LES CAHIERS DE L'ANALYSE DES DONNÉES

L. LEBART

Une procédure d'analyse lexicale écrite en langage FORTRAN

Les cahiers de l'analyse des données, tome 6, n° 2 (1981), p. 229-243

<http://www.numdam.org/item?id=CAD_1981__6_2_229_0>

© Les cahiers de l'analyse des données, Dunod, 1981, tous droits réservés.

L'accès aux archives de la revue « Les cahiers de l'analyse des données » implique l'accord avec les conditions générales d'utilisation (http://www.numdam.org/conditions). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme Numérisation de documents anciens mathématiques
http://www.numdam.org/

Les Cahiers de l'Analyse des Données
Vol. VI - 1981 - n° 2 - p. 229-243

UNE PROCÉDURE D'ANALYSE LEXICALE ÉCRITE EN LANGAGE FØRTRAN [ANA. LEX.]

par L. Lebart (1)

On va brièvement présenter le principe et les éléments techniques d'un petit programme portable (c'est-à-dire aisément adaptable à tout système de calcul) destiné à calculer et à représenter les profils lexicaux d'une série de textes.

On décrira successivement :

- la fonction et le champ d'application,
- les caractéristiques techniques (comportant le listage commenté du programme),
- le mode d'emploi.

1 *Fonction et champ d'application*

A l'origine, le programme présenté ici n'était qu'un module dans une chaîne de traitement des réponses libres aux questions ouvertes d'une enquête (cf. "Pratique de l'Analyse des Données", Tome 3 - LC3 - n° 4 [REP. LIBR.]).

Sa fonction est la suivante : à partir d'un recueil de NKLA textes mis sur support informatique en "format libre", sous forme littérale, sans altération ni codage préalable, il calcule la table de contingence croisant les NKLA textes et les MMAX formes graphiques* les plus fréquentes (NKLA est une donnée, MMAX est un paramètre), puis soumet cette table à une analyse des correspondances.

Le programme dénombre donc des chaînes de caractères séparées par une famille de séparateurs (ici : le blanc, le point, la virgule, l'apostrophe ; la liste peut évidemment être complétée ou changée) ; parmi les caractères, peuvent figurer non seulement des lettres, mais des chiffres, ce qui peut aider à transcrire commodément un alphabet quelconque.

Dans l'application aux traitements des réponses libres précitée, les textes étaient des regroupements de réponses, selon une partition des individus interrogés réalisée antérieurement.

Ces textes peuvent être, comme dans le cas d'applications maintenant classiques, des discours politiques, des tracts politiques ou syndicaux, des pièces de théâtre, des recueils de poèmes. Dans le jeu d'essai présenté ci-après, destiné uniquement à illustrer le fonctionnement du programme, on a mis seulement un poème (ou extrait de poème) par classe. Les résultats n'auront donc qu'une valeur anecdotique.

(*) N.B. Bien que par abus de langage nous appelions parfois ces formes des *mots*, on prendra garde que pour nous blanche est une autre forme que *blanches*, tandis que *avoir* (nom) est une même forme que *avoir*(verbe).

(1) *Maître de Recherche au CNRS. CREDOC - 142, rue du Chevaleret - 75013 Paris.*

Le programme s'applique cependant, dans son principe, à toute forme de relevé descriptif en clair (et en vrac!).

Exemple : - Listes nominales de conseils d'administration de grandes entreprises (pour étudier les membres communs).

- Tables des matières de lots de revues.

- Listes de noms latins d'insectes (avec répétition) relevés dans des parcelles, etc.

2 Caractéristiques techniques

2.1 Code : Le programme présenté est écrit en FORTRAN IV standard. Le sous-ensemble d'instructions utilisé le rend compatible avec les principaux matériels disponibles.

2.2 Structure générale, agencement des programmes : Le programme principal, ANALEX, constitue une interface avec une présentation courante des données de base : les textes sont tapés en format libre, séparés par une carte facile à identifier (ici, commençant par **** , le dernier texte étant terminé par une carte commençant par ====).

Le nombre de textes, et le nombre de cartes dans chaque texte sont calculés dans ANALEX.

Le sous programme principal CORLEX réserve la place des tableaux de travail et appelle les deux volets de calcul =

MOTS : qui calcule la table de contingence

PLUMK : qui réalise son analyse factorielle.

En fait, le corps de la procédure est constitué par le programme MOTS. On n'a pas jugé utile de publier à nouveau ici un programme d'analyse des correspondances. L'interface réalisée par PLUMK est suffisamment claire pour que les lecteurs puissent implémenter leur version à la suite (*).

Les deux utilitaires appelés et non publiés sont SHELK (Programme de tri classique) et CPLUM2 (Analyse des correspondances légère) , (voir "Technique de la Description Statistique" par L. Lebart, A. Morineau, N. Tabard - DUNOD - 1977).

2.3 Réservation de mémoire, limites : Le tableau principal (réservé dans le sous-programme CORLEX, instruction 620) doit surtout son importance au tableau LMOT du sous-programme MOTS ; on a en effet : 23500 éléments dans ce tableau, correspondant aux lettres de chacun des mots distincts retenus.

Sous cette forme, le programme ne dépend pas de la longueur des mots-machine (on utilise un mot-mémoire par lettre).

Le volume-mémoire peut être fortement réduit si l'on convertit le fichier, après avoir reconnu les mots distincts, en mots multiples de 4 caractères (pour les machines ayant 4 octets par mot adressable), en les complétant éventuellement avec des "blancs".

(*) Une implémentation de la procédure adaptée aux traitements des réponses libres fait partie du système SPADE1 (Système portable pour l'Analyse des Données d'Enquêtes) et peut être obtenue en écrivant au CESIA - 82, rue de Sèvres - 75007 Paris.

La version présentée ici est limitée à des tables de contingence (300 × 25). (On extrait les 300 mots les plus fréquents pour 25 classes, cf. ligne 1220, 1530, du sous-programme MOTS).

Il est facile de s'affranchir de ces limites en modifiant les dimensions correspondantes.

2.4 *Principe de fonctionnement* : Des places sont réservées à l'avance aux formes : le tableau NLETR (cf. instruction n° 1250) est un histogramme empirique des formes selon leur longueur : 30 formes d'une lettre, 90 de deux, 150 de trois, etc. . Cette distribution correspond aux observations faites sur des corpus formés de 4000 réponses à des questions ouvertes. Il va de soi que le tableau NLETR peut être modifié à propos d'autres applications, surtout s'il s'agit d'autres langues que le français.

Si l'on rencontre, par exemple, plus de 450 mots distincts ayant exactement 7 lettres, le 451-ème et les suivants seront ignorés (le dépassement est cependant signalé, ainsi que le nombre réel de mots de 7 lettres rencontrés). Il est ainsi possible de réajuster la distribution après un passage sur un corpus volumineux et original dans sa composition. On prendra garde que le paramètre LDIM (ligne 760 de CORLEX) doit majorer :

$$LDIM \geq \Sigma\{I \times NLETR(I) \mid I = 1, 16\}$$

Seules, les 16 premières lettres des mots sont prises en compte.

Les commentaires figurant dans le listage devraient permettre de suivre les différentes étapes du programme.

