

**MINISTERE DELEGUE A L'ENSEIGNEMENT  
SUPERIEUR ET A LA RECHERCHE**

**Direction générale de la Recherche et de l'Innovation**

Direction de la Stratégie

Mission de l'Information Scientifique et Technique

**Technologies de la langue en Europe  
Marché et tendances**

*Rapport final*

Table des matières :

<b>1.</b>	<b>Objectifs de l'étude.....</b>	<b>4</b>
<b>2.</b>	<b>Méthodologie.....</b>	<b>5</b>
2.1.	Exploitation des publications. ....	5
2.2.	Entretiens complémentaires auprès des acteurs de l'offre. ....	5
2.3.	Lissage des données.....	5
<b>3.</b>	<b>Périmètre de l'étude.....</b>	<b>6</b>
3.1.	Périmètre géographique.....	6
3.2.	Périmètre technologique.....	6
<b>4.</b>	<b>Phase 1 : Caractérisation du marché de l'offre en Europe et en France.....</b>	<b>8</b>
4.1.	Les indicateurs structurels et les facteurs conjoncturels du marché.....	8
4.1.1.	Les indicateurs structurels.....	9
4.1.1.1.	Les Produits Intérieurs Bruts de l'Union européenne.....	9
4.1.1.2.	Les Dépenses en TIC dans l'Union européenne et le monde.....	11
4.1.1.3.	La politique d'investissement en R&D dans les pays européens.....	12
4.1.1.4.	Les chiffres clés en France.....	13
4.1.1.4.1	La politique d'investissement en R&D en France.....	13
4.1.1.4.2	Le marché du logiciel en France.....	14
4.1.2.	Les facteurs conjoncturels.....	15
4.1.2.1.	Le stockage et la répartition des données dans les réseaux .....	15
4.1.2.2.	La place du secteur des NTIC dans l'économie européenne .....	15
4.1.2.3.	Fusions / acquisitions.....	16
4.1.2.4.	Les 6M et les usages numériques.....	17
4.2.	Marché européen des technologies de la langue (13 pays) : actualisation des données de l'étude conduite en 2003.....	19
4.2.1.	Identification et répartition des acteurs.....	19
4.2.2.	Evaluation du marché européen (13 pays).....	20
4.2.3.	Comparaison des résultats entre les chiffres de 2002 et de 2004.....	21
4.3.	Marché européen des technologies de la langue (25 pays) : intégration des 12 autres pays de l'Union européenne .....	24
4.3.1.	Identification et répartition des acteurs.....	24
4.3.2.	Evaluation du marché européen (25 pays).....	25
4.3.3.	Présence des acteurs sur le marché européen.....	27
4.3.4.	Valeur dégagée par le secteur des technologies linguistiques en Europe.....	28
4.4.	Marché français des technologies de la langue: actualisation des données de l'étude 2003.....	32
4.4.1.	Identification et répartition des acteurs.....	32
4.4.2.	Evaluation du marché.....	34
4.4.3.	Comparaison par rapport à 2002.....	37
4.4.4.	Constitution d'un échantillon de 46 sociétés représentatives du marché français des technologies de la langue et répartis par segments d'application.....	38

<b>5.</b>	<b>Phase 2 : France – analyse du noyau dur du marché .....</b>	<b>44</b>
	Constitution d'un échantillon de 20 sociétés représentatives du cœur du marché français des technologies de la langue .....	44
5.1.	Evolution globale du CA et du résultat de l'échantillon.....	45
5.2.	Evolution par segment d'application du cœur du marché.....	46
5.3.1.	Segment 1 : gestion du contenu.....	48
5.3.2.	Segment 2 : gestion des documents.....	49
5.3.3.	Segment 3 : traduction automatique.....	50
5.3.4.	Segment 4 : e-business.....	51
5.3.5.	Segment 5 : ressources linguistiques.....	52
5.3.6.	Segment 6 : production et édition.....	53
5.3.7.	Segment 7 : interfaces vocales et multimodales.....	54
5.3.8.	Segment 8 : e-learning.....	55
5.3.9.	Segment 9 : moteurs de recherche intelligents.....	56
	<b>Phase 3 : Tendances et perspectives du marché des technologies de la langue en France et en Europe.</b>	<b>57</b>
<b>6.</b>	<b>Analyse des tendances du marché français.....</b>	<b>57</b>
6.1.	Cartographie des acteurs.....	57
6.1.1.	Analyse concurrentielle et marketing du marché français.....	59
6.2.	Perspectives de croissance du marché des technologies de la langue en France.....	67
6.2.1.	Perspectives de croissance du marché français à l'horizon 2010.....	67
6.2.1.1.	Evolution globale du marché français.....	67
6.2.1.2.	Evolution des segments du marché français.....	69
6.2.2.	Evolution du cœur du marché français des technologies de la langue à l'horizon 2010.....	69
6.2.2.1.1.	Evolution globale du cœur de marché .....	69
6.2.2.1.2.	Evolution des segments du cœur de marché .....	72
6.3.	Tendances et perspectives du marché en Europe.....	73
6.3.1.	Impact des NTIC.....	73
6.3.2.	Perspectives de croissance du marché des outils linguistiques en Europe.....	74
6.3.2.1.	UE 25 pays.....	74
6.3.2.2.	UE 13 pays.....	77
6.3.3.	Evolution texte / vocal.....	79
6.3.3.1.	UE 25 pays.....	79
6.3.3.2.	UE 13 pays.....	79
<b>7.0.</b>	<b>Conclusions.....</b>	<b>80</b>

## **1. Objectifs de l'étude.**

Le présent document constitue le rapport final de l'étude de marché « Technologies de la langue en Europe : Marché et tendances » conduite en 2005-2006, réalisée à la demande de la Direction de la Stratégie du Ministère délégué à l'Enseignement supérieur et à la Recherche. Il s'agit de l'actualisation de l'étude précédemment conduite en 2003-2004, et portant sur les 13 pays de l'Union européenne (Allemagne, Autriche, Belgique, Danemark, Espagne, Finlande, France, Grèce, Irlande, Italie, Pays-Bas, Royaume-Uni, Suède), avec une extension aux 12 autres pays membres de l'UE, dont 10 nouveaux pays entrants en 2004 (Chypre, Estonie, Hongrie, Lettonie, Lituanie, Luxembourg, Malte, Pologne, Portugal, République tchèque, Slovaquie, Slovénie).

Les objectifs de cette étude consistent donc à qualifier la nature, la taille et le développement du marché des outils linguistiques axés sur le traitement des données textuelles ou vocales dans les différents pays européens, avec une attention particulière pour le marché français.

L'étude a été conduite en 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.

## **2. Méthodologie.**

Afin de pouvoir comparer les résultats, la même méthodologie que pour l'étude conduite en 2003 a été appliquée, à savoir :

- la recherche d'informations qualitatives et quantitatives, à partir des publications et des portails d'information sur Internet, et les sites web des principales sociétés du marché, des instituts d'études, et des associations professionnelles... ;
- les bases de données d'information économique et financière sur les entreprises ;
- l'extraction des indicateurs clés du marché ;
- des entretiens complémentaires réalisés auprès des acteurs du marché des outils linguistiques.

### **2.1. Exploitation des publications.**

L'analyse des publications existantes a permis :

- de confirmer la segmentation du marché définie lors de la précédente étude ;
- d'appréhender l'évolution des applications mises en œuvre ;
- d'extraire des différentes études économiques les indicateurs clés permettant de chiffrer le marché actuel et d'en préfigurer les tendances.

Les principales sources consultées figurent en annexe du présent document (annexe 1).

### **2.2. Entretiens complémentaires auprès des acteurs de l'offre.**

Les sociétés acteurs de l'offre (et de la demande) ont été contactées lors de différents salons professionnels, ou consultées au téléphone, pour :

- valider les informations ;
- repérer les tendances et envisager les perspectives.

### **2.3. Lissage des données.**

Plusieurs difficultés pour comparer les résultats dégagés d'une période à une autre ont dû être contournées :

- plusieurs sociétés n'ont pas communiqué leurs chiffres d'affaires ou les ont renseignés épisodiquement ;
- certaines sociétés ont communiqué des chiffres d'affaires européens consolidés (sans justifier d'une présence exhaustive à l'ensemble des pays européens) ;
- la volatilité des acteurs est assez importante : l'émergence de nouvelles sociétés et la disparition de certains acteurs (cessation d'activités, fusions et acquisitions dans le secteur) ne permettent pas l'établissement d'une photographie précise de ce marché ainsi que la définition claire du périmètre du marché des outils linguistiques.

Il a été décidé de suivre les sociétés dont les chiffres d'affaires étaient disponibles sur 5 ans de manière à obtenir une visibilité pertinente et plus fiable du marché. Par contre, l'analyse qualitative des acteurs a porté sur les sociétés présentes sur le marché en 2005-2006.

Un overlap de 60% sur le chiffre d'affaires global du marché des outils linguistiques a été retenu, tout comme lors de la précédente étude, afin de pas surestimer la valeur de ce marché. Certaines sociétés sont à la fois éditeurs de logiciels et offreurs de services.

Par ailleurs, des doublons de valeurs peuvent exister :

- les chiffres d'affaires de la majorité des acteurs ne sont pas exclusivement basés sur la vente de produits linguistiques : des services de formation, des technologies et des gestions de projets annexes au périmètre des outils linguistiques sont à écarter.
- certaines sociétés présentes sur plusieurs territoires déclarent des chiffres d'affaires consolidés dans un seul pays.

### **3. Périmètre de l'étude.**

Un périmètre géographique ainsi qu'un périmètre technologique ont été définis de manière à pouvoir identifier les acteurs positionnés sur des segments d'applications liés aux usages des technologies de la langue.

#### **3.1. Périmètre géographique.**

Cette étude porte sur :

- la France ;
- les 13 pays correspondant à l'étude 2004 : Allemagne, Autriche, Belgique, Danemark, Espagne, Finlande, France, Grèce, Irlande, Italie, Pays-Bas, Royaume-Uni, Suède ;
- les 25 pays de l'Union européenne.

#### **3.2. Périmètre technologique.**

La segmentation basée sur l'usage plutôt que sur les technologies, c'est-à-dire en fonction des modes d'utilisation et d'application des outils linguistiques, a été reprise pour la mise à jour de cette étude.

En effet, les typologies classiques basées sur une segmentation par brique technologique (analyseurs morphologiques, syntaxiques, sémantiques...) sont difficilement utilisables pour une étude de marché : plusieurs de ces briques technologiques peuvent être actives dans plusieurs segments, rendant ainsi la ventilation des chiffres d'affaires des acteurs du marché très difficile voire impossible.

Nous rappelons les 9 segments :

1. **Gestion du contenu** : knowledge management, veille, intelligence économique, analyse automatique de formulaires ou de textes (indexation, recherche de texte, visualisation cartographique, analyse).
2. **Gestion des documents** : gestion des archives et gestion multidocument, travail collaboratif assisté par ordinateur (catégorisation, métadonnées et taxonomies, workflow).
3. **Traduction** : traduction automatique et traduction assistée par ordinateur (traduction, multilinguisme, localisation).
4. **E-business** : portails Internet B2B et B2C, communautés d'internautes, accès aux catalogues (personnalisation des informations, constitution de terminologies).
5. **Ressources linguistiques texte et voix** : thésaurus, dictionnaires, corpus d'entraînement et outils associés (dictionnaires, thésaurus, grammaire, phonèmes et échantillons de voix).
6. **Production et édition** : correcteurs, aidé à la rédaction, génération, résumés, dictée vocale, text to speech (rédaction, résumés, correcteurs orthographiques, outils d'aide à la rédaction, générateurs de texte et de liens hypertextes).
7. **Interface vocale et multimodale** : call centers, kiosques et bornes multimédias, solutions embarquées et domotique, jeux vidéos et divertissement (automatisation, reconnaissance et synthèse vocale, navigation multimédia et interface visuelle).
8. **E-learning** : e-learning, applications pédagogiques, apprentissage des langues, aide au handicap (formation par Internet, outils pédagogiques multimédia, apprentissage des langues, accès à l'informatique).
9. **Moteurs de recherche intelligents** : outils de recherche pour Internet, Intranet et Extranet (traitement du langage naturel, web sémantique).

## 4. Phase 1 : Caractérisation du marché de l'offre en Europe

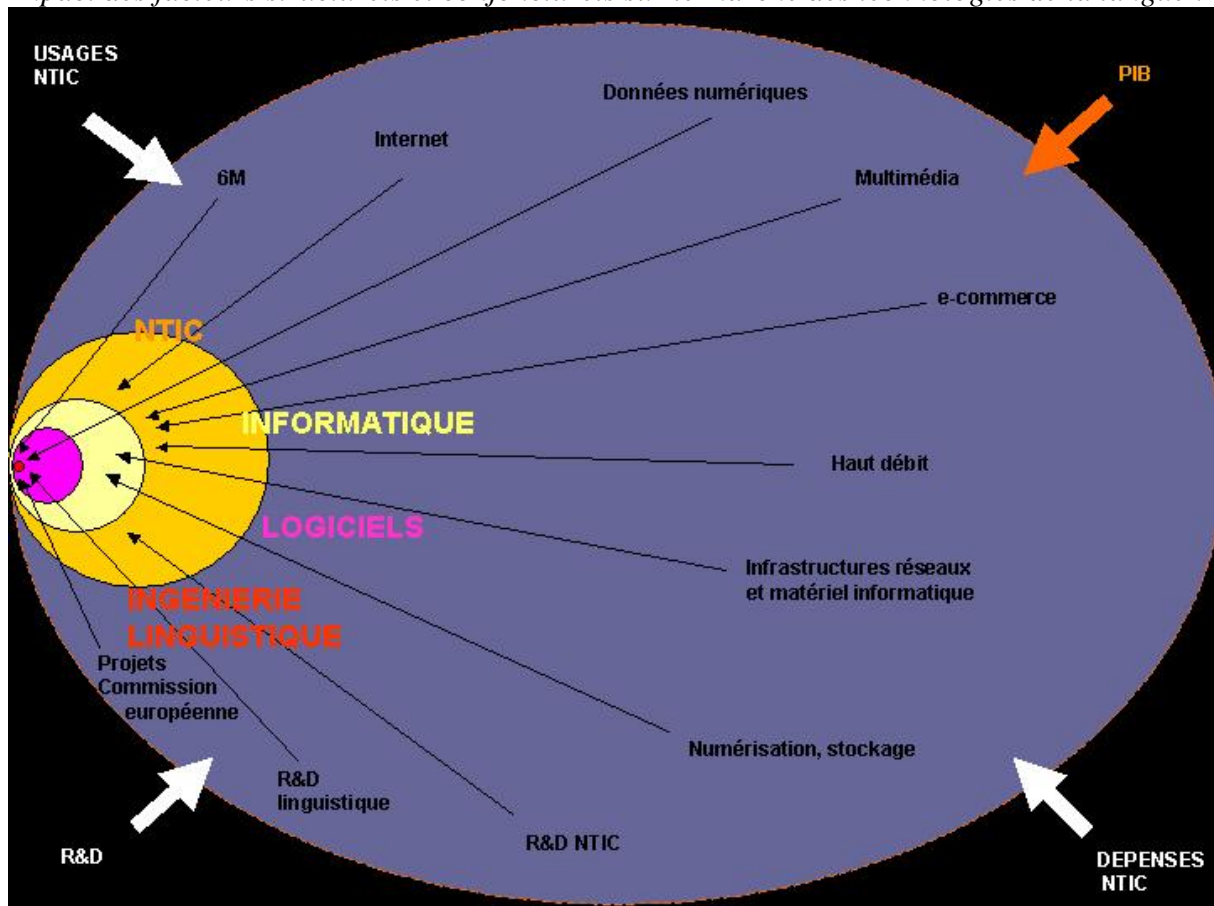
### 4.1. Les indicateurs structurels et les facteurs conjoncturels du marché des technologies de la langue.

Le marché des outils linguistiques est directement ou indirectement influencé par l'évolution des différents éléments structurels et conjoncturels des TIC :

- les dépenses en NTIC en Europe,
- les investissements en Recherche et Développement,
- le marché de l'informatique et des logiciels,
- la numérisation des données dans les réseaux,
- divers facteurs socio-économiques et culturels tels que le multilinguisme, Internet, la convergence numérique et les contenus multimédia...

Il s'agit de comparer et de vérifier les tendances évoquées en 2004.

*Impact des facteurs structurels et conjoncturels sur le marché des technologies de la langue :*





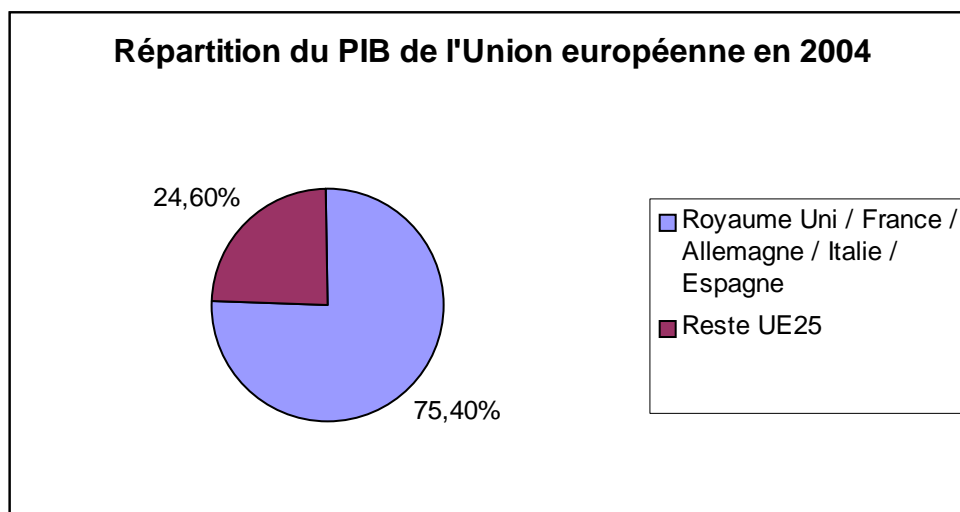
#### 4.1.1. Les indicateurs structurels.

##### 4.1.1.1. Les Produits Intérieurs Bruts de l'Union européenne.

Selon l'OCDE, les 25 principaux acteurs de l'Union Européenne comptabilisent une population de 460 millions de personnes en 2005 (dont plus 160 millions d'internautes, soit environ 40 % d'internautes) et 55% d'entre elles parlerait au moins une seconde langue. En 2002, les 13 pays étudiés comptaient 370 millions de personnes.

Selon l'OCDE,

- En 2002, la richesse produite par la zone Europe de 13 pays était de 7 778 milliards d'euros. L'Allemagne, la France, l'Italie et le Royaume-Uni représentent à eux seuls 73% du PIB des 13 pays. Les dépenses en NTIC (particuliers, administrations, entreprises) correspondent à 6,35% du PIB, soit 493 milliards d'euros.
- En 2004, la richesse produite par la zone Europe de 25 pays était de 10 400 milliards d'euros, avec un taux de croissance de 2,4% en 2004, et une inflation de 2,1%. Les 5 premiers acteurs du marché européen à 25 pays en terme de PIB sont l'Allemagne, la France, l'Italie, le Royaume-Uni et l'Espagne : ils représentent à eux seuls 75% du PIB des 25 pays.



Après calcul des parités de pouvoir d'achat (PPA), le PIB de la zone Europe des 25 pays se répartit ainsi (Source: Eurostat 2004):

PIB 2004 Union européenne	%	Milliards d'euros
Allemagne	22,00%	2288
Royaume-Uni	17,30%	1799,2
France	15,80%	1643,2
Italie	13,10%	1362,4
Espagne	7,20%	748,8
Pays-Bas	4,60%	478,4
Belgique	2,70%	280,8
Suède	2,70%	280,8
Autriche	2,30%	239,2
Pologne	2,10%	218,4
Danemark	1,90%	197,6
Grèce	1,50%	156
Finlande	1,40%	145,6
Irlande	1,30%	135,2
Portugal	1,30%	135,2
République Tchèque	0,80%	83,2
Hongrie	0,70%	72,8
Slovaquie	0,30%	31,2
Lituanie	0,20%	20,8
Slovénie	0,20%	20,8
Luxembourg	0,20%	20,8
Lettonie	0,10%	10,4
Estonie	0,10%	10,4
Chypre	0,10%	10,4
Malte	0,10%	10,4
TOTAL (milliards d'euros)	100,00%	10400

Il doit être noté que la zone Europe avec les 13 pays étudiés représente en 2004 93,80% du PIB dégagé par l'Union européenne à 25 pays.

#### 4.1.1.2. Les dépenses en TIC dans l'Union européenne et le monde.

- Technologies de l'information et de la communication

En 2004, le volume d'affaires lié à la dépense informatique dans le monde est de 965 milliards de dollars. Le marché doit atteindre les 1 200 milliards de dollars en 2008 (source IDC).

Le cabinet IDC a évalué le marché des technologies de l'information et de la communication (TIC) en Europe à 625 milliards d'euros en 2005. L'Europe constituerait le premier marché des TIC, mais sa part tend à diminuer en faveur des pays émergents asiatiques et des Etats-Unis où la croissance y est plus élevée.

Sur les deux secteurs composant la dénomination TIC, les technologies de l'information affichent la progression la plus rapide, à 3,8% contre 2,5% pour les télécoms. Le dynamisme des TI est à mettre à l'actif du marché des logiciels qui, à la faveur d'une forte demande en infrastructures systèmes et en solutions de sécurité, croît de 6,3% sur un an. Les services IT, quant à eux, atteignent une croissance de 5,3%. Du côté du matériel, les revenus pâtitent de la guerre des prix que se livrent les constructeurs et reculent de 0,3% pour les serveurs et de 1,2% pour les PC.

- Technologies de l'information

Selon Aberdeen Group, les dépenses en technologies de l'Information en Europe sont d'environ 402 milliards de dollars en 2004 (soit approximativement 65% des dépenses totales en NTIC), et de 410 milliards de dollars en 2005 :

Monde : évolution de l'investissement IT par zone (en milliards de dollars)						
Zone	2000	2001	2002	2003	2004	2005
Afrique/Moyen-Orient	30,9	32	34,1	35,9	37,8	39,8
Asie-Pacifique	250,6	251,8	258,1	274,5	292,4	310,9
Europe	375,6	377,4	384,6	393,1	402	410,5
Amérique du Sud	65,6	65,7	68,4	72,9	78,2	83,5
Amérique du Nord	477,4	475,7	494,6	522,3	550,2	578,2
<b>Total</b>	<i>1200,1</i>	<i>1202,6</i>	<i>1239,8</i>	<i>1298,7</i>	<i>1360,6</i>	<i>1422,9</i>
<b>Taux de croissance</b>		0,21%	3,09%	4,75%	4,77%	4,58%

Source Aberdeen Group

La croissance régionale observée a été de 1,91% en Europe en 2002: cette évolution positive se confirme en 2003 (+2,21%) et en 2004 (+2,26%) mais ces taux de croissance restent relativement faibles lorsqu'ils sont comparés aux taux de croissance constatés en Amérique du Nord. En effet, l'évolution de l'investissement IT dans la zone de l'Amérique du Nord a été de 3,97% en 2002, de 5,60% en 2003 et de 7,27% en 2004 soit 3 fois plus en moyenne que la croissance des dépenses en technologies de l'information en Europe.

Plusieurs facteurs socio-culturels dont le multilinguisme sont à la fois sources d'opportunité et de développement, mais ils peuvent aussi être perçus comme des contraintes supplémentaires freinant la dynamique de développement du marché. La situation comparée entre l'Europe et les Etats-Unis renforce cette idée dans la mesure où le secteur des NTIC y est plus dynamique, la pénétration des acteurs plus forte et les normes en vigueur plus rapide à mettre en œuvre...

#### **4.1.1.3. Investissements en R&D des pays européens.**

- Investissements en R&D en Europe

Les politiques d'investissement en R&D sont très différentes d'un pays à l'autre et marquent des positionnements stratégiques distincts :

- Positionnement global en Europe :

En 2004, l'Union européenne à 25 pays a dépensé près de 200 milliards d'euros au titre de la recherche et du développement. L'intensité de R&D, mesurée par la part des dépenses en pourcentage du PIB, s'est établie à 1,90% dans l'UE25 contre 1,92% en 2003 (proportion étant restée stable sur la période 1999-2004). L'intensité de R&D est restée nettement inférieure dans l'UE25 par rapport aux autres grandes économies. En 2003, les dépenses de R&D représentaient 2,59% du PIB aux Etats-Unis, 3,15% au Japon, tandis qu'elles comptaient pour 1,31% du PIB en Chine. En termes réels, les dépenses de R&D dans l'UE25 ont augmenté en moyenne de 1,3% par an entre 2001 et 2004, contre -0,1% aux Etats-Unis et +1,8% au Japon entre 2001 et 2003.

Plus de 60% du total des dépenses de R&D ont été réalisées par l'Allemagne (55 milliards d'euros), la France (36 milliards) et le Royaume-Uni (30 milliards). Le montant total des dépenses de R&D de ces pays est resté relativement stable au cours de la période 2001-2004, croissant modérément en termes nominaux en Allemagne et en France, moins de 3%, et enregistrant une faible baisse au Royaume-Uni. Les hausses les plus importantes en termes nominaux étaient observées en Estonie (19%), à Chypre (18%), en Lituanie (14%) et en Espagne (13%). En termes réels, les dépenses de R&D de l'UE-25 ont augmenté de 1,3% au cours de la même période. Au niveau national, les taux de croissance les plus élevés ont été en Estonie (16%), à Chypre (15%), et en Lituanie (12%).

Seuls 3 pays ont dépassé l'objectif de 3% de dépenses en R&D par rapport au PIB (fixé lors du Sommet de Lisbonne de 2000) : la Suède (3,74%), la Finlande (3,51%) et l'Islande (3,01%). En termes nominaux les dépenses de R&D dans l'UE à 25 pays ont enregistré une croissance de 2,7% par an entre 2001 et 2004.

- Répartition de l'investissement en R&D :

En 2004, dans l'UE à 25 pays, la R&D a été principalement financée par le secteur des entreprises avec 54% des fonds, contre 35% pour le secteur des administrations publiques.

Dans le secteur des entreprises en 2004, les intensités de R&D les plus élevées ont été observées dans les pays nordiques, la Finlande étant en tête avec 2,75%, suivi par la Suède (2,46%), le Danemark (1,81%) et l'Allemagne (1,75%). Dans l'UE-25, l'intensité de R&D du secteur des entreprises s'est élevée à 1,22% en 2004. Par rapport aux autres grandes économies, l'UE-25 s'est placée derrière le Japon (2,36%) et les Etats-Unis (1,79%) mais devant la Chine (0,82%).

Dans l'UE-25, le poids relatif des différentes sources de financement de la R&D est resté inchangé en 2004 par rapport à 2003. Le secteur des entreprises reste le secteur qui contribue le plus avec 54% du total : ce pourcentage reste néanmoins sensiblement inférieur à celui observé au Japon (75%), aux Etats-Unis (63%) et en Chine (60%). Le secteur des administrations publiques venait en seconde position, prenant à sa charge 35% des dépenses de R&D de l'UE-25, pour 31% en Chine, 30% aux Etats-Unis et seulement 18% au Japon. Dans toutes les grandes économies, le financement par l'étranger était mineur, moins de 2%, excepté dans l'UE-25 où il atteignait 9%.

- Comparaison entre l'Europe et les USA :

L'étude réalisée par MV Etudes et Conseil pour le compte du GFII sur l'évolution de la Recherche et du Développement dans les Sciences et les Technologies de l'Information et de la Communication montre que le montant total des investissements de R&D STIC des USA est le double que celui dégagé par l'Union européenne à 15 pays (dont Allemagne, Espagne, Finlande, France, Italie, Pays-Bas, Royaume-Uni, Suède). Cette tendance est à corréluer aux investissements en Technologies de l'information où la croissance observée aux USA est en moyenne 3 fois supérieure à la croissance des investissements européens. Au Japon, le montant des investissements en R&D STIC est au même niveau qu'en Europe. Par contre il doit être noté que l'effort public est plus important en Europe qu'au Japon.

#### **4.1.1.4. Les chiffres clés en France.**

##### **4.1.1.4.1. La politique d'Investissement en R&D en France.**

Le budget français consacré à la recherche scientifique, soit toutes les Dépenses Intérieures consacrées à la R&D (DIRD) en France, atteint 2,23 % du PIB en 2004. Ce chiffre situe la France dans la bonne moyenne mondiale et la place au-dessus de la moyenne de l'UE (1,99 %).

La part des dépenses consenties dans ce domaine par les entreprises sur le total des investissements R&D en France est de 52 % en 2004 (soit 1,36 % du PIB). Ce taux est relativement faible, comparé au poids des firmes allemandes (60 % du total des investissements R&D), des sociétés américaines (68 %) et des sociétés japonaises (72 % du secteur privé). Le tableau de répartition des investissements R&D entre le privé et le public figure en annexe du présent rapport.

#### 4.1.4.1.2. Le marché du logiciel

Après une reprise amorcée en 2004, le secteur français des logiciels et des services informatiques doit confirmer en 2005 son retour vers la croissance : selon Syntec Informatique, le chiffre d'affaires cumulé des éditeurs de logiciels et des services informatiques devrait croître de 6 à 7% en 2005 par rapport à 2004, pour atteindre environ 31,5 milliards d'euros avec 300 000 salariés. En 2004, la croissance était déjà de 4%, contre un recul de 5% en 2003 et de 3% en 2002. Les prévisions établies pour l'année 2006 resteraient positives avec des taux de croissance prévus entre 5 et 7%.

Selon l'étude Syntec-CXP, avec près de 2500 sociétés éditrices de logiciels, le marché français du logiciel emploie environ 60 000 personnes en France et 8000 à l'international en 2005 pour un chiffre d'affaires de 7,2 milliards d'euros. L'industrie du logiciel s'inscrit comme le principal moteur de la croissance du secteur informatique : entre 1999 et 2005, les dépenses en logiciel en France ont augmenté de 116%, soit deux fois plus rapidement que les dépenses en matériel (51%).

#### **Synthèse des facteurs structurels**

- Les dépenses en NTIC en Europe à 25 pays (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 (secteur privé=65%). Allemagne, France, et Royaume-Uni en représentent 120 milliards d'euros, soit 60% du total.

- La R&D orientée linguistique est estimée à 400 millions de dollars en 2002, et 800 millions de dollars en 2003 (source : SRC UK). La R&D orientée linguistique représente 1 à 25 % du CA des sociétés offeuses de technologie, et 1 à 30% de leur masse salariale. Pour rappel, le programme FP6 représente 3,2 milliards d'euros (2003-2006) et le budget HLT compte pour environ 400 millions d'euros. 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: 31,5 milliards d'euros en 2005. Marché français du logiciel : 7,2 milliards d'euros en 2005 avec 2500 entreprises comptabilisant 60 000 personnes en France et 8000 à l'international.

## 4.1.2. Les facteurs conjoncturels.

### 4.1.2.1. Le stockage et la répartition des données dans les réseaux

La capacité de stockage, à prix constant, suit la loi de Moore définie pour les processeurs : à l'instar du doublement de la capacité des processeurs tous les 18 mois, les capacités de stockage numérique doublent sur une même période. Le volume à stocker a augmenté de manière exponentielle, et le stockage constitue une problématique des entreprises en terme de coût (actuellement une grande partie du budget informatique) et de gestion des données stockées (conditions de conservation et de réutilisation).

