

G. VOLOCHINE

A. VOLOCHINE

**Étude comparée de textes russes : le Don Tranquille
et d'autres œuvres de M. A. Cholokhov ; les
nouvelles de F. D. Kriukov**

Les cahiers de l'analyse des données, tome 20, n° 1 (1995),
p. 7-26

http://www.numdam.org/item?id=CAD_1995__20_1_7_0

© Les cahiers de l'analyse des données, Dunod, 1995, 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/>

**ÉTUDE COMPARÉE DE TEXTES RUSSES:
LE DON TRANQUILLE ET D'AUTRES ŒUVRES
DE M. A. CHOLOKHOV;
LES NOUVELLES DE F. D. KRIUKOV**

[COMPAR. RUSSE]

*G. et A. VOLOCHINE**

Objet de l'étude

À la lecture du très volumineux ouvrage de M.A. CHOLOKHOV "Le Don paisible" (ТИХИЙ ДОН), notre attention a été attirée par la grande disparité de style et de richesse de vocabulaire d'un tome à un autre, d'un chapitre à un autre et parfois à l'intérieur d'un même chapitre (cf. Appendice).

Il était donc intéressant d'essayer de trouver par une méthode rigoureuse, exempte de toute subjectivité, les raisons possibles du fait constaté. La présente étude, s'appuyant sur l'analyse de correspondance, tente d'apporter quelque lumière sur ce problème.

1 Le corpus des textes traités

Notre souhait étant d'éliminer tout paramètre qui pourrait venir perturber les résultats d'analyse, les textes mentionnés ont été choisis pour leur homogénéité.

Ce sont tous des récits ou des nouvelles. Ils se rapportent à la période s'étendant de la grande guerre à la dernière guerre en passant par la révolution et la collectivisation des terres. Ils concernent, en général, des faits de guerre ou d'armes.

Tous mettent en scène le milieu cosaque du Don.

L'étude porte d'abord sur l'intégralité de l'œuvre "Le Don paisible" - entendez: le fleuve du Don - (ТИХИЙ ДОН). Dans le but d'établir si la disparité observée était particulière à cette œuvre, nous avons également inclus plusieurs autres œuvres de l'auteur: les deux tomes des "Terres défrichées" (ПОДНЯТАЯ ЦЕЛИНА), l'ouvrage "Ils ont combattu pour la patrie" (ОНИ

(*) G. Volochine: Agrégée de Russe ; A. Volochine: Ingénieur.

сражались за Родину); et une œuvre à part qu'est "Le Destin d'un Homme" (Судьба человека). Enfin, pour mettre éventuellement en valeur une différence d'écriture d'auteur, nous avons inclus cinq nouvelles de F.D. KRIUKOV. Le corpus global représente environ 747000 mots.

1.1 "Le Don paisible"

Cet ouvrage comprend 4 livres subdivisés en 8 parties, chacune d'entre-elle étant également subdivisée en chapitres (Édition de 1945). Il a été écrit en plusieurs étapes entre 1928 et 1940, avec de la part de l'auteur une rapidité d'écriture très variable. Le sujet en est une épopée de plusieurs familles cosaques du Don, couvrant la fin de la grande guerre (1914-1918), la révolution d'Octobre, la révolte des paysans-cosaques, leur défaite et l'instauration de l'ordre populaire (collectivisation).

Livre I : 1928 (3 parties soit 22% de l'œuvre) :

- 1-ère partie: 23 chapitres, ≈ 22000 mots ;
- 2-ème partie: 21 chapitres, ≈ 28000 mots ;
- 3-ème partie: 24 chapitres, ≈ 38000 mots

Livre II : 1929 (2 parties soit 22% de l'œuvre) :

- 4-ème partie: 21 chapitres, ≈ 40000 mots ;
- 5-ème partie: 31 chapitres, ≈ 47000 mots ;

Livre III: 1933 (1 partie soit 25% de l'œuvre) :

- 6-ème partie: 65 chapitres, ≈ 98000 mots ;

Livre IV: 1940 (2 parties soit 30% de l'œuvre)

- 7-ème partie: 29 chapitres, ≈ 73500 mots ;
- 8-ème partie: 18 chapitres, ≈ 47000 mots ;

Soit un total de quelque 393500 mots.

1.2 "Les Terres défrichées"

L'ouvrage comprend deux tomes subdivisés en 40 et 29 chapitres respectivement. Ils ont été écrits l'un en 1932 et l'autre en 1959. Le sujet en est la collectivisation et le développement des terres vierges dans les régions du Don et concerne toujours les mêmes populations.

Tome I : 40 chapitres, ≈ 97500 mots

Tome II : 29 chapitres, ≈ 107000 mots

Soit un total de ≈ 214500 mots

1.3 "Ils ont combattu pour la Patrie"

Écrit entre 1956 et 1957 cet ouvrage couvre, parmi les mêmes

populations, les troubles et les faits d'armes durant la dernière guerre.

26 chapitres, ≈ 60400 mots.

1.4 “Le Destin d'un Homme”

Ce récit est très particulier quant à la forme: il est écrit à la première personne et concerne des événements de la dernière guerre. Il a été écrit en 1959 et revu en 1967. Le récit compte quelque 10600 mots

1.5 Les Nouvelles de F.D. KRIUKOV

Ces cinq nouvelles ont été publiées entre 1903 et 1913; sauf “Kazatchka” (Казачка) publiée en 1896. Sans se rapporter à des faits d'armes, elles concernent les us et coutumes des cosaques du Don à l'époque préévolutionnaire: l'auteur était lui-même cosaque. Au total, ces nouvelles comprennent ≈ 68000 mots.