3 *Mode d'emploi et jeu d'essai*

Le programme ne lit qu'une carte paramètre (6 paramètres) (ligne 230 de ANALEX).

- NSEU = Seuil de fréquence minimale des formes.
- MMAX = Nombre maximal de formes finalement retenues.

Chacun de ces deux paramètres délimite le volume des sorties.

- MOMAX = délimite la longueur minimale des mots (en général 2 ou 3).
- NGRAF, NPAGE, NLIGR = désignent respectivement le nombre de graphiques, de pages, de lignes au niveau de l'analyse factorielle.

LE JEU D'ESSAI

- La première ligne correspond à la carte paramètre :

NSEU = 1 (les mots doivent apparaître au moins deux fois)
MMAX = 120 (on ne retiendra pas plus de 120 mots)
MOMAX = 2 (les mots retenus auront plus de 2 lettres)

Enfin l'unique graphique aura 1 page et 58 lignes.

- Les lignes suivantes constituent les textes : 5 poèmes ou extraits de poèmes sur le thème de l'automne de LAMARTINE, T. GAUTHIER, VERLAINE, C. BRUGNOI, BAUDELAIRE.

On donne comme exemple de sortie le tableau des profils lexicaux par classe (ici : par auteur) et le plan des deux premiers facteurs de l'analyse.

Remarques : Il est conseillé de sauvegarder le fichier LNOMB , qui contient les identificateurs des formes et la table de contingence, de façon à pouvoir refaire d'autres analyses où seraient désignées des formes et des classes supplémentaires.

La procédure fait intervenir quatre fichiers de travail ; LEM , LNOMB, NB1, NB2.

Les deux derniers (NB1 et NB2) peuvent être omis si l'on supprime l'appel de PLUMK dans CORLEX (cas où l'on procèderait en deux étapes séparées).

Références des textes du jeu d'essai :

TEXTE 1 : Extrait de "pensées des morts" de Lamartine (4 premières strophes - Harmonies)

TEXTE 2 : "Pensées d'automne" de Théophile Gauthier (Tarbes 1811-Neuilly 1872) (Premières poésies - 1983-1832)

TEXTE 3 : "Chanson d'automne" de Verlaine" (Poèmes saturniens)

TEXTE 4 : "Brouillards d'automne" de Charles Brugnot (1797-1831) (Poésies) (Ami d'Aloysius Bertrand, publié pour la première fois en 1834)

TEXTE 5 : "Chant d'automne" de Baudelaire (Fleurs du mal)