La nature des données évolue également : si les données structurées (comme les bases de données par exemple) connaissent une croissance annuelle de 4% en 2006, elles ne représentent que le dixième des informations qui composent le patrimoine numérique d'une organisation informatisée. Les 9/10e restants concernent la messagerie électronique, les fichiers multimédia et les présentations type PowerPoint, qui doublent de volume tous les deux mois pour une croissance cumulée de 6400% l'an (source IDC). Le Radicati Group estime que la taille moyenne d'un courrier électronique de nature professionnelle avec pièce jointe est actuellement de 300 Ko et que la part de ces courriels associés à des documents à valeur de preuve (contrats, propositions commerciales, factures, documents sociaux, etc.) a crû de 60% pour la seule année 2005.

### 4.1.2.2. La place du secteur des NTIC dans l'économie européenne.

Selon l'Institut de l'audiovisuel et des télécommunications en Europe (IDATE), en 2005, le marché mondial des TIC représente 2 681 Md€.

(milliards d' €)	2000	2001	2002	2003	2004	2005
Équipements de télécommunication	306	276	256	252	251	252
Services de télécommunication	732	811	859	916	971	1025
Matériels informatiques	276	252	228	228	244	263
Logiciels et services informatiques	498	511	505	528	573	623
Services audiovisuels	260	268	289	309	326	342
Électronique grand public	121	133	145	152	165	176
Total	2 194	2 251	2 283	2 385	2 530	2 681

Source Idate

Selon le cabinet anglo-saxon IDC, en 2006, les pays européens devraient totaliser près de 33,6% de la dépense mondiale dans les TIC (estimé par IDC à 2 027 milliards d'euros), devant les Etats-Unis avec 28% et le Japon avec 14,3%. La croissance devrait ralentir, après les années 2004 et 2005 où la hausse atteignait respectivement 3,8 et 4,1%.

Malgré tout, les TIC restent un relais de croissance en comparaison du PIB (Produit intérieur brut) des pays européens, situés à 2,5% environ en 2006. Structurellement, ce sont les services télécoms et réseaux qui s'octroient la plus grosse part du gâteau (44,3% des dépenses en TIC), devant les services informatiques (20,5%), le matériel informatique (12,2%) et les logiciels (11,1%). Mais cette réalité masque une évolution en profondeur des dépenses. Ainsi, la plus forte croissance structurelle en 2006 concerne les dépenses en logiciels (+6,3%), devant les services informatiques (+5,3%). En Europe, la France se situe plutôt dans la moyenne basse des investissements en TIC avec une croissance de 2,7% du marché en 2006, contre 3,1% pour l'ensemble des pays européens. Elle est notamment en retrait par rapport à l'Espagne (+5,2%) et le Royaume-Uni (+3,3%), mais rattrape l'Allemagne (+1,6%) et devance l'Italie (+2,0%).

Selon une étude conduite par l'EITO, avec Gfk et IDC, le marché des technologies de l'information et de la communication se montrerait en 2006 plus dynamique que l'ensemble du tissu économique des 25 membres de l'Union européenne (+2,5%). En France, la croissance atteindrait 2,7% et, contrairement à la majorité des grands pays d'Europe, devrait perdurer en 2007.

#### 4.1.2.3. Fusions / acquisitions.

Dans le secteur des NTIC, les rapprochements d'entreprises qu'ils soient commerciaux, technologiques ou stratégiques augmentent considérablement entre 2003 et 2004. Selon une étude de Regent Associates, les fusions et acquisitions du secteur IT en Europe ont augmenté de 70% entre le 1<sup>er</sup> semestre 2003 et le 1<sup>er</sup> semestre 2004 :

Secteur IT (Services IT, Logiciel, Distribution IT, Matériel, Services télécoms, Equipements télécoms, Technologie appliquée métier, Contenu et média, Autres)	Nombre de fusions et d'acquisitions au 1 <sup>er</sup> semestre 2003	Nombre de fusions et d'acquisitions au 1 <sup>er</sup> semestre 2004
Total	655	1118

Source Regent Associates

Concernant le marché des technologies de la langue, force est de constater que cette tendance à la consolidation des acteurs se vérifie : après une accalmie des opérations de fusions et d'acquisition en 2002, la reprise des opérations de fusions a repris et notamment en 2005 avec le rachat de Verity par Autonomy, puis avec le rachat de Hummingbird par Opentext, et Filenet par IBM, pour ne citer que les plus importantes en terme d'acquisition financière. La turbulence du marché des outils linguistiques pressentie lors de l'étude précédente est toujours constatée.



#### 4.1.2.4. Les 6M & les usages numériques.

Plusieurs marqueurs de tendances (intitulés les « 6M », pour Multimédia, Multilingue, Multimodal, Multi-usages, Multi-utilisateurs et Multi-Médiation), basés sur les usages et les applications déployées en NTIC, influent directement sur les tendances et les évolutions des segments issus du marché du logiciel, dans les 9 segments d'applications des technologies de la langue. Le préfixe « Multi » rend compte de l'interaction existant entre la communication, l'informatique et l'information, et du mouvement dynamique illustré par la convergence des usages, l'interaction des techniques et l'interopérabilité des systèmes.

##### Les 6M

- Multimédia: contenu texte, images, sons
- Multilingue: traduction automatique et assistée par ordinateur
- Multimodal: terminaux fixes, mobiles, bornes et kiosques multimédia, systèmes embarqués
  - Multi-formats: gestion multi-formats et multi-usages
  - Multi-utilisateurs: travail collaboratif, interactivité, e-learning
  - Multi-médiation : langage naturel textuel et vocal

Dans l'entreprise, le traitement des données informatiques est aujourd'hui :

- **multimédia** : les contenus numériques intégrés dans le cycle de l'information de l'entreprise sont de plus en plus complexes et multi-composites : texte, image fixes, images animées sont d'autant de contenus d'information à traiter (recherche, collecte, classement, catégorisation, indexation, organisation, structuration, valorisation, diffusion...).
- **multilingue** : la globalisation des échanges, l'internationalisation des entreprises, et l'utilisation massive de l'Internet et des réseaux, favorisent la pratique du multilinguisme dans l'entreprise avec la prise en compte de particularismes locaux et socio-culturels dans la communication d'entreprise à un niveau individuel (traduction automatique et/ou assistée par ordinateur) et à un niveau collectif (traduction et localisation des sites Internet des entreprises avec l'intégration des données locales)
- **multimodal** : l'avancée des technologies matériels (terminaux informatiques mobiles, téléphones portables multimédia, assistants personnels nomades, bornes interactives, systèmes embarqués) permet le développement généralisé de solutions logicielles multimodales définies pour des utilisateurs en postes fixe ou nomade.
- **multi-formats** : les solutions informatiques intègrent de plus en plus la gestion multi-usage et multi-format des contenus numériques. L'utilisation du XML pour le développement des applications permet la diffusion de contenus sources paramétrables pour des usages différents (personnalisation des fichiers en output : édition et production automatique de documents textuels et multimédia définie en fonction de profils d'utilisateurs) et dans des formats différents (possibilité d'exporter des fichiers au format texte, document, PDF, images GIF, TIF...).

- **multi-utilisateurs** : avec le travail collaboratif et l'interactivité découlant des NTIC, les applications ont intégré le traitement multi-utilisateur des données informatiques pour répondre aux besoins des applications de type workflow (collaboration d'un groupe de travail) ou pour favoriser la communication et l'interactivité dans l'apprentissage (e-learning pour les langues étrangères, pour les opérations de formation dans l'entreprises).
- **Multi-médiation** : le développement des technologies en langage naturel et vocal permet aujourd'hui le contrôle à distance de périphériques ; la communication système via téléphone portable favorise la mise en relation de personnes parlant des langues différentes.

### Synthèse des facteurs conjoncturels

- La croissance de la numérisation, du stockage et la répartition des données numériques dans les réseaux: la production mondiale d'information (en volume) : les bases structurées ont augmenté de 4% en 2006 et représentent environ 10% des données. Les données non structurées représentent 90% du total et augmente de 6400% par an (IDC).
- La place du secteur des NTIC dans l'économie mondiale: marché NTIC = 2681 milliards d'euros en 2005. Europe = 33,6 % du marché mondial des NTIC.
- Le multilinguisme: Europe = 450 millions de personnes, 160 millions d'internautes, 55% parlant au moins une seconde langue.
- L'élargissement de l'utilisation d'Internet mobile à haut débit
- Les nouveaux segments et Internet: gestion de contenu multimédia, interfaces interactives et vocales, e-business, e-learning, Internet mobile, moteurs de recherche et web sémantique...

## 4.2. Le marché UE 13 pays : actualisation des données de l'étude conduite en 2004.

Pour rappel, lors de l'étude conduite en 2004, 363 sociétés privées ont été identifiées dans les 13 pays européens étudiés : l'Allemagne, l'Autriche, la Belgique, le Danemark, l'Espagne, la Finlande, la France, la Grèce, l'Irlande, l'Italie, les Pays-Bas, le Royaume-Uni, et la Suède. L'actualisation de cette étude porte sur le nombre d'acteurs présents sur le marché, leur répartition géographique et, l'évaluation du marché européen des technologies de la langue basée sur l'analyse des chiffres d'affaires communiqués en 2004.

Les acteurs du marché sont positionnés à 80% sur le traitement des données textuelles et à 20% sur le traitement des données vocales. Le marché est estimé à 508 millions d'euros en 2002 ; le Royaume-Uni, la France, l'Allemagne et l'Italie font 60% du marché.

### 4.2.1. Identification et répartition des acteurs.

En 2006, 394 acteurs présents dans l'Union européenne à 13 pays et positionnés sur au moins un des neuf segments d'application des technologies de la langue ont été identifiés, soit :

UE - 13 PAYS	Etude conduite en 2006		Etude conduite en 2004	
	Sociétés recensées	%	Sociétés recensées	%
France	109	28%	96	26%
Allemagne	57	14%	57	16%
Royaume Uni	54	14%	49	13%
Italie	33	8%	31	9%
Espagne	23	6%	21	6%
Belgique	19	5%	21	6%
Pays Bas	18	5%	16	4%
Finlande	17	4%	17	5%
Suède	17	4%	11	3%
Danemark	15	4%	15	4%
Grèce	15	4%	15	4%
Autriche	9	2%	9	2%
Irlande	8	2%	5	1%
<b>TOTAL</b>	<b>394</b>	<b>100%</b>	<b>363</b>	<b>100%</b>

Notes :

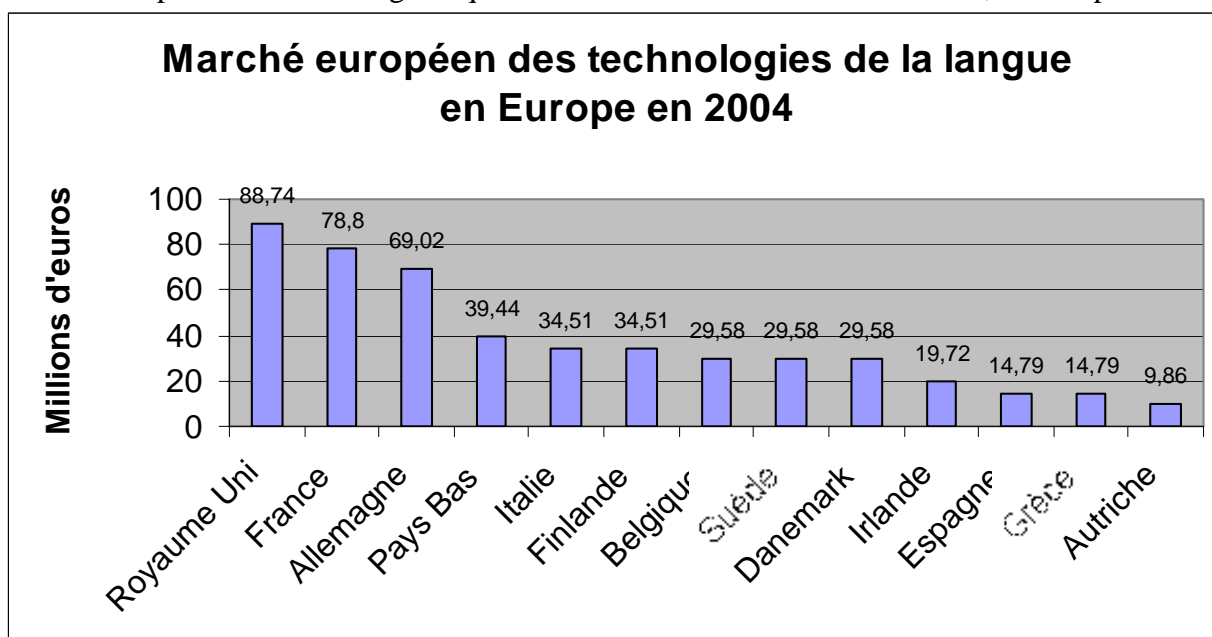
- Les données exploitées lors de l'étude 2004 proviennent des chiffres d'affaires officiels concernant l'année civile 2002. Les données exploitées lors de l'étude 2006 proviennent des chiffres d'affaires officiels concernant l'année civile 2004.
- Les chiffres d'affaires relatifs à l'année 2005 n'ont été communiqués que par un tiers des sociétés étudiées, expliquant le décalage de deux années entre l'année de réalisation de l'étude et l'année d'analyse des chiffres d'affaires des acteurs.

- Concernant le positionnement des acteurs par rapport au traitement des données numériques textuelles ou vocales, l'offre dans ces 13 pays d'Europe révèle en 2004 une progression du traitement des données vocales avec 30% des acteurs positionnés sur le vocal. En 2004, 70% des sociétés des sociétés sont positionnées sur le traitement des données textuelles exclusivement. En 2002, 80% des sociétés étaient positionnées sur le traitement des données textuelles exclusivement.

Pour l'analyse de l'évolution du secteur pour le marché européen, il a été procédé à un échantillonnage représentatif des acteurs qui ont communiqué leurs CA sur 5 années consécutives (2000 à 2004). La taille de l'échantillon pour l'Europe à 13 pays est de 92 sociétés.

#### 4.2.2. Evaluation du marché.

Selon nos estimations basées sur l'analyse des chiffres d'affaires des acteurs de l'offre, le marché européen des outils linguistiques est de 493 millions d'euros en 2004, ainsi répartis:



Les principaux pays leaders du secteur des technologies de la langue sont le Royaume-Uni, le France, l'Allemagne, les Pays-Bas, l'Italie et la Finlande. Ils représentent à eux seuls 70% du marché.

Les trois pays leaders du marché européen (Royaume-Uni, France et Allemagne) capitalisent près de la moitié du marché, avec une part de marché cumulée de 48%.

### 4.2.3. Comparaison des résultats entre les chiffres de 2002 et 2004.

Le tableau suivant présente les chiffres clés du marché européen à 13 pays, en 2002 et 2004 :

Europe - 13 pays	CA 2004	CA 2002
Nombre d'acteurs	394	363
Répartition Traitement texte / voix	70/30	80/20
Valeur dégagée (millions euros)	493	508

Le développement du marché des technologies de la langue se confirme sur la période 2002-2004 avec une hausse de 7.8% du nombre d'acteurs positionnés sur le marché.

La valeur dégagée par le secteur est passée de 508 millions d'euros en 2002 à 493 millions d'euros en 2004. Cependant, cette baisse observée n'est pas significative : elle doit être relativisée du fait que le nombre de sociétés qui communiquent leurs chiffres d'affaires est différent d'une année à l'autre (certaines sociétés ont disparu par rachats, fusions ou liquidation, et les nouveaux entrants n'ont pas forcément de chiffre d'affaires officialisé ou annuel la première année, d'autres n'ont tout simplement pas communiqué leurs chiffres).

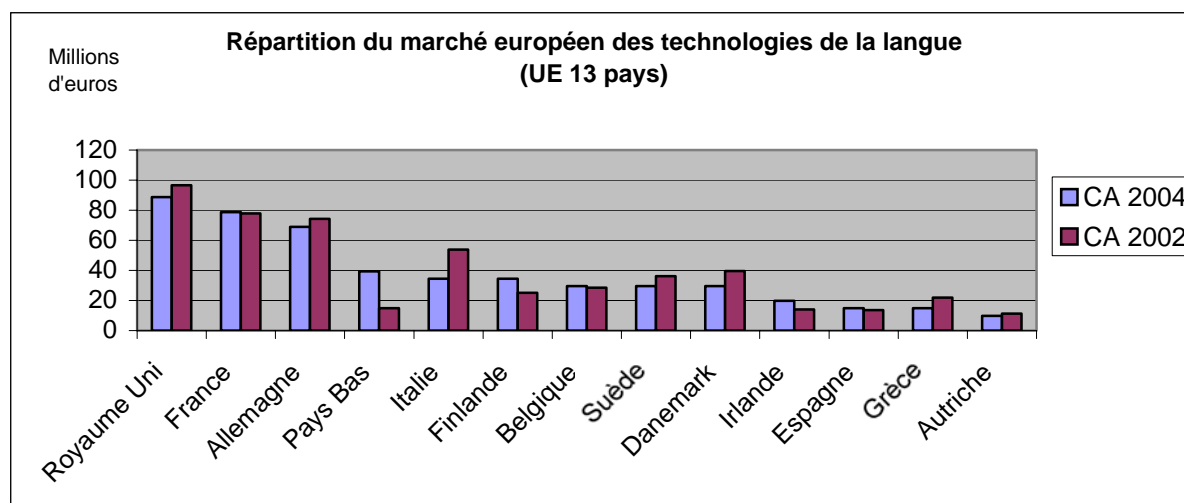
La répartition du marché européen des outils linguistiques a légèrement évolué entre 2002 et 2004, même si la plupart des pays leaders du marché conservent leurs positions : en 2004, Royaume-Uni, France, Allemagne, et Italie représentent 55% du marché européen. En 2002, ces 4 pays représentaient 59% du marché. Les pays leaders, tels que le Royaume-Uni, la France et l'Allemagne, ont appuyé leurs positions dominantes sur la période.

La part dégagée par l'Italie diminue, passant de 11% en 2002 à 7% en 2004 : cette tendance à la baisse peut s'expliquer par le fait que plusieurs sociétés n'ont pas déclaré leurs chiffres d'affaires 2004.

La part dégagée par les Pays-Bas augmente, passant de 3% en 2002 à 8% en 2004 : cette tendance à la hausse peut s'expliquer par le fait que la part du marché hollandais avait été estimée en 2002 (du fait de sources financières non communiquées). De même en 2004, un lissage de la valeur a été établi en prenant en compte :

- les chiffres d'affaires communiqués (qu'ils soient consolidés ou nationaux),
- le nombre d'acteurs en activité (sociétés avec une identité sociale, un site Internet ou un argumentaire commercial, et des chiffres d'affaires communiqués sur la période),
- la part relative estimée du nombre total d'acteurs hollandais référencés par les différents instituts (comme Elsnet) par rapport au total du nombre d'acteurs en Europe.

Comparaison de la répartition du marché européen des outils linguistiques entre l'année 2002 et l'année 2004 :



Les 3 pays leaders du marché en 2002 (Royaume-Uni, France, Allemagne) sont les mêmes en 2004. Ils représentent près de 50% du marché.

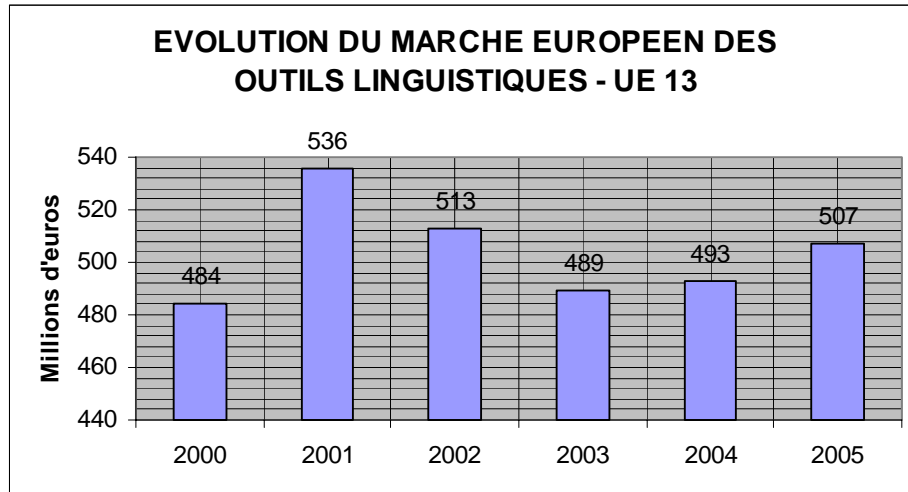
- Evolution du marché européen des technologies de la langue (Union européenne à 13 pays) de 2000 à 2005 :

Selon nos estimations, le marché européen des outils linguistiques (Europe 13 pays) reste stagnant sur la période 2000-2005 :

- en 2000: 484 millions d'euros
- en 2001: 536 millions d'euros
- en 2002: 513 millions d'euros (estimé à 508 millions lors de la première étude conduite en 2003)
- en 2003: 489 millions d'euros
- en 2004: 493 millions d'euros
- en 2005: 507 millions d'euros (prévisionnel)

L'évolution positive du marché européen des outils linguistiques en 2001 (zone Europe 13 pays) avec une hausse du marché de +10,71% en 2001 illustre l'ébullition du marché des NTIC sur la période 2000-2001. Les baisses constatées en 2002 (-4,21%) et en 2003 (-4,69%) confirment la tendance morose qu'a rencontré le secteur des NTIC, et plus particulièrement le marché du logiciel, après l'éclatement de la bulle Internet en 2001. L'année 2004 souligne le retour à la croissance pour le marché des technologies de la langue avec un taux de croissance positif de +0,76%. Ce marché demeure stable sur la période malgré les turbulences observées. L'année 2005 semble confirmer la hausse de 2004 avec un taux de croissance estimé à +2,81%.

Même si le marché a baissé sur la période, le marché a évolué positivement sur la période malgré les incidents économiques observés avec notamment l'éclatement de la bulle Internet 2001 et la tendance morose des années 2002 et 2003.



**Points faibles :**

- croissance turbulente des NTIC
- barrières socio-culturelles
- investissements importants pour le développement des technologies

**Points forts :**

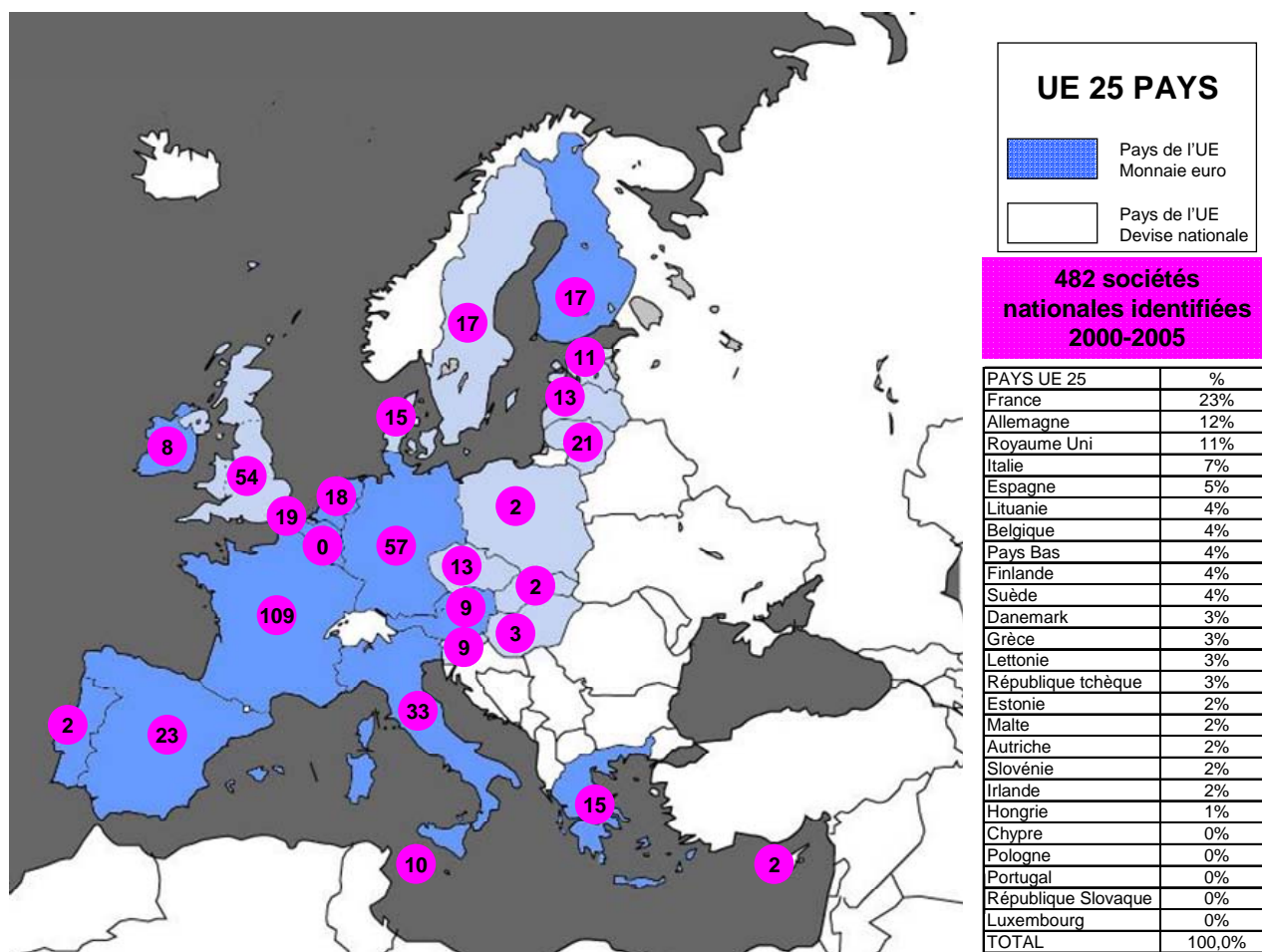
- convergence numérique de l'informatique, de l'information et de la communication
- poussée du traitement des données multimédia (vocales, images)
- maturité des usages

### 4.3. Marché européen des technologies de la langue (25 pays) : intégration des 12 autres pays de l'Union européenne

Les acteurs des 12 pays complémentaires (dont 10 nouveaux entrants) ont été intégrés après identification et qualification.

#### 4.3.1. Identification et répartition des acteurs

Il a été identifié 482 acteurs présents sur au moins un des neuf segments d'application des technologies de la langue, dans les 25 pays, dont 88 issus des 10 nouveaux pays entrants..



La répartition des acteurs européens par rapport au traitement des données numériques montre que 72% des acteurs sont positionnés exclusivement sur le traitement du texte et 28% sur le traitement des données vocales (et parfois textuelles). Notons que 12% des acteurs sont exclusivement positionnés sur le traitement des données vocales.

394 acteurs ont été identifiés dans l'Union européenne à 13 pays.

482 acteurs ont été identifiés dans l'Union européenne à 25 pays.



Plusieurs impacts existent sur le marché des technologies de la langue et participent à la poussée du vocal:

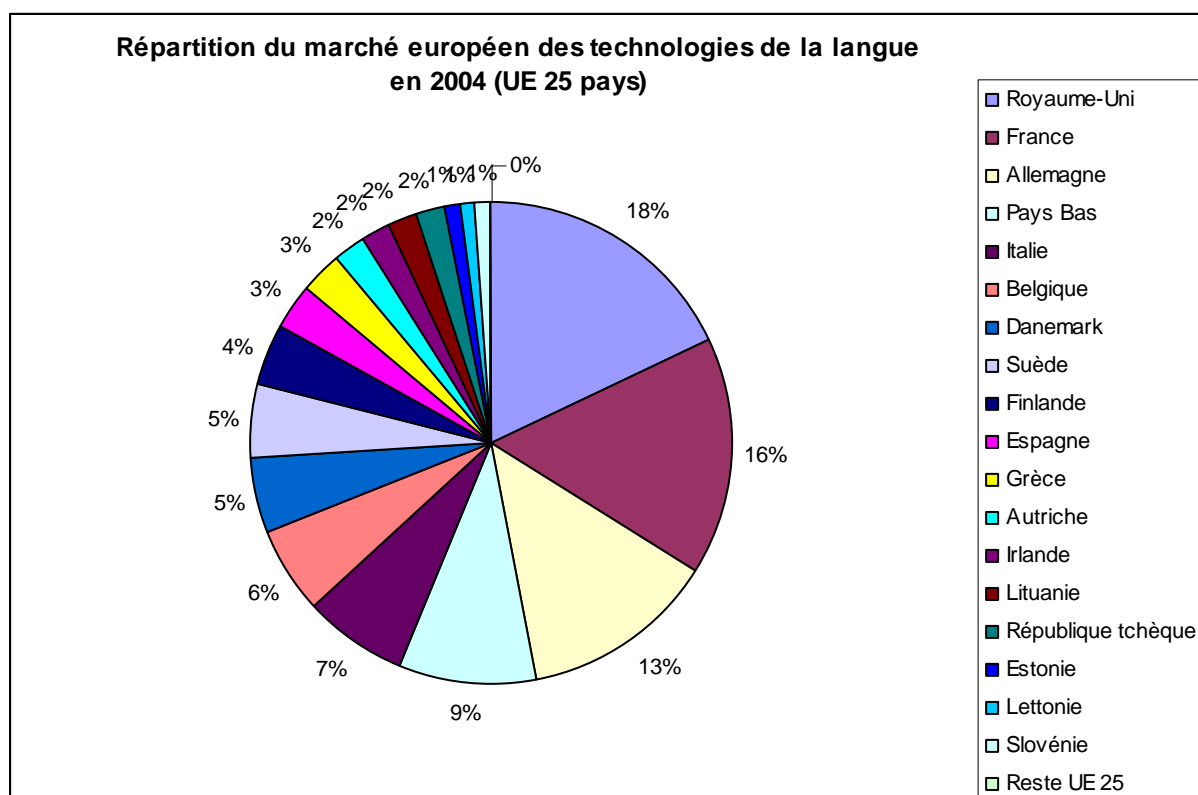
- L'impact technologique : interopérabilité des systèmes, émergence de nouvelles normes, voix sur réseaux (hausse des e-call centers)...
- L'impact économique : mondialisation des échanges...
- L'impact socio-culturel : multilinguisme, usages numériques...

#### 4.3.2. Evaluation du marché

Pour l'analyse de l'évolution du secteur pour le marché européen, il a été procédé à un échantillonnage représentatif des acteurs qui ont communiqué leurs CA sur 5 années consécutives (2000 à 2004). La taille de l'échantillon pour l'Europe à 25 pays est de 96 sociétés.

Sur les 482 sociétés identifiées comme offreurs de technologie linguistique, seules 45% ont communiquées leur chiffre d'affaires, soit 252 sociétés. L'évaluation du marché a été basée sur l'analyse de ces chiffres d'affaires communiqués.

Le chiffre d'affaires dégagé par le secteur des technologies de la langue en Europe (25 pays) est estimé à 610 millions d'euros en 2004, selon la répartition géographique suivante:



(Reste UE 25 : Chypre, Hongrie, Luxembourg, Malte, Pologne, Portugal, République slovaque).

Les pays leaders sont le Royaume-Uni, la France, l'Allemagne, et les Pays Bas : ces 4 pays leaders représentent 56% du marché européen des outils linguistiques en 2004.

