### Université de Montréal

Identification automatique du vocabulaire caractéristique du domaine de l'informatique fondée sur la comparaison de corpus

# par Chantal Lemay

Département de linguistique et de traduction Faculté des arts et des sciences

Mémoire présenté à la Faculté des études supérieures en vue de l'obtention du grade de Maître en traduction

Août 2003

© Chantal Lemay, 2003

# Université de Montréal Faculté des études supérieures

### Ce mémoire intitulé :

Identification automatique du vocabulaire caractéristique du domaine de l'informatique fondée sur la comparaison de corpus

présenté par :

**Chantal Lemay** 

a été évalué par un jury composé des personnes suivantes :

Nathan Ménard président-rapporteur

Marie-Claude L'Homme directrice de recherche

Patrick Drouin membre du jury

#### **RÉSUMÉ**

Notre travail s'inscrit dans le cadre de projets relevant du domaine du traitement automatique des langues, plus précisément de la terminologie computationnelle. Nous présentons deux méthodes visant à isoler de façon automatique le vocabulaire caractéristique du domaine de l'informatique. Ce vocabulaire englobe un vocabulaire fonctionnel et un vocabulaire spécialisé.

La plupart des travaux portant sur l'extraction d'unités lexicales à partir de corpus s'attachent à extraire la terminologie, plus spécifiquement les termes complexes du corpus. Notre travail a ceci de particulier qu'il s'intéresse à l'extraction d'unités lexicales simples, lesquelles ne sont pas nécessairement des termes, mais des unités pouvant appartenir à toutes les parties du discours.

Afin d'isoler notre vocabulaire, deux approches sont mises à contribution. Dans la première, nous opposons un corpus de référence, soit le corpus journalistique *Le Monde*, et un corpus d'analyse, soit un corpus composé de textes traitant de l'informatique. Dans la deuxième approche, nous opposons un corpus de référence, soit un corpus de textes sur l'informatique, et six sous-corpus d'analyse, chacun étant un sous-ensemble du corpus de référence et chacun représentant un sous-domaine de l'informatique.

La comparaison des corpus repose sur la technique des spécificités et sur l'application de modèles statistiques, lesquels sont intégrés à un logiciel d'acquisition automatique des termes : *TermoStat*.

Dans le but de valider le vocabulaire issu de nos deux méthodes de comparaison de corpus, nous avons recours à deux dictionnaires spécialisés de l'informatique.

Les résultats seront principalement appliqués à la construction d'un dictionnaire de base de l'informatique. Par ailleurs, ce travail nous permettra de proposer de nouvelles perspectives pour des outils d'extraction terminologique.

**Mots clés :** terminologie, terminologie computationnelle, comparaison de corpus, vocabulaire caractéristique, vocabulaire de base, spécificités, spécificités positives, statistique textuelle, langue de spécialité

#### **ABSTRACT**

This work lies within the scope of projects coming under the field of *natural language processing*, more specifically of *computational terminology*. In this work, two methods are used to automatically isolate the core vocabulary of the field of computer science. This vocabulary includes a functional vocabulary and a specialized vocabulary.

Most researchers who have addressed problems relating to information retrieval from corpora have attempted to extract the terminology, especially complex terms of a corpus. In this work, the extraction will be performed on simple lexical items, which are not inevitably terms, but units that might belong to any part of speech.

Two approaches contribute to isolate the vocabulary. In the first approach are opposed a reference corpus, namely, the journalistic corpus *Le Monde*, and an analysis corpus, here, a corpus made up of texts dealing with computer science. In the second approach are opposed a reference corpus, in this case, a corpus of texts related to computer science, and six analysis sub-corpora, each one being a subset of the reference corpus and each representing a sub-domain of computer science.

The comparison of corpora rests on the application of statistical techniques, which are integrated into an automatic term extraction software: *TermoStat*.

In order to validate the vocabulary resulting from our two methods of corpora comparison, we used two dictionaries specialized in computer science.

The results will be mainly applied to construct a dictionary of computer science. In addition, this work proposes new perspectives for terminological extraction tools.

**Keywords:** terminology, computational terminology, comparison of corpora, core vocabulary, specific vocabulary, specific words, quantitative linguistics, language for special purposes

# TABLE DES MATIÈRES

RÉSUMÉ	iii
ABSTRACT	v
TABLE DES MATIÈRES	vii
LISTE DES TABLEAUX	x
LISTE DES FIGURES	xii
LISTE DES SIGLES ET ABRÉVIATIONS	xiii
REMERCIEMENTS	xv
INTRODUCTION	1
CHAPITRE 1 : ÉTAT DE LA QUESTION	6
1.1 Stratification lexicale	7
1.1.1 Vocabulaire fondamental	7
1.1.2 Vocabulaire scientifique/technique général	10
1.1.2.1 Vocabulaire général d'orientation scientifique	10
1.1.2.2 Vocabulaire général d'orientation technique	12
1.1.2.3 Vocabulaire général d'enseignement scientifique	13
1.1.3 Langue de spécialité	13
1.2 Comparaison de corpus	16
1.2.1 Linguistique de corpus	17
1.2.2 Statistique textuelle	18
1.2.2.1 Muller (1977, 1979)	20
1.2.2.2 Lafon (1980)	21
1.2.2.3 Lebart et Salem (1994)	22
1.2.2.4 Nelson (2000)	22
1.2.2.5 Drouin (2002)	24
1.3 Logiciels d'analyse lexicale de corpus textuels	25
1.3.1 TermoStat – Patrick Drouin	25
1.3.2 LEXICO – André Salem	26
1.3.3 Hyperbase – Étienne Brunet	27
1.3.4 TALTAC – Sergio Bolasko	28

1.3.5 WordSmith Tools – Mike Scott	29
1.4 Conclusion	31
CHAPITRE 2 : MÉTHODOLOGIE	33
2.1 Présentation de la méthode 1	33
2.2 Constitution des corpus de M1	35
2.2.1 Corpus de référence	35
2.2.2 Corpus d'analyse	36
2.2.2.1 Provenance des textes	37
2.2.2.2 Choix des textes	38
2.2.2.3 Taille du corpus	38
2.3 Présentation de la méthode 2	39
2.4 Constitution des corpus de M2	39
2.4.1 Corpus de référence	39
2.4.2 Corpus d'analyse	40
2.4.2.1 Sous-domaines	40
2.5 Étiquetage morphosyntaxique	42
2.5.1 Format des textes	42
2.5.2 Nettoyage des textes du corpus informatique	43
2.5.3 Nettoyage des textes du corpus Le Monde	43
2.5.3.1 Prétraitement	44
2.5.3.1.1 Limites du prétraitement	46
2.5.4 L'étiqueteur WinBrill	47
2.5.4.1 Lemmatisation	50
2.6 Constitution des listes de référence	52
2.6.1 Regroupement d'étiquettes	54
2.7 Calcul des spécificités à l'aide de TermoStat	54
2.7.1 Loi normale	56
2.7.2 Valeurs-tests	57
2.8 Gestion des spécificités	57
2.8.1 Constitution des bases de données	57
2.8.2 Critères d'acceptation	60

2.9 Rés	sultats du calcul des spécificités	62
2.9.1	Spécificités positives ignorées	63
2.9.2	Spécificités positives retenues	64
2.9.3	Résumé des résultats obtenus suite au calcul des spécificités	65
CHAPITR	E 3 : ANALYSE ET VALIDATION DES RÉSULTATS	66
3.1 Ana	alyse des S+ obtenues avec M1	66
3.1.1	Vocabulaire spécifique de l'informatique selon M1	66
3.1.2	Répartition des catégories grammaticales	68
3.2 Ana	alyse des S+ obtenues avec M2	70
3.2.1	Vocabulaire spécifique aux sous-domaines de l'informatique	70
3.2.2	Répartition des catégories grammaticales	81
3.2.3	Vocabulaire spécifique de l'informatique selon M2	84
3.3 Val	idation des S+	86
3.3.1	Résultats de la validation – M1	89
3.3.2	Résultats de la validation – M2	93
3.3.3	Comparaison de M1 et M2 quant aux résultats de la validation	94
CONCLUS	SION	96
BIBLIOGE	RAPHIE	103
ANNEXE	1 – Liste des étiquettes	.xvi
ANNEXE	2 – Regroupement des étiquettes	xvii
ANNEXE	3 – Liste des S+ obtenues à l'aide de M1	viii
ANNEXE	4 – Liste des S+ obtenues à l'aide de M2xx	xvi
ANNEXE	5 – Vocabulaire répertorié dans <i>Ginguay</i> (lettres A, C et P)	lvi
ANNEXE	6 – Vocabulaire répertorié dans <i>Collin</i> (lettres A, C et P)	lxvi

# LISTE DES TABLEAUX

Tableau I. Taille des corpus d'analyse de M2	41
Tableau II. Structure des mots analysés après lemmatisation	52
Tableau III. Valeurs possibles des différents traits dans les textes lemmatisés	52
Tableau IV. Information générée par TermoStat	55
Tableau V. Nombre de formes comprises dans M1 et M2 avant nettoyage	62
Tableau VI. Nombre de spécificités positives à ignorer	63
Tableau VII. Exemples de formes ignorées dans l'analyse des résultats	64
Tableau VIII. Nombre de spécificités positives à analyser	64
Tableau IX. Résumé des résultats obtenus suite au calcul des spécificités	65
Tableau X. Liste des 25 premières S+ obtenues à l'aide de M1	67
Tableau XI. Répartition des catégories grammaticales – M1	69
Tableau XII. Liste des 25 premières S+ obtenues pour SC1	72
Tableau XIII. Liste des 25 premières S+ obtenues pour SC2	73
Tableau XIV. Liste des 25 premières S+ obtenues pour SC3	75
Tableau XV. Liste des 25 premières S+ obtenues pour SC4	76
Tableau XVI. Liste des 25 premières S+ obtenues pour SC5	77
Tableau XVII. Liste des 25 premières S+ obtenues pour SC6	79
Tableau XVIII. Répartition des catégories grammaticales – M2	82
Tableau XIX. Liste des 25 premières S+ obtenues pour l'ensemble des SC	85
Tableau XX. Résultats de la validation – M1	89
Tableau XXI. Répartition des catégories grammaticales selon la source	92
Tableau XXII. Résultats de la validation ignorant la catégorie <i>Autres</i>	93

Tableau XXIII.	Résultats de la validation – M2	.94
Tableau XXIV.	Comparaison des résultats de la validation entre M1 et M2	.95

# LISTE DES FIGURES

Figure 1.	Langues de spécialité et langue commune	15
Figure 2.	Illustration des spécificités de M1	34
Figure 3.	Illustration des spécificités positives de M2	41
Figure 4.	Exemple de regroupement d'étiquettes	54
Figure 5.	Exemple d'une requête de la base de données M2	60
Figure 6.	Échantillon du fichier de validation – M1	89

## LISTE DES SIGLES ET ABRÉVIATIONS

BEC Business English Corpus

BNC British National Corpus

CA corpus d'analyse

CR corpus de référence

F fréquence absolue

FF français fondamental

LSP langue de spécialité

M1 première méthode de comparaison

M2 seconde méthode de comparaison

PMC Published Materials Corpus

SC1 (sous-) corpus d'analyse Abc de l'informatique

SC2 (sous-) corpus d'analyse *Internet* 

SC3 (sous-) corpus d'analyse *Logiciels* 

SC4 (sous-) corpus d'analyse *Matériel* 

SC5 (sous-) corpus d'analyse *Programmation et réseaux* 

SC6 (sous-) corpus d'analyse Systèmes d'exploitation

S+ spécificité positive

VGES vocabulaire général d'enseignement scientifique

VGOS vocabulaire général d'orientation scientifique

VGOT vocabulaire général d'orientation technique

VT valeur-test

À Patrick

#### REMERCIEMENTS

À tout seigneur, tout honneur. Mes sincères remerciements sont adressés, tout d'abord, à ma directrice de recherche, Marie-Claude L'Homme, pour ses conseils, son encouragement, sa grande disponibilité, et toutes les discussions qui ont mené à bien ce travail. Je la remercie également pour son optimisme et sa bonne humeur qui m'ont encouragée et redonné confiance dans les moments de doute.

Ce travail n'aurait pu devenir ce qu'il est devenu sans l'aide de Patrick Drouin, professeur au département de linguistique et de traduction. Je lui dois beaucoup et lui suis particulièrement reconnaissante pour ses précieux conseils et son indispensable soutien technique.

Merci à mes fidèles amies Marie-Josée et Linda de m'avoir prêté une oreille bienveillante dans les hauts et les bas de ma deuxième vie estudiantine.

Que Michel et Cécile, pour l'inébranlable confiance qu'ils ont en leur fille, trouvent ici un témoignage de sa sincère reconnaissance.

Enfin et surtout, pour ses encouragements, son soutien financier et moral, son écoute, sa présence et sa patience de tous les instants, mes chaleureux remerciements vont à Patrick, qui n'avait probablement pas une idée très juste du nombre d'heures que j'allais invertir dans cette aventure.

#### INTRODUCTION

Notre recherche se situe en traitement automatique des langues, ou plus précisément en terminologie computationnelle, dans un contexte d'acquisition automatique d'informations lexicales. Dans ce travail, nous mettons de l'avant deux méthodes d'acquisition du vocabulaire caractéristique de l'informatique, fondées sur la technique des spécificités.

Les recherches menées en acquisition automatique se sont beaucoup intéressées à l'extraction de terminologie. Le but commun de ces travaux était d'obtenir, à partir d'un corpus, une liste de termes (dans la plupart des cas les unités nominales complexes) propres à un domaine.

Notre travail diffère de ces recherches en ce sens qu'il ne s'intéresse pas seulement aux formes nominales ou à la terminologie d'un domaine mais à toutes les parties du discours se démarquant dans un domaine spécialisé. Nous cherchons à isoler le vocabulaire caractéristique du domaine de l'informatique dans un but non seulement terminologique, mais également « pédagogique ». L'élaboration de notre vocabulaire vise sensiblement le même objectif que celui du *vocabulaire général d'orientation scientifique* (VGOS) de Phal, lequel avait pour but principal de permettre aux apprenants d'une langue ou d'une langue propre à un domaine des sciences, d'accéder aux études scientifiques. Notre vocabulaire est différent du VGOS en ceci qu'il ne s'agit pas d'un vocabulaire commun à plusieurs disciplines mais propre à une seule. Avec l'élaboration de notre vocabulaire, nous aimerions intéresser les terminologues et également toute personne voulant acquérir des

connaissances dans le domaine de l'informatique, pour des raisons personnelles ou professionnelles. Bien que dans cette recherche nous nous intéressions uniquement aux unités lexicales simples, il serait intéressant dans une étude ultérieure d'étendre le travail aux termes complexes.

Notre travail est le début d'un projet plus vaste du laboratoire ÉCLECTIK (Équipe de recherche en combinatoire lexicale, terminologie et informatique) de l'Université de Montréal. Ce projet vise à élaborer un dictionnaire de base de l'informatique ayant des points communs avec le *Dictionnaire d'apprentissage du français des affaires* (DAFA), lequel est un dictionnaire multilingue axé sur la production. Le DAFA s'adresse à un public d'apprenants de français langue étrangère mais peut également rendre service aux traducteurs et à un public francophone qui s'intéresse au français des affaires.

« Contrairement à ce que l'on croit souvent, ce n'est pas la terminologie proprement dite qui constitue la pierre d'achoppement pour les apprenants. Les termes spécialisés, généralement monosémiques, ne mettent guère à l'épreuve leur compétence [en] combinatoire lexicale. Le vocabulaire courant, dit 'simple', en revanche (p. ex. haut, élevé, salarié, etc.) pose des problèmes bien plus ardus, sans parler de l'interférence de la langue maternelle, source de nombreuses erreurs également. »

Binon et Verlinde (1998: 448)

Notre vocabulaire est établi par l'opposition d'un corpus d'analyse à un corpus de référence de taille plus importante. Nous voulons extraire toutes les formes du corpus d'analyse se démarquant par une fréquence significativement

plus élevée par rapport au corpus de référence. Notre vocabulaire sera établi au moyen d'une analyse lexicale effectuée à l'aide du logiciel *TermoStat* développé par Drouin (2002).

Nous avons adopté deux méthodes de comparaison de corpus et nous allons comparer les résultats obtenus avec chacune d'elles. Dans la première méthode, nous utilisons un corpus global hétérogène, c'est-à-dire un corpus composé de documents qui opposent deux types de discours; nous opposons un corpus d'analyse composé de textes du domaine de l'informatique et un corpus de référence de nature journalistique.

La deuxième méthode fait appel à un corpus global homogène. Nous opposons six corpus d'analyse, chacun étant un sous-ensemble du corpus de référence et chacun représentant un sous-domaine de l'informatique. Le corpus de référence est l'ensemble des six sous-corpus. Il s'agit en fait du corpus d'analyse de la première méthode de comparaison.

Enfin, les deux méthodes feront l'objet d'une évaluation. En effet, afin de valider le vocabulaire caractéristique obtenu, nous aurons recours à deux dictionnaires spécialisés de l'informatique.

Notre mémoire comporte trois chapitres :

Le premier chapitre présente en trois sections la position théorique de notre recherche. Dans la première section du chapitre, nous définissons les notions de

« vocabulaire fondamental », de « vocabulaire scientifique/technique général » et de « langue de spécialité » afin de situer le vocabulaire que nous cherchons à isoler parmi les différentes strates lexicales. La deuxième section du chapitre propose un survol des travaux fondés sur la comparaison de corpus. Nous consacrons la troisième section à la présentation de logiciels d'analyse lexicale de corpus textuels qui mettent l'accent sur l'identification automatique de spécificités lexicales.

Le deuxième chapitre expose en détail la méthodologie de notre recherche. Il est divisé en trois parties comprenant chacune un certain nombre de sections. La première partie du chapitre comprend les cinq sections suivantes : *Présentation de la méthode 1*, *Constitution des corpus de M1*, *Présentation de la méthode 2*, *Constitution des corpus de M2* et *Étiquetage morphosyntaxique*. Dans cette première partie nous présentons les deux méthodes de comparaison de corpus à l'étude. Nous y décrivons également la constitution des corpus de référence et d'analyse quant à la taille des corpus et à la provenance et au choix des textes. Ensuite, nous décrivons les différents traitements qu'ont subis les corpus avant l'étiquetage et la lemmatisation.

La deuxième partie du chapitre comporte les sections *Constitution des listes* de référence et Calcul des spécificités à l'aide de TermoStat. Dans cette partie du chapitre nous abordons plus en détail les techniques utilisées pour isoler le vocabulaire caractéristique de l'informatique. Nous décrivons le logiciel *TermoStat* et présentons de façon sommaire la loi normale, à laquelle fait appel le logiciel.

Enfin, la troisième partie de ce chapitre est composée des sections *Gestion* des spécificités et *Résultats du calcul des spécificités*. Cette section présente la méthode que nous avons élaborée pour gérer les spécificités recueillies. Nous y exposons les critères d'acceptation des spécificités positives pour ensuite rendre compte du nombre de spécificités positives ignorées et retenues.

Le troisième chapitre consiste en l'analyse et la validation des résultats. Dans un premier temps, nous livrons les résultats obtenus pour les deux méthodes à l'étude. Nous décrivons également les phénomènes observés dans notre vocabulaire en dégageant certaines observations quant au comportement des spécificités positives. Dans un deuxième temps, nous présentons le moyen que nous avons adopté pour valider nos résultats.

# **CHAPITRE 1: ÉTAT DE LA QUESTION**

Le présent travail a pour objectif d'isoler le vocabulaire caractéristique du domaine de l'informatique à partir de corpus de textes associés à ce domaine. Les mots identifiés comme faisant partie du vocabulaire caractéristique de l'informatique seront répartis dans chacune des strates du lexique de la langue de spécialité. Nous jugeons donc important de définir les différents vocabulaires, ou strates lexicales, et nous nous y employons à la section 1.1.

La section 1.2 présente les notions de « comparaison de corpus ». Nous y exposons les travaux de la statistique textuelle sur lesquels reposent les deux méthodes de comparaison que nous mettons de l'avant. Une partie des travaux que nous avons choisi de présenter a pour but de mettre en évidence les phénomènes liés à la variation de la fréquence des unités lexicales dans un sous-corpus par rapport à l'ensemble. L'autre partie des travaux cherche à mettre en opposition deux types de corpus (spécialisé et non spécialisé) afin de faire émerger les particularités lexicales du corpus spécialisé.

Enfin, nous présentons à la section 1.3 cinq logiciels d'analyse lexicale de corpus textuels, dont *TermoStat* (Drouin 2002), le logiciel que nous avons utilisé pour isoler le vocabulaire caractéristique de l'informatique. Ces logiciels ont été choisis car ils offrent tous la possibilité d'extraire les spécificités lexicales à partir de corpus.

## 1.1 Stratification lexicale

Les différents spécialistes (chercheurs, terminologues, lexicographes, linguistes et traducteurs) s'entendent pour dire qu'il n'existe encore aucune mesure permettant de délimiter sans équivoque *langue générale* (appelée également *langue usuelle* ou *langue commune*) et *langue de spécialité*.

Les langues de spécialité sont composées de textes spécialisés qui eux contiennent des mots (le vocabulaire usuel) et des termes (le vocabulaire spécialisé). Entre ces deux extrémités du spectre lexical se trouve un éventail de vocabulaires que nous définirons dans les paragraphes qui suivent.

#### 1.1.1 Vocabulaire fondamental

Des vocabulaires fondamentaux, ou vocabulaires de base, ont été établis dans plusieurs langues. Ces vocabulaires ont pour but principal de faciliter l'acquisition d'une langue pour un étranger. Ils sont composés des mots les plus fréquemment utilisés d'une langue (âge, couleur, beaucoup, parler, etc.) et permettent de comprendre et d'exprimer rapidement les idées et besoins courants de la vie quotidienne. Comme l'écrit Pottier (1973 : 539), contrairement à un trésor qui tend à réunir toutes les unités lexicales différentes qui sont employées par tous les membres d'une communauté linguistique, le vocabulaire fondamental reprend les unités lexicales communes à la majorité des membres d'une communauté linguistique. Les vocabulaires de base sont donc des sous-ensembles de la langue générale.

Les travaux menés sur les vocabulaires de base sont beaucoup plus nombreux en anglais qu'en français. Les premiers travaux en langue anglaise ont été entrepris dans la première moitié du  $20^{\rm e}$  siècle, alors qu'en français ils ont été élaborés dans la deuxième moitié du même siècle. Le vocabulaire de base le plus connu en anglais est sans doute le *Basic English* de Ogden.

De toutes les recherches qui ont été menées sur le vocabulaire fondamental du français, la plus achevée serait celle du *Français fondamental* (FF) de l'équipe dirigée par G. Gougenheim au Centre de Recherche et d'Étude pour la diffusion du français (CRÉDIF) (voir Gougenheim *et al.* 1964)<sup>1</sup>. Le FF offre sans doute la liste de fréquences la plus facile à utiliser et la plus populaire auprès des non-linguistes qui s'intéressent au langage : pédagogues, orthophonistes, psychologues, etc. Le FF a été élaboré à partir d'un corpus oral. Il comprend 1 475 mots différents dont 253 sont des mots grammaticaux. Le FF a été établi dans le but d'élaborer un lexique comprenant les mots les plus utiles et les plus fréquents de la langue française parlée pour permettre aux locuteurs d'une autre langue de communiquer dans la vie de tous les jours. Une fois ce vocabulaire de base établi,

« le C.R.E.D.I.F. a été amené à greffer sur ce 'tronc commun' que constitue le français fondamental (1<sup>er</sup> et 2<sup>e</sup> degrés), divers vocabulaires spécialisés dont le but est, une fois encore, de permettre à l'étudiant d'aller plus aisément à l'essentiel ».

(Gougenheim *et al.* 1964 : 13)

\_

<sup>&</sup>lt;sup>1</sup> Depuis, on a conçu des dictionnaires de fréquences plus riches en informations tirés de corpus écrits beaucoup plus imposants, comme celui du *Trésor de la langue française*.

C'est en greffant des vocabulaires spécialisés au français fondamental qu'est né le *vocabulaire général d'orientation scientifique* que nous allons définir à la section 1.1.2.1.

À la fin des années soixante-dix, des études ont également été menées sur le vocabulaire fondamental parlé québécois (voir Beauchemin et Martel 1979 et Beauchemin 1982). Beauchemin, en utilisant le FF comme base pour sa recherche, a élaboré un corpus de 300 000 mots de la langue parlée au Québec, provenant d'enregistrements d'enquêtes sur différents sujets, et en a recensé le vocabulaire commun à tous les répondants.

Bien que les travaux abondent en ce qui concerne le vocabulaire fondamental, les chercheurs ne sont pas tous d'accord sur la façon de le définir et ne proposent pas tous les mêmes listes de mots. C'est ainsi que Lee (2001) fait remarquer que les différentes façons de comprendre la notion de « vocabulaire fondamental » reflètent le manque de précision et les nombreux aspects du concept. D'après lui, il arrive que les auteurs de listes d'un vocabulaire fondamental aient des intérêts et des objectifs différents, ce qui expliquerait qu'ils obtiennent des listes différentes. Lee (2001 : 252-255) propose sept descriptions du vocabulaire fondamental, indépendantes les unes des autres. Dans le paragraphe ci-dessous, nous résumons la première, qui est celle qui correspond le plus à la façon dont nous abordons ce vocabulaire.

Ce sont les mots les plus fréquents de la langue commune. Les listes établies à partir de corpus de langue anglaise indiquent que les 700 mots les plus

fréquents représentent 70 % de tout texte de langue anglaise et les 2 500 mots les plus fréquents, en représentent 80 % (Kennedy 1998 : 96-97, cité dans Lee 2001 : 252). Dans un corpus de nature générale, en considérant les 100 mots les plus fréquents, on peut s'attendre à ce que 80 % d'entre eux soient des mots grammaticaux. Par contre, toujours selon Lee, plus le corpus est spécialisé ou homogène, plus il y aura de mots pleins parmi les mots les plus fréquents.

## 1.1.2 Vocabulaire scientifique/technique général

Le vocabulaire scientifique/technique général est celui qui est commun aux différents domaines scientifiques ou techniques. Il couvre une bonne part de tous les sous-domaines appartenant au domaine des sciences ou des techniques.

Le FF de Gougenheim avait été élaboré dans un but pédagogique; il devait permettre aux étrangers d'apprendre les rudiments de la langue française afin de pouvoir communiquer dans la vie de tous les jours. En langue de spécialité, des lexiques ont également été créés dans un but pédagogique. Ces lexiques veulent permettre aux apprenants d'une langue, ou d'une langue de spécialité, d'accéder aux études scientifiques ou techniques.

#### 1.1.2.1 Vocabulaire général d'orientation scientifique

Le vocabulaire général d'orientation scientifique (VGOS) est le premier vocabulaire scientifique/technique général. Il a été élaboré par une équipe du CRÉDIF dirigée par André Phal. Le VGOS, qui fait suite au FF, est un vocabulaire

intermédiaire entre le vocabulaire usuel et le vocabulaire des différentes spécialités. Il est composé majoritairement de mots de la langue générale mais utilisés à des fins d'expression scientifique et technique. On doit l'idée du VGOS à Michéa, un des auteurs du FF. Après avoir dégagé du FF 600 mots qui lui paraissaient être employés fréquemment dans la langue scientifique (action, augmenter, calcul, méthode, etc.), Michéa constata que ces mots formaient en quelque sorte le vocabulaire de base de l'expression scientifique (Phal 1971 : 8). Michéa donna le sens suivant au VGOS :

- « "vocabulaire général", à la fois parce qu'il fait partie du lexique général de la langue et parce qu'il est commun à toutes les spécialités considérées du point de vue spéculatif;
- « "d'orientation scientifique", à la fois parce que le sens des mots qui le constituent est orienté par les contextes spécialisés dans lesquels ils sont employés et parce qu'il fournit à quelqu'un qui en a la maîtrise, l'essentiel du bagage linguistique nécessaire à qui veut s'orienter vers des études scientifiques ou techniques. » (*ibid.* : 9)

Ce vocabulaire limité a été conçu pour faciliter aux étudiants, chercheurs et techniciens étrangers l'accès aux études scientifiques en langue française (*ibid.* : 7). Même si ce vocabulaire a été élaboré avant tout pour un public étranger, il peut aussi être utile à un public de langue française qui éprouverait des difficultés à saisir la forme sous laquelle on leur enseigne une matière (*ibid.* : 14).

Le corpus ayant servi à l'élaboration du VGOS est composé de textes écrits portant sur les mathématiques, les sciences physiques et les sciences naturelles. À

partir d'analyses de fréquence et de dispersion, Phal a isolé le vocabulaire commun aux textes composant son corpus.

## 1.1.2.2 Vocabulaire général d'orientation technique

Le vocabulaire général d'orientation technique (VGOT), élaboré par Normand Maillet, s'appuie d'abord sur les principes et méthodes de la statistique lexicale développés par Muller. Le VGOT est inspiré du VGOS et, tout comme ce dernier, il s'agit d'un vocabulaire établi à partir des analyses de fréquence et de dispersion. L'originalité du VGOT par rapport au VGOS est qu'il a été établi à partir d'un corpus d'échantillons de textes alors que le VGOS avait été élaboré à partir d'un corpus de manuels plein texte (Maillet 1998 : 261). De plus, le corpus ayant servi à l'élaboration du VGOT n'est pas composé de textes scientifiques comme l'est le VGOS mais de textes techniques québécois (100 échantillons de 2 500 mots chacun). Selon Maillet :

« Dans les textes techniques, nous observons une stratification lexicale allant des termes spécialisés propres à un domaine, aux mots de la langue générale, en passant par un espace intermédiaire qui permet l'expression technique. Cet espace renferme un vocabulaire commun à plusieurs domaines, une sorte de stock lexical situé entre la langue générale et la langue de spécialité. » (*ibid.* : 260)

Dans son travail, Maillet s'est attaché à faire ressortir les particularités du VGOT par rapport au VGOS. Il conclut que le VGOT se distingue du VGOS surtout par le vocabulaire et non pas par la structure ni par la taille (*ibid.* : 374).

#### 1.1.2.3 Vocabulaire général d'enseignement scientifique

Le vocabulaire général d'enseignement scientifique (VGES), élaboré par Lamria Chetouani, est un vocabulaire pédagogique de base inspiré des travaux du CREDIF. Les auteurs des listes du FF et du VGOS avaient écarté ce vocabulaire de leurs travaux. Le corpus ayant servi à l'élaboration du VGES est composé de textes provenant de la langue parlée et de la langue écrite, dans les domaines des sciences exactes et des sciences sociales. La préoccupation principale de Chetouani était de répertorier les éléments pédagogiques des textes à teneur scientifique (rappelons que, nous étudierons d'abord cela, je récapitule, etc.) pour en dégager le noyau dur du vocabulaire pédagogique. Le VGES vise donc à faciliter la transmission du contenu d'un discours scientifique (Chetouani 1989 : 84-85).

#### 1.1.3 Langue de spécialité

La notion de « langue de spécialité » fait référence à la terminologie propre à une situation de communication circonscrite à un domaine particulier. Malgré les nombreuses études qui ont été menées sur la langue de spécialité (LSP), il est encore difficile de bien la définir, les chercheurs n'étant pas toujours d'accord sur l'extension à donner à ce concept. Nous la définirons en ces termes : la langue de spécialité témoigne d'un usage particulier de la langue générale lorsqu'elle est

utilisée dans une situation de communication qui implique la transmission d'information relevant d'un champ d'expérience particulier. Les terminologues opposent souvent *LSP* et *langue générale*, cette dernière étant composée du vocabulaire commun utilisé et compris par les locuteurs d'une communauté linguistique.

« Les langues de spécialité sont les instruments de base de la communication entre spécialistes. La terminologie est l'aspect le plus important qui différencie non seulement les langues de spécialité de la langue générale, mais également les différentes langues de spécialité. »

(Cabré 1998 : 90)

Plusieurs définitions ont été données à la LSP et différentes dénominations en ont découlé. Kocourek (1991 : 16) dresse une liste des principales appellations : langue spéciale et langue technique (Vendryès 1968 : 276-277), langues spéciales (Saussure 1975 : 41), langue de groupes particuliers (Meillet 1975 : 247), langue des sciences (Fuchs 1966 : 675), langues techniques (Dauzat 1967 : 85), langue professionnelle (Wartburg 1963 : 116) et langues techniques et scientifiques (Quemada 1955 : 7).

Les auteurs s'entendent généralement pour dire que l'élément le plus caractéristique à la LSP est son lexique. La figure 1, tirée de Rondeau (1981 : 27), schématise l'ensemble lexical général d'une langue. La zone centrale correspond à la zone de la langue commune. La zone mitoyenne correspond à la frontière entre la zone de la langue commune et l'ensemble des zones des LSP. Dans cette zone, une

forme linguistique peut se retrouver à la fois dans les LSP et dans la langue commune. Rondeau (*ibid*. : 26) donne l'exemple de la forme *colique* se retrouvant à la fois dans la zone des LSP, où elle fait référence à une seule notion du domaine médical, et dans la zone de la langue commune où elle prend plusieurs sens. Les LSP se répartissent sur trois zones, soit la zone mitoyenne, la zone centrale des LSP et la zone des ensembles ultra-spécialisés. Selon nous, les VGOS et VGOT correspondraient à la zone mitoyenne.

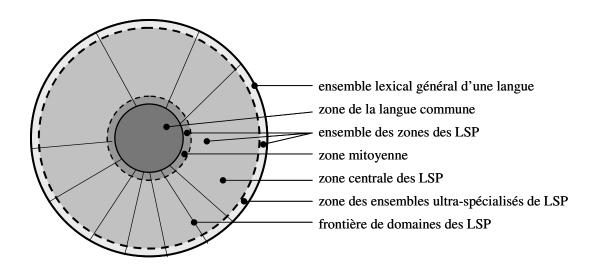


Figure 1. Langues de spécialité et langue commune

Enfin, parmi les termes qui composent la documentation spécialisée, certains sont empruntés à la langue commune (*mémoire* en informatique et *stratégie* en marketing). À l'inverse, des termes passent de la langue spécialisée à la langue générale lorsqu'ils entrent dans le vocabulaire des non-spécialistes (*imprimante* du domaine informatique et *idiotie* du domaine médical).

Nous venons de définir les différentes strates lexicales de la LSP. Nous croyons que le vocabulaire caractéristique que nous cherchons à isoler comprendra des termes spécialisés (excluant les termes hautement spécialisés) et un vocabulaire fonctionnel de l'informatique, lequel inclura des mots appartenant à la langue générale.

### 1.2 Comparaison de corpus

La comparaison de corpus est au centre même de notre recherche. C'est par la comparaison de deux corpus que nous ferons émerger le vocabulaire caractéristique de l'informatique.

D'après Rayson et Garside (2000), il existe deux principaux types de comparaison de corpus : la comparaison d'un échantillon de corpus à un corpus de plus grande taille et la comparaison de deux corpus ayant environ la même taille. Ces deux types de comparaison de corpus peuvent être utilisés pour extraire les mots-clés, lesquels différencieront un corpus de l'autre.

Dans le premier type de comparaison, le corpus de plus grande taille réfère au corpus de la norme. Selon Rayson et Garside, une attention particulière doit être portée à la représentativité, l'homogénéité à l'intérieur du corpus, la « comparabilité » des corpus et la fiabilité des tests statistiques mobilisés pour la comparaison.

Dans le deuxième type de comparaison, les corpus sont de même taille. L'objectif est de découvrir les caractéristiques qui distinguent les deux corpus.

Avant de présenter des travaux fondés sur la comparaison de corpus, nous décrivons d'abord la *linguistique de corpus* et la *statistique textuelle*, sur lesquelles s'appuient les travaux que nous présentons.

## 1.2.1 Linguistique de corpus

Sinclair définit le *corpus* de la façon suivante :

« Un corpus est une collection de données langagières qui sont sélectionnées et organisées selon des critères linguistiques explicites pour servir d'échantillon du langage. »

(Sinclair 1996 : 4, cité dans Habert 1997 : 11)

Dans le présent travail, le terme *corpus* référera toujours à la notion de « corpus électronique », c'est-à-dire une collection de textes informatisés. Les corpus électroniques sont encodés de façon standardisée et homogène pour permettre des recherches et des extractions « ouvertes » (Sinclair 1996 : 5).

Les premiers travaux en linguistique de corpus remontent à 1961. C'est à ce moment que le premier corpus électronique a été mis sur pied. Il s'agissait de ce qui s'appelle maintenant le corpus *Brown* (Garside *et al.* 1997 : 1). Le corpus a été complété en 1964 et comprenait alors un million de mots. Depuis les années 1990, la taille des corpus s'est accrue de façon considérable. Par exemple, le *British* 

National Corpus (BNC) comprend 100 millions de mots et le Bank of English en compte 300 millions. Les corpus ont maintenant la possibilité d'être annotés, c'est-à-dire que les mots peuvent être suivis d'informations morphologique, syntaxique, sémantique, etc. Leur popularité est due en partie au traitement automatique des langues naturelles, qui s'y intéresse désormais de plus en plus, notamment en mettant de l'avant des techniques permettant de cerner la terminologie des corpus.

D'après Bindi *et al.* (1994 : 29), l'analyse de corpus est le seul outil connu pouvant identifier et décrire les *sous-langages*. Par *sous-langage*, les auteurs font référence aux différentes utilisations d'une même langue dans différents contextes de communication, avec des locuteurs ayant différents objectifs de communication. Nous ne croyons pas risqué de croire qu'ils font référence à une notion très près de celle de «LSP». Selon eux, les corpus sont également les seules sources de données possibles permettant d'analyser les propriétés statistiques des différents éléments des langues et des sous-langages, soit la répartition, la fréquence, l'usage, etc.

#### 1.2.2 Statistique textuelle

La statistique textuelle découle de travaux d'un courant aux dénominations changeantes (Lebart et Salem 1994 : 7). La linguistique quantitative, manifeste dans les travaux des Allemands et des Hollandais dès la fin du 19<sup>e</sup> siècle, en est l'ancêtre. Le premier grand travail sur la fréquence des mots et des lettres a servi autant pour les analyses linguistiques que pour établir des listes de fréquence à des fins pédagogiques et aussi à des fins pratiques, telles la configuration des claviers

de machines à écrire et les méthodes de sténographie. Dans les années 1950, les Anglo-Saxons et les Français ont appliqué à leurs langues les mêmes méthodes et ont consolidé les bases de la *statistique linguistique*. Puis, dans les années 1970, on préfère parler de *lexicométrie* ou *statistique lexicale* en référence aux travaux spécifiques sur le vocabulaire et le lexique, notamment les études comparatives sur le vocabulaire des « grands auteurs ».

La statistique textuelle, par l'analyse de la fréquence et de la répartition des phénomènes, permet de dégager les constantes et les écarts lexicaux d'un corpus préétabli. Elle vise à extraire l'information essentielle, les structures signifiantes les plus fortes contenues dans un texte. Elle est un domaine de recherche très vaste dont nous nous contenterons de tracer les grandes lignes. La statistique textuelle a d'abord été étudiée sous l'angle de l'analyse de textes littéraires dont on analysait le comportement lexical. Puis, avec la vitesse de l'ordinateur et la documentation de plus en plus abondante des textes technoscientifiques, des recherches ont été entreprises dans les domaines spécialisés. Le facteur déclenchant a été l'analyse automatique de contenu des documents scientifiques. Comme nous l'avons déjà mentionné, nous voulons, dans le cadre de notre recherche, isoler le vocabulaire caractéristique d'un corpus spécialisé au moyen d'analyses statistiques. Pour atteindre cet objectif, nous aurons recours à un logiciel dont l'élaboration a intégré les résultats de travaux de statistique linguistique, lesquels sont présentés dans les sections qui suivent.

### 1.2.2.1 Muller (1977, 1979)

Muller a joué un grand rôle en statistique lexicale et ses travaux ont servi et continuent à servir de modèle aux chercheurs.

« Le calcul statistique permet de donner un contenu à la notion subjective de *vocabulaire caractéristique* d'un texte ou d'une partie d'un texte. Mais il va de soi qu'une fréquence observée ne saurait devenir 'caractéristique' que comparée à une fréquence théorique, donc par référence à un texte ou à un corpus plus étendu que celui qui est en question; cet ensemble plus grand, pris comme source du modèle théorique, peut contenir ou ne pas contenir le texte considéré. »

(Muller 1977 : 77)

Muller nomme vocabulaire caractéristique positif les formes dont on peut affirmer avec un niveau de confiance de 95 % qu'elles sont plus fréquentes dans le corpus d'analyse que dans le corpus de référence. Inversement, le vocabulaire caractéristique négatif est celui dont on peut affirmer, toujours avec un niveau de confiance de 95 %, que les formes sont moins fréquentes dans le corpus d'analyse que dans le corpus de référence. Par formes, nous faisons référence aux occurrences identiques dans un corpus de textes.

Afin d'isoler le vocabulaire spécifique à un auteur ou à une œuvre, Muller a fait appel au modèle hypergéométrique. Il utilisait une méthode qui permettait de comparer une œuvre et ses fragments quand ceux-ci formaient des unités

caractéristiques (Muller 1979 : 173), comme les différents rôles d'une pièce de théâtre ou le dialogue dans une œuvre narrative.

### 1.2.2.2 Lafon (1980)

La méthode utilisée par Lafon lui a permis d'obtenir non seulement le vocabulaire spécifique à un sous-corpus, comme l'avait fait Muller, mais également d'obtenir le vocabulaire de base de l'ensemble d'un corpus. C'est ainsi que Lafon utilise un grand corpus divisé en dix sous-corpus, chacun représentant un discours prononcé par Robespierre entre les mois de novembre 1793 et de juillet 1794. Il compare chacun des discours au corpus global afin d'isoler les spécificités lexicales du discours robespierriste, en reprenant l'idée de Muller sur le vocabulaire caractéristique positif ou négatif. Lafon emploie cependant les expressions spécificités lexicales positives ou spécificités lexicales négatives. Une fois les spécificités isolées, le vocabulaire commun aux dix sous-ensembles demeure, et c'est ainsi que Lafon découvre le noyau de vocabulaire du discours robespierriste (Lafon 1980).

Selon Lafon, (*ibid.* :128), la fréquence des formes au sein d'un corpus divisé en plusieurs parties est un problème important et très souvent traité en statistique linguistique. Il note que les instruments statistiques habituellement utilisés pour traiter ce problème, notamment le *khi2*, la *loi normale avec calcul d'écart réduit* pour les formes les plus fréquentes et la *loi de Poisson* pour les formes rares, ne constituent jamais que des approximations du calcul adéquat. Il a préféré, comme Muller l'avait fait avant lui, emprunter à la statistique les formules

de la distribution hypergéométrique, qui, selon lui, sont exactement adaptées à la population discrète des occurrences de vocabulaire et valides pour toute la gamme des fréquences rencontrées.

### 1.2.2.3 Lebart et Salem (1994)

À l'instar de Lafon, Lebart et Salem (1994) reprennent l'idée de Muller (1979) d'isoler les spécificités dans une tranche de corpus. Ces derniers ont eux aussi recours à un modèle hypergéométrique pour y parvenir. Aux notions de « spécificités positives » et de « spécificités négatives », ils ajoutent celle des « formes banales ». Il s'agit des formes qui ne présentent aucune spécificité. Dans la méthode utilisée par Lebart et Salem (1994 : 176), lorsqu'une forme est banale pour chacune des parties du corpus, cette forme appartient au vocabulaire de base du corpus.

### 1.2.2.4 Nelson (2000)

La thèse de doctorat de Nelson porte sur la comparaison de trois corpus. Le premier, *Published Materials Corpus* (PMC), est un corpus de 590 000 mots consistant en 33 manuels de cours et livres de référence publiés, traitant de l'anglais des affaires. Le deuxième corpus, *Business English Corpus* (BEC), est un corpus de 1 023 000 mots et est divisé en textes parlés (44 %) et écrits (56 %). Le corpus consiste en 28 *genres* catégorisés selon le langage employé pour « faire des affaires » (59 %) et le langage employé pour « parler affaires » (41 %). Finalement,

le troisième corpus, *British National Corpus* (BNC), est un corpus de deux millions de mots qui est utilisé à titre de corpus de référence.

La comparaison des corpus a été exécutée à l'aide du logiciel *WordSmith Tools*<sup>2</sup>. Nelson a d'abord comparé le corpus BEC au corpus BNC afin de faire émerger les mots suremployés et les mots sous-employés du corpus de l'anglais des affaires. En utilisant l'ensemble des mots suremployés dans le corpus BEC, Nelson a défini le champ lexical de la langue des affaires pour ensuite analyser le comportement de ces mots du point de vue grammatical et sémantique.

La deuxième comparaison effectuée dans ce travail est celle du corpus PMC au corpus BNC. À partir de cette comparaison, Nelson voulait voir en quoi les manuels et les livres publiés différaient de l'anglais général. En troisième lieu, il a comparé le corpus PMC au corpus BEC afin d'observer les différences entre deux corpus traitant de l'anglais des affaires. Il a découvert que le vocabulaire spécifique au corpus traitant des affaires en général (BEC) est très spécialisé et limité, et que le vocabulaire spécifique au corpus de manuels et livres de référence (PMC) l'est encore plus.

Les résultats obtenus par Nelson montrent que l'anglais des affaires diffère de façon significative de l'anglais général en ce qui a trait au lexique, à la « prosodie sémantique », au régime et au réseau sémantique. De la même façon, l'anglais des affaires des livres publiés diffère grandement de l'anglais des affaires « de la vie de tous les jours ».

۰

<sup>&</sup>lt;sup>2</sup> Nous décrivons le logiciel *WordSmith Tools* à la section 1.3.5.

### 1.2.2.5 Drouin (2002)

Dans sa thèse, Drouin a cherché à isoler et à cerner automatiquement le contenu d'un corpus spécialisé d'un point de vue terminologique. Les algorithmes mis en oeuvre par Drouin reposent sur l'hypothèse que l'on peut utiliser la thématique principale d'un corpus pour en identifier la terminologie. Cette thématique est identifiée à l'aide de tests statistiques (Drouin 2002 : 5).

L'approche adoptée par Drouin repose sur l'opposition de deux corpus : un corpus de référence composé de divers textes journalistiques et un corpus d'analyse formé de documents techniques appartenant au domaine des télécommunications (*ibid.*). Son approche a intégré des travaux utilisant la technique des spécificités mais à la différence des auteurs de ces travaux, il a adopté une perspective terminologique. Dans le cadre de sa recherche, Drouin a élaboré le logiciel *TermoStat* qui permet d'isoler les particularités lexicales et les termes d'un corpus d'analyse<sup>3</sup>.

Les travaux que nous venons de décrire se fondent sur des méthodes de la statistique textuelle. Dans la section qui suit, nous présentons des logiciels qui intègrent ces méthodes. Ces logiciels facilitent l'acquisition du vocabulaire caractéristique et permettent dans la plupart des cas de faire des analyses textuelles plus approfondies.

<sup>&</sup>lt;sup>3</sup> Nous présentons *TermoStat* à la section 1.3.1.

# 1.3 Logiciels d'analyse lexicale de corpus textuels

Les logiciels d'analyse lexicale se basent sur l'unité textuelle qu'est la forme pour estimer ses occurrences et comparer ses fréquences dans un sous-ensemble du corpus par rapport à un ou à plusieurs autres sous-ensembles, ou par rapport au corpus global. Nous présentons, dans les paragraphes qui suivent, cinq logiciels d'analyse de données textuelles qui permettent d'isoler le vocabulaire caractéristique d'un texte.

#### 1.3.1 TermoStat - Patrick Drouin

Le logiciel *TermoStat* a été élaboré dans le but de tester la stratégie d'acquisition automatique de termes conçue par Drouin dans le cadre de sa thèse.

TermoStat permet d'identifier les spécificités positives d'un corpus spécialisé pour en dégager la terminologie. Comme nous l'avons mentionné à la section 1.2.2.5, l'approche adoptée par Drouin repose sur l'opposition de deux types de corpus.

Afin de procéder à l'acquisition automatique des termes, le logiciel a recours à des tests statistiques qui ont pour objet de comparer le lexique du corpus d'analyse à celui du corpus de référence. Cette comparaison conduit à l'identification des unités lexicales qui possèdent une fréquence qui se démarque significativement de celle observée dans le corpus de référence : les *pivots lexicaux* spécialisés (PLS) (Drouin 2002 : 5).

La notion de « PLS » a été définie par Drouin dans le cadre de sa thèse. Les PLS constituent les formes (nominales et adjectivales seulement) les plus représentatives du lexique du corpus d'analyse. Cette notion chapeaute à la fois les notions de « vocabulaire spécialisé », de « VGOS », de « spécificité » et de « vocabulaire caractéristique » (*ibid.* : 141).