- ANNEXE I -

LISTAGE DU PROGRAMME

```
00010 C* * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *
00020 C                    PROGRAMME    ANALEX
00030 C* * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *
00040 C========  EXEMPLE DE PROGRAMME PRINCIPAL POUR LIRE LE TEXTE BRUT
00050 C========   ET LE COPIER SUR LE SUPPORT    LEM
00060 C* * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *
00070       DIMENSION NLIG(100),LIGNE(20)
00080       DATA  LSTAR / 4H****/ ,LEGAL / 4H====/
00090 C========LSTAR SEPARE LES CLASSES. LEGAL INDIQUE LA FIN DU TEXTE.
00100       COMMON /ENSOR/ LEC,IMP
00110       LEC = 5
00120       IMP = 6
00130       LNOMB = 10
00140       LEM = 11
00150       NB1 = 14
00160       NB2 = 15
00170 C========LEM,LNOMB, NB1 ET NB2 =QUATRE FICHIERS DE TRAVAIL
00180 C========*
00190       REWIND LEM
00200 C========*
00210 C========  LECTURE DE LA CARTE PARAMETRE UNIQUE
00220 C========*
00230       READ( LEC,200) NSEU,MMAX,MOMAX,NGRAF,NPAGE,NLIGR
00240       WRITE(IMP,300) NSEU,MMAX,MOMAX,NGRAF,NPAGE,NLIGR
00250 C========    .POUR LA SIGNIFICATION DES PARAMETRES ,VOIR  .CORLEX.
00260       DO 5 K = 1,100
00270     5 NLIG(K) = 0
00280       NKLA = 1
00290 C======== NKLA SERA LE NOMBRE DE TEXTES SEPARES PAR LSTAR=****
00300 C========. LECTURE DES CARTES EN 20 A4...
00310    10 CONTINUE
00320       READ(LEC,100) (LIGNE(L),L=1,20)
00330   100 FORMAT(  20A4 )
00340        IF( LIGNE(1)    .EQ.   LEGAL  )    GO TO 20
00350       IF(LIGNE(1) .EQ. LSTAR)              GO TO 15
00360       NLIG(NKLA) = NLIG(NKLA) + 1
00370 C======== IMPRESSION  DU TEXTE SUR LE SUPPORT LEM (ET SUR L IMPRIMANTE)
00380       WRITE(LEM,100)(LIGNE(L),L=1,20)
00390       WRITE( IMP, 400) NKLA, NLIG(NKLA),(LIGNE(L),L=1,20)
00400   400 FORMAT( 1H ,I2,5X,I4,5X,20A4)
00410       GO TO 10
00420   15 NKLA=NKLA+1
00430       GO TO 10
00440    20 CONTINUE
00450 C========*
00460       CALL CORLEX(NLIG,NKLA,80,NSEU,MMAX,MOMAX,LEM,LNOMB,
00470      1  NGRAF,NPAGE,NLIGR,NB1,NB2)
00480 C========*
00490   200 FORMAT( 20I4)
00500   300 FORMAT ( 1H ,///,6H NSEU= ,I4,5X,6H MMAX= ,I4,5X,6HMOMAX=,I4,//,
00510      1   15X,6HNGRAF=,I4,5X,6HNPAGE= ,I4,5X,6HNLIGR= ,I4,//)
00520       RETURN
00530       END
00540 C
```

```
00550       SUBROUTINE CORLEX(NLIG,NKLA,NCOL,NSEU,MMAX,MOMAX,LEM,LNOMB,
00560      1  NGRAF,NPAGE,NLIGR,NB1,NB2)
00570 C* * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *
00580 C* SOUS PROGRAMME PRINCIPAL DE CORRESPONDANCE LEXICALE
00590 C* * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *
00600 C*'    DATE DE MISE A JOUR = 6/03/80
00610 C* * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *
00620       DIMENSION Q(300500),NLIG(90)
00630       COMMON / ENSOR / LEC,IMP
00640 C* * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *
00650 C======== LE TEXTE EST SUR LE FICHIER LEM ,EN NCOL*A1
00660 C======== NKLA = NOMBRE DE CLASSES
00670 C======== NLIG(K) = NOMBRE DE LIGNES DE LA CLASSE K
00680 C======== NSEU = SEULS,LES MOTS APPARAISSANT PLUS DE NSEU FOIS SONT
00690 C======== RETENUS
00700 C======== MMAX  = ON NE RETIENDRA PAS PLUS DE MMAX MOTS POUR L ANALYSE
00710 C======== MOMAX = ON ELIMINE LES MOTS AYANT MOINS DE MOMAX LETTRES
00720 C======== NGRAF = NOMBRE DE PLANS FACTORIELS REPRESENTES
00730 C======== NPAGE = NOMBRE DE PAGES DES GRAPHIQUES
00740 C======== NLIGNE= NOMBRE DE LIGNES DES GRAPHIQUES
00750 C* * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *
00760       LDIM = 23800
00770     .  NDIM = 3350
00780 C======== NDIM MAJORE LE NOMBRE DE MOTS DIFFERENTS RETENUS
00790 C======== LDIM MAJORE LE NOMBRE DE LETTRES DE CES MOTS
00800 C======== (VOIR LE TABLEAU NLETR EN DATA DANS LE SOUS PROGRAMME MOTS)
00810       LDIM1 = LDIM + 1
00820       LDIM2 = LDIM1 + NDIM
00830 C======== CONSTRUCTION DE LA TABLE   MOTS-CLASSES
00840 C========*
00850       CALL MOTS(NLIG,NCOL,NSEU,Q(1),Q(LDIM1),Q(LDIM2),MMAX,MOMAX,LDIM,
00860      1  NDIM,NKLA,IMOT,LEM,LNOMB)
00870 C========*
00880       JMAX = NKLA +1
00890       IF(JMAX .LT. 6) JMAX = 6
00900       MJAX = JMAX + MMAX
00910       LQ2 = JMAX*JMAX + 1
00920       LQ3 = LQ2 + JMAX
00930       LQ4 = LQ3 + MJAX
00940       LQ5 = LQ4 + MJAX*6
00950       LQ6 = LQ5 + MJAX
00960       LQ7 = LQ6 + MJAX*3
00970       LQ8 = LQ7 + MJAX
00980 C======== ANALYSE DES CORRESPONDANCES DE LA TABLE FORMES X CLASSES
00990 C========*
01000       CALL PLUMK(IMOT,IMOT,NKLA,NKLA,LNOMB,JMAX,MJAX,
01010      1 Q(1),Q(LQ2),Q(LQ3),Q(LQ4),Q(LQ5),Q(LQ6),Q(LQ7),Q(LQ8),
01020      2  NGRAF,NPAGE,NLIGR,NB1,NB2)
01030 C========*
01040       RETURN
01050       END
01060 C
```

```
01065 C
01070       SUBROUTINE MOTS(NLIG,NCOL,NSEU,LMOT,NOMBR,NOMBK,MMAX,MOMAX,
01080      1 LDIM,NDIM,NKLA,IMOT,LEM,LNOMB)
01090 C* * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *
01100 C======== CONSTRUCTION DE LA TABLE DE CONTINGENCE CROISANT LES MOTS ET
01110 C======== LES  NKLA  CLASSES.
01120 C======== TEXTE SUR LE SUPPORT LEM ,EN 80A1...
01130 C======== EN ENTREE ,PARAMETRES OU TABLEAUX =LDIM,NDIM,NCOL,NSEU,MMAX
01140 C======== NLIG...DEFINIS DANS  .CORLEX. (NKLA =NOMBRE DE CLASSES)
01150 C======== EN SORTIE , IMOT EST LE NOMBRE DE MOTS RETENUS
01160 C========              LA TABLE (MOTS X CLASSES) EST COPIEE LIGNE
01170 C========              PAR LIGNE SUR LE SUPPORT LNOMB(FORMAT 1200)
01180 C* * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *
01190       DIMENSION LIG(80),KAR(20),LB(20),MOT(20,20),LMOT(LDIM),MO(20)