- Evolution du marché européen des technologies de la langue (Union européenne à 25 pays) de 2000 à 2005 :

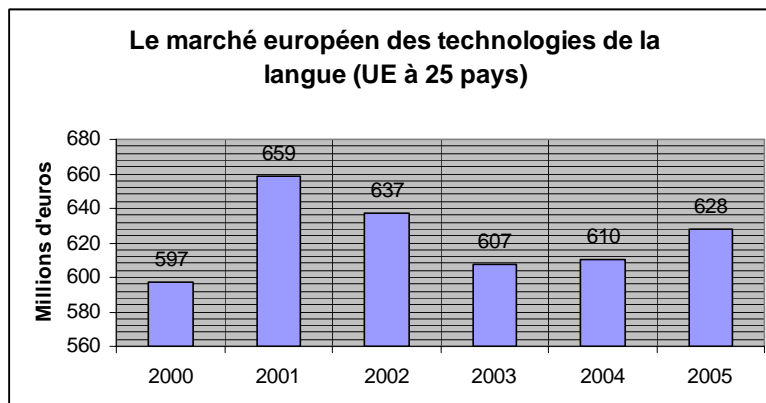
Le marché européen des technologies de la langue (UE à 25 pays) a représenté :

- en 2000: 596 millions d'euros
- en 2001: 658 millions d'euros
- en 2002: 637 millions d'euros
- en 2003: 607 millions d'euros
- en 2004: 610 millions d'euros
- en 2005: 628 millions d'euros (prévisionnel)

Remarques :

- Nous avons considéré que 80% du marché européen à 25 pays est assuré par l'Union européenne à 13 pays, ce qui a impliqué un coefficient correcteur de 15% sur la valeur globale du marché pour la prise en compte des acteurs non identifiés dans les 25 pays,

Années	UE 13	UE 25 – (échantillon 482 sociétés)	UE 25 (+15%)	% UE13 / UE25
2000	484	519	597	81,09 %
2001	536	573	659	81,34 %
2002	513	554	637	80,52 %
2003	489	528	607	80,53 %
2004	493	531	610	80,73 %
2005	507	546	628	80,74 %



Représentant ainsi moins de 20% du marché, les pays nouveaux entrants de l'Union européenne à 25 pays suivent l'évolution de l'Europe des 13 et la dynamique des pays leaders : l'évolution du marché européen des technologies de la langue à 25 pays suit donc la même tendance que celle qui a été constatée dans la zone de l'Union européenne à 13 pays.

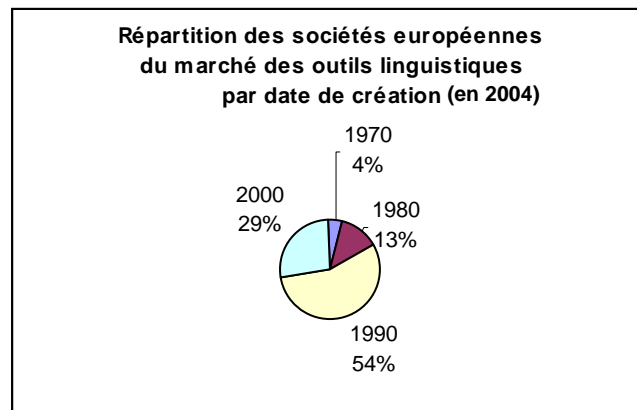
L'évolution du marché européen des outils linguistiques (zone Europe 25 pays) a été positive avec une hausse du marché de +10,37% en 2001, illustrant la bulle Internet en 2000-2001.

L'année 2004 souligne le retour à la croissance pour le marché des technologies de la langue avec un taux de croissance positif de +0,53%. L'année 2005 semble confirmer la hausse de 2004 avec un taux de croissance estimé à +2,72%. Ces taux de croissance restent inférieurs à ceux assurés par la zone de l'Union européenne à 13 pays.

#### 4.3.3. Présence des acteurs sur le marché européen.

En 20 ans, le nombre de sociétés créées a été multiplié par plus de 10, passant d'une trentaine à, 377 sociétés identifiées en Europe en 2003, à 394 sociétés identifiées en Europe (13 pays) en 2005. Sur la période 2000-2005, plus de 50 sociétés européennes spécialisées dans le traitement du langage ont disparu (liquidation ou rachat). Par ailleurs, les créations de sociétés ont considérablement augmenté dans les années 1990 et 2000: en 2004, sur 252 sociétés avec CA communiqués, 83% d'entre elles ont été créées après 1990.

La répartition illustrée dans le schéma ci-après montre la pérennité naissante des sociétés du marché, avec des existences de plus de 10 ans pour la plupart (année moyenne de création des 482 sociétés identifiées : 1994)



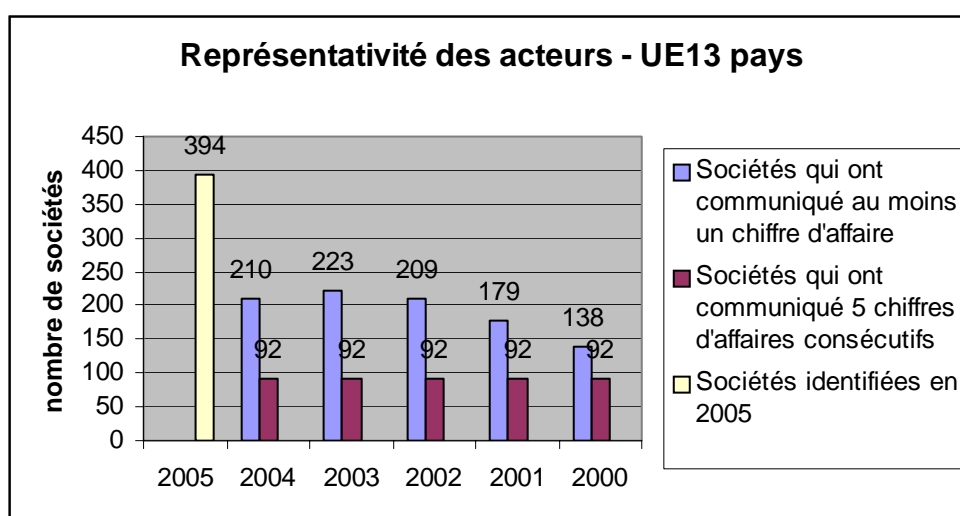
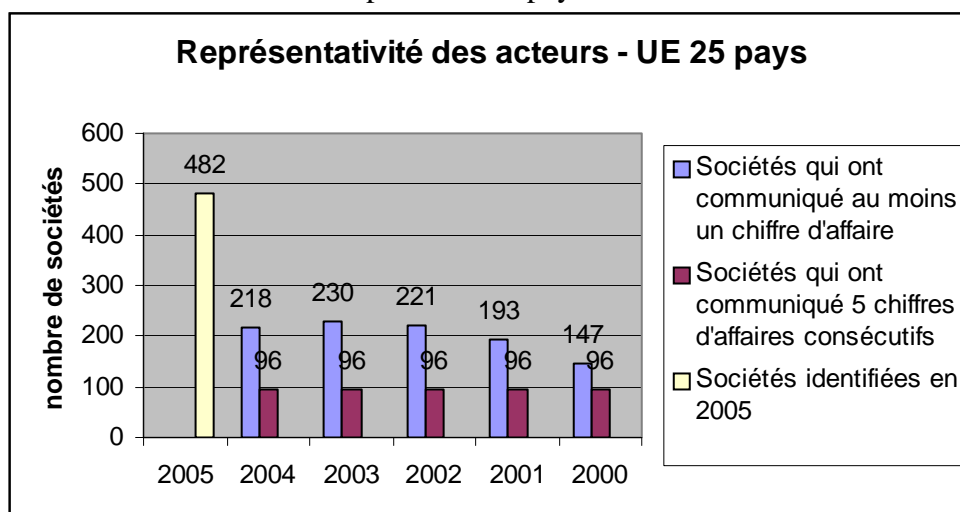
#### 4.3.4. Valeur dégagée par le secteur des technologies linguistiques en Europe.

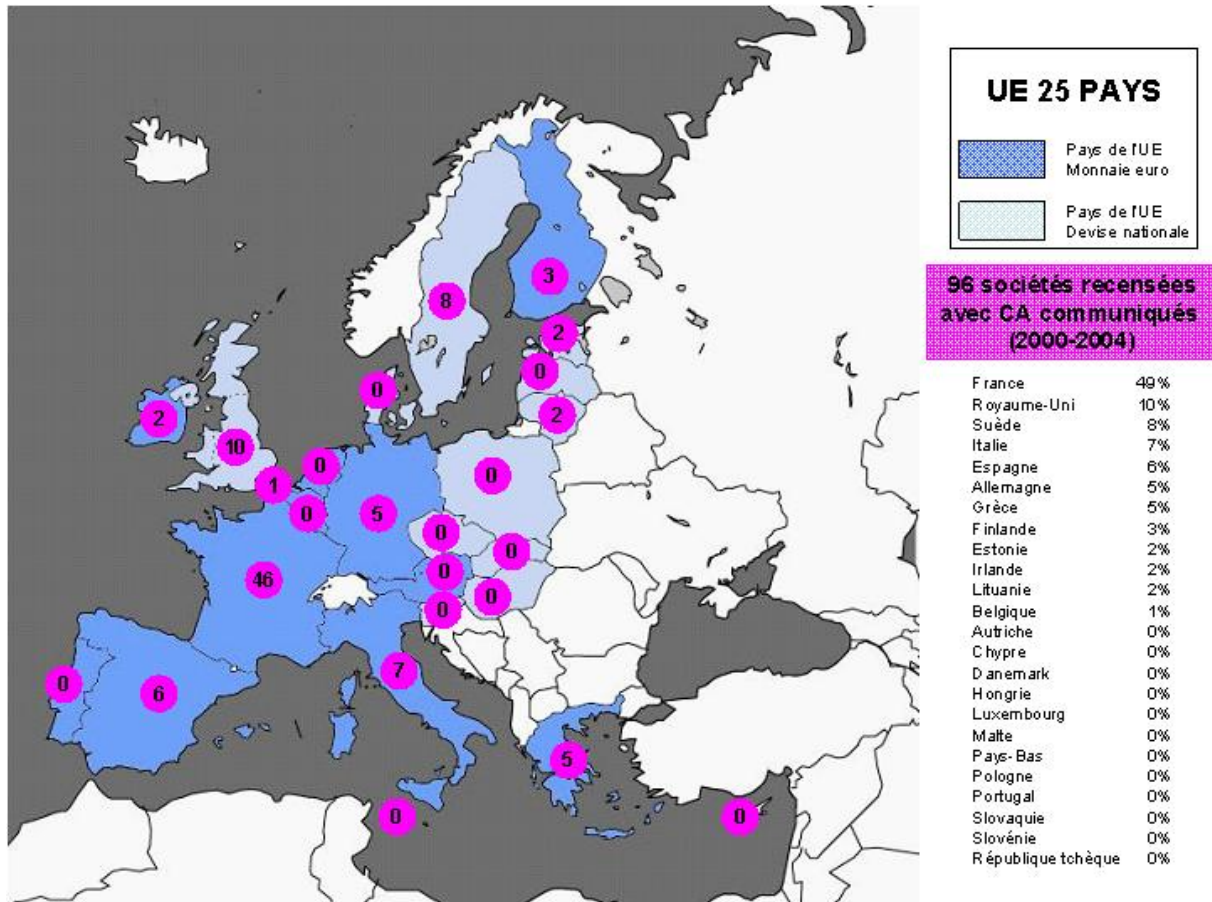
La valeur dégagée par le secteur des technologies de la langue en Europe a été estimée à partir d'un échantillon de sociétés européennes représentatives du secteur et ayant communiqué leurs chiffres d'affaires sur 5 années consécutives. Il s'agit avant tout de pouvoir dresser une photographie du marché à un instant  $t$ , comparable en  $t+1$ ,  $t+2$ ,  $t+3$ , et  $t+4$ .

- Constitution de l'échantillon

Compte tenu de la forte intensité concurrentielle sur le secteur, un échantillon représentatif du secteur a été constitué pour permettre une analyse fine et cohérente du marché sur la période.

Les données économiques et financières ont été recherchées sur les 482 sociétés identifiées : moins de la moitié des sociétés a communiqué au moins un chiffre d'affaire sur la période 2000-2004. Parmi ces acteurs, seules 92 sociétés présentes dans la zone de l'Union européenne à 13 pays ont communiqué 5 chiffres d'affaires consécutifs, et seulement 96 sociétés dans la zone de l'Union européenne à 25 pays.



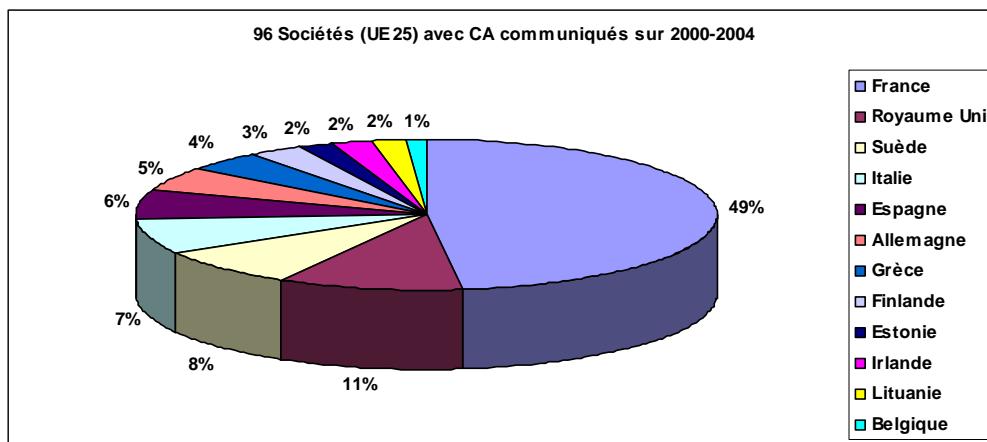


- Représentativité de l'échantillon

L'échantillon représentatif constitué de 96 sociétés européennes en provenance de 12 pays de l'Union européenne (France, Royaume-Uni, Suède, Italie, Espagne, Allemagne, Grèce, Finlande, Estonie, Lituanie, Belgique) présente les caractéristiques suivantes :

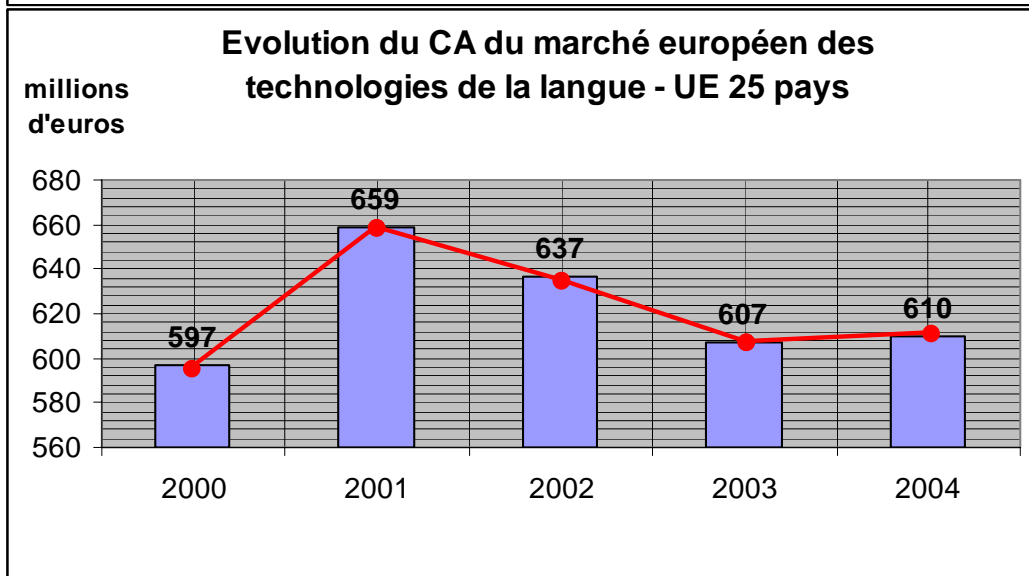
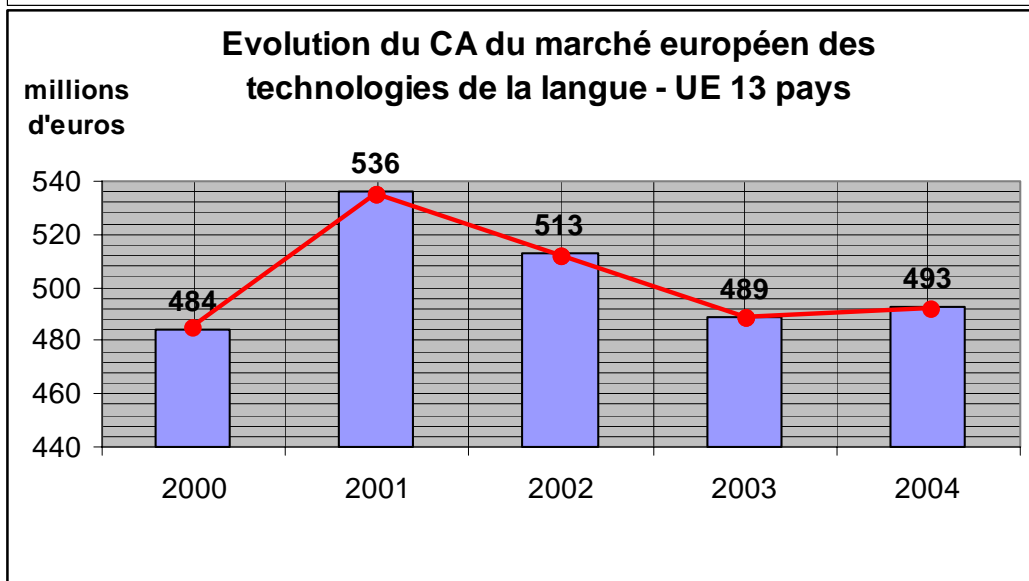
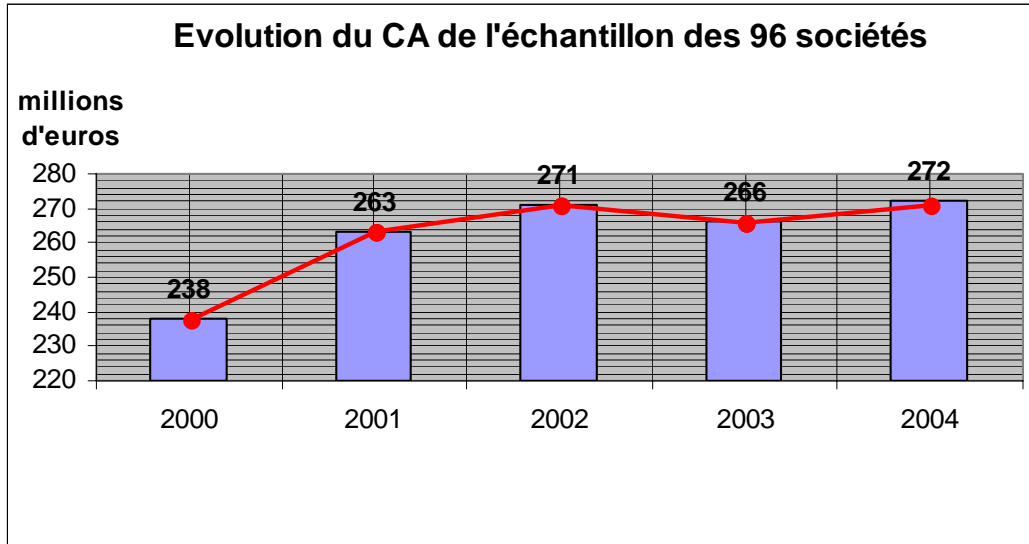
- **L'année moyenne de création** de l'échantillon est 1994. L'année moyenne de création de l'ensemble des acteurs identifiés en Europe à 25 pays est 1995 (les années moyennes de création des différents acteurs de l'échantillon se répartissent ainsi : Allemagne : 1991, Belgique : 1988, Espagne : 1991, Estonie : 1998, Finlande : 1995, France : 1992, Grèce : 1995, Irlande : 1996, Italie : 1985, Lituanie : 1994, Royaume-Uni : 1992, Suède : 1993).
- **18 nouvelles sociétés** de l'échantillon ont pénétré le marché entre 1998 et 2000 (et ont communiqué des chiffres d'affaires consécutifs sur la période) : plus précisément, deux sociétés ont été créées en 2000, huit sociétés en 1999 et huit en 1998.

- La **répartition texte / voix** est homogène à la situation dans la zone de l'Union européenne à 25 pays : la répartition des 96 acteurs en fonction du traitement textuel ou vocal des données numériques montre que 69 % sont positionnés exclusivement sur le texte, et 13 % sont positionnés sur la voix et 18% sur la voix et le texte.
- La répartition géographique des acteurs de l'échantillon souligne, comme lors de la précédente étude, la **sur-représentativité de la France** avec près de la moitié des sociétés de l'échantillon :



Les 96 sociétés européennes de l'échantillon sont principalement originaires des pays leaders du marché des outils linguistiques en Europe : Royaume-Uni, France, Allemagne et Italie représentent 69% des acteurs de l'échantillon. La France représente 47% de l'échantillon.

Il faut noter que l'évolution de la valeur dégagée par l'échantillon suit globalement l'évolution du marché européen des outils linguistiques.



L'analyse de la **valeur dégagée par l'échantillon suit globalement l'évolution de la valeur dégagée par le secteur** dans l'Union européenne à 25 pays sur la période 2000-2004 (hausse en 2001, baisse en 2003, hausse en 2004). Seule l'année 2002 diverge avec un léger taux de croissance du marché contre une baisse pour l'ensemble du marché.

Remarque : L'échantillon a connu en 2002 un taux de croissance de 2,80% au lieu d'une baisse attendue d'environ 3,3%. Cet effet positif pour 2002 est lié au fait que le marché français a été positif en 2002 et qu'il représente près de 50% de l'échantillon.

#### 4.4. Marché français des technologies de la langue: actualisation des données de l'étude conduite en 2004.

##### 4.4.1. Présence des sociétés en 2005 et répartition par segment.

En France, en 2005, il existe 109 sociétés identifiées dans le périmètre de l'ingénierie linguistique et des technologies de la langue. En 2003, 96 acteurs avaient été identifiés. Plusieurs des nouvelles sociétés sont issues d'anciens laboratoires de recherche issus de grands comptes privés et surtout publics (ex: Telisma et France Télécom, New Phenix et le CEA, ...). Plusieurs sociétés innovantes sont apparues, créées à partir de fonds exclusivement privés et/ou appuyées et financées par des octrois de subventions issues de programmes publics pour le développement de l'innovation et des NTIC.

Parmi ces 109 acteurs présents sur le marché, 105 sont présents en France et ont communiqué au moins un chiffre d'affaires sur la période 2000-2004. La liste des 109 acteurs présents en France est fournie en annexe.

La répartition des acteurs par segments d'application est illustrée dans le tableau ci-dessous :

CATEGORIES D'APPLICATIONS (2004)	Acteurs mono - catégories		Acteurs multi - catégories		Répartition totale des acteurs par catégorie		Répartition totale des acteurs par catégorie (rappel 2002)	
	Nombre	%	Nombre	%	Nombre	%	Nombre	%
1. Gestion du contenu	15	38%	42	23%	57	26%	48	26%
2. Gestion des documents	4	10%	30	17%	34	16%	29	16%
3. Traduction	3	8%	10	6%	13	6%	19	10%
4. E-business	1	3%	27	15%	28	13%	20	11%
5. Ressources linguistiques texte et voix	2	5%	12	7%	14	6%	14	8%
6. Production et édition	5	13%	11	6%	16	7%	11	6%
7. Interface vocale et multimodale	4	10%	13	7%	17	7%	14	8%
8. E-learning	3	8%	10	6%	13	6%	8	4%
9. Moteurs de recherche intelligents	2	5%	26	14%	28	13%	23	12%
Répartition des acteurs par activité mono ou multi catégories	39	36%	70	64%	109	100%	96	100%



39 sociétés parmi les 109, soit 36 % des sociétés françaises du marché des outils linguistiques, sont des acteurs spécialisés sur une seule catégorie 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%).

Comme en 2002, nous retrouvons la présence plus importante des sociétés positionnées dans les secteurs de la gestion de contenu (26 % des acteurs) et la gestion des documents (16% des acteurs). Il y a de légères différences pour les autres segments. Le tableau détaillé de la situation 2002 figure en annexe.

En 2005, l'analyse des années de création des sociétés françaises présentes sur le marché des technologies de la langue permet d'illustrer la jeunesse du marché sur le territoire national, même si la France compte plusieurs sociétés pionnières présentes en France depuis plusieurs décennies (comme par exemple Microsoft France en 1975, et Cimos en 1977). Plus de la moitié des acteurs du marché français ont été créés durant les années 1990 (57%). L'année moyenne de création des 109 acteurs est l'année 1995.

Près d'un quart des acteurs présents sur le marché français ont été créés dans les années 2000 et ont donc moins de 5 ans d'existence.

Parmi les nouveaux acteurs entrant sur le marché entre 2002 et 2005, nous pouvons notamment citer :

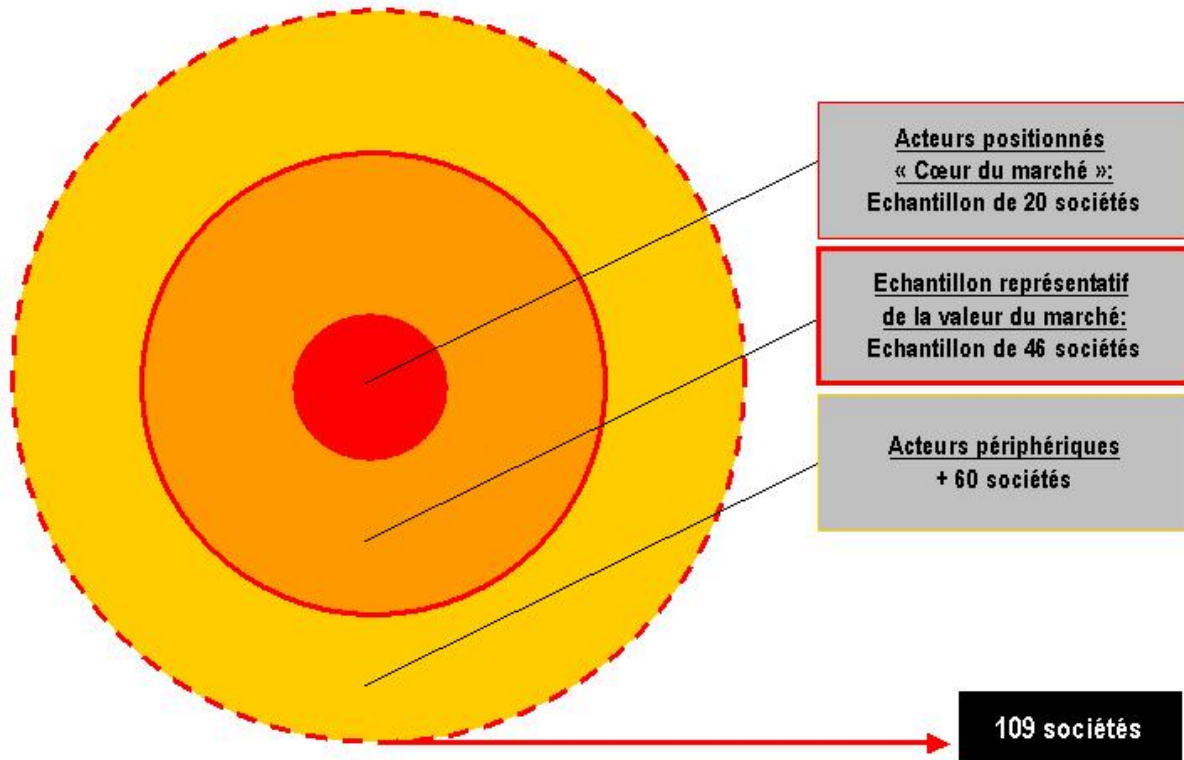
- en France : Evodia (2005), NewPhenix (2004), Polyspot (2004), AM2 Systems (2003), Linguanet France (2003), Ontologos Corp (2003), Arisem (2002), Autonomy France (2002), Babeling (2002), Electrel (2002) , Intellixir (2002), Ixxo (2002), Pertinence Mining (2002), Sinequa (2002), Tagmatica (2002) et Virtuoz (2002).
- En Europe : SPSS en Finlande (2003), Voice-In en Grèce (2003), Optiva en Lituanie (2004), Telephonetics au Royaume-Uni (2004), CIT en République tchèque (2001).

Parmi les acteurs disparus (absorbés, liquidés ou fusionnés, entre fin 2002 et 2005), nous pouvons citer entre autres :

- Rachats et fusions : Filenet France racheté par IBM, Hummingbird (Fulcrum) racheté par Open Text, Verity par Autonomy, Intraspect par Vignette, Némesia par le Groupe Nuages Blancs, Almas Ingénierie par ITESOFT, Mediapps par EverTeam, Trados / SDL, Elan / Speech / Acapela.
- Liquidations : Hyperwave en 2003, Crealog en 2003, Acktil en 2004

Par ailleurs, les 109 acteurs identifiés sont situés à différents niveaux sur le marché des outils linguistiques : certaines sociétés sont historiquement positionnées dans le cœur du marché de l'ingénierie linguistique, d'autres sont placées en périphérie du marché des technologies de la langue. Le schéma ci-après résume la distribution des acteurs en France en 2005 :

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



### 4.4.2. Evaluation du marché

Parmi ces 109 acteurs, près de 30 n'ont pas communiqué de chiffres d'affaires. Des acteurs tels que Microsoft ou Cap Gemini Innovation ont été cités et positionnés en tant qu'acteurs du secteur en France mais n'ont pas été intégrés en terme d'analyse des chiffres d'affaires : Développant de multiples activités annexes aux technologies de la langue, l'activité ingénierie linguistique ne peut être identifiée dans leurs chiffres d'affaires et la valeur n'a pas pu être extraite de leur CA global.

En 2004, le chiffre d'affaires dégagé par l'ensemble des acteurs sur le marché français est estimé à 78 millions d'euros, soit environ 16% du marché européen (Union européenne à 25 pays) et place la France au deuxième rang des pays leader après le Royaume-Uni.

Le marché français reste concentré : les 25 premières sociétés du marché français des outils linguistiques représenteraient près de 80% du CA global du secteur (chiffre à minimiser au regard des chiffres d'affaires très importants des 2 premiers acteurs : 39,6 millions d'euros pour Proteor et 23,3 millions pour Cril Technology). Le marché est donc concentré par la position dominante de quelques leaders et la fragmentation des acteurs les plus petits ou nouvellement arrivés.