Une fois identifiés, les PLS sont utilisés comme point de départ pour l'acquisition automatique des termes. Le processus d'acquisition exploite le concept de frontière de termes afin d'isoler les candidats-termes. Puis, afin de maximiser la qualité des résultats présentés par *TermoStat*, Drouin propose un indice permettant de représenter le caractère terminologique des candidats-termes : l'indice *iTer*, lequel prend en considération la fréquence et la longueur du candidat observables dans le corpus d'analyse (*ibid.* : iii).

#### 1.3.2 LEXICO - André Salem

Le logiciel *LEXICO* est un logiciel conçu pour le traitement lexicométrique de textes comportant plusieurs centaines de milliers d'occurrences. Le logiciel a d'abord été développé par André Salem et il est désormais maintenu par une équipe de l'*Université de la Sorbonne Nouvelle - Paris 3*.

Le logiciel comprend trois types d'outils : des outils d'exploration textuelle (segmentation, concordances, recherche de segments répétés), des outils d'analyse statistique (découpage en parties, spécificités) et des outils de navigation

lexicométrique (fonctionnalités qui permettent de se déplacer parmi les résultats produits par les différentes méthodes lexicométriques et le texte initial).

Le logiciel effectue d'abord une segmentation (opération qui consiste à délimiter des unités minimales dans un texte), puis il constitue la liste des formes présentes dans le corpus. Pour chacune de ces formes, il calcule le nombre de ses occurrences ainsi que son rang alphabétique. Pour comparer les variations dans l'usage du vocabulaire entre les différentes parties du corpus, l'outil d'analyse statistique est équipé d'un module appelé *Statistique par parties* qui opère une série de calculs statistiques et crée différents fichiers. L'un de ces fichiers, *Analyse des spécificités*, indique la ventilation des formes dans les parties et adjoint aux fréquences qui se démarquent un indice de spécificité interprété par les signes + ou – qui indiquent un suremploi ou un sous-emploi des formes (Lamalle *et al.* 2003).

# 1.3.3 Hyperbase – Étienne Brunet

Le logiciel *Hyperbase* était à l'origine destiné aux analyses lexicologiques et a été principalement utilisé dans les corpus littéraires. Son application s'est étendue aux textes juridiques, historiques, publicitaires et aux sondages ou enquêtes d'opinion. Ainsi, des sociologues, des psychologues, et des instituts de sondages utilisent *Hyperbase* pour l'analyse des enquêtes en texte libre ou des études de marché. Même s'il peut traiter rapidement et efficacement les textes courts, son rendement maximum est atteint dans les grands corpus (au delà du million de mots).

Le programme répond aux besoins classiques du traitement automatique des textes : index sélectifs ou systématiques, dictionnaires des fréquences, concordances, sélection de contextes élargis, cooccurrences, recherche des parties ou groupes de mots. Il permet également de définir un vocabulaire spécifique du corpus étudié (une pièce de Molière) par rapport à un corpus de référence (l'ensemble du théâtre de Molière).

Pour calculer les spécificités, *Hyperbase* utilise la *loi normale* ou le *modèle hypergéométrique* selon le choix de l'utilisateur (Brunet 2001).

## 1.3.4 TALTAC - Sergio Bolasko

D'après Bolasko (2000), malgré l'enrichissement récent des fonctions présentes dans les logiciels décrits aux paragraphes précédents, on n'a pas encore tracé une stratégie d'action pour le traitement des textes qui puisse assurer une bonne désambiguïsation des termes et, en même temps, une sélection du langage caractéristique dans le vocabulaire d'un corpus. Selon lui, la solution à ce problème se trouve dans la mise en correspondance directe d'instruments de types différents, c'est-à-dire : mesures statistiques et ressources linguistiques. Le logiciel *TALTAC* utilise plusieurs logiciels en parallèle. D'après son expérience, Bolasko note que l'utilisation simultanée d'*INTEX*<sup>4</sup>, de *TALTAC* et de *LEXICO* produit un enrichissement remarquable de l'analyse textuelle.

.

<sup>&</sup>lt;sup>4</sup> INTEX est un environnement de développement linguistique fondé sur la technologie des transducteurs à états finis dans lequel sont construites des descriptions formalisées des langues

Dans les étapes les plus usuelles de sa procédure, TALTAC utilise, comme nous venons de le mentionner, plusieurs logiciels en parallèle. La première étape est la *normalisation préliminaire* du corpus, c'est-à-dire le prétraitement. La deuxième étape est celle où les mesures lexicométriques de base sur le corpus permettent d'avoir une idée concrète de l'étendue du corpus et du vocabulaire, de la gamme des fréquences et du seuil éventuellement utile à l'analyse de contenu. La troisième étape, l'*identification des polyformes*, permet de désambiguïser certaines formes simples. La quatrième étape est celle de la *lemmatisation grammaticale* du texte et c'est la cinquième étape qui permet d'extraire le vocabulaire caractéristique du corpus par comparaison avec les dictionnaires de fréquence (Bolasko 2000).

#### 1.3.5 WordSmith Tools - Mike Scott

WordSmith Tools est une suite logicielle de six outils d'analyse textuelle. Ce progiciel commercialisé a été développé par Mike Scott (Scott, 1998), de l'Université de Liverpool.

Parmi les six outils disponibles, celui qui nous intéresse est l'outil permettant la comparaison de corpus. Il s'agit de l'outil *Keywords*, lequel identifie les mots-clés d'un texte à analyser, c'est-à-dire les mots qui caractérisent ce texte. Il s'agit des mots dont la fréquence est anormalement élevée par rapport aux textes d'un corpus plus général.

naturelles. INTEX permet entre autres le traitement de corpus dont les opérations typiques incluent l'indexation des textes et l'étude statistique des résultats (Silberztein 2000 : 8).

Keywords compare deux listes de mots, créées préalablement avec l'outil Wordlist. La première liste est la plus longue des deux listes et est considérée comme liste de référence. La deuxième est celle qui a été construite à partir du texte à analyser. Les mots-clés sont calculés en comparant la fréquence d'un mot dans la liste d'analyse avec la fréquence de ce mot dans la liste de référence en utilisant les tests statistiques du khi2 et de vraisemblance (Log Likelihood).

En plus de l'outil *Keywords*, *WordSmith Tools* offre cinq autres outils d'analyse textuelle, soit :

- Wordlist, un outil générant des listes de mots en ordre alphabétique et en ordre de fréquence. Cet outil permet également d'obtenir des statistiques sur les textes telles que le nombre total de mots, la longueur des mots, le nombre de phrases, etc. Les listes de mots obtenues sont ensuite utilisées par le programme Keywords;
- Concord, un concordancier permettant d'obtenir une liste de mots en contexte;
- *Splitter*, un outil permettant de diviser un gros fichier en plusieurs petits fichiers;
- Text Converter, un outil permettant d'apporter des modifications dans plusieurs fichiers à la fois;

 Viewer, un outil permettant de visualiser les fichiers en différents formats.

#### 1.4 Conclusion

Les travaux des auteurs précités, à l'exception de ceux de Nelson et de Drouin, reposaient principalement sur des textes littéraires ou politiques, ou encore utilisaient la technique des spécificités dans le cadre de l'analyse du discours.

Notre recherche se distingue de celle de ces auteurs en ceci que nous mettons en parallèle deux méthodes de comparaison de corpus se fondant sur la technique des spécificités: l'une opposant un corpus spécialisé (de nature informatique) à un corpus général (de nature journalistique), et l'autre opposant des sous-ensembles d'un corpus spécialisé (de nature informatique) à l'ensemble du corpus. Ainsi, la comparaison que nous effectuons dans la méthode 1 se fonde sur un corpus global hétérogène alors que celle de la méthode 2 se fonde sur un corpus global homogène.

La méthode 1 est inspirée des travaux de Drouin mais s'en distingue du fait que nous nous intéressons à l'ensemble du vocabulaire de l'informatique, c'est-à-dire à toutes les parties du discours, et non seulement à la terminologie d'un domaine présentée sous forme nominale ou adjectivale.

La méthode 2 se distingue des travaux effectués sur des corpus homogènes du fait que nous utilisons des sous-corpus issus d'un même domaine technique, soit

l'informatique. Phal (1971), par exemple, utilisait un corpus composé de textes appartenant aux trois grandes divisions traditionnelles des « sciences exactes » : mathématiques, sciences physiques (physique et chimie), et sciences naturelles. Quant à Maillet (1998), il utilisait un corpus composé de textes provenant de différents domaines techniques.

Nelson (2000) s'est intéressé aux deux types de comparaison de corpus : corpus homogènes et corpus hétérogènes. Toutefois, dans les deux cas, notre méthode diffère de la sienne en ceci que les textes composant ses corpus n'ont pas été étiquetés ni soumis à une lemmatisation. De plus, dans sa comparaison de corpus hétérogènes, Nelson s'intéressait aux comportements des spécificités positives ET négatives du point de vue grammatical et sémantique. Il a également comparé des corpus homogènes en ce sens que les deux corpus à l'étude relevaient du domaine des affaires. Toutefois, ces corpus étaient de taille et de registres différents.

# **CHAPITRE 2 : MÉTHODOLOGIE**

Ce chapitre présente la description de la démarche méthodologique adoptée pour établir notre vocabulaire caractéristique. Celle-ci repose sur deux méthodes différentes. Dans les deux cas, un logiciel de statistique textuelle a été utilisé afin d'isoler les *spécificités lexicales* des corpus d'analyse, c'est-à-dire les formes se démarquant par leur fréquence des formes du corpus de référence. Parmi celles-ci, nous nous intéressons aux formes qui se démarquent par leur fréquence élevée, soit les *spécificités positives*.

Notre travail visant à isoler le vocabulaire caractéristique de l'informatique à l'aide de deux méthodes de comparaison de corpus différentes, nous avons dû élaborer deux bases de données en parallèle. Nous avons importé les résultats du calcul des spécificités dans chacune des bases de données afin d'être en mesure d'observer et de comparer ces résultats pour chacune des méthodes à l'étude.

#### 2.1 Présentation de la méthode 1

La première méthode (M1) consiste à comparer les fréquences des unités lexicales d'un corpus d'analyse (CA) à celles d'un corpus de référence (CR) afin de faire émerger le vocabulaire propre au CA. Cette méthode s'inspire des travaux de Drouin (2002) dans lesquels sont comparées les fréquences des unités lexicales entre un CA technique et un CR de type journalistique afin de faire ressortir les spécificités lexicales du CA.

Dans cette méthode, nous postulons que le corpus d'analyse, de concert avec le corpus de référence, constitue un corpus global. À partir de cette hypothèse, nous tentons de vérifier, à l'aide du calcul des spécificités, si les fréquences du corpus d'analyse sont réparties normalement.

Le corpus d'analyse est un corpus spécialisé du domaine de l'informatique et le corpus de référence est un corpus de nature journalistique, soit le corpus *Le Monde*. Le logiciel de statistique textuelle utilisé pour isoler les spécificités, *TermoStat*, permet d'isoler trois types de formes : les spécificités positives, c'est-à-dire les formes qui sont suremployées dans le CA, les spécificités négatives, les formes qui sont sous-employées dans le CA, et les formes banales, celles dont le comportement n'est ni spécifiquement positif, ni spécifiquement négatif dans le CA. La figure 2 illustre ces trois types de formes. Seules les spécificités positives (S+) sont celles qui font l'objet de notre recherche.

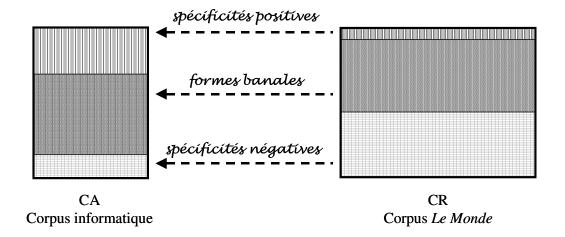


Figure 2. Illustration des spécificités de M1

### 2.2 Constitution des corpus de M1

# 2.2.1 Corpus de référence

Nous utilisons le corpus de référence comme base de comparaison afin de mesurer la déviance de notre corpus spécialisé.

Le corpus de référence de M1 est composé des articles parus dans le quotidien français *Le Monde* entre le 1<sup>er</sup> janvier 1993 et le 30 septembre 1993. Il s'agit d'un corpus de nature journalistique de 16 364 757 mots dont les textes portent sur des sujets variés. Ce corpus a été mis à la disposition des chercheurs par Jean Baudot du Groupe de recherche en sémantique, lexicologie et terminologie (GRESLET) de l'Université de Montréal. Ce corpus fait partie d'un corpus plus important, TEXTUM, utilisé pour la recherche en lexicographie par le GRESLET.

Sinclair (1996 : 10) considère qu'un corpus de référence devrait se composer de différents types de textes et donne l'exemple du corpus *The Bank of English* qui est divisé en plusieurs *genres*, dont : journaux, livres, revues, radio, prospectus et langage parlé familier, chacun des sous-corpus comptant plusieurs millions de mots et étant à leur tour divisés en de plus petits sous-corpus.

« Un corpus de référence est conçu pour fournir une information en profondeur sur une langue. Il vise à être suffisamment étendu pour représenter toutes les variétés pertinentes du langage et son

36

vocabulaire caractéristique, de manière à pouvoir servir de base à

des grammaires, des dictionnaires et d'autres usuels fiables »

(*ibid.* :10, cité dans Habert *et al.* 1997 : 144).

Bien que notre corpus de référence soit du seul type journalistique, nous

croyons qu'il pourra, par la variété et le grand nombre de ses articles, répondre à la

fonction de corpus de référence. Comme notre travail ne s'intéresse qu'à l'aspect

lexical des textes, nous sommes d'avis qu'il est moins important que notre corpus

de référence soit composé de textes dont les sources proviennent de différents

genres.

2.2.2 Corpus d'analyse

Le corpus d'analyse utilisé dans M1 est un corpus spécialisé composé de

textes portant sur l'informatique.

« Les corpus spécialisés sont limités à une situation de

communication, ou à un domaine. Parmi ces corpus, on trouve les

ensembles relevant de sous-langages que l'on trouve dans les

domaines scientifiques et techniques. »

(Habert et al. 1997: 144)

Le choix du domaine de l'informatique s'est imposé à nous pour deux

raisons. D'abord, le laboratoire ÉCLECTIK de l'Université de Montréal a déjà

réalisé de nombreux travaux portant sur ce domaine<sup>5</sup>, lesquels ont mené à

<sup>5</sup> La description de ces travaux est disponible sur le site de l'Éclectik :

http://www.ling.umontreal.ca/lhomme/eclectik.htm

l'élaboration d'un corpus informatique très intéressant pour notre recherche. La deuxième raison est celle de notre expérience personnelle, l'informatique étant notre spécialisation en tant que traductrice.

#### 2.2.2.1 Provenance des textes

Des corpus, étiquetés ou non, sont disponibles pour la recherche. Toutefois, il s'agit pour la plupart de corpus de nature dite *générale*, les corpus spécialisés étant plutôt rares. Un centre de recherche de l'Université Lancaster, *University Center for Computer Corpus Research on Language* (UCREL 2003), ainsi que l'*Association Européenne pour les Ressources Linguistiques* (ELRA 2002) dressent la liste des quelques corpus spécialisés disponibles. Parmi ceux-ci, un seul est issu du domaine de l'informatique, soit le corpus *IBM Manuals Treebank*. Ce corpus de 800 000 mots est composé de documents rédigés en anglais et ne comporte que des textes provenant de manuels d'informatique.

Dans le cadre de notre recherche, puisque nous visons à établir le vocabulaire caractéristique de l'informatique, le corpus utilisé doit être composé de textes traitant des différents sous-domaines de l'informatique. C'est pour cette raison que nous avons eu recours au corpus d'informatique du laboratoire ÉCLECTIK. Les textes qui constituent ce corpus ont été rédigés entre 1988 et 2003, ils sont de nature variée et proviennent d'ouvrages imprimés et de sites Internet spécialisés en informatique.

#### 2.2.2.2 Choix des textes

Une autre raison de notre intérêt pour le corpus informatique de l'ÉCLECTIK est qu'il répond à nos critères : les textes sont rédigés en français; ils sont de différents niveaux de langue; ils sont rédigés par différents auteurs, ce qui permet d'empêcher la surreprésentation d'un terme ou d'une formulation propre à un auteur; et finalement, ils couvrent plusieurs sous-domaines de l'informatique. Ce dernier critère vise à permettre l'élaboration d'un vocabulaire caractéristique le plus complet possible.

# 2.2.2.3 Taille du corpus

Lorsque nous avons commencé à travailler sur notre mémoire, le corpus informatique comptait environ 400 000 mots répartis en 45 textes. Nous avons alimenté le corpus pour atteindre un total de 53 textes totalisant 600 034 mots.

Selon Pearson (1998 : 51), la taille des corpus spécialisés est moins importante que celle des corpus de nature plus générale. Toutefois, les corpus spécialisés se doivent d'être très représentatifs du domaine étudié. Par conséquent, plus le corpus sera volumineux, plus il risque d'être représentatif. Elle ajoute (*ibid.* : 57) que les chercheurs qui compilent les corpus spécialisés croient, sans toutefois en avoir la certitude, que les phénomènes recherchés apparaîtront avec une fréquence suffisante dans les corpus de petite taille pour leur permettre d'obtenir des résultats qui soient satisfaisants. En outre, selon Sinclair (1996 : 9), un corpus spécialisé peut se permettre d'être beaucoup moins volumineux qu'un

corpus de référence car, en raison de la sélection rigoureuse des textes qui le composent, de son lexique spécialisé et de l'homogénéité de sa composition, il contient rapidement toutes les propriétés inhérentes au domaine spécialisé en question. Pour ces raisons, nous avons jugé que la taille de notre corpus d'informatique était suffisante pour le type de traitement auquel le corpus est soumis.

### 2.3 Présentation de la méthode 2

La deuxième méthode (M2) s'inspire des travaux de Lafon (1980) dans lesquels il comparait un certain nombre de discours de Robespierre à l'ensemble des discours pour dégager deux types de vocabulaire : d'abord le vocabulaire spécifique à chacun des discours, et ensuite le vocabulaire de base du discours robespierriste.

M2 vise à opposer des sous-ensembles d'un corpus informatique à l'ensemble du corpus. Dans cette méthode, le corpus global n'est donc pas hétérogène mais homogène.

### 2.4 Constitution des corpus de M2

# 2.4.1 Corpus de référence

Le corpus de référence utilisé dans M2 n'est pas composé de textes de nature générale comme celui de M1 mais de textes spécialisés du domaine de l'informatique. Il correspond exactement au corpus d'analyse de M1.

# 2.4.2 Corpus d'analyse

Dans M2, nous avons établi six corpus d'analyse (SC1 à SC6) représentant un sous-domaine de l'informatique. Chacun des sous-corpus est un sous-ensemble du CR.

À l'instar de Lafon, nous voulons dégager le vocabulaire spécifique à chacun des sous-corpus et le vocabulaire de base de l'ensemble des sous-corpus. Toutefois, le vocabulaire de base du domaine de l'informatique ne sera pas composé des formes banales comme l'était le vocabulaire de base du discours robespierriste mais de la totalité des formes spécifiques à chacun des sous-corpus.

### 2.4.2.1 Sous-domaines

Nous avons réparti les 53 textes composant le corpus informatique en six sous-domaines de l'informatique (les six corpus d'analyse de M2), soit *Abc de l'informatique*, *Internet*, *Logiciels*, *Matériel*, *Programmation et réseaux* et *Systèmes d'exploitation*. Tous les corpus ne comptent pas le même nombre de textes et de mots. Toutefois, le programme utilisé pour le calcul des fréquences tient compte de ces différences dans son algorithme. Le tableau I présente la taille de chacun des sous-corpus.

	Taille des corpus	
Corpus d'analyse	Nombre de textes	Nombre de mots
Abc de l'informatique (SC1)	8	116 821
Internet (SC2)	12	102 972
Logiciels (SC3)	4	78 412

	Taille des corpus	
Corpus d'analyse	Nombre de textes	Nombre de mots
Matériel (SC4)	5	41 816
Programmation et réseaux (SC5)	11	38 909
Systèmes d'exploitation (SC6)	13	221 104
Total	53	600 034

Tableau I. Taille des corpus d'analyse de M2

De même que pour M1, la comparaison a été effectuée à l'aide du logiciel *TermoStat*. Les S+ isolées pour chacun des sous-corpus représentent le vocabulaire spécifique à un sous-domaine de l'informatique alors que l'ensemble des S+ retenues devrait représenter le vocabulaire caractéristique de l'informatique en général (voir la figure 3).

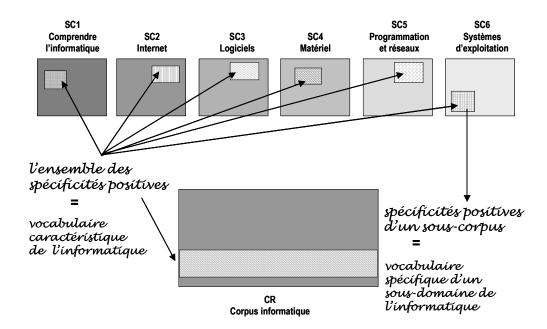


Figure 3. Illustration des spécificités positives de M2

# 2.5 Étiquetage morphosyntaxique

Afin de maximiser la précision du calcul des fréquences et des spécificités, il était important d'étiqueter et de lemmatiser les textes du corpus informatique et ceux du corpus *Le Monde*. Tous les textes ont été étiquetés à l'aide du logiciel *WinBrill*, lequel est la version Windows de l'étiqueteur de Brill. *WinBrill* a été adapté pour Windows par le laboratoire ATILF (Analyse et Traitement Informatique de la Langue Française) qui est né de l'Institut National de la Langue Française (INaLF).

Nous décrivons dans les sections qui suivent les différents traitements que doivent subir les textes avant l'étape de l'étiquetage morphosyntaxique.

### 2.5.1 Format des textes

Avant d'être étiquetés, tous les textes doivent être convertis au format ASCII. Les textes du corpus *Le Monde* et les textes originaux du corpus informatique étaient déjà dans ce format. Nous avons également converti au format ASCII les huit textes que nous avons ajoutés au corpus informatique. Afin de pouvoir mettre sur support informatique les textes dont la source était un ouvrage écrit, les premiers chercheurs à avoir alimenté le corpus informatique avaient numérisé ces textes et les avaient sauvegardés en format ASCII.

# 2.5.2 Nettoyage des textes du corpus informatique

Le nettoyage auquel a été soumis le corpus informatique consiste en la suppression de la source apparaissant au début de chaque texte ainsi que de la bibliographie, s'il y avait lieu, à la fin des textes. Tous les textes ont également été soumis au correcteur d'orthographe du logiciel *Word* afin d'éliminer les coquilles, les fautes d'orthographe et les erreurs dues à la numérisation. Voici un exemple de la source apparaissant au début de chaque texte :

```
(Page consultée le 30 mai 2002). " Mettre en place un réseau local ", en ligne. Adresse URL : http://www.iptsos.com/formation/windows/reseaulocal/reseaulocal1.p hp [resloc.txt] 1345 mots
```

# 2.5.3 Nettoyage des textes du corpus Le Monde

Le corpus *Le Monde* à l'état brut contenait des caractères de contrôle que nous avons voulu éliminer pour ne pas fausser les résultats. Le corpus se présentait comme un seul fichier ASCII comportant des millions de mots. Nous l'avons divisé en plusieurs parties, puis il nous a suffi d'ouvrir le corpus avec *TextPad* et de faire une recherche sur ces codes afin de les effacer. Nous avons réussi à éliminer tous les caractères qui se trouvent entre les codes <99><0> et <0></60> inclusivement, à l'exception du titre. Voici un exemple des caractères de contrôle :

```
<99><0>
Date 01/01/1993 Page: 13 Taille de l'article: 1 col
Titre: Les plumes de l'ange Pour un texte écrit par Pasolini en
harmoniques avec son film " Théorème ", Baudoin invente de belles
assonances dessinées
</0><60>
```

Nous n'avons pas soumis le corpus *Le Monde* au correcteur d'orthographe en raison du trop grand nombre de mots. De plus, nous avons cru qu'il était peu probable de rencontrer un grand nombre de fautes dans ce type de corpus. Les guillemets anglais ont également été éliminés des deux corpus. Nous en expliquons la raison à la section 2.5.3.1.1.

#### 2.5.3.1 Prétraitement

Afin d'être en mesure d'étiqueter les textes, un prétraitement spécifique à l'étiquetage a dû être effectué sur les deux corpus. L'étiqueteur de Brill offre une option de prétraitement mais dans la version de *WinBrill* que nous avons utilisée cette option n'est pas opérationnelle. Nous avons donc élaboré une macrocommande à l'aide du logiciel *Word* en nous fondant sur les normes requises au prétraitement fournies dans Lecomte (1998 : 28). La macro-commande effectue les opérations suivantes :

 les textes sont organisés de façon à n'avoir qu'une seule phrase par ligne;

- les signes de ponctuation sont séparés du mot précédent par un espace;
- l'apostrophe demeure avec le mot qui la précède et est séparée par un espace du mot qui la suit;
- aucune phrase ne commence par un espace;
- aucune ligne n'est vide;
- les lettres majuscules en début de phrase sont converties en minuscules.

De plus, chaque unité d'un terme complexe (*pomme de terre*, *disque dur*) a été étiquetée individuellement alors que les mots composés (*avant-midi, double-cliquer*) ont été étiquetés comme un seul mot. Voici l'exemple d'un texte ayant été soumis au prétraitement :

cependant , l' usage des ordinateurs est loin d'être limité au seul calcul numérique .

par exemple , au sein d' une entreprise , ils peuvent être utilisés à diverses fins : gestion des stocks de matières premières et de produits finis , gestion de la comptabilité , circulation de l' information entre les différents niveaux de l' entreprise par le biais d' un système de messagerie électronique , aide au travail de secrétariat par la mise à disposition de facilité pour l' édition et l' impression du courrier , de rapports , ou pour le classement et la maintenance de l' ensemble des dossiers clients , etc .

### 2.5.3.1.1 Limites du prétraitement

La macro-commande du prétraitement comporte certaines limites. D'abord, les majuscules ne sont pas très bien gérées. Lorsqu'une majuscule se trouve à une autre position qu'à la première lettre d'une phrase, elle n'est pas convertie en minuscule. Ceci résulte souvent en une erreur d'étiquetage car le mot est étiqueté comme nom propre alors que ça ne devrait pas toujours être le cas. De plus, le prétraitement ne donne pas de bons résultats pour les mots exceptionnels comportant une apostrophe tels que aujourd'hui et presqu'île. L'apostrophe étant séparée par un espace du mot qui la suit, ces formes sont scindées en deux et deviennent aujourd', hui, presqu' et île. Ainsi, de ces quatre formes une seule est analysée, soit *île*, les trois autres étant inexistantes. Une autre limite du prétraitement concerne les guillemets anglais. Ceux-ci ne sont pas reconnus par la macro-commande car il s'agit de caractères réservés à Visual Basic, le programme utilisé par Word lors de la création des macros. Ainsi, les guillemets anglais encadrant un mot ne sont pas séparés de ce mot par un espace, ce qui implique souvent une erreur d'étiquetage. Par exemple, l'étiqueteur ne parvient pas à analyser correctement un mot tel « "souris" », le mot étant en fait « souris ». Nous n'avons pu résoudre le problème de la lettre majuscule mais celui des guillemets anglais l'a été en soumettant tous les textes à l'option de recherche de TextPad et en éliminant chaque occurrence des guillemets anglais.

### 2.5.4 L'étiqueteur WinBrill

WinBrill fonctionne selon un mode d'« auto-apprentissage ». Dans la phase d'apprentissage, un échantillon de corpus est étiqueté manuellement et est soumis à un algorithme qui, entre autres, génère un lexique et une table contenant tous les mots trouvés dans le corpus avec les étiquettes qui y sont associées (L'Homme et al., 2002).

L'apprentissage se fait en deux phases. Dans la première phase, il y a apprentissage des règles lexicales qui permet d'attribuer l'étiquette la plus probable pour les mots inconnus du lexique. Dans la deuxième phase, il y a apprentissage des modèles de transformations contextuelles (Lecomte 1998 : 26).

Dans le cadre de notre recherche, nous utilisons une version de *WinBrill* déjà entraînée avec le corpus *FranText*<sup>6</sup>. Nous ignorons donc l'étape de l'apprentissage et passons directement à l'étape de l'étiquetage, laquelle se fait également en deux phases avec l'utilisation de l'information produite par le processus d'apprentissage. Dans un premier temps, le logiciel consulte son lexique et affecte à chaque mot du texte une étiquette correspondant à sa partie du discours la plus probable. Si le mot se trouve au lexique, c'est l'étiquette du lexique qui est attribuée. Si le mot ne se trouve pas au lexique, deux étiquettes par défaut sont prévues : l'étiquette NNP (nom propre) sera attribuée à tous les mots commençant par une majuscule, et NN (nom commun) sera attribuée dans tous les autres cas

<sup>6</sup> FranText est un corpus de l'ATILF de plus de 80 millions de mots, constitué surtout de textes littéraires qui s'échelonnent du XVI<sup>e</sup> au XX<sup>e</sup> siècle.

(*ibid.* : 31). Des règles lexicales sont ensuite appliquées afin d'affiner l'attribution des étiquettes par défaut. Dans un deuxième temps, le système revient sur les premières affectations, examine le contexte local à l'aide de règles contextuelles, et corrige éventuellement les étiquettes qui avaient été attribuées précédemment. À la fin de cette seconde étape, chaque mot aura reçu une étiquette correspondant à sa classe en contexte (par exemple, un mot qui avait d'abord été étiqueté comme adverbe pourra se retrouver, en contexte, étiqueté comme un nom ou un pronom) (*ibid.* : 3).

Lorsque l'étiquetage est terminé, le texte a la forme suivante :

cela/PRO:sg signifie/VCJ:sg que/SUB des/DTN:pl machines/SBC:pl clientes/SBC:pl (/( des/DTC:pl machines/SBC:pl faisant/VNCNT partie/SBC:sg du/DTC:sg réseau/SBC:sg )/) contactent/VCJ:pl un/DTN:sg serveur/SBC:sg ,/, une/DTN:sg machine/SBC:sg généralement/ADV très/ADV puissante/ADJ:sg en/PREP terme/SBC:sg de/PREP capacités/SBC:pl d'/PREP entrée-sortie/VNCFF ,/, qui/REL leur/PRV:pl fournit/VCJ:sg des/DTC:pl services/SBC:pl ./.

Chacun des mots de l'exemple précédent s'est vu attribuer une étiquette représentant sa partie du discours. Pour les catégories où il y a lieu de l'indiquer, le

nombre, précédé des deux-points, suit la partie du discours. Les étiquettes attribuées aux mots de l'exemple précédent sont expliquées ci-dessous<sup>7</sup> :

( parenthèse ouvrante

) parenthèse fermante

, virgule

. point

sg singulier

pl pluriel

ADJ adjectif (sauf participe passé)

ADV adverbe

DTC déterminant contracté

DTN déterminant de groupe nominal

PREP préposition

PRO pronom

PRV pronom supporté par le verbe

REL pronom, adjectif ou adverbe relatif

SBC substantif (nom commun)

SUB subordonnant

VCJ verbe conjugué (autre que avoir et être)

VNCFF verbe non conjugué à l'infinitif (autre que avoir et être)

VNCNT verbe non conjugué en –ant (autre que avoir et être)

<sup>7</sup> La liste complète est présentée à l'annexe 1.

.

#### 2.5.4.1 Lemmatisation

La lemmatisation est l'opération par laquelle les formes d'un texte sont réduites à leur forme canonique, c'est-à-dire leur forme non fléchie. Par exemple :

systèmes --> système;

intéressantes --> intéressant;

sauvegardait --> sauvegarder.

Pour notre recherche, il était préférable d'étiqueter les textes avec lemmatisation car le fait de regrouper les unités lexicales selon leur lemme nous permettait de comparer les véritables occurrences d'une unité lexicale (par exemple, les occurrences *exécutait*, *exécutent* et *exécuter* sont regroupées sous le lemme *exécuter*) et non seulement les formes, comme se limitent à le faire plusieurs logiciels.

Le lemmatiseur utilisé par *WinBrill* est l'analyseur flexionnel du français pour des corpus étiquetés *FLEMM*, qui est un programme *Perl* effectuant l'analyse morphologique flexionnelle de textes français préalablement étiquetés.

Les étiquettes attribuées aux mots des textes lemmatisés, en plus d'indiquer la partie du discours, donnent également les informations suivantes : le genre et le nombre pour les adjectifs, les déterminants et les participes; le nombre pour les noms; le genre, le nombre, la personne et le cas pour les pronoms; le nombre, la personne, le temps, le mode et le groupe de conjugaison pour les verbes.

L'étiquetage avec lemmatisation a la forme suivante :

cela/PRO:3p:m:s:\_/cela signifie/VCJ:{{1|3}p:s:pst:{indlsubj}|2p: s:pst:imper}/signifier:1g que/SUB/que des/DTN:\_:p/un machines/ SBC:\_:p /machine clientes/SBC:\_:p/cliente (/(des/DTC:\_:p/du machines/SBC:\_:p/machine faisant/VNCNT:m:s/faire:3g partie/ SBC:f:s/partie du/DTC:m:s/du réseau/SBC:\_:s/réseau )/)contactent/VCJ:3p:p:pst:{indlsubj}/contacter:1g un/DTN:m:s/un serveur/SBC:\_:s/serveur ,/, une/DTN:m:s/un machine/SBC: \_:s/machine généralement/ADV/ généralement très/ADV/très puissante/ADJ:f:s/puissant en/PREP/en terme/SBC:\_:s/terme de/ PREP/de capacités/SBC:\_:p/capacité d'/PREP/de entrée-sortie/ SBC:\_:s/entrée-sortie ,/, qui/REL:\_:\_/qui leur/ PRV:3p:\_:p:d/ lui fournit/VCJ:3p:s:{pstlps}:ind/fournir:2g des/DTC:\_:p/du services/ SBC:\_:p/service ./.

Le tableau II résume la structure des mots étiquetés avec lemmatisation :

Catégorie grammaticale (étiquette)	Format
verbes (VCJ)	mot fléchi/étiquette:personne:nombre:temps:mode/ lemme:groupe/
participes (VPAR, VNCNT, ADJ1PAR, ADJ2PAR, EPAR, APAR, ANCNT, ENCNT) adjectifs (ADJ) noms (SBC) déterminants (DTN, DTC) pronoms relatifs (REL)	mot fléchi/étiquette:genre:nombre/lemme/

Catégorie grammaticale (étiquette)	Format
pronoms personnels (PRO, PRV)	mot fléchi/étiquette:personne:genre:nombre:cas/lemme
autres catégories	mot/étiquette/mot

Tableau II. Structure des mots analysés après lemmatisation<sup>8</sup>

Le tableau III présente l'ensemble des valeurs possibles que peut prendre chaque trait :

Traits	Valeurs possibles	Signification
personne	1p, 2p, 3p, _	1 <sup>re</sup> pers., 2 <sup>e</sup> pers., 3 <sup>e</sup> pers., valeur non spécifiée
genre	m, f, _	masculin, féminin, valeur non spécifiée
nombre	s, p, _	singulier, pluriel, valeur non spécifiée
temps	pst, impft, fut, ps	présent, imparfait, futur, passé simple
mode	ind, subj, cond, imper	indicatif, subjonctif, conditionnel, impératif
groupe	1g, 2g, 3g	1 <sup>er</sup> groupe, 2 <sup>e</sup> groupe, 3 <sup>e</sup> groupe
cas	n, a, d, o, _	nominatif, accusatif, datif, oblique, valeur non spécifiée
Note : les analyses ambiguës sont factorisées en tant qu'ensembles disjoints délimités par		

Note : les analyses ambiguës sont factorisées en tant qu'ensembles disjoints délimités par « { » et « } », et séparés par « | ».

Tableau III. Valeurs possibles des différents traits dans les textes lemmatisés

# 2.6 Constitution des listes de référence

Une fois les corpus étiquetés, nous avons créé une liste de référence par CR. Cette liste est présentée sous forme de fichier ASCII et c'est la liste, et non le corpus, qui servira ensuite de point de comparaison. Ce fichier est beaucoup moins lourd que le corpus même. À titre d'exemple, le corpus de référence de M1 est de 300 Mo alors que la liste ne compte que 3 Mo. Elle consiste en une liste de toutes

<sup>8</sup> Les informations comprises dans les tableaux II et III ont été recueillies dans le fichier d'aide du logiciel *WinBrill* version 0.4.

les occurrences du corpus sous leur forme lemmatisée, accompagnées de la partie du discours et de la fréquence. Pour obtenir cette liste, il nous a suffi de lancer le programme Unix *build\_ref\_list.exe* conçu par Drouin (2002). Ce petit programme calcule la fréquence de chaque occurrence du corpus, selon le lemme.

Ainsi, pour M1, nous avons obtenu une liste de référence pour le corpus *Le Monde* et pour M2, nous avons obtenu six listes de référence, pour chacun des six CA. La raison pour laquelle nous devions créer six listes de référence pour M2 et non une seule est que, lors de son analyse, *TermoStat* incorpore le CA au corpus de référence. Ainsi, si nous avions créé une seule liste à partir des six sous-corpus, toutes les occurrences du corpus à analyser auraient été doublées.

Voici un échantillon d'une liste de référence, où la première colonne représente la forme lemmatisée, la deuxième colonne représente la partie du discours et la troisième colonne, la fréquence absolue des occurrences du CR :

affichage	SBC	175
affiche	SBC	561
afficher	VB	724
afficher	ADJPAR	324
affichette	SBC	26
afficheur	SBC	9
affichiste	SBC	4
affiliation	SBC	27
affilier	ADJPAR	91
affilier	VB	6
affilié	SBC	11

# 2.6.1 Regroupement d'étiquettes

Le programme *build\_ref\_list.exe* fonctionne avec les étiquettes utilisées par *WinBrill*. Toutefois, comme *WinBrill* peut attribuer différentes étiquettes pour une même catégorie, nous avons demandé à Patrick Drouin de faire une modification à son programme, laquelle permettrait de regrouper certaines étiquettes. Puisque le calcul se base sur le lemme, les indications de genre et de nombre inscrites à la suite de l'étiquette ne sont plus pertinentes. Nous présentons à la figure 4 un exemple des étiquettes pour lesquelles nous avons demandé un regroupement<sup>9</sup>.

Étiquette WinBrill	Signification	Nouvelle étiquette
ACJ:sg	« avoir » (verbe ou aux.) conjugué sing.	
ACJ:pl	« avoir » (verbe ou aux.) conjugué plur.	
ANCFF	« avoir » (verbe ou aux.) infinitif	VBA
ANCNT	« avoir » (verbe ou aux.) en -ant	VDA
APAR:sg	« avoir » (verbe ou aux.) PP verbal sing. après « avoir »	
APAR:pl	« avoir » (verbe ou aux.) PP verbal plur. après « avoir »	

Figure 4. Exemple de regroupement d'étiquettes

# 2.7 Calcul des spécificités à l'aide de TermoStat

Nous avons utilisé le logiciel *TermoStat*, conçu par Drouin dans le cadre de sa thèse, pour isoler le vocabulaire caractéristique de l'informatique, et ce pour les deux méthodes à l'étude. *TermoStat* fait appel à des techniques statistiques pour isoler les phénomènes lexicaux en opposant le lexique du CA au lexique du CR. Ainsi, les déviances significatives du comportement du lexique dans le corpus

•

<sup>&</sup>lt;sup>9</sup> La liste complète est présentée à l'annexe 2.

d'analyse sont recensées automatiquement (Drouin 2002 : 175). Ces déviances significatives sont appelées *spécificités*. Les spécificités sont dites *positives* lorsqu'une unité lexicale dans un CA a une fréquence supérieure à sa fréquence établie à partir d'observations faites dans le CR. Les spécificités sont dites *négatives* lorsque les formes sont beaucoup moins fréquentes dans le CA, et on parle de *formes banales* lorsque les formes ne se démarquent pas par leur fréquence d'un corpus à l'autre. Bien que ce logiciel permette d'isoler les termes (simples ou complexes) d'un corpus, l'utilisation que nous en faisons se limite aux spécificités <sup>10</sup>. De plus, nous nous intéressons seulement aux spécificités positives, lesquelles, d'après nous, représenteront le vocabulaire caractéristique de l'informatique.

Parmi les informations générées par *TermoStat*, nous nous intéressons aux suivantes :

Information	Description
Forme	Forme telle que recensée dans le texte
Partie du discours	Partie du discours attribuée par l'étiqueteur
Fréquence absolue	Fréquence de la forme dans le CA et dans le CR
Valeur-test	Valeur-test obtenue à partir du calcul des spécificités

Tableau IV. Information générée par TermoStat

Les algorithmes utilisés par *TermoStat* font appel à la loi normale. De plus, la *valeur-test* est utilisée pour recenser les spécificités positives.

-

<sup>10</sup> Il s'agit ici d'unités simples seulement.

#### 2.7.1 Loi normale

La loi normale est utilisée pour vérifier si un phénomène a une distribution normale dans un échantillon d'un ensemble. Drouin (2002 : 142) écrit que la loi normale permet d'identifier à la fois les formes qui sont significativement fréquentes et celles qui sont significativement peu fréquentes dans un corpus d'analyse par rapport à un corpus de référence. Nous présentons ci-après la formule de la loi normale.

#### Loi normale

$$P(k_{ia}-0.5 \le X \le k_{ia}+0.5) = P \frac{\left(k_{ia}-0.5\right) - \left(\frac{k_{a}k_{i\bullet}}{k}\right)}{\sqrt{k_{a}\left(\frac{k_{i\bullet}}{k}\right)\left[1 - \frac{k_{i\bullet}}{k}\right]}} \le N(0,1) \le \frac{\left(k_{ia}+0.5\right) - \left(\frac{k_{a}k_{i\bullet}}{k}\right)}{\sqrt{k_{a}\left(\frac{k_{i\bullet}}{k}\right)\left[1 - \frac{k_{i\bullet}}{k}\right]}}$$

où:

P correspond à la probabilité,

N correspond à la loi normale centrée réduite,

k correspond à la fréquence,

 $k_a$  correspond à la taille du corpus de référence (en nombre de mots),

 $k_{ia}$  correspond à la forme i dans le CA,

 $k_{ir}$  correspond à la forme i dans le CR,

 $k_{i\bullet}$  correspond à la fréquence i dans le CA et le CR (les formes étant numérotées de 1 à n).

#### 2.7.2 Valeurs-tests

« Lebart et Salem (1994 : 317) décrivent la valeur-test comme une quantité permettant d'apprécier la signification de la position d'un élément illustratif sur un axe factoriel. En d'autres termes, il s'agit de la position d'un point sur la courbe normale. »

(Drouin 2002 : 147)

Afin de ne retenir que les formes très significatives, nous adoption le seuil de 3,09 pour les valeurs-tests. Ce seuil permet d'assurer qu'il n'y a qu'une chance sur 1 000 que la fréquence observée dans le texte du corpus d'analyse soit due au hasard. Un seuil de 1,96 correspond à une probabilité de 5 % alors qu'un seuil de 2,33 correspond à une probabilité de 1 %.

### 2.8 Gestion des spécificités

Dans le but de gérer de façon efficace les spécificités obtenues à l'aide de *TermoStat*, nous avons constitué deux bases de données, une pour M1 et une autre pour M2, à l'aide du logiciel *Access*. *Access* est un logiciel qui permet d'exploiter un volume important de données, de créer des liens entre différentes entrées et d'effectuer des requêtes en combinant plusieurs éléments de recherche.

### 2.8.1 Constitution des bases de données

Les deux bases de données comportent des différences mais elles possèdent également plusieurs champs communs. Les deux sont composées de quatre champs correspondant aux informations obtenues par *TermoStat*, soit la forme telle que

recensée dans le texte par *TermoStat*, la partie du discours attribuée par l'étiqueteur, la fréquence de la forme dans le corpus d'analyse et la valeur obtenue par le calcul des spécificités. À ces quatre champs, trois autres s'ajoutent :

ID: numéro séquentiel de l'entrée attribué automatiquement par Access;

Code: code que nous avons attribué à chacune des formes recensées pour en permettre une meilleure analyse. Ils se décrivent comme suit:

> ACR acronyme, abréviation et certains mots spécifiques aux commandes ou à la programmation, tel que « autoexec »

ANG mot anglais

forme qui répond à nos critères d'acceptation<sup>11</sup> BON

CAR chiffre arabe ou romain et date

ERR forme inexistante et forme brachygraphique composée de lettres, de chiffres et de symboles, telles *b*2, *d*21, {*f*10, etc.

FAU forme comportant une faute d'orthographe

<sup>&</sup>lt;sup>11</sup> Les critères d'acceptation sont donnés à la section 2.8.2.

LEM forme pour laquelle il y a eu une erreur lors de la lemmatisation

SYM symbole et ponctuation

Étiquetage : ce champ comporte les valeurs « oui » ou « non » selon que l'unité lexicale correspondant au code BON est bien étiquetée ou mal étiquetée.

La base de données M2 comporte un champ supplémentaire appelé *Sous-corpus* dans lequel est affiché le nom du sous-corpus d'où provient la S+. L'ajout de ce champ permet d'inclure les six corpus d'analyse de M2 dans une seule base de données. Par la suite, nous avons créé des requêtes nous permettant d'obtenir les informations propres à chacun des sous-domaines de l'informatique.

La figure 5 présente un exemple d'une requête faite à partir de la base de données M2. Dans cette requête, nous avons demandé l'affichage des valeurs-tests supérieures ou égales à 3,09, des fréquences absolues supérieures à 2 et du sous-corpus *Internet*.

ID	Code	Forme	Etiquetage	PdeD	Fréq abs	Valeur-test	Sous-corpus
21336	BON	intel	non	ADJ	11	4,37473	internet
21342	BON	intellectuel	oui	ADJ	6	3,47404	internet
21444	BON	interconnexion	oui	SBC	57	13,5674	internet
21450	BON	intercontinental	oui	ADJ	7	4,77666	internet
21491	BON	intérêt	oui	SBC	27	3,4948	internet
21558	BON	internatif	non	SBP	6	4,35251	internet
21562	BON	international	oui	ADJ	64	14,3728	internet
21569	BON	internationalisat	oui	SBC	26	9,57877	internet
21574	FAU	internaut		PRO	5	3,88406	internet
21577	BON	internaute	oui	SBC	32	10,7287	internet
21579	LEM	internautes		SBP	4	3,3543	internet
21591	BON	internet	oui	SBC	786	42,8164	internet
21593	BON	internet	non	C00	17	7,53884	internet
21612	BON	interopérabilité	oui	SBC	7	3,60019	internet
21666	BON	interréseau	oui	SBC	4	3,3543	internet
21736	BON	intranet	oui	SBP	9	5,53024	internet
21757	BON	introduction	oui	SBC	16	3,33372	internet
21784	BON	intuitif	oui	ADJ	6	3,13612	internet
21876	BON	investissement	oui	SBC	30	6,78227	internet
21930	ACR	ip		SBP	171	9,11294	internet
21975	ACR	irc		SBP	25	9,64279	internet
22079	ERR	issir		ADJPAR	14	3,47936	internet
22091	BON	italie	oui	SBP	5	3.37512	internet

Figure 5. Exemple d'une requête de la base de données M2

## 2.8.2 Critères d'acceptation

Afin de ne tenir compte dans notre analyse que de certaines formes répondant à des critères bien précis que nous décrivons dans la présente section, nous avons révisé manuellement la catégorisation faite par l'étiqueteur.

Avant d'observer de plus près les résultats obtenus, nous avons décidé d'éliminer certaines formes de notre analyse. Nous avons alors attribué des codes aux spécificités positives. Nous avons codé les formes utiles à notre analyse ainsi que celles dont nous ne tenons pas compte. Nous marquons ces dernières à des fins statistiques seulement. Les formes dont nous ne voulons pas tenir compte dans notre analyse sont celles dont les codes sont ACR, ANG, CAR, ERR, FAU, LEM, et SYM.

Les formes retenues sont les S+, c'est-à-dire les formes dont la valeur-test est supérieure ou égale à 3,09 et la fréquence absolue supérieure deux 12. Nous ne tenons pas compte des formes dont la fréquence absolue est inférieure à trois car la plupart de ces formes ne répondent pas à nos critères. De plus, nous avons cru qu'une forme possédant une fréquence si peu élevée présentait moins d'intérêt pour notre recherche. Parmi les S+, les formes retenues pour l'analyse sont celles dont valeur du champ *Code* est « BON » et la valeur du champ *Étiquetage* est « oui ». Pour se voir attribuer le code « BON », une S+ doit respecter les critères suivants :

- il doit s'agir d'un mot français;
- le mot ne doit pas être formé par un procédé de réduction (sigle, acronyme, abréviation et autre forme brachygraphique);
- la forme ne doit pas être un chiffre ou une date;
- la forme ne doit pas être un symbole ou une ponctuation;
- la forme doit être écrite sans faute d'orthographe;
- la forme ne doit contenir aucune erreur d'étiquetage ou de lemmatisation.