01200       DIMENSION NLETR(16),LCUM(21),KCUM(21),NOMBR(NDIM),
01210      1 LADR(500),NMOT(20),NEFF(20),LONG(500),IV(500),NADR(500),
01220      2 NLIG(90),NOMBK(NDIM),TAB(300,25),FMARG(25),POURC(25)
01230       COMMON / ENSOR / LEC,IMP
01240 C======== TABLEAU DES EFFECTIFS MAXIMA DE MOTS SELON LEUR LONGUEUR..
01250       DATA NLETR /30,90,150,250,350,420,450,420,350,250,150,
01260      1 100,60,50,40,20 /
01270 C========*
01280 C======== NLETR(K) = NOMBRE DE MOTS DISTINCTS DE K LETTRES
01290 C======== LB(K) = LISTE DES SEPARATEURS ACCEPTES
01300       DATA LB(1)/1H /,LB(2)/1H,/,LB(3)/1H./,LB(4)/1H'/
01310 C========*
01320       LCUM(1) = 0
01330       KCUM(1) = 0
01340       NLMAX = 16
01350       DO 10 K = 1,NLMAX
01360       NEFF(K) = 0
01370       LCUM(K+1) = LCUM(K) + NLETR(K)*K
01380       KCUM(K+1) =KCUM(K) +NLETR(K)
01390    10 CONTINUE
01400 C========*
01410       IF(LEM.NE.LEC) REWIND LEM
01420       REWIND LNOMB
01430 C========*
01440       MAX = KCUM(NLMAX + 1)
01450       DO 20 K=1,MAX
01460    20 NOMBR(K)=0
01470 C========*
01480       NTOT = LCUM(NLMAX +1)
01490       DO 30 NNN= 1,NTOT
01500    30   LMOT(NNN) = LB(1)
01510 C======== GRANDE BOUCLE SUR LES NKLA CLASSES (BOUCLE 160 )
01520 C======== ICI,NKLA NE DOIT PAS DEPASSER 25
01530       IF( NKLA .GT. 25)    NKLA = 25
01540       DO 160 KLA = 1,NKLA
01550       DO 40 K=1,MAX
01560    40 NOMBK(K) = 0
01570       NLIGK = NLIG(KLA)
01580       IF(NLIGK .EQ. 0)  GO TO 150
01590 C======== GRANDE BOUCLE SUR LES NLIG(KLA) LIGNES DE TEXTE (BOUCLE 140 )
01600       DO 140 I = 1,NLIGK
01610       READ(LEM,1000) (LIG(K),K=1,NCOL)
01620       M = 1
01630       KM = 0
01640       MLIG = 0
01650 C======== STOCKAGE,DANS MOT,DES MOTS DE LA LIGNE I    (BOUCLE 60)
01660       DO 60 K = 1,NCOL
01670         LK=LIG(K)
01680       IF(LK.NE.LB(1).AND.LK.NE.LB(2).AND.LK.NE.LB(3).AND.LK.NE.LB(4))
01690      1    GO TO 50
01700       IF(MLIG .EQ. 1) M = M + 1
01710       KM =0
01720       MLIG = 0
01730       GO TO 60
01740    50 CONTINUE
01750       MLIG = 1
01760       KM = KM + 1
01770       IF( KM .GT. NLMAX) KM = NLMAX
01780       MOT(KM,M) = LIG(K)
01790 C======== MOT(KM,M)=CARACTERE KM DU MOT M  DE LA LIGNE I.
01800 C========  CE MOT M A POUR LONGUEUR KAR(M)
01810        KAR(M) = KM
01820    60 CONTINUE
01830       IF(MLIG .EQ.0 ) M = M - 1
01840       IF(M .LE.0)    GO TO 140
01850 C======== RECHERCHE DES FREQUENCES OU COMPLETION DU THESAURUS..
```

```
01860 C========*
01870       DO 130 MM =1,M
01880       KM = KAR(MM)
01890       IF(KM .LE. MOMAX .OR. KM.GT. NLMAX )   GO TO 130
01900 C======== REJET DES MOTS AYANT MOINS DE MOMAX LETTRES
01910       NEFM1 = NEFF(KM)
01920       N1 = LCUM(KM) + 1
01930       IF( NEFF(KM) .GT.  NLETR(KM)  ) NEFM1 = NLETR(KM)
01940       N2 = N1 + NEFM1*KM -1
01950       IF(N2 .LT. N1)     GO TO 90
01960        DO 80 NM=N1,N2,KM
01970       NM1 = NM + KM - 1
01980       DO 70 NNN=NM,NM1
01990       IF(LMOT(NNN) .NE. MOT(NNN - NM + 1,MM))   GO TO 80
02000    70 CONTINUE
02010       GO TO 120
02020    80 CONTINUE
02030    90 CONTINUE
02040       NEFF(KM) = NEFF(KM) + 1
02050       IF(NEFF(KM) .GT. NLETR(KM) )  GO TO 110
02060       N3 = N2 + 1
02070       N4 = N3 + KM - 1
02080       LLL=KCUM(KM)
02090       KKK=NEFF(KM)
02100       NOMBR(LLL+KKK) = 1
02110       NOMBK(LLL+KKK) = 1
02120 C======== LES MOTS NOUVEAUX SONT DEPOSES DANS L ORDRE DE LEUR
02130 C======== APPARITION DANS LE VOLET DU THESAURUS LMOT CORRESPONDANT
02140 C======== A LEURS NOMBRES DE LETTRES.(REPERE PAR KCUM...).
02150       DO 100  NNN = N3,N4
02160   100 LMOT(NNN) = MOT(NNN - N3 + 1,MM)
02170   110 CONTINUE
02180       GO TO 130
02190   120 NMOZ = (NM-N1)/KM+1
02200       LLL=KCUM(KM)
02210       NOMBR(LLL+NMOZ) = NOMBR(LLL + NMOZ) + 1
02220       NOMBK(LLL+NMOZ) = NOMBK(LLL + NMOZ) + 1
02230   130 CONTINUE
02240   140 CONTINUE
02250   150 WRITE(LNOMB)(NOMBK(KZK),KZK=1,MAX)
02260   160 CONTINUE
02270 C========  FIN  DE BOUCLE SUR LES CLASSES
02280 C======== CALCUL D UN NOUVEAU SEUIL NSE2 SI MMAX EST TROP PETIT
02290       NSE2 = NSEU
02300   170   IM = 0
02310       DO 180 J = 1,MAX
02320       IF(NOMBR(J) .LE. NSE2)  GO TO 180
02330       IM = IM + 1
02340   180 CONTINUE
02350       IF(IM .GT. MMAX) NSE2 = NSE2 + 1
02360       IF(IM .GT. MMAX)   GO TO 170
02370 C========*
02380       IND1 = 1
02390       IMOT = 0
02400       IMOTD = 0
02410       IMOTT = 0
02420 C======== BILAN DU TRI ( BOUCLE 210 )
02430 C======== TASSEMENT DES VECTEURS NOMBR ET LMOT
02440       DO 210 KM = 1,NLMAX
02450       NMOT(KM) = 0
02460       NL = NEFF(KM)
02470       IF( NL .GT. NLETR(KM) )  NL = NLETR(KM)
02480       IMOTD = IMOTD + NL
02490       IF(NL .EQ. 0 ) GO TO 210
02500       DO 200 J=1,NL
02510       IND = KCUM(KM) + J
02520       IMOTT = IMOTT + NOMBR(IND)
02530       IF(NOMBR(IND) .LE. NSE2)   GO TO 200
02540       NMOT(KM) = NMOT(KM) + 1
02550       IMOT = IMOT + 1
02560       NOMBR(IMOT) = NOMBR(IND)
02570       NADR(IMOT) = IND
02580       IND3 = LCUM(KM) + (J-1)*KM + 1
02590       IND4 = IND3 + KM - 1
02600       IND2 = IND1 + KM - 1
02610       DO 190 L=IND3,IND4
02620       LL = L - IND3
02630       LMOT(IND1 + LL) = LMOT(L)
02640       LADR(IMOT) = IND1
02650       LONG(IMOT) = KM
02660   190 CONTINUE
02670       IND1 = IND1 + KM
02680   200 CONTINUE
02690   210 CONTINUE
```

```
02700 C======== CHARGEMENT DE TAB   (BOUCLE 230)
02710       REWIND LNOMB
02720       DO 230 KLA = 1,NKLA
02730       READ(LNOMB)(NOMBK(KZK),KZK=1,MAX)