2 Préparation des données et enchaînement des analyses

2.1 Principe de l'étude

Quelques études préliminaires, faites sur la 1ère partie du “Don paisible”, des chapitres du 1er tome des “Terres défrichées” et trois nouvelles de F.D. КРИУКОВ, ont montré l'influence notoire des mots thématiques sur les résultats d'analyse. Pour éviter l'intrusion de ces effets dans les résultats de la présente étude, il a été décidé d'éliminer tous les mots qui, d'une façon ou d'une autre, pouvaient avoir une teinture thématique.

En résumé ont été éliminés:

Tous les noms propres;

Tous les substantifs;

Les adjectifs qualificatifs ou d'appartenance;

Les pronoms personnels;

Les verbes, à l'exception de deux d'entre eux, (БЫТЬ: être) et (СТАНОВИТЬСЯ: devenir), conservés sous toutes leurs formes graphiques;

Les numéraux sauf (ОДИН: un) très particulier en langue russe (comme dans bien d'autres langues) ;

Les catégories conservées sont:

Les prépositions

Les adverbes et formes adverbiales des adjectifs

Les conjonctions

Les pronoms démonstratifs sous toutes leurs formes

Les adjectifs possessifs sous toutes leurs formes

L'étude préliminaire a montré que dans l'ordre décroissant les mots les

plus fréquents se trouvent en très grande majorité être ceux des catégories conservées (pour plus de précision, cf. *infra*, §2.3).

Pour simplifier les références nous utiliserons par la suite les symboles suivants:

- D1 à D8 pour les 8 parties du “Don paisible”;
- Ts1 et Ts2 pour les deux tomes de “Terres défrichées”;
- R pour “Ils ont combattu pour la Patrie”;
- S pour “Le Destin d'un Homme”;
- K1 à K5 pour les cinq nouvelles de F.D. KRIUKOV;

Nous avons pris la partie comme unité d'analyse quelle que soit sa longueur. “Le Destin d'un Homme” est pris comme un tout. Enfin dans une partie préliminaire chaque nouvelle de F.D. KRIUKOV a été prise comme unité; pour la suite, toutes ces nouvelles ont été groupées en un seul texte, désigné par K§.

2.2 Codage et enregistrement

Tous les programmes d'analyse ont été écrits en QBASIC de Microsoft.

Comme ce langage n'accepte pas les caractères cyrilliques il a fallu d'abord définir une norme de transcription des textes, en caractères du tableau ASCII étendu. Pour éviter au maximum les causes d'erreur à l'enregistrement, cette norme a été conçue de façon à serrer au plus près les réactions naturelles.

Ainsi le (x, kh guttural) a été transcrit en ‘x’ latin, le (г) en ‘g’, le (в, prononcé v en russe) en ‘v’; ... etc.

Exception faite de la 1ère partie du Don, de 5 chapitres de Tselina I, et de 3 nouvelles de F.D. KRIUKOV enregistrés intégralement, tous les autres enregistrements ne concernent que les mots outils dont les catégories ont été définies plus haut.

Une des difficultés rencontrées concerne l'appartenance d'un mot à telle ou telle catégorie. Tandis que “dolgo” (ДОЛГО : longtemps) ne pose pas de difficulté, un mot comme “tixo” (ТИХО) peut avoir deux sens: “silencieux”, et c'est un attribut; ou bien “sans bruit”, et c'est un adverbe; dans ce dernier cas on doit l'enregistrer et non dans le premier. Il est donc nécessaire d'être vigilant à l'enregistrement et de tenir compte du texte.

2.3 Procédure de choix des formes de mots

Pour les mots déclinables, (ainsi que pour les rares formes verbales conservées) on a distingué les formes par cas, genre et nombre: e.g. (какой: quel), (какая: quelle), (какие: quels). Il s'agit donc, à strictement parler, non de mots, mais de formes de mots; mais on dira, en bref: ‘mot’.

Après avoir déterminé les occurrences de tous les mots outil enregistrés par unité retenue, on a recherché par programme tous les mots outil (les formes) communs à tous les textes sauf "Le Destin d'un Homme" (Судьба), trop court en comparaison avec les autres textes. À l'aide de cette liste et toujours par programme on a calculé, par texte, les occurrences des mots de la liste. De la sorte le tableau des correspondances ne contient pas d'occurrence nulle sauf pour certains mots dans (Судьба).

La procédure décrite ici a abouti à une liste de 224 formes de mots outil existant au moins une fois dans chaque texte retenu sauf dans (Судьба) dont l'étendue est faible. Le nombre total de ces occurrences est de 236283 soit pour un corpus de 747000 mots (cf. *supra*, §1), une couverture d'environ 32%.

2.4 Remarque: poids des mots outil dans les textes

Le dénombrement des mots outil est précis. Celui de l'ensemble des mots des textes a été fait de la façon suivante:

Pour chaque édition, a été déterminé par échantillonnage le nombre moyen des mots par ligne complète ainsi que sa variance.

Par texte, on a dénombré le nombre de lignes complètes. Le nombre de mots s'en déduit.

Les résultats de ces dénombrements sont mentionnés au §1. On peut déjà noter que pour le "Don paisible" le poids des mots outil passe de 32,4% dans les premières parties à 43,6% pour la 8-ème partie. Par contre, pour les autres œuvres de M.A. CHOLOKHOV ce paramètre oscille entre 42,4% (R) et 44,7% (S) exception faite pour Ts1(36,3%). Au contraire chez F.D. KRIUKOV ce nombre varie peu (de 35 à 38,8%) et reste faible.

2.5 Enchaînement de l'exposé

Aux §§3 et 4, on présente les résultats d'analyses factorielles et de classifications, en considérant exclusivement l'ensemble des textes. Dans la suite, on considère les contributions des mots et l'on précise le détail des subdivisions en chapitres de certaines parties d'œuvres de M.A. CHOLOKHOV.