<sup>12</sup> Nous ne tenons pas compte des formes dont la fréquence absolue est inférieure à trois car la plupart de ces formes ne répondent pas à nos critères. De plus, nous avons cru qu'une forme

possédant une fréquence si peu élevée présentait moins d'intérêt pour notre recherche.

## 2.9 Résultats du calcul des spécificités

Une fois tous les champs de la base de données complétés, nous avons observé les données en différentes étapes. Nous illustrons dans les paragraphes qui suivent les résultats obtenus.

Le nombre total de formes (S+, S- et formes banales) obtenues s'élève à 22 421 pour M1. Les formes de M2 sont réparties comme suit : SC1 compte 7 277 formes, SC2 en compte 8 850, SC3 : 6 254, SC4 : 4 949, SC5 : 4 382, et SC6 en compte 10 581. Le tableau V rend compte de ces résultats et précise également le nombre de spécificités positives comprises dans chacune des méthodes. Les spécificités positives feront l'objet de plus amples observations dans les sections qui suivent.

	Méthode 1			Méth	ode 2		
	wethode i	SC1	SC2	SC3	SC4	SC5	SC6
Nb total de formes	22 421	7 277	8 850	6 254	4 949	4 382	10 581
Formes banales et spécificités négatives (formes dont F < 3 et VT < 3,09)	<u>17 428</u>	6 818	7 956	5 582	4 312	3 890	10 059
Spécificités positives (formes dont F > 2 et VT ≥ 3,09)	4 993	459	894	672	637	492	522

Tableau V. Nombre de formes comprises dans M1 et M2 avant nettoyage

# 2.9.1 Spécificités positives ignorées

Les spécificités positives ignorées dans notre analyse des résultats sont celles dont le code est ACR, ANG, CAR, ERR, FAU, LEM ou SYM. À titre d'information, nous rendons compte de ces S+ dans le tableau VI.

	Méthode 1	Méthode 2						
	Wicthode 1	SC1	SC2	SC3	SC4	SC5	SC6	
Nombre de ACR	329	29	67	20	26	19	44	
Nombre de ANG	803	16	55	32	41	67	68	
Nombre de CAR	119	15	15	14	41	5	6	
Nombre de ERR	992	25	89	137	71	60	52	
Nombre de FAU	102	4	4	14	2	5	1	
Nombre de LEM	106	2	16	34	0	7	13	
Nombre de SYM	33	8	7	7	11_	9	9	
Total	2 484	99	253	258	192	172	193	

Tableau VI. Nombre de spécificités positives à ignorer

Pour chacun des codes apparaissant au tableau VI, nous donnons dans le tableau suivant quelques exemples de spécificités positives dont nous ne tenons pas compte dans l'analyse des résultats.

Code	Forme	Fréq. abs. CA	Valeur-test
	ansi	8	6,23
ACR	autoexec	15	5,25
	sgbd	23	10,79
	сору	41	12,75
ANG	command	22	7,00
	prompt	13	7,42
	5	104	3,90
CAR	8-21-85	7	5,76
	vii	3	3,32

Code	Forme	Fréq. abs. CA	Valeur-test
	t~	4	4,06
ERR	lotir	9	6,66
	{f10	3	3,32
	activitÉ	3	3,32
FAU	disquett	24	5,77
	donnees	3	3,32
	déplacez	6	5,25
LEM	prenons	7	5,76
	commandes	6	4,73
	}	363	41,43
SYM	?	165	9,75
	<	34	3,10

Tableau VII. Exemples de formes ignorées dans l'analyse des résultats

# 2.9.2 Spécificités positives retenues

Les S+ sur lesquelles porte notre analyse sont celles dont la valeur du champ *Code* est « BON » (les formes qui répondent à nos critères d'acceptation). De ces formes ont été déduites les formes qui avaient été mal étiquetées par *WinBrill*. Nous présentons dans le tableau VIII le nombre de spécificités positives sur lequel portera notre analyse.

	Méthode 1	Méthode 2					
	Methode 1	SC1	SC2	SC3	SC4	SC5	SC6
Nombre de BON	2 509	360	641	414	445	320	329
Nombre de BON mal étiquetés	341	13_	46	47	25	35	27
Nombre de formes qui correspondent à nos critères	2 168	347_	595	367	420	285	302

Tableau VIII. Nombre de spécificités positives à analyser

# 2.9.3 Résumé des résultats obtenus suite au calcul des spécificités

En résumé, M1 compte 22 421 spécificités dont 4 993 spécificités positives parmi lesquelles 2 484 ont été éliminées car elles ne correspondaient pas à nos critères d'acceptation. Il reste donc 2 509 spécificités positives dont 341 étaient mal étiquetées. Finalement, 2 168 spécificités positives sont analysées au chapitre suivant. Le tableau IX illustre ces résultats et également ceux obtenus pour M2.

	Méthode 1	Méthode 2					
	Methode i	SC1	SC2	SC3	SC4	SC5	SC6
Nb total de formes	22 421	7 277	8 850	6 254	4 949	4 382	10 581
moins : Nb de formes dont F < 3 et VT < 3,09	<u>17 428</u>	6 818	7 956	5 582	4 312	3 890	10 059
Nb de formes dont F > 2 et VT ≥ 3,09	4 993	459	894	672	637	492	522
moins : Total code à éliminer	2 484	99	253	258	192	172	193
= : Nombre de BON	2 509	360	641	414	445	320	329
moins : Nombre de BON mal étiquetés	341	13_	46_	47	25	35	27
Nombre de formes qui correspondent à nos critères	2 168	347	595	367	420	285	302

Tableau IX. Résumé des résultats obtenus suite au calcul des spécificités

# **CHAPITRE 3: ANALYSE ET VALIDATION DES RÉSULTATS**

Notre analyse porte sur les S+ dont les valeurs-tests sont supérieures ou égales à 3,09. Ce seuil permet d'assurer qu'il y a moins d'une chance sur 1 000 que la fréquence observée dans le corpus d'analyse soit le fruit du hasard. Dans notre recherche, nous sommes parties du postulat que les S+ obtenues à l'aide des deux méthodes représentent le vocabulaire caractéristique de l'informatique. Dans la section 3.1.1, nous observons les S+ obtenues à l'aide de M1. Dans la section 3.2.1, nous observons le vocabulaire spécifique à chacun des sous-domaines de l'informatique obtenu à l'aide de M2. Nous portons ensuite notre attention sur le vocabulaire de l'ensemble des sous-corpus car l'union des formes obtenues pour chacun des SC devrait représenter le vocabulaire caractéristique de l'informatique pour M2. Finalement, nous validons nos résultats à l'aide de deux dictionnaires, puis nous comparons le vocabulaire caractéristique obtenu pour chacune des deux méthodes.

#### 3.1 Analyse des S+ obtenues avec M1

#### 3.1.1 Vocabulaire spécifique de l'informatique selon M1

Nous avons vu au chapitre précédent qu'après l'application de nos critères d'acceptation, 2 168 S+ ont été retenues parmi les 22 421 formes recensées dans notre corpus d'analyse (voir tableau IX de la section 2.9.3). Le tableau X présente la liste des 25 formes les plus spécifiques parmi les 2 168 S+ retenues<sup>13</sup>. Elles

-

<sup>&</sup>lt;sup>13</sup> La liste complète des S+ obtenues à l'aide de M1 est présentée à l'annexe 3.

apparaissent en ordre décroissant de valeur-test puisque c'est cette dernière qui détermine la spécificité d'une forme.

Forme	PdeD	Fréq. abs. CR	Fréq. abs. CA	Valeur-test
fichier	SBC	193	3956	315,54
commande	SBC	1020	1902	177,53
internet	SBC	2	1102	170,62
serveur	SBC	97	1166	168,29
utiliser	VB	1708	1993	163,32
utilisateur	SBC	119	1117	162,82
logiciel	SBC	194	1166	161,68
option	SBC	745	1486	158,92
ordinateur	SBC	508	1283	153,73
système	SBC	5360	2699	143,55
configuration	SBC	101	845	140,70
répertoire	SBC	328	1003	139,75
touche	SBC	348	855	124,92
disquette	SBC	22	609	124,51
windows	SBP	27	613	124,41
votre	DT	1689	1365	121,36
disque	SBC	959	1093	120,17
réseau	SBC	2289	1511	119,17
imprimante	SBC	16	537	117,29
mémoire	SBC	1672	1240	112,42
clavier	SBC	127	579	111,16
donnée	SBC	641	840	109,16
linux	NNP	0	442	108,08
connexion	SBC	45	486	108,07
adresse	SBC	372	701	107,94

Tableau X. Liste des 25 premières S+ obtenues à l'aide de M1

Dans la liste qui précède, nous notons une forte majorité de substantifs en comparaison des autres catégories grammaticales. En effet, on y observe 21 noms communs, deux noms propres, un verbe et un déterminant. Dans la section

suivante, nous décrivons la répartition des catégories grammaticales pour la totalité des S+ retenues.

Selon M1, la plupart des formes présentes dans le tableau précédent font partie du vocabulaire spécialisé de l'informatique. Effectivement, ces formes semblent toutes être très représentatives du domaine de l'informatique, sauf une, soit le déterminant *votre*. On peut également s'interroger sur la présence du verbe *utiliser*.

En élargissant notre observation à la liste complète, nous notons que plus la valeur-test diminue, plus la concentration de termes spécialisés (*serveur*, *configuration*, *disquette*, *imprimante*) tend à diminuer pour faire place au *vocabulaire fonctionnel* de l'informatique, c'est-à-dire les mots typiques à une « manière d'écrire » en informatique et indispensables à la compréhension des textes. Ces unités lexicales s'articulent autour du vocabulaire spécialisé (*fonctionner*, *nombre*, *version*, *afficher*, *particularité*, etc.).

Afin de déterminer si nous sommes réellement en présence du vocabulaire de l'informatique, nous avons adopté une méthode de validation. Nous expliquons à la section 3.3 la démarche entreprise pour valider notre vocabulaire.

# 3.1.2 Répartition des catégories grammaticales

Le tableau suivant présente la répartition des catégories grammaticales pour la totalité des S+ retenues dans M1.

Catégorie grammaticale	Proportion (%)
Adjectifs	17,98
Adverbes	5,12
Noms communs	44,67
Noms propres	5,99
Participes passés	9,96
Pronoms	0,41
Verbes	13,79
Mots grammaticaux	2,07
Coordonnants	0,14
Déterminants	0,37
Interjections	0,14
Particules	0,18
Préfixes	0,28
Prépositions	0,55
Relatifs (pro. ou adj.)	0,09
Subordonnants	0,32

Tableau XI. Répartition des catégories grammaticales – M1

Comme nous l'avions noté pour les 25 premières S+, les noms communs sont majoritaires lorsque nous observons la totalité des S+. Ils couvrent presque la moitié des formes, soit 44,67 %. De plus, nous avons remarqué que la proportion de noms communs par rapport aux autres catégories grammaticales, bien qu'encore majoritaire, est moins importante en fin de liste qu'en début de liste. La deuxième catégorie grammaticale la mieux représentée est celle des adjectifs avec 17,98 % des S+. Suivent ensuite les verbes (13,79 %), les participes passés (9,96 %), les noms propres (5,99 %), et finalement les adverbes (5,12 %). Les noms communs, les adjectifs, les verbes et les participes passés représentent aussi bien le vocabulaire spécialisé (connexion, interface, virtuel, binaire, cliquer, configurer, formaté, connecté) que le vocabulaire fonctionnel (information, puissance, libre, obsolète, insérer, créer, utilisé, géré) de l'informatique. Les noms propres comprennent en très grande majorité des termes spécifiques à l'informatique (Unix,

Adobe, Telnet), alors que les adverbes font partie exclusive du vocabulaire fonctionnel (manuellement, simplement, normalement). Les coordonnants, déterminants, interjections, particules, préfixes, prépositions, pronoms, relatifs (adjectifs et pronoms) et subordonnants sont des catégories pour lesquelles nous avons recensé très peu de S+. Nous avons groupé ces dernières catégories grammaticales sous le générique « mots grammaticaux ».

Nous constatons que le vocabulaire spécialisé de l'informatique semble surtout être représenté par les substantifs alors que le vocabulaire fonctionnel de l'informatique semble être représenté par toutes les catégories grammaticales, et surtout par les substantifs, les adjectifs et les verbes. Ensemble, ces deux types de vocabulaire forment le vocabulaire caractéristique de l'informatique.

#### 3.2 Analyse des S+ obtenues avec M2

Dans la présente section, nous présentons d'abord les S+ obtenues pour chacun des SC, lesquelles représentent le vocabulaire spécifique à chacun des sous-domaines de l'informatique que nous avons définis. L'union des S+ de chacun des six SC représente le vocabulaire caractéristique de l'informatique, que nous allons décrire à la fin de la présente section.

## 3.2.1 Vocabulaire spécifique aux sous-domaines de l'informatique

Nous présentons dans cette section le vocabulaire spécifique à chacun des sous-domaines de l'informatique représentés par les six sous-corpus de M2. Nous

rappelons que le nombre de S+ est différent d'un SC à l'autre. Les sous-corpus 1 à 6 comptent 347, 595, 367, 420, 285 et 302 spécificités positives respectivement. Le vocabulaire des sous-corpus ne sera pas soumis à une validation de façon individuelle mais l'ensemble du vocabulaire des sous-corpus le sera.

Les tableaux XII à XVII présentés dans les pages qui suivent rendent compte des 25 premières S+ retenues pour chacun des sous-domaines de l'informatique<sup>14</sup>. À la suite de chaque tableau, nous faisons part de nos observations sur les 25 premières S+ de chaque sous-domaine, puis à la fin de la section, nous présentons les généralisations que nous pouvons formuler sur l'ensemble des SC. À la section 3.2.2, nous exposons les phénomènes observés pour la totalité des spécificités.

SC1 (Abc de l'informatique)									
Forme	PdeD	Fréq. abs. CR	Fréq. abs. CA	Valeur-test					
mémoire	SBC	503	737	31,90					
ordinateur	SBC	720	563	19,81					
disque	SBC	611	482	18,44					
unité	SBC	327	302	16,19					
appuyer	VB	100	168	15,97					
texte	SBC	331	298	15,83					
touche	SBC	484	371	15,83					
disquette	SBC	329	280	14,79					
lecteur	SBC	223	226	14,78					
un	DT	13 125	4219	14,67					
microprocesseur	SBC	12	80	14,55					
processeur	SBC	125	150	13,12					
wordperfect	SBP	2	49	12,24					
dur	ADJ	218	182	11,75					
carte-mère	SBC	6	48	11,41					

<sup>&</sup>lt;sup>14</sup> La liste complète des S+ obtenues à l'aide de M2 est présentée à l'annexe 4.

SC1 (Abc de l'informatique)								
Forme	PdeD	Fréq. abs. CR	Fréq. abs. CA	Valeur-test				
bit	SBC	209	173	11,38				
exploitation	SBC	259	196	11,37				
principal	ADJ	235	183	11,21				
servir	VB	154	143	11,14				
écran	SBC	487	286	11,01				
magnétique	ADJ	22	59	10,76				
temps	SBC	348	222	10,50				
segment	SBC	8	40	9,87				
vif	ADJ	39	65	9,84				
rigide	ADJ	58	58	9,65				

Tableau XII. Liste des 25 premières S+ obtenues pour SC1

Nous constatons que la liste des 25 premières formes recensées pour le sous-domaine *Abc de l'informatique* sont effectivement susceptibles de se trouver dans les textes s'adressant à des non-initiés de l'informatique et dans les textes traitant de l'informatique en général (*lecteur*, *ordinateur*, *disquette*, etc.). Nous avons observé dans cette liste 16 noms communs, cinq adjectifs, deux verbes, un nom propre et un déterminant. On peut se questionner sur la présence du déterminant *un* et du verbe *servir* qui semblent de nature plus générale que les autres formes. Quant à la présence des adjectifs *dur* et *vif*, elle peut s'expliquer par la présence des substantifs *disque* et *mémoire* avec lesquels ils forment des termes complexes.

Le tableau suivant présente les 25 premières S+ obtenues pour SC2 :

SC2 (Internet)								
Forme	PdeD	Fréq. abs. CR	Fréq. abs. CR Fréq. abs. CA					
internet <sup>15</sup>	SBC	316	786	42,82				
site	SBC	64	306	30,00				
recherche	SBC	109	346	29,88				
web <sup>16</sup>	SBC	71	303	29,38				
électronique	ADJ	97	202	20,68				
france	SBP	42	154	20,40				
service	SBC	311	303	18,88				
moteur	SBC	22	115	18,53				
courrier	SBC	45	136	18,50				
message	SBC	178	215	17,63				
infrastructure	SBC	19	89	16,05				
document	SBC	243	228	16,04				
public	ADJ	40	100	15,19				
langue	SBC	23	82	14,76				
opérateur	SBC	62	111	14,62				
technologie	SBC	87	127	14,59				
français	ADJ	40	93	14,39				
international	ADJ	8	64	14,37				
zope	SBP	1	53	14,04				
serveur	SBC	764	402	13,90				
interconnexion	SBC	7	57	13,57				
formulaire	SBC	7	55	13,28				
fibre	SBC	2	49	13,28				
national	ADJ	11	57	12,96				
netscape	SBP	3	48	12,94				

Tableau XIII. Liste des 25 premières S+ obtenues pour SC2

Les 25 premières formes recensées dans SC2 sont très représentatives du sous-domaine de l'Internet. On y rencontre des termes tels *internet*, *site*, *recherche*, *web*, etc. Notons également les noms propres *Zope* et *Netscape*, le premier étant un

1

<sup>&</sup>lt;sup>15</sup> Nous voulons expliquer la présence de ce mot dans la liste malgré le fait que l'étiqueteur lui ait attribué l'étiquette de « nom commun ». Nous avons décidé de ne pas considérer cette attribution comme une erreur d'étiquetage puisque ce mot est considéré comme nom propre ET comme nom commun selon certaines sources dont le *Petit Larousse illustré*.

<sup>16</sup> *idem* 

serveur Web et le second un navigateur Web. On trouve dans cette liste 17 noms communs, cinq adjectifs et trois noms propres. On peut s'interroger sur la présence du nom propre *France*, des adjectifs *français*, *national* et *international* et du nom commun *langue*. La présence de deux textes faisant partie du SC *Internet* « Problèmes reliés à l'internationalisation d'Internet » et « Internet : pour quelle société? » pourrait être une partie de la réponse. On devine également la présence de termes complexes dans le corpus tels *courrier/formulaire électronique*, *langue française* et *moteur de recherche*.

Le tableau qui suit présente les 25 premières S+ retenues dans SC3 :

SC3 (Logiciels)						
Forme	PdeD	Fréq. abs. CR	Fréq. abs. CA	Valeur-test		
étendue	SBC	3	75	19,87		
cellule	SBC	34	94	18,50		
disquette	SBC	361	248	18,46		
unité	SBC	382	247	17,76		
phase	SBC	39	82	16,27		
presser	VB	9	51	15,02		
logiciel	SBC	823	343	15,01		
examiner	VB	11	50	14,50		
figure	SBC	64	80	13,76		
enregistrement	SBC	96	94	13,54		
micro-informatique	ADJ	3	37	13,49		
colonne	SBC	68	78	13,15		
programmation	SBC	125	101	12,82		
projet	SBC	69	76	12,77		
micro-ordinateur	SBC	51	65	12,47		
équipe	SBC	24	47	12,03		
entité	SBC	16	41	11,93		
progiciel	SBC	34	34	11,59		
test	SBC	106	84	11,56		
erreur	SBC	200	116	11,30		

SC3 (Logiciels)							
Forme	PdeD Fréq. abs. CR Fréq. abs. CA						
chiffrier	SBC	1	25	11,30			
façon	SBC	231	126	11,29			
entrer	VB	106	82	11,27			
donnée	SBC	612	228	10,92			
sous-répertoire	SBC	40	50	10,82			

Tableau XIV. Liste des 25 premières S+ obtenues pour SC3

La liste des 25 premières formes spécifiques au sous-domaine *Logiciels* comporte 21 noms communs, trois verbes et un adjectif. Nous y observons des termes relatifs à la notion de « logiciel » en général, tels *progiciel* et *logiciel*, ou spécifique à un type de logiciel en particulier, tels *cellule*, *donnée* et *colonne* qui appartiennent soit à un logiciel de chiffrier électronique, soit à un logiciel de base de données. Toutefois, certaines formes semblent de nature plus générale, telles *phase*, *façon* et *examiner*. On devine la présence des termes complexes suivants : *unité de disquette* et *enregistrement de données*.

Le tableau qui suit présente les 25 premières S+ retenues dans SC4 :

SC4 (Matériel)						
Forme	PdeD	Valeur-test				
imprimante	SBC	337	200	26,92		
papier	SBC	77	104	25,95		
epson	SBP	2	44	22,79		
laser	SBC	41	62	20,55		
imprimant	ADJ	29	52	19,56		
feuille	SBC	61	67	19,55		
étiquette	SBC	21	43	18,25		
aiguille	SBC	19	41	18,00		
imprimer	VB	104	73	17,39		

	SC4 (Matériel)						
Forme	orme PdeD Fréq. abs. CR Fréq. abs. 0		Fréq. abs. CA	Valeur-test			
ruban	SBC	20	37	16,54			
continu	ADJ	22	37	16,19			
trait	SBC	19	35	16,06			
broche	SBC	3	24	15,96			
matriciel	ADJ	10	29	15,84			
micro	SBC	31	40	15,75			
impression	SBC	182	84	15,35			
borne	SBC	9	27	15,35			
cartouche	SBC	14	29	14,93			
consigne	SBC	0	16	13,79			
cordon	SBC	2	17	13,36			
réglette	SBC	0	14	12,84			
amstrad	SBP	0	14	12,84			
tabulation	SBC	20	26	12,63			
mais	COO	1039	200	12,52			
maître	SBC	15	23	12,37			

Tableau XV. Liste des 25 premières S+ obtenues pour SC4

La liste des 25 premières S+ du sous-domaine *Matériel* compte 18 noms communs, deux noms propres, trois adjectifs, un verbe et un coordonnant (mais), dont la présence peut susciter un questionnement. Cette liste comporte surtout des termes relatifs aux imprimantes. Ceci s'explique de la façon suivante : bien qu'il n'y ait qu'un seul texte traitant des imprimantes sur les cinq textes composant SC4, à lui seul ce texte couvre les trois quarts du SC, soit 30 049 mots sur 41 816. Notons également la présence du nom propre *Amstrad* qui peut paraître inappropriée, mais il s'agit en fait d'une firme ayant fabriqué des ordinateurs au milieu des années 1980. On note la présence d'unités lexicales susceptibles de composer des termes complexes tels *imprimante* et *laser*, *imprimante* et *aiguille*,

*imprimante* et *matriciel*, et également *trait* et *continu*, dont on peut questionner la présence dans cette liste.

Le tableau qui suit présente les 25 premières S+ retenues dans SC5 :

SC5 (Programmation et réseaux)						
Forme	PdeD	Fréq. abs. CR	Fréq. abs. CA	Valeur-test		
implémentation	SBC	15	55	23,53		
algorithme	SBC	42	61	20,90		
synchronisation	SBC	48	63	20,69		
remplissage	SBC	6	38	20,58		
contrat	SBC	14	44	20,56		
réseau	SBC	1219	292	19,68		
réservoir	SBC	0	28	19,12		
mobile	SBC	1	25	17,63		
intranet	SBC	2	24	16,85		
architecture	SBC	133	72	16,03		
migration	SBC	7	25	15,62		
topologie	SBC	6	23	15,09		
problème	SBC	353	110	14,58		
serveur	SBC	965	201	14,46		
java	SBP	30	34	14,46		
recoupement	SBC	1	17	14,24		
mobile	ADJ	15	26	14,06		
langage	SBC	361	104	13,43		
entreprise	SBC	163	66	13,20		
client	SBC	333	96	12,90		
coder	VB	13	22	12,81		
plate-forme	SBC	24	27	12,80		
classe	SBC	69	41	12,54		
méthodologie	SBC	6	17	12,33		
œuvre	SBC	1	13	12,21		

Tableau XVI. Liste des 25 premières S+ obtenues pour SC5

La liste des 25 premières S+ du sous-domaine *Programmation et réseaux* comporte 22 noms communs, un nom propre, un adjectif et un verbe. Nous avons groupé les sous-domaines *programmation* et *réseaux* en un seul sous-domaine car les notions s'y rapportant étaient souvent abordées dans les mêmes textes. On y rencontre des termes ayant trait à la programmation tels que *algorithme*, *langage*, *java*, *coder*, etc. et des termes relatifs aux réseaux tel que *implémentation*, *réseau*, *intranet*, *serveur*, etc. Bien que cette liste présente des termes spécifiques au sous-domaine, on y trouve également plusieurs formes plus générales (*contrat*, *problème*, *recoupement*, *entreprise*, *classe* et *méthodologie*). De plus, la présence de l'adjectif *mobile* et du nom *œuvre* peut s'expliquer du fait que ce sont des unités comprises dans des termes complexes (*mise en œuvre*, *informatique mobile*, *plate-forme mobile*, etc.).

Voici maintenant les 25 premières S+ retenues dans SC6.

SC6 (Systèmes d'exploitation)						
Forme	PdeD	Fréq. abs. CR	Fréq. abs. CA	Valeur-test		
option	SBC	228	1258	30,76		
fichier	SBC	1568	2388	24,92		
recommander	ADJPAR	20	496	22,40		
noyau	SBC	22	438	20,82		
linux	NNP	20	422	20,50		
configuration	SBC	185	660	20,02		
utiliser	VB	731	1262	19,85		
activer	VB	54 368		17,22		
се	DT	3329	3158	16,33		
police	SBC	65	328	15,38		
script	SBC	23	255	15,20		
réponse	SBC	120	393	15,02		
connexion	SBC	115	371	14,51		

SC6 (Systèmes d'exploitation)						
Forme	PdeD	Fréq. abs. CR	Fréq. abs. CA	Valeur-test		
utiliser	ADJPAR	360	644	14,50		
paquet	SBC	53	279	14,30		
n'	ADV	983	1169	13,70		
suivant	ADJ	265	512	13,55		
répertoire	SBC	385	618	13,16		
permettre	VB	1275	1349	12,67		
périphérique	ADJ	90	285	12,63		
fonctionnalité	SBC	71	259	12,59		
support	SBC	163	365	12,39		
installation	SBC	67	246	12,29		
paramètre	SBC	132	323	12,15		
être	VBE	10065	7260	11,95		

Tableau XVII. Liste des 25 premières S+ obtenues pour SC6

Le sous-domaine *Systèmes d'exploitation* est celui qui comporte le plus grand nombre de catégories grammaticales différentes pour les 25 premières S+. On y rencontre 14 noms communs, quatre verbes, deux participes passés, deux adjectifs, un nom propre, un déterminant et un adverbe. Notons qu'il s'agit du seul SC où sont présents participes passés et adverbe dans la liste des 25 premières S+. On trouve dans ce SC des termes appartement au vocabulaire des systèmes d'exploitation tels que *noyau*, *Linux*, *connexion*, *installation*, etc. On rencontre également des mots appartenant à la langue générale ou au vocabulaire fonctionnel de l'informatique tels que l'adverbe n', le déterminant *ce*, le verbe *être*, le participe passé *utilisé* et l'adjectif *suivant*. Nous remarquons aussi que les fréquences absolues des formes de SC6 sont beaucoup plus élevées que les fréquences absolues observées dans les autres sous-corpus. Ceci s'explique par la taille de SC6 qui est beaucoup plus importante que celle des autres sous-corpus. Notons

également la présence d'unités lexicales susceptibles de faire partie d'associations privilégiées ou de termes complexes tels que *activer* la *connexion*, *installation* recommandée et paramètre de noyau.

Suite à l'observation des six tableaux précédents, des généralisations peuvent être dégagées :

- la catégorie des substantifs est celle pour laquelle nous obtenons le plus grand nombre de S+;
- pour chacun des sous-corpus, la majorité des formes recensées fait partie du vocabulaire spécialisé propre à chaque sous-domaine et la minorité appartient au vocabulaire fonctionnel de l'informatique;
- les fréquences sont élevées. Aucun SC ne présente une fréquence absolue inférieure à 13 et la quasi-totalité des fréquences sont supérieures à ce nombre;
- dans tous les SC, nous avons repéré des unités dont on peut supposer qu'elles sont des composantes d'un terme complexe. Par exemple, dans SC1 on note la présence des substantifs disque et mémoire et des adjectifs dur et vif qui sont les unités formant les termes complexes disque dur et mémoire vive. Dans les autres SC, nous remarquons la présence des unités simples formulaire et électronique, site et web, imprimante et laser, langage et java, etc.

Ceci nous laisse supposer que dans un travail où ne seraient pas pris en compte seulement les termes simples, ces termes complexes auraient été repérés;

• nous avons également observé que certains termes ont été empruntés à la langue générale. Par exemple, *routine* et *mémoire* pour SC1, *site* et *moteur* pour SC2, *cellule* et *fonction* pour SC3, *aiguille* et *cartouche* pour SC4, *architecture* et *migration* pour SC5 et finalement, *noyau* et *démon* pour SC6.

### 3.2.2 Répartition des catégories grammaticales

Les tableaux de la section précédente nous ont permis de constater qu'en début de liste, les substantifs sont les plus représentés parmi toutes les catégories grammaticales. Ce fait est également vrai pour la totalité des S+. De plus, après avoir étendu nos observations à la totalité des S+, nous avons noté que tous les sous-corpus tendent à avoir une proportion de formes appartenant à un vocabulaire spécialisé plus élevée en début de liste et une proportion de formes appartenant à un vocabulaire fonctionnel de l'informatique plus élevée en fin de liste. Par ailleurs, nous avons remarqué que ce fait est plus marquant dans le sous-corpus *Matériel* que dans les autres sous-corpus. Le tableau XVIII présente la répartition des catégories grammaticales dans chaque SC.

Catégorie	Proportion (%)						
grammaticale	SC1	SC2	SC3	SC4	SC5	SC6	
Adjectifs	11,53	18,32	15,80	16,90	15,09	11,92	
Adverbes	4,61	1,51	4,63	5,24	2,11	10,26	
Noms communs	51,87	61,01	58,58	51,90	54,04	46,36	
Noms propres	2,59	6,05	1,63	3,33	6,67	3,64	
Participes passés	8,65	4,54	5,72	6,67	5,61	6,95	
Pronoms	1,44		0,27	1,19	0,35	0,66	
Verbes	14,70	6,55	10,90	11,67	12,98	15,56	
Mots grammaticaux	4,61	2,02	2,45	3,10	3,16	4,64	
Coordonnants	1,15	0,17	0,54	0,71	0,35		
Déterminants	0,58	0,84	1,09	0,48	1,40	1,32	
Interjections			0,27	0,48			
Particules	0,29				0,35	0,33	
Préfixes	0,29					0,33	
Prépositions	1,73	1,01	0,27	0,95	0,70	1,32	
Relatifs (pro. ou adj.)			0,27		0,35		
Subordonnants	0,58			0,48		1,32	

Tableau XVIII. Répartition des catégories grammaticales – M2

De l'observation du tableau XVIII, nous constatons les faits suivants :

- les noms communs ont la plus forte proportion de S+ dans tous les SC;
- la deuxième catégorie grammaticale la plus représentée est celle des adjectifs pour les sous-corpus SC2, SC3, SC4 et SC5, et les verbes pour les sous-corpus SC1 et SC6. Ces deux catégories comprennent des formes appartenant aux vocabulaires spécialisé et fonctionnel de l'informatique;
- le sous-corpus *Internet* est celui qui se démarque le plus des autres sous-corpus; c'est celui qui a la plus forte proportion de noms

communs et d'adjectifs mais c'est également celui où les adverbes, les participes passés, les pronoms, les mots grammaticaux et particulièrement les verbes sont les moins nombreux. Il y 50 % moins de verbes dans SC2 que dans tout autre corpus. Sur les 39 verbes de SC2, nous considérons que seulement sept d'entre eux appartiennent au vocabulaire spécialisé du domaine de l'informatique, soit *naviguer*, *télécharger*, *indexer*, *explorer*, *héberger*, *encoder* et *sécuriser*;

- aucun sous-corpus ne comporte plus de 9 % de participes passés. Les formes présentes dans cette catégorie grammaticale se trouvent surtout dans le vocabulaire fonctionnel de l'informatique mais quelques-unes font également partie du vocabulaire spécialisé;
- les adverbes de tous les sous-corpus appartiennent au vocabulaire fonctionnel de l'informatique. Ils sont peu nombreux dans tous les sous-corpus sauf dans SC6 où ils représentent 10 % des S+;
  - les noms propres sont également peu nombreux dans tous les sous-corpus ; pour chacun ils représentent moins de 7 % des S+. Dans tous les sous-corpus, à l'exception de SC3, on rencontre des noms propres très spécifiques du sous-domaine en question : *Pentium* et *Wordperfect* pour SC1, *Netscape* et *Yahoo* pour SC2, *Canon* et *Apple* pour SC4, *Oracle* et *Java* pour SC5, et enfin, *Windows* et *Unix* pour SC6. Seulement six noms propres font partie de SC3 et un seul est

caractéristique du sous-domaine *Logiciels*, soit *Cobol*. Les autres formes sont des noms propres de personne : *Paul*, *Armand*, *Robert*, *Sandra* et *Artur*.

Pour donner suite à ces observations, il serait intéressant d'étudier plus en profondeur les textes composant les différents SC afin de comprendre pourquoi le comportement des S+ diffère d'un SC à l'autre. Toutefois, ce travail dépasse les limites de notre mémoire.

# 3.2.3 Vocabulaire spécifique de l'informatique selon M2

Nous avons présenté à la section 3.2.1 des listes partielles de S+ apparaissant en ordre décroissant de valeur-test, la première S+ de la liste étant la forme se démarquant le plus dans le corpus d'analyse. Ces formes font partie du vocabulaire spécifique de chaque sous-domaine de l'informatique. La présente section rend compte, pour M2, du vocabulaire caractéristique de l'informatique dans son ensemble. Il s'agit des S+ de tous les SC réunis. Le tableau qui suit présente les 25 premières S+ de l'ensemble des SC.

Forme	PdeD	Fréq. abs. CR	Fréq. abs. CA	V-test	sc
internet	SBC	316	786	42,82	internet
mémoire	SBC	503	737	31,90	abc
option	SBC	228	1258	30,76	syst. expl.
site	SBC	64	306	30,00	internet
recherche	SBC	109	346	29,88	internet
web	SBC	71	303	29,38	internet
imprimante	SBC	337	200	26,92	matériel
papier	SBC	77	104	25,95	matériel

Forme	PdeD	Fréq. abs. CR	Fréq. abs. CA	V-test	sc
fichier	SBC	1568	2388	24,92	syst. expl.
implémentation	SBC	15	55	23,53	prog. rés.
epson	SBP	2	44	22,79	matériel
recommander	ADJPAR	20	496	22,40	syst. expl.
algorithme	SBC	42	61	20,90	prog. rés.
noyau	SBC	22	438	20,82	syst. expl.
synchronisation	SBC	48	63	20,69	prog. rés.
électronique	ADJ	97	202	20,68	internet
remplissage	SBC	6	38	20,58	prog. rés.
contrat	SBC	14	44	20,56	prog. rés.
laser	SBC	41	62	20,55	matériel
linux	NNP	20	422	20,50	syst. expl.
france	SBP	42	154	20,40	internet
configuration	SBC	185	660	20,01	syst. expl
étendue	SBC	3	75	19,87	logiciels
utiliser	VB	731	1262	19,85	syst. expl.
ordinateur	SBC	720	563	19,81	abc

Tableau XIX. Liste des 25 premières S+ obtenues pour l'ensemble des SC

Les formes en caractères gras, au nombre de neuf, sont celles qui se trouvent aussi parmi les 25 premières S+ de M1. Grâce à une commande SQL lancée à partir du logiciel *Access*, nous avons été en mesure de calculer la proportion de S+ communes à M1 et M2 pour la totalité des S+. Ainsi, il y a 59 % des formes retenues dans M1 qui font partie de la liste des formes retenues dans M2. Après avoir observé ces résultats, nous avons noté que plus la valeur-test d'une spécificité est élevée, plus il y a de chance que cette forme soit commune aux deux méthodes. En effet, sur les 25 formes les plus spécifiques de M2 (tableau XIX), 23 formes font partie des S+ de M1 (en observant l'ensemble des S+ de M1). Les deux formes de la liste qui ne se trouvent pas sur la liste totale des S+ de M1 sont *contrat* et *France*. Par ailleurs, sur les 25 formes les moins spécifiques de M2, nous

avons observé qu'à peine un peu plus de la moitié des formes (14 S+) se trouvaient parmi les S+ de M1. Nous avons calculé le pourcentage des formes communes aux deux méthodes à titre d'information seulement car notre analyse ne s'intéresse pas à savoir quels sont les termes communs aux deux méthodes mais plutôt à connaître quelle méthode est la plus apte à faire émerger le vocabulaire propre au domaine de l'informatique. Dans les sections qui suivent, nous rendons compte de la démarche adoptée afin de valider les S+ retenues dans chacune des deux méthodes.

#### 3.3 Validation des S+

Afin de valider le vocabulaire issu de nos deux méthodes de comparaison de corpus, nous avons eu recours à deux dictionnaires spécialisés de l'informatique, soit *Le Bilingue de l'informatique* (Collin) et le *Dictionnaire français-anglais d'informatique* (Ginguay).

Collin est un dictionnaire français-anglais, anglais-français publié chez Peter Collin. Dans le cadre de notre travail, nous avons utilisé la nouvelle édition revue et corrigée de 1996. Ce dictionnaire comprend, outre les mots du langage courant employés dans le quotidien de l'informatique et de la micro-informatique, le vocabulaire des ordinateurs (matériel, logiciels, programmation, périphériques, réseaux, etc.), celui des innovations plus récentes (CD ROM, messagerie électronique, Internet, etc.) et celui des applications multiples se rapportant au traitement des données (télécommunications, infographie, PAO, etc.).

Ginguay est un dictionnaire français-anglais traitant de la bureautique, de la télématique et de la micro-informatique (programmation, matériel, logiciels, applications courantes). Nous avons utilisé la sixième édition, 2<sup>e</sup> tirage avec mise à jour de 1998.

Nous avons choisi ces deux dictionnaires car tous deux répertorient le vocabulaire propre aux différents sous-domaines définis à partir de notre corpus informatique<sup>17</sup>. De plus, tout comme le vocabulaire isolé à l'aide de nos méthodes, les entrées répertoriées dans les deux dictionnaires correspondent à une unité simple. Toutefois, lorsque cette unité simple peut se combiner à une ou plusieurs autres unités pour former un terme complexe, ce nouveau terme est inscrit sous cette entrée. Par exemple, sous l'entrée adressage de Collin, on note les termes complexes adressage abrégé, adressage absolu, adressage binaire, etc. Enfin, bien que les unités simples abrégé, absolu et binaire font également l'objet d'une entrée, tous les adjectifs membres de termes complexes ne sont pas nécessairement répertoriés dans les dictionnaires. C'est le cas des adjectifs compilé et compressé qui se combinent pourtant avec un grand nombre de formes nominales.

Puisqu'il est difficile, voire impossible, de trouver un moyen de validation qui soit parfait, nous partons du postulat que la vérification avec les dictionnaires spécialisés est la meilleure façon d'entériner nos résultats et de voir dans quelle mesure les résultats peuvent être utilisés dans la pratique terminologique. De plus, nous n'avons pas cru nécessaire de valider la totalité du vocabulaire que nous avons

<sup>&</sup>lt;sup>17</sup> Nous avions envisagé l'utilisation de banques de terminologie mais pour des raisons techniques cette validation a été mise de côté.

isolé à l'aide de nos deux méthodes. Ainsi, nous avons validé les mots commençant par les lettres A, C et P. Notre choix s'est porté sur ces trois lettres car ce sont par elles que débutent le plus grand nombre de mots faisant partie de notre vocabulaire. De plus, nous estimons que les résultats obtenus à partir d'une telle validation seront représentatifs de l'ensemble du vocabulaire isolé au moyen de nos méthodes.

La validation consiste en une vérification manuelle de notre vocabulaire caractéristique de l'informatique avec le vocabulaire répertorié dans les deux dictionnaires. Tous les mots des deux dictionnaires dont la première lettre est *A*, *C* ou *P* ont été saisis à l'ordinateur dans un fichier *Excel*. Puis, toujours de façon manuelle, les mots recensées par M1 et par M2 ont été alignés avec ceux des dictionnaires dans deux fichiers distincts. Nous voulons préciser que les acronymes répertoriés dans les dictionnaires n'ont pas été saisis car ceux-ci n'avaient pas été retenus selon nos critères d'acceptation des S+. La figure 6 présente un échantillon du fichier de validation dans lequel nous pouvons observer l'alignement des S+ avec les termes des dictionnaires *Ginguay* et *Collin*<sup>18</sup>.

\_

<sup>&</sup>lt;sup>18</sup> La liste complète des termes répertoriés dans *Ginguay* débutant par les lettres *A*, *C* et *P* est présentée à l'annexe 5 et celle des termes répertoriés dans *Collin* est présentée à l'annexe 6. Sont également indiquées dans ces deux annexes, les S+ qui sont présentes dans les dictionnaires.

S+ retenues	PdeD	Ginguay	Collin
adobe	SBP		Adobe
adressable	ADJ	adressable	adressable
adressage	SBC	adressage	adressage
adresse	SBC	adresse	adresse
aération	SBC	aération	
affectation	SBC	affectation	affectation
affecter	VB	affecter	affecter
affichage	SBC	affichage	affichage
affiche	SBC		affiche
afficher	VB	afficher	afficher
afficher	ADJPAR		
afficheur	SBC	afficheur	

Figure 6. Échantillon du fichier de validation – M1

#### 3.3.1 Résultats de la validation – M1

Une fois notre vocabulaire aligné avec celui de *Ginguay* et celui de *Collin*, nous avons calculé quelle en était la proportion validée dans les dictionnaires. Les résultats de la validation pour M1 sont présentés dans le tableau XX.

Le *bruit* est le nom donné aux formes retenues dans notre vocabulaire caractéristique qui ne sont pas répertoriées dans le dictionnaire, alors que le *silence* est le nom donné aux unités lexicales figurant dans un dictionnaire mais qui ne font pas partie de notre vocabulaire caractéristique.

	Nombre de	S+au die	ctionnaire	Bruit	Silence
	formes	(nombre)	(%)	(%)	(%)
S+	646				
Ginguay	1 290	346	53,56	46,44	73,16
Collin	1 327	390	60,37	39,63	70,45

Tableau XX. Résultats de la validation - M1

Parmi la liste des 2 168 S+ que nous avons analysées, 646 d'entre elles ont été soumises à une validation, soit les S+ débutant par les lettres *A*, *C* ou *P*. Les termes commençant par ces trois lettres répertoriés dans les dictionnaires sont au nombre de 1 290 dans *Ginguay* et de 1 327 dans *Collin*. Après avoir aligné les S+ avec les termes des deux dictionnaires tel que montré à la figure 6, nous avons été en mesure de valider 346 S+ dans *Ginguay*, soit 53,56% des S+, et 390 dans *Collin*, soit 60,37 % des S+. Nous avons obtenu 46,44 % de bruit et 73,16 % de silence pour la validation avec *Ginguay* contre 39,63 % de bruit et 70,45 % de silence pour la validation avec *Collin*. Notre vocabulaire est légèrement plus près du vocabulaire répertorié dans *Collin* que de celui répertorié dans *Ginguay*.

Par ailleurs, nous avons voulu voir quelle incidence aurait eu un seuil de 1,96 pour les valeurs-tests plutôt qu'un seuil de 3,09. Précisons qu'avec un seuil de 1,96, il y a une chance sur 100 que les fréquences observées soient dues au hasard. Le calcul effectué pour un seuil de 1,96 a été fait pour les mots commençant par la lettre A seulement. Suite à ce nouveau calcul, nous avons pu noter que le bruit augmentait (passant de 44,58 % à 61,94 % pour *Ginguay* et de 46,27 % à 53,73 % pour *Collin*) alors que le silence diminuait (passant de 77,34 % à 74,88 % pour *Ginguay* et de 74,07 % à 69,38 % pour *Collin*). Ces nouveaux résultats justifient notre choix initial d'avoir opté pour un seuil de 3,09.

Nous pouvons fournir quelques explications concernant le silence : les mots répertoriés dans les dictionnaires ne figurant pas dans notre vocabulaire caractéristique se trouvent quand même pour la plupart dans notre corpus. En fait,

si ces mots ne se trouvent pas parmi les S+ c'est qu'ils font partie des formes banales ou des spécificités négatives. Prenons l'exemple du nom absence et du verbe parler, lesquels figurent au Ginguay. Il est permis de s'interroger sur le statut terminologique de ces mots et sur les raisons pour lesquelles les lexicographes les ont retenus. Suite à l'application de la technique des spécificités à notre corpus informatique, le nom absence est classé parmi les formes banales avec une valeurtest de -1,68, alors que le verbe parler se trouve parmi les spécificités négatives avec une valeur-test de -4,77. Le silence s'expliquerait également en raison des erreurs dues à l'étiquetage. En effet, plusieurs mots ont été étiquetés comme noms propres en raison du prétraitement exigé par l'étiqueteur WinBrill, lequel éprouve de la difficulté à gérer les majuscules. Évidemment, les mots qui ont été étiquetés « nom propre » par erreur n'ont pas été lemmatisés. Puisque nous avons éliminé de notre analyse toutes les erreurs d'étiquetage, certains des mots qui ont été laissés de côté auraient certainement fait partie de notre vocabulaire caractéristique s'ils n'avaient pas été écartés de notre analyse.

Nous croyons que le bruit obtenu peut s'expliquer du fait que les dictionnaires sont composés essentiellement de formes nominales, verbales et adjectivales, alors que les S+ isolées par nos méthodes présentent un certain pourcentage de formes appartenant à d'autres catégories grammaticales telles que participe passé, adverbe, préposition, etc. (voir le tableau XXI). Une autre explication serait qu'avec notre méthode les formes isolées sont susceptibles de faire partie soit de la tête, soit du modifieur d'un terme complexe alors que les

dictionnaires ne répertorient souvent que la tête, le modifieur se trouvant sous l'entrée de celle-ci.

Catégorie	S+	Ginguay	Collin		
grammaticale	(lettres A, C et P)				
Noms communs	44,43 %	63,07 %	55,98 %		
Adjectifs	16,56 %	19,78 %	25,00 %		
Verbes	13,93 %	16,21 %	14,85 %		
Autres	25,08 %	0,93 %	4,17 %		

Tableau XXI. Répartition des catégories grammaticales selon la source

Les S+ comptent 25,08 % de formes de la catégorie *Autres*, c'est-à-dire toutes les catégories grammaticales autres que celles des noms communs, des adjectifs et des verbes. Moins de 1 % des termes présents dans *Ginguay* font partie de la catégorie « Autres » alors qu'il y a 4,17 % des termes de cette catégorie dans *Collin*. Ceci pourrait expliquer le fait qu'un plus grand nombre de S+ soit présent dans *Collin* que dans *Ginguay*.

Après avoir noté que la catégorie *Autres* était beaucoup plus importante dans notre vocabulaire que dans celui des dictionnaires, nous avons voulu calculer à nouveau la proportion de notre vocabulaire validée dans les dictionnaires mais seulement pour les trois catégories *Noms communs*, *Adjectifs* et *Verbes*.

	Nombre de	S+au dictionnaire		Bruit	Silence
	formes	(nombre)	(%)	(%)	(%)
S+	486				
Ginguay	1 223	323	66,46	33,54	73,59
Collin	1 185	340	69,96	30,04	71,31

Tableau XXII. Résultats de la validation ignorant la catégorie Autres

En ne conservant que les catégories grammaticales des noms communs, verbes et adjectifs, nous observons les faits suivants :

- le nombre de S+ passe de 646 à 486;
- la proportion de notre vocabulaire se trouvant dans les dictionnaires a subi une hausse de 9,9 % dans *Ginguay* et de 9,6 % dans *Collin*;
- le bruit a diminué de 12,9 % dans *Ginguay* et de 9,5 % dans *Collin*;
- la proportion de silence a subi une légère hausse dans les deux cas, soit 0,4 % dans *Ginguay* et 0,7 % dans *Collin*.