Répartition du CA français du marché des outils linguistiques par segment de marché :

<b>Segments d'application</b>	<b>2004</b>	<b>2002</b>
<b>1. Gestion du contenu :</b>	<b>17%</b>	<b>21%</b>
<b>2. Gestion des documents :</b>	<b>22%</b>	<b>27%</b>
<b>3. Traduction:</b>	<b>5%</b>	<b>13%</b>
<b>4. E-business:</b>	<b>10%</b>	<b>10%</b>
<b>5. Ressources linguistiques texte et voix :</b>	<b>3%</b>	<b>3%</b>
<b>6. Production et édition :</b>	<b>3%</b>	<b>2%</b>
<b>7. Interface vocale et multimodale :</b>	<b>17%</b>	<b>11%</b>
<b>8. Applications pédagogiques et e-learning :</b>	<b>13%</b>	<b>5%</b>
<b>9. Moteurs de recherche intelligents :</b>	<b>10%</b>	<b>8%</b>
<b>TOTAL</b>	<b>100%</b>	<b>100%</b>

**- Gestion des documents (segment 2) :**

Les revenus dégagés les plus importants concernent le segment de la GED : la mise en place des solutions sollicitent des temps plus longs que pour les systèmes de gestion de contenu ou de e-business. Les budgets alloués sont donc plus importants que dans d'autres segments du marché et les mises en place sont souvent planifiées sur plusieurs mois (12-24 mois en moyenne).

**- Gestion du contenu (segment 1) :**

La gestion du contenu est en France la deuxième source de revenus du marché des outils linguistiques avec la présence de multiples sociétés sur le segment. Plusieurs sociétés de taille moyenne voire modeste sont présentes sur ce segment pour deux raisons principales :

- Les modules d'aide à la décision et d'analyse de texte sont souvent assez rapide à intégrer dans les systèmes ou se vendent même clé en main.
- Les PME du secteur ont donc la possibilité d'offrir leurs gammes de produits linguistiques facilement adaptables aux besoins des entreprises de petites mais aussi de grandes tailles.

**- Interfaces vocales et multimodales et e-learning (segments 7 et 8) :**

Le segment du traitement vocal est en pleine croissance et augmente sa part de marché sur la période mais le marché français reste légèrement en retard par rapport à ses voisins européens, du fait des pratiques et des usages de la demande française traditionnellement encrée sur le traitement du texte. Les leaders européens du traitement de la voix sont présents en France mais il existe un réel potentiel de développement du segment des applications vocales et multimodales dans la mesure où l'intensité concurrentielle semble moins prononcée qu'au Royaume-Uni ou en Allemagne. Ce segment semble de plus en plus lié au segment du e-learning qui présente des solutions d'apprentissage des langues ou d'aide au handicap avec des interfaces vocales et multimédia, d'où la tendance parallèle à la hausse suivie par le segment 8 passant de +5% à +13%.

**- E-business et moteurs de recherche intelligents (segments 4 et 9) :**

Les segments du e-business et des moteurs de recherche intelligents restent fortement corrélés à l'évolution de l'économie Internet en France : leurs marges de progression sont donc liées au déploiement du haut-débit, à la multiplication et à la généralisation des usages et des utilisateurs, à la mise en place de services web, et au paiement en ligne...

**- Traduction automatique (segment 3) :**

La valeur dégagée par le segment de la traduction automatique semble baisser sur la période 2002-2004 mais cette tendance doit être modulée au regard des éléments suivants :

- certains acteurs identifiés en 2002 n'ont pas été réintroduits dans le panorama des acteurs français des technologies de la langue en 2004 : il s'agit entre autres de la société d'ingénierie linguistique Acktil aujourd'hui disparue (qui commercialisait le logiciel Trados, aujourd'hui repris par SDL International), de la société Lucid'IT en dépôt de bilan, des sociétés Alshare et Julia Emily Software et Lucid'IT apparemment absentes du marché français (même si rérérencées par l'EAMT)
- plusieurs acteurs du segment n'ont pas communiqué leurs chiffres d'affaires
- des acteurs positionnés sur d'autres segments (et qui n'ont pas été catégorisés dans le segment 3) n'offrent pas à proprement parlé des « techniquement » des solutions de traduction automatique mais des aides au choix des termes métiers et des dictionnaires multilingues, ce qui a pu minimiser l'effet valeur de ce segment

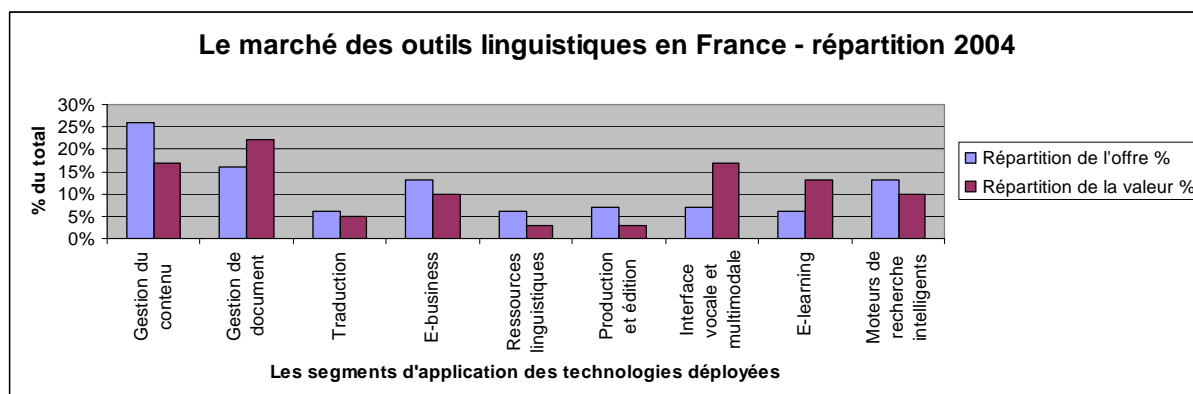
**- Ressources linguistiques et Production/Édition (segments 5 et 6) :**

Les deux autres segments du marché des outils linguistiques (ressources linguistiques, et production et édition) sont des domaines d'applications qui s'apparentent encore aujourd'hui à des niches de marché. Cependant il faut aussi noter :

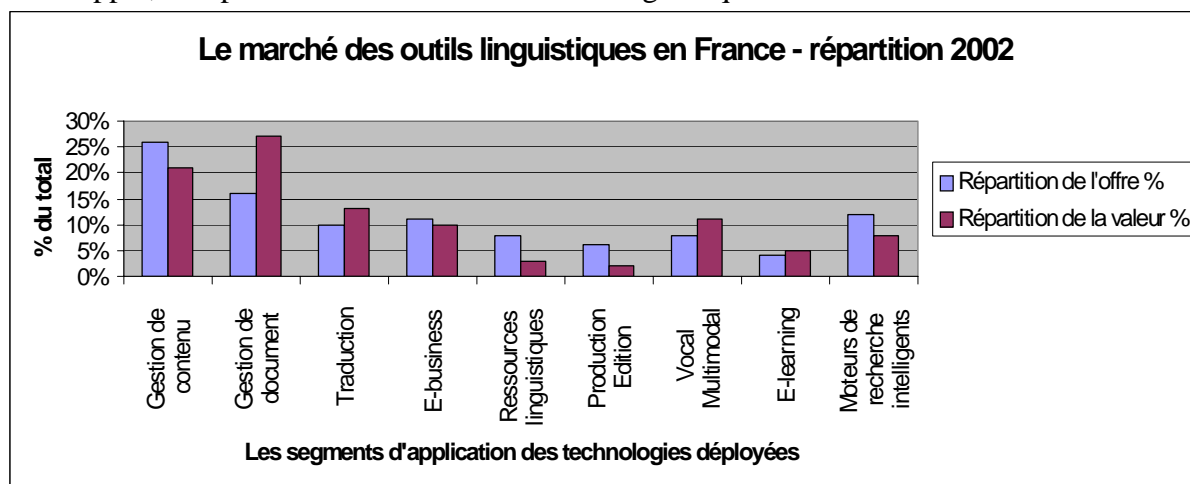
- qu'une multitude de solutions et de produits logiciels n'appartenant pas à ce segment proposent aussi des fonctionnalités d'édition ou de production intégrées au système complexe de Gestion de Contenu ou de GED par exemple,
- que les ressources linguistiques associées aux solutions ou systèmes sont souvent proposées en accompagnement ou en services complémentaires à la mise en place par les éditeurs sans communiquer pour autant sur un positionnement en tant que fournisseur de ressources linguistiques.

#### 4.4.3. Comparaison par rapport à 2002.

Répartition du marché des outils linguistiques 2004 :



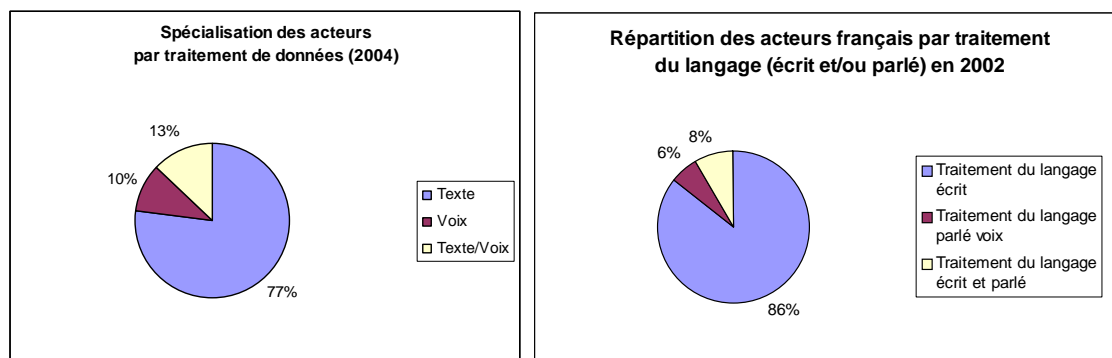
Pour rappel, la répartition du marché des outils linguistiques 2002 était :



En ce qui concerne le nombre d'acteurs, en 2003 86% des sociétés leaders du marché français sont positionnées exclusivement sur le traitement des données textuelles contre 77% en 2005.

Même si le texte reste encore aujourd'hui le moteur de la valeur du marché des outils linguistiques en France, les revenus générés par la voix sont en expansion importante et continue.

En se basant sur les leaders du marché en France, la prédominance du traitement du texte sur le marché telle qu'observée en France en 2002 (86% des sociétés leaders du marché français étaient positionnées exclusivement sur le traitement des données textuelles) semble se réduire : en effet, 77% des sociétés leaders du marché français sont positionnées exclusivement sur le texte en 2004.



#### 4.4.4. Constitution d'un échantillon de 46 acteurs représentatifs du marché français et répartis par segments d'application.

Une analyse comparative par segment sur 5 ans a été conduite à partir de la constitution d'un échantillon représentatif de 46 sociétés leaders du marché français, ayant communiqué leurs chiffres d'affaires sur une période consécutive de 5 ans (de 2000 à 2005). L'évolution sur 5 ans pour chaque segment a été estimée en fonction de la valeur dégagée du chiffre d'affaires et du résultat.

Pour l'analyse de l'évolution du marché français, la taille de l'échantillon est de 46 sociétés. La liste des 46 acteurs de l'échantillon figure en annexe.

Concernant les 46 acteurs présents sur le marché français en 2004, la répartition des acteurs par segments d'application est illustrée dans le tableau ci-dessous :

CATEGORIES D'APPLICATIONS	Acteurs mono - catégories		Acteurs multi - catégories		Total des acteurs	Répartition des acteurs par catégorie
		%		%		
1. Gestion du contenu	5	33%	18	21%	23	23%
2. Gestion des documents	4	27%	13	15%	17	17%
3. Traduction	2	13%	7	8%	9	9%
4. E-business	0	0%	12	14%	12	12%
5. Ressources linguistiques texte et voix	0	0%	6	7%	6	6%
6. Production et édition	1	7%	3	3%	4	4%
7. Interface vocale et multimodale	2	13%	8	9%	10	10%
8. E-learning	1	7%	6	7%	7	7%
9. Moteurs de recherche intelligents	0	0%	13	15%	13	13%
Répartition des acteurs par activité mono ou multi catégories	15	39%	31	61%	46	

Les caractéristiques suivantes confirment la représentativité de l'échantillon.

- 36% pour l'échantillon des 46 sociétés (contre 39 % des 109 sociétés françaises du marché des outils linguistiques) sont des acteurs spécialisés sur un seul segment d'application.
- 76% des sociétés sont positionnées sur le traitement du texte, 15% sur le traitement des données textuelles et vocales et, 9% sur le traitement de la voix exclusivement.
- La date moyenne de création est 1994 pour l'échantillon contre 1995 pour les 109 sociétés françaises.
- En 2004, les sociétés les plus présentes sur le marché des outils linguistiques sont celles traitant la gestion du contenu (23% contre 22% en 2002), la gestion des documents (17% contre 19%), les moteurs de recherche intelligents (13 % contre 10%), puis les spécialistes du e-business (12% contre 10%), les sociétés spécialisées dans les interfaces vocales et multimodales (10%), et les sociétés de traduction par ordinateur (9% contre 10%).
- En 2004, en terme de potentiel de CA, les sociétés de GED (22%), d'applications modales (21%), d'elearning (17%) et de gestion de contenu (14% contre 12%), représentent près de 75 % de la valeur dégagée par l'échantillon.

**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 marché est donc concentré du fait d'une part de la position dominante de quelques leaders et d'autre part de la dispersion des acteurs les plus petits ou nouvellement arrivés. 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.

**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. Le CA moyen de l'échantillon des 46 sociétés est de 4 millions d'euros en 2004.

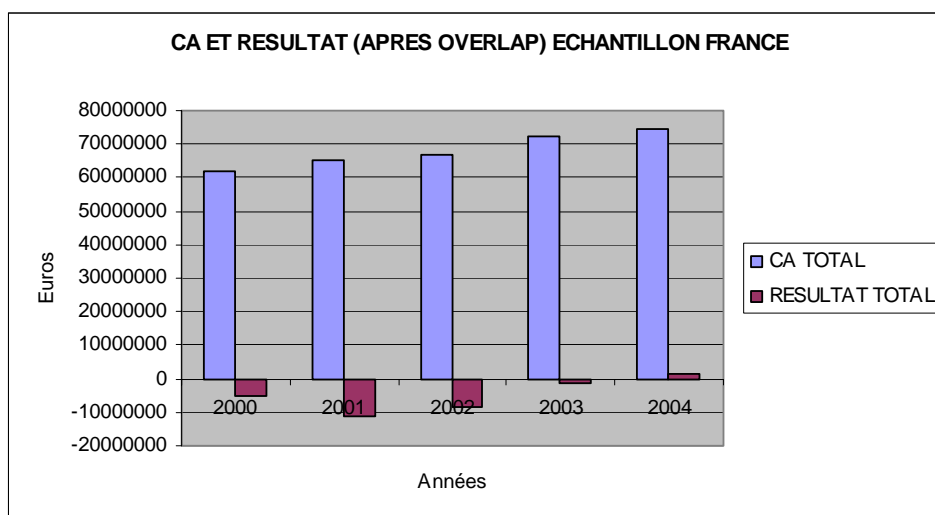
### Evolution globale du chiffre d'affaires et du résultat de l'échantillon :

Les chiffres d'affaires et les résultats des 46 sociétés de l'échantillon progressent positivement sur la période 2000 à 2004 :

Echantillon 46 sociétés	2000	2001	2002	2003	2004
Chiffre d'affaires (millions euros)	61,9	65,2	66,7	72,3	74,2
Taux de croissance du CA		5,32%	2,32%	8,43%	2,63%
Résultat (millions euros)	-3,5	-7,5	-5,7	-0,7	0,8
Rentabilité économique	-5,80%	-11,53%	-8,58%	-1,07%	1,16%

La valeur dégagée par les 46 sociétés augmente de 2000 à 2004, passant de 61,9 millions d'euros à 74,2 millions d'euros, avec des taux de croissance positifs et consécutifs sur la période.

Le marché français des technologies de la langue a moins subi la crise de 2001 que le marché européen : le taux de croissance moyen sur la période a été de 4.68%.

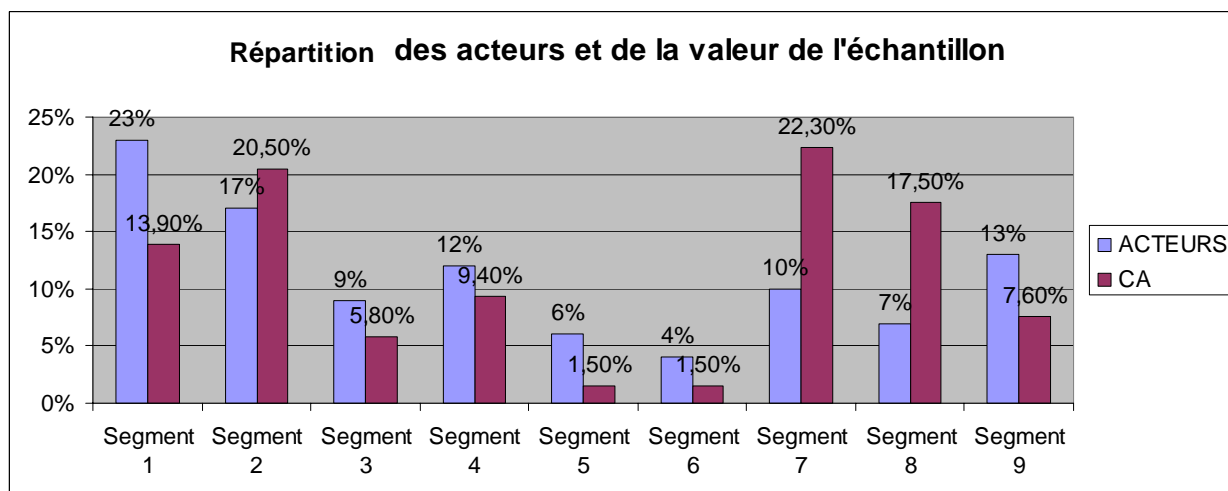


Après des années difficiles (2001 et 2002 surtout), la rentabilité économique du secteur est revenue à l'équilibre et semble même progresser positivement en 2004.

En effet, le marché français a du affronter la réalité difficile d'une rentabilité économique rarement présente (la rentabilité économique  $Re$  est le ratio  $\text{Résultat} / \text{CA}$ ), même si celle-ci est devenue positive pour la première fois en 2004 avec une  $Re$  de 1.16%.

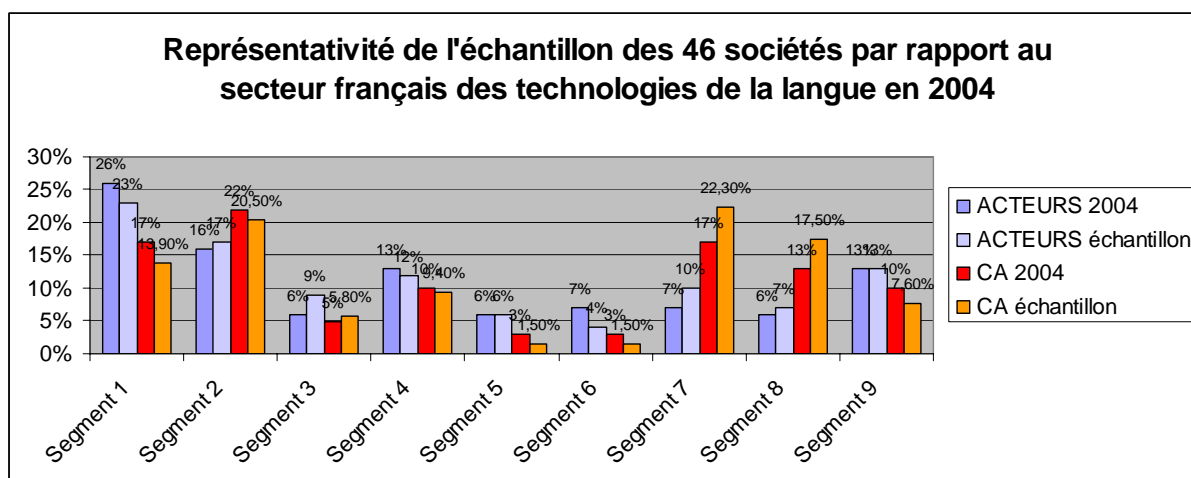
L'attractivité du marché est à distinguer par segments d'application : les segments Interfaces vocales et multimodales ainsi que le e-learning sont les deux segments où un faible nombre d'acteurs dégage une valeur marché importante, que ce soit en termes absolus ou relatifs par rapport aux autres segments.



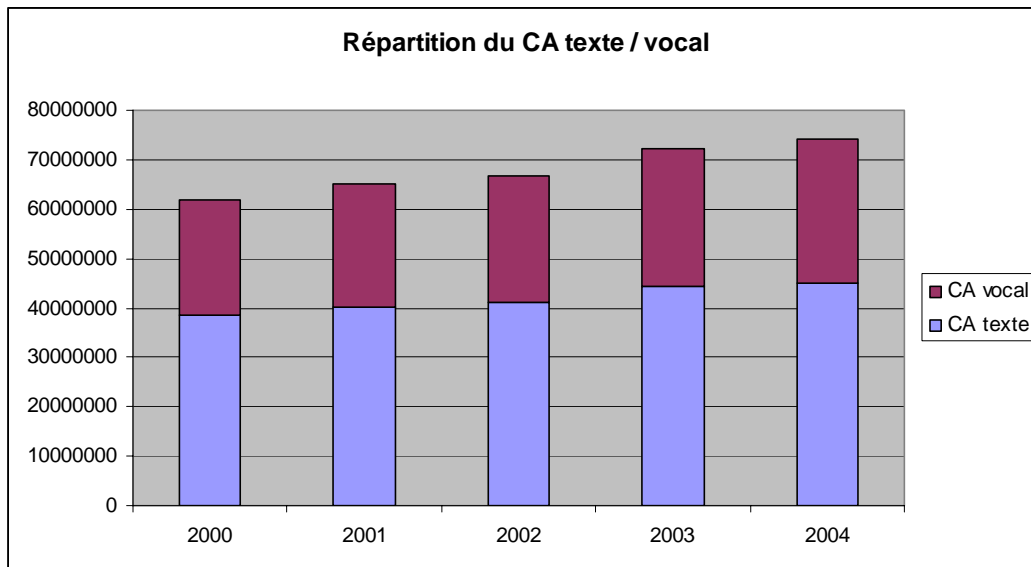


Indicateurs	Année	SG1 / TOTAL	SG2 / TOTAL	SG3 / TOTAL	SG4 / TOTAL	SG5 / TOTAL	SG6 / TOTAL	SG7 / TOTAL	SG8 / TOTAL	SG9 / TOTAL	TOTAL	Taux de croissance
CA	2000	13,5%	25,2%	6,1%	9,6%	1,4%	1,0%	22,4%	17,5%	3,4%	61,9	
	2001	12,8%	24,4%	4,4%	10,7%	1,5%	1,1%	22,1%	17,3%	5,9%	65,2	5,32%
	2002	12,8%	22,2%	5,0%	10,1%	1,6%	1,3%	22,9%	17,9%	6,2%	66,7	2,32%
	2003	12,8%	21,9%	5,7%	9,9%	1,4%	1,5%	22,4%	17,7%	6,6%	72,4	8,43%
	2004	13,9%	20,5%	5,8%	9,4%	1,5%	1,5%	22,3%	17,5%	7,6%	74,3	2,63%
Résultat	2000	35,2%	37,3%	-5,9%	37,3%	1,7%	-1,2%	-4,0%	-4,1%	3,8%	-3,6	
	2001	37,4%	28,0%	3,0%	34,3%	0,6%	-0,5%	-3,2%	-3,3%	3,7%	-7,5	-109,28%
	2002	35,3%	20,5%	28,5%	25,2%	0,3%	-0,2%	-3,7%	-7,8%	1,9%	-5,7	23,85%
	2003	149,7%	4,4%	118,6%	39,7%	10,5%	4,5%	-14,5%	-40,8%	65,0%	-0,8	86,44%
	2004	-16,9%	-24,4%	71,0%	-52,1%	1,1%	-1,4%	35,0%	44,7%	43,0%	0,9	210,88%

A titre de comparaison, le tableau ci-dessous illustre la représentativité de l'échantillon par rapport au secteur :

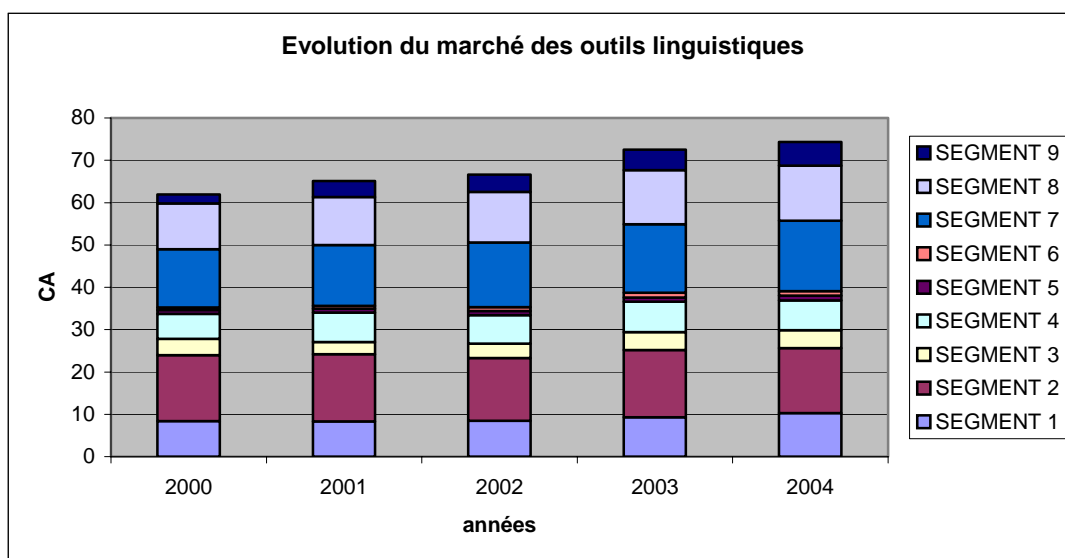


L'analyse de la répartition du CA de l'échantillon concernant le traitement des données textuelles et des données vocales met en lumière la prédominance du traitement textuel sur le marché français représentant environ 60 % de la valeur dégagée sur la période. La répartition texte/vocal de l'échantillon suit la même que celle du marché.



Répartition du CA	2000	2001	2002	2003	2004
CA texte	62,40%	61,67%	61,42%	61,41%	60,56%
CA vocal	37,60%	38,33%	38,58%	38,59%	39,44%

Concernant la répartition du CA par segment d'application, l'histogramme suivant résume sur la période 2000-2004 l'évolution de la valeur dégagée par chaque segment : 4 segments représentent plus de 70% de la valeur totale dégagée sur la période. Il s'agit du segment 1 « Gestion de contenu », du segment 2 « Gestion de documents », du segment 7 « Applications vocales et multimodales », et du segment 8 « E-learning ».



**Tableau de synthèse de l'échantillon des 46 sociétés françaises :**

Années	2000	2001	2002	2003	2004
CA (M euros)	61,9	65,2	66,7	72,3	74,2
Taux de croissance CA		5,32%	2,32%	8,43%	2,63%
Taux de croissance CA cumulé (2004-2000)	19,87%				
Résultat (M d'euros)	-3,5	-7,5	-5,7	-0,7	0,8
Taux de croissance Résultat		-114,29%	24,00%	87,72%	214,29%
Taux de croissance Résultat cumulé (2004-2000)	122,86%				
Effectifs totaux	1 468	1 467	1 419	1 533	1 563
Taux de croissance Effectifs totaux (2004-2000)		-0,07%	-3,27%	8,03%	1,96%
Taux de croissance cumulé Effectifs totaux (2004-2000)	6,47%				
Effectifs moyens par société	33	33	32	35	36

## 5. Phase 2 : France – Analyse du noyau dur du marché

### 5.1. Constitution d'un échantillon représentatif de 20 sociétés du cœur du marché des outils linguistiques.

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 de l'ingénierie linguistique en France et positionnées dans le noyau dur du marché avec des chiffres d'affaires communiqués sur 5 années consécutives.

L'échantillon est constitué des sociétés suivantes :

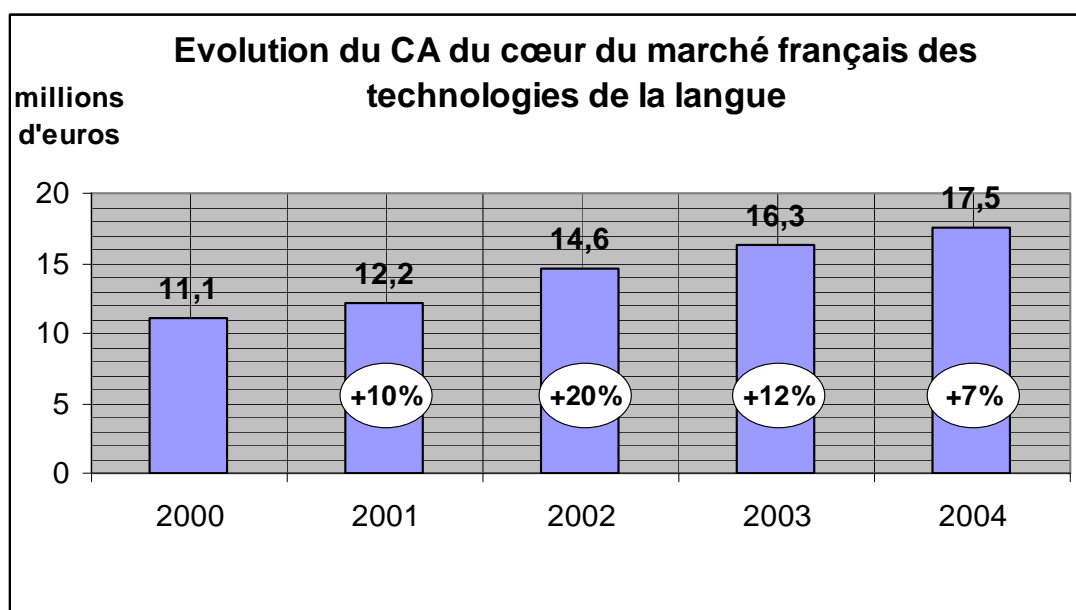
SOCIETES	T/V	Capital	Date de création	SEGMENTS HLT
ALOGIC	T	2 462 445	1999	1 2 9
AURALOG	T/V	1 546 173	1987	7 8
DIGIMIND	T	44 440	1998	1 9
EXALEAD	T	75 387	2000	1 4 9
GRIMMERSOFT	T	100 000	1995	1
LE SPHINX DEVELOPPEMENT	T	100 000	1994	1
LINGWAY	T	253 000	2001	1 9
MEDIAPPS	T	1 433 913	1999	1 2 4
MEMODATA	T	20 580	1989	5
MONDECA	T	440 500	1999	1 4 5
NEURO-CONCEPT	T/V	160 000	1990	8
NOEMATICS	T	75 900	1995	1 9
PERTIMM	T	7 622	1997	1 9
SOFTISSIMO	T/V	56 667	1986	3 5 6 7
SPSS	T	1 038 499	1993	1 9
SYNAPSE DEVELOPPEMENT	T	53 357	1994	3 5 6 9
SYSTRAN	T	15 087 000	1986	3
TELISMA	V	274 711	2000	7
TEMIS France	T	191 157	2000	1
VECSYS	V	1 011 750	1979	7

Sinequa, Arisem, Pertinence Mining, New Phenix et Go Albert, sociétés faisant également partie du noyau dur du marché des outils linguistiques en France, n'ont pas été introduites dans l'échantillon, leur date de création ne permettant pas d'avoir toutes les données socio-économiques sur la période 2000-2004.