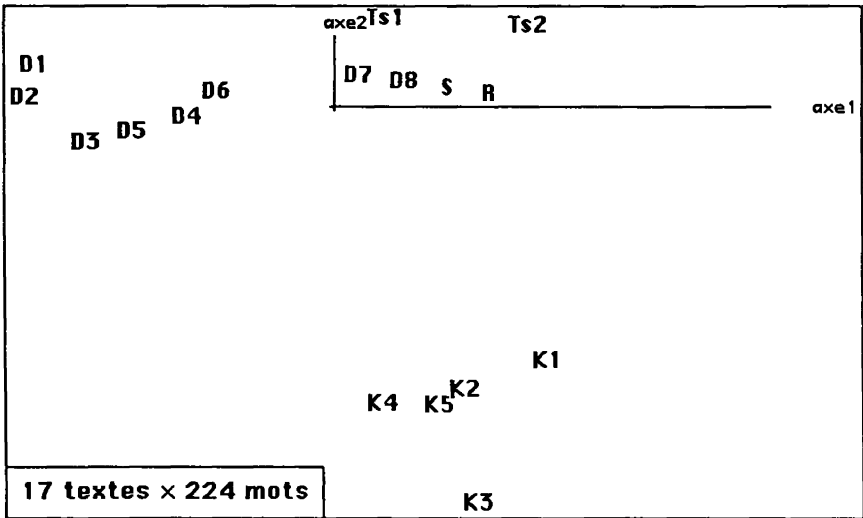
3 Analyses de correspondances

3.1 Croisement de 17 textes avec 224 mots

224 formes x 17 textes

trace :	7.066e-2									
rang :	1	2	3	4	5	6	7	8	9	10
lambda :	230	123	58	43	37	32	28	25	25	22 e-4
taux :	3261	1740	819	614	529	448	397	354	352	310 e-4
cumul :	3261	5001	5820	6434	6963	7411	7808	8162	8514	8824 e-4

Cette première analyse est instructive à plusieurs titres. Le premier axe concentre 32,6% de l'inertie. Le 2-ème avec 17,4% ne représente qu'à peine un peu plus que la moitié du 1-er. Le 3-ème axe avec 8,2% ne vaut pas la moitié du 2ème: ce fait se traduit d'ailleurs sur le nuage des textes où on trouve pratiquement tous les textes voisins du 1-er axe sauf les nouvelles de F.D. KRIUKOV.



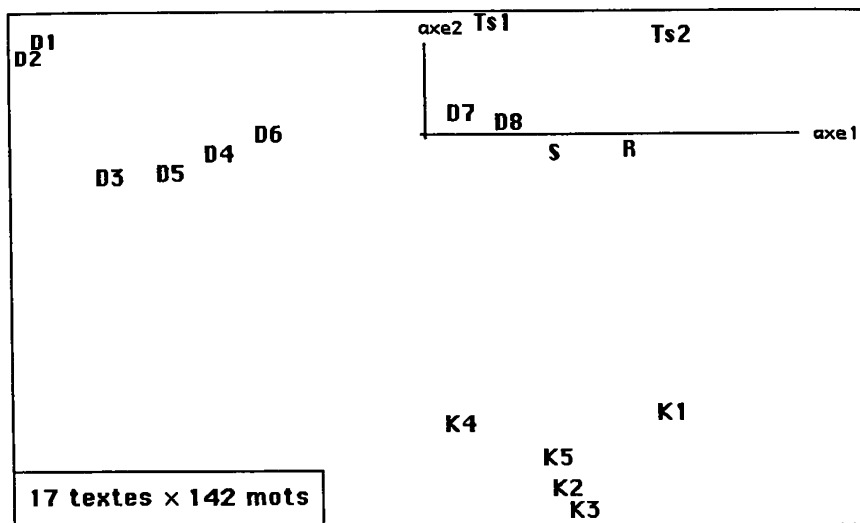
L'examen du graphique montre que les 8 parties du Don s'étalent sur un arc allant des valeurs négatives du 1er axe aux valeurs positives, ce qui est bien curieux, alors que les autres textes de M.A. CHOLOKHOV restent concentrées aux valeurs positives du 1er axe. Pratiquement les 6 premières parties du Don sont sur le demi-axe ($F1 < 0$) alors que D7 et D8 viennent rejoindre les autres œuvres de M.A. CHOLOKHOV à droite de l'axe vertical dans le 1-er quadrant. De même les 5 nouvelles de F.D. KRIUKOV sont bien groupées vers la partie négative du 2-ème axe. Nous aurons une autre remarque à ajouter, au §6, après l'examen de la correspondance avec les mots.

A ce point l'examen du tableau des occurrences nous montre qu'au delà d'une certaine limite on trouve de plus en plus d'occurrences faibles par texte. Pour éviter de se trouver confronté à des effets de bruit aléatoire nous avons décidé de limiter le tableau à un nombre de mots dont la somme des occurrences pour l'ensemble des textes considérés dépasse 235, soit un peu plus de 1% du montant de la fréquence maximale (mot "i", $\mu = \text{et}$). Le tableau est ainsi limité à 142 mots. Par la suite, ayant constaté que les 5 nouvelles de F.D. KRIUKOV restent bien groupées, on a cumulé ces cinq textes en un seul, totalisant l'ensemble de leurs occurrences.

