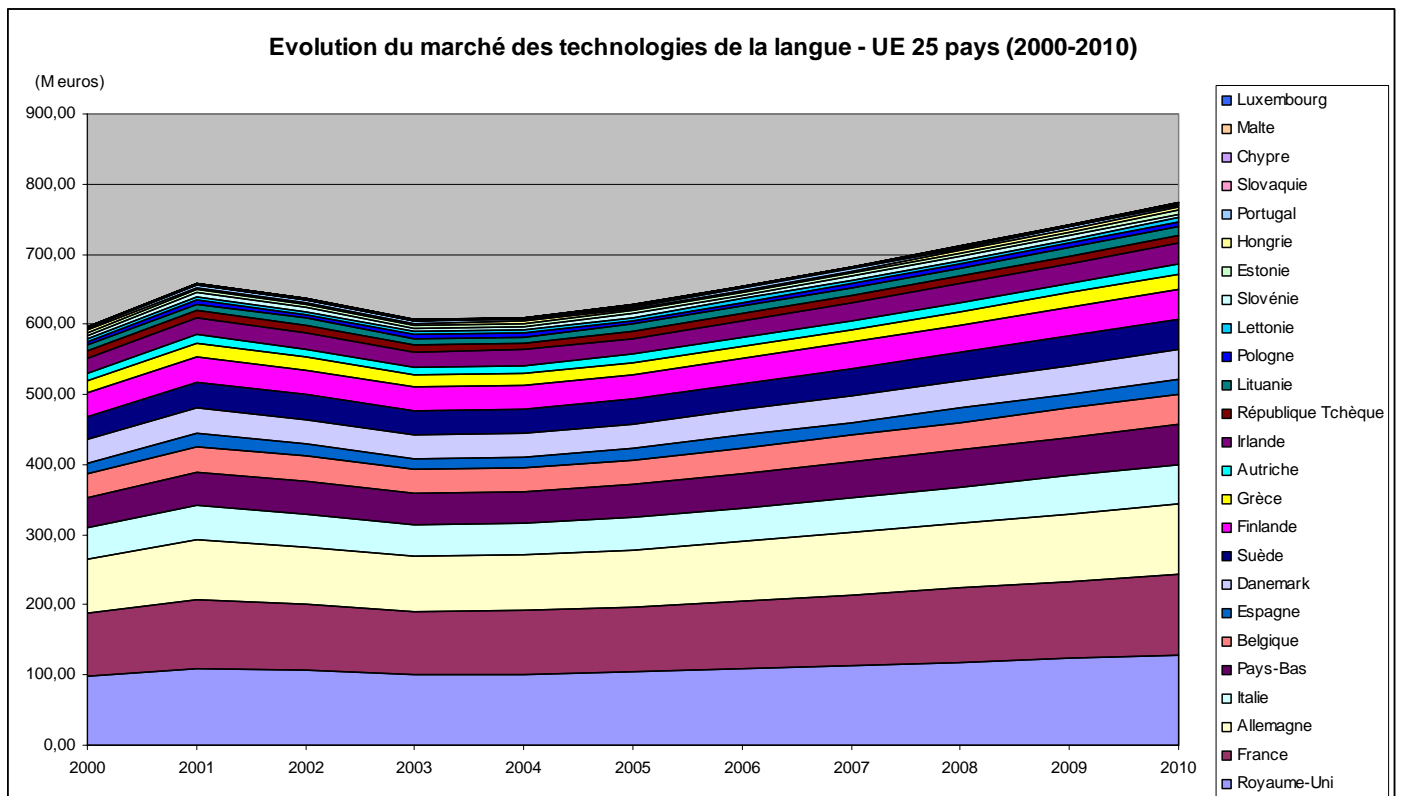
Les taux de croissance annuels moyens sur la période 2001-2004 mettent en lumière une tendance positive marquée avec des taux de croissance à 2 chiffres pour ces 4 segments d'application. Le segment 9 fait figure de segment vedette en ce sens où le taux de croissance annuel moyen est trois supérieur à la moyenne du secteur (31,42% contre 12,08%) et le taux de croissance cumulé a représenté + 179.31% sur la période. Les segments gestion de contenu (+94.24% en cumulé sur la période) et les applications vocales et multimodales (+69.55% en cumulé sur la période) sont elles aussi en forte croissance. Ces 3 segments ont augmenté leur part de marché sur la période.



### 3.3. EVOLUTION DU MARCHE EUROPEEN DES TECHNOLOGIES DE LA LANGUE.

Le marché des technologies de la langue dans la zone de l'Union européenne à 25 pays devrait évoluer de 610 millions d'euros en 2004 à 628 millions d'euros en 2005 et se monter à 774 millions d'euros en 2010.

La courbe de tendance illustrée par le graphique ci-après met une nouvelle fois en évidence la chute du marché après 2001 et sa reprise progressive avec un retour au niveau antérieur entre 2005 et 2007. Les années 2008-2010 s'annoncent comme des années fastes en terme de croissance de chiffres d'affaires mais également en terme de rentabilité économique dans la mesure où les résultats dégagés par les entreprises du secteur semblent revenir à l'équilibre voire à dégager des bénéfices dès l'année 2005 pour certains acteurs.



L'analyse de l'évolution des 9 segments sur la période 2000-2010 en France a été réalisée à partir de l'étude d'un échantillon représentatif du marché. La valeur dégagée par le marché sera située entre 100 millions (hypothèse basse) et 115 millions d'euros environ en 2010 (hypothèse haute).