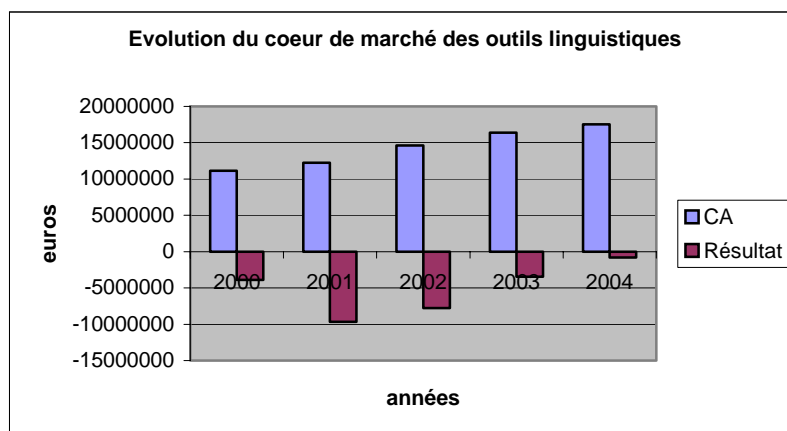
## 5.2. Evolution globale du CA et du résultat de l'échantillon

L'échantillon des 20 sociétés a généré un chiffre d'affaires de 17,5 millions d'euros en 2004. Le taux de croissance moyen sur la période a été de 12,08% et le taux cumulé sur 4 ans (2004-2001) a été de 57,24%.

CA	2000	2001	2002	2003	2004	taux de croissance moyen (2004-2001)	Taux de croissance cumulé (2004-2001)
TOTAL	11,1	12,2	14,6	16,3	17,5		
TAUX DE CROISSANCE		9,93%	19,63%	11,82%	6,93%	12,08%	57,24%



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.



La répartition de la valeur dégagée à partir du traitement des données textuelles et vocales reste stable sur la période : 70% de la valeur dégagée par le cœur du marché des technologies de la langue en France est assurée par les outils linguistiques traitant les données textuelles.

### 5.3. Evolution par segment d'application du cœur du marché

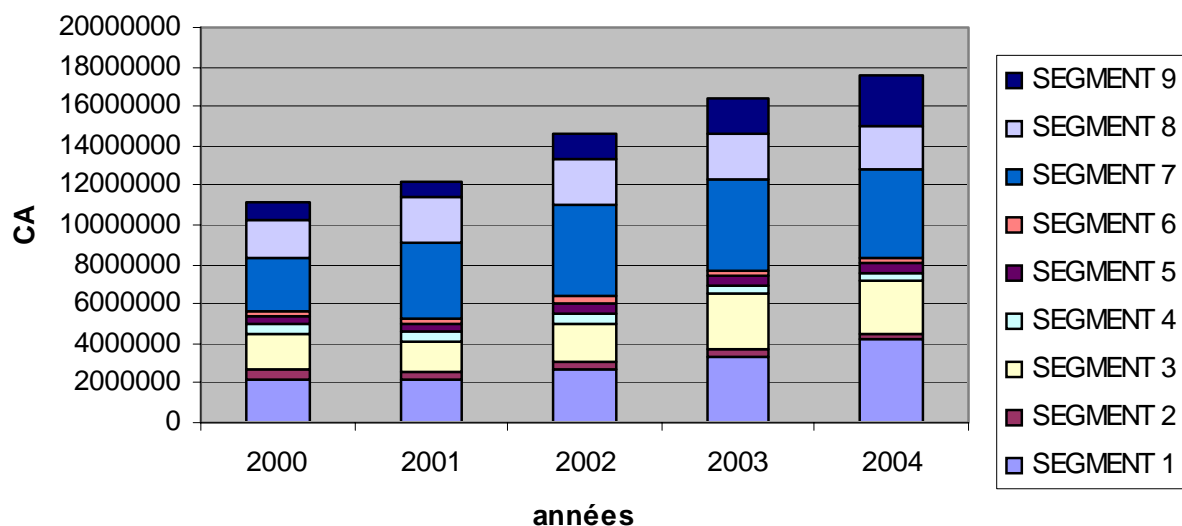
Le tableau ci-dessous résume l'évolution du chiffre d'affaires de chacun des 9 segments du marché des technologies de la langue sur la période. Il intègre aussi les taux de croissance annuels respectifs pour chaque segment, ainsi que les taux de croissance moyens et cumulés de 2001, 2002, 2003 et 2004.

CA	2000	2001	2002	2003	2004	taux de croissance moyen (2004-2001)	Taux de croissance cumulé (2004-2001)
Gestion de contenu	2198691	2167938	2638852	3325620	4270757		
TAUX DE CROISSANCE		-1,40%	21,72%	26,03%	28,42%	18,69%	94,24%
Gestion des documents	525 076	454 660	472 582	391 522	246 660		
TAUX DE CROISSANCE		-13,41%	3,94%	-17,15%	-37,00%	-15,91%	-53,02%
Traduction automatique	1 772 831	1 475 793	1 851 322	2 799 741	2 645 921		
TAUX DE CROISSANCE		-16,76%	25,45%	51,23%	-5,49%	13,61%	49,25%
E-business	539 219	498 667	558 180	445 206	405 797		
TAUX DE CROISSANCE		-7,52%	11,93%	-20,24%	-8,85%	-6,17%	-24,74%
Ressources linguistiques	339 174	399 927	496 351	421 348	449 072		
TAUX DE CROISSANCE		17,91%	24,11%	-15,11%	6,58%	8,37%	32,40%
Production / Edition	265 537	316 593	395 061	324 616	336 321		
TAUX DE CROISSANCE		19,23%	24,79%	-17,83%	3,61%	7,45%	26,66%
Applications vocales / multimodales	2 640 158	3 841 699	4 585 925	4 653 498	4 476 304		
TAUX DE CROISSANCE		45,51%	19,37%	1,47%	-3,81%	15,64%	69,55%
E-learning	1 942 523	2 227 482	2 341 339	2 238 748	2 130 475		
TAUX DE CROISSANCE		14,67%	5,11%	-4,38%	-4,84%	2,64%	9,68%
Moteurs de recherche	912 547	858 514	1 304 175	1 774 648	2 548 813		
TAUX DE CROISSANCE		-5,92%	51,91%	36,07%	43,62%	31,42%	179,31%
TOTAL	11135756	12 241 273	14 643 787	16 374 947	17 510 120		
TAUX DE CROISSANCE		9,93%	19,63%	11,82%	6,93%	12,08%	57,24%

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 fois 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.

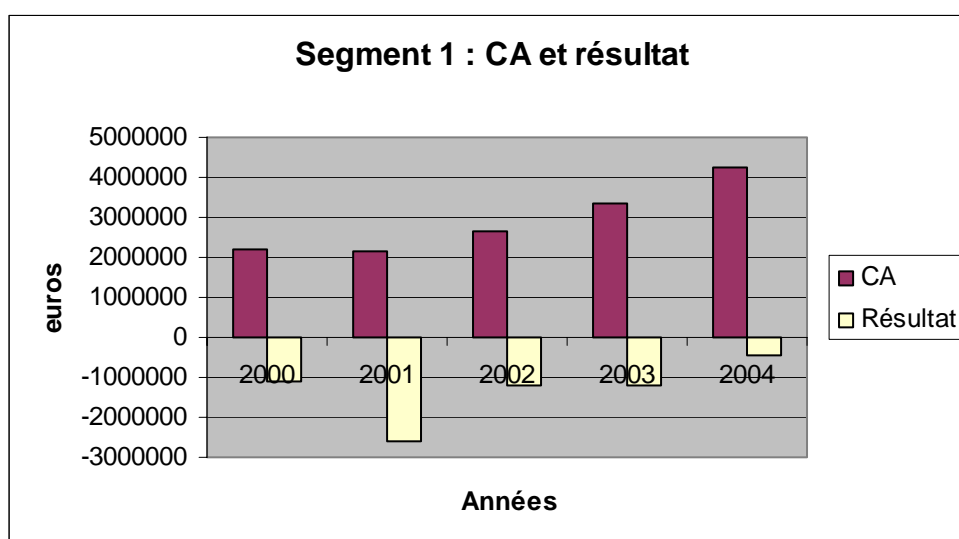
### Evolution du marché des technologies de la langue 2000-2004



REPARTITION DU CA	PART DE MARCHE MOYENNE (de 2000 à 2004)
Gestion de contenu	20,03%
Gestion des documents	3,09%
Traduction automatique	14,57%
E-business	3,55%
Ressources linguistiques	2,97%
Production & Edition	2,31%
Applications vocales & multimodales	28,08%
E-learning	15,49%
Moteurs de recherche	9,90%
TOTAL 20 SOCIETES	100,00%

### 5.3.1. Segment 1 : gestion de contenu

Dans le cœur du marché de l'ingénierie linguistique, le segment 1 « Gestion de contenu » a connu une évolution marquée sur la période 2000-2004 avec une forte progression de son chiffre d'affaires passant de 2,19 millions d'euros à 4,27 millions d'euros en 2004, soit un taux de croissance cumulé de 94,24% (18,69% par an en moyenne). La part de marché du segment évolue positivement sur la période passant de 19,74 % à 24,39%.



Les applications phares concernant la gestion de contenu sont principalement orientées autour de l'analyse du contenu textuel même si certaines applications commencent à intégrer l'analyse de contenu plus multimédia (documents sonores, images) et multilingues :

- dispositifs de veille Internet et d'intelligence économique
- analyse automatique de formulaires ou de textes
- visualisation cartographique

Les sociétés représentatives du cœur du marché des technologies de la langue et proposant des applications en matière de gestion de contenu sont entre autres les sociétés suivantes : Alogic, Grimmersoft, Le Sphinx Développement, Temis France, Mediapps, Mondeca, Exalead, Digimind, Lingway, Noematics, Pertimm, Spss.

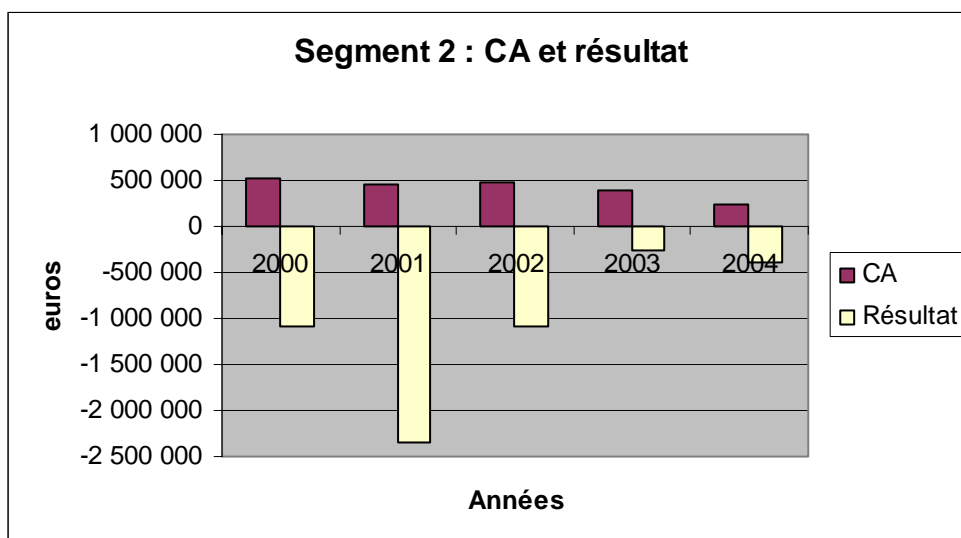
Evolution des effectifs du segment représenté par l'échantillon du cœur de marché (effectifs cumulés des sociétés du segment):

Segments	Effectifs	2000	2001	2002	2003	2004
	Effectifs totaux	138	149	132	161	177
Segment 1	Effectifs moyens / société positionné sur le segment	12	12	11	13	15



### 5.3.2. Segment 2 : Gestion des documents

Le segment 2 « Gestion des documents » est passé d'un chiffre d'affaires de 525 076 euros en 2000 à 214 660 euros en 2004. La forte baisse du segment sur la période 2000-2004 avec un taux de croissance moyen de  $-15,91\%$  (et un taux de croissance cumulé de  $-53,02\%$ ) est à minimiser du fait du très faible nombre d'acteurs du cœur du marché des outils linguistiques positionné sur le segment de la gestion des documents. La part de marché du segment évolue négativement sur la période passant de  $4,72\%$  à  $1,41\%$ .



Les applications phares concernant la gestion de documents sont principalement orientées autour de la structuration et de la catégorisation des documents textuels même si les dispositifs mis en place intègrent de plus en plus les documents multicomposites et audiovisuels (films, images fixes et animées, sons) :

- gestion des archives et des documents numériques
- travail collaboratif assisté par ordinateur (catégorisation, workflow).

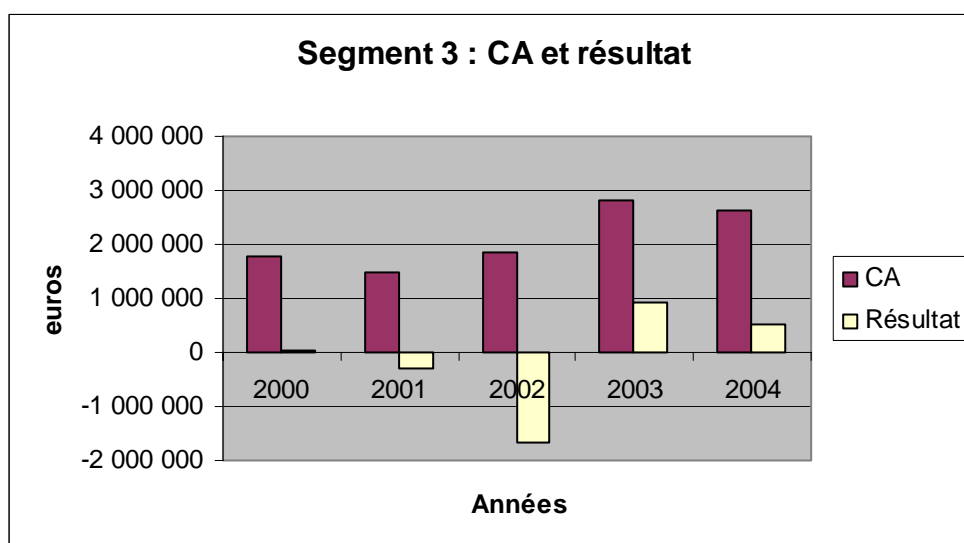
Les sociétés représentatives du cœur du marché des technologies de la langue et proposant des applications liées à la gestion de documents sont Alogic et Mediapps (récemment absorbé par la société de GED et de Portail Ever Team).

Evolution des effectifs du segment représenté par l'échantillon du cœur de marché (effectifs cumulés des sociétés du segment):

Segments	Effectifs	2000	2001	2002	2003	2004
	Effectifs totaux	85	93	44	32	24
Segment 2	Effectifs moyens / société positionné sur le segment	43	47	22	16	12

### 5.3.3. Segment 3 : Traduction automatique

Le segment 3 « Traduction automatique et assistée par ordinateur » a connu une évolution non linéaire sur la période 2000-2004 avec une progression de son chiffre d'affaires de 49,25%, passant de 1,7 millions d'euros à 2,6 millions d'euros en 2004. Le taux de croissance moyen sur la période a été de 13.61%. La part de marché du segment baisse légèrement mais reste stable sur la période passant de 15,92% à 15,11%.



Les applications phares concernant la traduction automatique sont principalement orientées autour de la traduction de documents textuels et la localisation d'information (sites Internet, documentations techniques et documents de présentation localisés) :

- traduction de texte
- localisation de contenu multilingue (traduction de sites Internet multilingues, informations locales)
- traduction vocale (vocal – texte –vocal).

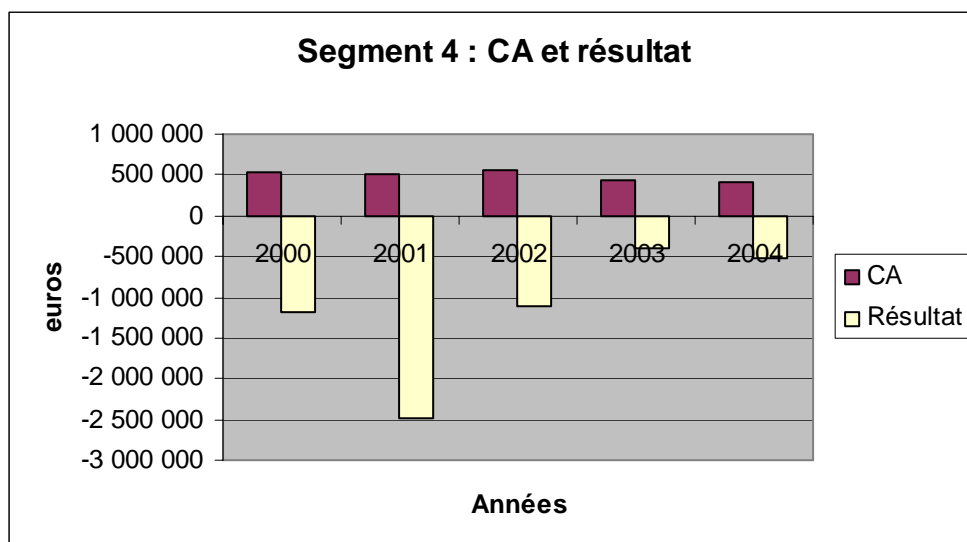
Les sociétés représentatives du cœur du marché des technologies de la langue et proposant des applications liées à la traduction automatique sont Softissimo, Synapse Developpement, et Systran.

Evolution des effectifs du segment représenté par l'échantillon du cœur de marché (effectifs cumulés des sociétés du segment):

Segments	Effectifs	2000	2001	2002	2003	2004
Segment 3	Effectifs totaux	25	17	43	43	41
	Effectifs moyens / société positionné sur le segment	8	6	14	14	14

### 5.3.4. Segment 4 : E-business

Le segment 4 « E-Business » a connu une évolution stable mais négative de son chiffre d'affaires sur la période 2000-2004, avec un taux de croissance moyen de  $-6,17\%$  et un taux de croissance cumulé de  $-24,74\%$  sur 4 ans (2001 à 2004). Le chiffre d'affaires dégagé est passé de 0,53 millions d'euros à 0,4 millions d'euros en 2004. La part de marché du segment évolue négativement sur la période passant de 4,84% à 2,32%.



Les applications phares concernant le E-Business sont principalement les applications intelligentes optimisant les portails Internet et l'accès aux catalogues (personnalisation des informations, constitution de terminologies)

- portails B2B et B2C
- Sites de communautés d'internautes

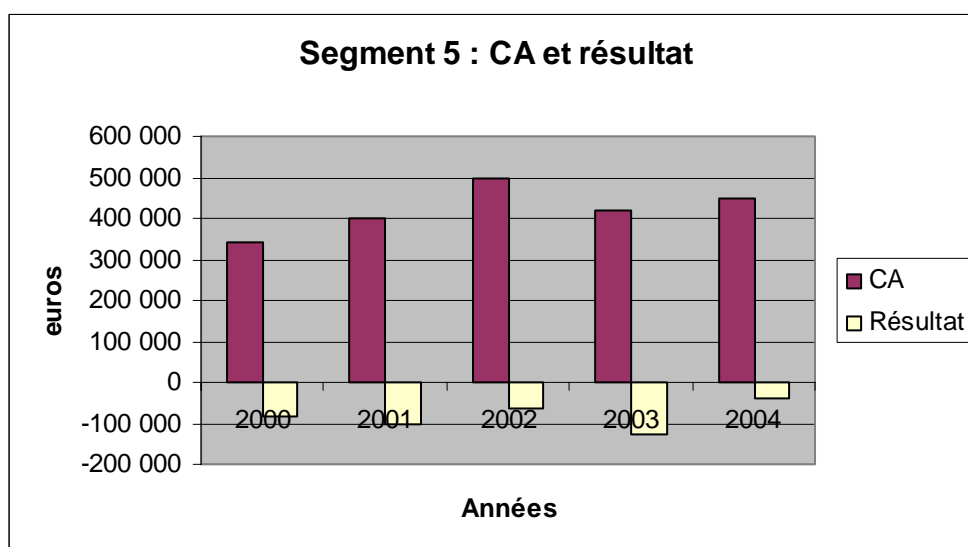
Les sociétés représentatives du cœur du marché des technologies de la langue et proposant des applications liées au E-Business sont entre autres Exalead, Mediapps, et Mondeca.

Evolution des effectifs du segment représenté par l'échantillon du cœur de marché (effectifs cumulés des sociétés du segment):

Segments	Effectifs	2000	2001	2002	2003	2004
	Effectifs totaux	82	89	56	46	48
Segment 4	Effectifs moyens / société positionné sur le segment	27	30	19	15	16

### 5.3.5. Segment 5 : Ressources linguistiques

Le segment 5 « Ressources linguistiques » a connu une évolution positive mais non linéaire sur la période 2000-2004 avec une progression annuelle moyenne de son chiffre d'affaires de 8,37% (+32.40% en taux de croissance cumulé de 2001 à 2004), passant de 0,33 millions d'euros à 0,45 millions d'euros en 2004. La part de marché du segment évolue négativement sur la période passant de 3,05 % à 2,56%.



Les applications phares concernant les ressources linguistiques texte et voix sont principalement :

- les thésaurus et les dictionnaires
- les outils associés (dictionnaires, thésaurus, grammaire, phonèmes et échantillons de voix).

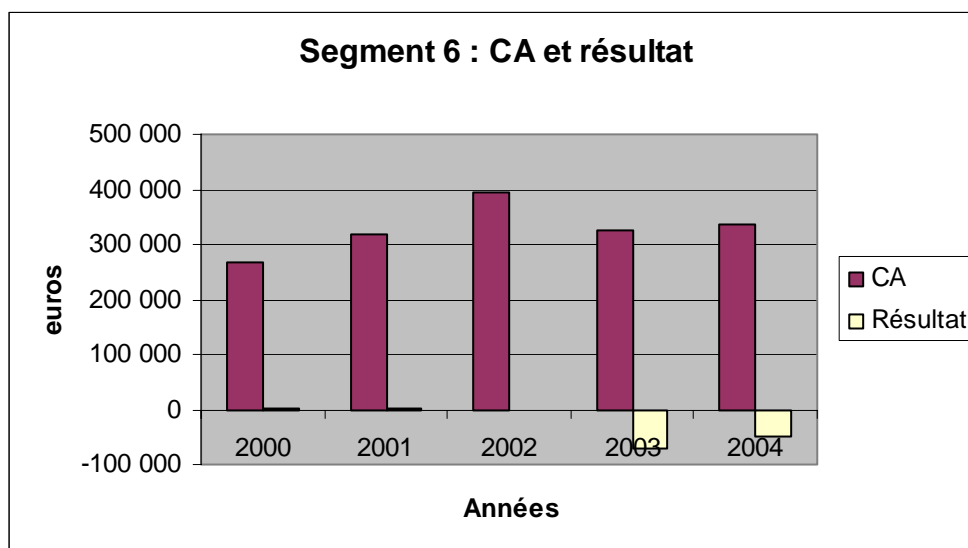
Les sociétés représentatives du cœur du marché des technologies de la langue et proposant des applications liées aux ressources linguistiques sont entre autres Memodata, Mondeca, Softissimo, et Synapse Développement.

L'évolution des effectifs du segment représenté par l'échantillon du cœur de marché (effectifs cumulés des sociétés du segment):

Segments	Effectifs	2000	2001	2002	2003	2004
	Effectifs totaux	4	20	22	22	27
Segment 5	Effectifs moyens / société positionné sur le segment	1	5	6	6	7

### 5.3.6. Segment 6 : Production et édition.

Le segment 6 « Production et Edition » a connu une évolution marquée sur la période 2000-2004 avec une progression moyenne annuelle de son chiffre d'affaires de 7,45%, passant de 0,26 millions d'euros à 0,33 millions d'euros en 2004. La part de marché du segment évolue négativement sur la période passant de 2,38 % à 1,92%.



Les applications phares relatives au segment « Production et Edition » concernent notamment :

- les outils d'aide à la rédaction (correcteurs orthographiques, production de résumés, générateurs de texte et de liens hypertextes)
- les outils de dictée vocale et de text-to-speech

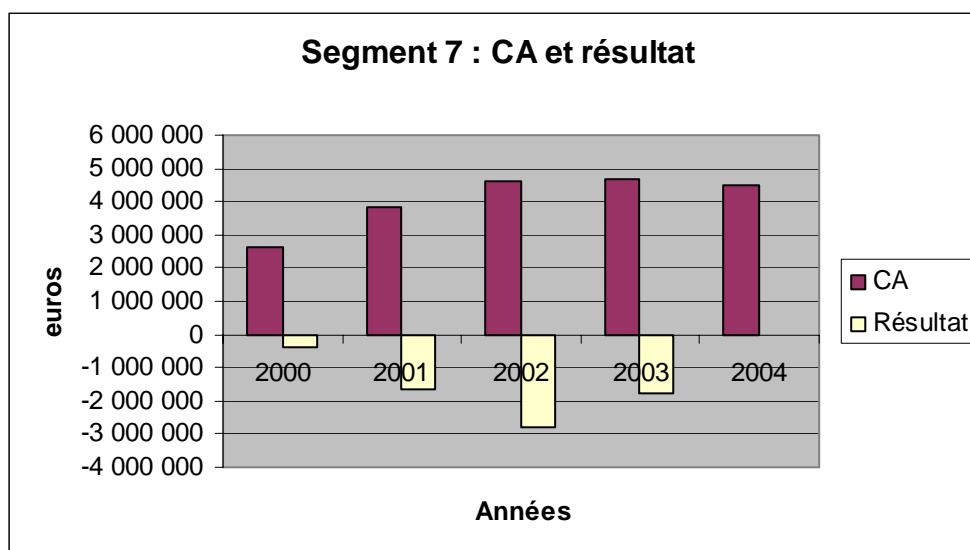
Les sociétés représentatives du cœur du marché des technologies de la langue et proposant des applications liées au segment « Production et Edition » sont entre autres Softissimo et Synapse Développement.

L'évolution des effectifs du segment représenté par l'échantillon du cœur de marché (effectifs cumulés des sociétés du segment):

Segments	Effectifs	2000	2001	2002	2003	2004
Segment 6	Effectifs totaux	0	17	17	17	15
	Effectifs moyens / société positionné sur le segment	0	9	9	9	8

### 5.3.7. Segment 7 : Interfaces vocales et multimodales

Le segment 7 « Interfaces vocales et multimodales » a connu une évolution positive marquée sur la période 2000-2004 avec une progression annuelle moyenne de son chiffre d'affaires de 15,64% (+ 69,55% en cumulé sur la période), passant de 2,6 millions d'euros à 4,4 millions d'euros en 2004. La part de marché du segment évolue positivement sur la période passant de 23,71 % à 25,56%.



Les applications phares relatives au segment « Interfaces vocales et multimodales » concernent notamment :

- les call centers avec dispositifs intelligents
- les kiosques et bornes multimédias,
- les solutions embarquées et la domotique,
- les jeux vidéos et de divertissement (automatisation, reconnaissance et synthèse vocale, navigation multimédia et interface visuelle).

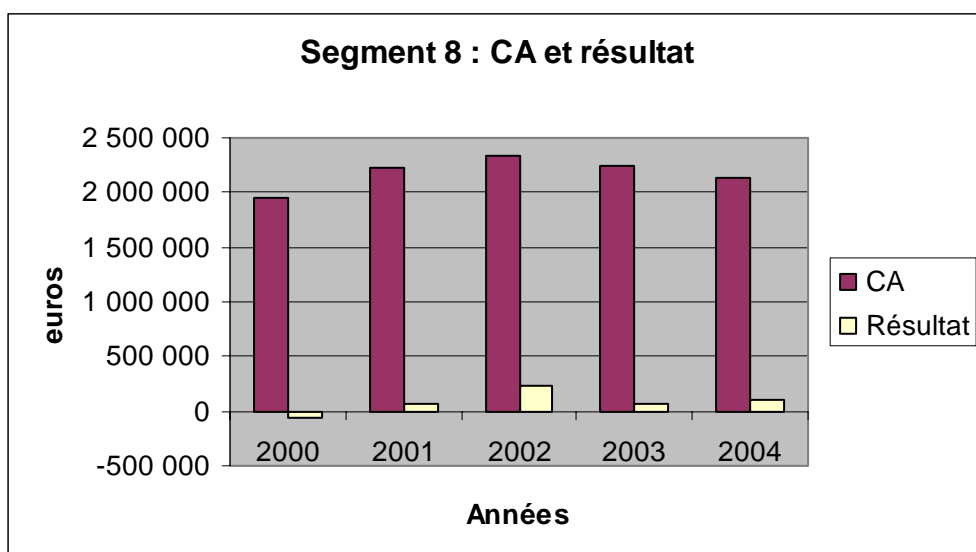
Les sociétés représentatives du cœur du marché des technologies de la langue et proposant des applications liées aux « Interfaces vocales et multimodales » sont entre autres Auralog, Softissimo, Telisma (partenariat d'intégration à Prosodie), et Vecsys.

L'évolution des effectifs du segment représenté par l'échantillon du cœur de marché (effectifs cumulés des sociétés du segment):

Segments	Effectifs	2000	2001	2002	2003	2004
Segment 7	Effectifs totaux	161	216	246	227	210
	Effectifs moyens / société positionné sur le segment	40	54	62	57	53

### 5.3.8. Segment 8 : E-learning

Le segment 8 « E-learning » a connu une évolution plutôt stable mais non linéaire sur la période 2000-2004 avec une progression moyenne annuelle de son chiffre d'affaires de 2,64% sur la période, passant de 1,9 millions d'euros à 2,1 millions d'euros en 2004. Le taux de croissance cumulée sur la période 2000-2004 a été de 9,68%. La part de marché du segment évolue négativement sur la période passant de 17,44 % à 12,17%.



Les applications phares relatives au segment e-learning concernent notamment les applications suivantes :

- les applications pédagogiques et le e-learning,
- l'apprentissage des langues,
- l'aide au handicap (formation par Internet, outils pédagogiques multimédia, apprentissage des langues, accès à l'informatique).

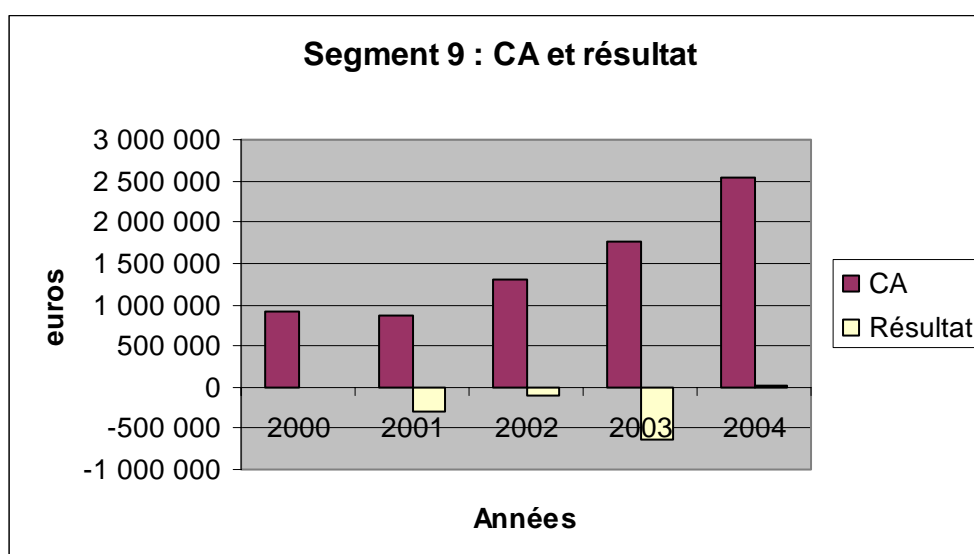
Les sociétés représentatives du cœur du marché des technologies de la langue et proposant des applications liées au e-learning sont entre autres Auralog, Neuro-Concept.