3.2 Croisement de 17 textes avec 142 mots

142 formes x 17 textes	:	trace	:	5.594e-2							
rang	:	1	2	3	4	5	6	7	8	9	10
lambda	:	215	80	50	36	27	24	20	20	18	16 e-4
taux	:	3852	1425	890	641	478	435	363	355	318	283 e-4
cumul	:	3852	5277	6167	6809	7287	7722	8085	8441	8758	9042 e-4

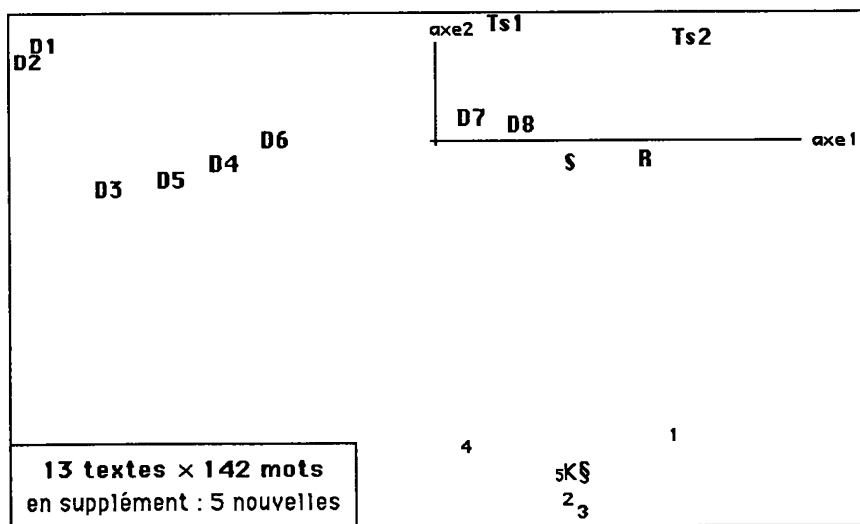
Cette analyse confirme la précédente sans modification notable. Le 1-er axe avec 38,5% de part d'inertie a pris encore plus d'importance au détriment



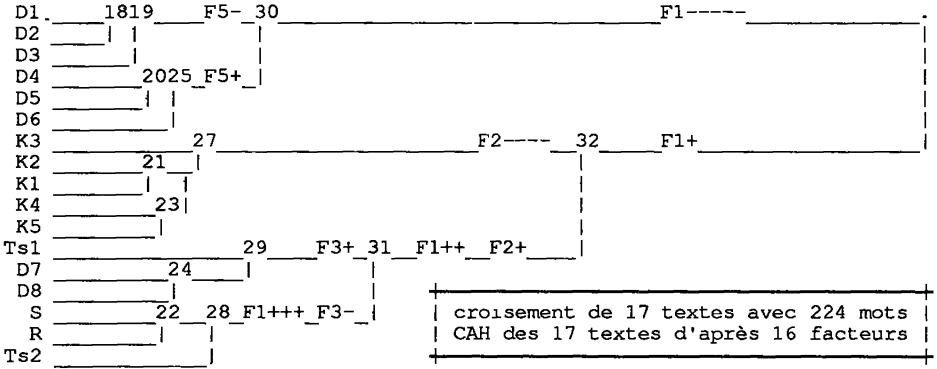
du 2-ème axe qui passe à 14,3%, les deux premiers axes totalisant 52,8% de l'inertie. Les groupements des textes en trois amas restent les mêmes.

3.3 Croisement de 13 textes avec 142 mots

Mêmes conclusions que dans le cas précédent. Tout au plus D1 et D2 se sont-ils beaucoup rapprochés. Le 1-er axe représente à lui seul 43,1% de l'inertie et avec le 2-ème axe à 16% l'ensemble traduit 59% de l'inertie, le 3ème axe étant proche de l'aléa. Dans la suite, au §6, on précisera le rôle des mots d'après cette configuration, soit 13 textes croisés avec 142 mots.



4 Classification Ascendante Hiérarchique



4.1 CAH des 17 textes d'après le croisement avec 224 mots

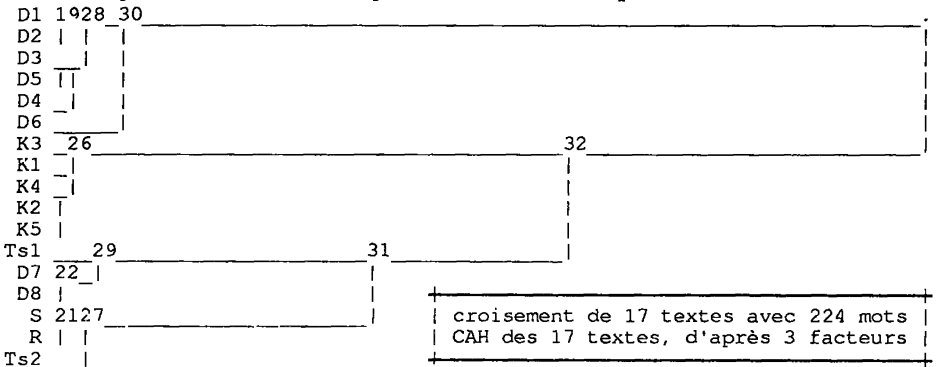
4.1.1 CAH dans l'espace engendré par les 16 axes

Prenant en compte la totalité des facteurs (16 facteurs) le graphe montre d'un coté une branche 30, où sont groupées les 6 premières parties du "Don" (D1 à D6); et, de l'autre, la branche 32, divisée entre 31, où est le reste des œuvres de CHOLOKHOV, et 27, composée des 5 nouvelles de F.D. KRIUKOV.

Le premier groupe se subdivise en deux sous-groupes: {D1, D2, D3} et {D4, D5, D6}. De même dans le 2-ème groupe on distingue d'une part {D7, D8, Ts1} et d'autre part {Ts2, R, S}. Il est bizarre de constater que dans une œuvre comme le "Don", traitant d'un seul sujet, dans un environnement stable 70% du texte se sépare complètement du reste de l'œuvre et des autres textes du même auteur; d'autant plus que la séparation s'observe au plus haut niveau. Ce schéma général ne fait d'ailleurs que confirmer l'image du plan (1, 2).