Concernant le cœur du marché des technologies de la langue en France, le marché est estimé à 45,8 millions d'euros en 2010 (estimation haute).

Les taux de croissance constatés dans le cœur du marché des outils linguistiques sont plus élevés que ceux dégagés par la moyenne du marché (12,08% de 2000 à 2004 pour le noyau dur du marché contre 4,70 pour la moyenne globale).

De même, la rentabilité économique du cœur du marché est souvent inférieure à celle dégagée par le marché : les chiffres d'affaires dégagés par les acteurs du noyau dur sont souvent moins élevés que la moyenne des CA dégagés par des acteurs structurellement plus importants et positionnés à la périphérie du marché, ce qui peut expliquer ce différentiel de croissance.

La maturité de certains segments d'application et la stabilité des acteurs du noyau dur du marché sur une période de 5 ans illustre la nouvelle ère digitale, véritable levier de croissance pour le secteur.

Les perspectives d'évolution du marché des outils linguistiques dépendent avant tout de son intensité concurrentielle qui est de plus en plus forte : l'évolution de l'offre, de la demande et de l'environnement technologique engendre des mouvements de concentration du marché, une mise en concurrence systématique des technologies et des outils disponibles.

Depuis ces dix dernières années, les fusions et acquisitions au sein du marché sont de plus en plus fréquentes et s'expliquent par :

- la frénésie du marché de l'Internet dans les années 1990 avec la course au premier entrant qui a permis de développer des partenariats et de valider des fusions entre acteurs concurrents ;
- le marasme économique du secteur des NTIC dans le début des années 2000 avec l'éclatement de la bulle Internet et la perte de valeur de plusieurs acteurs de l'offre ;
- le positionnement stratégique multinational des offreurs de technologie sur un marché des outils linguistiques qui devient global (interconnexion des acteurs européens et américains), aujourd'hui concrétisé par la pénétration des acteurs américains sur le marché européen ;
- la volonté de contrôle des technologies par la demande, enclin à racheter les sociétés qui ont développé des outils de gestion des connaissances dans un domaine hautement stratégique.

La dynamique du marché se confirme vers une logique de la demande : intégration des outils au sein d'architectures existantes, capitalisation de l'investissement, demande permanente de l'évolution des technologies avec la généralisation du multimédia et de l'Internet mobile, et du besoin de la communication multilingue.

L'offre est toujours diffuse et disparate : la taille des offreurs de technologie est très variée, et leur représentativité nationale n'est pas toujours facile à identifier. De plus, d'une année sur l'autre, le périmètre est flottant voire turbulent : plusieurs sociétés disparaissent (liquidation, rachat ou fusion), d'autres se créent (*spin off* de sociétés grands comptes, de centres de recherche, *start up*).

Les années 2006-2010 pourraient représenter le cercle vertueux de croissance attendu par les différents protagonistes du secteur.

## CONCLUSION

La convergence des technologies est fortement marquée dans le domaine de la communication. L'informatique, les télécommunications et l'audiovisuel sont fédérés par la numérisation. Le contenu numérique concerne désormais le texte, le son et les images. L'ère du multimédia se concrétise avec la mise en oeuvre notamment des kiosques interactifs, des téléphones mobiles de nouvelle génération, des consoles et des terminaux de divertissement et des systèmes embarqués. Les connexions réseaux haut débit se généralisent. Les applications concernent les intranets documentaires multimédia, la gestion de contenu multilingue, le *e-business*, le vocal/multimodal, le *e-learning*, la traduction automatique...

Différents facteurs ont participé ces dernières années au développement du marché européen des technologies de la langue :

- les différents plans de numérisation publics et privés ;
- les réseaux informatiques à très haut débit et sans fil ;
- la mondialisation accrue du marché des technologies de la langue avec l'augmentation des partenariats internationaux (stratégiques, commerciaux, technologiques) et la hausse induite des besoins en traduction instantanée, qu'elle soit vocale ou textuelle, liée aux facteurs socio-culturels dont le multilinguisme ;
- l'Open source, à la fois :
  - opportunité de développement : de plus en plus stables, évolutives et puissantes, les solutions Open source restent gratuites et permettent de capitaliser des modèles et des standards reconnus.
  - et, contrainte concurrentielle obligeant à la baisse des coûts : les développements associés sont parfois lourds et longs, ce qui peut être un frein au développement d'outils open source européens et compétitifs dans la mesure où les pays émergents qui maîtrisent ces langages informatiques, capitalisent un savoir-faire technique, et ont su acquérir des compétences reconnues, peuvent devenir de sérieux concurrents pour la réalisation de projets multilingues et le développement de technologies de la connaissance à moindre coût.

Plusieurs facteurs incitent les décideurs à intégrer dans l'entreprise des solutions innovantes pour la gestion intelligente des contenus numériques /

- la convergence numérique des technologies de l'informatique, de l'information et de la communication ;
- l'intégration des contenus multimédias ;
- l'Internet sans fil et haut débit ;
- les applications du web 2.0. et du web 3.0. favorisant la contribution éditoriale et sociale des utilisateurs-prescripteurs de contenus et de services.

La progression de l'usage des NTIC laisse présager un élargissement du marché des outils linguistiques vers le grand public. Le besoin est ressenti d'engager des actions *marketing* pour optimiser l'adéquation de l'offre à la demande.

Bureau van Dijk Information Management - Martine DEJEAN : [mde@bvdim.com](mailto:mde@bvdim.com) - Antoine RAULIN : [ara@bvdim.com](mailto:ara@bvdim.com)