#### 3.3.2 Résultats de la validation - M2

Nous présentons pour M2 la proportion de S+ qui a été validée dans les deux dictionnaires. Les données présentées sont celles des six sous-corpus une fois ces corpus regroupés selon les modalités décrites à la section 3.2.3.

	Nombre de	S+ au di	ctionnaire	Bruit	Silence	
	formes	(nombre)	(%)	(%)	(%)	
S+	673					
Ginguay	1 290	328	48,74	51,26	74,57	
Collin	1 327	371	55,13	44,87	72,04	

Tableau XXIII. Résultats de la validation – M2

Nous avons soumis les S+ de M2 dont la première lettre était A, C ou P à la validation. Les résultats obtenus indiquent que 48,74 % des S+ sont répertoriées dans Ginguay et 55,13 % sont répertoriées dans Collin. Nous avons obtenu 51,26 % de bruit et 74,57 % de silence pour la validation avec Ginguay contre 44,87 % de bruit et 72,04 % de silence pour la validation avec Collin. Pour M2 comme pour M1, notre vocabulaire est plus près de celui de Collin que de celui de Ginguay.

### 3.3.3 Comparaison de M1 et M2 quant aux résultats de la validation

En partant du postulat que la vérification avec les dictionnaires spécialisés est la meilleure façon d'entériner nos résultats, la comparaison des deux méthodes nous permet de réaliser que les résultats de M1 sont meilleurs que ceux obtenus avec M2, ou du moins sont-ils plus près d'une certaine pratique lexicographique. Le tableau XXIV présente la comparaison des résultats.

	S+ au dic	tionnaire	Br	uit	Silence		
	M1	M2	M1	M2	M1	M2	
Ginguay	53,56 %	48,74 %	46,44 %	51,26 %	73,18 %	74,57 %	
Meilleur résultat	+4,82 %		-4,82 %		-1,39 %		

	S+ au dic	tionnaire	Br	uit	Silence		
	M1	M2	M1	M2	M1	M2	
Collin	60,37 %	55,13 %	39,63 %	44,87 %	70,61 %	72,04 %	
Meilleur résultat	+5,24 %		-5,24 %		-1,43 %		

Tableau XXIV. Comparaison des résultats de la validation entre M1 et M2

L'analyse de la validation avec *Ginguay* indique qu'après avoir comparé les S+ de M1 et de M2, le nombre de S+ au dictionnaire est supérieur de 4,82 % dans M1, et on note 4,82 % moins de bruit et 1,39 % moins de silence dans M1 que dans M2. La validation avec *Collin* donne sensiblement les même résultats; le nombre de S+ dans ce dictionnaire est supérieur de 5,24 % dans M1, et il y a 5,24 % moins de bruit et 1,43 % moins de silence dans M1 que dans M2.

En tenant compte de la méthode de validation choisie, M1 (un corpus spécialisé est opposé à un corpus de nature journalistique) est plus près d'une certaine pratique lexicographique que M2 (des sous-corpus spécialisés sont opposés à l'ensemble du corpus) car les formes retenues dans M1 sont présentes en plus grand nombre dans les dictionnaires que les formes retenues dans M2. De plus, les proportions de bruit et de silence sont plus faibles dans M1 que dans M2.

Comme nous l'avons constaté à la section 3.2.3, les listes produites pour M1 et pour M2 sont fort différentes. Dans un travail qui ferait suite à celui-ci, il serait intéressant de combiner les listes produites par les deux méthodes afin de vérifier de combien augmenterait la proportion de S+ présentes dans les dictionnaires.

#### CONCLUSION

L'objectif du présent travail était d'isoler de façon automatique le vocabulaire propre au domaine de l'informatique. Deux méthodes fondées sur la comparaison de corpus et la statistique lexicale, plus précisément sur la technique des spécificités, ont été mises à contribution dans le cadre de cette recherche. Dans la première méthode, un corpus journalistique (le corpus de référence) a été opposé à un corpus spécialisé du domaine de l'informatique (le corpus d'analyse). Dans la deuxième méthode, nous avons opposé le même corpus informatique, qui était utilisé cette fois comme corpus de référence, à des sous-ensembles de ce corpus, chaque sous-ensemble représentant un corpus d'analyse.

Afin d'atteindre notre objectif, nous avons étudié les travaux des auteurs qui se sont penchés sur la statistique lexicale et dont l'objectif principal était de dégager les formes hautement spécifiques d'un échantillon de corpus. Nous avons également examiné différents logiciels d'analyse lexicale de corpus textuels, notamment *TermoStat*, élaboré par Drouin (2002), qui est le logiciel adopté par nos deux méthodes de comparaison de corpus.

Dans l'exposé de la méthodologie, nous avons défini les principaux concepts liés à notre travail, soit les notions de « corpus de référence », de « corpus d'analyse », de « listes de référence », de « spécificités positives » et de « calcul des spécificités ».

La méthode 1 de notre recherche, effectuée sur un corpus global hétérogène, bien qu'elle ait été inspirée des travaux de Drouin, a ceci d'original qu'elle vise à isoler l'ensemble du vocabulaire de l'informatique, c'est-à-dire toutes les parties du discours, alors que Drouin ne s'intéressait qu'à la terminologie présentée sous forme nominale ou adjectivale.

La méthode 2 de notre recherche tire son originalité du fait que les six souscorpus d'analyse sont issus d'un même domaine technique, soit l'informatique, où chaque sous-corpus représente un sous-domaine de cette spécialité. Jusqu'à maintenant, les travaux en LSP effectués sur des corpus homogènes et fondés sur la technique des spécificités étaient menés sur des corpus combinant plusieurs disciplines.

Ajoutons qu'au meilleur de notre connaissance, la technique des spécificités n'avait pas encore été utilisée pour isoler le vocabulaire caractéristique du domaine de l'informatique.

Pour chacune des méthodes, nous avons étiqueté les corpus à l'aide du logiciel *WinBrill*, en ayant préalablement activé l'option de lemmatisation. Par la suite, nous avons dressé une liste de référence, laquelle a été utilisée comme point de comparaison. La comparaison des corpus a été effectuée à l'aide du logiciel *TermoStat*, qui fait appel à la loi normale. Nous avons isolé les spécificités positives dont la valeur-test était supérieure ou égales à 3,09, ce seuil signifiant qu'il y a une chance sur 1 000 que les fréquences observées soient le fruit du hasard. Nous avons également établi un seuil de 2 quant à la valeur absolue, ce qui

nous a permis d'éliminer un grand nombre de mots ne correspondant pas à nos critères.

Afin de procéder à l'analyse des résultats, nous avons importé nos données dans le logiciel *Access*. Nous avons alors défini des critères d'acceptation quant aux spécificités à retenir. Ainsi, n'ont pas été retenus les mots anglais, les acronymes, les chiffres, les abréviations, les symboles, les ponctuations et également les mots comportant des erreurs d'étiquetage ou de lemmatisation.

L'analyse des résultats de M1 nous a permis d'observer que les S+ dont les valeurs-tests étaient les plus élevées étaient principalement des substantifs faisant partie du vocabulaire spécialisé de l'informatique, et que plus la valeur-test diminuait, s'ajoutaient aux formes nominales des formes adjectivales et verbales faisant partie d'un vocabulaire fonctionnel de l'informatique. Les noms communs, les adjectifs, les verbes et les participes passés comportent autant de termes appartenant au vocabulaire spécialisé que de mots appartenant au vocabulaire fonctionnel. Les noms propres sont en majorité des termes très spécifiques et tous les adverbes font partie du vocabulaire fonctionnel. Enfin, les mots grammaticaux forment un faible pourcentage de la totalité des catégories grammaticales, soit 2.07 %.

L'analyse des résultats de M2 nous a permis d'observer les résultats pour chacun des sous-domaines de l'informatique que nous avions définis. Nous avons d'abord concentré notre analyse sur les 25 premières spécificités de chaque sous-corpus. Ainsi, nous avons noté que pour chacun d'eux, les substantifs étaient en

plus grand nombre que les autres catégories grammaticales et qu'ils étaient très représentatifs du domaine de l'informatique. Nous avons également noté la présence d'unités lexicales susceptibles de composer des termes complexes. De plus, chaque sous-corpus comportait une minorité de formes dont on pouvait questionner la présence.

Pour M2 comme pour M1, lorsque nos observations portaient sur la totalité des S+, nous avons pu noter que tous les SC tendaient à comprendre une proportion plus élevée de termes spécialisés en début de liste qu'en fin de liste et une proportion plus élevée de termes fonctionnels en fin de liste qu'en début de liste.

Par ailleurs, pour la plupart des sous-corpus, nous avons noté des particularités quant aux catégories grammaticales, entre autres :

- le sous-domaine Abc de l'informatique est celui qui a le plus grand nombre de participes passés;
- le sous-domaine portant sur l'Internet est celui qui se démarque le plus des autres SC en ce sens qu'il a la plus forte proportion de noms communs et d'adjectifs, mais la plus faible proportion d'adverbes, de participes passés, de pronoms, de mots grammaticaux et particulièrement de verbes;
- le sous-domaine Logiciels est celui dont le nombre de noms propres est le moins élevé;

- les sous-domaines *Matériel* et *Programmation et réseaux* ne se démarquent pas particulièrement des autres SC ; et
- le sous-domaine des systèmes d'exploitation est celui où les adverbes et les verbes sont les plus nombreux et où les noms communs sont en moins grand nombre.

Dans un travail qui donnerait suite à la présente recherche, il serait sans doute intéressant d'observer plus en détail les textes constituant les corpus afin d'être en mesure de mieux interpréter le comportement des formes propres à chaque sous-domaine.

Nous avons calculé la proportion de S+ communes à M1 et M2 et nous avons constaté avec étonnement que seulement 59 % des S+ de M1 sont également des S+ de M2. Par ailleurs, nous avons constaté que plus la valeur-test d'une S+ est élevée, plus il y a de chances que cette S+ soit commune aux deux méthodes.

On voudra peut-être, dans un prolongement à ce travail, vérifier l'influence que pourrait avoir sur les résultats le fait de combiner le vocabulaire retenu par chacune des deux méthodes.

Enfin, afin de valider notre vocabulaire, nous avons eu recours à deux dictionnaires de l'informatique, soit *Le Bilingue de l'informatique* de Collin et le *Dictionnaire français-anglais d'informatique* de Ginguay. Nous avons validé les

mots dont la première lettre était *A*, *C* ou *P* en supposant que les résultats seraient représentatifs de l'ensemble du vocabulaire.

À la lumière des résultats de la validation, nous avons constaté que la première méthode à l'étude était plus près d'une certaine pratique lexicographique que la deuxième méthode. En effet, M1 est la méthode dans laquelle le plus grand nombre de S+ figure aux dictionnaires. C'est avec cette méthode également que le bruit et le silence se trouvent en plus faible proportion.

Nous avons interprété la présence du bruit du fait que les dictionnaires sont composés principalement de formes nominales, verbales et adjectivales, alors que notre vocabulaire comporte des formes appartenant à un plus grand nombre de catégories grammaticales, et également du fait que les deux unités lexicales faisant partie d'un terme complexe ne se trouvent pas nécessairement toutes deux dans les dictionnaires. En effet, dans de nombreux cas, seule la tête du terme complexe fait l'objet d'une entrée dans les dictionnaires.

En ce qui concerne le silence, nous avons expliqué sa présence du fait que certains termes des dictionnaires, bien qu'ils ne se trouvent pas dans notre vocabulaire, se trouvent toutefois dans notre corpus parmi les formes banales ou les spécificités négatives. Par ailleurs, nous croyons que la proportion de silence aurait été moins élevée n'eût été des erreurs d'étiquetage.

Il vaudrait sans doute la peine de refaire cette étude pour un domaine spécialisé d'une autre nature. En effet, les méthodes, bien qu'elles aient été testées sur un corpus d'informatique, peuvent s'appliquer à d'autres domaines spécialisés. Nous sommes d'avis que la ligne de démarcation entre vocabulaire spécialisé et vocabulaire fonctionnel serait plus nette dans un domaine où il y a une plus forte proportion de termes hautement spécialisés, par exemple dans le domaine médical, que dans un domaine qui est de plus en plus accessible, comme celui de l'informatique.

Nous espérons que ce travail fera l'objet de développements ultérieurs car de nouvelles avenues sont envisageables :

- validation du vocabulaire avec des banques de terminologie puisque les pratiques ne sont pas les mêmes;
- analyse qualitative des formes retenues (retour au contexte);
- adaptation à d'autres langues pour lesquelles on dispose des mêmes outils;
- extraction de néologismes (opposition d'un corpus spécialisé très récent à un corpus de même nature moins récent);

Enfin, pour élaborer un dictionnaire de base de l'informatique qui soit le plus complet possible, il serait nécessaire d'élargir la recherche aux termes complexes.

#### **BIBLIOGRAPHIE**

- BEAUCHEMIN, Normand et Pierre MARTEL (1979): *Vocabulaire fondamental du québécois parlé : index de fréquence*. Coll. « Recherches sociolinguistiques dans la région de Sherbrooke : Document de travail », xvi + 253 p.
- BINDI, Remo, CALZOLARI, Nicoletta, MONACHINI, Monica, PIRRELLI, Vito et Antonio ZAMPOLLI (1994): «Corpora and Computational Lexica: Integration of Different Methodologies of Lexical Knowledge Acquisition», *Literary and Linguistic Computing*, vol. 9, n° 1, pp. 29-46.
- BINON, Jean et Serge VERLINDE (1998): « Le *Dictionnaire d'apprentissage du français des affaires*: un nouveau type de dictionnaire d'apprentissage multilingue axé sur la production », *Langage et l'homme*, vol. 33, n° 4, pp. 437-449.
- BOLASCO, Sergio (2000): « TALTAC : un environnement pour l'exploitation de ressources statistiques et linguistiques dans l'analyse textuelle. Un exemple d'application au discours politique », JADT 2000, DSGSSAR Università di Roma, La Sapienza.
  - http://www.cavi.univ-paris3.fr/lexicometrica/jadt/jadt2000/pdf/80/80.pdf
- BRUNET, Étienne (2001): « Le logiciel Hyperbase », Institut national de la langue française (Nice, France). <a href="http://www.uottawa.ca/academic/arts/astrolabe/articles/art0017.htm/Hyperbase0">http://www.uottawa.ca/academic/arts/astrolabe/articles/art0017.htm/Hyperbase0</a> <a href="http://www.uottawa.ca/academic/arts/astrolabe/articles/art0017.htm/Hyperbase0">http://www.uottawa.ca/academic/arts/astrolabe/articles/art0017.htm/Hyperbase0</a>
- CABRÉ, Maria Teresa (1998): *La terminologie : théorie, méthode et applications*, traduit du catalan et adapté par Monique C. Cormier et John Humbley, Ottawa/Paris, Les Presses de l'Université d'Ottawa/André Colin, 322 p.
- CHETOUANI, Lamria (1989): « Un nouveau type de vocabulaire de base : le VGES : de la quantification à l'usage pédagogique » *Cahiers de Lexicologie*, vol. 54, pp. 83-103.
- COLLIN, S. M. H., LAURENDEAU, Françoise et Bernard MOUGET (1996): *Le Bilingue de l'informatique : dictionnaire français-anglais, anglais-français.* Coll. « Peter Collins », Middlesex, Peter Collin Pub, 494 p.
- DAUZAT, Albert (1967): *Tableau de la langue française; origine, évolution, structure actuelle,* Coll. « Petite collection Payot », Paris, Payot, 295 p.

- DROUIN, Patrick (2002): Acquisition automatique des termes: l'utilisation des pivots lexicaux spécialisés, thèse présentée à l'Université de Montréal, 274 p.
- ELRA (2002): European Language Resources Association, http://www.icp.grenet.fr/ELRA/home.html
- FUCHS, Max (1966): « La langue des sciences », BRUNOT, Ferdinand et al. *Histoire de la langue française des origines à nos jours*, (1966-1972), t. VI, 1<sup>re</sup> partie, fasc. 2<sup>e</sup>, Paris, Colin, pp. 523-675.
- GARSIDE, Roger, LEECH, Geoffrey N. et Tony McENERY (1997): *Corpus annotation: linguistic information from computer text corpora*, New York, N.Y. Longman, 281 p.
- GINGUAY, Michel (1998): Dictionnaire français-anglais d'informatique: bureautique, télématique, micro-informatique, 6<sup>e</sup> éd., 2<sup>e</sup> tirage avec mise à jour, Paris, InterEditions, 224 p.
- GOUGENHEIM, Georges, MICHÉA, René, RIVENC, Paul et Aurélien SAUVAGEOT (1964): L'Elaboration du français fondamental (1er degré): étude sur l'etablissement d'un vocabulaire et d'une grammaire de base, Coll. « Linguistique appliquée », Paris, Didier, 302 p.
- HABERT, Benoît, NAZARENKO, Adeline et André SALEM (1997): Les linguistiques de corpus, coll. « U. Linguistique », Paris, A. Colin, 240 p.
- KENNEDY, Graeme D. (1998): An Introduction to Corpus Linguistics, London, Longman, 315 p.
- KOCOUREK, Rostislav (1991): La langue française de la technique et de la science : vers une linguistique de la langue savante, 2<sup>e</sup> éd. augm., Wiesbaden, Allemagne, O. Brandstetter, xviii + 327 p.
- LAFON, Pierre (1980) : « Sur la variabilité de la fréquence des formes dans un corpus » *MOTS*, n° 1, pp. 128-165.
- LAMALLE, Cédric, MARTINEZ, William, FLEURY, Serge et André SALEM (2003): « LEXICO 3 », SYLED CLA2T, Université de la Sorbonne Nouvelle Paris 3, 50 p. http://www.cavi.univ-paris3.fr/ilpga/ilpga/tal/lexicoWWW/
- LAROUSSE (2001): Le Petit Larousse illustré, Paris, Larousse, 1786 p.

- LEBART, Ludovic et André SALEM (1994): *Statistique textuelle*, Paris, Dunod, viii + 342 p.
- LECOMTE, Josette (1998): Le catégoriseur Brill14-JL5 / WinBrill-0.3, entraîné pour le Français à l'INaLF, Rapport interne INaLF, Nancy. http://www.inalf.fr/winbrill/BRILL14-JL5\_WinBrill.doc
- LEE, David Y. W. (2001): « Defining Core Vocabulary and Tracking its Distribution across Spoken and Written Genres: Evidence of a Gradience of Variation from the British National Corpus », *Journal of English Linguistics*, vol. 29, n° 3, pp. 250-278.
- L'HOMME, Denis, L'HOMME, Marie-Claude et Chantal LEMAY (2002): « Benchmarking the performance of two part-of-speech (POS) taggers for terminological purposes », *Knowledge Organization*.
- MAILLET, Normand (1998) : Caractérisation lexicale de la langue technique québécoise, Université de Sherbrooke, Projet de thèse, 375 p.
- MEILLET, Antoine (1975) : *Linguistique historique et linguistique générale*, Paris, Champion.
- MULLER, Charles (1977): Principes et méthodes de statistique lexicale, Coll. « Langue Linguistique Communication », Paris, Hachette, 206 p.
- MULLER, Charles (1979): Langue française et linguistique quantitative, Éditions Slatkine, Genève, 470 p.
- NELSON, Mike (2000): A Corpus-Based Study of Business English and Business English Teaching Materials, Unpublished PhD Thesis, Manchester, University of Manchester. http://www.kielikanava.com/thesis.html
- PEARSON, Jennifer (1998): *Terms in Context*, Coll. « Studies in Corpus Linguistics », Amsterdam/Philadelphia, John Benjamins Publishing, xii + 243 p.
- PHAL, André (1976): Vocabulaire général d'orientation scientifique (VGOS): part du lexique commun dans l'expression, Paris, CREDIF, Didier, 128 p.
- POTTIER, Bernard (1973): *Le langage*, Coll. « Dictionnaires du savoir moderne », Paris, Denoël, 544 p.

- QUEMADA, Bernard (1955): *Introduction à l'étude du vocabulaire médical* (1600-1710), Coll. « Université de Besançon, France. Annales littéraires. Langues et littérature françaises », Paris, Les Belles Lettres, 198 p.
- RAYSON, Paul et Roger GARSIDE (2000): «Comparing Corpora using Frequency Profiling», UCREL Research Group, Computing Department, Lancaster University. http://www.comp.lancs.ac.uk/computing/users/paul/publications/rg\_acl2000.pdf
- RONDEAU, Guy (1981): *Introduction à la terminologie*, Montréal, Centre éducatif et culturel, 227 p.
- SAUSSURE, Ferdinand de (1975): Cours de linguistique générale, Coll. « Payothèque », Paris, Payot, 510 p.
- SCOTT, Mike (1998): *WordSmith Tools Manual*, Oxford, Oxford University Press, 149 p. <a href="http://www.lexically.net/wordsmith/version3/manual.pdf">http://www.lexically.net/wordsmith/version3/manual.pdf</a>
- SILBERZTEIN, Max (2000): *Manuel Intex* http://grelis.univ-fcomte.fr/intex/downloads/Manuel.pdf
- SINCLAIR, John (1996): *Preliminary recommendations on Corpus Typology*, Rap. tech., EAGLES (Expert Advisory Group on Language Engineering Standards), mai 1996, CEE, 13 p. <a href="http://www.ilc.cnr.it/EAGLES96/browse.html">http://www.ilc.cnr.it/EAGLES96/browse.html</a>
- UCREL (2003): *University Centre for Computer Corpus Research on Language*, University of Lancaster, United Kingdom. http://www.comp.lancs.ac.uk/computing/research/ucrel/.
- VANNESTE, Alex (1988): « Pour une description statistique de la structure morphologique du lexique », *Travaux de Linguistique*, n° 16, pp. 123-144.
- VENDRYÈS, Joseph (1968): *Le langage; introduction linguistique à l'histoire*, Coll. « L'Evolution de l'humanité », Paris : Michel, 444 p.
- WARTBURG, Walther von (1963): *Problèmes et méthodes de la linguistique*, traduit de l'allemand par Pierre Maillard, 2<sup>e</sup> éd. Paris, PUF, 350 p.

## ANNEXE 1 – Liste des étiquettes

Étiquette WinBrill	Signification
ABR	Abréviation
ADJ:sg	Adjectif (sauf PP) sing.
ADJ:pl	Adjectif (sauf PP) plur.
ADJ1PAR:sg	PP adjectif sing. derrière auxiliaire « être »
ADJ1PAR:pl	PP adjectif plur. derrière auxiliaire « être »
ADJ2PAR:sg	PP adjectif sing. non derrière auxiliaire
ADJ2PAR:pl	PP adjectif plur. non derrière auxiliaire
ADV	Adverbe
CAR	Cardinal (chiffres ou lettres)
COO	Coordonnant
DTN:sg	Déterminant de gr. nominal, sing.
DTN:pl	Déterminant de gr. nominal, plur.
DTC:sg	Déterminant contracté, sing.
DTC:pl	Déterminant contracté, plur.
FGW	Mot étranger
INJ	Interjection ou onomatopée
PFX	Préfixe (détaché)
PREP	Préposition
PRV:sg	Pronom sing. supporté par le verbe
PRV:pl	Pronom plur. supporté par le verbe
PRO:sg	Autre pronom, sing.
PRO:pl	Autre pronom, plur.
PUL	Particule
REL	Relatif (pron., adj. ou adv.)
SUB	Subordonnant
SUB\$	code par défaut pour « que »
SBC:sg	Substantif sing. (nom commun)
SBC:pl	Substantif plur. (nom commun)
SBP:sg	Substantif (nom propre sing. ou titre)
SBP:pl	Substantif (nom propre plur. ou titre)
SBP?:sg	Substantif (nom propre probable)
SYM	Symboles ou signes mathématiques
ACJ:sg	« avoir » (verbe ou aux.) conjugué sing.
ACJ:pl	« avoir » (verbe ou aux.) conjugué plur.
ANCFF	« avoir » (verbe ou aux.) infinitif
ANCNT	« avoir » (verbe ou aux.) en -ant
APAR:sg	« avoir » (verbe ou aux.) PP verbal sing. après « avoir »
APAR:pl	« avoir » (verbe ou aux.) PP verbal plur. après « avoir »
ECJ:sg	« être » (verbe ou aux.) conjugué sing.
ECJ:pl	« être » (verbe ou aux.) conjugué plur.
ENCFF	« être » (verbe ou aux.) infinitif
ENCNT	« être » (verbe ou aux.) en -ant
EPAR:sg	« être » (verbe ou aux.) PP verbal sing. après « avoir »
VCJ:sg	autre verbe conjugué, sing.
VCJ:pl	autre verbe conjugué, plur.
VNCFF	autre verbe non conj. infinitif
VNCNT	autre verbe non conj. en -ant
VPAR:sg	autre verbe PP sing. après « avoir »
VPAR:pl	autre verbe PP plur. après « avoir »

# ANNEXE 2 – Regroupement des étiquettes

Étiquette WinBrill	Signification	Étiquette des lemmes
ABR	Abréviation	ABR
ADJ:sg	Adjectif (sauf PP) sing.	
ADJ:pl	Adjectif (sauf PP) plur.	ADJ
ADJ1PAR:sg	PP adjectif sing. derrière auxiliaire « être »	
ADJ1PAR:pl	PP adjectif plur. derrière auxiliaire « être »	1,5,5,5
ADJ2PAR:sg	PP adjectif sing. non derrière auxiliaire	ADJPAR
ADJ2PAR:pl	PP adjectif plur. non derrière auxiliaire	7
ADV	Adverbe	ADV
CAR	Cardinal (chiffres ou lettres)	CAR
COO	Coordonnant	COO
DTN:sg	Déterminant de gr. nominal, sing.	
DTN:pl	Déterminant de gr. nominal, plur.	
DTC:sg	Déterminant contracté, sing.	- DT
DTC:pl	Déterminant contracté, plur.	7
FGW	Mot étranger	FGW
INJ	Interjection ou onomatopée	INJ
PFX	Préfixe (détaché)	PFX
PREP	Préposition	PREP
PRV:sg	Pronom sing. supporté par le verbe	
PRV:pl	Pronom plur. supporté par le verbe	PRO
PRO:sg	Autre pronom, sing.	T PRO
PRO:pl	Autre pronom, plur.	
PUL	Particule	PUL
REL	Relatif (pron., adj. ou adv.)	REL
SUB	Subordonnant	SUB
SUB\$	code par défaut pour « que »	SUB\$
SBC:sg	Substantif sing. (nom commun)	SBC
SBC:pl	Substantif plur. (nom commun)	360
SBP:sg	Substantif (nom propre sing. ou titre)	
SBP:pl	Substantif (nom propre plur. ou titre)	SBP
SBP?:sg	Substantif (nom propre probable)	
SYM	Symboles ou signes mathématiques	SYM
ACJ:sg	« avoir » (verbe ou aux.) conjugué sing.	
ACJ:pl	« avoir » (verbe ou aux.) conjugué plur.	
ANCFF	« avoir » (verbe ou aux.) infinitif	VBA
ANCNT	« avoir » (verbe ou aux.) en -ant	] '5/
APAR:sg	« avoir » (verbe ou aux.) PP verbal sing. après « avoir »	<u> </u>
APAR:pl	« avoir » (verbe ou aux.) PP verbal plur. après « avoir »	
ECJ:sg	« être » (verbe ou aux.) conjugué sing.	
ECJ:pl	« être » (verbe ou aux.) conjugué plur.	<u> </u>
ENCFF	« être » (verbe ou aux.) infinitif	VBE
ENCNT	« être » (verbe ou aux.) en -ant	
EPAR:sg	« être » (verbe ou aux.) PP verbal sing. après « avoir »	
VCJ:sg	autre verbe conjugué, sing.	
VCJ:pl	autre verbe conjugué, plur.	_
VNCFF	autre verbe non conj. infinitif	VB
VNCNT	autre verbe non conj. en -ant	J ' '
VPAR:sg	autre verbe PP sing. après « avoir »	
VPAR:pl	autre verbe PP plur. après « avoir »	

## ANNEXE 3 - Liste des S+ obtenues à l'aide de M1

Forme	PdeD	Fréq. abs.	Valeur- test	Forme	PdeD	Fréq. abs.	Valeur- test
fichier	SBC	3956	315,535	application	SBC	784	75,9382
commande	SBC	1902	177,526	donner	ADJPAR	875	75,227
internet	SBC	1102	170,618	bios	SBP	208	73,8553
serveur	SBC	1166	168,286	taper	VB	297	72,3874
utiliser	VB	1993	163,323	accès	SBC	697	72,108
utilisateur	SBC	1117	162,823	moniteur	SBC	232	71,5013
logiciel	SBC	1166	161,681	carte	SBC	872	71,4387
option	SBC	1486	158,922	périphérique	SBC	228	71,3522
ordinateur	SBC	1283	153,731	ceci	PRO	342	71,1974
système	SBC	2699	143,549	cliquer	VB	191	70,9395
configuration	SBC	845	140,704	paquet	SBC	332	70,6471
répertoire	SBC	1003	139,748	information	SBC	1102	70,4916
touche	SBC	855	124,923	spécifier	VB	213	69,3334
disquette	SBC	609	124,514	syntaxe	SBC	213	69,1809
windows	SBP	613	124,412	exemple	SBC	1171	68,3981
votre	DT	1365	121,361	langage	SBC	465	67,5176
disque	SBC	1093	120,167	standard	ADJ	242	67,5114
réseau	SBC	1511	119,167	connecter	ADJPAR	192	67,2591
imprimante	SBC	537	117,294	connecter	VB	188	67,1733
mémoire	SBC	1240	112,42	pouvoir	VB	3946	64,8044
clavier	SBC	579	111,162	outil	SBC	418	64,2729
donnée	SBC	840	109,156	os	SBP	176	64,0698
linux	NNP	442	108,083	code	SBC	614	63,994
connexion	SBC	486	108,072	virtuel	ADJ	239	63,9928
adresse	SBC	701	107,938	bouton	SBC	204	63,6336
caractère	SBC	1096	107,309	format	SBC	322	63,6234
permettre	VB	2624	107,059	octet	SBC	154	63,4364
utiliser	ADJPAR	1004	105,988	mode	SBC	688	63,3405
recommander	ADJPAR	516	104,764	variable	SBC	164	62,1212
interface	SBC	412	100,964	partition	SBC	334	61,8221
machine	SBC	895	100,871	configurer	VB	141	60,6616
bit .	SBC	382	100,04	section	SBC	378	58,6391
web	SBC	374	99,4001	binaire	ADJ	158	58,5364
paramètre	SBC	455	98,5955	démarrage	SBC	190	57,1614
graphique	ADJ	501	98,376	encodage	SBC	123	56,842
support	SBC	528	97,259	désirer	VB	263	56,8187
module	SBC	442	96,1238	affichage	SBC	227	56,5036
contenir	VB	685	93,4505	exploitation	SBC	455	56,3634
fonctionnalité	SBC	330	92,3067	unité	SBC	629	55,9797
protocole	SBC	541	89,87	OU to market a l	COO	3548	55,0604
programme	SBC	1570	88,7556	terminal	SBC	146	54,8394
activer	VB	422	87,2824	A	PREP	113	54,462
ligne	SBC	1171	86,4238	utilisation	SBC	402	54,141
menu	SBC	471 772	86,0654	contenu	SBC	358	53,9528
écran	SBC	773 1479	85,8078	lecteur	SBC	449	53,5229
nom	SBC SBC	_	85,5273	etc	ADV SBC	581 111	53,1538
noyau		460 777	84,88	synchronisation	SBC		52,7103
suivant périphérique	ADJ ADJ	777 375	84,6229 82,6337	taille	ADJ	363	52,5919 52,2338
périphérique	SBC	375 275	82,6337 82,4885	terminal afficher	VB	193 359	52,2338 51,0435
processeur script	SBC	275 278	82,4665 82,2186	compilateur	SBC	103	51,9435 51,7042
	SBC	276 875	81,2405	imprimer	VB	103	51,7042 51,5
type unix	SBP	238	77,1635	spécifier	ADJPAR	114	51,3 51,3944
souris	SBC	308	77,1033	défaut	SBC	373	50,8426
curseur	SBC	230	76,4645	accéder	VB	303	50,7969
33,000	555	_00	. 5, . 5 . 5	1 2200001		555	33,, 000

être	VBE	17325	50,7618	descripteur	SBC	58	38,1464
base	SBC	681	50,46	ibm	SBP	170	38,1375
modem	SBC	105	50,4474	activer	ADJPAR	65	38,1063
algorithme	SBC	103	50,4146	liste	SBC	453	38,08
extension	SBC	281	50,3165	gérer	VB	292	38,0772
xwindow	SBP	96		9	SBC	346	
			50,1574	ressource			37,9853
électronique	ADJ	299	49,7248	bus	SBC	137	37,9837
programmeur	SBC	95	49,3375	paquetage	SBC	59	37,8283
automatiquement	ADV	187	49,1912	postscript	SBP	55	37,8111
exécuter	VB	210	48,9594	pc	SBC	302	37,7505
sous-répertoire	SBC	90	48,5471	nombre	SBC	797	37,7287
icône	SBC	134	48,1071	chaque	DT	790	37,6369
port	SBC	420	47,2507	archiver	VB	65	37,5654
modifier	VB	382	46,8569	zope	SBP	54	37,4592
utilisateur	ADJ	108	46,799	carte-mère	SBC	54	37,4592
dur	ADJ	400	46,7512	passerelle	SBC	114	37,4563
message	SBC	393	46,3166	ci-dessus	ADV	98	37,3221
fonctionner	VB	345	45,9459	interpréteur	SBC	53	37,104
réponse	SBC	513	45,9303	barrette	SBC	59	36,8894
débit	SBC	144	45,8657	authentification	SBC	55	36,7464
appeler	ADJPAR	348	45,8516	formatage	SBC	54	36,7341
client	SBC	429	45,7712	site	SBC	370	36,7015
	SBC	623	45,3521	compilation	SBC	89	36,665
page	ADJ	191		valeur	SBC	511	,
numérique			44,9963				36,6127
fréquence	SBC	181	44,7876	enregistrement	SBC	190	36,563
exécution	SBC	297	44,5991	champ	SBC	307	36,5218
stocker	ADJPAR	119	44,3275	boîtier	SBC	64	36,4445
copier	VB	127	44,0857	fournir	ADJPAR	200	36,4002
donc	ADV	1210	44,0338	netscape	SBP	51	36,3831
abréviation	SBC	78	43,9567	wordperfect	SBP	51	36,3831
effectuer	VB	366	43,4704	insérer	VB	111	36,207
sélectionner	VB	136	43,4464	majuscule	SBC	64	36,195
programmation	SBC	226	43,3041	console	SBC	78	36,1927
requête	SBC	191	43,1938	fournisseur	SBC	171	36,1285
disponible	ADJ	302	43,0628	variable	ADJ	152	36,1168
imprimant	ADJ	81	42,9721	gestionnaire	SBC	145	35,8883
microprocesseur	SBC	92	42,8968	description	SBC	141	35,4737
gestion	SBC	529	42,8484	amorçage	SBC	49	35,2628
document	SBC	471	42,7927	exécutable	ADJ	49	35,2628
implémentation	SBC	70	42,7437	graveur	SBC	76	35,2064
micro-ordinateur	SBC	116	42,6199	appuyer	VB	268	35,0813
stocker	VB	109	42,3652	navigateur	SBC	97	34,9733
généralement	ADV	291	42,2698	logiciel	ADJ	48	34,8855
version	SBC	389	42,2090		SUB	485	34,7722
	SBC			lorsque	SBC		
fenêtre		227	42,1371	tabulation		46 46	34,515
différent	ADJ	664	41,365	pointeur	SBC	46	34,515
routage	SBC	71	41,2129	initialisation	SBC	46	34,515
si 	SUB	1807	41,1376	horizontal	ADJ	84	34,4651
fonction	SBC	652	40,4911	vertical	ADJ	98	34,4249
laser	SBC	103	40,0229	informatique	ADJ	235	34,3744
spécification	SBC	71	39,8331	architecture	SBC	205	34,3445
horloge	SBC	108	39,7325	traitement	SBC	321	34,2068
correctement	ADV	130	39,4705	sérier	VB	58	34,0439
truetype	SBP	59	39,1872	connecteur	SBC	51	33,9447
démon	SBC	134	39,0192	java	SBP	64	33,9274
ce	DT	6487	38,9467	mécanisme	SBC	225	33,7976
installation	SBC	313	38,9141	chargement	SBC	81	33,6683
courrier	SBC	181	38,6811	compiler	VB	49	33,4954
modèle	SBC	410	38,6686	passe	SBC	160	33,4838
possible	ADJ	580	38,5868	epson	SBP	46	33,3614
définir	VB	326	38,2447	utilitaire	SBC	43	33,344
échappement	SBC	76	38,191	définir	ADJPAR	198	33,3367
сопаррешеш	350	70	50,131	GOITHI	יזטטו עוז	130	00,0007

				1			
compatible	ADJ	120	33,2926	un	DT	17344	28,9783
interconnexion	SBC	64	33,1204	contrôleur	SBC	92	28,8928
mot-clé	SBC	51	33,0274	technologie	SBC	214	28,8396
		42					
identificateur	SBC		32,9445	manuellement	ADV	40	28,7891
clic	SBC	43	32,9347	utilitaire	ADJ	74	28,7302
onglet	SBC	43	32,9347	nécessaire	ADJ	439	28,6469
boucle	SBC	99	32,5582	mot	SBC	455	28,59
charger	VB	168	32,551	vérifier	VB	163	28,4948
	SBP						
unicode		41	32,54	noter	VB	234	28,4535
niveau	SBC	544	32,4288	apple	SBP	68	28,4493
en-tête	SBC	58	32,3938	formulaire	SBC	62	28,4412
erreur	SBC	316	32,3359	afficher	ADJPAR	130	28,4271
correspondre	VB	273	32,2511		VB	90	28,4161
				sauvegarder			
utilisable	ADJ	73	32,2055	compiler	ADJPAR	34	28,2382
configurer	ADJPAR	42	32,1305	corba	SBP	31	28,1779
ethernet	SBP	40	32,1304	micro-informatique	ADJ	40	28,1664
nécessiter	VB	179	32,1221	vitesse	SBC	196	28,0998
processus	SBC	366	32,0365	passant	ADJ	59	28,0801
•							
copie	SBC	145	32,0066	caractéristique	SBC	112	28,0361
composant	ADJ	126	31,9999	balayage	SBC	39	28,027
externe	ADJ	110	31,9076	messagerie	SBC	57	27,9492
cas	SBC	806	31,803	contenant	ADJ	52	27,8856
développeur	SBC	40	31,7069	distant	ADJ	69	27,8655
codage	SBC	60	31,6782	logique	ADJ	174	27,7814
samba	SBP	51	31,6364	translation	SBC	32	27,7245
instruction	SBC	340	31,5483	yahoo	SBP	30	27,704
ci-dessous	ADV	101	31,5463	corel	SBP	30	27,704
courant	ADJ	213	31,4548	redirection	SBC	30	27,704
progiciel	SBC	43	31,4243	routeur	SBC	30	27,704
accessible	ADJ	143	31,422	visualiser	VB	50	27,6521
entrée-sortie	SBC	38	31,2952	déplacer	VB	137	27,5874
démarrer	VB	120	31,2836	feuille	SBC	128	27,5368
impression	SBC	266	31,2605	source	SBC	341	27,5129
	ADJPAR	154			SBC	146	27,3083
décrire			31,2149	colonne			
composant	SBC	83	31,2031	majuscule	ADJ	36	27,3053
pour	PREP	6448	31,1119	assembleur	SBC	31	27,2454
standard	SBC	69	30,8686	spécifique	ADJ	181	27,2253
désactiver	VB	38	30,8614	adressage	SBC	30	27,2198
arborescence	SBC	38	30,8614	matériel	SBC	221	27,2112
			,				
table	SBC	265	30,8243	environnement	SBC	315	27,0783
signal	SBC	141	30,7981	plusieurs	DT	738	27,064
texte	SBC	629	30,7725	générique	ADJ	61	27,0207
rigide	ADJ	89	30,7312	installer	VB	253	26,9947
test	SBC	190	30,6186	maintenance	SBC	79	26,9592
	ADV	284	30,5527	documentation	SBC	93	26,9492
directement							
générer	VB	94	30,5197	clé	SBC	149	26,9132
raciner	VB	36	30,4371	ruban	SBC	57	26,9112
exister	VB	444	30,3149	on	PRO	2652	26,816
disposer	VB	375	30,0965	adobe	SBP	28	26,731
rafraîchissement	SBC	39	30,0901	fournir	VB	256	26,7196
télécharger	VB	35	29,9989	combinaison	SBC	85	26,5994
réglage	SBC	52	29,6625	aiguille	SBC	60	26,5397
visualisation	SBC	42	29,6604	mail	SBC	32	26,4843
tableur	SBC	35	29,5481	vidéo	ADJ	103	26,4427
guillemet	SBC	48	29,4214	altavista	SBP	27	26,231
matriciel	ADJ	39	29,3579	baud	SBC	27	26,231
simple	ADJ	434	29,3496	émulateur	ADJ	27	26,231
audio	PFX	61	29,2606	suffire	VB	288	26,0933
magnétique	ADJ	81	29,2078	physique	ADJ	173	25,9456
internaute	SBC	33	29,1026	coder	VB	35	25,8634
pouce	SBC	105	29,056	signifier	VB	213	25,8429
	SBC	64	,		VB VB		
largeur	300	U <del>1</del>	29,0069	déterminer	۷D	157	25,8341

, .	40.0040		05 70 40	I	000	- 4	00 00 10
exécuter	ADJPAR	119	25,7849	ghostscript	SBP	21	23,0042
espacement	SBC	28	25,7562	implémenter	VB	21	23,0042
dynamiquement	ADV	26	25,7213	référencer	ADJPAR	26	22,986
intranet	SBC	26	25,7213	pipe	SBC	39	22,9786
chiffrier	SBC	26	25,7213	multimédia	ADJ	44	22,9777
opérateur	SBC	173	25,7107	transactionnel	ADJ	28	22,9764
générer	ADJPAR	46	25,6415	analogique	ADJ	35	22,961
contenir	ADJPAR	104	25,6332	moteur	SBC	137	22,9151
élément	SBC	298	25,5648	détecter	VB	64	22,8883
préfixe	SBC	27	25,2407	remplissage	SBC	44	22,8284
arborescent	ADJ	27	25,2407	effacer	VB	113	22,8029
normalement	ADV	109	25,2311	chargeur	SBC	33	22,69
séquence	SBC	109	25,2311	couramment	ADV	53	22,6329
paramétrer	VB	26	25,2044	communiquer	VB	105	22,6085
contextuel	ADJ	26	25,2044	papier	SBC	181	22,6078
formater	VB	25	25,2013	superviseur	SBC	26	22,6024
téléchargement	SBC	25	25,2013	défilement	SBC	23	22,5828
copier	ADJPAR	41	25,0702	cependant	ADV	398	22,5679
attribut	SBC	53		lister	VB	29	
			25,0003				22,5429
topologie	SBC	29	24,98	formater	ADJPAR	22	22,4926
étendue	SBC	78	24,9537	relier	ADJPAR	78	22,4524
coder	ADJPAR	47	24,8452	calcul	SBC	137	22,4465
sauvegarde	SBC	98	24,8437	staroffice	SBP	20	22,4214
choisir	VB	348	24,8115	commutateur	SBC	20	22,4214
définition	SBC	169	24,76	paramétrage	SBC	20	22,4214
	VB	69		initialiser	VB	20	22,4214
manipuler			24,7208				
cache	SBC	49	24,7169	appletalk	SBP	20	22,4214
explorateur	SBP	25	24,6751	câble	SBC	112	22,4081
telnet	SBP	24	24,6703	index	SBC	54	22,3845
opérande	SBC	24	24,6703	simplement	ADV	250	22,3736
filtrage	SBC	38	24,6031	falloir	VB	921	22,331
tampon	SBC	46	24,5979	opération	SBC	478	22,2925
•	SBC	86	24,5631	amovible	ADJ	24	22,269
emplacement							
spécial	ADJ	298	24,5094	séquentiel	ADJ	24	22,269
distribution	SBC	232	24,4959	cartouche	SBC	43	22,263
communication	SBC	315	24,4239	compresser	ADJPAR	30	22,2011
paragraphe	SBC	62	24,3941	explorer	SBP	31	22,1849
case	SBC	66	24,2683	domino	SBP	23	22,1246
compatibilité	SBC	44	24,2669	taper	ADJPAR	23	22,1246
graphe	SBC	24	24,1344	fonctionnement	SBC	214	22,1099
usenet	SBP	23	24,1277	modification	SBC	169	22,1074
			•				
hypertexte	ADJ	23	24,1277	matériel	ADJ	136	22,0643
multiprogrammation	SBC	23	24,1277	destination	SBC	115	21,9566
sous-menu	SBC	23	24,1277	aide	SBC	450	21,9542
sous-réseau	SBC	23	24,1277	supprimer	VB	161	21,9475
indiquer	ADJPAR	123	24,0385	suffixe	SBC	20	21,8402
brancher	VB	42	23,9391	sous-commande	SBC	19	21,8231
arithmétique	ADJ	46	23,8061	débogage	SBC	19	21,8231
recherche	SBC	455	23,798	manuel	ADJ	65	21,8142
						35	
performance	SBC	174	23,7692	évoluer	ADJPAR		21,6986
utile	ADJ	153	23,7395	gérer	ADJPAR	102	21,6131
numérotation	SBC	24	23,629	optimiser	VB	40	21,5714
symbole	SBC	173	23,5984	désirer	ADJPAR	47	21,5689
chapitre	SBC	153	23,588	saisie	SBC	69	21,5486
créer	VB	414	23,5617	pentium	SBP	26	21,5465
renommer	VB	27	23,531	segment	SBC	48	21,5072
posséder	VB	192	23,3915	d'	PREP	12008	21,4191
				-			
imprimer	ADJPAR	76	23,3369	positionner	VB CBC	34	21,4164
intel	SBP	48	23,1955	indicatif	SBC	21	21,4037
stockage	SBC	72	23,0646	valide	ADJ	45	21,3654
numéro	SBC	287	23,021	annuaire	SBC	50	21,3184
prédéfini	ADJ	21	23,0042	écriture	SBC	145	21,2204
•			•	•			•

encapsuler	VB	18	21,208	quicktime	SBP	15	19,2454
				•		15	
conséquent	ADJ	78 42	21,1796	diode	SBC		19,2454
cliente	SBC	43	21,1729	astérisque	SBC	15	19,2454
envoyer	VB	179	21,0509	cisco	SBP	15	19,2454
capacité	SBC	224	20,9631	image	SBC	448	19,2372
identifier	VB	98	20,9614	bloc	SBC	111	19,1899
entrée	SBC	297	20,9393	optimisation	SBC	24	19,1803
États-unis	SBP	51	20,9137	flux	SBC	101	19,1555
parallèle	ADJ	109	20,881	façon	SBC	357	19,0739
broche	SBC	27	20,8767	explication	SBP	20	19,0448
structure	SBC	242	20,8482	additionnel	ADJ	35	19,0396
usager	SBC	84	20,8225	programmable	ADJ	18	19,0235
afin	PUL	383	20,7696	sauvegarder	ADJPAR	34	19,0113
oracle	SBP	19	20,6761	couleur	SBC	204	18,9655
alphanumérique	ADJ	18	20,5993	besoin	SBC	347	18,9295
applicatif	ADJ	18	20,5993	supporter	ADJPAR	44	18,8439
mégaoctet	SBC	17	20,5745	intermédiaire	SBC	116	18,7812
client-serveur	ADJ	17	20,5745	également	ADV	511	18,7117
flashage	SBC	17	20,5745	recopier	VB	23	18,5916
ajout	SBC	37	20,5438	référencer	VB	15	18,5874
tâche	SBC	189	20,5182	sécuriser	ADJPAR	18	18,5662
consulter	VB	94	20,4957	plupart	SBC	295	18,5657
cylindre	SBC	47	20,4937	par	PREP	5182	18,5482
-	ADJ		,	•	SBC	14	
minuscule		64	20,4739	réglette		14	18,5453
fiabilité	SBC	46	20,4114	connectivité	SBC		18,5453
capable	ADJ	194	20,3528	bogue	SBC	14	18,5453
manuel	SBC	54	20,3482	cryptologie	SBC	14	18,5453
adaptateur	SBC	25	20,3455	arpanet	SBP	14	18,5453
branchement	SBC	27	20,3282	œuvre	SBC	14	18,5453
multimédia	SBC	20	20,3148	conversion	SBC	63	18,4156
cellule	SBC	128	20,2963	diagramme	SBC	17	18,348
puce	SBC	64	20,293	manière	SBC	347	18,2346
perforer	ADJPAR	22	20,2878	baser	ADJPAR	102	18,227
simultanément	ADV	87	20,2714	navigation	SBC	55	18,2016
trouver	VB	667	20,2155	synchroniser	VB	16	18,1517
barre	SBC	125	20,1639	sous-système	SBC	16	18,1517
servir	VB	297	20,1293	couche	SBC	68	18,1214
notion	SBC	128	20,1291	solution	SBC	272	18,0632
gravage	SBC	17	19,9506	provenir	VB	120	18,0115
virgule	SBC	31	19,9358	police	SBC	393	17,8764
rediriger	VB	16	19,921	relocalisation	SBC	14	17,868
encapsulation	SBC	16	19,921	itération	SBC	14	17,868
optique	ADJ	61	19,9183	usage	SBC	182	17,8578
longueur	SBC	89	19,8215	incrément	SBC	13	17,8178
séparer	ADJPAR	100	19,7998	terminaison	SBC	13	17,8178
sur	PREP	4223	19,7964	proxy	SBC	13	17,8178
lien	SBC	191	19,7371	encapsuler	ADJPAR	13	17,8178
approprier	ADJPAR	74	19,6968	enfichable	ADJ	13	17,8178
décimal	ADJ	19	19,6785	nanoseconde	SBC	13	17,8178
accéder	ADJPAR	19	19,6785	méthode	SBC	188	17,7844
éditeur	SBC	163	19,5285	redémarrer	VB	41	17,7597
	ADJPAR				SBC	26	
optimiser		18	19,512	graphique			17,7089
baser	VB CDD	36	19,4957	réutilisation	SBC	16	17,6502
net	SBP	23	19,4909	problème	SBC	463	17,6404
service	SBC	614	19,4589	lorsqu'	SUB	250	17,5963
indiquer	VB	425	19,4486	dédier	ADJPAR	54	17,5864
validation	SBC	30	19,4282	fonctionnel	ADJ	43	17,579
possibilité	SBC	235	19,3902	particulier	ADJ	284	17,5285
anglais	SBC	188	19,3848	forme	SBC	339	17,5084
principal	ADJ	418	19,3824	conseiller	ADJPAR	37	17,49
infrastructure	SBC	108	19,3212	syntaxique	ADJ	15	17,436
pixel	SBC	15	19,2454	devoir	VB	1551	17,3932

présentement	ADV	25	17,3582	local	ADJ	420	15,8265
•							
interruption	SBC	68	17,3048	accentuer	ADJPAR	45	15,8231
micro	SBC	71	17,297	simplifier	VB	44	15,782
règle	SBC	245	17,2846	fibre	SBC	51	15,7722
_	SBC	77	17,28	suivant	PREP	76	15,6992
registre							
encre	SBC	48	17,2512	calculatrice	SBC	12	15,6939
facilement	ADV	93	17,2155	accumulateur	ADJ	12	15,6939
molette	SBC	13	17,1196	faisceau	SBC	32	15,6165
partageable	ADJ	13	17,1196	tube	SBC	49	15,5909
inclure	VB	76	17,1028	multilingue	ADJ	19	15,572
maximal	ADJ	53	17,096	pratique	ADJ	141	15,5481
	SBC	23				82	
téléphonie			17,0742	varier	VB		15,5394
ventilateur	SBC	23	17,0742	modéliser	VB	11	15,5197
antislash	SBC	12	17,0596	encoche	SBC	11	15,5197
plantage	SBC	12	17,0596	entendre	ADJPAR	158	15,5147
tramage	SBC	12	17,0596	lecture	SBC	169	15,4997
acrobat	SBP	12	17,0596	adapter	ADJPAR	98	15,4864
coolmon	SBP	12	17,0596	normaliser	ADJPAR	20	15,486
milliseconde	SBC	12	17,0596	sortie	SBC	146	15,4753
rechercher	VB	102	17,0136	forum	SBC	46	15,4623
interactif	ADJ	40	17,0061	powerpoint	SBP	10	15,4326
certain	DT	547	17,006	compuserve	SBP	10	15,4326
car	COO	482	16,9947	cyberbavardage	SBC	10	15,4326
robustesse	SBC	18	16,9915	patche	SBC	10	15,4326
autre	ADJ	1096	16,9284	sous-programme	SBC	10	15,4326
	ADJPAR	117	16,8765	pica	SBC	10	15,4326
partager							
technique	ADJ	261	16,8354	encoder	ADJPAR	10	15,4326
norme	SBC	117	16,8245	implémenter	ADJPAR	10	15,4326
facultatif	ADJ	31	16,7665	manipulation	SBC	62	15,4034
	PFX	14				13	
pseudo			16,6931	interagir	VB		15,3994
précédemment	ADV	58	16,662	python	SBP	19	15,323
décrire	VB	101	16,6127	choix	SBC	275	15,322
extraire	VB	39	16,6084	insérer	ADJPAR	28	15,3156
	SBC						
raccourci		29	16,6027	cible	SBC	97	15,2972
mémoriser	ADJPAR	13	16,4907	développement	SBC	344	15,2912
éditer	VB	56	16,4391	émulation	SBC	25	15,2694
statique	ADJ	24	16,4028	transférer	VB	56	15,1754
ainsi	ADV	837	16,3938	plateforme	SBC	18	15,1736
consister	VB	155	16,3414	mémoriser	VB	18	15,1736
étiquetage	SBC	18	16,3255	enfoncer	ADJPAR	31	15,1645
alphabet	SBC	27	16,322	téléphonique	ADJ	100	15,1376
expéditeur	SBC	17	16,3058	calculer	VB	45	15,0669
nativement	ADV	11	16,2663	domaine	SBC	259	15,0339
paintbrush	SBP	11	16,2663	destinataire	SBC	32	15,0228
routeur	ADJ	11	16,2663	tout	DT	1795	15,0174
algorithmique	ADJ	11	16,2663	bande	SBC	135	15,0058
simm	SBP	11	16,2663	valider	VB	33	14,9981
inode	SBC	11	16,2663	convivial	ADJ	31	14,9637
seconde	SBC	130	16,1278	via	PREP	90	14,9395
tableau	SBC	170	16,0963	comporter	VB	137	14,9232
inconvénient	SBC	59	16,0765	transfert	SBC	122	14,9173
filtre	SBC	33	16,0641	sélection	SBC	89	14,8826
complet	ADJ	139	16,0591	chaîner	VB	15	14,8704
matrice	SBC	24	16,0323	accumulateur	SBC	14	14,8558
sous-ensemble	SBC	15	16,0198	vidéo	SBC	79	14,8386
simuler	VB	28	15,9651	relationnel	ADJ	26	14,8186
automatique	ADJ	82	15,9131	correspondant	ADJ	113	14,8021
				•			
substitution	SBC	46	15,8677	sous	PREP	719	14,7831
uniquement	ADV	98	15,8563	récupérer	VB	87	14,7478
partage	SBC	132	15,8412	phase	SBC	121	14,7192
apparaître	VB	254	15,8347	étendre	ADJPAR	67	14,6885
pagination	SBC	19	15,8315	chaîne	SBC	271	14,6696