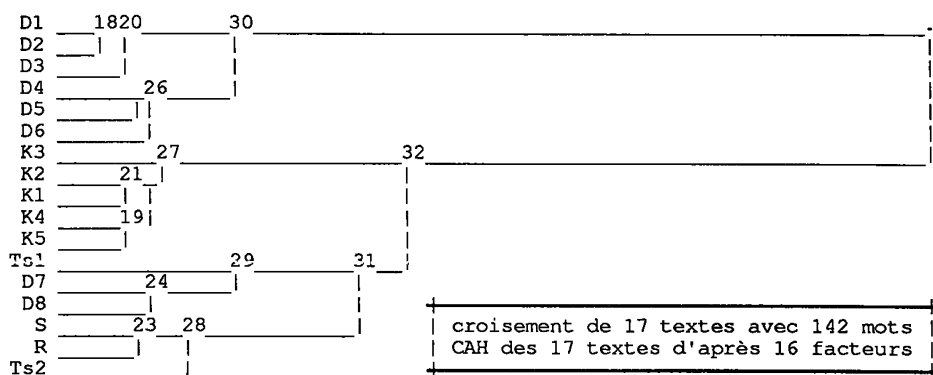
4.1.2 CAH dans l'espace engendré par les axes 1 à 3

Compte tenu de la remarque antérieure sur le poids des axes le calcul de

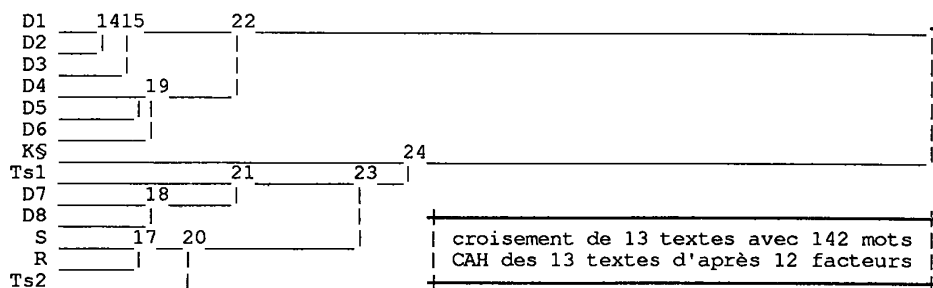


la CAH a été refait sur 3 facteurs. La partition en 4 classes définie par les 3 nœuds les plus hauts est la même: en effet, comme l'indique l'étiquetage de l'arbre, les dichotomies sont créées par les facteurs 1 à 3. Nous retrouvons ainsi les mêmes conclusions avec toutefois une variante: l'observation de la grande densité des 4 classes, 30, 29 27 et 26, dont les textes s'agrègent à un niveau très bas (les écarts dont rendent compte les axes 4 à 16 étant éliminés dans la présente CAH; notamment les contributions des axes 4 et 5 que signalent l'étiquetage); particulièrement {D1, D2}, {D7, D8}, {R, S} {K1 à K5} ont une grande proximité dans l'espace engendré par les axes 1 à 3. On notera l'éloignement relatif de D6 d'avec les 5 premières parties du "Don".

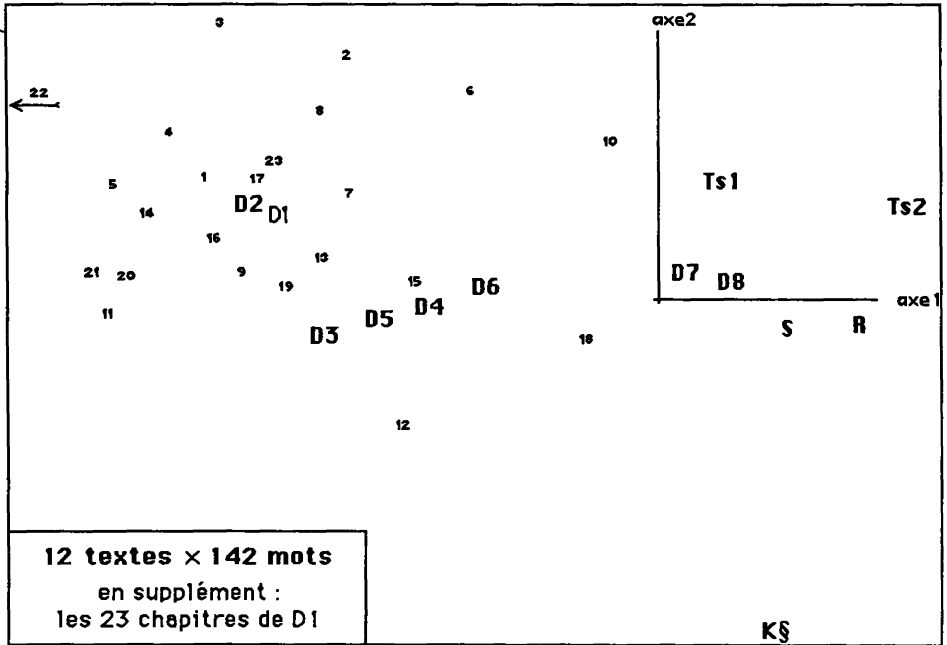
4.2 CAH des textes d'après le croisement avec 142 mots



Que les nouvelles {K1 à K5} soient cumulées ou non, en prenant en compte l'ensemble des facteurs, ou en n'en conservant que 3, nous retrouvons les mêmes caractéristiques: un premier sous-groupe des œuvres de CHOLOKHOV, D1 à D6; un 2-ème avec {D7, D8, Ts1}; un 3-ème avec {Ts2, R, S}; et enfin, à part, KRIUKOV.