L'évolution des effectifs du segment représenté par l'échantillon du cœur de marché (effectifs cumulés des sociétés du segment):

Segments	Effectifs	2000	2001	2002	2003	2004
	Effectifs totaux	141	118	143	128	116
Segment 8	Effectifs moyens / société positionné sur le segment	71	59	72	64	58

### 5.3.9. Segment 9 : Moteurs de recherche intelligents

Le segment 9 « Moteurs de recherche intelligents » a connu une évolution très marquée sur la période 2000-2004 avec une progression annuelle moyenne de son chiffre d'affaires de 31,42%, passant de 0,9 million d'euros à 2,5 millions d'euros en 2004. Le taux de croissance cumulée sur la période 2000-2004 a été de 179,31%. La part de marché du segment évolue positivement sur la période passant de 8,19 % à 14,56%.



Les applications phares concernant les moteurs de recherche intelligents sont notamment :

- les outils de recherche pour Internet, Intranet et Extranet liés au traitement du langage naturel et au, web sémantique

Les sociétés représentatives du cœur du marché des technologies de la langue et proposant des applications liées aux moteurs de recherche intelligents sont : Alogic, Digimind, Exalead, Lingway, Noematics, Pertimm, Spss, Synapse Développement.

L'évolution des effectifs du segment représenté par l'échantillon du cœur de marché (effectifs cumulés des sociétés du segment):

Segments	Effectifs	2000	2001	2002	2003	2004
	Effectifs totaux	50	51	81	107	117
Segment 9	Effectifs moyens / société positionné sur le segment	6	6	10	13	15

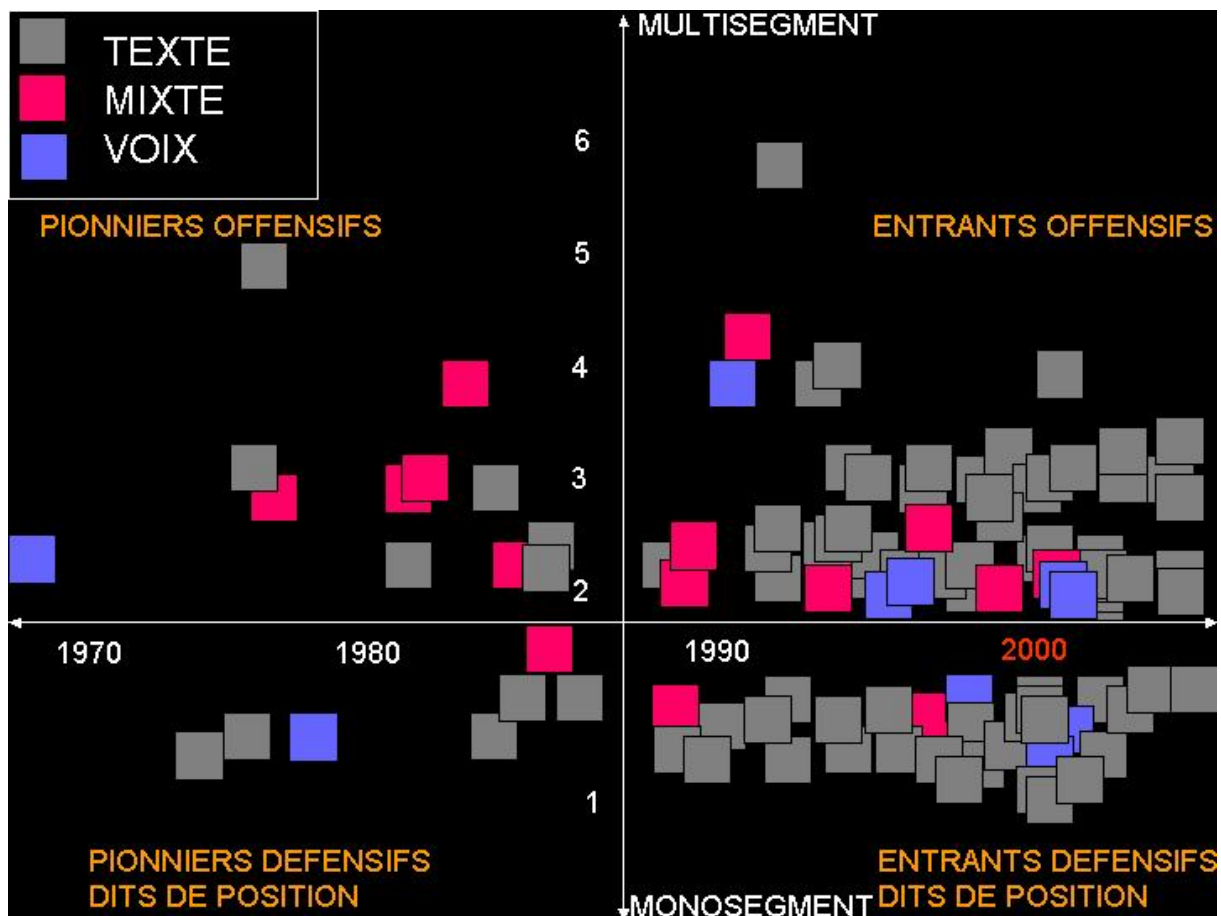


**6. Phase 3 : Tendances et perspectives du marché des technologies de la langue en France et en Europe.**

**6.1. Analyse des tendances du marché français**

**6.1.1. Cartographie des acteurs**

- **Cartographie des 109 acteurs du marché des technologies de la langue par année de création / spécialisation :**



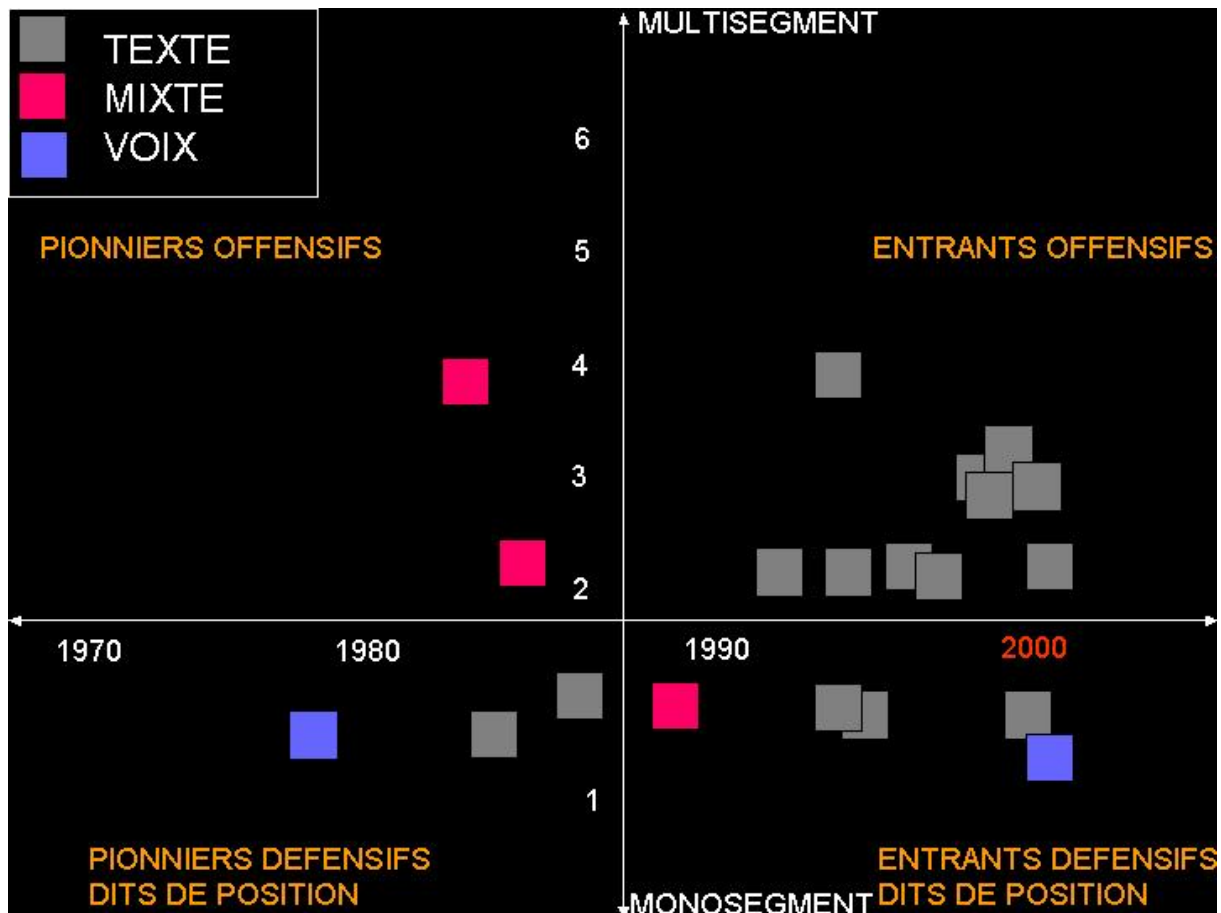
La création de sociétés dans le secteur des technologies de la langue est toujours dynamique : la majorité des nouveaux entrants adoptent une stratégie offensive dans la mesure où la plupart des acteurs se positionnent sur 2 voire 3 segments d'applications plutôt qu'un seul.

Toutefois, le développement des outils linguistiques semble encore principalement orienté en France vers le traitement des données textuelles exclusivement : le nombre de sociétés positionnées sur le traitement mixte des données textuelles et vocales n'a pas considérablement augmenté depuis 2000. Plusieurs rapprochements et rachats au niveau européen peuvent expliquer ce constat en ce sens où le nombre d'acteurs concernés et

positionnés sur le traitement des données vocales diminue par concentration des compétences et des technologies développées par des PME qui sont de plus en plus regroupés (par rachat ou fusion notamment) au sein d'une entité unique souvent définie au niveau européen (holding, consortium, ...) favorisant ainsi la mise sur le marché de produits mixtes et multilingues.

Le nombre de sociétés positionnées sur le traitement des données vocales est passé de 20% en 2002 à 30% en 2005. Cette tendance devrait s'accroître compte tenu de l'intégration des données voix numérisées via les réseaux.

- **Cartographie des 20 acteurs du cœur de marché des outils linguistiques par année de création / spécialisation :**



La majorité des nouveaux entrants adoptent une stratégie plutôt offensive dans la mesure où la plupart des acteurs se positionnent généralement sur 2 d'applications.

Le cœur du marché des technologies de la langue en France est marqué par des sociétés de l'ingénierie linguistique le plus souvent positionnées sur le traitement des données textuelles exclusivement. Le ratio reste néanmoins dans la tendance européenne puisque la répartition

des acteurs était de 80/20 en 2002 au niveau de l'Union européen à 13 pays, de 70/30 en 2004 au niveau de l'Union européenne à 25 pays, et de 75/25 dans le noyau dur du marché français des technologies de la langue.

### 6.1.2. Analyse concurrentielle et marketing du marché français

Le positionnement concurrentiel et stratégique des segments du marché de 2000 à 2004 a été réalisé à partir du mapping stratégique des 9 segments.

L'analyse des taux de croissance annuels moyens des chiffres d'affaires dégagés par les 9 segments et de leurs parts de marché moyenne sur la période rapporté à la moyenne théorique d'un segment de marché (soit pdmm / 11,11%) permet d'identifier sur la période 2001-2004:

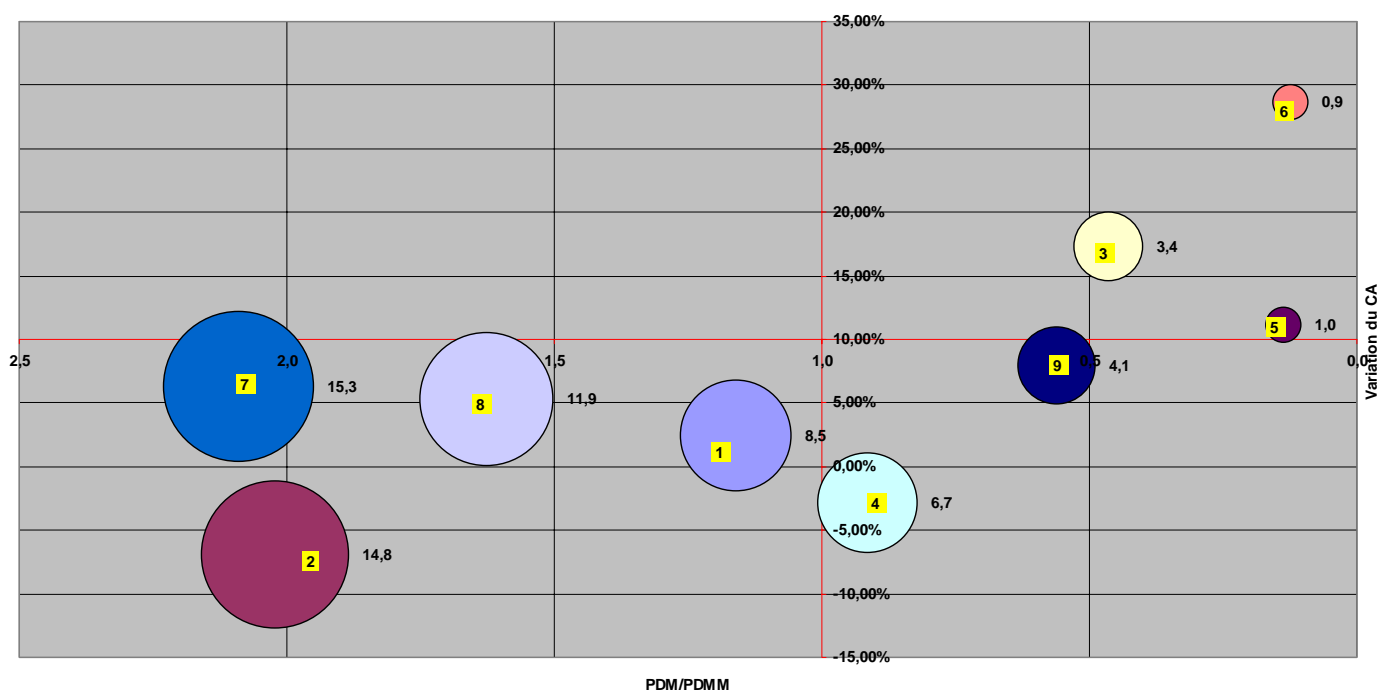
- les segments matures (taux de croissance stable mais faible, part de marché pondérée à la moyenne importante) ;
- les segments vedettes (fort taux de croissance, part de marché pondérée à la moyenne importante) ;
- les segments émergents (fort taux de croissance, part de marché pondérée à la moyenne faible) ;
- les segments suiveurs (faible taux de croissance, part de marché pondérée à la moyenne faible).

Les axes rouges représentent en abscisse le taux de croissance à 10% et en ordonnées le ratio part de marché du segment / part de marché moyenne égal à 1 (>1 à gauche, <1 à droite). La taille des bulles est proportionnelle aux CA dégagés par les segments.

#### • Mapping stratégique des 9 segments en 2004 (marché des technologies de la langue) :

##### Année 2002 (variation 2002-2001) :

Evolution des segments (2001-2002)



Durant l'année 2002, trois segments ont dépassé les 10% de croissance. Ces 3 segments présentent également des parts de marché relative à la moyenne inférieures à 1 : il s'agit des segments Production et édition, Ressources linguistiques et Traduction automatique. L'impact de ces segments sur l'évolution du marché et l'accroissement de sa valeur est négligeable. Les segments significatifs de l'année 2002 sont entre autres les segments :

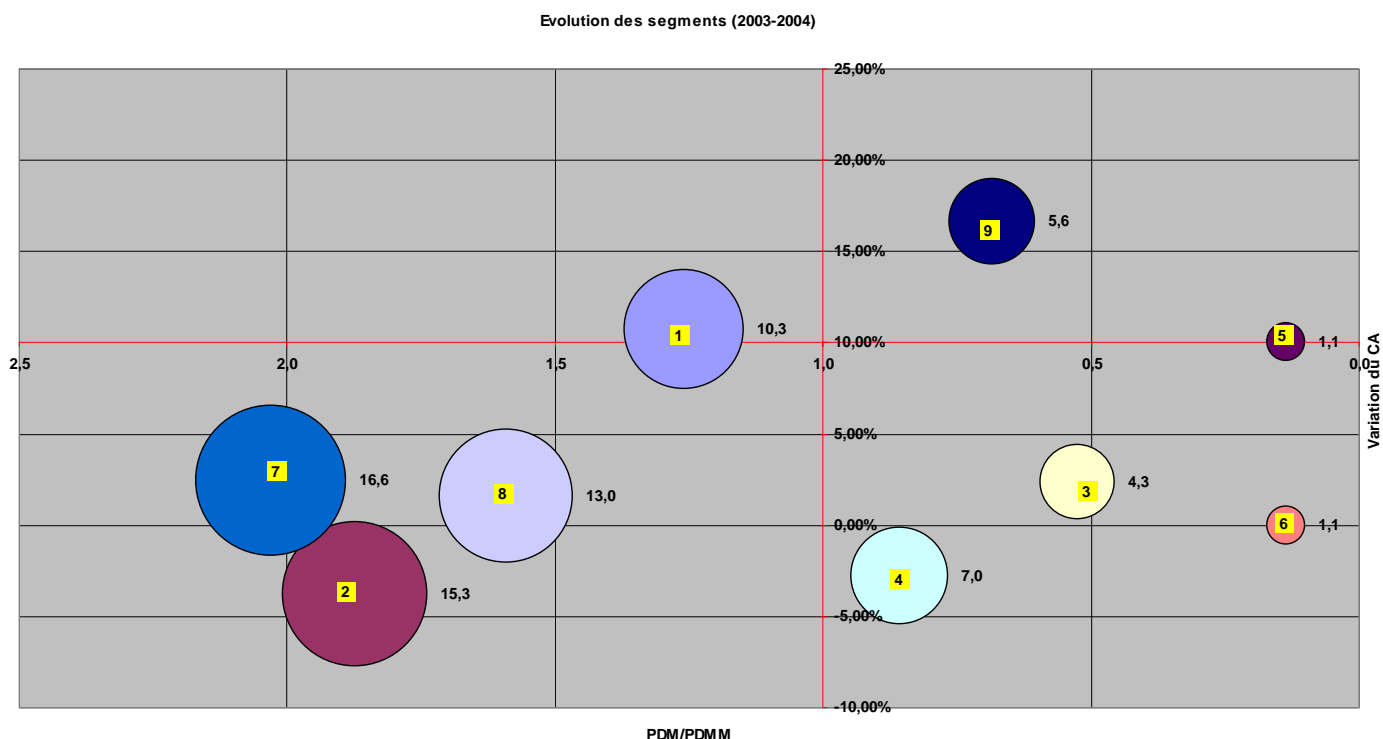
- gestion de documents
- applications vocales et multimodales
- e-learning

Les segments gestion de contenu et e-business sont présents mais dégagent une valeur moins importante.

Le segment moteurs de recherche reste encore très négligeable dans le secteur des technologies de la langue.

Le segment traduction automatique a connu une forte croissance sur la période (environ 20%). Le chiffre d'affaire du segment Production et Edition augmente de près 30% sur la période et celui des ressources linguistiques de plus de 10%. Ces deux segments présentent des parts de marché relatives encore très négligeables.

#### Année 2004 (variation 2004-2003) :



Durant l'année 2004, trois segments ont dépassé les 10% de croissance. Il s'agit des segments Gestion de contenu, moteurs de recherche et Production/Édition. Le segment gestion de contenu a une part de marché relative à la moyenne supérieure à 1 : l'évolution de ce segment a un impact conséquent sur le marché des technologies de langue. Les autres segments impactant le marché sont notamment les segments :

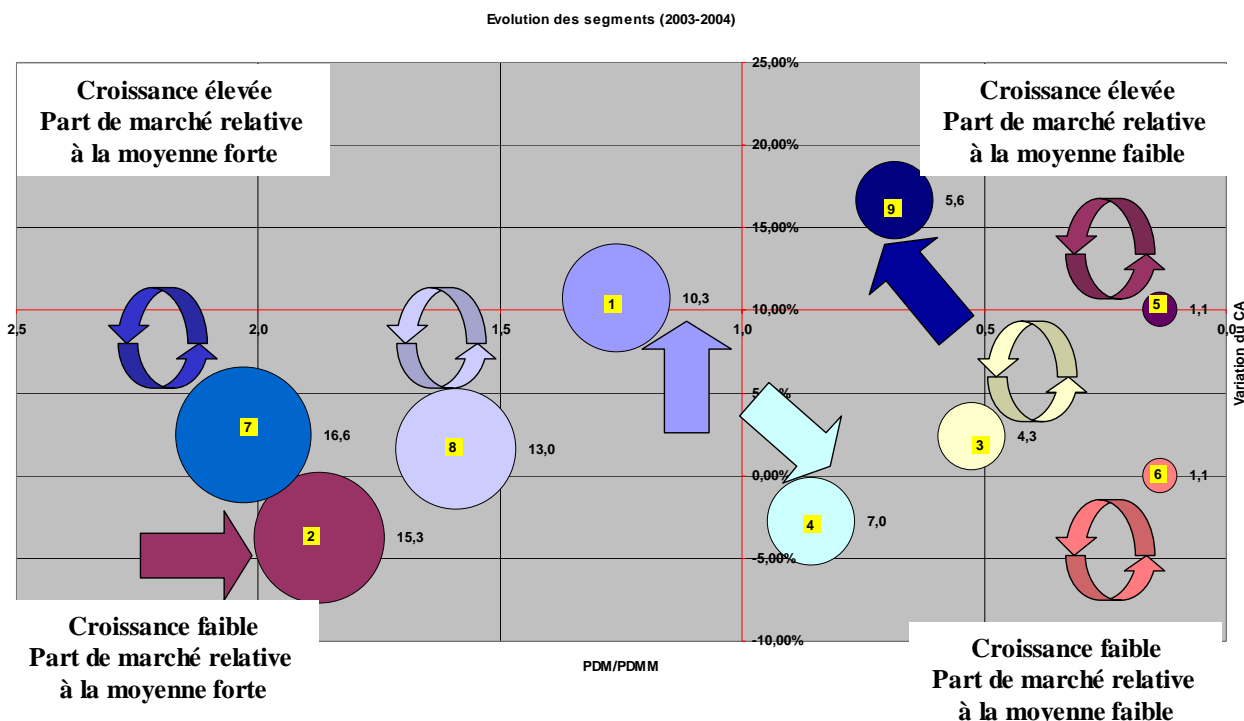
- gestion de documents
- applications vocales et multimodales
- e-learning

Le segment e-business reste stagnant sur la période 2004 par rapport à 2002. Le segment traduction automatique accuse une baisse son chiffre d'affaires en 2004 par rapport à 2003.

Le segment moteurs de recherche reste encore peu important mais sa croissance à deux chiffres se confirme en 2004. Les segments Production / Édition et ressources linguistiques conservent des parts de marché relatives encore très négligeables.

**Evolution fléchée des segments (2000-2004) :**

**Tendances du marché des technologies de la langue**



Les segments ayant des caractéristiques communes relatives à leur croissance ont été regroupés :

**Applications matures du marché des technologies de la langue:**

- Segment 7 : Applications vocales et multimodales
- Segment 8 : e-learning
- Segment 1: Gestion de contenu
- Segment 2 : Gestion des documents

**Segments émergents :**

- Segment 9 : Moteurs de recherche intelligents

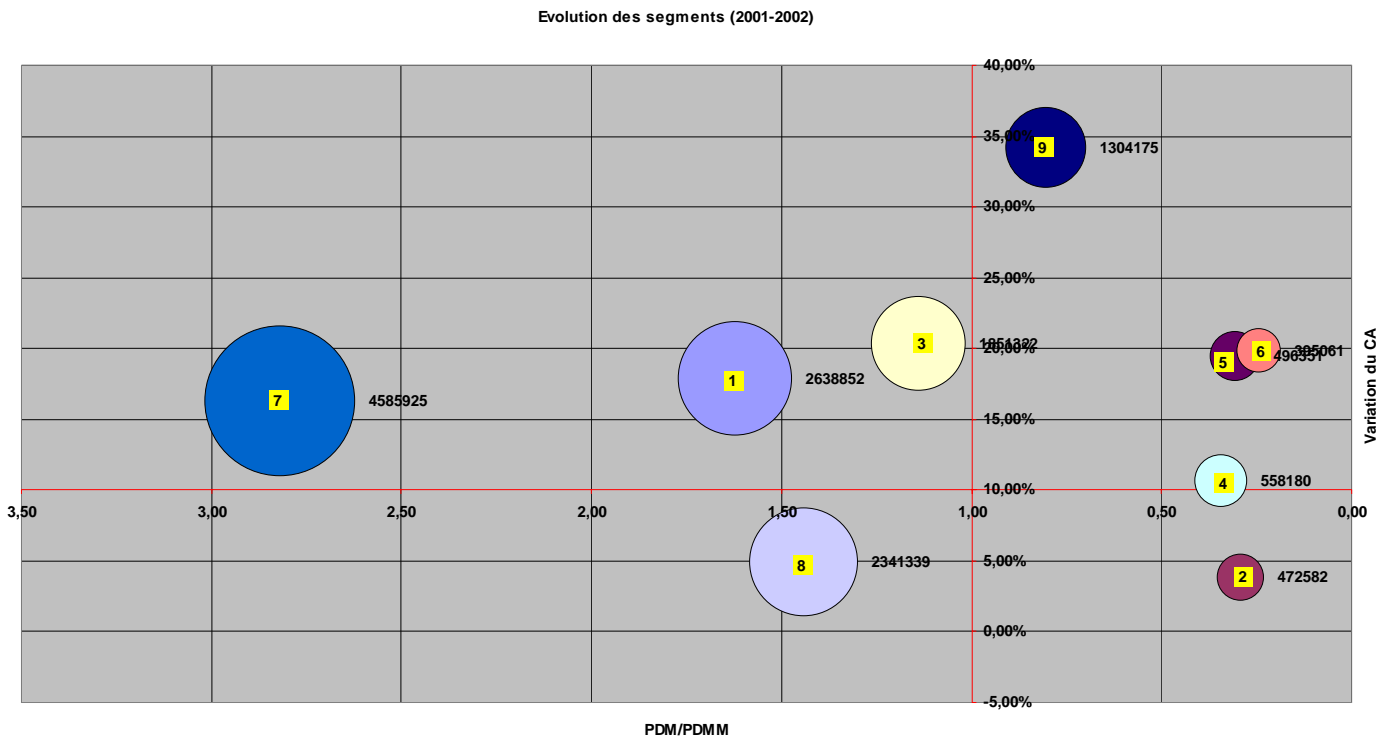
**Segments suiveurs :**

- Segments 5 : ressources linguistiques
- Segment 6 : Production et édition
- Segment 4: e-Business

<b>Evolution du marché des technologies de la langue (2004-2000)</b>	<b>Taux de croissance du Chiffre d'affaires</b>	<b>Taux de croissance de la part de marché relative à la moyenne</b>
<b>Gestion de contenu</b>	<b>22,62%</b>	<b>2,16%</b>
<b>Gestion des documents</b>	<b>-1,92%</b>	<b>-18,29%</b>
<b>Traduction automatique</b>	<b>13,16%</b>	<b>-5,73%</b>
<b>E-business</b>	<b>18,64%</b>	<b>-1,16%</b>
<b>Ressources linguistiques</b>	<b>22,22%</b>	<b>1,82%</b>
<b>Production &amp; Edition</b>	<b>83,33%</b>	<b>52,74%</b>
<b>Applications vocales et multimodales</b>	<b>20,29%</b>	<b>0,21%</b>
<b>E-learning</b>	<b>20,37%</b>	<b>0,28%</b>
<b>Moteurs de recherche</b>	<b>166,67%</b>	<b>122,16%</b>

- Mapping stratégique des 9 segments en 2004 (cœur de marché de l'ingénierie linguistique)

Année 2002 (variation 2002-2001) :



Durant l'année 2002, six segments ont dépassé les 10% de croissance : il s'agit des segments e-business, Production et Edition, Ressources linguistiques, Traduction automatique, Gestion de contenu et Applications vocales et multimodales. Parmi ces 6 segments, trois présentent également des part de marché relative à la moyenne supérieures à 1 (Traduction automatique, Gestion de contenu et Applications vocales et multimodales) : leur impact sur l'évolution du cœur du marché et l'accroissement de sa valeur est conséquent. L'autre segment important du marché en 2002 est le e-learning.

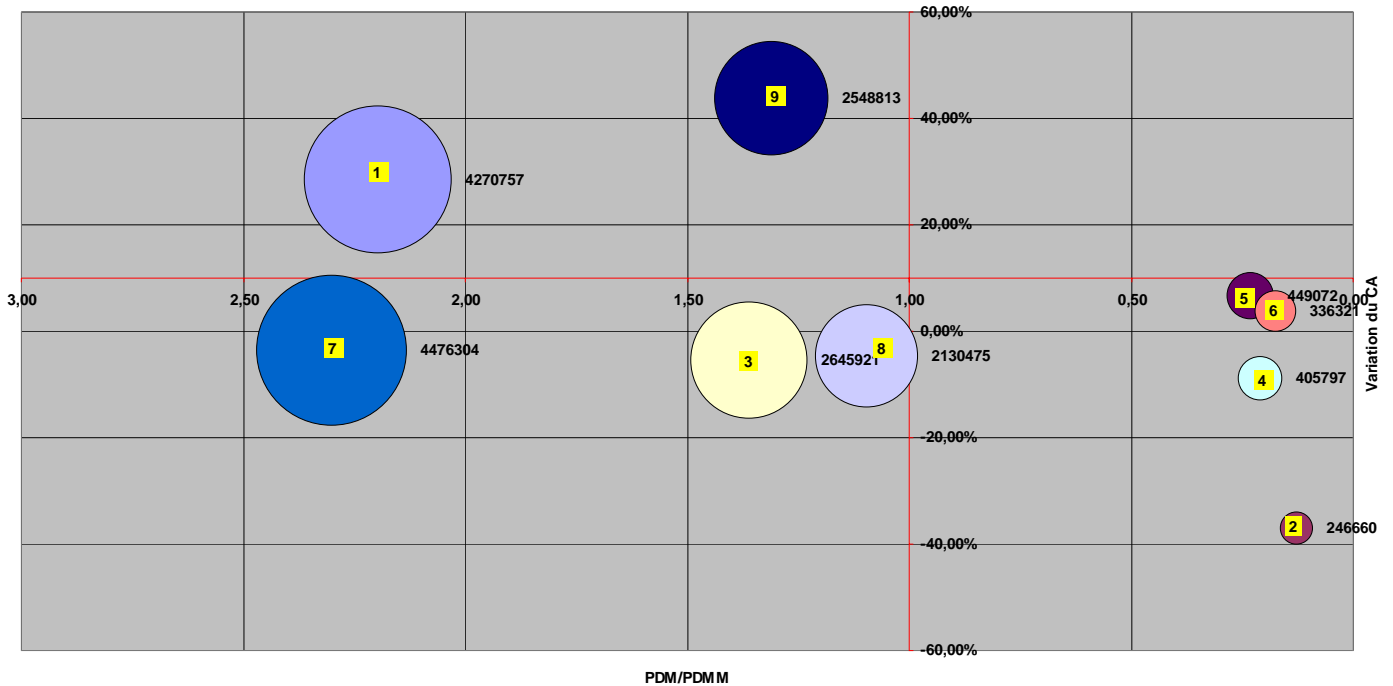
Les segments e-business, Gestion des documents, Production et Edition, Ressources linguistiques restent des segments marginaux sur la période.