02740       DO 220 J=1,IMOT
02750       IND = NADR(J)
02760   220 TAB(J,KLA) = NOMBK(IND)
02770   230 CONTINUE
02780       REWIND LNOMB
02790 C========IMPRESSION DE TAB  (BOUCLE 260 )
02800       WRITE(IMP,1100)
02810       DO 260 KE = 1,IMOT
02820       IA = LADR(KE)
02830       KA = LONG(KE)
02840       DO 240 L = 1,20
02850   240 MO(L) = 0
02860       DO 250 L = 1,KA
02870   250 MO(L) = LMOT(IA + L - 1)
02880 C========COPIE DE LA TABLE (MOTS X CLASSES) SUR LNOMB,ET EDITION
02890 C========EN VUE DE L ANALYSE DES CORRESPONDANCES PAR CPLUMK..
02900       WRITE(LNOMB,1200)(MO(L),L=1,20),(TAB(KE,K),K=1,NKLA)
02910       WRITE(IMP,1300) KE ,(MO(L),L=1,20),(TAB(KE,K) ,K=1,NKLA)
02920   260 CONTINUE
02930 C========  RECHERCHE DES MOTS LES PLUS FREQUENTS
02940       CALL SHELK( IMOT, NOMBR, IV)
02950       NEDI = IMOT
02960       WRITE(IMP,1400)
02970       WRITE(IMP,1350)(NEFF(KM),KM=1,NLMAX)
02980       WRITE(IMP,1900)(NLETR(KM),KM=1,NLMAX)
02990       WRITE(IMP,1500) IMOTT,IMOTD,NSE2,IMOT
03000       RIMOT = IMOTT
03010       WRITE(IMP,1600)(KT,KT=1,NKLA)
03020       WRITE(IMP,1700)
03030       DO 280 K = 1,NKLA
03040       FMARG(K) = 0
03050       DO 270 KE = 1,IMOT
03060   270 FMARG(K) = FMARG(K) + TAB(KE,K)
03070       IF( FMARG(K) .LT.1.E-6 ) FMARG(K) = 1.E-6
03080   280 CONTINUE
03090 C========EDITION DES FREQUENCES RELATIVES (PROFILS LEXICAUX DES CLASSES)
03100       DO 320 KE=1,NEDI
03110       DO 290 L =1,20
03120   290 MO(L) = 0
03130       KK=IMOT- KE + 1
03140       IM = IV(KK)
03150       IA = LADR( IM)
03160       KA = LONG(IM)
03170       DO 300 L = 1,KA
03180   300 MO(L) = LMOT( IA + L - 1)
03190       FK = NOMBR(KK)/RIMOT
03200       FK = 100.*FK
03210       DO 310 K = 1,NKLA
03220   310 POURC(K) = (TAB(IM,K)/FMARG(K))*100.
03230       WRITE(IMP,1800) NOMBR(KK),FK,(MO(L),L=1,20),(POURC(K),K=1,NKLA)
03240   320 CONTINUE
03250 C======== FORMATS
03260  1000 FORMAT(80A1)
03270  1100 FORMAT(1H1,30H TABLEAU DES EFFECTIFS                ,//)
03280  1200 FORMAT(20A1,25F6.0)
03290  1300 FORMAT(1X,I6,3X,20A1,3X,24F4.0)
03300  1350 FORMAT(//,21H EFFECTIFS OBSERVES   ,16I6)
03310  1400 FORMAT(1H ,///,36H EFFECTIFS SELON LE NOMBRE DE LETTRE   )
03320  1500 FORMAT(1H1,25H    NOMBRE TOTAL DE MOTS= ,I8,/,
03330     1 26H NOMBRE DE MOTS DISTINCTS= ,I8,/,
03340     1 26H       SEUIL DE FREQUENCE =  ,I8,/,
03350     1 26H   NOMBRE DE MOTS RETENUS= ,I8,/,1H ,130(1H-) )
03360  1600 FORMAT(1X,10HFREQUENCES,2X,12HPOURCENTAGES,2X,
03370     1 8HLIBELLES,13X,20I4)
03380  1700 FORMAT(1H ,130(1H-),/)
03390  1800 FORMAT( 1X,I8,7X,F7.2,2X,20A1,2X,(1H ,20F4.1))
03400  1900 FORMAT(/,21H EFF.MAXIMAUX PREVUS  ,16I6,//)
03410       RETURN
03420       END
```

```
03430 C
03440       SUBROUTINE PLUMK(ICARD,ITOT,JCARD,JTOT,LNOMB,JMAX,MJAX,
03450      1 S,D,P,COORD,CO,ID,U,V,NGRAF,NPA,NLIG,NB1,NB2)
03460 C* * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *
03470 C======== SOUS PROGRAMME APPELLANT UNE ANALYSE DES CORRESPONDANCES
03480 C========DU TABLEAU FIGURANT,IDENTIFICATEUR EN TETE,SUR LNOMB,SELON
03490 C========LE FORMAT 1000.
03500 C========LES IDENTIFICATEURS DES LIGNES SONT LES 12 PREMIERES LETTRES
03510 C========DES MOTS .LES COLONNES(CLASSES) SONT SIMPLEMENT NUMEROTEES.
03511 C======== ICARD,ITOT =LIGNES ACTIVES,SUPPLEMENTAIRES
03512 C======== JCARD,JTOT =COLONNES ACTIVES,SUPPLEMENTAIRES
03513 C======== JMAX,MJAX =PARAMETRES DE DIMENSIONS
03520 C* * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *
03530       COMMON  / BBB / NB,NC,NGUS,NBEL,NPER,NLEX,NGRI
03531 C========CE COMMON TRANSMET LES FICHIERS NB(DONNEES) ET NGRI (FICHIER
03532 C========DE TRAVAIL) AU SOUS PROGRAMME CPLUM2 ,QUI PEUT ETRE REMPLACE
03533 C========PAR UN UTILITAIRE ANALOGUE...
03540       DIMENSION S(JMAX,JMAX),D(JMAX),P(MJAX),COORD(MJAX,6),CO(MJAX)
03550       DIMENSION ID(MJAX,3),U(MJAX),V(MJAX),IZ(25)
03560 C========LISTE DES IDENTIFICATEURS DE CLASSES POUR LES GRAPHIQUES..
03570       DATA IZ /4H *1*,4H *2*,4H *3*,4H *4*,4H *5*,4H *6*,4H *7*,
03580      1 4H *8*,4H *9*,4H*10*,4H*11*,4H*12*,4H*13*,4H*14*,4H*15*,
03590      1 4H*16*,4H*17*,4H*18*,4H*19*,4H*20*,4H*21*,4H*22*,4H*23*,
03600      1 4H*24*,4H*25* /
03610       DATA LBLAN/4H    /
03620       KFAC = 6
03630 C======== PARAMETRES DE LA REPRESENTATION GRAPHIQUE
03640         IF( NGRAF .EQ. 0 )      RETURN
03650         IF( NPA .EQ. 0 )        NPA = 1
03660 C========  NB  ET  NGRI  SONT DEUX FICHIERS DE TRAVAIL INTERNES.
03670       NB = NB1
03680       NGRI = NB2
03690 C========LES MOTS ET LA TABLE SONT LUS SUR LE FICHIER LNOMB
03700       REWIND LNOMB
03710       REWIND NB
03720       DO 1 I=1,ITOT
03730       JJ = I + JTOT
03740       READ(LNOMB,1000) (ID(JJ,K),K=1,3), (U(L),L=1,JTOT)
03750  1000 FORMAT(3A4,8X,25F6.0)
03760       WRITE(NB)(U(L),L=1,JTOT)
03770 C======== LE TABLEAU A ANALYSER EST COPIE SANS FORMAT SUR NB.
03780     1 CONTINUE
03790       DO 10 J=1,JTOT
03800       ID(J,1) = IZ(J)
03810       ID(J,2) = LBLAN
03820    10 ID(J,3)=LBLAN
03830 C======== LE PROGRAMME CPLUM2 D ANALYSE DES CORRESPONDANCES LEGERE
03840 C======== EST IDENTIQUE AU SOUS PROGRAMME CPLUM DU LIVRE--TECHNIQUE DE
03850 C======== LA DESCRIPTION STATISTIQUE--A CECI PRES QU IL AUTORISE DES
03860 C======== IDENTIFICATEURS PLUS LONGS POUR LES GRAPHIQUES
03870       CALL CPLUM2(MJAX,JMAX,KFAC,ICARD,ITOT,JCARD,JTOT,NGRAF,
03880      1 S,D,P,COORD,CO,ID,U,V,NPA,NLIG)
03890 C========
03900       RETURN
03910       END
```