On n'a pas précisé d'étiquetage, parce que les contributions des axes aux dichotomies successives sont les mêmes qu'au §4.1 (avec 224 formes de mots).



5 Analyses prenant en compte le détail des chapitres

Les diverses constatations faites plus haut ne permettent pas d'expliquer clairement la disparité de style et de vocabulaire mentionnée au début. Pourquoi chez un auteur constate-t-on une telle variation d'utilisation de certains mots outil (comme par exemple les prépositions de lieu, cf. *infra*) et ce entre une partie et une autre d'un seul ouvrage, et non dans trois autres productions du même auteur? Pourquoi le nuage des six premières parties du "Don" ne se rattache-t-il absolument pas au reste de son œuvre?

Nous avons eu la curiosité de reprendre les analyses faites plus haut mais en éclatant successivement chaque partie en ses chapitres. Comme les occurrences des mots y sont forcément très faibles l'analyse a consisté à prendre en principal tous les textes sans la partie considérée et en adjoignant les chapitres de celle-ci comme des variables supplémentaires.

[On aurait pu, de façon équivalente, mettre en principal les 13 textes, ou parties, avec, en supplément, les chapitres d'un seul: en procédant ainsi, on aurait gardé le même fond de plan (1, 2) qu'au §3.3; mais on n'aurait pas multiplié les vues du nuage des parties].

Ensuite, la constitution des nuages a été vérifiée par des C.A.H., dont les résultats sont cités ici brièvement, mais non publiés.

On a procédé à la même analyse pour toutes les parties du "Don" ainsi que pour les deux tomes {Ts1, Ts2} des "Terres" et pour "Patrie" (R). On a trouvé, pour D1 et D2, un nuage des chapitres concentré autour du bloc {D1...D6}. Puis pour D4 à D7 le nuage s'étale de plus en plus et migre vers le pôle opposé de l'axe 1. Pour D8, T1, T2 et R, il se concentre de nouveau, mais autour de ces dernières œuvres. À noter que l'ensemble des nouvelles de F.D. KRIUKOV (sigle K§) reste toujours à part et bien séparé de toutes les autres œuvres, vers (F2<0).

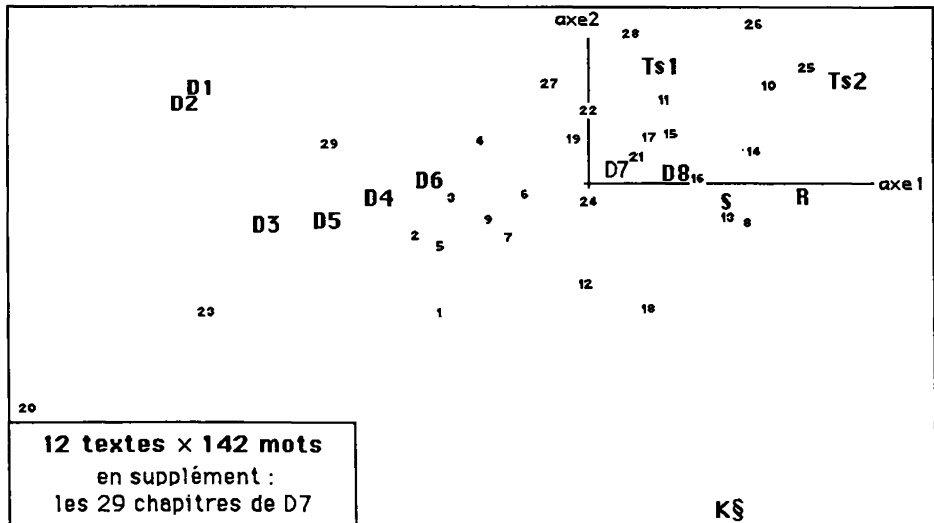
Le graphique du plan (1, 2) est présenté pour trois cas: D1, D7, Ts2.

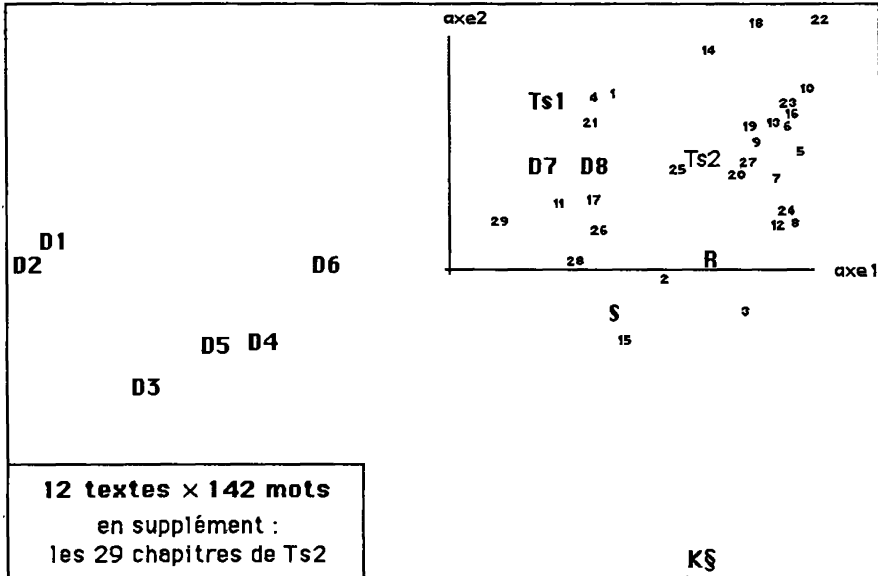
5.1 La partie D1 du "Don" est éclatée en ses 23 chapitres

On constate que la nébuleuse des chapitres de D1 englobe {D2...D6}. La C.A.H. (non publiée) montre une très forte densité de ce nuage et également sa séparation totale du reste de l'œuvre, sauf le chapitre 10. Il faut prendre en compte dans cet examen la taille du chapitre considéré et ne pas donner trop de signification à tel ou tel écart si le nombre de mots outil est notablement trop faible. Le chapitre 10 contient 756 occurrences de mots outil, à comparer à 1559 pour le plus fourni (chapitre 16) et à 101 pour le plus pauvre (ch. 22; lequel, vraisemblablement pour cette raison, a un profil excentrique: F1<<0).