Le segment moteurs de recherche reste très négligeable dans le secteur des technologies de la langue présente un taux de croissance deux fois supérieur au taux de croissance du 2 segment le plus dynamique du marché en 2002.

Le segment traduction automatique a connu une forte croissance sur la période (environ 20%).

**Année 2004 (variation 2004-2003) :**

Evolution des segments (2003-2004)



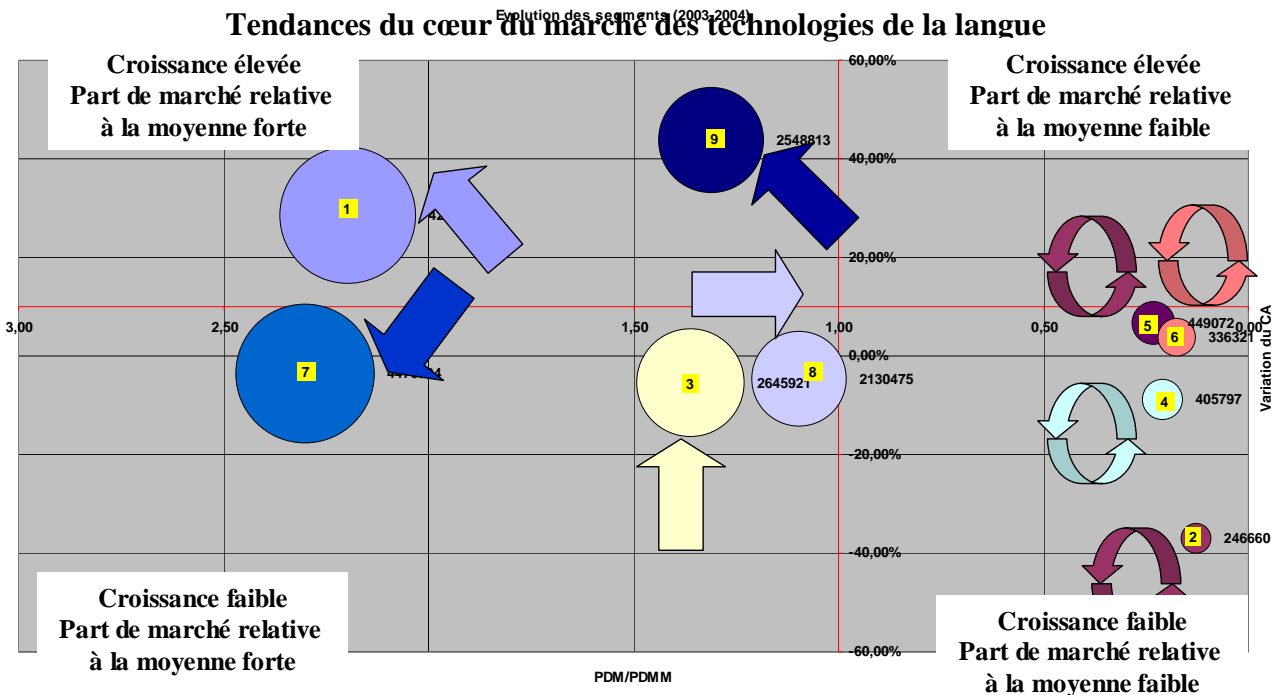
Durant l'année 2004, deux segments ont dépassé les 10% de croissance. Il s'agit des segments Gestion de contenu, et moteurs de recherche. Ces 2 segments gestion de contenu ont des parts de marché relatives à la moyenne supérieures à 1 : la croissance de ces segment influe directement le cœur du marché des technologies de langue. Les autres segments impactant le marché sont aussi :

- les applications vocales et multimodales
- la traductio automatique
- le e-learning

Les segment e-business, gestion de documents, production et édition, ressources linguistiques restent stagnants en 2004 par rapport à 2002. leur impact sur le cœur de marché est négligeable car leurs parts de marché relatives à la moyenne sont largement inférieures à 1 en 2004.



**Evolution fléchée des segments (2000-2004) :**



**Applications matures du cœur du marché des technologies de la langue:**

- Segment 7 : Applications multimodales
- Segment 3 : Traduction automatique et assistée par ordinateur
- Segment 8: e-learning

**Applications vedettes du cœur du marché des technologies de la langue:**

- Segment 1: Gestion de contenu
- Segment 9 : Moteurs de recherche

**Segments suiveurs :**

- Segments 2 : GED
- Segments 4 : e-Business
- Segment 5 : ressources linguistiques
- Segment 6 : Production et édition

Dans le cœur du marché des technologies de la langue, les segments matures et les segments vedettes sont clairement identifiés avec des indicateurs très significatifs souvent à 2 chiffres.

<b>Evolution du cœur de marché des technologies de la langue (2004-2000)</b>	<b>Taux de croissance du Chiffre d'affaires</b>	<b>Taux de croissance de la part de marché relative à la moyenne</b>
<b>Gestion de contenu</b>	<b>94,24%</b>	<b>23,53%</b>
<b>Gestion des documents</b>	<b>-53,02%</b>	<b>-70,13%</b>
<b>Traduction automatique</b>	<b>49,25%</b>	<b>-5,08%</b>
<b>E-business</b>	<b>-24,74%</b>	<b>-52,14%</b>
<b>Ressources linguistiques</b>	<b>32,40%</b>	<b>-15,80%</b>
<b>Production &amp; Edition</b>	<b>26,66%</b>	<b>-19,45%</b>
<b>Applications vocales et multimodales</b>	<b>69,55%</b>	<b>7,83%</b>
<b>E-learning</b>	<b>9,68%</b>	<b>-30,25%</b>
<b>Moteurs de recherche</b>	<b>179,31%</b>	<b>77,63%</b>

Sur la période 2000-2004, les tendances communes au marché des technologies de la langue et à son cœur sont notamment :

- la hausse du nombre d'acteurs du marché des technologies de la langue :
  - 377 acteurs dans l'Europe des 13 en 2003, 482 dans l'Europe des 25 en 2005
  - 96 acteurs en France en 2003, 109 dans l'Europe des 25 en 2005
- Le pourcentage important d'acteurs positionnés sur au moins 2 segments d'application : 64% des acteurs en 2002 et en 2004.
- L'accroissement du nombre de sociétés spécialisées sur le traitement des données vocales : 20% des acteurs en 2002, 30% en 2004.
- L'augmentation du nombre de fusions, acquisitions et de partenariats stratégiques, industriels, technologies et commerciaux.

En 2004 et pour les années à venir, les applications innovantes qui peuvent pousser la croissance des segments du marché sont notamment:

- Interfaces vocales et communication homme machine (,lancement de commandes sans les mains
- Applications multimodales pour améliorer la qualité des services (système de transport intelligent, robotique et domotique, bornes interactives) : 50% des automobiles américaines sont équipées de reconnaissance vocales en 2007: 15% des américains rédigeant des SMS dans leur voiture...
- Le développement des call centers hébergés sur Internet (e-call centers):
- Le développement d'analyseurs de contenu (emails, Intranet, sites Internet)les applications embarquées avec synthèse vocale de qualité-
- La reconstruction de voix pathologiques à partir d'échantillons de quelques minutes de paroles
- ...

## 6.2. Perspectives de croissance du marché en France.

### 6.2.1. Evolution du marché français à l'horizon 2010

#### 6.2.1.1. Evolution globale du marché français

Calculé à partir de l'évolution du taux de croissance du chiffre d'affaires des 46 acteurs français de l'échantillon, le taux de croissance moyen du marché des technologies de la langue en France pour la période 2005-2010 est de 4.70%.

Le marché des technologies de la langue dans la zone de l'Union européenne à 25 pays est estimé à 507 millions d'euros en 2005 et à 625 millions d'euros en 2010. Ces estimations sont considérées comme des estimations basses.

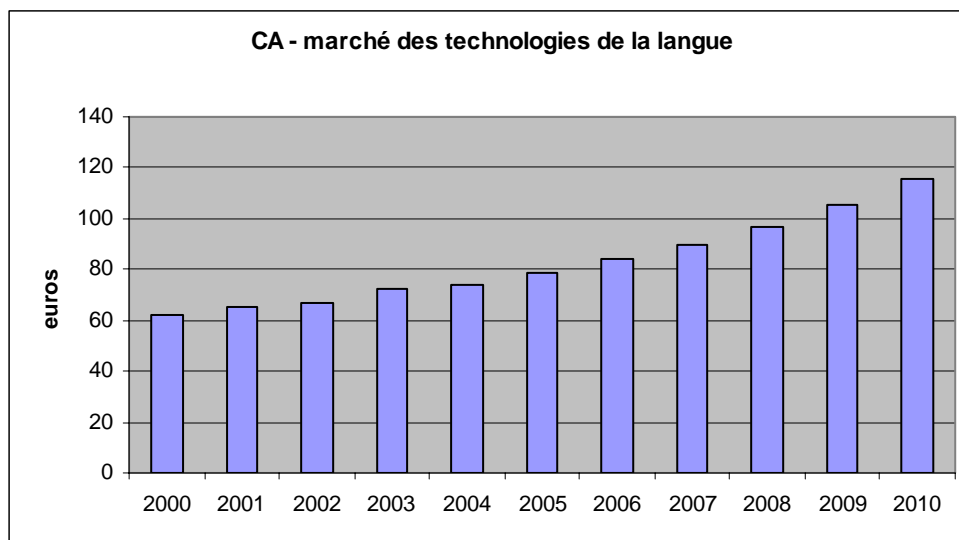
Les pays leaders du marché resteront à leur position relative : Royaume-Uni, France, Allemagne et Italie seront les pays où le nombre d'acteurs et la valeur dégagée seront les plus élevés.

Les taux de croissance des années antérieures ont été calculés à partir de l'échantillon représentatif de 46 sociétés positionnées sur le marché des technologies de la langue :

CA	2000	2001	2002	2003	2004	MOY	VAR2004-2000
SEGMENT 1	8,4	8,3	8,5	9,3	10,3		
TAUX DE CROISSANCE		-1,19%	2,41%	9,41%	10,75%	5,35%	22,62%
SEGMENT 2	15,6	15,9	14,8	15,9	15,3		
TAUX DE CROISSANCE		1,92%	-6,92%	7,43%	-3,77%	-0,33%	-1,92%
SEGMENT 3	3,8	2,9	3,4	4,2	4,3		
TAUX DE CROISSANCE		-23,68%	17,24%	23,53%	2,38%	4,87%	13,16%
SEGMENT 4	5,9	6,9	6,7	7,2	7		
TAUX DE CROISSANCE		16,95%	-2,90%	7,46%	-2,78%	4,68%	18,64%
SEGMENT 5	0,9	0,9	1	1	1,1		
TAUX DE CROISSANCE		0,00%	11,11%	0,00%	10,00%	5,28%	22,22%
SEGMENT 6	0,6	0,7	0,9	1,1	1,1		
TAUX DE CROISSANCE		16,67%	28,57%	22,22%	0,00%	16,87%	83,33%
SEGMENT 7	13,8	14,4	15,3	16,2	16,6		
TAUX DE CROISSANCE		4,35%	6,25%	5,88%	2,47%	4,74%	20,29%
SEGMENT 8	10,8	11,3	11,9	12,8	13		
TAUX DE CROISSANCE		4,63%	5,31%	7,56%	1,56%	4,77%	20,37%
SEGMENT 9	2,1	3,8	4,1	4,8	5,6		
TAUX DE CROISSANCE		80,95%	7,89%	17,07%	16,67%	30,65%	166,67%
TOTAL	61,9	65,1	66,6	72,5	74,3		
TAUX DE CROISSANCE		5,17%	2,30%	8,86%	2,48%	4,70%	20,03%

Les taux de croissance moyens des 9 segments oscillent entre -0.33% et +30.65% sur la période 2000-2010. Nous avons appliqué à chaque segment le taux moyen constaté sur les années antérieures (de 2000 à 2004). Les perspectives dégagées sont des estimations hautes

(évaluation haute du marché : environ 115 millions d'euros en 2010, estimations basses : 100 millions d'euros en 2010).

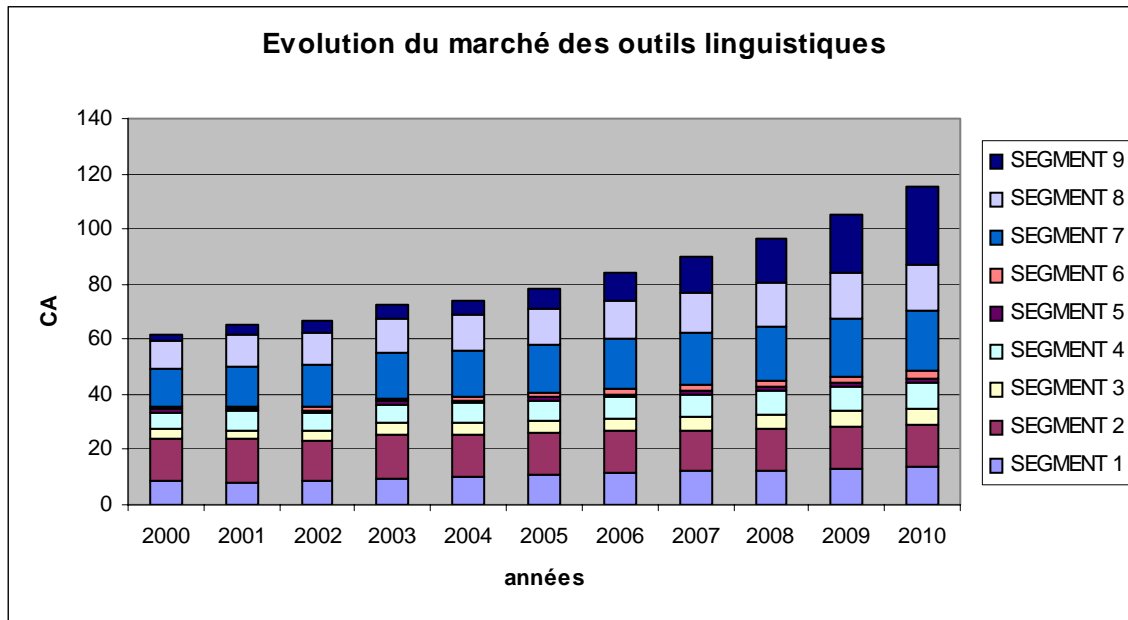


Le tableau suivant détaille les chiffres d'affaires et les taux de croissance estimés du marché français à l'horizon 2010 :

CA	2005	2006	2007	2008	2009	2010
SEGMENT 1	10,85	11,43	12,04	12,69	13,36	14,08
TAUX DE CROISSANCE		5,35%	5,35%	5,35%	5,35%	5,35%
SEGMENT 2	15,25	15,20	15,15	15,10	15,05	15,00
TAUX DE CROISSANCE		-0,33%	-0,33%	-0,33%	-0,33%	-0,33%
SEGMENT 3	4,51	4,73	4,96	5,20	5,45	5,72
TAUX DE CROISSANCE		4,87%	4,87%	4,87%	4,87%	4,87%
SEGMENT 4	7,33	7,67	8,03	8,41	8,80	9,21
TAUX DE CROISSANCE		4,68%	4,68%	4,68%	4,68%	4,68%
SEGMENT 5	1,16	1,22	1,28	1,35	1,42	1,50
TAUX DE CROISSANCE		5,28%	5,28%	5,28%	5,28%	5,28%
SEGMENT 6	1,29	1,50	1,76	2,05	2,40	2,80
TAUX DE CROISSANCE		16,87%	16,87%	16,87%	16,87%	16,87%
SEGMENT 7	17,39	18,21	19,07	19,98	20,92	21,91
TAUX DE CROISSANCE		4,74%	4,74%	4,74%	4,74%	4,74%
SEGMENT 8	13,62	14,27	14,95	15,66	16,41	17,19
TAUX DE CROISSANCE		4,77%	4,77%	4,77%	4,77%	4,77%
SEGMENT 9	7,32	9,56	12,49	16,31	21,31	27,85
TAUX DE CROISSANCE		30,65%	30,65%	30,65%	30,65%	30,65%
TOTAL	78,70	83,79	89,73	96,74	105,13	115,26
TAUX DE CROISSANCE		5,93%	6,46%	7,09%	7,82%	8,67%

### 6.2.1.2. Evolution des segments du marché français

L'évolution des segments est illustrée par le graphique suivant :

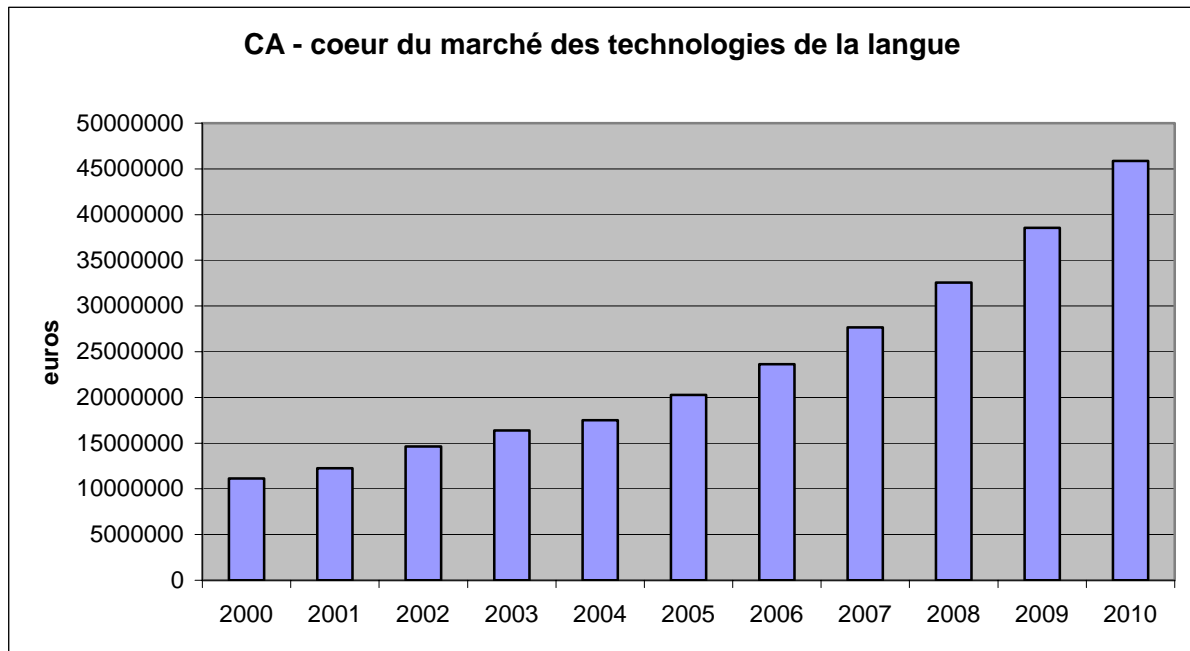


Compte tenu des taux de croissance antérieur, le segments 9 (moteurs de recherche intelligents) représenterait près d'un ¼ de la valeur du marché en 2010 avec chiffre d'affaires estimé de 27,85 millions d'euros.

## 6.2.2. Evolution du cœur du marché français des technologies de la langue à l'horizon 2010

### 6.2.2.1. Evolution globale du cœur du marché français

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 taux de croissance des années antérieures ont été calculés à partir de l'échantillon représentatif de 20 sociétés positionnées dans le noyau dur du marché des technologies de la langue. Compte tenu de l'extrême sensibilité du cœur du marché, les taux de croissance négatifs sur la période 200-2004 concernant le segments 2 (gestion des documents) et le segment 4 (e-business) n'ont pas été retenus dans la mesure où les indicateurs structurels et conjoncturels des NTIC ne déterminent pas spécifiquement une baisse de la valeur du marché de la GED et du e-business pour les années à venir. Dans une position neutre, nous avons retenu un taux de 0% qui a peu de conséquence sur la valeur globale du noyau dur du marché en ce sens où les segments 2 et 4 y sont moins représentés.

Les taux de croissance des années antérieures ont été calculés à partir de l'échantillon représentatif de 20 sociétés positionnées sur le marché des technologies de la langue :

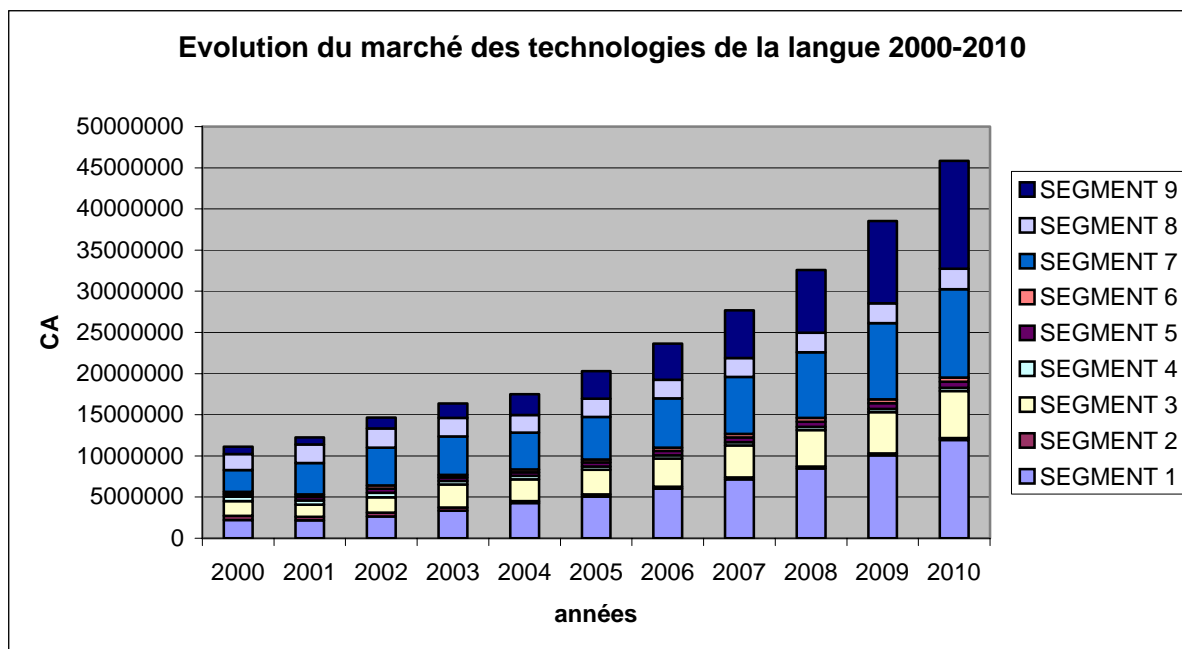
CA	2000	2001	2002	2003	2004	MOY	VAR2004-2000
SEGMENT 1	2198691	2167938	2638852	3325620	4270757		
TAUX DE CROISSANCE		-1,40%	21,72%	26,03%	28,42%	18,69%	94,24%
SEGMENT 2	525 076	454 660	472 582	391 522	246 660		
TAUX DE CROISSANCE		-13,41%	3,94%	-17,15%	-37,00%	-15,91%	-53,02%
SEGMENT 3	1 772 831	1 475 793	1 851 322	2 799 741	2 645 921		
TAUX DE CROISSANCE		-16,76%	25,45%	51,23%	-5,49%	13,61%	49,25%
SEGMENT 4	539 219	498 667	558 180	445 206	405 797		
TAUX DE CROISSANCE		-7,52%	11,93%	-20,24%	-8,85%	-6,17%	-24,74%
SEGMENT 5	339 174	399 927	496 351	421 348	449 072		
TAUX DE CROISSANCE		17,91%	24,11%	-15,11%	6,58%	8,37%	32,40%
SEGMENT 6	265 537	316 593	395 061	324 616	336 321		
TAUX DE CROISSANCE		19,23%	24,79%	-17,83%	3,61%	7,45%	26,66%
SEGMENT 7	2 640 158	3 841 699	4 585 925	4 653 498	4 476 304		
TAUX DE CROISSANCE		45,51%	19,37%	1,47%	-3,81%	15,64%	69,55%
SEGMENT 8	1 942 523	2 227 482	2 341 339	2 238 748	2 130 475		
TAUX DE CROISSANCE		14,67%	5,11%	-4,38%	-4,84%	2,64%	9,68%
SEGMENT 9	912 547	858 514	1 304 175	1 774 648	2 548 813		
TAUX DE CROISSANCE		-5,92%	51,91%	36,07%	43,62%	31,42%	179,31%
TOTAL	11135756	12 241 273	14 643 787	16 374 947	17 510 120		
TAUX DE CROISSANCE		9,93%	19,63%	11,82%	6,93%	12,08%	57,24%

Le tableau suivant détaille les chiffres d'affaires et les taux de croissance estimés du cœur du marché français à l'horizon 2010 :

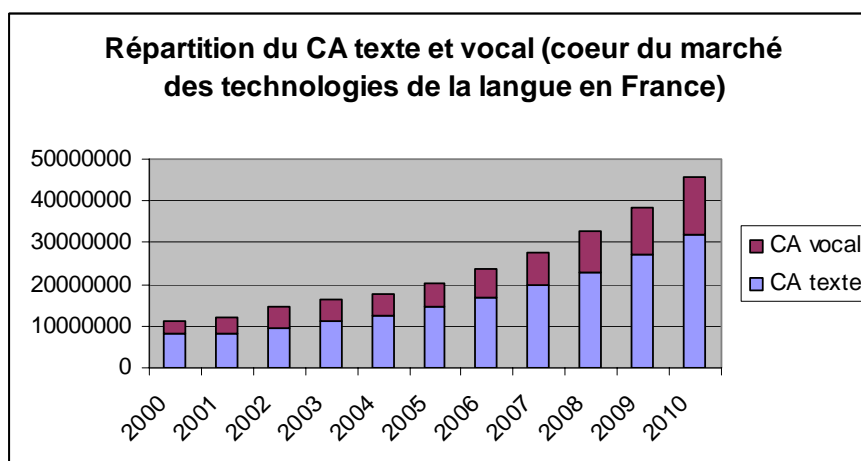
CA	2005	2006	2007	2008	2009	2010
SEGMENT 1	5069048,83	6016557,73	7141175,42	8476007,16	10060346,24	11940830,71
TAUX DE CROISSANCE		18,69%	18,69%	18,69%	18,69%	18,69%
SEGMENT 2	246660,00	246660,00	246660,00	246660,00	246660,00	246660,00
TAUX DE CROISSANCE		0,00%	0,00%	0,00%	0,00%	0,00%
SEGMENT 3	3005938,95	3414942,84	3879597,96	4407476,50	5007180,97	5688484,40
TAUX DE CROISSANCE		13,61%	13,61%	13,61%	13,61%	13,61%
SEGMENT 4	405797,00	405797,00	405797,00	405797,00	405797,00	405797,00
TAUX DE CROISSANCE		0,00%	0,00%	0,00%	0,00%	0,00%
SEGMENT 5	486672,13	527420,45	571580,58	619438,16	671302,79	727509,97
TAUX DE CROISSANCE		8,37%	8,37%	8,37%	8,37%	8,37%
SEGMENT 6	361365,95	388275,94	417189,84	448256,88	481637,41	517503,70
TAUX DE CROISSANCE		7,45%	7,45%	7,45%	7,45%	7,45%
SEGMENT 7	5176266,28	5985682,08	6921666,70	8004011,78	9255603,80	10702908,00
TAUX DE CROISSANCE		15,64%	15,64%	15,64%	15,64%	15,64%
SEGMENT 8	2186735,28	2244481,24	2303752,12	2364588,19	2427030,78	2491122,32
TAUX DE CROISSANCE		2,64%	2,64%	2,64%	2,64%	2,64%
SEGMENT 9	3349698,35	4402237,06	5785503,37	7603418,18	9992556,27	13132406,83
TAUX DE CROISSANCE		31,42%	31,42%	31,42%	31,42%	31,42%
TOTAL	20288182,77	23632054,33	27672922,98	32575653,84	38548115,27	45853222,92
TAUX DE CROISSANCE		<b>15,87%</b>	<b>16,48%</b>	<b>17,10%</b>	<b>17,72%</b>	<b>18,95%</b>

### 6.2.2.2. Evolution des segments du coeur du marché français

L'évolution des segments est illustrée par le graphique suivant :



Compte tenu des taux de croissance antérieur, le segments 9 (moteurs de recherche intelligents) représenterait plus d'un ¼ de la valeur du marché en 2010 avec chiffre d'affaires estimé de 13,1 millions d'euros, suivi de près par le segment 1 (gestion de contenu) estimé à 11,9 millions d'euros en 2010 et le segment 7 (applications vocales et multimodales) estimé à 10,7 millions d'euros.



Le traitement linguistique des données vocales augmente légèrement sur la période avec une valeur dégagée estimée à environ 30% de la valeur globale du coeur du marché en 2010.

Années	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
CA texte	73,47%	65,41%	65,61%	69,32%	72,44%	71,99%	71,53%	71,06%	70,59%	70,12%	69,64%
CA vocal	26,53%	34,59%	34,39%	30,68%	27,56%	28,01%	28,47%	28,94%	29,41%	29,88%	30,36%



### **6.3. Tendances et perspectives du marché en Europe**

#### **6.3.1. Impact des NTIC**

L'évolution du marché des technologies de la langue est étroitement associée à l'évolution du marché des NTIC et plus précisément à celles du marché de l'informatique et des logiciels. L'analyse conjoncturelle du marché des outils linguistiques implique de prendre en compte la dynamique des facteurs suivants :

- la maturité des usages en NTIC
- l'évolution du PIB
- les dépenses en NTIC
- la R&D nationale et communautaire

L'observatoire européen des technologies de l'information (EITO) considère que l'industrie des NTIC est entré dans un nouveau cycle de développement fort et significatif poussé par les facteurs conjoncturels suivants : la convergence numérique des technologies de l'informatique, de l'information et de la communication avec les contenus multimédias et interactifs, l'Internet protocole (IP) universellement reconnu, les nouvelles technologies sans fil, le haut débit, et les applications innovantes du web 2.0. et du web 3.0. favorisant à la contribution éditoriale et sociale des utilisateurs désormais considérés comme des internautes actifs prescripteurs de contenus et de services. La gestion des contenus numériques (données textuelles et vocales) intègre de plus en plus la mobilité des usages et des pratiques des utilisateurs finaux avec la mise en place de services intégrés en réseaux et souvent proposés en web services avec des modèles économiques basés sur une tarification à l'usage.

Selon l'EITO, la croissance du marché des NTIC devrait être de 7% à 12% par an pour les nouveaux pays entrants de l'Union européenne à 25 pays, quant le taux de croissance des pays pionniers de l'Union s'oriente à des taux compris entre 2% et 7% pour la période 2006-2008.