				I			
encoder	VB	10	14,6579	tableur	ADJ	9	13,7457
microordinateur	SBC	10	14,6579	dupliquer	VB	9	13,7457
détecter	ADJPAR	23	14,6361	désactiver	ADJPAR	9	13,7457
archivage	SBC	16	14,6337	interligne	SBC	9	13,7457
<u> </u>				_		9	
déploiement	SBC	62	14,5893	cosse	SBC		13,7457
rapide	ADJ	163	14,5812	informatique	SBC	54	13,7404
compression	SBC	33	14,5639	intégrer	ADJPAR	72	13,6921
				<u> </u>			
nommage	SBC	9	14,5518	calculateur	SBC	13	13,6917
condensateur	SBC	9	14,5518	emmagasiner	VB	13	13,6917
	SBP	9		localiser	VB	24	
palmos			14,5518				13,6561
intranet	SBP	9	14,5518	programmer	VB	47	13,6393
caper	VB	9	14,5518	méthodologie	SBC	23	13,6337
				<u> </u>			
codepson	SBP	9	14,5518	recompiler	VB	8	13,6148
cobol	SBP	9	14,5518	multitâche	ADJ	8	13,6148
macro	SBC	9	14,5518	syntaxe	NN	8	13,6148
booléen	ADJ	9	14,5518	conversationnel	ADJ	8	13,6148
cœur	NN	9	14,5518	émuler	VB	8	13,6148
redirigé	ADJ	9	14,5518	binette	SBC	8	13,6148
solaris	SBP	9	14,5518	netware	SBP	8	13,6148
séparateur	SBC	9	14,5518	fortran	SBP	8	13,6148
supporter	VB	93	14,5378	napster	SBP	8	13,6148
canal	SBC	84	14,5331	kilo-octet	SBC	8	13,6148
						8	
traçant	ADJ	13	14,4803	printmaster	SBP		13,6148
transistor	SBC	19	14,4181	télécharger	ADJPAR	8	13,6148
même	ADJ	979	14,4159	méga-octet	SBC	8	13,6148
secondaire	ADJ	75	14,4115	portabilité	SBC	8	13,6148
dynamique	ADJ	67	14,3447	double-cliquer	VB	8	13,6148
très	ADV	861	14,3264		SBP	8	
				javascript			13,6148
défectueux	ADJ	24	14,3058	liaison	SBC	99	13,5939
transmission	SBC	84	14,2833	tester	VB	46	13,5834
			,				
décompresser	VB	11	14,2682	ensemble	SBC	341	13,5773
structuration	SBC	17	14,245	classiquement	ADV	18	13,5747
portable	ADJ	33	14,2359	préférable	ADJ	36	13,5314
•							
placer	VB	137	14,21	critère	SBC	103	13,5053
interne	ADJ	129	14,2091	importer	VB	152	13,5041
détection	SBC	28	14,1831	édition	ADJ	14	13,4656
amstrad	SBP	14	14,117	positionner	ADJPAR	14	13,4656
fente	SBC	13	14,0718	relativement	ADV	88	13,4052
vérification	SBC	52	14,0535	problématique	SBC	28	13,4039
existant	ADJ	79	14,0454	tiret	SBC	10	13,3793
réaliser	VB	202	14,0169	tape	SBC	10	13,3793
				•			
mathématique	ADJ	35	14,0037	maxi	ADJ	10	13,3793
alias	ADV	33	14,0002	alphabétique	ADJ	23	13,3758
bêta	SBC	10	13,9798	modifier	ADJPAR	71	13,3518
interopérabilité	SBC	10	13,9798	borne	SBC	36	13,3463
flasher	VB	10	13,9798	entrer	VB	188	13,3453
invariant	SBC	10	13,9798	migrer	VB	13	13,337
performant	ADJ	50	13,9761	technique	SBC	101	13,2226
éventuellement	ADV	69	13,9626	qualité	SBC	236	13,2165
implicite	ADJ	34	13,9605	déconnecter	VB	14	13,1673
quel	DT	263	13,9419	italique	ADJ	15	13,0951
piste	SBC	115	13,9185	indexer	VB	16	13,087
•							
sélectionner	ADJPAR	44	13,9015	espace	SBC	209	13,0761
classique	ADJ	155	13,897	brochage	SBC	9	13,0495
étiquette	SBC	64	13,8903	déconnexion	SBC	9	13,0495
monochrome	ADJ	20	13,8878	picot	SBC	9	13,0495
effectuer	ADJPAR	119	13,882	structurer	ADJPAR	30	13,0381
électron	SBC	26	13,8568	cathodique	ADJ	23	13,0092
voici	PREP	119	13,8566	optionnel	ADJ	13	13,0047
celui-ci	PRO	393	13,7919	librairie	SBC	63	12,9978
digital	ADJ	24	13,7806	figure	SBC	144	12,9836
modélisation	SBC	17	13,7687	oblique	ADJ	19	12,9454

				•			
lequel	REL	540	12,9075	distance	SBC	105	12,0603
fiable	ADJ	35	12,9063	adéquatement	ADV	8	12,0588
							,
boire	VB	63	12,8954	développeur	ADJ	8	12,0588
contenant	SBC	14	12,8851	centrage	SBC	8	12,0588
appelant	ADJ	14	12,8851	remplacer	VB	116	12,0273
local	SBP	12	12,879	premièrement	ADV	17	12,0257
adhésif	ADJ	12		1 -	SBC	18	,
			12,879	recoupement			12,0238
perl	SBP	12	12,879	intéressant	ADJ	86	11,9829
presser	VB	60	12,8678	associer	ADJPAR	96	11,9679
personnaliser	VB	16	12,867	asynchrone	ADJ	10	11,9194
redondance	SBC	15	12,8485	vista	SBP	10	11,9194
bureautique	ADJ	10	12,8425	tridimensionnel	ADJ	10	11,9194
activation	SBC	11	12,8118	sous-section	SBC	9	11,8998
silicium	SBC	11	12,8118	ergonomie	SBC	9	11,8998
cadencer	ADJPAR	11	12,8118	cocher	VB	9	11,8998
allouer	VB	21	12,7969	macromolécule	SBC	9	11,8998
	ADV	28			SBC	12	
inversement			12,795	couplage			11,8924
envoyer	ADJPAR	73	12,7879	brancher	ADJPAR	26	11,8851
encreur	SBC	8	12,7736	détailler	ADJPAR	43	11,8779
dimensionnement	SBC	8	12,7736	étape	SBC	136	11,826
lettre	SBC	238	12,7365	boîte	SBC	79	11,8103
écrire	ADJPAR	107	12,6671	concevoir	ADJPAR	84	11,7829
chemin	SBC	164	12,6522	enchaînement	SBC	28	11,7378
graphisme	SBC	20	12,6359	qwerty	SBP	7	11,7288
concept	SBC	76	12,6216	calibrage	SBC	7	11,7288
switche	SBC	7	12,6094	paginer	ADJPAR	7	11,7288
relocalisable	ADJ	7	12,6094	poster	VB	, 16	11,7126
				1 .			
chiffrement	SBC	7	12,6094	littéral	ADJ	16	11,7126
partitionner	VB	7	12,6094	bas	SBC	81	11,7058
réinitialiser	VB	7	12,6094	racine	SBC	55	11,6876
multi-tâche	ADJ	7	12,6094	tablette	SBC	14	11,67
préfixer	VB	<i>.</i> 7	12,6094	accolade	SBC	14	11,67
							,
préemptif	ADJ	7	12,6094	supposer	VB	91	11,6413
défragmentation	SBC	7	12,6094	compteur	SBC	20	11,6175
paramétrable	ADJ	7	12,6094	condenser	ADJPAR	13	11,6094
multi-utilisateur	ADJ	7	12,6094	recharger	VB	12	11,6038
e-mail	SBC	7	12,6094	respectivement	ADV	78	11,5994
		7					
pare-feu	SBC		12,6094	avantage	SBC	130	11,5953
email	SBC	7	12,6094	cela	PRO	498	11,5913
coprocesseur	SBC	7	12,6094	interpréteur	ADJ	6	11,5185
objet	SBC	242	12,5612	azerty	SBP	6	11,5185
verticalement	ADV	12	12,5277	défragmenter	VB	6	11,5185
développer	ADJPAR	84	12,5223	implémenté	ADJ	6	11,5185
		-				-	
circuit	SBC	81	12,5079	autocad	SBP	6	11,5185
répandre	ADJPAR	53	12,4827	antivirus	SBC	6	11,5185
mobile	SBC	26	12,4713	effaçable	ADJ	6	11,5185
lancer	VB	230	12,465	multiprogrammer	ADJPAR	6	11,5185
interaction	SBC	25	12,4473	adressable	ADJ	6	11,5185
synchroniser	ADJPAR	13	12,3987	séquencement	SBC	6	11,5185
déconseiller	ADJPAR	13	12,3987	scout	SBP	6	11,5185
optimal	ADJ	22	12,3623	déroulant	ADJ	6	11,5185
aspect	SBC	110	12,3623	multitâche	SBC	6	11,5185
compresser	VB	10	12,3587	photoshop	SBP	6	11,5185
	SBC	10		hypertexte		6	
duplication			12,3587		SBC		11,5185
abonné	SBC	42	12,3492	moleter	ADJPAR	6	11,5185
complexité	SBC	59	12,3284	client-serveur	SBC	6	11,5185
souvent	ADV	360	12,297	modificateur	ADJ	6	11,5185
présentation	SBC	92	12,2706	initialisation	SBP	6	11,5185
libellé	SBC	12	12,1997	signet	SBC	6	11,5185
invite	SBC	17	12,1933	téléchargeable	ADJ	6	11,5185
haut	SBC	111	12,1398	collaboratif	ADJ	6	11,5185
complexe	ADJ	110	12,1024	signification	SBC	50	11,4931

				Las			
auxiliaire	ADJ	32	11,424	ligner	VB	13	10,3425
voyant	SBC	9	11,4169	éteindre	VB	34	10,3397
appareil	SBC	110	11,3947	saule-pleureur	SBC	5	10,3161
ensuite	ADV	218	11,369	interrelié	ADJ	5	10,3161
vectoriel	ADJ	11	11,3636	scalaire	ADJ	5	10,3161
convertir	VB	28	11,3379	double-clic	SBC	5	10,3161
bibliographique	ADJ	12	11,3318	journaliser	ADJPAR	5	10,3161
				*			
libre	ADJ	166	11,3159	econet	SBP	5	10,3161
natif	ADJ	26	11,2821	redémarrer	ADJPAR	5	10,3161
licence	SBC	49	11,2217	extranet	SBC	5	10,3161
plate-forme	SBC	51	11,2064	pagemill	SBP	5	10,3161
exactement	ADV	71	11,206	sybex	SBP	5	10,3161
suivre	VB	219	11,1826	susmentionné	ADJ	5	10,3161
entité	SBC	57	11,1808	superutilisateur	SBC	5	10,3161
jet	SBC	34	11,1674	point-virgule	SBC	5	10,3161
grâce	SBC	253	11,1601	modularisation	SBC	5	10,3161
_	SBC	71			SBC	5	
bibliothèque			11,1185	mémoire-cache			10,3161
actif	ADJ	134	11,1104	worldcom	SBP	5	10,3161
relier	VB	66	11,0666	sous-clé	SBC	5	10,3161
recalculer	VB	7	10,9955	sous-chaîne	SBC	5	10,3161
vers	PREP	384	10,9933	soulignement	SBC	5	10,3161
concepteur	SBC	24	10,9576	lister	ADJPAR	5	10,3161
lire	ADJPAR	40	10,9426	postcondition	SBC	5	10,3161
modulaire	ADJ	13	10,9324	aironet	SBP	5	10,3161
flèche	SBC	30	10,9178	powerarchiver	SBP	5	10,3161
	VB	28	10,9046	dbase	SBP	5	10,3161
rajouter							
ce	PRO	2646	10,893	récursif	ADJ	5	10,3161
résolution	SBC	124	10,8787	paramétrique	ADJ	5	10,3161
simultané	ADJ	26	10,8695	imprimable	ADJ	5	10,3161
installer	ADJPAR	135	10,8357	multiplexage	SBC	5	10,3161
intégrer	VB	80		sélénium	SBC	5	
			10,8328				10,3161
identification	SBC	33	10,8064	troncature	SBC	5	10,3161
pavé	SBC	34	10,7609	vierge	ADJ	26	10,3041
correction	SBC	36	10,7395	personnaliser	ADJPAR	21	10,2955
variante	SBC	29	10,7395	commutation	SBC	11	10,2649
	ADJ	321	10,7344		ADJ	24	10,2459
important				typique			
mobile	ADJ	41	10,7331	notification	SBC	17	10,2429
cycle	SBC	83	10,7267	session	SBC	98	10,2114
réel	ADJ	166	10,715	lire	VB	188	10,1532
limiter	ADJPAR	111	10,7004	lui	PRO	18175	10,1479
guantité	SBC	61	10,6827	fiche	SBC	35	10,1319
•							
introduire	ADJPAR	40	10,6628	alimentation	SBC	38	10,1081
transmettre	ADJPAR	50	10,6072	ultérieurement	ADV	30	10,0696
réalisable	ADJ	18	10,5996	apache	SBP	14	10,0464
ergonomique	ADJ	6	10,5932	astuce	SBC	20	10,0371
sémantique	SBC	9	10,5858	mini-ordinateur	SBC	8	9,99118
migration	SBC	32	10,5803	excel	SBP	8	9,99118
inutile	ADJ	66	10,5729	correctif	ADJ	12	9,97426
simplifier	ADJPAR	26	10,548	câblage	SBC	12	9,97426
machiner	VB	11	10,5149	caractéristique	ADJ	22	9,93978
inutiliser	ADJPAR	10	10,4933		SBC	65	
				transaction			9,91945
panne	SBC	42	10,485	appeler	VB	255	9,87521
évolutivité	SBC	8	10,4203	élémentaire	ADJ	39	9,86758
router	VB	8	10,4203	interconnecter	VB	6	9,84276
caroll	SBP	8	10,4203	atari	SBP	6	9,84276
		53		débrancher		6	
compliquer	ADJPAR		10,4108		VB		9,84276
gamme	SBC	66	10,3978	numéroter	VB	6	9,84276
favori	SBC	20	10,3878	cryptage	SBC	7	9,83275
ponctuation	SBC	12	10,3821	internationalisation	SBC	27	9,83135
segmentation	SBC	12	10,3821	identifier	ADJPAR	45	9,81358
	ADJPAR	89		distinct	ADJ	39	
constituer			10,3779				9,80572
crochet	SBC	14	10,3608	distinguer	VB	71	9,79346

équiper	ADJPAR	51	9,7851	transparent	ADJ	34	9,16744
numéroter	ADJPAR	12	9,78404	orienter	ADJPAR	32	9,12277
redondant	ADJ	10	9,66782	inutilisable	ADJ	13	9,07407
verrouillage	SBC	13	9,66134	rubrique	SBC	34	9,06899
composante	SBC	34	9,65105	incompatible	ADJ	28	9,05439
entier	SBC	16		-	ADJ	11	
			9,64619	servant			9,01617
réservoir	SBC	28	9,63442	partir	VB	322	8,99651
puisqu'	SUB	102	9,61907	mémorisation	SBC	10	8,98267
assertion	SBC	14	9,61311	comparer	VB	55	8,96645
placer	ADJPAR	128	9,61304	algol	SBP	4	8,96028
demander	ADJPAR	76	9,61234	sync	SBP	4	8,96028
dépendre	VB	106	9,61139	cocher	ADJPAR	4	8,96028
•	VB					4	,
subdiviser		8	9,60399	valué	ADJ		8,96028
modifiable	ADJ	8	9,60399	hébergeur	SBC	4	8,96028
clignoter	VB	8	9,60399	hexadécimal	SBC	4	8,96028
décoder	VB	12	9,60202	sous-dossier	SBC	4	8,96028
standardisation	SBC	11			ADV	4	
			9,59318	statiquement			8,96028
diffusion	SBC	99	9,58539	realaudio	SBP	4	8,96028
familiariser	VB	16	9,53637	complexifier	VB	4	8,96028
abonnement	SBC	35	9,52466	hyperlien	SBC	4	8,96028
limitation	SBC	45		économiseur	SBC	4	
			9,48969				8,96028
suivante	SBC	12	9,42761	hypermédia	ADJ	4	8,96028
terminer	VB	83	9,41478	hexadécimal	ADJ	4	8,96028
représenter	VB	217	9,4064	adapteur	SBC	4	8,96028
hétérogène	ADJ	17	9,40589	surbrillance	SBC	4	8,96028
_							
entrant	ADJ	11	9,39152	anti-virus	SBC	4	8,96028
routine	SBC	26	9,37726	activex	SBP	4	8,96028
offrir	VB	167	9,37547	fureteur	SBC	4	8,96028
combinatoire	SBC	7	9,36003	x-window	SBP	4	8,96028
tournevis	SBC	7	9,36003	récapituler	ADJPAR	4	8,96028
dièse	SBC	7	9,36003	uploader	VB	4	8,96028
réutilisable	ADJ	7	9,36003	typage	SBC	4	8,96028
rapidité	SBC	33	9,35285	bibl	ADJ	4	8,96028
interconnecté	ADJ	5	9,34074	défragmenteur	SBC	4	8,96028
paramétrer	ADJPAR	5	9,34074	encrer	ADJPAR	4	8,96028
					VB	4	
bidirectionnel	ADJ	5	9,34074	enficher			8,96028
classifier	VB	5	9,34074	centronics	SBP	4	8,96028
tronquer	VB	5	9,34074	digitalisation	SBC	4	8,96028
séparateur	ADJ	5	9,34074	upgrader	VB	4	8,96028
mainteneur	SBC	5	9,34074	désactivé	ADJ	4	8,96028
						4	
analyseur	SBC	5	9,34074	prérequis	SBC		8,96028
lexicographique	ADJ	5	9,34074	interréseau	SBC	4	8,96028
symbologie	SBC	5	9,34074	partagiciel	ADJ	4	8,96028
pertinent	ADJ	27	9,33931	kilooctet	SBC	4	8,96028
caractériser	ADJPAR	30	9,33428	journalisation	SBC	4	8,96028
gratuit	ADJ	54	9,32366	paginable	ADJ	4	8,96028
coût	SBC	183	9,32129	multimode	ADJ	4	8,96028
enregistrer	VB	101	9,31811	publisher	SBP	4	8,96028
répondre	VB	222	9,31535	multi-terminal	ADJ	4	8,96028
disponibilité	SBC	28		prédéfinir	ADJPAR	4	
			9,31418				8,96028
multiplicateur	ADJ	9	9,30052	logitech	SBP	4	8,96028
requérir	VB	41	9,27902	lisp	NNP	4	8,96028
prestataire	SBC	15	9,25995	multi-tête	ADJ	4	8,96028
ajuster	VB	23	9,23976	partitionnement	SBC	4	8,96028
	ADV	24			ADJ	4	
impérativement			9,22416	multi-processeur			8,96028
novell	SBP	6	9,21735	multi-utilisateur	SBC	4	8,96028
latence	SBC	6	9,21735	paint	SBP	4	8,96028
optimum	SBC	6	9,21735	contrôle	SBC	245	8,94664
horizontalement	ADV	6	9,21735	propriétaire	ADJ	33	8,94099
reilly	SBP	6	9,21735	périodique	SBC	7	8,94083
personnalisation	SBC	10	9,19806	usuel .	ADJ	13	8,93935
coordonnée	SBC	10	9,19806	raccourcir	ADJPAR	13	8,93935

signaler	VB	67	8,93607	principalement	ADV	60	8,32219
raccorder				1 -	VB		
	ADJPAR	8	8,93045	traiter		94	8,31434
décompression	SBC	8	8,93045	uniforme	ADJ	17	8,31043
afficheur	SBC	8	8,93045	trait	SBC	54	8,28912
	ADJ	16	8,92919	manipuler	ADJPAR	16	8,23312
sous-jacent							
transférer	ADJPAR	38	8,9267	samba	SBC	10	8,22551
remarquer	VB	91	8,9246	francisation	SBC	6	8,22437
endroit	SBC	74	8,87778	segmenter	ADJPAR	6	8,22437
corriger	VB	42	8,87633	correspondance	SBC	56	8,21229
redémarrage	SBC	26	8,85555	chacun	PRO	237	8,20381
naviguer	VB	28	8,84734	changer	VB	189	8,18231
					SBC		,
lisible	ADJ	19	8,83178	portion		20	8,17338
chariot	SBC	15	8,82981	hiérarchique	ADJ	22	8,16229
positionnement	SBC	15	8,82981	assemblage	SBC	22	8,16229
son	SBC	69	8,79799	tri	SBC	23	8,15194
masque	SBC	41	8,7966	dans	PREP	5205	8,14891
extensible	ADJ	9	8,78777	suite	SBC	231	8,13658
quartz	SBC	9	8,78777	résultant	ADJ	9	8,13118
normal	ADJ	93	8,7859	raccordement	SBC	9	8,13118
munir	ADJPAR	32	8,74425	télécom	SBC	9	8,13118
quelconque	ADJ	42	8,73402	maximum	ADJ	20	8,11697
attente	SBC	94	8,7251	canon	SBP	8	8,10771
			,				
soit	COO	206	8,71872	traduction	SBC	57	8,08877
représenter	ADJPAR	64	8,7185	caractériser	VB	39	8,08125
facile	ADJ	89	8,70386	marguerite	SBC	10	8,05809
prédéterminer	ADJPAR	6	8,68507	gauche	ADJ	44	8,02835
•				•			
bip	SBC	6	8,68507	métropolitain	ADJ	22	8,01636
sécurisation	SBC	6	8,68507	dimension	SBC	72	7,98572
revendeur	SBC	16	8,65467	interpréter	VB	42	7,9853
	ADJ	8	,		ADJPAR	110	
pleureur			8,63465	situer			7,97943
ultérieur	ADJ	26	8,59212	regrouper	ADJPAR	36	7,97752
standardiser	VB	5	8,57699	рс	NN	29	7,9668
recharger	ADJPAR	5	8,57699	archiver	ADJPAR	5	7,95676
	SBC	5			SBC	5	
homme-machine			8,57699	micro-processeur			7,95676
magnétiser	ADJPAR	5	8,57699	métrologie	SBC	5	7,95676
reprogrammable	ADJ	5	8,57699	coupe-feu	SBC	5	7,95676
réellement	ADV	53	8,57102	restriction	SBC	41	7,95556
	SBC	7		contrôler	VB	88	7,93955
démontage			8,5655				,
macro	PFX	7	8,5655	classeur	SBC	9	7,93651
concentrique	ADJ	9	8,55552	gnome	SBP	4	7,93048
portail	SBC	14	8,52721	multi	PFX	4	7,93048
vif	ADJ	104	8,52012	emmagasiner	ADJPAR	4	7,93048
				-		-	
localisation	SBC	21	8,51719	remontage	SBC	4	7,93048
marqueur	SBC	12	8,51399	traduisible	ADJ	4	7,93048
résoudre	VB	78	8,4982	monitoring	SBC	4	7,93048
mise	SBC	274	8,49215	pliure	SBC	4	7,93048
indépendamment	ADV	19	8,49072	amiga	SBP	4	7,93048
exhaustif	ADJ	21	8,46052	schématiser	ADJPAR	4	7,93048
constituer	VB	217	8,44038	initialiser	ADJPAR	4	7,93048
souhaitable	ADJ	34	8,42442	microinformatique	SBC	4	
	-			•			7,93048
recouvrement	SBC	14	8,42237	commuter	ADJPAR	4	7,93048
effacer	ADJPAR	25	8,41854	addition	SBC	22	7,92157
toutefois	ADV	167	8,39095	nappe	SBC	22	7,92157
	SBC	58					
déplacement			8,38715	basique	ADJ	7	7,91856
sémaphore	SBC	8	8,36134	abstraire	VB	7	7,91856
intercaler	VB	8	8,36134	boire	ADJPAR	10	7,89804
interactivité	SBC	11	8,35859	point	SBC	385	7,89224
décomposer	VB	15	8,34725	raccorder	VB	12	7,88734
complètement	ADV	68	8,33803	restreindre	ADJPAR	31	7,88511
lancer	SBC	9	8,33711	vérifier	ADJPAR	19	7,87816
intuitif	ADJ	9	8,33711	saule	SBC	8	7,87141
montage	SBC	51	8,32693	sécuriser	VB	8	7,87141
montage	050	51	0,02030	000011001	۷۵	J	1,01141

incorrect	ADJ	8	7.87141	routable	ADJ	3	7,37485
lancement	SBC	63	7,87108	internet	SBP	3	7,37485
					_		
réception	SBC	31	7,85569	préprocesseur	SBC	3	7,37485
temporaire	ADJ	48	7,84541	multitraitement	SBC	3	7,37485
télécommunication	SBC	63	7,84523	photodiode	SBC	3	7,37485
agenda	SBC	20	7,84515	datagramme	SBC	3	7,37485
•							,
bloc-note	SBC	6	7,82012	modulariser	VB	3	7,37485
buse	SBC	6	7,82012	granularité	SBC	3	7,37485
réplication	SBC	6	7,82012	séquentiellement	ADV	3	7,37485
deuxièmement	ADV	11	7,79828	courriel	SBC	3	7,37485
lisibilité	SBC	12	7,7735	eicon	SBP	3	7,37485
					_		
sémantique	ADJ	12	7,7735	boguer	ADJPAR	3	7,37485
remplacer	ADJPAR	52	7,77264	item	SBC	3	7,37485
heureusement	ADV	45	7,75964	multi-usager	ADJ	3	7,37485
plan	ADJ	33	7,7557	œil	SBC	3	7,37485
centraliser	ADJPAR	19	7,70991	disque-dur	SBC	3	7,37485
				'			
fuseau	SBC	8	7,65047	décoder	ADJPAR	3	7,37485
flécher	ADJPAR	8	7,65047	décommenter	VB	3	7,37485
progressiste	SBC	7	7,63666	biunivoque	ADJ	3	7,37485
abréger	ADJPAR	7	7,63666	frontpage	SBP	3	7,37485
<del>_</del>	SBC	, 61		microseconde	SBC	3	
fabricant			7,62892				7,37485
arrière	ADV	38	7,62837	pointbase	SBP	3	7,37485
visible	ADJ	44	7,62712	activable	ADJ	3	7,37485
documenter	ADJPAR	15	7,60005	traçabilité	SBC	3	7,37485
déconnecter	ADJPAR	10	7,59793	listage	SBC	3	7,37485
				sous-schéma	SBC	3	
partie	SBC	356	7,58783				7,37485
trier	VB	14	7,58504	extensibilité	SBC	3	7,37485
diagonale	SBC	9	7,57696	reconnecter	SBC	3	7,37485
malheureusement	ADV	47	7,56181	cliquable	ADJ	3	7,37485
cristal	SBC	20	7,53956	méta-assembleur	SBC	3	7,37485
	SBC	28		désactif	ADJ	3	
expérimentation			7,53025				7,37485
verrou	SBC	20	7,49063	cliquer	ADJPAR	3	7,37485
rangement	SBC	6	7,46133	cabochon	SBC	3	7,37485
visualiser	ADJPAR	6	7,46133	switcher	VB	3	7,37485
export	SBC	10	7,45691	préfixer	ADJPAR	3	7,37485
choisir	ADJPAR	76	7,44935	mime	NN	3	7,37485
certain	ADJ	212	7,4486	multisynchrone	ADJ	3	7,37485
convertisseur	SBC	5	7,43922	tutoriel	ADJ	3	7,37485
perforation	SBC	5	7,43922	outlook	SBP	3	7,37485
incorporer	VB	11	7,42783	fenêtrage	SBC	3	7,37485
conditionnel	ADJ	14	7,42093	clarisworks	SBP	3	7,37485
			,				
intégrité	SBC	30	7,41968	micro-instruction	SBC	3	7,37485
balise	SBC	9	7,41037	commande	NN	3	7,37485
enregistrer	ADJPAR	74	7,40415	mnémonique	ADJ	3	7,37485
tier	SBC	3	7,37485	multiplate-forme	ADJ	3	7,37485
décompacter	VB	3	7,37485	extranet	SBP	3	7,37485
•		3				3	
arcnet	SBP		7,37485	didacticiel	SBC		7,37485
dreamweaver	SBP	3	7,37485	appleshare	SBP	3	7,37485
helvetica	SBP	3	7,37485	complétion	SBC	3	7,37485
pricewaterhousecoop	oers	SBP	3	instancier7,37485	VB	3	7,37485
aster	SBP	3	7,37485	masquage	SBC	3	7,37485
						3	
économiseur	ADJ	3	7,37485	métaprogrammation	SBC		7,37485
luminophore	ADJ	3	7,37485	trinitron	SBP	3	7,37485
dutilleul	SBP	3	7,37485	inférence	SBC	3	7,37485
interblocage	SBC	3	7,37485	méta-index	SBC	3	7,37485
dissipateur	SBC	3	7,37485	factoriel	ADJ	3	7,37485
	ADJ				VB	3	
auto-adhésif		3	7,37485	sur-dimensionner			7,37485
mono-utilisateur	ADJ	3	7,37485	démonter	VB	18	7,34997
coreldraw	SBP	3	7,37485	vide	ADJ	60	7,34632
magnétisable	ADJ	3	7,37485	électrique	ADJ	51	7,34065
démarrable	ADJ	3	7,37485	affecter	VB	51	7,34065
cryptographie	SBC	3	7,37485	nécessairement	ADV	38	7,2879
σιγρισμιπο	350	J	1,51405	I IIOOGSSAII GIIIGIIL	7D V	50	1,2013

clef	SBC	23	7,26824	détail	SBC	80	6,79716
proprement	ADV	35	7,26409	volatil	ADJ	12	6,79482
tester	ADJPAR	14	7,26353	cordon	SBC	19	6,77483
note	SBP	9	7,25159	québec	SBP	30	6,7613
sinon	ADV	9	7,25159	thermique	ADJ	13	6,75928
quasiment	ADV	47	7,23133	informaticien	SBC	13	6,75928
robot	SBC	16	7,24706	certainement	ADV	55	6,74955
allumer	ADJPAR	17		précéder	ADJPAR	24	
fixe	ADJEAN	54	7,20328 7,196	authentifier	VB	8	6,73085
-	ADJPAR	40		apparition	SBC	56	6,72547
supprimer	SBC	40 42	7,19308 7,19145	1	SBC	87	6,7206
envoi				comportement			6,71988
top résultat	SBC	10	7,19094	spécialiser	ADJPAR	92	6,71633
	SBC	290	7,17001	générateur	SBC	10	6,71405
notice	SBC	15	7,16976	ci-contre	ADV	16	6,71063
emmêlement	SBC	4	7,16298	renvoyer	VB	56	6,69584
vidage	SBC	4	7,16298	transiter	VB	22	6,69581
flash	SBP	4	7,16298	encombrement	SBC	11	6,68938
eudora	SBP	4	7,16298	comprendre	VB	227	6,6761
terminologie	SBC	12	7,15531	adéquat	ADJ	20	6,65268
validité	SBC	22	7,14688	bien	ADV	852	6,64819
résumé	SBC	17	7,14653	but	SBC	88	6,63008
fait	SBC	370	7,14256	contacter	VB	15	6,6023
catégorie	SBC	84	7,14163	évolutif	ADJ	11	6,59563
acronyme	SBC	6	7,13982	démarrer	ADJPAR	6	6,58496
continu	SBC	7	7,13755	textuel	ADJ	8	6,56885
autoriser	VB	81	7,13199	caldera	SBP	4	6,56077
horaire	ADJ	27	7,12824	inria	SBP	4	6,56077
tv	SBC	31	7,11817	grippage	SBC	4	6,56077
préférence	SBC	49	7,11607	vidéo-conférence	SBC	4	6,56077
arrière	SBC	28	7,10977	oxyder	ADJPAR	4	6,56077
inférieur	ADJ	70	7,08182	défectuosité	SBC	4	6,56077
pile	SBC	20	7,07336	décodage	SBC	4	6,56077
dire	ADJPAR	92	7,0723	maintenant	ADV	162	6,55997
compléter	VB	38	7,06513	surcharge	SBC	9	6,55523
minimal	ADJ	28	7,05251	continuellement	ADV	12	6,54661
souple	ADJ	30	7,04056	complémentaire	ADJ	44	6,54597
inclure	ADJPAR	22	7,02976	initial	ADJ	59	6,54225
phosphore	SBC	5	6,99818	aperçu	SBC	13	6,5403
traceur	SBC	5	6,99818	identique	ADJ	47	6,53056
alvéole	SBC	5	6,99818	intermédiaire	ADJ	33	6,52492
transmettre	VB	48	6,98663	quel	PRO	103	6,52171
faciliter	VB	65	6,98366	éviter	VB	167	6,50888
volée	SBC	17	6,9809	fastidieux	ADJ	10	6,49888
note	SBC	107	6,97988	examiner	VB	61	6,49457
équipement	SBC	107	6,9729	gratuitement	ADV	24	6,47987
micro-informatique	SBC	12	6,97053	parfaitement	ADV	67	6,47738
combiner	VB	19	6,95527	différencier	VB	13	6,47016
contrainte	SBC	61	6,95024	ajouter	VB	246	6,45678
décaler	VB	10	6,94409	récupérer	ADJPAR	22	6,45243
affectation	SBC	23	6,93589	obsolète	ADJ	16	6,44477
correspondant	SBC	24	6,92852	tache	SBC	19	6,4358
manette	SBC	7	6,9149	ingénierie	SBC	17	6,42605
photocopieuse	SBC	7	6,9149	réinstaller	VB	8	6,41979
indicateur	ADJ	21	6,90305	indicatif	ADJ	10	6,39634
ordinal	ADJ	8	6,89035	technologique	ADJ	48	6,39424
certification	SBC	16	6,8793	recouper	VB	12	6,39039
bureautique	SBC	6	6,84932	vider	VB	25	6,3867
rangée	SBC	10	6,8271	classification	SBC	15	6,37487
pointer	VB	21	6,82414	volume	SBC	97	6,35904
expression	SBC	99	6,82082	introduire	VB	56	6,35841
neumann	SBP	9	6,81622	variété	SBC	36	6,34926
temps	SBC	5 570	6,79838	exploiter	VB	42	6,3482
.5	350	5.5	3,. 0000	I cubicitoi			5,5 .02

				•			
baladeur	SBC	6	6,34288	ha	INJ	5	5,98183
mini	ADJ	6	6,34288	réutiliser	ADJPAR	5	5,98183
isolément	ADV	6	6,34288	mégabit	SBC	5	5,98183
accessibilité	SBC	6	6,34288	date	SBC	142	5,97607
impair	ADJ	6	6,34288	structurer	VB	14	5,97224
	ADJPAR	17		extrêmement	ADV	53	
abonner			6,33299				5,96727
portatif	ADJ	7	6,33092	concerner	VB	184	5,96344
standardiser	ADJPAR	7	6,33092	systématiquement	ADV	35	5,96177
administrateur	SBC	46	6,32617	écrire	VB	208	5,96147
		-					
mentionner	ADJPAR	19	6,31548	essayer	VB	90	5,9531
delà	PUL	9	6,31397	puis	ADV	305	5,92049
terminer	ADJPAR	44	6,30342	centrer	ADJPAR	20	5,91911
bourre	SBC	3	6,29308	sériel	ADJ	6	5,91375
numération	SBC	3	6,29308	câbler	VB	6	5,91375
invariant	ADJ	3	6,29308	découper	ADJPAR	17	5,89817
				•			
initialisé	ADJ	3	6,29308	typiquement	ADV	13	5,89442
nortel	SBP	3	6,29308	artur	SBP	8	5,88762
proche	NN	3	6,29308	souder	VB	9	5,88096
fonctionnellement	ADV				SBC	63	5,85191
		3	6,29308	propriété			
max	ADJ	3	6,29308	saut	SBC	25	5,85002
multiprocesseur	SBC	3	6,29308	réutiliser	VB	7	5,84419
euro	SBC	3	6,29308	densité	SBC	20	5,81938
							,
cranter	ADJPAR	3	6,29308	sonore	ADJ	27	5,78974
varier	ADJPAR	26	6,28306	impératif	ADJ	20	5,78667
déterministe	ADJ	5	6,28021	symbolique	ADJ	53	5,77139
rapidement	ADV	106	6,27928	piloter	VB	17	5,73699
pince	SBC	8	6,27766	physiquement	ADV	17	5,73699
classer	VB	25	6,27425	succinctement	ADV	6	5,72209
	ADJ	53			SBC	6	
égal			6,21757	copyright			5,72209
que	SUB\$	3068	6,21362	seul	ADJ	531	5,72085
conceptuel	ADJ	15	6,21289	illimiter	ADJPAR	16	5,69996
obligatoirement	ADV	18	6,21222	intercontinental	ADJ	7	5,69886
desserte	SBC	13	6,20264	ruche	SBC	7	5,69886
rectangle	SBC	10	6,20047	précis	ADJ	84	5,67618
parallèle	SBC	14	6,19676	norton	SBP	4	5,66108
créer		122			ADJ	4	,
	ADJPAR		6,1955	adaptatif			5,66108
réserver	ADJPAR	71	6,19069	pointiller	ADJPAR	4	5,66108
obtenir	VB	224	6,19027	informationnel	ADJ	4	5,66108
	SUB	2287	6,18956	rotation	SBC	14	5,65974
que							
température	SBC	23	6,18333	glisser	VB	39	5,65949
offre	SBC	106	6,17533	émettre	ADJPAR	32	5,65632
dessus	ADV	20	6,16154	texture	SBC	8	5,65355
requérir	ADJPAR	26	6,15215	divers	ADJ	139	5,63292
avérer	VB	29	6,14921	évidemment	ADV	74	5,62076
état	SBC	232	6,14056	cohérent	ADJ	29	5,61141
conception	SBC	69	6,12406	similaire	ADJ	33	5,5805
effet	SBC	393	6,11322	effacement	SBC	12	5,57829
suivre	ADJPAR	83	6,09931	prise	SBC	112	5,56658
voir	VB	582	6,09505	zéro	SBC	33	5,5629
fixer	VB	68	6,09054	ample	ADJ	14	5,56138
introduction	SBC	39	6,08131	circonflexe	ADJ	3	5,54487
lavoie	SBP	4	6,07076	lecture-écriture	SBC	3	5,54487
serrage	SBC	4	6,07076	pliage	SBC	3	5,54487
				1 0			
cocher	SBC	4	6,07076	plain	ADJ	3	5,54487
rétroaction	SBC	4	6,07076	multifonction	ADJ	3	5,54487
complétude	SBC	4	6,07076	retendre	VB	3	5,54487
temporairement	ADV	17	6,06634	pan-européen	ADJ	3	5,54487
effectivement	ADV	51	6,02747	redimensionner	VB	3	5,54487
initialement	ADV	33	6,02181	électriquement	ADV	3	5,54487
symétrique	ADJ	10	6,0158	fermant	ADJ	3	5,54487
lent	ADJ	47	6,0061	interprétable	ADJ	3	5,54487
frappe	SBC	25	6,00456	testeur	SBC	3	5,54487
• •				•			

				•			
soustraction	SBC	6	5,54328	francophone	SBC	8	5,04786
stéréo	SBC	6	5,54328	esclave	SBC	21	5,02989
						14	
pointage	SBC	6	5,54328	inverser	ADJPAR		5,02279
socle	SBC	15	5,53848	dérégulation	SBC	9	5,01047
lors	ADV	351	5,53066	possédant	ADJ	4	5,0076
bulle	SBC			•	VB	4	5,0076
		18	5,52981	aiguiller			,
convier	VB	53	5,52625	photocopieur	SBC	4	5,0076
moduler	VB	10	5,51869	uniformiser	VB	4	5,0076
	SBC	9			SBC	4	
multiple .			5,50172	gaine			5,0076
comparaison	SBC	39	5,49627	interlude	SBC	4	5,0076
risc	SBP	5	5,47198	verticale	SBC	4	5,0076
énoncé	SBC	12	5,46012	powerbook	SBP	3	4,98522
				1 -			
classer	ADJPAR	34	5,44593	remplaçable	ADJ	3	4,98522
semblable	ADJ	30	5,44358	clignotement	SBC	3	4,98522
minitel	SBC	8	5,43699	conceptuellement	ADV	3	4,98522
		7				3	
dactylographie	SBC		5,42922	rechargeable	ADJ		4,98522
relatif	ADJ	82	5,42298	latéralement	ADV	3	4,98522
décharger	VB	13	5,40033	bristol	SBC	3	4,98522
dessin	SBC	68	5,38102	veiner	ADJPAR	3	4,98522
volt	SBC	6	5,37587	vérificateur	SBC	3	4,98522
goulot	SBC	6	5,37587	héberger	VB	14	4,98144
grandement	ADV	11	5,37156	balayer	ADJPAR	12	4,97525
				,			
traitant	ADJ	10	5,36914	coefficient	SBC	12	4,97525
logiquement	ADV	19	5,33638	bonjour	INJ	7	4,95902
compaq	SBP	9	5,3287	majeur	ADJ	77	4,95838
						14	
compléter	ADJPAR	22	5,32625	pertinence	SBC		4,94058
ranger	VB	20	5,32468	puisque	SUB	123	4,93995
numéris	SBP	4	5,3113	recommencer	VB	23	4,93308
numérisé	ADJ	4	5,3113	administrateur	ADJ	15	4,93226
			,				
rouleau	SBC	11	5,30842	asymétrique	ADJ	6	4,93053
combiner	ADJPAR	13	5,29956	exponentiel	ADJ	6	4,93053
exécutant	SBC	10	5,29701	abonner	VB	6	4,93053
			,				
échanger	VB	33	5,2891	satellite	ADJ	12	4,92556
améliorer	VB	74	5,28305	conduit	SBC	13	4,92405
abstraction	SBC	20	5,26683	classe	SBC	110	4,91124
unicité	SBC	5	5,25136	indispensable	ADJ	76	4,90924
ranger	ADJPAR	19	5,24429	descendant	ADJ	11	4,89857
lumineux	ADJ	29	5,23886	explorateur	SBC	11	4,89857
débutant	SBC	12	5,23573	dépendant	ADJ	19	4,89582
	SBC	6		édition	SBC	127	
instrumentation			5,21861				4,88679
produire	VB	125	5,21157	spécifiquement	ADV	12	4,87663
fréquemment	ADV	26	5,19807	réactivité	SBC	5	4,86266
filtrer	VB	14	5,19327	plus	ADV	2898	4,85465
				1.			
préalablement	ADV	15	5,16264	déconseiller	VB	7	4,8534
plage	SBC	38	5,13493	agir	VB	315	4,84755
gênant	ADJ	15	5,1232	embarquer	ADJPAR	16	4,80297
conserver	ADJPAR	23	5,1182	restreindre	VB	18	4,80127
impliquer	VB	53	5,11294	résiduel	ADJ	6	4,79802
extraction	SBC	10	5,09028	piratage	SBC	6	4,79802
retouche	SBC	10	5,09028	signifiant	ADJ	9	4,79277
	VB	59		fur	PUL	20	
désigner			5,08781				4,78004
pourvoir	ADJPAR	21	5,08149	transposition	SBC	10	4,77432
primitif	ADJ	20	5,07056	traiter	ADJPAR	39	4,76997
complet	SBC	6	5,07047	motif	SBC	50	4,76944
•							
soudure	SBC	7	5,06904	césure	SBC	7	4,75187
où	ADV	110	5,06627	équation	SBC	15	4,75102
acquisition	SBC	47	5,06264	caoutchouc	SBC	13	4,75095
-	ADJPAR				VB	43	
interpréter		30	5,06011	remplir			4,74501
inverse	ADJ	27	5,04963	puissance	SBC	87	4,7437
décodeur	SBC	5	5,0491	éjection	SBC	4	4,7403
fractionner	VB	5	5,0491	entrelacer	ADJPAR	4	4,7403
	VB	5				4	
retaper	۷D	J	5,0491	rabat	SBC	4	4,7403