5.2 La partie D7 du "Don" est éclatée en ses 29 chapitres

Là il y a un changement complet: l'ensemble des chapitres de D7 se répartissent sur toute l'étendue du graphe le long du premier axe. La C.A.H. (non publiée) montre que 9 chapitres se rattachent au groupe {D1...D6}; les autres allant avec le reste des textes de CHOLOKHOV.





5.3 Le tome Ts2 des "Terres défrichées est éclaté en 29 chapitres

De nouveau l'aspect du graphe est différent: l'ensemble des chapitres se concentre au voisinage de D7, D8, T, R, et S, ce qui est illustré par la C.A.H. où on trouve le bloc {D1...D6} complètement séparé du reste des textes.

6 Textes et mots

6.1 Approche globale: examen du plan (1, 2)

A ce stade il serait intéressant de représenter sur un même graphique les 13 textes et les 142 mots. Une première tentative nous a produit un amas inextricable de points voisins de l'origine et dû à beaucoup de mots à contribution faible. Nous avons préféré figurer les textes avec les 30 mots apportant aux axes 1 et 2 les plus fortes contributions.

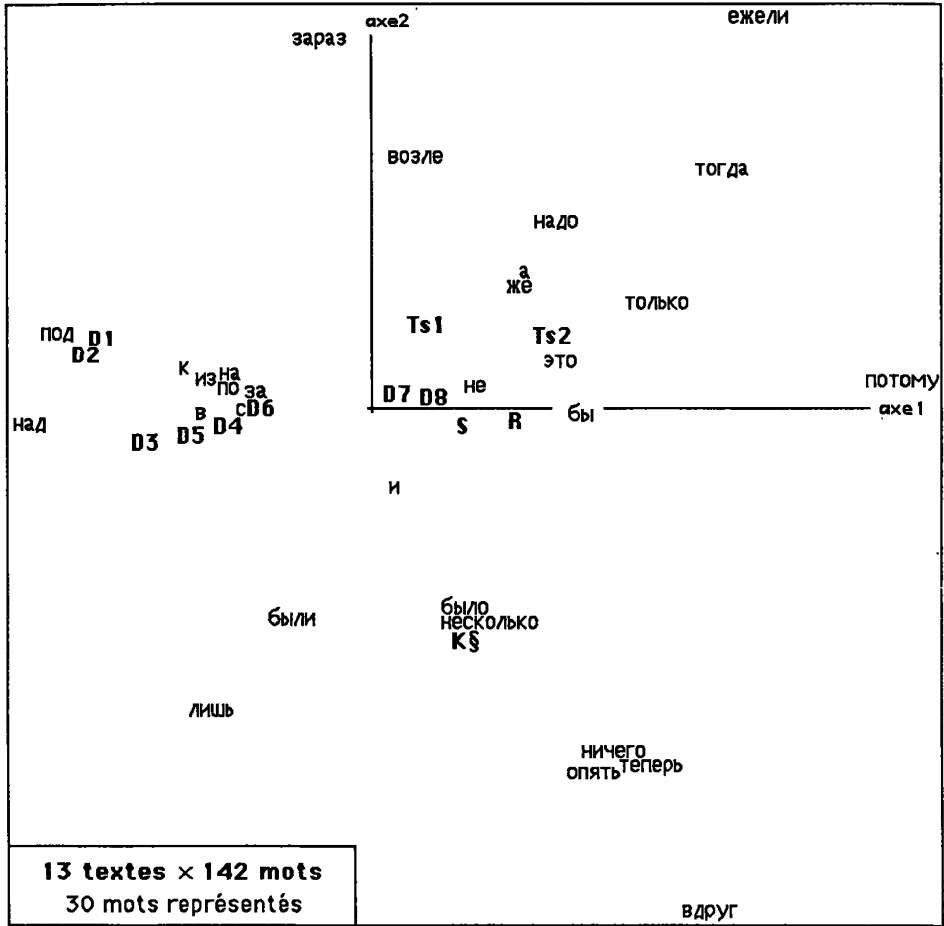
Une particularité attire notre attention: du côté du 1-er axe négatif on trouve {D1...D6} dans une nébuleuse avec 9 mots:

В на с к по за из под над

ce sont tous des prépositions se rapportant principalement à des désignations de lieu. Du côté de la partie positive du 1-er axe on trouve:

не а же это было только тогда потому

(не: négation) (а: mais) (это: démonstratif) (только: seulement) (тогда: alors) (потому: parce): rien de particulier à noter: toutes les catégories de mots outil sont représentées.



Sur la partie négative du 2-ème axe on a :

было теперь были ничего лишь опять вдруг

(вдруг: soudain) (опять: de nouveau) (теперь: maintenant) (лишь: seulement) (было были: formes passées de 'être'); prédominance d'adverbes, qui évoque les nouvelles de K.