#### **Faiblesses :**

Dans la zone européenne des 25 pays, le marché des NTIC, et plus spécifiquement celui du logiciel et des technologies de la langue, reste fragmenté avec une majorité de petits acteurs. Les investissements européens en R&D restent insuffisants en comparaison des chiffres de l'industrie des USA : l'investissement européen ne correspond en valeur relative qu'à la moitié des investissements réalisés aux USA. Le nombre limité de sociétés « start up » et l'insuffisance du nombre de partenariats et d'alliances internationales sont deux autres facteurs qui illustrent les faiblesses actuelles du marché.

#### **Forces :**

L'appropriation généralisée de l'Internet et la révolution qu'il engendre, couplé aux pratiques de plus en plus élargies du multilinguisme en Europe, est un facteur incontestable du développement du marché des technologies de la langue depuis 2000. La convergence numérique avec le développement des contenus multimédia et les nouvelles applications issues des technologies open source et du web 2.0/ 3.0 pousse les acteurs du marché à développer et industrialiser des applications robustes pour apporter aux sociétés demandeuses des solutions favorisant la connaissance partagée et collaborative au sein de l'entreprise.

## 6.3.2. Perspectives de croissance du marché des outils linguistiques en Europe

### 6.3.2.1. UE 25 pays

A partir du lissage des indicateurs de répartition suivants :

- Répartition du PIB de la zone Europe à 25 pays,
- Répartition de la valeur moyenne dégagée au niveau du marché européen des technologies de la langue par chaque pays,
- Répartition du nombre d'acteurs spécialisés dans le secteur,

Il a été possible de déterminer les taux moyens de répartition du marché des technologies de la langue dans la zone de l'Union européenne à 25 pays pour la période 2005-2010 :

PAYS UE 25	pib	pdmhlt%	acteurs	Répartition CA futur non lissé	Répartition CA futur lissé
<b>Royaume-Uni</b>	<b>17,30%</b>	<b>18,00%</b>	<b>11,00%</b>	<b>17,55%</b>	<b>16,63%</b>
<b>France</b>	<b>15,80%</b>	<b>16,00%</b>	<b>23,00%</b>	<b>16,16%</b>	<b>14,78%</b>
<b>Allemagne</b>	<b>22,00%</b>	<b>13,00%</b>	<b>12,00%</b>	<b>15,87%</b>	<b>12,93%</b>
<b>Italie</b>	<b>13,10%</b>	<b>7,00%</b>	<b>7,00%</b>	<b>8,97%</b>	<b>7,39%</b>
<b>Pays-Bas</b>	<b>4,60%</b>	<b>9,00%</b>	<b>4,00%</b>	<b>7,42%</b>	<b>7,39%</b>
<b>Belgique</b>	<b>2,70%</b>	<b>6,00%</b>	<b>4,00%</b>	<b>4,87%</b>	<b>5,54%</b>
<b>Espagne</b>	<b>7,20%</b>	<b>3,00%</b>	<b>5,00%</b>	<b>4,42%</b>	<b>2,77%</b>
<b>Danemark</b>	<b>1,90%</b>	<b>5,00%</b>	<b>3,00%</b>	<b>3,94%</b>	<b>5,54%</b>
<b>Suède</b>	<b>2,70%</b>	<b>4,00%</b>	<b>4,00%</b>	<b>3,58%</b>	<b>5,54%</b>
<b>Finlande</b>	<b>1,40%</b>	<b>4,00%</b>	<b>4,00%</b>	<b>3,16%</b>	<b>5,54%</b>
<b>Grèce</b>	<b>1,50%</b>	<b>3,00%</b>	<b>3,00%</b>	<b>2,52%</b>	<b>2,77%</b>
<b>Autriche</b>	<b>2,30%</b>	<b>2,00%</b>	<b>2,00%</b>	<b>2,10%</b>	<b>1,85%</b>
<b>Irlande</b>	<b>1,30%</b>	<b>2,00%</b>	<b>2,00%</b>	<b>1,77%</b>	<b>3,70%</b>
<b>République Tchèque</b>	<b>0,80%</b>	<b>2,00%</b>	<b>3,00%</b>	<b>1,65%</b>	<b>1,64%</b>
<b>Lituanie</b>	<b>0,20%</b>	<b>2,00%</b>	<b>4,00%</b>	<b>1,48%</b>	<b>1,47%</b>
<b>Pologne</b>	<b>2,10%</b>	<b>0,30%</b>	<b>0,20%</b>	<b>0,88%</b>	<b>0,87%</b>
<b>Lettonie</b>	<b>0,10%</b>	<b>1,00%</b>	<b>3,00%</b>	<b>0,77%</b>	<b>0,77%</b>
<b>Slovénie</b>	<b>0,20%</b>	<b>1,00%</b>	<b>2,00%</b>	<b>0,77%</b>	<b>0,77%</b>
<b>Estonie</b>	<b>0,10%</b>	<b>1,00%</b>	<b>2,00%</b>	<b>0,74%</b>	<b>0,74%</b>
<b>Hongrie</b>	<b>0,70%</b>	<b>0,30%</b>	<b>1,00%</b>	<b>0,45%</b>	<b>0,45%</b>
<b>Portugal</b>	<b>1,30%</b>	<b>0,00%</b>	<b>0,20%</b>	<b>0,43%</b>	<b>0,43%</b>
<b>Slovaquie</b>	<b>0,30%</b>	<b>0,20%</b>	<b>0,20%</b>	<b>0,23%</b>	<b>0,23%</b>
<b>Chypre</b>	<b>0,10%</b>	<b>0,10%</b>	<b>0,20%</b>	<b>0,10%</b>	<b>0,10%</b>
<b>Malte</b>	<b>0,10%</b>	<b>0,10%</b>	<b>0,20%</b>	<b>0,10%</b>	<b>0,10%</b>
<b>Luxembourg</b>	<b>0,20%</b>	<b>0,00%</b>	<b>0,00%</b>	<b>0,06%</b>	<b>0,06%</b>
<b>TOTAL</b>	<b>100,00%</b>	<b>100,00%</b>	<b>100,00%</b>	<b>100,00%</b>	<b>100,00%</b>

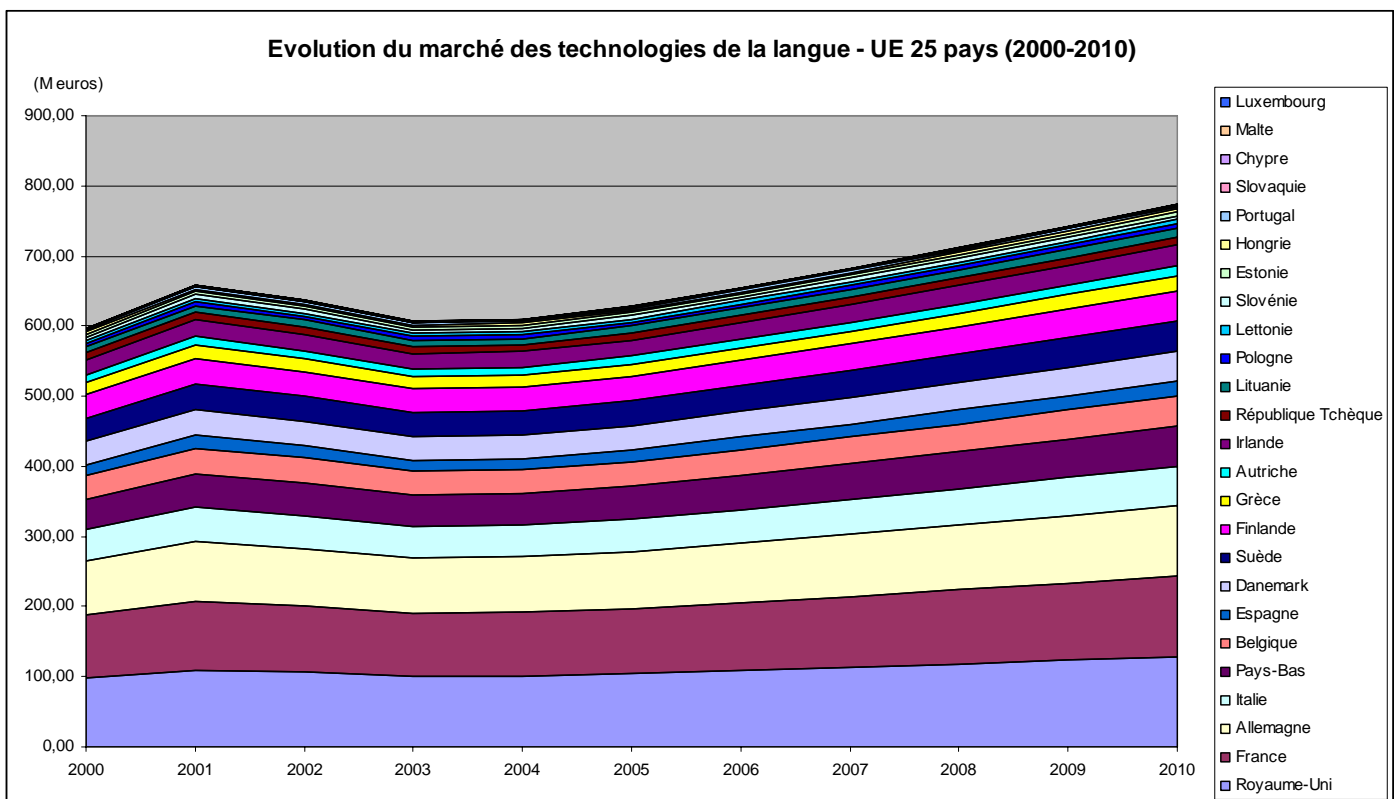
Le taux de croissance moyen du marché des technologies de la langue en Europe (25 pays) qui a été retenu pour la période 2005-2010 est basé sur la moyenne européenne constatée pour la période 2000-2004 avec un lissage pour la prise en compte de la dynamique de croissance du secteur des NTIC et du marché du logiciel dans la zone européenne à 25 pays. Le taux de croissance moyen retenu est de 4.27%. L'analyse du taux de croissance 2005 calculé à partir de l'analyse des chiffres d'affaires 2005 communiqués par les acteurs du marché préconisait un taux de croissance de 2,95% pour 2005 qui a été conservé. Les chiffres présentés ci-dessous sont à considérer comme des estimations lissées sur la période 2000-2010. Ils permettent principalement de dégager les tendances de croissance du secteur et de rendre visible les principaux acteurs du marché sur la période.

PAYS UE 25 PAYS	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
Royaume-Uni	99,30	109,64	105,90	100,92	101,44	104,43	108,89	113,55	118,39	123,44	128,72
France	88,27	97,46	94,14	89,70	90,17	92,83	96,79	100,93	105,23	109,72	114,41
Allemagne	77,23	85,28	82,37	78,49	78,90	81,22	84,69	88,31	92,08	96,01	100,11
Italie	38,62	42,64	41,18	39,25	39,45	40,61	42,35	44,16	46,04	48,00	50,06
Pays-Bas	44,13	48,73	47,07	44,85	45,08	46,41	48,40	50,47	52,62	54,86	57,21
Belgique	33,10	36,55	35,30	33,64	33,81	34,81	36,30	37,85	39,46	41,15	42,91
Espagne	16,55	18,27	17,65	16,82	16,91	17,41	18,15	18,92	19,73	20,57	21,45
Danemark	33,10	36,55	35,30	33,64	33,81	34,81	36,30	37,85	39,46	41,15	42,91
Suède	33,10	36,55	35,30	33,64	33,81	34,81	36,30	37,85	39,46	41,15	42,91
Finlande	38,62	42,64	41,18	39,25	39,45	40,61	42,35	44,16	46,04	48,00	50,06
Grèce	16,55	18,27	17,65	16,82	16,91	17,41	18,15	18,92	19,73	20,57	21,45
Autriche	11,03	12,18	11,77	11,21	11,27	11,60	12,10	12,62	13,15	13,72	14,30
Irlande	22,07	24,37	23,53	22,43	22,54	23,21	24,20	25,23	26,31	27,43	28,60
République Tchèque	9,77	10,75	10,48	9,99	10,00	10,30	10,73	11,20	11,67	12,17	12,70
Lituanie	8,75	9,64	9,40	8,95	8,97	9,24	9,64	10,05	10,47	10,92	11,39
Pologne	5,21	5,73	5,59	5,32	5,34	5,50	5,73	5,97	6,23	6,49	6,77
Lettonie	4,56	5,01	4,90	4,65	4,66	4,80	5,01	5,22	5,45	5,68	5,92
Slovénie	4,56	5,01	4,90	4,65	4,66	4,80	5,01	5,22	5,45	5,68	5,92
Estonie	4,38	4,82	4,70	4,48	4,49	4,62	4,81	5,02	5,24	5,46	5,69
Hongrie	2,67	2,93	2,86	2,73	2,73	2,81	2,93	3,05	3,19	3,31	3,46
Portugal	2,54	2,80	2,73	2,60	2,61	2,69	2,80	2,92	3,04	3,18	3,30
Slovaquie	1,36	1,50	1,46	1,39	1,39	1,44	1,50	1,57	1,62	1,69	1,77
Chypre	0,59	0,65	0,63	0,61	0,61	0,63	0,65	0,67	0,71	0,74	0,77
Malte	0,59	0,65	0,63	0,61	0,61	0,63	0,65	0,67	0,71	0,74	0,77
Luxembourg	0,35	0,39	0,38	0,37	0,37	0,38	0,39	0,41	0,42	0,45	0,46
TOTAL UE 25	597,00	659,00	637,00	607,00	610,00	628,00	654,80	682,80	711,90	742,30	774,00
Taux de croissance		10,39%	-3,34%	-4,71%	0,49%	2,95%	4,27%	4,27%	4,27%	4,27%	4,27%

Le marché des technologies de la langue dans la zone de l'Union européenne à 25 pays est estimé à 628 millions d'euros en 2005 et à 774 millions d'euros en 2010. Compte tenu des taux de croissance moyens relativement faibles, ces estimations sont considérées comme des estimations basses.

Le correcteur de 15% (pour la prise en compte des acteurs non identifiés au niveau de la zone de l'Union européenne à 25 pays) a été réparti à l'ensemble des 25 pays européens étudiés. Les pays leaders du marché resteront à leur position relative : Royaume-Uni, France, Allemagne et Italie seront les pays où le nombre d'acteurs et la valeur dégagée seront les plus élevés.

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.



Basée sur la même méthodologie que pour l'analyse de l'Union européenne à 25 pays, les taux de répartition nationaux ont été établis à partir du lissage des indicateurs de répartition suivants :

- Répartition du PIB de la zone Europe à 25 pays,
- Répartition de la valeur moyenne dégagée au niveau du marché européen des technologies de la langue par chaque pays,
- Répartition du nombre d'acteurs spécialisés dans le secteur,

Les taux moyens de répartition du marché des technologies de la langue dans la zone de l'Union européenne à 13 pays pour la période 2005-2010 sont :

PAYS UE 13	pib	pdmhlt%	acteurs	Répartition CA futur non lissé	Répartition CA futur lissé
<b>Royaume-Uni</b>	<b>18,44%</b>	<b>19,57%</b>	<b>13,10%</b>	<b>19,01%</b>	<b>18%</b>
<b>France</b>	<b>16,84%</b>	<b>17,39%</b>	<b>27,38%</b>	<b>17,50%</b>	<b>16%</b>
<b>Allemagne</b>	<b>23,45%</b>	<b>14,13%</b>	<b>14,29%</b>	<b>17,19%</b>	<b>14%</b>
<b>Italie</b>	<b>13,97%</b>	<b>7,61%</b>	<b>8,33%</b>	<b>9,72%</b>	<b>8%</b>
<b>Pays-Bas</b>	<b>4,90%</b>	<b>9,78%</b>	<b>4,76%</b>	<b>8,04%</b>	<b>8%</b>
<b>Belgique</b>	<b>2,88%</b>	<b>6,52%</b>	<b>4,76%</b>	<b>5,27%</b>	<b>6%</b>
<b>Espagne</b>	<b>7,68%</b>	<b>3,26%</b>	<b>5,95%</b>	<b>4,79%</b>	<b>6%</b>
<b>Danemark</b>	<b>2,03%</b>	<b>5,43%</b>	<b>3,57%</b>	<b>4,27%</b>	<b>6%</b>
<b>Suède</b>	<b>2,88%</b>	<b>4,35%</b>	<b>4,76%</b>	<b>3,88%</b>	<b>6%</b>
<b>Finlande</b>	<b>1,49%</b>	<b>4,35%</b>	<b>4,76%</b>	<b>3,42%</b>	<b>4%</b>
<b>Grèce</b>	<b>1,60%</b>	<b>3,26%</b>	<b>3,57%</b>	<b>2,73%</b>	<b>3%</b>
<b>Autriche</b>	<b>2,45%</b>	<b>2,17%</b>	<b>2,38%</b>	<b>2,27%</b>	<b>3%</b>
<b>Irlande</b>	<b>1,39%</b>	<b>2,17%</b>	<b>2,38%</b>	<b>1,92%</b>	<b>2%</b>
Total UE 13	100,00%	100,00%	100,00%	100,00%	100,00%

Dans la mesure où 80% de la valeur dégagée par le secteur est assuré par les acteurs issus de la zone de l'Union européenne à 13 pays, le taux de croissance moyen du marché des technologies de la langue en Europe (13 pays) qui a été retenu pour la période 2005-2010 est de 4.27%. L'analyse du taux de croissance 2005 calculé à partir de l'analyse des chiffres d'affaires 2005 communiqués par les acteurs du marché préconisait un taux de croissance de 2,95% pour 2005 qui a été conservé.

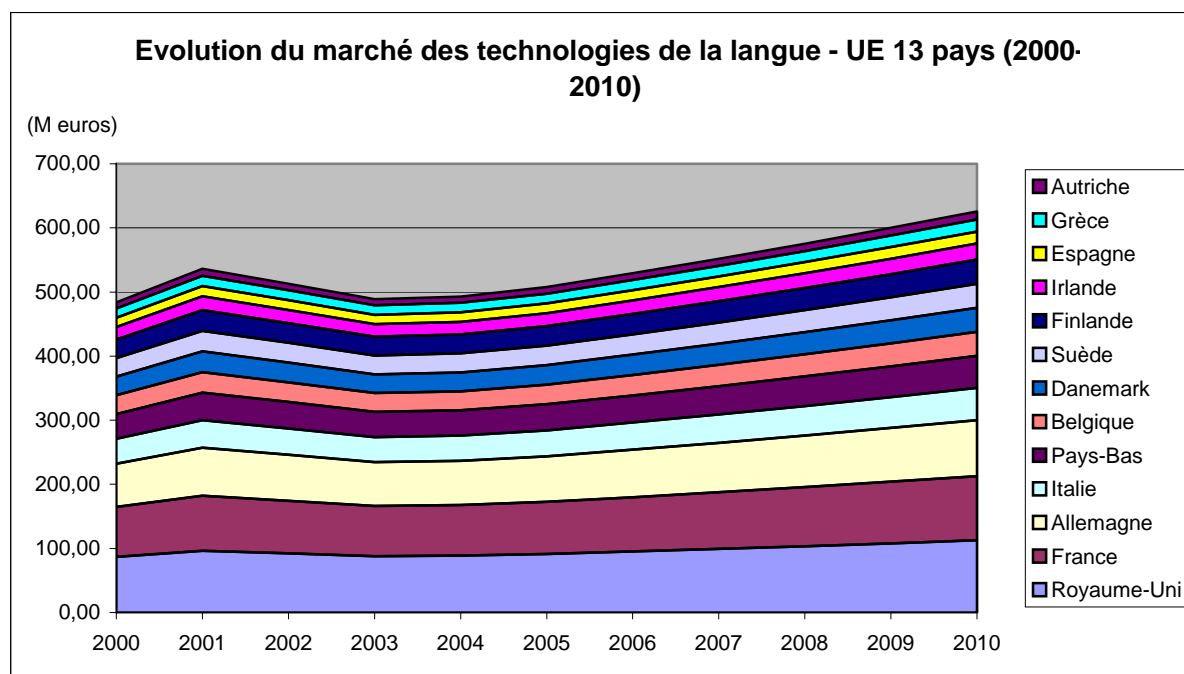
**Le marché des technologies de la langue dans la zone de l'Union européenne à 13 pays est estimé à 507 millions d'euros en 2005 et à 625 millions d'euros en 2010. Ces estimations sont considérées comme des estimations basses.**

Les pays leaders du marché resteront à leur position relative : Royaume-Uni, France, et Allemagne seront les pays où le nombre d'acteurs et la valeur dégagée seront les plus élevés.

Le tableau ci-après détaille l'évolution du marché par pays jusqu'à 2010.

PAYS UE 13	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
<b>Royaume-Uni</b>	87,12	96,48	92,34	88,02	88,74	91,36	95,26	99,33	103,57	107,99	112,60
<b>France</b>	77,44	85,76	82,08	78,24	78,88	81,21	84,68	88,29	92,06	95,99	100,09
<b>Allemagne</b>	67,76	75,04	71,82	68,46	69,02	71,06	74,09	77,25	80,55	83,99	87,58
<b>Italie</b>	38,72	42,88	41,04	39,12	39,44	40,60	42,34	44,14	46,03	48,00	50,04
<b>Pays-Bas</b>	38,72	42,88	41,04	39,12	39,44	40,60	42,34	44,14	46,03	48,00	50,04
<b>Belgique</b>	29,04	32,16	30,78	29,34	29,58	30,45	31,75	33,11	34,52	36,00	37,53
<b>Danemark</b>	29,04	32,16	30,78	29,34	29,58	30,45	31,75	33,11	34,52	36,00	37,53
<b>Suède</b>	29,04	32,16	30,78	29,34	29,58	30,45	31,75	33,11	34,52	36,00	37,53
<b>Finlande</b>	29,04	32,16	30,78	29,34	29,58	30,45	31,75	33,11	34,52	36,00	37,53
<b>Irlande</b>	19,36	21,44	20,52	19,56	19,72	20,30	21,17	22,07	23,02	24,00	25,02
<b>Espagne</b>	14,52	16,08	15,39	14,67	14,79	15,23	15,88	16,55	17,26	18,00	18,77
<b>Grèce</b>	14,52	16,08	15,39	14,67	14,79	15,23	15,88	16,55	17,26	18,00	18,77
<b>Autriche</b>	9,68	10,72	10,26	9,78	9,86	10,15	10,58	11,04	11,51	12,00	12,51
<b>Total UE 13</b>	484	536	513	489	493	507,54	529,22	551,81	575,38	599,94	625,56
<b>Taux de croissance</b>		0,11%	-0,04%	-0,05%	0,01%	2,95%	4,27%	4,27%	4,27%	4,27%	4,27%

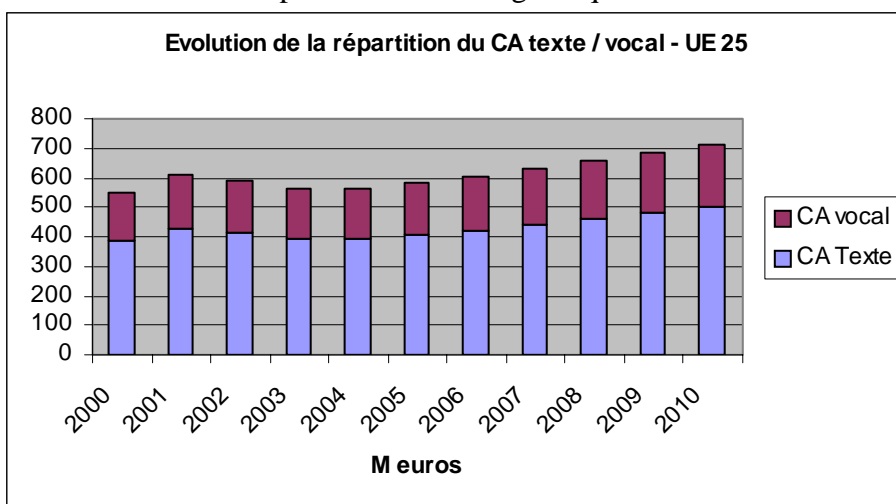
La courbe de tendance illustrée par le graphique ci dessous souligne la reprise du marché et le retour au niveau de 2001 entre 2005 et 2007. Les années 2008-2010 s'annoncent comme des années à fort potentiel de croissance et possiblement accompagnées d'une rentabilité économique positive.



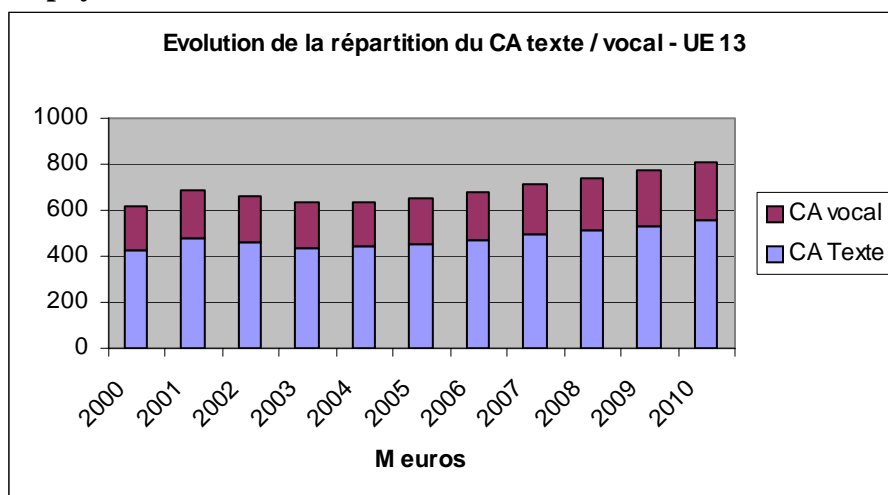
### 6.3.3. Evolution Texte / Vocal

#### 6.3.3.1. UE 25 pays

A partir de la répartition des acteurs issus de l'échantillon et de la répartition des chiffres d'affaires communiqués, la ventilation de la valeur entre les revenus issus du traitement des données textuelles et des données vocales évoluent dans un rapport stable dans le temps. Ici représenté par un ratio 1/3 / 2/3, les revenus issus du traitement des données vocales devraient prendre encore plus d'importance. Parfois surestimée par les analystes du secteur, la part du vocal pourrait représenter jusqu'à 2/3 du marché. Les chiffres ci-dessous en sont encore très éloignés car la dynamique de pénétration du vocal a pu être moins forte que le prévoaient les études prospectives le plus souvent anglo-saxonnes et donc aussi plus structurellement mature en implantation de solutions vocales innovantes (reconnaissance et synthèse vocale intégrées aux solutions de eCRM vocal, eCall centers, analyse des formulaires textuels et des messages vocaux numérisés). De plus les revenus issus du traitement des données vocales sont souvent associés aux activités des leaders du secteur des télécommunications et n'ont pas été intégrés dans l'évaluation du marché européen des outils linguistiques.



#### 6.3.3.2. UE 13 pays



## 7.0. Conclusions

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 de connaissance dans un domaine hautement stratégique.

La baisse du marché peut être expliquée par l'effet de l'éclatement de la bulle Internet 2001 et par des facteurs directement liés à l'évolution du marché des outils linguistiques :

- baisse du prix des solutions technologiques
- baisse du temps de mise en place des solutions
- baisse du CA des acteurs du fait de la concurrence accrue liée à l'augmentation du nombre d'acteurs nouvellement positionnés sur le marché
- la demande en ressources et outils linguistiques est fortement corrélée aux pratiques numériques des entreprises : la mise en place de solutions innovantes intégrant notamment le traitement linguistique des données textuelles et vocales se justifie pour répondre aux enjeux de la gestion des contenus numériques et à leur valorisation.

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 d'évolution des technologies du fait de la généralisation du multimédia, 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. 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*). La comparaison d'une année à l'autre est donc délicate.



Le retour à la croissance du marché semble se vérifier depuis 2004 et la rentabilité économique marque un retour à l'équilibre pour plusieurs sociétés en 2005. Les années 2006-2010 pourraient représenter le cercle vertueux de croissance attendus par les différents protagonistes du secteur.

Les années 2000 marquent l'installation de la société de la communication avec le traitement numérique du texte, de l'image, de la voix, de la vidéo en phase industrielle. La convergence des technologies de l'informatique, des télécommunications et audiovisuelles est aujourd'hui fédérée par la numérisation. L'ère du multimédia se concrétise avec le déploiement de kiosques multimédias, des téléphones mobiles de nouvelle génération type « photophones », des consoles et des terminaux de divertissement, de la domotique. Les connexions réseaux haut débit se généralisent. Les applications portent sur 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.

Force est de constater que la dichotomie existant entre les données textuelles et vocales s'amenuise voire disparaît dans la mesure où le contenu de l'information étant devenu numérique, il s'affranchit de son support : dématérialisé par la numérisation, un fichier digital textuel peut facilement être converti en fichier audio et vice versa. L'opposition Texte / Vocal deviendrait alors caduque. Ainsi du fait de la numérisation, on ne traite que du numérique

Le positionnement des acteurs vers la voix devrait s'orienter vers un équilibre autour de 50% : certains acteurs resteront historiquement positionnés sur le traitement des données vocales car ils ont développé des approches spécifiques pour son traitement.

Un cercle vertueux de la croissance du marché des technologies de la langue semble possible en ce sens où les usages numériques peuvent pousser l'offre et la demande grâce aux couples gagnants des contenus numériques suivants :

- Texte → Voix → Texte
- Images → Texte
- Voix → Texte
- Vidéo → Voix, Images, Texte

L'analyse des SMS peut être traitée par conversion au format vocal des données textuelles le plus souvent tapées phonétiquement. L'analyse des messages vocaux voire vidéos sera bientôt une réalité intégrée aux solutions de CRM pour permettre de prendre en compte les différents canaux relationnels vers sa clientèle (email, courriers, téléphone, site Internet, vidéo...).

L'analyse des données images peut elle aussi drainer des traitements textuels (recherche en langage naturel sur les légendes et les descripteurs, analyse contextuelle des mots présents autour de l'image si éditée en page web). La recherche de contenu visuel par similarité est aussi couplée à l'analyse des mots et des textes rattachés aux images numériques.

De plus, la standardisation multiformat poussée par le XML et le web 3.0 devrait accélérer ces dynamiques et permettre de capitaliser les possibles technologiques d'un univers NTIC de plus en plus ouvert et multipoint.

Les usages des NTIC et les technologies linguistiques semblent devoir se déployer plus particulièrement dans les secteurs suivants :

- Santé : domaine médical ;
- Marketing : relation client et e-call centers ;
- Transport, Tourisme, Industrie : Robotique/Domotique, interopérabilité des systèmes ;

- Information Scientifique et Technique : Open access, ontologies ;
- Les informations géolocalisées et les SIG : interfaces intelligentes et intuitives pour rechercher des informations textuelles ou images avec géolocalisation (plans de ville, cartes, informations locales, statistiques,...) ;
- Les portails collaboratifs orientés vers le web 2.0. et le web 3.0 pour la mise en place de dispositifs intégrant la participation active des utilisateurs (éditoriale, sociale, prescripteur...) : des interfaces type wikipedia se créent pour la réalisation de vocabulaires métiers élaborés par des réseaux d'experts et des spécialistes des domaines (ontologies, thésaurus, dictionnaires, lexiques métiers intégrés aux systèmes d'information, aux bases de connaissance, aux catalogues e-business...).