				•			
miniaturisation	SBC	4	4,7403	particulier	SBC	55	4,4612
exporter	ADJPAR	12	4,73412	gras	ADJ	22	4,45633
•							
plaque	SBC	23	4,71225	intégralité	SBC	15	4,44645
intrinsèquement	ADV	5	4,68995	découpe	SBC	6	4,43864
enduire	ADJPAR	5	4,68995	luminosité	SBC	6	4,43864
	-						
zone	SBC	175	4,6788	descriptif	ADJ	6	4,43864
arbitrairement	ADV	6	4,67225	latéral	ADJ	8	4,4032
ci	PUL	6		réaliser	ADJPAR	113	4,40259
-			4,67225		-	_	,
intégration	SBC	64	4,66002	virtuellement	ADV	7	4,38152
puissant	ADJ	67	4,65739	repérer	VB	21	4,36769
•	SBC	9	4,6563	personnel	ADJ	103	4,36736
récepteur				'			
étranglement	SBC	9	4,6563	vocabulaire	SBC	18	4,36486
indexer	ADJPAR	9	4,6563	regrouper	VB	60	4,36089
ex	PFX	19	4,65318	capture	SBC	11	4,34696
livrer	ADJPAR	35	4,64302	valider	ADJPAR	8	4,33243
coupler	ADJPAR	8	4,62743	dépendance	SBC	21	4,30264
impossible	ADJ	73	4,61926	noter	ADJPAR	11	4,3014
-							
charge	SBC	154	4,60276	repérer	ADJPAR	12	4,30025
impact	SBC	43	4,59484	récapitulatif	ADJ	4	4,28846
carbone	SBC	9	4,59045	dépanner	VB	4	4,28846
fréquent	ADJ	31	4,58269	fragmentation	SBC	9	4,28242
généralité	SBC	7	4,56002	lot	SBC	30	4,28142
absolument	ADV	37	4,55461	agréable	ADJ	19	4,27923
			,		ADJPAR	26	
paradigme	SBC	6	4,55264	déplacer			4,26791
endommager	VB	8	4,5506	ascendant	ADJ	8	4,2635
phosphorescent	ADJ	3	4,54456	aléatoire	ADJ	17	4,25275
enfoncement	SBC	3	,		ADJ	5	
			4,54456	exploitable			4,23857
télétel	SBP	3	4,54456	numérisation	SBC	5	4,23857
nécessitant	ADJ	3	4,54456	idem	ADV	6	4,22574
intercaler	ADJPAR	3	4,54456	global	ADJ	75	4,20881
	-			· ·	-		
déclaratif	ADJ	3	4,54456	panneau	SBC	24	4,20551
supercalculateur	SBC	3	4,54456	arrière-plan	SBC	8	4,19631
interpersonnel	ADJ	3	4,54456	expérimental	ADJ	23	4,18585
emboîtement	SBC	3	4,54456	croisillon	SBC	3	4,18477
stimuli	SBC	3	4,54456	lexical	ADJ	3	4,18477
polygone	SBC	3	4,54456	amplificateur	ADJ	3	4,18477
différence	SBC	85	4,53725	polymorphisme	SBC	3	4,18477
extraire	ADJPAR	17	4,52933	synchronique	ADJ	3	4,18477
kit	SBC	5	4,52928	entreposage	SBC	3	4,18477
-	SBC	5	,		SBC	39	,
réceptacle			4,52928	remplacement			4,1843
compétiteur	SBC	5	4,52928	phrase	SBC	49	4,17862
surface	SBC	57	4,52796	catalogue	SBC	38	4,15696
résident	ADJ	9	4,52609	présent	ADJ	160	4,14772
		-					
composer	VB	48	4,52455	préciser	ADJPAR	33	4,1423
dater	VB	38	4,51016	agrandissement	SBC	7	4,13537
référence	SBC	76	4,50842	robuste	ADJ	8	4,13078
glissière	SBC	4	4,50234	graver	VB	8	4,13078
codification	SBC	4	4,50234	analogie	SBC	8	4,13078
récepteur	ADJ	4	4,50234	redéfinir	ADJPAR	6	4,12605
					-		
fourreau	SBC	4	4,50234	relâchement	SBC	9	4,11267
hé	INJ	4	4,50234	marquage	SBC	5	4,10631
coaxial	ADJ	4	4,50234	contigu	ADJ	5	4,10631
	SBC	4		affirmative	SBC	5	
allumage			4,50234				4,10631
originellement	ADV	4	4,50234	monolithique	ADJ	5	4,10631
éjecter	VB	4	4,50234	nombreux	ADJ	216	4,09758
particularité	SBC	17	4,50073	intentionner	ADJPAR	4	4,09464
•							
éliminer	VB	29	4,48815	récupérable	ADJ	4	4,09464
égal	SBC	11	4,48787	étiqueter	VB	4	4,09464
dénomination	SBC	11	4,48787	énumérer	ADJPAR	4	4,09464
							•
garantir	VB .	53	4,48628	déchargement	SBC	4	4,09464
central	ADJ	181	4,48584	suffisamment	ADV	45	4,09257
incorporer	ADJPAR	8	4,47589	acheminer	ADJPAR	10	4,08241
		-	.,	1			.,

				1			
périodique	ADJ	10	4,08241	studio	SBP	11	3,76412
indicateur	SBC	17	4,07069	paire	SBC	11	3,76412
minimum	SBC	38	4,06335	occasionnellement	ADV	4	3,75531
			,				
défaillance	SBC	20	4,05835	résidant	ADJ	4	3,75531
conforme	ADJ	36	4,04184	ajuster	ADJPAR	5	3,75186
maximum	SBC	38	4,04019	alors	ADV	554	3,74758
			,				
évident	ADJ	54	4,02515	proposer	ADJPAR	73	3,71593
jeu	SBC	196	4,02432	globalisation	SBC	8	3,71294
nommer	VB	51	4,02139	héberger	ADJPAR	7	3,70593
	VB	114	4,01764	magnétoscope	SBC	10	3,6846
poser							
facilité	SBC	26	4,01711	recommandation	SBC	25	3,66233
final	ADJ	58	3,99391	double	ADJ	85	3,65986
sun	SBP	12	3,99139	limiter	VB	70	3,65578
							,
liquide	ADJ	15	3,991	médium	SBC	5	3,64565
hewlett-packard	SBP	5	3,98159	superposer	ADJPAR	5	3,64565
dorsal	ADJ	5	3,98159	électromagnétique	ADJ	5	3,64565
normalisation	SBC	17	3,97053	obtenir	ADJPAR	46	3,6455
							,
déterminer	ADJPAR	35	3,96462	adapter	VB	42	3,64057
simulation	SBC	11	3,95961	grandeur	SBC	23	3,63209
télématique	SBC	10	3,94361	distinction	SBC	22	3,62893
			,				,
utilité	SBC	23	3,92692	combinatoire	ADJ	3	3,62441
inventer	ADJPAR	17	3,92144	radial	ADJ	3	3,62441
menu	ADJ	11	3,91948	usuellement	ADV	3	3,62441
généralisation	SBC	12	,		ADJPAR	17	
O .			3,91869	renvoyer			3,61733
techniquement	ADV	12	3,91869	émetteur	SBC	11	3,61642
amplificateur	SBC	4	3,91775	conserver	VB	81	3,61596
antériorité	SBC	4	3,91775	rigoureusement	ADV	9	3,61223
				•			,
mégarde	SBC	4	3,91775	expliciter	ADJPAR	4	3,60532
constituant	SBC	4	3,91775	consultable	ADJ	4	3,60532
minimiser	VB	13	3,90056	présenter	VB	206	3,60046
indexation	SBC	10	3,89878	planter	VB	16	
			,	1 '			3,5999
géostationnaire	ADJ	3	3,88295	améliorer	ADJPAR	20	3,5958
décloisonnement	SBC	3	3,88295	facteur	SBC	48	3,59423
illustré	SBC	3	3,88295	référer	VB	22	3,59413
							,
abrégé	SBC	3	3,88295	auquel	REL	122	3,5897
capteur	ADJ	3	3,88295	simplification	SBC	12	3,57858
renseigner	VB	11	3,87988	hétérogénéité	SBC	7	3,57804
évoluer	VB	48	3,87394	respectif	ADJ	21	3,57485
							,
composer	ADJPAR	60	3,87348	ratio	SBC	13	3,5608
réglementaire	ADJ	20	3,87096	épaisseur	SBC	13	3,5608
coller	VB	20	3.87096	dispositif	SBC	70	3,55641
amorcer	VB	20	,	mineur	ADJ	35	3,55215
			3,87096				,
onéreux	ADJ	13	3,86816	persistance	SBC	12	3,54653
familièrement	ADV	5	3,86365	intelligible	ADJ	5	3,54453
distribuer	ADJPAR	37	3,84853	habituellement	ADV	19	3,54411
incompatibilité	SBC	7	3,84092	téléphone	SBC	50	3,54054
primordial	ADJ	13	3,83608	ici	ADV	247	3,53511
trouver	ADJPAR	43	3,83569	hébergement	SBC	17	3,52858
constructeur	SBC	56	3,83565	omettre	VB	11	3,51026
assembler	VB	8	3,82565	payant	ADJ	14	3,50325
coûteux	ADJ	33	3,81862	point	ADV	78	3,48616
incitation	SBC	22	3,80736	gérant	SBC	16	3,4829
affiche	SBC	39	3,79256	francophone	ADJ	31	
							3,47777
consulter	ADJPAR	18	3,79132	acheminer	VB	14	3,47675
reproduire	VB	23	3,78566	finaliser	ADJPAR	4	3,46616
historique	SBC	7	3,77249	phénoménal	ADJ	4	3,46616
•							
centrer	VB	7	3,77249	comparativement	ADV	4	3,46616
équiper	VB	16	3,77039	possesseur	SBC	4	3,46616
repérage	SBC	8	3,76868	vôtre	PRO	7	3,45656
vis	SBC	8	3,76868	confidentialité	SBC	7	3,45656
tarification	SBC	8	3,76868	clone	SBC	6	3,45332
répétitif	ADJ	10	3,76836	vouloir	ADJPAR	26	3,45143

sommaire différé	SBC SBC	5 5	3,44807 3,44807	récupération non	SBC ADV	14 452	3,29692 3,2746
super	ADJ	9	3,43523	intervalle	SBC	14	3,272
nature	SBC	107	3,41994	précédent	ADJ	120	3,27066
convenir	VB	28	3,40879	associer	VB	40	3,26875
électroniquement	ADV	3	3,39922	télécopieur	SBC	5	3,26765
simuler	ADJPAR	3	3,39922	exact	ADJ	32	3,26475
franciser	ADJPAR	3	3,39922	superposer	VB	8	3,25923
omettre	ADJPAR	3	3,39922	enfin	ADV	241	3,2388
apostrophe	SBC	3	3,39922	maillon	SBC	9	3,22915
concevable	ADJ	7	3,39805	rébarbatif	ADJ	4	3,21516
propagation	SBC	7	3,39805	bouter	VB	4	3,21516
série	SBC	134	3,39787	imbrication	SBC	4	3,21516
sinon	PREP	59	3,39343	hiérarchie	SBC	29	3,21497
offrir	ADJPAR	41	3,38684	hyper	PFX	3	3,20039
orthographe	SBC	6	3,38167	cambouis	SBC	3	3,20039
imbriquer	ADJPAR	6	3,38167	mathématiquement	ADV	3	3,20039
amorce	SBC	10	3,37218	souplesse	SBC	21	3,1962
potentiellement	ADV	10	3,37218	énormément	ADV	10	3,19307
relâcher	VB	10	3,37218	organisationnel	ADJ	5	3,18305
flot	SBC	20	3,35616	amplification	SBC	5	3,18305
étiqueter	ADJPAR	5	3,35589	vous	SBP	6	3,17957
parallélisme	SBC	5	3,35589	différer	VB	16	3,1743
émergent	ADJ	8	3,35332	primaire	ADJ	36	3,17012
diamètre	SBC	8	3,35332	dernière	SBC	42	3,16509
participant	ADJ	8	3,35332	appliquer	VB	73	3,13833
fortement	ADV	53	3,34755	déroulement	SBC	19	3,13288
pilote	SBC	34	3,34511	biais	SBC	25	3,13265
paye	SBC	4	3,33648	sillon	SBC	6	3,1161
aération	SBC	4	3,33648	lancer	ADJPAR	92	3,11365
synthétiser	VB	4	3,33648	connaissance	SBC	62	3,11091
cacher	VB	61	3,3344	concis	ADJ	5	3,10181
procurer	VB	25	3,3313	valoir	ADJPAR	4	3,10127
tracer	VB	18	3,31246	socioculturel	ADJ	4	3,10127
propre	ADJ	185	3,30751	stationnaire	ADJ	4	3,10127
notation	SBC	14	3,29692	confort	SBC	19	3,09755
brièvement	ADV	14	3,29692	inventaire	SBC	17	3,0947

## ANNEXE 4 – Liste des S+ obtenues à l'aide de M2

## SC1 – Abc de l'informatique

Forme	PdeD	Fréq. abs.	Valeur- test	Forme	PdeD	Fréq. abs.	Valeur- test
mémoire	SBC	737	31,8951	jet	SBC	25	6,9507
ordinateur	SBC	563	19,807	appareil	SBC	54	6,93458
disque	SBC	482	18,4371	mourir	ADJPAR	31	6,92127
unité	SBC	302	16,1886	pouce	SBC	52	6,87064
appuyer	VB	168	15,9702	superviseur	SBC	21	6,8626
texte	SBC	298	15,8269	imprimante	SBC	174	6,74775
touche	SBC	371	15,8267	système	SBC	680	6,73547
disquette	SBC	280	14,7883	ce	PRO	668	6,72806
lecteur	SBC	226	14,7757	mégaoctet	SBC	16	6,7014
un	DT	4219	14,6717	petit	ADJ	112	6,65802
microprocesseur	SBC	80	14,5548	encre	SBC	30	6,59411
processeur	SBC	150	13,1188	relocalisation	SBC	14	6,52682
wordperfect	SBP	49	12,2422	munir	ADJPAR	23	6,51925
dur	ADJ	182	11,7476	sur	PREP	1008	6,48631
carte-mère	SBC	48	11,4087	entrée-sortie	SBC	25	6,28842
bit	SBC	173	11,3835	chiffre	SBC	50	6,28257
exploitation	SBC	196	11,3652	nanoseconde	SBC	13	6,26694
principal	ADJ	183	11,2143	arborescent	ADJ	20	6,21321
servir	VB	143	11,1396	copier	VB	56	6,19052
écran	SBC	286	11,0131	déplacer	VB	59	6,16434
magnétique	ADJ	59	10,762	glisser	VB	25	6,13663
temps	SBC	222	10,4982	interne	ADJ	56	6,06465
segment	SBC	40	9,86564	cycle	SBC	41	6,05638
vif	ADJ	65	9,8363	taper	VB	104	6,00948
rigide	ADJ	58	9,65255	simultanément	ADV	42	5,96932
seconde	SBC	73 50	9,38224	plus	ADV	706	5,96361
franc	SBC	52	9,21154	généralement	ADV SBC	102	5,96052
capacité couleur	SBC SBC	104 97	9,07174 9,01285	fréquence génération	SBC	71 35	5,94187 5,94179
implicite	ADJ	30	8,89425	extension	SBC	99	5,92316
dessin	SBC	46	8,86797	animation	SBC	17	5,90377
surface	SBC	41	8,82763	double	ADJ	41	5,88899
chargeur	SBC	29	8,71025	lumineux	ADJ	20	5,83128
OU	COO	918	8,65228	instruction	SBC	114	5,81712
multiprogrammation	SBC	23	8,51764	deux	CAR	290	5,80941
d'	PREP	2741	8,40466	prix	SBC	72	5,77028
assembleur	SBC	27	8,33115	fenêtre	SBC	83	5,76416
secteur	SBC	73	8,26904	secondaire	ADJ	37	5,73185
dispositif	SBC	44	8,09308	puis	ADV	104	5,72802
curseur	SBC	98	7,88125	wordperfect	COO	11	5,71217
connecteur	SBC	35	7,79882	paintbrush	SBP	11	5,71217
barrette	SBC	38	7,67659	taille	SBC	119	5,69211
lui	PRO	3987	7,63608	ventilateur	SBC	17	5,68201
piste	SBC	59	7,6333	boîtier	SBC	33	5,67825
posséder	VB	84	7,5456	externe	ADJ	48	5,63787
appeler	ADJPAR	129	7,38354	ibm	SBP	66	5,63406
vitesse	SBC	84	7,34215	vouloir	ADJPAR	18	5,52908
enregistrement	SBC	81	7,15587	plateau	SBC	18	5,52908
central	ADJ	78	7,1213	routine	SBC	18	5,52908
traitement	SBC	119	7,08738	appeler	VB	89	5,51659
dram	SBP	21	7,08674	sauvegarder	VB	41	5,49054
clavier	SBC	187	6,95317	pièce	SBC	17	5,47231

				•			
rapport	SBC	68	5,43641	relocalisable	ADJ	7	4,40097
chaleur	SBC	11	5,34163	multimédia	SBC	13	4,36199
programme	SBC	399	5,31842	espace	SBC	69	4,3614
séance	SBC	13	5,31805	vectoriel	ADJ	9	4,34541
ami	SBC	12	5,31531	silicium	SBC	9	4,34541
occuper	VB	38	5,30999	palette	SBC	8	4,34266
tête	SBC	48	5,30823	photographie	SBC	8	4,34266
interruption	SBC	33	5,2947	emplacement	SBC	35	4,34057
transfert	SBC	50	5,28452	souris	SBC	94	4,3329
						-	
lorsque	SUB	146	5,25952	haut	SBC	42	4,27968
formatage	SBC	28	5,23987	bloc	SBC	42	4,27968
caractériser	ADJPAR	19	5,2389	rayon	SBC	12	4,27169
octet	SBC	59	5,20974	pc	SBC	92	4,26717
			,			26	
son	SBC	33	5,20308	déplacement	SBC		4,22899
pointeur	SBC	25	5,1951	bus	SBC	49	4,22775
alors	ADV	162	5,16873	fonction	SBC	175	4,22512
former	VB	24	5,10318	distinguer	VB	30	4,21756
condensateur	SBC	9	5,09817	circuit	SBC	33	4,21396
assistance	SBC	14	5,0965	entrer	ADJPAR	20	4,1978
pagination	SBC	14	5,0965	marquer	ADJPAR	15	4,19556
stocker	VB	45		clic	SBC	21	
			5,05461				4,19252
cadencer	ADJPAR	10	5,02879	cliquer	VB	63	4,15283
illustration	SBC	12	5,02116	désigner	VB	26	4,13556
				_			
effacer	VB	46	5,01155	entrer	VB	62	4,11705
consister	VB	58	4,9754	gagner	VB	17	4,11521
soit	COO	72	4,95908	famille	SBC	18	4,09428
assemblage	SBC	15	4,93732	courant	SBC	23	4,0923
rapide	ADJ	60	4,93039	mais	COO	305	4,08007
•							
aussi	ADV	199	4,91055	tension	SBC	21	4,07809
varier	VB	36	4,89038	transistor	SBC	12	4,05654
densité	SBC	14	4,8688	carte	SBC	223	4,05136
image	SBC	133	4,85127	implanter	ADJPAR	9	4,03305
commande	SBC	463	4,8004	poussière	SBC	9	4,03305
métal	SBC	8	4,76207	sillon	SBC	6	4,00841
boule	SBC	8	4,76207	latence	SBC	6	4,00841
printmaster	SBP	8	4,76207	adressable	ADJ	6	4,00841
travail	SBC	116	4,74303	multiprogrammer	ADJPAR	6	4,00841
	ADJ	20		à	PREP	2239	
préférable			4,719				4,00378
kilo	SBC	11	4,70955	jusqu'	PREP	60	4,00086
tirage	SBC	9	4,69702	électrique	ADJ	23	3,9902
recouvrir	ADJPAR	9	4,69702	magnétoscope	SBC	8	3,98029
coefficient	SBC	10	4,68734	puce	SBC	27	3,9783
volatil	ADJ	10	4,68734	enregistrer	ADJPAR	30	3,97732
		-					
amateur	SBC	10	4,68734	rapidité	SBC	17	3,97558
stockage	SBC	32	4,67042	formater	ADJPAR	13	3,97086
enfoncer	ADJPAR	18	4,66739	ainsi	ADV	214	3,96369
moins	ADV	134	4,6208	flécher	ADJPAR	7	3,96074
moment	SBC	43	4,62039	mini-ordinateur	SBC	7	3,96074
ex	PFX	13	4,57652	brancher	VB	20	3,96047
allocation	SBC	17	4,56862	façon	SBC	103	3,96009
tampon	SBC	23	4,52672	tandis	PUL	19	3,95898
ordre	SBC	65	4,52465	ordinaire	ADJ	11	3,95283
parler	VB	46	4,52217	lecture	SBC	56	3,94098
•							
tube	SBC	24	4,52067	écrire	VB	66	3,93094
endroit	SBC	32	4,5043	contrôleur	SBC	35	3,92065
apprendre	VB	19	4,4774	afficher	VB	103	3,90246
horloge	SBC	42	4,46608	produire	VB	44	3,88589
celui-ci	PRO	116	4,45971	impression	SBC	80	3,85279
coupure	SBC	11	4,43594	instrument	SBC	10	3,85073
comporter	VB	50	4,42141	positionner	VB	17	3,84101
flèche	SBC	17		cher	ADJ	28	
			4,41128				3,83874
mutuel	ADJ	7	4,40097	autre	PRO	123	3,80522

défilement	SBC	13	3,79158	donner	ADJPAR	215	3,3863
traiter	VB	35	3,7877	calculatrice	SBC	8	3,37876
qualité	SBC	72	3,77152	segmentation	SBC	8	3,37876
maintenant	ADV	55	3,76661	affichage	SBC	67	3,35685
environ	SBC	9	3,75246	équiper	ADJPAR	21	3,35543
partageable	ADJ	9	3,75246	remplacer	VB	39	3,35023
alimenter	ADJPAR	9	3,75246	commander	VB VB	11	3,34836
						118	
gauche	ADJ	20	3,73639	exister	VB		3,3429
musique	SBC	15	3,72684	taper	ADJPAR	12	3,31898
étendre	ADJPAR	27	3,72651	auxiliaire	ADJ	15	3,31386
copier	ADJPAR	19	3,72341	pentium	SBP	13	3,30655
conserver	VB	31	3,71026	simultané	ADJ	13	3,30655
éditeur	SBC	53	3,68757	éliminer	VB	14	3,30595
pourquoi	ADV	37	3,67386	contenir	VB	172	3,3057
former	ADJPAR	12	3,66599	sauvegarde	SBC	34	3,30236
pourvoir	ADJPAR	12	3,66599	certain	DT	141	3,29792
repérer	VB	12	3,66599	qu'	SUB\$	281	3,29726
intel	SBP	21	3,64974	trop	ADV	47	3,29601
venir	VB	53	3,64186	accomplir	VB	9	3,26551
gestionnaire	SBC	48	3,6281	sophistiquer	ADJPAR	9	3,26551
combien	ADV	13	3,62169	restitution	SBC	7	3,26355
avoir	VBA	973	3,61381	professionnel	SBC	7	3,26355
vidéo	SBC	30	3,6012	amorce	SBC	7	3,26355
employer	VB	14	3,59603	report	SBC	7	3,26355
confirmer	VB	14	3,59603	copie	SBC	46	3,25161
scanner	SBC	14	3,59603	vrai	ADJ	25	3,25094
multiplicateur	ADJ	7	3,58715	dire	ADJPAR	32	3,21171
enceindre	ADJPAR	7	3,58715	constructeur	SBC	22	3,21022
gigaoctet	SBC	7	3,58715	mentionner	VB	10	3,20324
refermer	VB	7	3,58715	marge	SBC	15	3,18646
majuscule	ADJ	17	3,58572	différence	SBC	30	3,18463
paragraphe	SBC	25	3,57831	pointer	VB	11	3,17139
alvéole	SBC	5	3,57469	perdre	ADJPAR	14	3,16983
refroidissement	SBC	5	3,57469	bande	SBC	43	3,16437
télécopieur	SBC	5	3,57469	marche	SBC	13	3,15983
risc	SBP	5	3,57469	intensité	SBC	6	3,15941
cadence	SBC	5	3,57469	volumineux	ADJ	6	3,15941
sybex	SBP	5	3,57469	allumer	VB	6	3,15941
enduire	ADJPAR	5	3,57469	demain	ADV	6	3,15941
photographique	ADJ	5	3,57469	téléviseur	SBC	6	3,15941
structure	SBC	72	3,55429	fermeture	SBC	6	3,15941
coprocesseur	SBC	6	3,54431	lit	SBC	6	3,15941
crayon	SBC	6	3,54431	graver	ADJPAR	6	3,15941
contour	SBC	6	3,54431	diviser	ADJPAR	12	3,15903
dactylographie	SBC	6	3,54431	séquentiel	ADJ	12	3,15903
chauffer	VB	6	3,54431	micro-ordinateur	SBC	38	3,13977
portatif	ADJ	6	3,54431	seulement	ADV	59	3,13636
air	SBC	11	3,53656	marque	SBC	24	3,12832
aucun	PRO	13	3,46028	lire	VB	56	3,12511
vide	ADJ	24	3,45876	emmagasiner	VB	8	3,12384
grâce	SBC	74	3,45605	doigt	SBC	8	3,12384
fond	SBC	16	3,4523	implicitement	ADV	8	3,12384
pavé	SBC	16	3,4523	impulsion	SBC	8	3,12384
contre	PREP	40	3,44915	poudre	SBC	8	3,12384
voilà	PREP	15	3,44613	argent	SBC	8	3,12384
trois	CAR	78	3,44081	aussitôt	ADV	8	3,12384
insérer	VB	78 38	3,44061 3,4191	thermique	ADV	8	3,12384
	VB VB	55		sortir	VB	6 22	
imprimer	SBC		3,41519		ADJ	22 17	3,12349
invite		10 47	3,40311	interactif	MDJ	1 /	3,12267
sortie	SBC	47	3,39154	l			

### SC2 – Internet

Forme	PdeD	Fréq. abs.	Valeur- test	Forme	PdeD	Fréq. abs.	Valeur- test
internet	SBC	786	42,8164	corba	SBP	30	10,3594
site	SBC	306	29,9967	requête	SBC	93	10,2526
recherche	SBC	346	29,8776	québec	SBP	29	10,1698
web	SBC	303	29,3781	priver	ADJPAR	47	10,0563
électronique	ADJ	202	20,6802	altavista	SBP	27	10,0395
france	SBP	154	20,4047	messagerie	SBC	42	10,0068
service	SBC	303	18,8849	technique	ADJ	113	9,92524
moteur	SBC	115	18,5328	société	SBC	59	9,82923
courrier	SBC SBC	136 215	18,4953 17,6287	protocole social	SBC ADJ	190 29	9,79367 9,69919
message infrastructure	SBC	89	16,0484	destinataire	SBC	29	9,69919
document	SBC	228	16,0367	internationalisation	SBC	26	9,57877
public	ADJ	100	15,1919	trafic	SBC	37	9,55999
langue	SBC	82	14,7627	francophone	ADJ	28	9,4994
opérateur	SBC	111	14,6211	page	SBC	208	9,48137
technologie	SBC	127	14,5945	communauté	SBC	33	9,47733
français	ADJ	93	14,3868	mondial	ADJ	30	9,46709
international	ADJ	64	14,3728	fournisseur	SBC	82	9,4602
zope	SBP	53	14,0354	mot	SBC	163	9,33177
serveur	SBC	402	13,8983	navigation	SBC	39	9,33124
interconnexion	SBC	57	13,5674	norme	SBC	63	9,31612
formulaire	SBC	55	13,2824	croissance	SBC	29	9,26633
fibre	SBC	49	13,2767	usenet	SBP	23	9,22906
national	ADJ	57	12,9597	publication	SBC	26	9,08745
netscape	SBP	48	12,9415	adresse	SBC	223	9,06469
transport	SBC	56	12,8146	million	SBC	59	8,97358
télécommunication	SBC	53	12,5207	déployer	VB	23	8,94948
réseau	SBC	466	12,4467	expérimentation	SBC	25	8,87479
offre	SBC	72	12,3268	produit	SBC	56	8,7702
pays	SBC	71	12,2919	initiative	SBC	30	8,70094
débit	SBC	87	12,2464	canada	SBP	24	8,65727
acteur	SBC	51	12,0662	organisme	SBC	26	8,63939
déploiement	SBC	51	12,0662	communication	SBC	119	8,57328
europe	SBP	58	12,0233	réglementaire	ADJ	20	8,57121
navigateur	SBC	66	11,809	gratuit	ADJ	36	8,49666
domaine	SBC	125	11,8009	monde	SBC	89	8,42284
usage	SBC	98	11,6691	régional	ADJ	24	8,42233
économique	ADJ	69 66	11,6291	favoriser	VB	25 52	8,42135
diffusion	SBC SBC	66 39	11,6052	téléphonique dollar	ADJ SBC	53 30	8,39267 8,35562
abonné satellite	SBC	35 35	11,516 11,4905	multilingue	ADJ	19	8,34045
université	SBC	41	11,4905	abonnement	SBC	27	8,25303
forum	SBC	41	11,4636	autoroute	SBC	21	8,23063
États-unis	SBP	43	11,2652	accentuer	ADJPAR	31	8,08347
développement	SBC	146	11,0369	et	COO	2127	8,07301
sujet	SBC	65	10,9771	métropolitain	ADJ	20	7,99423
entreprise	SBC	109	10,8463	administration	SBC	45	7,88909
information	SBC	343	10,8432	expéditeur	SBC	17	7,8587
équipement	SBC	66	10,8399	législation	SBC	17	7,8587
européen	ADJ	44	10,7914	industriel	ADJ	26	7,84834
internaute	SBC	32	10,7287	personne	SBC	77	7,83703
français	SBC	60	10,6435	concurrence	SBC	25	7,81132
boucle	SBC	62	10,6426	opportunité	SBC	18	7,79119
américain	ADJ	43	10,6226	culture	SBC	22	7,73576
yahoo	SBP	30	10,6069	depuis	PREP	67	7,70541
discussion	SBC	37	10,5874	étiquetage	SBC	17	7,53884

région	SBC	22	7,51743	conduit	SBC	12	6,13009
supporter	VB	47	7,51633	amérique	SBP	12	6,13009
annuaire	SBC	31	7,37331	émergence	SBC	14	6,11671
institution	SBC	15	7,34559	encourager	VB	14	6,11671
	SBC	15	7,34559	débat	SBC	13	6,10927
citoyen universitaire	ADJ	15	7,34559	optique	ADJ	31	6,08723
commercial	ADJ	42	7,34339	formation	SBC	20	
	SBC	16	7,28003	participer	VB	17	6,05938 6,01727
hébergement abonner	ADJPAR	16	7,27804	stratégique	ADJ	17	6,01727
retard	SBC	25	7,26993	voie	SBC	29	5,99237
accès	SBC	202	7,20333	avérer	VB	19	5,97538
recommandation	SBC	202	7,24865	étude	SBC	24	5,97453
publicité	SBC	20	7,24865	paris	SBP	16	5,95251
culturel	ADJ	17	7,24194	échange	SBC	37	5,9505
hypertexte	ADJ	19	7,23221	science	SBC	15	5,89449
tirer	VB	35	7,20115	presse	SBC	15	5,89449
année	SBC	62	7,08037	indexation	SBC	10	5,87145
portail	SBC	14	7,07518	communiste	SBC	10	5,87145
gouvernement	SBC	14	7,07518	cyberbavardage	SBC	10	5,87145
cryptologie	SBC	14	7,07518	en	PREP	1521	5,87009
arpanet	SBP	14	7,07518	signature	SBC	14	5,84593
terrestre	ADJ	14	7,07518	marché	SBC	54	5,82738
évolution	SBC	50	7,06316	télécharger	VB	21	5,82491
laboratoire	SBC	15	7,00792	indexer	VB	13	5,81085
publicitaire	ADJ	16	6,97453	connectivité	SBC	12	5,79547
parti	SBC	18	6,97337	enseignement	SBC	12	5,79547
éducation	SBC	17	6,96518	protection	SBC	32	5,79461
technologique	ADJ	29	6,95479	format	SBC	100	5,78114
catalogue	SBC	25	6,94052	matière	SBC	35	5,7478
an	SBC	43	6,84531	téléphonie	SBC	16	5,73457
banque	SBC	20	6,81425	disponibilité	SBC	18	5,70763
desserte	SBC	13	6,7941	outil	SBC	122	5,68944
investissement	SBC	30	6,78227	situation	SBC	39	5,67816
article	SBC	49	6,76135	solution	SBC	87	5,66894
canal	SBC	41	6,75149	local	ADJ	122	5,63511
radio	SBC	26	6,71338	offrir	VB	60	5,62589
grand	SBC	17	6,70617	index	SBC	27	5,56439
normalisation	SBC	15	6,69738	du	DT	2964	5,5604
mois	SBC	37	6,65916	contribution	SBC	13	5,53605
origine	SBC	39	6,56964	sous-marin	ADJ	9	5,53024
bibliothèque	SBC	36	6,5673	audiovisuel	ADJ	9	5,53024
telnet	SBP	18	6,50598	capital	SBC	9	5,53024
bibliographique	ADJ	12	6,50097	dérégulation	SBC	9	5,53024
satellite	ADJ	12	6,50097	macromolécule	SBC	9	5,53024
académique	ADJ	12	6,50097	booléen	ADJ	9	5,53024
commerce	SBC ADJ	28	6,48176	conférence	SBC SBP	16 12	5,52857
canadien flot	SBC	13 16	6,43532 6,42983	eric	SBC	12	5,49111 5,49111
son	DT	442	6,42303	prestataire obstacle	SBC	12	5,49111
robot	SBC	14	6,40939	archive	SBC	12	5,49111
voix	SBC	19	6,3535	conduire	VB	30	5,48694
géographique	ADJ	18	6,29099	standardisation	SBC	10	5,47227
oeuvre	SBC	34	6,28187	allemagne	SBP	10	5,47227
usager	SBC	39	6,22902	participation	SBC	11	5,46608
webstar	SBP	11	6,19412	molécule	SBC	11	5,46608
progressiste	ADJ	11	6,19412	ratio	SBC	11	5,46608
communiste	ADJ	11	6,19412	antenne	SBC	11	5,46608
historique	ADJ	16	6,18373	réglementation	SBC	11	5,46608
naviguer	VB	19	6,16008	client	SBC	122	5,39417
accueil	SBC	15	6,14344	nouveau	ADJ	160	5,38174
phrase	SBC	27	6,13973	organisation	SBC	49	5,36468
convergence	SBC	12	6,13009	routeur	SBC	18	5,36158

tálánhana	SBC	O.E.	E 0/170	Lananyma	AD I	0	4 75040
téléphone		25	5,34178	anonyme	ADJ	9	4,75042
utilisation	SBC	115	5,29883	vie	SBC	29	4,74476
travers	SBC	33	5,28413	reproduire	VB	14	4,73615
loi	SBC	13	5,28163	place	SBC	73	4,73395
publier	ADJPAR	13	5,28163	carbone	SBC	8	4,7322
plateforme	SBC	13	5,28163	québécois	ADJ	8	4,7322
idée	SBC	29	5,2813	indexer	ADJPAR	8	4,7322
transactionnel	ADJ	17	5,25518	solaris	SBP	8	4,7322
certification	SBC	12	5,21232	alcatel	SBP	8	4,7322
tiers	SBC	12	5,21232	entre	PREP	206	4,68296
compétence	SBC	18	5,19937	auprès	ADV	15	4,67832
au	DT	740	5,19616	contexte	SBC	21	4,62842
contenu	SBC	104	5,19423	localisation	SBC	13	4,6164
fourniture	SBC	8	5,16693	tarif	SBC	11	4,61379
audiovisuel	SBC	8	5,16693	couvrir	VB	11	4,61379
toile	SBC	8			CAR	11	
			5,16693	cent	-		4,61379
minitel	SBC	8	5,16693	logiciel	SBC	269	4,56591
saule	SBC	8	5,16693	traditionnel	ADJ	23	4,56016
binette	SBC	8	5,16693	liste	SBC	120	4,55324
administré	SBC	8	5,16693	concepteur	SBC	14	4,55117
moteur	ADJ	8	5,16693	politique	ADJ	14	4,55117
territoire	SBC	8	5,16693	câbler	ADJPAR	10	4,51576
pleureur	ADJ	8	5,16693	maîtrise	SBC	10	4,51576
office	SBP	11			SBC	10	,
			5,15561	pertinence			4,51576
entreprendre	ADJPAR	16	5,14785	population	SBC	10	4,51576
média	SBC	16	5,14785	payant	ADJ	10	4,51576
vigueur	SBC	10	5,11872	leur	DT	200	4,51307
financement	SBC	10	5,11872	corps	SBC	16	4,48736
marqueur	SBC	10	5,11872	développer	VB	30	4,46022
tridimensionnel	ADJ	9	5,11436	demander	ADJPAR	30	4,46022
télématique	SBC	9	5,11436	acrobat	SBP	9	4,4276
transit	SBC	9	5,11436	carnet	SBC	9	4,4276
concentration	SBC	9	5,11436	responsabilité	SBC	9	
							4,4276
transposition	SBC	9	5,11436	dynamique	SBC	9	4,4276
gratuitement	ADV	15	5,03987	officiel	ADJ	13	4,42121
principalement	ADV	27	4,95531	milliard	SBC	13	4,42121
transmission	SBC	34	4,92286	mettre	ADJPAR	48	4,41466
développer	ADJPAR	34	4,92286	chercher	VB	39	4,40787
diffuser	VB	11	4,87295	téléchargement	SBC	14	4,37568
notice	SBC	11	4,87295	accent	SBC	11	4,37473
explorer	VB	13	4,8238	cuivre	SBC	8	4,35728
accord	SBC	13	4,8238	spécialiser	ADJPAR	34	4,35558
	SBC	31		fédéral	ADJEAN	6	
centre			4,81455				4,35251
chercheur	SBC	10	4,80207	élève	SBC	6	4,35251
intercontinental	ADJ	7	4,77666	francisation	SBC	6	4,35251
transatlantique	ADJ	7	4,77666	pseudonyme	SBC	6	4,35251
confidentialité	SBC	7	4,77666	consommateur	SBC	6	4,35251
périodique	SBC	7	4,77666	contractuel	ADJ	6	4,35251
éducatif	ADJ	7	4,77666	marchand	ADJ	6	4,35251
audience	SBC	7	4,77666	partenariat	SBC	6	4,35251
enseignant	ADJ	7	4,77666	terme	SBC	57	4,3493
mensuel	ADJ	7			ADJ	18	
			4,77666	original			4,34717
acheminement	SBC	7	4,77666	thématique	ADJ	7	4,32048
enseignant	SBC	7	4,77666	nord	SBC	7	4,32048
planétaire	ADJ	7	4,77666	francophone	SBC	7	4,32048
progressiste	SBC	7	4,77666	moyen	SBC	49	4,31696
pôle	SBC	7	4,77666	thème	SBC	21	4,29549
nantes	SBP	7	4,77666	normaliser	ADJPAR	12	4,28845
visiteur	SBC	7	4,77666	proposer	VB	57	4,2643
informatique	ADJ	, 72	4,76409	distance	SBC	37	4,24809
	SBC	9			ADJ	84	
commutation			4,75042	disponible			4,24553
mars	SBC	9	4,75042	politique	SBC	22	4,22491

ouverture	SBC	24	4,20659	postal	ADJ	9	3,8759
	VB	17					
contribuer			4,20616	salle	SBC	9	3,8759
analogique	ADJ	17	4,20616	américain	SBC	9	3,8759
transaction	SBC	26	4,2049	militaire	ADJ	6	3,87177
économie	SBC	15	4,19705	consensus	SBC	6	3,87177
absence	SBC	20	4,15754	email	SBC	6	3,87177
	SBC					6	
succès		20	4,15754	telnet	SBC		3,87177
atout	SBC	11	4,15302	hertzien	ADJ	6	3,87177
janvier	SBC	11	4,15302	héberger	ADJPAR	6	3,87177
•							
progrès	SBC	11	4,15302	guerre	SBC	6	3,87177
frais	SBC	9	4,13806	lucratif	ADJ	6	3,87177
limiter	ADJPAR	38	4,12107	privé	SBC	6	3,87177
passant	ADJ	24	4,11641	offrir	ADJPAR	18	3,86874
comparable	ADJ	12	4,09396	groupe	SBC	79	3,8614
				• .			
devenir	VB	46	4,06571	évoluer	VB	20	3,84466
résultat	SBC	80	4,06437	production	SBC	17	3,83344
				· .			
digital	ADJ	13	4,06247	clé	SBC	46	3,82582
compression	SBC	16	4,0604	atteindre	VB	29	3,82229
échanger	VB	16	4,0604	hétérogène	ADJ	10	3,79407
pertinent	ADJ	14	4,04973	notre	DT	60	3,77751
consortium	SBC	8	4,02856	échelle	SBC	11	3,75311
			,				
exclure	VB	8	4,02856	néanmoins	ADV	14	3,75247
descendant	ADJ	8	4,02856	membre	SBC	14	3,75247
montant	ADJ	8	4,02856	école	SBC	8	3,73648
codage	SBC	24	4,02804	raccorder	VB	8	3,73648
poster	VB	10	4,01526	constitution	SBC	8	3,73648
•						8	
autorité	SBC	10	4,01526	évidence	SBC		3,73648
envoyer	VB	54	3,98561	nombreux	ADJ	61	3,71706
expérience	SBC	20	3,94627	relation	SBC	30	3,67613
consommateur	ADJ	7	3,93416	impact	SBC	18	3,65031
sol	SBC	7	3,93416	pratique	SBC	20	3,64884
raccordement	SBC	7	3,93416	film	SBC	9	
			,				3,63663
comité	SBC	7	3,93416	enjeu	SBC	9	3,63663
participant	SBC	7	3,93416	ville	SBC	9	3,63663
balise	SBC	7	3,93416	contenir	ADJPAR	34	3,60509
par	PREP	1022	3,91878	encoder	VB	7	3,60019
récemment	ADV	13	3,89685	illicite	ADJ	7	3,60019
titre	SBC	40	3,8947	interopérabilité	SBC	7	3,60019
candidat	SBC	5	3,88406	pédagogique	ADJ	7	3,60019
différé	SBC	5	•	élargir	VB	7	3,60019
			3,88406				
dorsal	ADJ	5	3,88406	coder	ADJPAR	19	3,59704
coupe-feu	SBC	5	3,88406	standard	SBC	25	3,59184
•	SBC	5			SBC	13	
conversation		-	3,88406	consultation			3,58898
worldcom	SBP	5	3,88406	métier	SBC	10	3,58872
massivement	ADV	5	3,88406	applicatif	ADJ	10	3,58872
saule-pleureur	SBC	5	3,88406	transparence	SBC	10	3,58872
royaume-uni	SBP	5	3,88406	directive	SBC	10	3,58872
planète	SBC	5	3,88406	viser	VB	11	3,57151
-							
pagemill	SBP	5	3,88406	obligation	SBC	11	3,57151
numérisation	SBC	5	3,88406	mécanisme	SBC	62	3,55096
multiplexage	SBC	5	3,88406	sur	PREP	833	3,53684
angleterre	SBP	5	3,88406	trouver	VB	155	3,53628
médium	SBC	5	3,88406	largement	ADV	19	3,49909
législatif	ADJ	5	3,88406	intérêt	SBC	27	
							3,4948
troncature	SBC	5	3,88406	négociation	SBC	6	3,47404
globe	SBC	5	3,88406	émergent	ADJ	6	3,47404
universitaire	SBC	5	3,88406	novembre	SBC	6	3,47404
exportation	SBC	5	3,88406	napster	SBP	6	3,47404
explosion	SBC	5	3,88406	public	SBC	6	3,47404
exciter	VB	5	3,88406	relais	SBC	6	3,47404
héberger	VB	9	3,8759	apport	SBC	6	3,47404
fax	SBC	9	3,8759	intellectuel	ADJ	6	3,47404
iux	350	5	5,0755	Intollootagi	, 100	0	U, T I TUT

sécuriser textuel globalisation foyer dans industrie mille massif binaire confiance parcourir annoncer exploiter asymétrique guise italie décisif	VB ADJ SBC PREP SBC CAR ADJ ADJ SBC VB ADJPAR VB ADJ SBC SBP ADJ SBC	6 6 6 6 1012 12 9 46 11 11 10 17 5 5 5	3,47404 3,47404 3,47404 3,47404 3,44349 3,41802 3,41673 3,41673 3,41618 3,40035 3,40035 3,39718 3,3853 3,37512 3,37512 3,37512 3,37512	patrimoine pionnier poster rachat realaudio sprint subsidiarité majoritairement colossal continental concertation démocratique vidéo-conférence déréglementation énumérer commuter domination	SBC SBC SBP SBP SBC ADV ADJ ADJ SBC ADJ SBC ADJ SBC ADJPAR ADJPAR ADJPAR SBC SBC	4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4	3,3543 3,3543 3,3543 3,3543 3,3543 3,3543 3,3543 3,3543 3,3543 3,3543 3,3543 3,3543 3,3543 3,3543 3,3543
stage réussite	SBC SBC	5 5	3,37512 3,37512	vinton décentralisation	SBP SBC	4 4	3,3543 3,3543
rémunération	SBC	5	3,37512	universel	ADJ	14	3,35073
vitrine hypertexte	SBC SBC	5 5	3,37512 3,37512	introduction auteur	SBC SBC	16 34	3,33372 3,3214
reilly	SBP	5	3,37512	réflexion	SBC	7	3,30671
baladeur	SBC	5	3,37512	administratif	ADJ	7	3,30671
promouvoir coordination	VB SBC	5 5	3,37512 3,37512	routeur avril	ADJ SBC	7 7	3,30671
préoccupation	SBC	5	3,37512	proposer	ADJPAR	7 25	3,30671 3,29649
continent	SBC	5	3,37512	conforme	ADJ	15	3,27966
mutation	SBC	5	3,37512	objet	SBC	64	3,27532
nomade	ADJ	5	3,37512	sein	SBC	18	3,24781
nord-américain	ADJ	5	3,37512	collectif	ADJ	8	3,23604
bavardage	SBC	4	3,3543	déployer	ADJPAR	8	3,23604
bulletin	SBC	4	3,3543	télévision	SBC	8	3,23604
artère	SBC	4	3,3543	hôte	SBC	10	3,21777
alerte attractif	ADJ ADJ	4 4	3,3543 3,3543	naissance défi	SBC SBC	9 9	3,21341 3,21341
banalisation	SBC	4	3,3543	propre	ADJ	51	3,20806
modalité	SBC	4	3,3543	propriété	SBC	22	3,16981
fourreau	SBC	4	3,3543	scientifique	ADJ	15	3,16637
fureteur	SBC	4	3,3543	facteur	SBC	18	3,15352
grande-bretagne	SBP	4	3,3543	effort	SBC	18	3,15352
hébreu	SBC	4	3,3543	direct	SBC	6	3,13612
hyperlien	SBC	4	3,3543	intuitif	ADJ	6	3,13612
hypermédia chambre	ADJ SBC	4 4	3,3543 3,3543	pouvoir	SBC ADJPAR	6 6	3,13612
trésor	SBC	4	3,3543	mener engendrer	VB	6	3,13612 3,13612
interréseau	SBC	4	3,3543	financer	ADJPAR	6	3,13612
libéralisation	SBC	4	3,3543	classeur	SBC	6	3,13612
londres	SBP	4	3,3543	union	SBC	6	3,13612
forfait	SBC	4	3,3543	baptiser	ADJPAR	6	3,13612
métropole	SBC	4	3,3543	champ	SBC	77	3,13493
tronçon	SBC	4	3,3543	objectif	SBC	19	3,13012
mondialisation	SBC SBP	4 4	3,3543	élever version	ADJPAR SBC	31 94	3,12529
nancy parental	ADJ	4	3,3543 3,3543	oui	ADV	94 14	3,1122 3,10742
parution	SBC	4	3,3543	Juli	, (D v	17	5,10772