- ANNEXE II -

JEU DE DONNEES D'ESSAI

```
   1 120   2   1   1  58
VOILA LES FEUILLES SANS SEVE ,QUI TOMBENT SUR LE GAZON
VOILA LE VENT QUI S ELEVE ,ET GEMIT DANS LE VALLON
VOILA L ERRANTE HIRONDELLE,QUI RASE DU BOUT DE L AILE,
L EAU DORMANTE DES MARAIS.....
VOILA L ENFANT DES CHAUMIERES,QUI GLANE SUR LA BRUYERE,
LE BOIS TOMBE DES FORETS.....
C EST LA SAISON OU TOUT TOMBE,SOUS LES COUPS REDOUBLES DES VENTS
UN VENT QUI VIENT DE LA TOMBE ,MOISSONNE AUSSI LES VIVANTS
ILS TOMBENT ALORS PAR MILLE,COMME LA PLUME INUTILE
QUE L AIGLE ABANDONNE AUX AIRS
LORSQUE DES PLUMES NOUVELLES VIENNENT RECHAUFFER SES AILES
A L APPROCHE DE L HIVER
C EST ALORS QUE MA PAUPIERE,VOUS VIT PALIR ET MOURIR
TENDRE FRUIT QU A LA LUMIERE,DIEU N A PAS LAISSE MURIR
QUOIQUE JEUNE SUR LA TERRE,JE SUIS DEJA SOLITAIRE,
PARMI CEUX DE MA SAISON,
ET QUAND JE DIS EN MOI MEME,OU SONT CEUX QUE TON COEUR AIME
JE REGARDE LE GAZON.
C EST UN AMI DE L ENFANCE ,OU AUX JOURS SOMBRES DU MALHEUR
NOUS PRETA LA PROVIDENCE ,POUR APPUYER NOTRE COEUR
IL N EST PLUS NOTRE AME EST VEUVE,IL NOUS SUIT DANS NOTRE EPREUVE
ET NOUS DIT AVEC PITIE
AMI SI TON AME EST PLEINE,DE TA JOIE OU DE TA PEINE,
QUI PORTERA LA MOITIE...
****
L AUTOMNE VA FINIR,AU MILIEU DU CIEL TERNE,
DANS UN CERCLE BLAFARD ET LIVIDE QUE CERNE
UN NUAGE PLOMBE,LE SOLEIL DORT.DU FOND
DES ETANGS REMPLIS D EAU MONTE UN BROUILLARD QUI FOND
COLLINES,CHAMPS,HAMEAUX DANS UNE MEME TEINTE.
SUR LES CARREAUX LA PLUIE EN LARGES GOUTTES TINTE.
LA FROIDE BISE SIFFLE,UN SOURD FREMISSEMENT
SORT DU SEIN DES PORTES.LES OISEAUX TRISTEMENT
MELANT LEUR CRIS PLAINTIFS AUX CRIS DES BETES FAUVES
SAUTENT DE BRANCHE EN BRANCHE A TRAVERS LES BOIS CHAUVES,
ET SEMBLENT AUX BEAUX JOURS ENVOLES DIRE ADIEU.
LE PAUVRE PAYSAN SE RECOMMANDE A DIEU,
CRAIGNANT UN HIVER RUDE,ET MOI,DANS LES VALLEES,
QUAND JE VOIS LE GAZON SOUS LES BRANCHES GELEES
DISPARAITRE ET MOURIR,JE REVIENS A PAS LENTS
M ASSEOIR LE COEUR NAVRE PRES DES TISONS BRULANTS,
ET LA JE ME SOUVIENS DU SOLEIL DE SEPTEMBRE
QUI DONNAIT A LA GRAPPE UN JAUNE REFLET D AMBRE,
DES POMMIERS DU CHEMIN PLIANT SOUS LEUR FARDEAU,
ET DU TREFLE FLEURI,PITTORESQUE RIDEAU
S ETENDANT A LONG PLIS SUR LA PLAINE RAYEE,
ET DE LA ROUTE ETROITE EN SON MILIEU FRAYEE,
ET SURTOUT DES BLEUFTS ET DES COQUELICOTS,
POINTS DE POURPRE ET D AZUR DANS L OR DES BLES EGAUX.
****
LES SANGLOTS LONGS
DES VIOLONS
DE L AUTOMNE
BLESSENT MON COEUR
D UNE LANGUEUR
MONOTONE.
TOUT SUFFOCANT
ET BLEME,QUAND
SONNE L HEURE,
JE ME SOUVIENS
DES JOURS ANCIENS
ET JE PLEURE.
ET JE M EN VAIS
AU VENT MAUVAIS
QUI M EMPORTE
DECA,DELA,
PAREIL A LA
FEUILLE MORTE.
****
```

L HERBE SE FANE DANS LES PRES,
LES JOURS DE SOLEIL SONT PASSES
LES FEUILLES JAUNES ET POURPREES
JONCHENT LES SENTIERS EFFACES.
LE VOILA DONC MON BEL AUTOMNE.
VOILA SON CHARME MONOTONE,
SES JOURS MORNES COMME DES NUITS
QUI LAISSENT VAGUEMENT A L AME
TELS QUE LES BAISERS D UNE FEMME,
LA VOLUPTE...LES LONGS ENNUIS.
LA BRUME AU COL DES MONTS SERPENTE,
PAREILLE AUX CRINS BLANCS D UN CHEVAL.
OH.LAISSEZ MOI SUIVRE MA PENTE,
COMME LA SOURCE AU FOND DU VAL
LAISSEZ MOI SEUL.LAISSEZ QUE J ERRE,
PAUVRE HIRONDELLE PASSAGERE
QUI REVIENT VISITER SON NID,
DE L,ETANG AUX DORMANTS RIVAGES,
JUSQU AU FRONT DES ROCHERS SAUVAGES,
OU LA TERRE AU CIEL BLEU S UNIT.
****
BIENTOT NOUS PLONGERONS DANS LES FROIDES TENEBRES
ADIEU,VIVE CLARTE DE NOS ETES TROP COURTS.
J ENTENDS DEJA TOMBER AVEC DES CHOCS FUNEBRES
LE BOIS RETENTISSANT SUR LE PAVE DES COURS.
TOUT L HIVER VA RENTER DANS MON ETRE.COLERE,
HAINE,FRISSON,HORREUR,LABEUR DUR ET FORCE,
ET COMME LE SOLEIL DANS SON ENFER POLAIRE
MON COEUR NE SERA PLUS QU UN BLOC ROUGE ET GLACE.
J ECOUTE EN FREMISSANT CHAQUE BUCHE QUI TOMBE,
L ECHAFAUD QU ON BATIT N A PAS D ECHO PLUS SOURD.
MON ESPRIT EST PAREIL A LA TOUR QUI SUCCOMBE,
SOUS LES COUPS DU BELIER INFATIGABLE ET LOURD.
IL ME SEMBLE,BERCE PAR CE CHOC MONOTONE,
QU ON CLOUE EN GRANDE HATE UN CERCUEIL QUELQUE PART
POUR QUI..C ETAIT HIER L ETE,VOICI L AUTOMNE
CE BRUIT MYSTERIEUX SONNE COMME UN DEPART.
=====

- ANNEXE III -

PRINCIPALES SORTIES CORRESPONDANT AU JEU D'ESSAI (pour contrôle)

TABLEAU DES EFFECTIFS

		1	2	3	4	5
1	LES	3.	5.	1.	6.	2.
2	QUI	6.	2.	1.	2.	3.
3	SUR	3.	2.	0.	0.	1.
4	EAU	1.	1.	0.	0.	0.
5	DES	5.	8.	2.	3.	2.
6	EST	6.	0.	0.	0.	1.
7	PAR	1.	0.	0.	0.	1.
8	QUE	3.	1.	0.	2.	0.
9	AUX	2.	2.	0.	2.	0.
10	SES	1.	0.	0.	1.	0.
11	PAS	1.	1.	0.	0.	1.
12	MOI	1.	1.	0.	2.	0.
13	TON	2.	0.	0.	0.	0.
14	AMI	2.	0.	0.	0.	0.
15	AME	2.	0.	0.	1.	0.
16	UNE	0.	1.	1.	1.	0.
17	SON	0.	1.	0.	2.	1.
18	MON	0.	0.	1.	1.	3.
19	VENT	2.	0.	1.	0.	0.
20	DANS	2.	4.	0.	1.	3.
21	BOIS	1.	1.	0.	0.	1.
22	TOUT	1.	0.	1.	0.	1.
23	SOUS	1.	2.	0.	0.	1.
24	DIEU	1.	1.	0.	0.	0.
25	DEJA	1.	0.	0.	0.	1.
26	CEUX	2.	0.	0.	0.	0.
27	MEME	1.	1.	0.	0.	0.
28	SONT	1.	0.	0.	1.	0.
29	NOUS	3.	0.	0.	0.	1.
30	POUR	1.	0.	0.	0.	1.
31	PLUS	1.	0.	0.	0.	2.
32	AVEC	1.	0.	0.	0.	1.
33	CIEL	0.	1.	0.	1.	0.
34	FOND	0.	2.	0.	1.	0.
35	LEUR	0.	2.	0.	0.	0.
36	CRIS	0.	2.	0.	0.	0.
37	PRES	0.	1.	0.	1.	0.
38	VOILA	4.	0.	0.	2.	0.
39	GAZON	2.	1.	0.	0.	0.
40	TOMBE	3.	0.	0.	0.	1.
41	COUPS	1.	0.	0.	0.	1.
42	ALORS	2.	0.	0.	0.	0.
43	COMME	1.	0.	0.	2.	2.
44	HIVER	1.	1.	0.	0.	1.
45	TERRE	1.	0.	0.	1.	0.
46	QUAND	1.	1.	1.	0.	0.
47	COEUR	2.	1.	1.	0.	1.
48	JOURS	1.	1.	1.	2.	0.
49	NOTRE	3.	0.	0.	0.	0.
50	SOURD	0.	1.	0.	0.	1.
51	ADIEU	0.	1.	0.	0.	1.
52	LONGS	0.	0.	1.	1.	0.
53	SONNE	0.	0.	1.	0.	1.
54	SAISON	2.	0.	0.	0.	0.
55	MOURIR	1.	1.	0.	0.	0.
56	MILIEU	0.	2.	0.	0.	0.
57	SOLEIL	0.	2.	0.	1.	1.
58	PAUVRE	0.	1.	0.	1.	0.
59	PAREIL	0.	0.	1.	0.	1.
60	TOMBENT	2.	0.	0.	0.	0.
61	AUTOMNE	0.	1.	1.	1.	1.
62	BRANCHE	0.	2.	0.	0.	0.
63	LAISSEZ	0.	0.	0.	3.	0.
64	FEUILLES	1.	0.	0.	1.	0.
65	SOUVIENS	0.	1.	1.	0.	0.
66	MONOTONE	0.	0.	1.	1.	1.
67	HIRONDELLE	1.	0.	0.	1.	0.

TABLEAU DE CONTINGENCE

(Formes classées par longueurs)

EFFECTIFS SELON LE NOMBRE DE LETTRE

EFFECTIFS OBSERVES	0	0	29	64	79	63	55	29	12	10	4	2	0	0	0
EFF.MAXIMAUX PREVUS	30	90	150	250	350	420	450	420	350	250	150	100	60	50	40

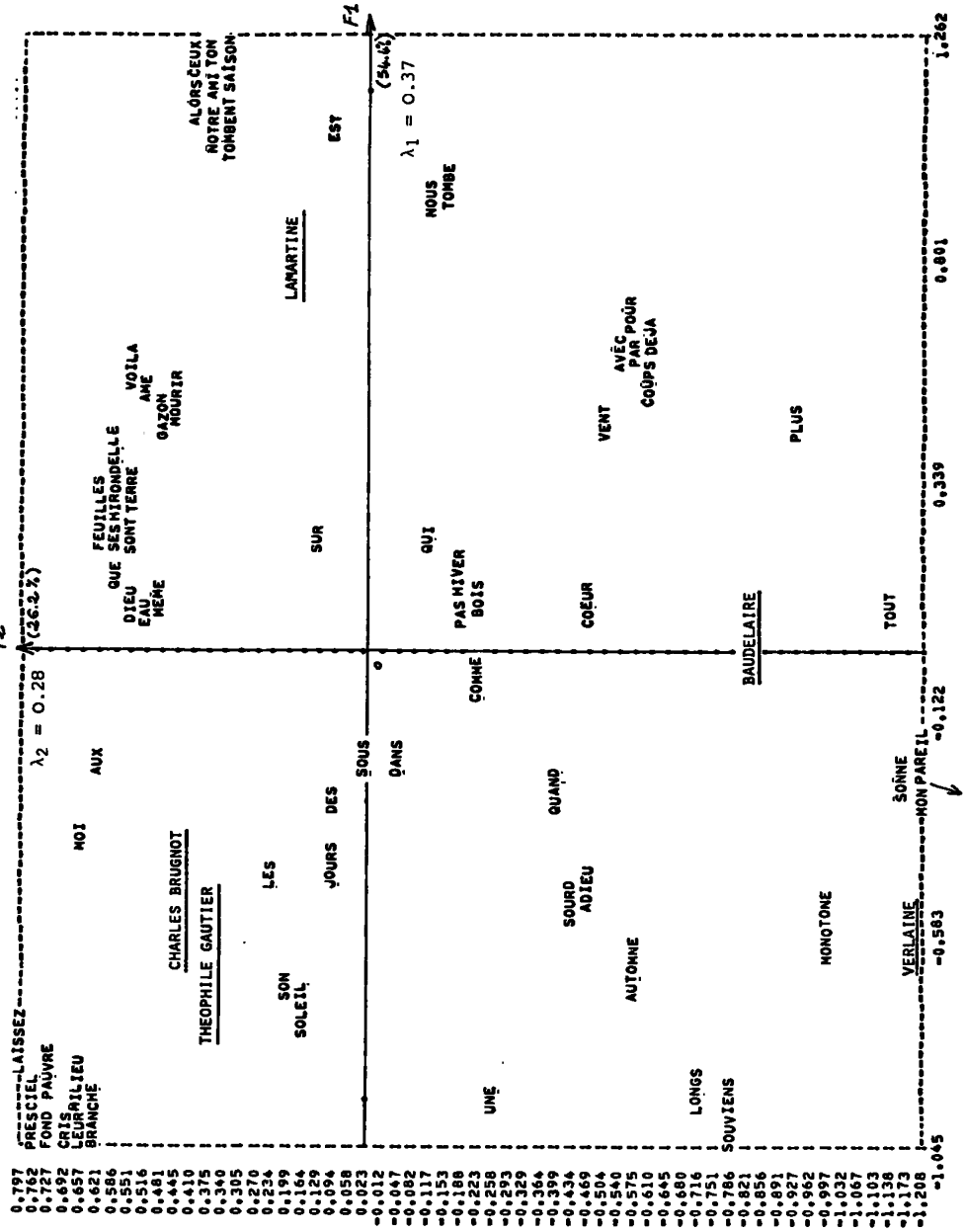
NOMBRE TOTAL DE MOTS= 528
NOMBRE DE MOTS DISTINCTS= 347
SEUIL DE FREQUENCE = 1
NOMBRE DE MOTS RETENUS= 67