# SC3 – Logiciels

Forme	PdeD	Fréq. abs.	Valeur- test	Forme	PdeD	Fréq. abs.	Valeur- test
étendue	SBC	75	19,8745	présenter	VB	68	7,6376
cellule	SBC	94	18,5	base	SBC	164	7,62981
disquette	SBC	248	18,4584	indicatif	SBC	16	7,5885
unité	SBC	247	17,7622	atelier	SBC	14	7,46014
phase	SBC	82	16,2724	inventaire	SBC	14	7,46014
presser	VB	51	15,0242	tableur	SBC	21	7,32819
logiciel	SBC	343	15,0076	courbe	SBC	12	7,04977
examiner	VB	50	14,5031	avoir	VBA	765	7,03572
figure	SBC	80	13,7602	libellé	SBC	11	7,03498
enregistrement	SBC	94	13,5423	accumulateur	ADJ	11	7,03498
micro-informatique	ADJ	37	13,4938	écran	SBC	174	6,9382
colonne	SBC	78	13,1519	processus	SBC	97	6,82747
programmation	SBC	101	12,8178	stock	SBC	15	6,75236
projet	SBC	76	12,7726	structure	SBC	71	6,72973
micro-ordinateur	SBC	65	12,4665	nominal	ADJ	9	6,66294
équipe	SBC	47	12,0347	architecte	SBC	16	6,66284
entité	SBC	41	11,9325	illustrer	ADJPAR	11	6,65771
progiciel	SBC	34	11,5947	productivité	SBC	17	6,60625
test	SBC	84	11,5555	développement	SBC	91	6,58954
erreur	SBC	116	11,3024	décision	SBC	18	6,40544
chiffrier	SBC	25	11,2951	champ	SBC	82	6,33966
façon	SBC	126	11,2908	date	SBC	47	6,33417
entrer	VB CBC	82 228	11,2706	recouvrement	SBC	11	6,31793
donnée	SBC SBC		10,9162	axe	SBC VB	15 71	6,31581
sous-répertoire enchaînement	SBC	50 25	10,8239	apparaître	SBP	71 8	6,29388
	SBC	25 362	10,7462 10,5239	artur commis	SBC	8	6,22809 6,22809
programme instruction	SBC	116	10,3239	unitaire	ADJ	9	6,20554
graphe	SBC	22	10,4209	cycle	SBC	32	6,14112
sous-commande	SBC	19	10,2203	structuration	SBC	12	6,13186
activité	SBC	54	9,90012	ingénierie	SBC	12	6,13186
conception	SBC	39	9,6585	individu	SBC	16	6,08875
coût	SBC	72	9,534	usager	SBC	32	6,06461
parc	SBC	24	9,52445	taper	VB	78	6,02171
diagramme	SBC	17	9,45256	abstraction	SBC	13	6,02063
souhaitable	ADJ	25	9,3763	conceptuel	ADJ	11	6,00941
maintenance	SBC	41	9,23189	registre	SBC	30	6,00144
incitation	SBC	19	9,08761	case	SBC	27	5,96552
défaillance	SBC	18	9,08217	commande	SBC	349	5,96256
génie	SBC	20	8,8535	contenu	SBC	89	5,89967
présenter	ADJPAR	42	8,7776	critère	SBC	36	5,87242
état	SBC	80	8,73125	cohérence	SBC	9	5,80664
formule	SBC	39	8,71209	incorporer	VB	9	5,80664
donner	ADJPAR	211	8,6889	analyse	SBC	28	5,77605
accumulateur	SBC	14	8,51346	combinatoire	SBC	7	5,76115
modèle	SBC	117	8,36258	fractionnement	SBC	7	5,76115
présentement	ADV	19	8,30584	libeller	ADJPAR	7	5,76115
contrôle	SBC	80	8,19316	opérande	SBC	14	5,75685
langage	SBC	127	8,19127	surcharge	SBC	8	5,75017
incrément	SBC	13	8,17665	sauver	VB	8	5,75017
sûreté	SBC	15	8,12428	validation	SBC	16	5,74895
programmeur	SBC	42	8,10018	arbre	SBC	15	5,74377
graphique	SBC	19	8,07293	littéral	ADJ	11	5,72729
relation	SBC	39	7,89682	copier	VB	41	5,72583
fonction	SBC	161	7,89057	transaction	SBC	26	5,71687
curseur	SBC	74	7,73377	logique	ADJ	51	5,6667

				I			
qui	REL	683	5,62139	correctif	ADJ	8	4,66354
comptable	ADJ	10	5,58608	défectueux	ADJ	12	4,63893
effort	SBC	21	5,57243	ouvrage	SBC	12	4,63893
on	PRO	458	5,53032	point	SBC	85	4,62812
		41		finir	ADJPAR		
organisation	SBC		5,51011			9	4,59535
modélisation	SBC	11	5,46772	coordonner	VB	9	4,59535
intégration	SBC	25	5,4647	sous-ensemble	SBC	9	4,59535
constituant	ADJ	9	5,45402	disque	SBC	201	4,53666
appareil	SBC	36	5,43877	compilateur	SBC	31	4,52327
			,				
auxiliaire	ADJ	16	5,43729	expérience	SBC	18	4,48636
exercice	SBC	22	5,42841	byte	SBC	7	4,47368
employé	SBC	12	5,40832	quel	DT	62	4,46266
. ,				•			
théorie	SBC	12	5,40832	chapitre	SBC	41	4,44907
agenda	SBC	12	5,40832	correction	SBC	15	4,42648
abstraire	ADJPAR	10	5,30238	addition	SBC	11	4,41711
copie	SBC	43	5,26762	essentiel	ADJ	24	4,41158
sémaphore	SBC	7	5,25995	supposer	VB	28	4,41045
instrumentation	SBC	6	5,25394	durée	SBC	26	4,40431
informatiser	ADJPAR	6	5,25394	particulièrement	ADV	20	4,40274
ou	COO	587	5,22228	vérification	SBC	19	4,39176
couverture	SBC	14	5,21678	protéger	ADJPAR	13	4,38955
humain	ADJ	21	5,15125	principal	ADJ	89	4,38805
copier	ADJPAR	18	5,14562	important	ADJ	72	4,3858
parent	SBC	9	5,13877	mission	SBC	8	4,3793
•							
scénario	SBC	9	5,13877	habituellement	ADV	10	4,37618
rigide	ADJ	30	5,1177	relationnel	ADJ	12	4,31371
formater	VB	13	5,01984	assurer	VB	49	4,30005
effacer	VB	35	5,00541	nature	SBC	31	4,2967
déroulement	SBC	11	5,00439	ordinal	ADJ	6	4,29182
		12			ADJPAR	6	
caractéristique	ADJ		5,00092	incorporer	-		4,29182
rendement	SBC	8	4,98101	acte	SBC	6	4,29182
sport	SBC	8	4,98101	ci-dessus	ADV	29	4,26841
•							
événement	SBC	27	4,91852	déplacer	VB	37	4,26502
réalisation	SBC	18	4,91315	successif	ADJ	11	4,24387
	DT					94	
du		2267	4,88752	exploitation	SBC		4,21436
brièvement	ADV	9	4,85424	composante	SBC	14	4,21044
débuter	VB	13	4,85074	fiabilité	SBC	17	4,18496
frontière	SBC	7	4,8374	arithmétique	ADJ	17	4,18496
voyager	VB	7	4,8374	fondamental	ADJ	17	4,18496
file	SBC	26	4,8262	stockage	SBC	23	4,16242
=							
sous-menu	SBC	12	4,81485	ensemble	SBC	74	4,15633
satisfaire	VB	15	4,81402	résiduel	ADJ	5	4,13603
		11				5	
enfant	SBC		4,79602	séquencement	SBC		4,13603
construire	VB	20	4,7715	raisonner	VB	5	4,13603
approfondir	ADJPAR	6	4,72617	signe	SBC	24	4,129
				•			
abstraire	VB	6	4,72617	adopter	ADJPAR	8	4,1224
aptitude	SBC	6	4,72617	inverser	ADJPAR	8	4,1224
vie	SBC	24	4,7135	vouloir	VB	76	4,11809
fractionner	VB	5	4,69411	bon	ADJ	69	4,115
ha	INJ	5	4,69411	achat	SBC	15	4,07739
symbologie	SBC	5	4,69411	rôle	SBC	33	4,07361
équivalence	SBC	5	4,69411	verrouiller	ADJPAR	4	4,06163
automate	SBC	5	4,69411	attentivement	ADV	4	
							4,06163
mainteneur	SBC	5	4,69411	sandra	SBP	4	4,06163
métrologie	SBC	5	4,69411	complétude	SBC	4	4,06163
•							
micro-processeur	SBC	5	4,69411	armand	SBP	4	4,06163
compétiteur	SBC	5	4,69411	fidélité	SBC	4	4,06163
difficulté	SBC	22	4,68589	rétroaction	SBC	4	4,06163
couplage	SBC	8	4,66354	un	DT	2505	4,02847
sémantique	ADJ	8	4,66354	insérer	ADJPAR	12	4,01876
formulation	SBC	8	4,66354	nouveau	ADJ	116	3,98628
observation	SBC	8	4,66354	schéma	SBC	13	3,98506

	000	_	0.00404	La seco	45.	4.0	0.4007
cobol	SBP	6	3,92464	initial	ADJ	18	3,4337
représentatif	ADJ	6	3,92464	représentation	SBC	16	3,42022
adresser	ADJPAR	6	3,92464	etc	ADV	108	3,41114
phénomène	SBC	11	3,9245	fabrication	SBC	9	3,40564
syntaxique	ADJ	8	3,88832	trier	VB	7	3,39055
période	SBC	12	3,88106	extrême	ADJ	7	3,39055
livraison	SBC	7	3,87306	conseil	SBC	12	3,38396
dictionnaire	SBC	7	3,87306	vente	SBC	12	3,38396
terminologie	SBC	7	3,87306	identification	SBC	12	3,38396
tâche	SBC	45	3,84349	considérable	ADJ	10	3,37684
	ADV	10	3,8333	résulter	VB	11	3,37256
auparavant							
nombre	SBC	146	3,80506	opérer	VB	11	3,37256
entrée	SBC	64	3,79678	maintenant	ADV	39	3,36243
vierge	ADJ	11	3,77668	augmenter	VB	26	3,33093
. ~							
traitement	SBC	68	3,77432	évolutif	ADJ	6	3,32975
protéger	VB	15	3,76003	humain	SBC	5	3,32368
lot	SBC	12	3,74915	implanter	VB	5	3,32368
préparer	VB	12	3,74915	mortalité	SBC	3	3,31985
						5	
feuille	SBC	33	3,73471	multi-usager	ADJ	3	3,31985
contre	PREP	31	3,72173	pondération	SBC	3	3,31985
cause	SBC	19	3,71134	observateur	SBC	3	3,31985
besoin	SBC	72	3.70853	lucidité	SBC	3	3,31985
			-,				
et	COO	1503	3,70647	micro-instruction	SBC	3	3,31985
aérien	ADJ	5	3,69119	anticipation	SBC	3	3,31985
expansion	SBC	5	3,69119	paul	SBP	3	3,31985
prépondérant	ADJ	5	3,69119	imputable	ADJ	3	3,31985
• •					-		
vue	SBC	20	3,67892	saboter	VB	3	3,31985
préparer	ADJPAR	8	3,67356	sinistre	ADJ	3	3,31985
facture	SBC	8	3,67356	sous-schéma	SBC	3	3,31985
transition	SBC	10	3,67289	stimuli	SBC	3	3,31985
						3	
objectif	SBC	17	3,63227	synchronique	ADJ		3,31985
efficacité	SBC	13	3,62526	tarte	SBC	3	3,31985
régime	SBC	7	3,61983	absurdité	SBC	3	3,31985
allure	SBC	7	3,61983	voler	VB	3	3,31985
bâtir	VB	7	3,61983	analytique	ADJ	3	3,31985
rentabilité	SBC	6			ADV	3	
			3,60775	extraordinairement			3,31985
inévitable	ADJ	6	3,60775	consistance	SBC	3	3,31985
rigoureux	ADJ	6	3,60775	illustré	SBC	3	3,31985
robert	SBP	6	3,60775	formellement	ADV	3	3,31985
soigneusement	ADV	6	3,60775	dualité	SBC	3	3,31985
						3	,
duplication	SBC	6	3,60775	diagnostiquer	VB		3,31985
attachement	SBC	6	3,60775	continuel	ADJ	3	3,31985
opération	SBC	93	3,60219	fautif	ADJ	3	3,31985
suffixe	SBC	9	3,5714	immanquablement	ADV	3	3,31985
monochrome	ADJ	9	3,5714	emboîtement	SBC	3	3,31985
						5	
terminer	ADJPAR	15	3,56347	entreposage	SBC	3	3,31985
insérer	VB	29	3,56007	chiffrage	SBC	3	3,31985
chargement	SBC	23	3,55918	biunivoque	ADJ	3	3,31985
le	DT	7759	3,55457	bref	ADV	3	3,31985
construction	SBC	14	3,53539	nécessairement	ADV	13	3,3015
tableau	SBC	40	3,53456	dépendre	VB	27	3,28845
populaire	ADJ	10	3,52102	contrainte	SBC	18	3,28342
séquentiel	ADJ	10	3,52102	associer	ADJPAR	25	3,26922
•	SBC	24				38	
comportement			3,49294	réel	ADJ		3,26516
ajuster	ADJPAR	4	3,46951	fonctionnel	ADJ	14	3,24276
susmentionné	ADJ	4	3,46951	affiche	SBC	13	3,20042
interrelié	ADJ	4	3,46951	élémentaire	ADJ	13	3,20042
	ADJ	4			SBC	13	
déterministe			3,46951	comparaison			3,20042
décodeur	SBC	4	3,46951	résultat	SBC	59	3,1883
introuvable	ADJ	4	3,46951	équation	SBC	7	3,18129
répliquer	VB	4	3,46951	extrêmement	ADV	16	3,17153
transformer	VB	15	3,46922	mathématique	ADJ	12	3,1624
		. •	3,	1	0		J, . J

modification	SBC	38	3,15177	logiquement	ADV	8	3,11975
documentation	SBC	24	3,15117	stabilité	SBC	8	3,11975
produire	VB	30	3,14243	détecter	ADJPAR	9	3,10191
élément	SBC	60	3,13461				

## SC4 – Matériel

Forme	PdeD	Fréq. abs.	Valeur- test	Forme	PdeD	Fréq. abs.	Valeur- test
imprimante	SBC	200	26,9233	mètre	SBC	13	9,84254
papier	SBC	104	25,9476	dur	ADJ	79	9,8089
epson	SBP	44	22,7884	port	SBC	81	9,69149
laser	SBC	62	20,5546	adhésif	ADJ	10	9,59645
imprimant	ADJ	52	19,5618	pouce	SBC	33	9,46937
feuille	SBC	67	19,554	caroll	SBP	8	9,41489
étiquette	SBC	43	18,2538	encreur	SBC	8	9,41489
aiguille	SBC	41	17,9985	pince	SBC	8	9,41489
imprimer	VB	73	17,3942	légende	SBC	11	9,36974
ruban	SBC	37	16,541	hélas	INJ	12	9,27125
continu	ADJ	37	16,1903	mère	SBC	26	9,19735
trait	SBC	35	16,0587	liaison	SBC	31	9,13982
broche	SBC	24	15,9608	haut	SBC	33	9,06118
matriciel	ADJ	29	15,8443	bord	SBC	23	8,99174
micro	SBC	40	15,7503	court	ADJ	18	8,94371
impression	SBC SBC	84 27	15,3475	vertical	ADJ VB	30 8	8,82588
borne cartouche	SBC	29	15,3454 14,9301	souder manuel	SBC	o 21	8,7894 8,76332
consigne	SBC	16	13,7934	calibrage	SBC	7	8,71636
cordon	SBC	17	13,3564	nationalité	SBC	7	8,71636
réglette	SBC	14	12,8385	façade	SBC	7	8,71636
amstrad	SBP	14	12,8385	switche	SBC	7	8,71636
tabulation	SBC	26	12,6258	soudure	SBC	7	8,71636
mais	COO	200	12,5205	ligne	SBC	158	8,69418
maître	SBC	23	12,3667	largeur	SBC	23	8,68054
code	SBC	121	12,1549	boîtier	SBC	23	8,68054
câble	SBC	41	11,8907	cristal	SBC	12	8,6787
bac	SBC	13	11,8152	en	PREP	706	8,6072
tableau	SBC	52	11,7209	cylindre	SBC	19	8,54278
ce	PRO	336	11,562	italique	ADJ	10	8,38105
imprimer	ADJPAR	32	11,5597	plat	ADJ	10	8,38105
nappe	SBC	16	11,429	saut	SBC	13	8,26717
neuf	ADJ	18	11,2819	encadrement	SBC	7	8,06108
traçant	ADJ	12	11,2716	centrage	SBC	7	8,06108
caoutchouc	SBC	12	11,2716	panne	SBC	17	8,05672
condenser	ADJPAR	12	11,2716	dater	VB	16	8,01841
tête	SBC SBC	39 15	10,9474	petit friction	ADJ SBC	58 6	7,96403
esclave pica	SBC	10	10,9193 10,6775	anarchie	SBC	6 6	7,95812 7,95812
maxi	ADJ	10	10,6775	tracé	SBC	6	7,95812
у	PRO	134	10,6132	élite	SBC	6	7,95812
gras	ADJ	15	10,6126	moleter	ADJPAR	6	7,95812
liquide	ADJ	12	10,3584	très	ADV	120	7,93596
socle	SBC	12	10,3584	fil	SBC	15	7,84015
côté	SBC	35	10,2914	astuce	SBC	11	7,82247
zéro	SBC	18	10,1649	sérier	VB	20	7,81578
prise	SBC	36	10,0815	barrette	SBC	20	7,71527
codepson	SBP	9	10,0658	primaire	ADJ	15	7,68697
brochage	SBC	9	10,0658	codage	SBC	20	7,61699
picot	SBC	9	10,0658	ibm	SBP	38	7,60885
cosse	SBC	9	10,0658	graveur	SBC	23	7,60638

aller	VB	87	7,57096	arrière	SBC	10	5,49274
annuler	ADJPAR	7	7,51301	alors	ADV	72	5,48711
brancher	VB	16	7,46585	rouleau	SBC	6	5,48377
encre	SBC	17	7,31022	huit	CAR	13	5,48363
entraîner	ADJPAR	6	7,26908	tension	SBC	13	5,48363
photocopieuse	SBC	6			SBC	8	
			7,26908	décalage			5,45018
démonstration	SBC	6	7,26908	imprimable	ADJ	4	5,40958
loger	VB	6	7,26908	grec	ADJ	4	5,40958
césure	SBC	6	7,26908	secrétariat	SBC	4	5,40958
bas	SBC	23	7,22281	conducteur	SBC	4	5,40958
proportionnel	ADJ	10	7,20712	alcool	SBC	4	5,40958
sélénium	SBC	5	7,12203	perforation	SBC	4	5,40958
		5		1 .		4	
mâle 	ADJ		7,12203	acheteur	SBC		5,40958
liasse	SBC	5	7,12203	voiler	VB	4	5,40958
remarque	SBC	9	7,09234	encombrant	ADJ	4	5,40958
page	SBC	89	7,06142	scalaire	ADJ	4	5,40958
					-	9	
commutateur	SBC	10	6,96624	hauteur	SBC		5,36443
mécanique	ADJ	10	6,96624	graphique	ADJ	66	5,36164
centrer	ADJPAR	10	6,96624	coller	ADJPAR	5	5,35347
espacement	SBC	12	6,94004	draft	SBC	5	5,35347
•			,			5	
passage	SBC	20	6,90159	canon	SBP		5,35347
programmer	VB	16	6,86712	horizontal	ADJ	19	5,33592
abord	SBC	29	6,84166	marque	SBC	16	5,32991
normal	ADJ	24	6,81289	aussi	ADV	87	5,32548
_			,				
caractère	SBC	134	6,80034	final	ADJ	15	5,30174
macintosh	SBP	17	6,77752	second	ADJ	21	5,29114
alimentation	SBC	14	6,77604	vider	VB	9	5,20382
édition	SBC	29	6,74123	large	ADJ	13	5,19146
usure	SBC	6	6,70727	alimenter	VB	7	5,17776
latéral	ADJ	6	6,70727	bloquer	ADJPAR	6	5,17492
puis	ADV	51	6,51285	métallique	ADJ	6	5,17492
chariot	SBC	8	6,40368	centimètre	SBC	6	5,17492
			,				
découpe	SBC	5	6,39488	tour	SBC	16	5,11351
secondaire	ADJ	20	6,36056	temporaire	ADJ	13	5,09941
word	SBP	11	6,21639	surtout	ADV	24	5,08542
serrage	SBC	4	6,17867	imbiber	ADJPAR	3	5,07454
. •		4					
grippage	SBC		6,17867	bourre	SBC	3	5,07454
rabat	SBC	4	6,17867	switcher	VB	3	5,07454
piège	SBC	4	6,17867	pâle	ADJ	3	5,07454
teinter	ADJPAR	4	6,17867	retendre	VB	3	5,07454
gaine	SBC	4	6,17867	confectionner	VB	3	5,07454
•							,
traduisible	ADJ	4	6,17867	galet	SBC	3	5,07454
étain	SBC	4	6,17867	croisée	SBC	3	5,07454
pliure	SBC	4	6,17867	garnir	VB	3	5,07454
coton	SBC	4	6,17867	pliage	SBC	3	5,07454
modèle	SBC	61	6,1547	croisillon	SBC	3	5,07454
graphisme	SBC	9	6,11001	authentique	ADJ	3	5,07454
optionnel	ADJ	7	5,96149	témoin	SBC	3	5,07454
longueur	SBC	21	5,85729	luxe	SBC	3	5,07454
_			,				
sous-programme	SBC	6	5,83401	traction	SBC	3	5,07454
mention	SBC	6	5,83401	polygone	SBC	3	5,07454
entraînement	SBC	6	5,83401	tutoriel	ADJ	3	5,07454
chose	SBC	27	5,83206	veiner	ADJPAR	3	5,07454
				cabochon			
électron	SBC	10	5,80252		SBC	3	5,07454
concevoir	VB	10	5,80252	plasma	SBC	3	5,07454
conclusion	SBC	9	5,71432	juxtaposition	SBC	3	5,07454
deux	CAR	121	5,6843	auto-adhésif	ADJ	3	5,07454
genre	SBC	18	5,61185	dépannage	SBC	3	5,07454
colonne	SBC	28	5,55976	rail	SBC	3	5,07454
fabricant	SBC	16	5,55971	tracteur	SBC	3	5,07454
auteur	SBC	23	5,52604	bristol	SBC	3	5,07454
essai	SBC	10	5,49274	lumière	SBC	8	5,07018

				1			
un	DT	1355	5,05	figurer	VB	10	4,29131
masse	SBC	11	5,02137	ou	COO	309	4,28655
	SBC	13	5,0098		ADV	19	
tube			,	là			4,28334
avoir	VBA	391	5,00472	hé	INJ	3	4,2641
dire	VB	49	4,97901	colorer	ADJPAR	3	4,2641
	SBC	5			SBC	3	4,2641
manoeuvre			4,96025	remontage			,
deviner	VB	5	4,96025	résidant	ADJ	3	4,2641
interligne	SBC	5	4,96025	voisin	ADJ	3	4,2641
_	SBC	5			VB	3	
fer			4,96025	étiqueter			4,2641
parité	SBC	7	4,95983	aiguiller	VB	3	4,2641
départ	SBC	14	4,95367	centronics	SBP	3	4,2641
	SBC					3	,
branchement		9	4,90686	amateur	ADJ		4,2641
enfichable	ADJ	6	4,89948	barrer	ADJPAR	3	4,2641
couper	ADJPAR	6	4,89948	amplement	ADV	3	4,2641
•	ADJPAR	6	,		ADJ	3	4,2641
disposer			4,89948	simpliste			
quand	SUB	24	4,87334	quatre	CAR	21	4,26373
bout	SBC	13	4,83745	gamme	SBC	14	4,24154
	ADJPAR	4		derrière	PREP	6	
traire			4,83164	1.			4,22047
bureautique	SBC	4	4,83164	baud	SBC	8	4,16993
buse	SBC	4	4,83164	cause	SBC	13	4,15649
atari	SBP	4	4,83164	à	PREP	831	
			,				4,10381
défragmenter	VB	4	4,83164	savoir	VB	41	4,10357
homologue	ADJ	4	4,83164	perforer	ADJPAR	7	4,08155
volt	SBC	4	4,83164	atteindre	ADJPAR	7	4,08155
cela	PRO	62	4,72629	coûter	VB	12	4,07345
éteindre	VB	10	4,71587	signe	SBC	15	4,03593
maison	SBC	6	4,65146	découper	ADJPAR	6	4,03112
				•			
variante	SBC	9	4,63765	gauche	SBC	19	4,01549
octet	SBC	26	4,6279	intercaler	VB	4	3,99966
rangée	SBC	5	4,62312	réfléchir	VB	4	3,99966
surprise	SBC	5	4,62312	auto	SBC	4	3,99966
dix	CAR	11	4,60701	relier	ADJPAR	15	3,98041
fiche	SBC	10	4,60387	calcul	SBC	22	3,97722
	SBP	15	4,57973	taille	SBC	45	
apple <sub>.</sub>							3,96877
seconde	SBC	23	4,57909	sept	CAR	7	3,93738
carte	SBC	95	4,54475	cathodique	ADJ	7	3,93738
titre	SBC	22	4,52712	annuler	VB	7	3,93738
environ	ADV	17	4,50029	milieu	SBC	7	3,93738
tout	PRO	36	4,49426	minuscule	ADJ	13	3,89393
double	ADJ	17	4,44541	long	ADJ	22	3,86816
avec	PREP	187	4,44125	dépasser	VB	12	3,86103
écran	SBC	85	4,37912	couper	VB	6	3,85599
prix	SBC	29	4,3789	tracer	VB	6	3,85599
•		4	4,37452	s'	SUB		3,85335
qwerty	SBP					25	
roue	SBC	4	4,37452	opter	VB	5	3,83746
vôtre	PRO	4	4,37452	électronique	SBC	5	3,83746
quinze	CAR	4	4,37452	ligner	VB	5	3,83746
kilomètre	SBC	4	4,37452	poudre	SBC	5	3,83746
voyant	ADJ	4	4,37452	épaisseur	SBC	5	3,83746
stylo	SBC	4	4,37452	arriver	VB	18	3,82007
levier	SBC	4	4,37452	près	ADV	15	3,81895
silence	SBC	4	4,37452	circuit	SBC	15	3,81895
angle	SBC	5	4,32924	même	ADV	35	3,81806
•		5					
tomber	VB	5	4,32924	bref	ADJ	8	3,81291
verre	SBC	5	4,32924	entièrement	ADV	7	3,80117
rouge	ADJ	5	4,32924	suppression	SBC	10	3,75179
•							
paire	SBC	5	4,32924	longtemps	ADV	10	3,75179
processeur	SBC	38	4,32897	chapitre	SBC	23	3,73531
équiper	ADJPAR	12	4,30125	inconvénient	SBC	12	3,72703
somme	SBC	12	4,30125	séparer	ADJPAR	17	3,70675
pas	SBC	8	4,30058	fréquent	ADJ	8	3,704
moins	ADV	56	4,29198	accidentellement	ADV	3	3,69714
			-	•			-

bizarre         ADJ         3         3,69714 pourtant presque         ADV         8         3,3093 presque           phosphore         SBC         3         3,69714 presque         ADV         10         3,2973 presque           fragile         ADJ         3         3,69714 presque         ADJPAR         3         3,2684 presque           provisoirement         ADV         3         3,69714 presque         ADJPAR         3         3,2684 presque           provisoirement         ADJ         3         3,69714 presque         ADJPAR         3         3,2684 presque           provisoirement         ADJ         3         3,69714 presque         ADJPAR         3         3,2684 presque           provisoirement         ADJ         3         3,69714 presque         4         ADJPAR         3         3,2684 presque           provisoirement         ADJ         3         3,69714 presque         4         ADJPAR         3         3,2684 presque           provisoirement         ADJ         4         3,68387 presque         4         ADJPAR         3         3,2684 presque           provisoirement         ADJ         4         3,68387 presque         4         ADJPAR         3         3,2684 presque
diagonale SBC 4 3,68387 azerty SBP 3 3,2684
,
,
manque SBC 5 3,62807 dessus SBC 4 3,174
engager VB 5 3,62807 interactivité SBC 4 3,174
compatible ADJ 19 3,60614 peu ADV 27 3,1670
car COO 54 3,59867 servir VB 35 3,1654
cher ADJ 13 3,593 programmeur SBC 15 3,1512
cas SBC 82 3,58005 excellent ADJ 6 3,1389 avantage SBC 20 3,57149 ordinaire ADJ 5 3,1024
avantage         SBC         20         3,57149 ordinaire         ordinaire         ADJ         5         3,1024 rassembler           destiner         ADJPAR         15         3,56547 rassembler         VB         5         3,1024 rassembler
plateau SBC 7 3,54962 enlever VB 5 3,1024
légèrement ADV 7 3,54962 calculer VB 9 3,100
bloquer VB 6 3,54132 prêt ADJ 9 3,100
appletalk SBP 6 3,54132 série SBC 19 3,096
chance SBC 8 3,49909 moyen ADJ 10 3,0955

# SC5 – Programmation et réseaux

Forme	PdeD	Fréq. abs.	Valeur- test	Forme	PdeD	Fréq. abs.	Valeur- test
implémentation	SBC	55	23,5282	coder	VB	22	12,809
algorithme	SBC	61	20,8992	plate-forme	SBC	27	12,7996
synchronisation remplissage	SBC	63	20,6869	classe	SBC	41	12,544
	SBC	38	20,5786	méthodologie	SBC	17	12,3307
contrat réseau	SBC SBC SBC	44 292 28	20,5596 19,6802 19,1221	tramage	SBC SBC VB	13 12 12	12,2057 12,1934
réservoir mobile intranet	SBC SBC SBC	25 24	17,6266 16,8543	recouper implémenter poste	VB VB SBC	16 41	12,1934 12,156 12,0738
architecture	SBC	72	16,0315	oracle	SBP	15	11,9929
migration	SBC	25	15,6191	problématique	SBC	18	11,68
topologie	SBC	23	15,086	calcul	SBC	43	11,3275
problème	SBC	110	14,5802	aspect	SBC	37	11,042
serveur	SBC	201	14,4628	caper	VB	9	10,3957
java	SBP	34	14,4592	palmos	SBP		10,3957
recoupement	SBC	17	14,2377	approche	SBC	19	10,08
mobile	ADJ	26	14,0569	cliente	SBC	20	10,0454
langage	SBC	104	13,4292	énoncé	SBC	10	9,91983
entreprise	SBC	66	13,204	invariant	SBC	9	9,78193
client	SBC	96	12,9011	solution	SBC	59	9,7819

roland lan grand niveau algorithmique problématique assertion patrick partenaire directeur	SBP SBP ADJ SBC ADJ ADJ SBC SBP SBC SBC SBC	8 11 100 91 9 8 10 7 11	9,72381 9,69266 9,39921 9,36042 9,25015 9,08305 9,04825 9,00276 8,9815 8,9815	manière flashage administration communication donnée province hétérogénéité anneau transiter public	SBC SBC SBC SBC SBC SBC SBC SBC VB ADJ	52 8 20 48 99 5 5 5 9	6,15813 6,1162 6,11559 6,03347 6,02693 6,02595 6,02595 6,02595 5,94534 5,81551
ressource	SBC	65	8,93349	coût	SBC	32	5,73513
base grille	SBC SBC	103 12	8,84476 8,71867	généralement offrir	ADV VB	44 30	5,71102 5,70844
volume	SBC	28	8,49503	chef	SBC	6	5,68813
garantie méthode	SBC SBC	12 42	8,46408 8,43816	rédiger sécurité	VB SBC	7 35	5,62555 5,60355
centraliser	ADJPAR	11	8,37914	contresens	SBC	4	5,59602
calculateur	SBC	9	8,36805	michel	SBP	4	5,59602
affectation application	SBC SBC	12 109	8,22509 8,15354	dilemme extranet	SBC SBC	4 4	5,59602 5,59602
fraction	SBC	8	8,06281	journaliser	ADJPAR	4	5,59602
modéliser	VB	8	8,06281	politique	ADJ	9	5,58856
bancaire	ADJ	10	8,0264	diffusion	SBC	21	5,58712
terminal	ADJ DT	41 344	7,95326 7,83845	exécution participant	SBC ADJ	44 5	5,56463 5,54698
au donner	ADJPAR	115	7,73596	protocole	SBC	68	5,52362
ibm	SBP	37	7,71385	documentation	SBC	20	5,51415
département	SBC	9	7,6591	ordinateur	SBC	133	5,47596
proposition informatique	SBC ADJ	11 45	7,6245 7,53964	réutilisation héritage	SBC SBC	7 6	5,38343 5,37267
eugène	SBP	5	7,35304	accès	SBC	81	5,30568
postcondition	SBC	5	7,35712	dutilleul	SBP	3	5,24408
modularisation	SBC	5	7,35712	préprocesseur	SBC	3	5,24408
paramétrique eiffel	ADJ SBP	5 5	7,35712 7,35712	supercalculateur méthodologique	SBC ADJ	3 3	5,24408 5,24408
contrainte	SBC	19	7,33712	modulariser	VB	3	5,24408
migrer	VB	8	7,27585	attachmate	SBP	3	5,24408
central	ADJ	37	7,26818	recomposition	SBC	3	5,24408
complexité	SBC SBC	18 10	7,02153 7	frontal goulet	SBC SBC	3 3	5,24408 5,24408
responsable bénéfice	SBC	10	7	invariant	ADJ	3	5,24408
égal	SBC	7	6,87547	concerter	ADJPAR	3	5,24408
cisco	SBP	8	6,64232	interfacer	SBC	3	5,24408
java besoin	SBC SBC	5 54	6,61244 6,58103	euro outlook	SBC SBP	3 3	5,24408 5,24408
effectif	ADJ	12	6,56918	narval	SBP	3	5,24408
maillon	SBC	6	6,45775	vulnérable	ADJ	3	5,24408
résoudre	VB	20	6,45231	technicité	SBC	3	5,24408
exception multi-terminal	SBC ADJ	14 4	6,44212 6,38351	métaprogrammation passerelle	SBC SBC	3 22	5,24408
incitatif	ADJ	4	6,38351	client-serveur	ADJ	7	5,21871 5,16111
vidage	SBC	4	6,38351	spécialiser	ADJPAR	19	5,15992
audrey	SBP	4	6,38351	logistique	SBC	5	5,1451
dénominateur typage	SBC SBC	4 4	6,38351 6,38351	spécification final	SBC ADJ	16 14	5,10437 5,04676
emmêlement	SBC	4	6,38351	administrer	VB	4	5,00478
symbian	SBP	4	6,38351	parisien	ADJ	4	5,00478
primitif	ADJ	9	6,34908	lier	ADJPAR	16	4,97452
terme un	SBC DT	34 1326	6,3174 6,26683	tâche programmation	SBC SBC	30 34	4,95958 4,95843
proposer	VB	34	6,24582	sécuriser	ADJPAR	7	4,95584
mémoriser	ADJPAR	7	6,18366	robustesse	SBC	7	4,95584

				1			
développement	SBC	46	4,95178	biais	SBC	7	3,84967
entier	ADJ	17	4,93434	être	VB	7	3,84967
action	SBC	29	4,92625	symptôme	SBC	3	3,83492
égal	ADJ	13	4,91291	sémantique	SBC	4	3,83245
local	ADJ	53	4,88512	négliger	VB	4	3,83245
retenir	ADJPAR	6	4,83839	cœur	NN	4	
	-						3,83245
flasher	VB	5	4,80077	difficile	ADJ	17	3,82531
statique	ADJ	8	4,78472	reporter	VB	5	3,78593
authentification	SBC	13	4,75428	ensemble	SBC	40	3,73497
routeur	SBC	9	4,7204	abstraction	SBC	6	3,7074
mise	SBC	38			ADJ	8	
			4,71863	propriétaire			3,68259
reposer	VB	10	4,71152	évoluer	VB	10	3,64048
architecte	SBC	8	4,63727	constructeur	SBC	11	3,62521
chez	PREP	12	4,60896	représenter	VB	28	3,61117
notre	DT	31	4,60887	celui-ci	PRO	44	3,60774
équation	SBC	6	4,60877	agir	VB	37	3,59229
agenda	SBC	7	4,58796	abandonner	VB	5	3,59199
et	COO	778	4,57557	messagerie	SBC	11	3,55964
côté	SBC	19	4,55432	potentiellement	ADV	4	3,55548
aussi	ADV	78	4,54944	implémenter	ADJPAR	4	3,55548
		-					
dépôt	SBC	4	4,53755	lequel	REL	56	3,50537
individuel	ADJ	10	4,50348	signification	SBC	10	3,49511
assistant	SBC	5	4,50079	appliquer	ADJPAR	7	3,49363
orienter	ADJPAR	9	4,48073	précis	ADJ	14	3,47325
				fournir		31	
positif	ADJ	7	4,42198		VB		3,44423
nature	SBC	19	4,41638	concept	SBC	13	3,43277
opératoire	ADJ	3	4,41454	désigner	VB	11	3,43267
attaquer	VB	3	4,41454	synchroniser	VB	5	3,41444
facturer	ADJPAR	3	4,41454	facture	SBC	5	3,41444
					VB		
coaxial	ADJ	3	4,41454	bénéficier		8	3,39885
dominique	SBP	3	4,41454	siège	SBC	3	3,39712
frontal	ADJ	3	4,41454	réplication	SBC	3	3,39712
journalisation	SBC	3	4,41454	sécurisation	SBC	3	3,39712
affirmation	SBC	3	4,41454	implémenté	ADJ	3	3,39712
énoncer	ADJPAR	3	4,41454	coexister	VB	3	3,39712
moyen	SBC	24	4,41089	codage	SBC	11	3,37116
alphabet	SBC	8	4,36448	code	SBC	61	3,32378
étape	SBC	22	4,29897	studio	SBP	4	3,31345
consister	VB	24	4,27202	routeur	ADJ	4	3,31345
naturellement	ADV	7	4,26618	immense	ADJ	4	3,31345
	SBC	8		étudier	VB	6	
productivité			4,23784				3,29833
collection	SBC	6	4,20601	considérer	VB	11	3,25184
modélisation	SBC	6	4,20601	dater	VB	8	3,22581
notification	SBC	6	4,20601	à	PREP	765	3,19022
concret	ADJ	6	4,20601	atteindre	VB	13	3,18406
marketing	SBC	4	4,15471	susceptible	ADJ	7	3,18346
~				· •			
du	DT	1119	4,12403	pc	SBC	34	3,17882
transition	SBC	7	4,11945	argument	SBC	6	3,17705
stable	ADJ	8	4,00142	élémentaire	ADJ	8	3,14352
rendez-vous	SBC	5	3,99927	tandis	PUL	8	3,14352
étoile	SBC	5	3,99927	comparaison	SBC	8	3,14352
				coder	ADJPAR		
relier	VB	13	3,99623			9	3,14162
autant	ADV	20	3,99083	alphanumérique	ADJ	5	3,09946
partager	ADJPAR	19	3,98867	métier	SBC	5	3,09946
performance	SBC	25	3,96484	perspective	SBC	5	3,09946
centre	SBC	14	3,93963	câblage	SBC	4	3,09909
				_			
présentation	SBC	16	3,92817	persistance	SBC	4	3,09909
objectif	SBC	11	3,90292	décoder	VB	4	3,09909
ensemble	ADV	9	3,86464	découler	VB	4	3,09909
débogage	SBC	6	3,86197	introduire	VB	10	3,09896
décimal	ADJ	6	3,86197				
			•	•			

SC6 – Systèmes d'exploitation

Forme	PdeD	Fréq. abs.	Valeur- test	Forme	PdeD	Fréq. abs.	Valeur- test
option	SBC	1258	30,7561	numéro	SBC	191	8,38639
fichier	SBC	2388	24,924	manière	SBC	218	8,08162
recommander	ADJPAR	496	22,402	règle	SBC	166	8,05274
noyau	SBC	438	20,8244	périphérique	SBC	157	8,04061
linux	NNP	422	20,5008	unix	SBP	162	8,01503
configuration	SBC	660	20,0158	menu	SBC	277	7,98751
utiliser	VB	1262	19,8464	virtuel	ADJ	162	7,95917
activer	VB	368	17,2156	souris	SBC	196	7,84445
ce	DT	3158	16,3312	ligne	SBC	589	7,81007
police	SBC SBC	328 255	15,3814	paquetage droit	SBC SBC	58 126	7,75461
script réponse	SBC	393	15,2034 15,0156	échappement	SBC	68	7,64993 7,55391
connexion	SBC	371	14,5141	console	SBC	69	7,50713
utiliser	ADJPAR	644	14,495	démarrage	SBC	131	7,34986
paquet	SBC	279	14,3034	capable	ADJ	133	7,338
n'	ADV	1169	13,7039	monter	VB	58	7,20001
suivant	ADJ	512	13,5514	destination	SBC	89	7,18662
répertoire	SBC	618	13,1612	lancer	VB	150	7,16162
permettre	VB	1349	12,6671	archiver	VB	59	7,14391
périphérique	ADJ	285	12,6303	descripteur	SBC	54	7,03087
fonctionnalité	SBC	259	12,59	gestion	SBC	291	7,02166
support	SBC	365	12,3884	pouvoir	VB	1705	7,00133
installation	SBC	246	12,2909	mot-clé	SBC	49	6,9307
paramètre	SBC	323	12,1513	variable	ADJ	107	6,85578
être	VBE	7260	11,9548	fournir	ADJPAR	132	6,85374
utilisateur distribution	SBC SBC	649 194	11,9475	indiquer carte	ADJPAR SBC	91 440	6,81189
module	SBC	309	11,8348 11,5994	interface	SBC	234	6,80924 6,78777
partition	SBC	246	11,2042	spécifier	VB	138	6,78401
démon	SBC	126	10,9586	librairie	SBC	56	6,78282
variable	SBC	144	10,8182	devoir	VB	727	6,7614
machine	SBC	519	10,6313	plupart	SBC	178	6,73718
nom	SBC	785	10,5551	lancer	ADJPAR	73	6,72127
donc	ADV	663	10,5306	lilo	NNP	42	6,68771
ceci	PRO	242	10,4529	partage	SBC	95	6,67888
commande	SBC	965	10,2866	défaut	SBC	214	6,63868
spécial	ADJ	216	10,2434	abréviation	SBC	64	6,56977
simplement	ADV	189	10,1896	général	ADJ	200	6,56523
valeur	SBC	325	10,1163	passe	SBC	109	6,56198
chaîne lui	SBC PRO	199 7434	10,0222 9,98615	décrire	ADJPAR SBC	105 39	6,44691 6,43495
syntaxe	SBC	162	9,507	pipe utilisable	ADJ	60	6,36936
ne	ADV	1178	9,4344	spécifier	ADJPAR		6,34419
windows	SBP	365	9,42361	session	SBC	73	6,14618
pas	ADV	1309	9,40376	disposer	VB	209	6,13098
installer	VB	183	9,39041	substitution	SBC	42	6,03411
système	SBC	1274	9,22234	indiquer	VB	229	5,90055
terminal	SBC	121	9,21363	revanche	SBC	79	5,88892
configurer	VB	116	8,93269	processus	SBC	202	5,88414
port	SBC	264	8,91128	classique	ADJ	101	5,84558
encodage	SBC	105	8,90201	note	SBC	76	5,8367
également	ADV	308	8,87082	terminal	ADJ	119	5,7318
Si	SUB	886	8,82368	passerelle	SBC	79	5,72389
entendre section	ADJPAR SBC	125 240	8,80714 8,65391	réaliser identificateur	VB SBC	123 38	5,68618 5,66581
graphique	ADJ	2 <del>4</del> 0 297	8,41487	connecter	VB	36 116	5,6655
a. apquo	, .50	_0,	5,	1 30111100101			5,5555

routage	SBC	55	5,61933	verrou	SBC	20	4,51792
pour	PREP	2624	5,61538	staroffice	SBP	20	4,51792
dossier	SBC	73	5,61089	graveur	SBC	52	4,51515
corel	SBP	30	5,60891	détail	SBC	54	4,50093
choisir montage	VB SBC	190 43	5,55201 5,53965	conversion parfaitement	SBC ADV	45 47	4,48825
deuxième	ADJ	97	5,53905	lors	ADV	179	4,46159 4,45919
translation	SBC	31	5,50994	caractère	SBC	489	4,43706
doute	SBC	75	5,50861	manuellement	ADV	32	4,42602
cependant	ADV	212	5,50378	ci-dessous	ADV	64	4,39115
supprimer	VB	101	5,45	serveur	SBC	516	4,37299
compilation	SBC	64	5,44612	etc	ADV	276	4,36605
linux	SBC	32	5,4222	provenir	VB	73	4,34124
afin	PUL ADV	204 74	5,39657	intermédiaire	SBC	71 275	4,33231
normalement alias	ADV	31	5,35001 5,32067	possible bureau	ADJ SBC	68	4,32613 4,32083
source	SBC	184	5,29962	générer	VB	60	4,30471
répondre	VB	129	5,28	pilote	SBC	28	4,28643
plage	SBC	34	5,27413	barre	SBC	75	4,27959
libre	ADJ	102	5,26141	automatiquement	ADV	104	4,27176
le	DT	21137	5,24615	classiquement	ADV	18	4,26656
mineur	ADJ	32	5,24209	onglet	SBC	33	4,24499
ici	ADV	140	5,20378	suffire	VB	149	4,23796
définition	SBC	103	5,20248	hors	PREP	49	4,23772
que installer	SUB\$ ADJPAR	1293 86	5,18895 5,16574	seul chemin	ADJ SBC	253 93	4,22478 4,22405
désirer	VB	147	5,1618	liste	SBC	220	4,22024
amorçage	SBC	40	5,11457	filtrage	SBC	30	4,1998
local	ADJ	217	5,11324	démarrer	VB	72	4,19009
explorateur	SBP	25	5,0926	clavier	SBC	272	4,14847
flux	SBC	68	5,05016	protocole	SBC	256	4,13943
lien	SBC	112	5,01046	icône	SBC	78	4,09287
postscript	SBP	43	5,00863	groupe	SBC	149	4,06743
gérer	ADJPAR	68 71	4,96553	résolution	SBC	73 127	4,05367
utilisateur définir	ADJ VB	71 173	4,95487 4,91314	table balayage	SBC SBC	137 30	4,05085 4,04889
respectivement	ADV	55	4,88252	horaire	ADJ	23	4,03062
compiler	VB	39	4,87804	restriction	SBC	31	4,01881
nommer	ADJPAR	58	4,85146	gérer	VB	148	3,97053
aide	SBC	226	4,78815	complètement	ADV	45	3,95385
noter	VB	130	4,77723	évidemment	ADV	48	3,9463
guillemet	SBC	38	4,7768	cd-rom	SBC	68	3,94423
filtre	SBC	29	4,74425	cd	SBC	76	3,94067
standard	ADJ SBC	133	4,7063	description redémarrage	SBC	80 22	3,91278
redirection conséquent	ADJ	27 54	4,70207 4,69505	environnement	SBC SBC	157	3,90115 3,87972
plan	SBC	55	4,68604	garde	SBC	26	3,87961
quasiment	ADV	37	4,67395	secret	SBC	15	3,85919
falloir	VB	422	4,66739	activer	ADJPAR	43	3,85809
mode	SBC	325	4,64174	complémentaire	ADJ	32	3,85582
licence	SBC	38	4,64152	votre	DT	584	3,83594
courant	ADJ	119	4,62831	compresser	ADJPAR	24	3,79524
mon	DT	31	4,62766	moniteur	SBC	120	3,78995
définir	ADJPAR ADJPAR	112 109	4,6206 4,5953	contrôleur correctement	SBC ADV	56 74	3,78671 3,78638
connecter contenir	VB	323	4,5935	réaliser	ADJPAR	66	3,78111
chat	SBC	32	4,57912	ajouter	VB	126	3,77479
charge	SBC	91	4,57893	modifier	VB	184	3,73991
générer	ADJPAR	36	4,56942	gravage	SBC	16	3,73364
unicode	SBP	33	4,53595	uniforme	ADJ	16	3,73364
sous-réseau	SBC	22	4,52561	directement	ADV	142	3,72325
favori	SBC	20	4,51792	bogue	SBC	14	3,71358

pseudo propriétaire nommer raccourci générique impérativement	PFX SBC VB SBC ADJ ADV	14 27 35 23 40 20	3,71358 3,68839 3,68492 3,66483 3,65595 3,63112	exporter ethernet accéder combinaison plan réseau	ADJPAR SBP VB SBC ADJ SBC	12 28 146 50 24 629	3,40396 3,37889 3,32578 3,32062 3,30318 3,29393
placer	ADJPAR	72	3,62992	majeur	ADJ	46	3,28479
bien	ADV	375	3,61369	ample	ADJ	13	3,27109
adresse	SBC	314	3,61364	enfin	ADV	119	3,26165
que	SUB	939	3,61157	host	SBC	16	3,25454
parallèle	ADJ	63	3,60551	rafraîchissement	SBC	27	3,25353
rediriger	VB	15	3,58564	simplifier	ADJPAR	20	3,25176
contextuel	ADJ	21	3,57646	gestionnaire	SBC	77	3,24527
motif	SBC	34	3,57286	inode	SBC	11	3,23824
enregistrer	VB	59	3,56739	framer	VB	11	3,23824
fixer	VB	43	3,55229	dialogue	SBC	38	3,23198
où	ADV	63	3,53149	horizontal	ADJ	49	3,22558
inversement	ADV	22	3,53096	constater	VB	36	3,22487
nécessaire	ADJ	205	3,50504	exécutable	ADJ	32	3,22237
redémarrer	VB	29	3,50167	réellement	ADV	34	3,22136
anglais	SBC	98	3,49174	prendre	VB	164	3,17008
heureusement	ADV	31	3,47588	sous	PREP	314	3,1625
déterminer	VB	84	3,46532	puisqu'	SUB	57	3,16217
profondeur	SBC	24	3,46117	question	SBC	145	3,15086
effectuer	VB	174	3,46039	appartenir	VB	40	3,143
créer	VB	194	3,45627	baser	VB	25	3,1365
symbolique	ADJ	35	3,44879	contenant	ADJ	33	3,10635
code	SBC	276	3,44246	administrateur	SBC	30	3,10472
administrateur	ADJ	14	3,4317	fixer	ADJPAR	32	3,10456
relativement	ADV	52	3,42336	proxy	SBC	12	3,1029
dans	PREP	2048	3,42265	démonter	VB	15	3,09585
antislash	SBC	12	3,40396	veille	SBC	15	3,09585
coolmon	SBP	12	3,40396	série	SBC	71	3,09161