FREQUENCES	POURCENTAGES	LIBELLES	1	2	3	4	5
20	3.79	DES	5.7	13.6	11.8	6.7	5.0
17	3.22	LES	3.4	8.5	5.9	13.3	5.0
14	2.65	QUI	6.9	3.4	5.9	4.4	7.5
10	1.89	DANS	2.3	6.8	0.0	2.2	7.5
7	1.33	EST	6.9	0.0	0.0	0.0	2.5
6	1.14	VOILA	4.6	0.0	0.0	4.4	0.0
6	1.14	QUE	3.4	1.7	0.0	4.4	0.0
6	1.14	AUX	2.3	3.4	0.0	4.4	0.0
6	1.14	SUR	3.4	3.4	0.0	0.0	2.5
5	0.95	JOURS	1.1	1.7	5.9	4.4	0.0
5	0.95	COMME	1.1	0.0	0.0	4.4	5.0
5	0.95	MON	0.0	0.0	5.9	2.2	7.5
5	0.95	COEUR	2.3	1.7	5.9	0.0	2.5
4	0.76	SOLEIL	0.0	3.4	0.0	2.2	2.5
4	0.76	NOUS	3.4	0.0	0.0	0.0	2.5
4	0.76	TOMBE	3.4	0.0	0.0	0.0	2.5
4	0.76	AUTOMNE	0.0	1.7	5.9	2.2	2.5
4	0.76	SON	0.0	1.7	0.0	4.4	2.5
4	0.76	MOI	1.1	1.7	0.0	4.4	0.0
4	0.76	SOUS	1.1	3.4	0.0	0.0	2.5
3	0.57	LAISSEZ	0.0	0.0	0.0	6.7	0.0
3	0.57	VENT	2.3	0.0	5.9	0.0	0.0
3	0.57	HIVER	1.1	1.7	0.0	0.0	2.5
3	0.57	BOIS	1.1	1.7	0.0	0.0	2.5
3	0.57	NOTRE	3.4	0.0	0.0	0.0	0.0
3	0.57	QUAND	1.1	1.7	5.9	0.0	0.0
3	0.57	FOND	0.0	3.4	0.0	2.2	0.0
3	0.57	GAZON	2.3	1.7	0.0	0.0	0.0
3	0.57	UNE	0.0	1.7	5.9	2.2	0.0
3	0.57	AME	2.3	0.0	0.0	2.2	0.0
3	0.57	PAS	1.1	1.7	0.0	0.0	2.5
3	0.57	PLUS	1.1	0.0	0.0	0.0	5.0
3	0.57	TOUT	1.1	0.0	5.9	0.0	2.5
3	0.57	MONOTONE	0.0	0.0	5.9	2.2	2.5
2	0.38	MILIEU	0.0	3.4	0.0	0.0	0.0
2	0.38	HIRONDELLE	1.1	0.0	0.0	2.2	0.0
2	0.38	LEUR	0.0	3.4	0.0	0.0	0.0
2	0.38	MOURIR	1.1	1.7	0.0	0.0	0.0
2	0.38	TOMBENT	2.3	0.0	0.0	0.0	0.0
2	0.38	PAUVRE	0.0	1.7	0.0	2.2	0.0
2	0.38	PAREIL	0.0	0.0	5.9	0.0	2.5
2	0.38	CEUX	2.3	0.0	0.0	0.0	0.0
2	0.38	SONT	1.1	0.0	0.0	2.2	0.0
2	0.38	COUPS	1.1	0.0	0.0	0.0	2.5
2	0.38	TERRE	1.1	0.0	0.0	2.2	0.0
2	0.38	SONNE	0.0	0.0	5.9	0.0	2.5
2	0.38	LONGS	0.0	0.0	5.9	2.2	0.0
2	0.38	ADIEU	0.0	1.7	0.0	0.0	2.5
2	0.38	SOURD	0.0	1.7	0.0	0.0	2.5
2	0.38	DEJA	1.1	0.0	0.0	0.0	2.5
2	0.38	AVEC	1.1	0.0	0.0	0.0	2.5
2	0.38	BRANCHE	0.0	3.4	0.0	0.0	0.0
2	0.38	POUR	1.1	0.0	0.0	0.0	2.5
2	0.38	AMI	2.3	0.0	0.0	0.0	0.0
2	0.38	TON	2.3	0.0	0.0	0.0	0.0
2	0.38	MEME	1.1	1.7	0.0	0.0	0.0
2	0.38	ALORS	2.3	0.0	0.0	0.0	0.0
2	0.38	SES	1.1	0.0	0.0	2.2	0.0
2	0.38	DIEU	1.1	1.7	0.0	0.0	0.0
2	0.38	SAISON	2.3	0.0	0.0	0.0	0.0
2	0.38	PAR	1.1	0.0	0.0	0.0	2.5
2	0.38	PRES	0.0	1.7	0.0	2.2	0.0
2	0.38	CRIS	0.0	3.4	0.0	0.0	0.0
2	0.38	EAU	1.1	1.7	0.0	0.0	0.0
2	0.38	CIEL	0.0	1.7	0.0	2.2	0.0
2	0.38	SOUVIENS	0.0	1.7	5.9	0.0	0.0
2	0.38	FEUILLES	1.1	0.0	0.0	2.2	0.0

TABLEAU DES PROFILS LEXICAUX

(formes classées par fréquences totales)

FIGURE I


F2
(36.2 %)
λ2 = 0.28
F1
(54.6 %)
λ1 = 0.37
LAISSEZ
PRESCIEL
FOND PAUVRE
CRIS
LEUR AILLIEU
BRANCHE
MOI
AUX
FEUILLES
QUE SES HIRONDELLE
SONT TERRE
DIEU
EAU
MÊME
VOILA
AME
GAZON
MOURIR
ALORS CEUX
NOTRE AMI TON
TOMBENT SAISON
CHARLES BRUGNOT
THEOPHILE GAUTIER
LES
SON
SOLEIL
JOURS
DES
SOUS
LAMARTINE
SUR
EST
NOUS
TOMBE
QUI
PAS HIVER
BOIS
DANS
COMME
QUAND
SOURD
ADIEU
AUTOMNE
UNE
LONGS
SOUVIENS
COEUR
VENT
AVEC
PAR POUR
COUPS DEJA
BAUDELAIRE
PLUS
TOUT
VERLAINE
MONOTONE
SONNE
MON PAREIL
0.797
0.762
0.727
0.692
0.657
0.621
0.586
0.551
0.516
0.481
0.445
0.410
0.375
0.340
0.305
0.270
0.234
0.199
0.164
0.129
0.094
0.058
0.023
-0.012
-0.047
-0.082
-0.117
-0.153
-0.188
-0.223
-0.258
-0.293
-0.329
-0.364
-0.399
-0.434
-0.469
-0.504
-0.540
-0.575
-0.610
-0.645
-0.680
-0.716
-0.751
-0.786
-0.821
-0.856
-0.891
-0.927
-0.962
-0.997
-1.032
-1.067
-1.103
-1.138
-1.173
-1.208
-1.245
-0.583
-0.122
0.339
0.801
1.262