### ANNEXE 5 – Vocabulaire répertorié dans Ginguay (lettres A, C et P)

(les mots en caractères gras représentent les S+)

#### Méthode 1

À abaissé abaissement abaisser abandon abandonner abaque abîmé abonné aboutir abrégé abrégé abréger abriter abscisse absence absolu abstrait accédant accéder accélérateur accélération accéléré accélérer accentué acceptation accepter accepteur accès accessibilité accessible accessoire accolade accordéon accro accrochage accroissement accueil accueillir accumulateur accumulation accumuler accusé accuser acheminement acheminer acheteur acoustique acoustique acquérir acquisition acquittement Actes

acteur actif actionnement actionner actionneur activer activité actualisable actualisation actualiser acyclique adaptabilité adaptateur adaptatif adaptation addition additionnel additionner additionneur additionneursoustracteur adjacent adjonction administrateur administratif administration adopté adopter adressable adressage adresse adresser aération aérer aérographe aéroporté affaiblissement affectable affectation affecter affichable affichage afficher afficheur affolement agence agencement agenda agent agrandir agrandissement agréer

agrégat

agrément agricole aide aide-comptable aide-opérateur aiguillage aiguille aiguiller aimant aimantation aimanter ajournement ajout ajouter ajustage ajustement aiuster alarme album aléatoire aléatoirement algèbre algébrique algébriquement algol algorithme algorithmique aliassage alignement alianer alimentation alimenter aller aller-retour allocation allophone allouer allumage allumer alphabet alphabétique alphagéométrique alphagraphique alphamosaïque alphanumérique alphaphotographiqu altération altérer alternance alternat alternatif

alterner

alvéole amateur ambiance ambiant améliorer aménagement amener amont amorçable amorçage amorce amorcer amortisseur amovibilité amovible amplificateur amplification amplitude analogique analysable analyse analyser analyseur analyste analysteprogrammeur anamorphose ancien ancrage anglais angle angulaire animatique animer anisynchrone anneau annexe annexer annonce annoncer annuaire annulable annulaire annulation annuler anomalie anormal antémémoire antenne antibourrage anticipation anticipé

anticrénelage

antiéblouissant antimaculage antiopisation antiopiser antiparasites antirebond antiréfléchissant antireflet antislash antistatique antiviral antivirus apostrophe appareil apparition appartenance appauvrissement appel appelable appelant appelant appeler appeleur appellation applicabilité applicatif application appliquer appoint apposer apprenant apprentissage approche approvisionnement approvisionner appui appuyer après-vente aptitude arbitrage arborescence arborescent arbre arcade arc-en-ciel architecture architecturer archivabilité archivable archivage archiver

archives

archiviste

ardoise arête argument argumentaire arithmétique arithmétique armement armer armoire arrêté arrêter arrière arriéré arrière-guichet arrière-plan arrivée arriver arrondi arrondi arrondir arrondissage arrondissement arroser arséniure artère article articulation articulé artificiel artisanal arythmique ascendant ascenseur assemblable assemblage assembler assembleur assembleuse asservi asservir asservissement assimilable assimiler assistance assistant assisté associatif association assortiment assortir assurance astérisque asymétrique asynchrone asynchronisme atelier atmosphère atomique attaque attaquer atteint attendus

attente attention attribuer attribut attribution audionumérique audiotypiste audit auditabilité auditer auditeur augmenter auteur authentification authentifier auto-adaptatif auto-adaptation autochargeable autochargement autochargeur autociblé autocode autocodeur autocommutateur autocomplémenteur autoconfiguration autoconnexion autocontrôle autocopiant autocorrecteur autocorrection autocorrélation autodécrémentation autodétecteur autodiagnostic autodocumentant auto-édition auto-étalonnage auto-générateur autogramme auto-incrémentation auto-indexation automate automaticien automaticité automation automatique automatique automatisable automatisation automatisé automatiser automatisme autonome autonomie autopsie autoréparateur autoreproducteur autorestauré autorisation autoriser

autosurveillance

capter

autotranslatable autovérification auxiliaire aval avaler avance avancé avancée avancement avancer avant avant-dernier avant-garde avant-plan avertisseur avis axe axé axer axial cabestan cabinet câblage câble câblé câbler câbleuse cablo-opérateur cabochon cache caché cache-cache cacher cadence cadencement cadencer cadrage cadran cadratin cadre cadré cadrer cadreur cahier caisse caissier caisson calcinateur calcul calculable calculateur calculatrice calculer calculette calendrier camembert canal canaliser cap capacité capot

capteur capture capuchon capuchonnage capuchonner caractère caractéristique carbone carboné caret carnet caroll carotte carrosserie carte carter cartographie carton cartothèque cartouche cas cascade case casier casse cassette cassettothèque catalogue cataloguer cathodique cavalier cellule centrage central central centralisation centraliser centre centré centrer cercle certificateur certification certifier certifieur cerveau césure chaînage chaîne chaîner chambrer champ changement changer changeur chapelet charge chargé chargeable chargement

charger

chargeur

chariot châssis chausse-pied chef chemin cheminement chenille chenillette chèque chéquier chercheur chercheur cheval chevauchement chevaucher chevrons chien chiffrage chiffre chiffre-clé chiffrement chiffrer chiffrier chine chronogramme chronologique chronométrage chronométrer cible cina cinquième circuit circulaire circulant circulation circuler citation citer clair claquage claquer classement classer classeur classique clausier clavier clavier-écran claviste clé clic clicher client clientèle clignotement clignoter climatisation climatiser climatiseur cliquage cliquer cloche

cloisonnement cloisonner clonable clonage clone cloner clore club coaxial codage code codé codec coder codet codeur codéveloppeur codification codifier coefficient coeur coffret cogniticien cognitique cohérence cohérent coin coïncidence collationnement collationner collecte collecter collecteur collectif collection collectique collerette colleuse collision colloque colonnage colonne coloriage colorier combinaison combinatoire combiné combiner comité commande commandé commander commentaire commerce commercial commercialisable commercialisation commercialiser commercique commun communauté communicant

communication communiqué communiquer commutable commutateur commutation commuté commuter compacité compact compactage compacter compagnie comparaison comparateur comparatif comparer compartiment compartimentage compartimenter compatibilité compatible compétitif compétitivité compilable compilateur compilation compiler complément complémentaire complémentation complet complexe composant composé composer composeur composition compresser compresseur compression comprimer comptabilisation comptabiliser comptabilité comptable comptable comptage compte compte-cartes compter compte-rendu compteur compteurdécompteur comptométrice concaténation concaténer concédant concentrateur

concentration

constituant

concentrer

concepteur conception concession concessionnaire concevoir concordance concu concurrence concurrencer concurrent condensateur condensation condensé condenser condition conditionnel conditionnellement conditionnement conditionner conducteur conduit conduite confectionner conférence confetti confidentialité configurabilité configurable configurateur configuration configurer conflit conforme conformité conglomérat conjonction connectabilité connectable connecté connecter connecteur connectique connectivité connexe connexion connexionisme consécutif conseil conseiller conservation conserver consignateur consignation consigne consigner console consommables consommateur consommation consommer constante

constituer constitution constructeur construire consultable consultant consultation consulter contact contenu contexte contigu continu contournement contourner contradiction contradictoire contraste contrat contreargumentaire contre-vérification contre-vérifier contrôlabilité contrôlable contrôle contrôlé contrôler contrôleur convention convergence conversation conversationnel conversion convertir convertisseur convivial convivialement convivialiste convivialité convolution convolutionnel coordonnées copie copier copieur coprocesseur corbeille cordon coresidence corps correcteur correctif correction corrélation correspondance correspondant correspondant correspondre corriger côté

coton-tige

cotraitement couche coupe couper couper/coller coupe-rives couplage couple coupler coupleur coupure courant courant courbe courbure couronne courrier courroie cours course courtage court-circuit court-circuiter courtier cousu coût couteau cracher crayon cravon-lecteur créateur création crédit-bail créditer créé créer crénage créneau crénelage creux crible cristal critère critique crochet crochu croisé croisement croiser croissance croissant cru crucial cryogénie cryonégique cryptage crypter culture cumul cumulande cumulateur cumulatif

cumulation cumulativement cumuler curseur cybernéticien cybernétique cycle cyclique cylindre page paginable pagination paginer pagineur paie paiement pair pair paire palette palier palpeur panachage panier panne panneau panoramique panoramiquer papier papillotement paquet paquetage parabole parabolique paragraphe para-informatique parallèle parallélépipède parallélisation paralléliser parallélisme paramétrable paramétrage paramètre paramétré paramétrer paramétrique parasite parasurtenseur paravent parc parcage parcourir parcours parenthésage parenthèse parité parler parleur parole parquer

part

partage partagé partageable partager partenariat partie partiel partition partitionnement partitionner pas passage passation passe-partout passer passerelle passe-vues passif pastille pâte patienter patte pause pavé paye paysage pellicule pelure pendule pénétration pénétrer pense-bête pénurie percée percer perceur percussion percuter péremption perfectionnement perfectionner perfo perforateur perforation perforatrice perforé perforer perforeuse performance performant perfostyl péri péri-informaticien péri-informatique périminitélie période périphérie périphérique périphérique

péritélévision

permanence

plan

perluète

permanent permutable permutation permuter persistance personnalisable personnalisation personnalisé personnaliser personnel perte pertinence perturbateur perturbation perturber pétale petit phare phase phobie phonique photocomposeuse photodétecteur photodiode photoélectrique photolecture photostyle physionomie physique piano pianoter pico-informatique pico-ordinateur picoseconde picot pictogramme pièce pied piège piégeage piéger pile pilotable pilotage pilote piloter pince pinceau pinceur pipeline piratable piratage pirate pirater piraterie piste pistolet pivot pixel place plage

plancher planétaire planification planifié planifier planning plantage plante planter plaque plateau plate-forme plein plénum pléthore pli pliage plier plieuse pliure plongeur plot plume poche pochette poids poignée poinçon poinconner poinconneuse point pointage pointe pointer pointeur pointillé pointu polarisation polarisé polarité police polonais polyligne polymarque polyphasé polytrait polyvalent pompage ponctuation pondération pondéré pondérer pont-routeur port portabilité portable portage portatif porte porte-balais porte-copie

portefeuille

porte-monnaie porter porteur ajd porteur porteuse portrait pose position positionnable positionnel positionnement positionner possesseur possibilité poste posté postmarquage postmarquer postmarqueuse postprocesseur poubelle pourcentage poursuite poussoir pouvoir pragma pratique préalable préanalyse préannonce précâblage précâblé précâbler précalcul précharger précision précompilateur précompilation précompiler prédéfini prédictif prédiffusé préenregistrer préfacturation préfixé préformatage préformater préformé préimpression préimprimé prélecture prélèvement prélever prémarquage prémarquer prématriçage premier prémonter prendre préparation préperforation préperforer

prépositionnement prépositionner préprocesseur préprogrammé prérégler prérogative préséance présence présentation présenter presse presse-bande presse-cartes presse-papiers presser prestataire prêt prétraitement prétraiter prétrier

preuve préventif préventique primitif primitive principal prioritaire priorité prise prisonnier privatif privatique privatisation privatiser privé privilégié prix proactif problème procédé

procédure processeur processus producteur producticien productif production productique produire produit profane professionnel profil profiler progiciel programmabilité programmable programmateur programmathèque programmation

programmatique programme programmé programmer programmétrie programmeur progresser progressif progression projet projeter propagé propager propre propriétaire prospect protection protégé protéger protocole

prototypage provisoire pseudo-aléatoire pseudo-code pseudo-instruction pseudonyme pseudo-opération public public publication publicité publipostage puce puissance puits pupitrage pupitre pupitrer pupitreur

#### Méthode 2

abaissé abaissement abaisser abandon abandonner abaque abîmé abonné aboutir abrégé. abrégé abréger abriter abscisse absence absolu abstrait accédant accéder accélérateur accélération accéléré accélérer accentué acceptation accepter accepteur accès accessibilité accessible accessoire accolade accordéon accro accrochage accroissement accueil accueillir accumulateur accumulation accumuler accusé accuser acheminement acheminer acheteur acoustique acoustique acquérir acquisition acquittement **Actes** acteur actif actionnement actionner actionneur activer

activité actualisable actualisation actualiser acyclique adaptabilité adaptateur adaptatif adaptation addition additionnel additionner additionneur additionneursoustracteur adjacent adjonction administrateur administratif administration adopté adopter adressable adressage adresse adresser aération aérer aérographe aéroporté affaiblissement affectable affectation affecter affichable affichage afficher afficheur affolement agence agencement agenda agent agrandir agrandissement agréer agrégat agrément agricole aide aide-comptable aide-opérateur aiguillage aiguille aiguiller aimant aimantation

aimanter

ajournement

ajout ajouter ajustage ajustement ajuster alarme album aléatoire aléatoirement algèbre algébrique algébriquement algol algorithme algorithmique aliassage alignement aligner alimentation alimenter aller aller-retour allocation allophone allouer allumage allumer alphabet alphabétique alphagéométrique alphagraphique alphamosaïque alphanumérique alphaphotographiqu altération altérer alternance alternat alternatif alterner alvéole amateur ambiance ambiant améliorer aménagement amener amont amorçable amorçage amorce amorcer amortisseur amovibilité amovible

amplificateur

amplification

amplitude

analogique analysable analyse analyser analyseur analyste analysteprogrammeur anamorphose ancien ancrage anglais angle angulaire animatique animer anisynchrone anneau annexe annexer annonce annoncer annuaire annulable annulaire annulation annuler anomalie anormal antémémoire antenne antibourrage anticipation anticipé anticrénelage antiéblouissant antimaculage antiopisation antiopiser antiparasites antirebond antiréfléchissant antireflet antislash antistatique antiviral antivirus apostrophe appareil apparition appartenance appauvrissement appel appelable appelant appelant appeler appeleur appellation

applicabilité applicatif application appliquer appoint apposer apprenant apprentissage approche approvisionnement approvisionner appui appuyer après-vente aptitude arbitrage arborescence arborescent arbre arcade arc-en-ciel architecture architecturer archivabilité archivable archivage archiver archives archiviste ardoise arête argument argumentaire arithmétique arithmétique armement armer armoire arrêté arrêter arrière arriéré arrière-guichet arrière-plan arrivée arriver arrondi arrondi arrondir arrondissage arrondissement arroser arséniure artère article articulation articulé artificiel artisanal

arythmique ascendant ascenseur assemblable assemblage assembler assembleur assembleuse asservi asservir asservissement assimilable assimiler assistance assistant assisté associatif association assortiment assortir assurance astérisque asymétrique asynchrone asynchronisme atelier atmosphère atomique attaque attaquer atteint attendus attente attention attribuer attribut attribution audionumérique audiotypiste audit auditabilité auditer auditeur augmenter auteur authentification authentifier auto-adaptatif auto-adaptation autochargeable autochargement autochargeur autociblé autocode autocodeur autocommutateur autocomplémenteur autoconfiguration autoconnexion autocontrôle autocopiant

autocorrecteur

cache

autocorrection autocorrélation autodécrémentation autodétecteur autodiagnostic autodocumentant auto-édition auto-étalonnage auto-générateur autogramme auto-incrémentation auto-indexation automate automaticien automaticité automation automatique automatique automatisable automatisation automatisé automatiser automatisme autonome autonomie autopsie autoréparateur autoreproducteur autorestauré autorisation autoriser autosurveillance autotranslatable autovérification auxiliaire aval avaler avance avancé avancée avancement avancer avant avant-dernier avant-garde avant-plan avertisseur avis axe axé axer axial cabestan cabinet câblage câble câblé câbler câbleuse cablo-opérateur cabochon

caché cache-cache cacher cadence cadencement cadencer cadrage cadran cadratin cadre cadré cadrer cadreur cahier caisse caissier caisson calcinateur calcul calculable calculateur calculatrice calculer calculette calendrier camembert canal canaliser cap capacité capot capter capteur capture capuchon capuchonnage capuchonner caractère caractéristique carbone carboné caret carnet caroll carotte carrosserie carte carter cartographie carton cartothèque cartouche cas cascade case casier casse cassette cassettothèque catalogue cataloguer

cathodique

cavalier cellule centrage central central centralisation centraliser centre centré centrer cercle certificateur certification certifier certifieur cerveau césure chaînage chaîne chaîner chambrer champ changement changer changeur chapelet charge chargé chargeable chargement charger chargeur chariot châssis chausse-pied chef chemin cheminement chenille chenillette chèque chéquier chercheur chercheur cheval chevauchement chevaucher chevrons chien chiffrage chiffre chiffre-clé chiffrement chiffrer chiffrier chine chronogramme chronologique chronométrage chronométrer cible cinq

cinquième circuit circulaire circulant circulation circuler citation citer clair claquage claquer classement classer classeur classique clausier clavier clavier-écran claviste clé clic clicher client clientèle clignotement clignoter climatisation climatiser climatiseur cliquage cliquer cloche cloisonnement cloisonner clonable clonage clone cloner clore club coaxial codage code codé codec coder codet codeur codéveloppeur codification codifier coefficient coeur coffret cogniticien cognitique cohérence

cohérent

coïncidence

collationner

collationnement

coin

collecte collecter collecteur collectif collection collectique collerette colleuse collision colloque colonnage colonne coloriage colorier combinaison combinatoire combiné combiner comité commande commandé commander commentaire commerce commercial commercialisable commercialisation commercialiser commercique commun communauté communicant communication communiqué communiquer commutable commutateur commutation commuté commuter compacité compact compactage compacter compagnie comparaison comparateur comparatif comparer compartiment compartimentage compartimenter compatibilité compatible compétitif compétitivité compilable compilateur compilation compiler complément complémentaire

complémentation complet complexe composant composé composer composeur composition compresser compresseur compression comprimer comptabilisation comptabiliser comptabilité comptable comptable comptage compte compte-cartes compter compte-rendu compteur compteurdécompteur comptométrice concaténation concaténer concédant concentrateur concentration concentrer concepteur conception concession concessionnaire concevoir concordance conçu concurrence concurrencer concurrent condensateur condensation condensé condenser condition conditionnel conditionnellement conditionnement conditionner conducteur conduit conduite confectionner conférence confetti confidentialité configurabilité

configurable

configurateur

configuration

configurer conflit conforme conformité conglomérat conjonction connectabilité connectable connecté connecter connecteur connectique connectivité connexe connexion connexionisme consécutif conseil conseiller conservation conserver consignateur consignation consigne consigner console consommables consommateur consommation consommer constante constituant constituer constitution constructeur construire consultable consultant consultation consulter contact contenu contexte contigu continu contournement contourner contradiction contradictoire contraste contrat contreargumentaire contre-vérification contre-vérifier contrôlabilité contrôlable contrôle contrôlé contrôler contrôleur convention

convergence conversation conversationnel conversion convertir convertisseur convivial convivialement convivialiste convivialité convolution convolutionnel coordonnées copie copier copieur coprocesseur corbeille cordon coresidence corps correcteur correctif correction corrélation correspondance correspondant correspondant correspondre corriger côté coton-tige cotraitement couche coupe couper couper/coller coupe-rives couplage couple coupler coupleur coupure courant courant courbe courbure couronne courrier courroie cours course courtage court-circuit court-circuiter courtier cousu coût couteau cracher crayon crayon-lecteur

créateur création crédit-bail créditer créé créer crénage créneau crénelage creux crible cristal critère critique crochet crochu croisé croisement croiser croissance croissant cru crucial cryogénie cryonégique cryptage crypter culture cumul cumulande cumulateur cumulatif cumulation cumulativement cumuler curseur cybernéticien cybernétique cycle cyclique cylindre page paginable pagination paginer pagineur paie paiement pair pair paire palette palier palpeur panachage panier panne panneau panoramique panoramiquer

papier

papillotement

paquet paquetage parabole parabolique paragraphe para-informatique parallèle parallélépipède parallélisation paralléliser parallélisme paramétrable paramétrage paramètre paramétré paramétrer paramétrique parasite parasurtenseur paravent parc parcage parcourir parcours parenthésage parenthèse parité parler parleur parole parquer part partage partagé partageable partager partenariat partie partiel partition partitionnement partitionner pas passage passation passe-partout passer passerelle passe-vues passif pastille pâte patienter patte pause pavé paye paysage pellicule pelure pendule

pénétration

pénétrer pense-bête pénurie percée percer perceur percussion percuter péremption perfectionnement perfectionner perfo perforateur perforation perforatrice perforé perforer perforeuse performance performant perfostyl péri péri-informaticien péri-informatique périminitélie période périphérie périphérique périphérique péritélévision perluète permanence permanent permutable permutation permuter persistance personnalisable personnalisation personnalisé personnaliser personnel perte pertinence perturbateur perturbation perturber pétale petit phare phase phobie phonique photocomposeuse photodétecteur photodiode photoélectrique photolecture photostyle physionomie

physique

piano

pianoter pico-informatique pico-ordinateur picoseconde picot pictogramme pièce pied piège piégeage piéger pile pilotable pilotage pilote piloter pince pinceau pinceur pipeline piratable piratage pirate pirater piraterie piste pistolet pivot pixel place plage plan plancher planétaire planification planifié planifier planning plantage plante planter plaque plateau plate-forme plein plénum pléthore pli pliage plier plieuse pliure plongeur plot plume poche pochette poids poignée poinçon poinçonner poinçonneuse

point pointage pointe pointer pointeur pointillé pointu polarisation polarisé polarité police polonais polyligne polymarque polyphasé polytrait polyvalent pompage ponctuation pondération pondéré pondérer pont-routeur pool port portabilité portable portage portatif porte porte-balais porte-copie portefeuille porte-monnaie porter porteur ajd porteur porteuse portrait pose position positionnable positionnel positionnement positionner possesseur possibilité poste posté postmarquage postmarquer postmarqueuse postprocesseur poubelle pourcentage poursuite poussoir pouvoir pragma pratique préalable

préanalyse

préannonce précâblage précâblé précâbler précalcul précharger précision précompilateur précompilation précompiler prédéfini prédictif prédiffusé préenregistrer préfacturation préfixé préformatage préformater préformé préimpression préimprimé prélecture prélèvement prélever prémarquage prémarquer prématriçage premier prémonter prendre préparation préperforation préperforer prépositionnement prépositionner préprocesseur préprogrammé prérégler prérogative préséance présence présentation présenter presse presse-bande presse-cartes presse-papiers presser prestataire prêt prétraitement prétraiter prétrier preuve préventif préventique primitif primitive principal prioritaire priorité prise

prisonnier
privatif
privatique
privatisation
privatiser
privé
privilégié
prix
proactif
problème
procédé
procédure
processeur
processus

producteur
producticien
production
productique
produire
produit
profane
professionnel
profiler
progiciel
programmabilité
programmable

programmateur programmathèque programmation programme programme programmer programmetrie programmeur progresser progressif progression projet projeter propagé
propager
propre
propriétaire
prospect
protection
protégé
protéger
protocole
prototypage
provisoire
pseudo-aléatoire
pseudo-code
pseudo-instruction

pseudonyme
pseudo-opération
public
public
publication
publicité
publipostage
puce
puissance
puits
pupitrage
pupitre
pupitrer
pupitreur

### ANNEXE 6 - Vocabulaire répertorié dans Collin (lettres A, C et P)

(les mots en caractères gras sont également des S+)

#### Méthode 1

abandonné abandonner abaque aberration abîmé abîmer abonné abrégé abréviation abscisse absolu absorber accédant accéder accélérateur accélérateur accélération accéléré accent accentué acceptable acceptation accepté accepter accès accessible accessoire accidentel accolades accommodation accompagnement accord accordé accordéon accorder accordeur accrochage accroché accrocher accumulateur accumuler accusé accuser achat acheminement acheminer acheter achèvement achever achromatique Ackerman acoustique acquisition

Acrobat acronyme actif actinique actionné actionner activation activé activer activité actualisation actuel acuité acutance adaptable adaptateur adaptatif adaptation adapté adapter additif addition additionnel additionner additionneur additionneursoustracteur adhésif adhésif adiacent administrateur Adobe adoucir adressable adressage adresse adresser aérien aérographe affaiblir affaiblissement affectable affectation affecter affichage affiche afficher affirmatif affirmation agence agenda

agent

agglomérer

agrandir agrandissement agréer agrégat agrégé aide aider aigu aiguillage aiguille aimant aimanté aire ajouté ajouter ajustement aiuster alarme album albumen alcalin alcool aléatoire alerte alerter alerteur algèbre algébrique Algol algorithme algorithmique alias aligné alignement alimentation alimenter aller aller v. allocation allonger allophone alloué allouer allumer alphabet alphabétique alphabétiquement alphagéométrique alphamosaïque alphanumérique alphaphotographiqu altération

altérer alternance alternat alternateur alternatif alternative alterné alterner ambiant ambigu ambiguïté ambisonique âme amélioration amélioré améliorer **Amiga** amont amorçage amorce amorcer amortir amortissement amovible ampère ampli amplificateur amplification amplifié amplifier amplitude analogique analogiquenumérique analogue analyse analyser analyseur analyste anamorphose anastigmate ancêtre ancrage angle angström animé animer anneau annexe annexer annonciateur

annonciateur

annotation

annoter annuaire annulaire annulation annuler anode anomale anomalie anormal anormalement antémémoire antenne antérieur antérieurement anthropomorphique anti anticipation anticipé anticopie anticrénelage anti-écho antiparasite antirebond antireflet antislash antistatique antivirus aperçu apochromatique apogée apostrophe apparaître appareil apparence apparié appel appeler appendice Apple appleshare appletalk application appliquer apprécier apprendre apprentissage approbation approché approuver approximatif approximation

approximativement

appuyer

après après-vente arabe arbitrage arborescence arborescent arbre archétype Archimède architecture archivage archivé archiver archives archiviste ARCNEt arête argument armature armé armer arrangement arrêter arrière arrière-plan arrivée arrondi arrondir arrondissement arséniure article artificiel ascendant ascenseur aspect assemblage assembler assembleur assembleuse assertion asservi assigné assigner assistance assistant assisté assister associatif association associé associer assortir assourdir assourdissement assurance assurer astable astérisque astigmatisme astucieux asymétrique asynchrone

Atari atelier atmosphère atmosphérique atome atomique attachement attacher attaque atteindre attente attention atténuation atténuer atterrissage attribuer attribut attribution audible audience audimat audio audiocassette audioconférence audiofréquence audionumérique audiovisuel audit augmentation augmenté augmenter auteur authentification authentifier authentique auto auto-adaptable auto-adaptatif autocommutateur autocontrôle autocopiant autocorrecteur autodétecteur autodiagnostic autodidacte auto-enrichissant auto-entretenu automate automatique automatisation automatisé automatiser autonome autonomie autopositif autopsie autorafraîchi autoréponse autorisation autorisé autoriser

autorité

autoroute autotest auxiliaire avalanche avance avancé capot avancer avant avarie avertir avertissement avertisseur aveugle avion avis avorter caret axe carré **AZERTY** carré azimut carte cabestan cabine carton câblage câble câblé câbler câblogramme cache case caché cacher casse cacheté cadence cadrage cadran cadratin cadre cadrer cadre-réponse cage cahier caisse cent calcul centi calculable calculateur calculatrice calculé calculer calculette calendrier calibrage centre calibre calibrer callier calligraphie cercle cambridge camembert caméra Caméscope canal canalisation candela

canon

canonique

capable capacitance capacité capactif capitale capteur capture capturer caractère caractéristique carbone carboné cardinal cardioïde cartésien cartonnage cartonné cartonner cartouche cascade casque cassette catalogue cataloguer catastrophe cathode cathodique cavalier cédille cellulaire cellule centile centimètre centrage central central centrale centralisé centrer Centronics céramique certain césure chaînage citer chaîne clair chaîné clairement chaîner clapet clarté chaleur classe chambre champ classement

changement changer changeur chapelet chapitre char charbon charge chargement charger chargeur chariot chasse châssis chaud chauffer chef chemin cheminement chemise chercher chevauchement chevaucher chiffrage chiffre chiffré chiffrement chiffrer chimique choc choix chroma chromaticité chromatique chrominance chronique chronologique chronomètre chronométrie chute cible cinéaste cinéma cinématographie cinématographique cinq cinquième circonflexe circuit circuiterie circulaire circulant circulation circuler citation

classer classeur classification clause clavardage clavier claviste clé clef clic cliché client clianotant clianotant clianotement clignoter clip clipart clip-art cliquer clone clore clôture coaxial COBOL codage code codé coder codeur coercibilité coffret cogniticien cohérence cohérent coin coincement coincer coïncidence collationnement collationner collecte collecteur collectif collection coller collision colonne colophon colorier combinaison combinatoire combiné combiner commande commandé commencement commencer commentaire commercial commercialisation commercialiser

commun communauté communication communiqué communiquer commutateur commutation commuté compact compactage compacté compacter Compag comparable comparaison comparateur comparer compatibilité compatible compensateur compilateur compilation compiler complément complémentaire complémenté complet complètement compléter complexe complexité compliqué comporter composant composé composer composite compositeur composition compresser compresseur compression comprimer compris comptabilité comptage compte compter compteur CompuServe concaténation concaténé concaténer concave concentrateur concentrateurdéconcentrateur concentration concentrer concepteur conception conceptuel

concevoir concu concurrence concurrent condensateur condensation condensé condenser condition conditionnel conditionnement conditionner conducteur conductibilité conductible conduction conduire conduit conduite cône conférence confetti confiance confidentialité confidentiel configuration configurer confirmation confirmer conforme conformer conique conionction connaissance connecté connecter connecteur connectif connectique connectivité connexe connexion conscient consécutif consécutivement conseil conserver console consommateur consommation consommer constant constante constituer constructeur constructif construction construire consultation consulter contact correcteur

contacter

contenir contention contenu contexte contextuel contigu continu continuation continuel continuellement continuité contraint contrainte contraire contrarotatif contrastant contraste contrasté contrat contre contre-épreuve contrefaçon contre-jour contre-rotation contrôlable contrôle contrôlé contrôler contrôleur convenir convention conventionnel conversation conversationnel conversion convertibilité convertible convertir convertisseur convexe convivial convivialité convoyeur coopératif coordination coordonnée coordonner copie copier copieur co-processeur copyright coquille CORAL corbeille cordon co-résident coroutine corpus correct

correction

correspondance correspondant correspondre corriger corrompre corrompu corruption cosmique côté couche couler couleur coulomb coup coupe coupe-circuit couper couper/coller couplage coupler coupleur coupure courant courant courbe Courier courir courrier course court court-circuit court-circuiter coussinet couture couverture crash crasse crayon créateur création crédit créer crénage créneau crénelage crête creux cristal critère critique crochet croisé croissant croix cryptage crypter cryptographie cryptographique cuivre cumul cumulande cumulateur

cumulatif curatif curseur cvan-magentaiaune-noir cybernétique cvcle cyclique cylindre packageur page pagination paginer pair pair paire pair-impair pâle palette palier pâlir panier panne panneau panoplie Pantone papier papier-calque papillotement paquet par parabolique paragraphe paraître parallèle paramétrable paramétrage paramètre paramétré paramétrer paramétrique parasite paravent parc parcage parcourir parent parenthèse parité parole parquer partage partagé partager partial particule particulier partie partiel partiellement partition

parution parvenir pas **PASCAL** passage passager passe passe-bande passe-bas passe-haut passer passerelle passionné pastille patch pause pavé payer pays paysage peindre peinture pellicule pelliculé pelliculer pelure pendant pense-bête **Pentium** pépin perceptible perche perdre perdu père péremption perforateur perforateurcompteur perforation perforatrice perforé perforer performance performant périgée périmé période périodique périodiquement périphérique permanence permanent perméabilité permettre permis permis permission permutable permutation

persistance persistant personnalisé personnaliser personne personnel personnel perte pertinence pertinent perturbation péta péta-octet petit phase phone phonème phonétique phosphore phosphoré phosphorescence phosphorescent photo photocellule photocomposé photocomposeuse photocomposition photoconducteur photoconductivité photocopie photocopier photocopieur photodiode photoélectricité photoélectrique photoémission photographie photographier photographique photographiquemen photogravure photo-litho photolithographie photomécanique photométrie photon photosensible photostat phototransistor photovoltaïque physique pic pica pico picofarad picoseconde picot pièce

pied

pile

piézoélectrique

polaire

permuté

permuter

paru

piller PILOT pilotage pilote piloté piloter pinceau piqué piratage pirate piraté pirater piste pitch pixel placard place placement placer plage plan plan planar planche planète planification planifier planning plantage planter plaque plaquette plasma plat plateau plate-forme platine plein pleurage pli pliage plier plieuse pliure plomb plongeant plume plus poche pochette pochoir poids poignée poincon point pointage pointe pointeur pointillé pointillé point-virgule

polarisation polarisé polarité polaroïde police policé pollution polonais polymérisé polynomial polyvalent pomme pompage ponctuation pondération pondéré pondérer pont port portabilité portable portage portance portatif porte portée porteur porteuse pose positif positif position positionnel positionnement positionner possibilité possible post postal poste poster post-processeur **PostScript** postsynchronisé

postsynchronisation potentiel potentiomètre poubelle pouce pourcentage poursuite poussé PowerBook pratique préallocation préampli préamplificateur préamplifier préavis précaution précédemment précédent

précéder précis préciser précision précompilé préconditionner préconiser prédéfini prédéterminé prédicat prééditer préemptif préenregistré préenregistrer préétabli préférentiel préfixe préfixé préformaté préimprimé prélever prématuré prématurément premier prémixage préparation préprocesseur

préproduction préprogrammé préprogrammer préréglé pré-sélectionné présent présentation présenter présentoir presse presse-papiers presseur pression prêt prêt prétraitement prétraiter préventif prévention prévisualisation prévisualiser primaire primitif primitive principal principalement prioritaire

priorité prise privatif privé privilège privilégié problème procédé procédure processeur processus productif production productique produire produit produit profession professionnel profil profondeur progiciel programmable programmateur programmathèque programmation programme

programmé programmer programmeur progressif projecteur projection projet projeter PROLOG promouvoir propagation propagé propager proportion proportionnel proportionnellement propriétaire propriété prospectus protecteur protection protégé protéger protocole prototypage prototype

provenant

provenir provisoire pseudo pseudo-aléatoire pseudocode pseudo-instruction pub public public publication publicitaire publicité publier publipostage puce puissance puits pupitre pupitreur pur purée purger pylône pyramidal

#### Méthode 2

abandonné abandonner abaque aberration abîmé abîmer abonné abrégé abréviation abscisse absolu absorber accédant accéder accélérateur accélérateur accélération accéléré accent accentué acceptable acceptation accepté accepter accès accessible accessoire accidentel accolades accommodation accompagnement accord accordé accordéon accorder accordeur accrochage accroché accrocher accumulateur accumuler accusé accuser achat acheminement acheminer acheter achèvement achever achromatique Ackerman acoustique acquisition Acrobat acronyme actif actinique actionné actionner

activation activé activer activité actualisation actuel acuité acutance adaptable adaptateur adaptatif adaptation adapté adapter additif addition additionnel additionner additionneur additionneursoustracteur adhésif adhésif adiacent administrateur Adobe adoucir adressable adressage adresse adresser aérien aérographe affaiblir affaiblissement affectable affectation affecter affichage affiche afficher affirmatif affirmation agence agenda agent agglomérer agrandir agrandissement agréer agrégat agrégé aide aider aigu aiguillage aiguille aimant aimanté

aire ajouté ajouter ajustement aiuster alarme album albumen alcalin alcool aléatoire alerte alerter alerteur algèbre algébrique Algol algorithme algorithmique alias aligné alignement alimentation alimenter aller aller v. allocation allonger allophone alloué allouer allumer alphabet alphabétique alphabétiquement alphagéométrique alphamosaïque alphanumérique alphaphotographiqu altération altérer alternance alternat alternateur alternatif alternative alterné alterner ambiant ambigu ambiguïté ambisonique âme amélioration

amélioré

améliorer

Amiga

amont

amorçage amorce amorcer amortir amortissement amovible ampère ampli amplificateur amplification amplifié amplifier amplitude analogique analogiquenumérique analogue analyse analyser analyseur analyste anamorphose anastigmate ancêtre ancrage angle angström animé animer anneau annexe annexer annonciateur annonciateur annotation annoter annuaire annulaire annulation annuler anode anomale anomalie anormal anormalement ansi antémémoire antenne antérieur antérieurement anthropomorphique anti anticipation anticipé anticopie anticrénelage anti-écho antiparasite antirebond

antireflet antislash antistatique antivirus aperçu apochromatique apogée apostrophe apparaître appareil apparence apparié appel appeler appendice Apple appleshare appletalk application appliquer apprécier apprendre apprentissage approbation approché approuver approximatif approximation approximativement appuyer après après-vente arabe arbitrage arborescence arborescent arbre archétype Archimède architecture archivage archivé archiver archives archiviste **ARCNEt** arête argument armature armé armer arrangement arrêter arrière arrière-plan arrivée arrondi arrondir arrondissement

arséniure article artificiel ascendant ascenseur aspect assemblage assembler assembleur assembleuse assertion asservi assigné assigner assistance assistant assisté assister associatif association associé associer assortir assourdir assourdissement assurance assurer astable astérisque astigmatisme astucieux asymétrique asynchrone Atari atelier atmosphère atmosphérique atome atomique attachement attacher attaque atteindre attente attention atténuation atténuer atterrissage attribuer attribut attribution audible audience audimat audio audiocassette audioconférence audiofréquence audionumérique audiovisuel audit augmentation

auamenté augmenter auteur authentification authentifier authentique auto auto-adaptable auto-adaptatif autocommutateur autocontrôle autocopiant autocorrecteur autodétecteur autodiagnostic autodidacte auto-enrichissant auto-entretenu automate automatique automatisation automatisé automatiser autonome autonomie autopositif autopsie autorafraîchi autoréponse autorisation autorisé autoriser autorité autoroute autotest auxiliaire avalanche avance avancé avancer avant avarie avertir avertissement avertisseur aveugle avion avis avorter axe **AZERTY** azimut cabestan cabine câblage câble câblé câbler câblogramme cache

caché

cacher

cacheté cadence cadrage cadran cadratin cadre cadrer cadre-réponse cage cahier caisse calcul calculable calculateur calculatrice calculé calculer calculette calendrier calibrage calibre calibrer callier calligraphie cambridge camembert caméra Caméscope canal canalisation candela canon canonique capable capacitance capacité capactif capitale capot capteur capture capturer caractère caractéristique carbone carboné cardinal cardioïde caret carré carré carte cartésien carton cartonnage cartonné cartonner cartouche cascade case casque

casse

cassette choc catalogue cataloguer catastrophe cathode cathodique cavalier cédille cellulaire cellule cent centi centile centimètre centrage central central centrale centralisé centre centrer Centronics céramique cercle certain césure chaînage chaîne chaîné chaîner chaleur chambre champ changement changer changeur chapelet chapitre char charbon charge chargement charger chargeur chariot chasse châssis chaud chauffer chef chemin cheminement chemise chercher chevauchement chevaucher chiffrage chiffre chiffré chiffrement chiffrer chimique

choix chroma chromaticité chromatique chrominance chronique chronologique chronomètre chronométrie chute cible cinéaste cinéma cinématographie cinématographique cinq cinquième circonflexe circuit circuiterie circulaire circulant circulation circuler citation citer clair clairement clapet clarté classe classement classer classeur classification clause clavardage clavier claviste clé clef clic cliché client clignotant clignotant clignotement clignoter clip clipart clip-art cliquer clone clore clôture coaxial COBOL codage code codé coder

codeur coercibilité coffret cogniticien cohérence cohérent coin coincement coincer coïncidence collationnement collationner collecte collecteur collectif collection coller collision colonne colophon colorier combinaison combinatoire combiné combiner commande commandé commencement commencer commentaire commercial commercialisation commercialiser commun communauté communication communiqué communiquer commutateur commutation commuté compact compactage compacté compacter Compag comparable comparaison comparateur comparer compatibilité compatible compensateur compilateur compilation compiler complément complémentaire complémenté complet complètement compléter

complexe complexité compliqué comporter composant composé composer composite compositeur composition compresser compresseur compression comprimer compris comptabilité comptage compte compter compteur CompuServe concaténation concaténé concaténer concave concentrateur concentrateurdéconcentrateur concentration concentrer concepteur conception conceptuel concevoir conçu concurrence concurrent condensateur condensation condensé condenser condition conditionnel conditionnement conditionner conducteur conductibilité conductible conduction conduire conduit conduite cône conférence confetti confiance confidentialité confidentiel configuration configurer confirmation confirmer

conforme conformer conique conjonction connaissance connecté connecter connecteur connectif connectique connectivité connexe connexion conscient consécutif consécutivement conseil conserver console consommateur consommation consommer constant constante constituer constructeur constructif construction construire consultation consulter contact contacter contenir contention contenu contexte contextuel contiqu continu continuation continuel continuellement continuité contraint contrainte contraire contrarotatif contrastant contraste contrasté contrat contre contre-épreuve contrefacon contre-jour contre-rotation contrôlable contrôle contrôlé contrôler

contrôleur

convenir convention conventionnel conversation conversationnel conversion convertibilité convertible convertir convertisseur convexe convivial convivialité convoyeur coopératif coordination coordonnée coordonner copie copier copieur co-processeur copyright coquille CORAL corbeille cordon co-résident coroutine corpus correct correcteur correction correspondance correspondant correspondre corriger corrompre corrompu corruption cosmique côté couche couler couleur coulomb coup coupe coupe-circuit couper couper/coller couplage coupler coupleur coupure courant courant courbe Courier courir courrier course

court court-circuit court-circuiter coussinet couture couverture crash crasse crayon créateur création crédit créer crénage créneau crénelage crête creux cristal critère critique crochet croisé croissant croix cryptage crypter cryptographie cryptographique cuivre cumul cumulande cumulateur cumulatif curatif curseur cyan-magentajaune-noir cybernétique cvcle cyclique cylindre packageur page pagination paginer pair pair paire pair-impair pâle palette palier pâlir panier panne panneau panoplie Pantone papier papier-calque

papillotement

photocopie paquet pépin planter positif par photocopier plaque perceptible positif photocopieur parabolique perche plaquette position paragraphe perdre photodiode plasma positionnel paraître perdu photoélectricité plat positionnement parallèle photoélectrique plateau positionner père paramétrable péremption photoémission plate-forme possibilité photographie possible paramétrage perforateur platine paramètre perforateurphotographier plein post photographique postal paramétré compteur pleurage perforation photographiquemen paramétrer pli poste paramétrique pliage poster perforatrice parasite perforé photogravure plier post-processeur photo-litho plieuse **PostScript** paravent perforer parc performance photolithographie pliure postsynchronisation photomécanique postsynchronisé parcage performant plomb parcourir périgée photométrie plongeant potentiel périmé parent photon plume potentiomètre parenthèse période photosensible plus poubelle parité périodique photostat poche pouce parole périodiquement phototransistor pochette pourcentage parquer périphérique photovoltaïque pochoir poursuite physique poussé partage permanence poids poignée PowerBook pic partagé permanent pica poinçon pratique partager perméabilité pico préallocation partial permettre point particule picofarad pointage préampli permis particulier picoseconde préamplificateur permis pointe permission picot pointeur préamplifier partie partiel permutable pièce pointillé préavis partiellement permutation pied pointillé précaution partition piézoélectrique point-virgule précédemment permuté permuter pile précédent paru polaire piller parution persistance polarisation précéder PILOT précis parvenir persistant polarisé pilotage pas personnalisé polarité préciser **PASCAL** personnaliser pilote polaroïde précision piloté passage personne police précompilé piloter préconditionner passager personnel policé personnel pinceau pollution préconiser passe prédéfini passe-bande perte piqué polonais passe-bas pertinence piratage polymérisé prédéterminé pirate polynomial prédicat passe-haut pertinent perturbation piraté polyvalent prééditer passer passerelle péta pirater pomme préemptif passionné péta-octet piste pompage préenregistré pastille petit pitch ponctuation préenregistrer phase pixel pondération préétabli patch préférentiel phone placard pondéré pause phonème place préfixe pavé pondérer phonétique placement préfixé payer pont pays phosphore placer port préformaté plage phosphoré portabilité préimprimé paysage peindre phosphorescence plan portable prélever peinture phosphorescent plan portage prématuré pellicule planar portance prématurément photo pelliculé photocellule planche portatif premier pelliculer photocomposé planète prémixage porte pelure photocomposeuse planification portée préparation photocomposition préprocesseur pendant planifier porteur photoconducteur pense-bête planning préproduction porteuse Pentium photoconductivité plantage préprogrammé pose

préprogrammer préréglé pré-sélectionné présent présentation présenter présentoir presse presse-papiers presseur pression prêt prêt prétraitement prétraiter préventif prévention prévisualisation prévisualiser primaire primitif

primitive principal principalement prioritaire priorité prise privatif privé privilège privilégié problème procédé procédure processeur processus productif production productique produire produit produit

profession professionnel profil profondeur progiciel programmable programmateur programmathèque programmation programme programmé programmer programmeur progressif projecteur projection projet projeter PROLOG promouvoir prompt

propagation propagé propager proportion proportionnel proportionnellement propriétaire propriété prospectus protecteur protection protégé protéger protocole prototypage prototype provenant provenir provisoire pseudo pseudo-aléatoire

pseudocode pseudo-instruction bub public public publication publicitaire publicité publier publipostage puce puissance puits pupitre pupitreur pur purée purger pylône pyramidal