Tandis que sur la partie positive du 2-ème axe on relève :

а же надо возле тогда зараз ежели

(зараз: à la fois) (ежели: 'si' conditionnel) (возле: auprès de) (тогда: alors) (надо: il faut)...; mots qui peuvent être les outils d'un discours et non seulement d'un récit.

c	CAH des 142 mots d'après leur correspondance avec 13 textes
268	и одна почти весь ли своего видно была что-то много все было
263	чуть словно назад о лишь вместо своих был один всех около долго были
245	ведь сейчас если
230	опять ничего теперь вдруг
266	того ещё со до там куда как у что том при эту будет тем для то
261	уж чего этом молча об
70	зараз
246	потом тоже стал всё снова
262	как-то свою какая вот даже всю но этой когда этого несколько тихо быть или ни чем без чём
264	сколько бы это нет да ну есть можно какой так чтоб не
267	мой очень сразу потому этот себя тут уже только больше такой пока про
269	ежели надо хоть тогда будем возле какие же буду такое почему сам а
240	над из-под рядом под через тот
250	перед своей то' кто против другой за
265	одной с из к после себе будто где во от по на мимо в

6.2 Correspondance entre classes de mots et classes de textes

Les remarques précédentes sont faites en considérant le graphe des deux premiers axes. Ce point de vue peut être critiquable car dans l'espace les points des textes et des mots peuvent révéler des proximités différentes. Aussi nous avons voulu établir un tableau d'attractions/répulsions de tous ces points dans l'espace.

L'ensemble I des 142 mots a été soumis à la CAH, et l'on a retenu une partition en un ensemble I_q de 15 classes, dont le graphique arborescent donne l'explication en termes de facteurs: on lit, par exemple, que la classe i230:

опять ничего теперь вдруг

déjà remarquée au § précédent, est caractérisée par une valeur très fortement négative de F₂; tandis que la classe i245 (agrégée à i230 pour former i273):

ведь сейчас если

tout en ayant (F₂<0), a aussi pour caractéristique (F₃<<0) et (F₅<0).

correspondance entre I_q = classes de mots et J_q = classes de textes

	i268	i263	i245	i230	i266	i261	i70	i246	i262	i264	i267	i269	i240	i250	i265
j15	3112	587	79	111	2585	310	40	225	956	2259	562	1216	725	926	8504
j19	7369	2178	312	302	6320	600	137	812	2813	5685	1556	3384	1335	1865	17526
j13	3975	829	162	417	2294	277	6	435	1585	2922	761	1120	260	813	5181
j9	3701	819	177	196	3657	447	150	500	1716	4054	860	2666	586	1092	7555
j18	5655	1330	81	264	4447	804	139	969	2295	5664	1353	2918	598	1232	10736
j17	3594	677	231	247	3128	250	0	609	1615	3366	1197	2066	324	651	5852
j10	5083	893	235	305	4676	447	76	683	2741	5428	1907	3829	485	1053	8907

268	_____	277	279	_____	F2--	282	
263	F5+	_____		_____			282: F1+ (263 et 70 exceptés)
245	F5+	273	_____				245: F3---
230	_____						230: F2----
266	_____	280	281	_____			266 ≈ CdG
261	_____	271	_____				261: F3++
70	_____						70: F1- F2+++ F3+++
246	_____	274					246: F3+ F4+
262	_____	270					
264	_____						
267	F3-	276	_____				
269	F2++						
240	_____	275	_____				F1---
250	272				240: F1-----	250: F3---	
265	_____				265: F1-----		

CAH des 142 mots d'après leur correspondance avec 13 textes

Sans prétendre publier le tableau des données lui-même, on peut en donner un aperçu par le tableau de cumul croisant une partition Jq des 13 textes avec la partition Iq des 142 mots. On lit, e.g., dans ce tableau, que la classe j13, réduite au seul texte de F.D. KRIUKOV, comprend, au total, 417 occurrences des quatre mots de la classe i230.

Ce nombre est le plus élevé de la colonne i230, mais il est encore plus remarquable si on considère son poids relatif dans le profil de la ligne j13.

L'étiquetage de la classification des 13 textes (partagés en 9 classes) prend en compte, dans la mesure du possible, l'ensemble des caractères du profil de chaque classe de textes sur l'ensemble des 15 classes de mots.

On lit, notamment, que j22, qui n'est autre que le bloc {D1...D6} du "Don tranquille", est fortement lié à i265 (où sont précisément des prépositions de lieu); le lien étant le plus fort avec j15 = {D1, D2, D3}; mais un peu plus faible avec j19 = {D4, D5, D6}, qui est également marqué i263+, pour signaler une certaine association avec la classe i263.

La classe i269, où sont notamment les mots:

а же надо возле тогда ежели

remarqués dans le plan (1, 2) vers (F2>0) sert à étiqueter Ts1, et surtout Ts2. Quant à (эраз: à la fois), absent ou quasi absent dans {K, S, R} (cf. tableau) il constitue une classe de la partition définie par les 14 nœuds les plus hauts.

D1	D2	D3	15	265++++	_____	22	265+++	_____
D4	D5	D6	19	263+	_____			
K\$			13	230+++	_____	268++	_____	24
Ts1			9	269+	_____	70++	21	23
D7	D8		18	261++	246+	_____		
S	R		17	246+	_____	20	_____	
Ts2			10	269++	_____			

CAH des 13 textes d'après leur correspondance avec 142 mots

7 Conclusion provisoire

L'analyse des correspondances montre ici tout son potentiel. Elle a permis de dégager des relations bien déterminées entre les textes qu'on y a soumis; relations qu'on n'aurait pas pu mettre en évidence autrement.

On peut également souligner que les résultats obtenus ont pu l'être en utilisant quasi exclusivement des formes de mots outil, à l'exclusion des autres catégories de mots. Comme mentionné dans [TEXTES RUSSES], in *CAD*, Vol.XIX, n°1, 1994, l'utilisation, par exemple, de pronoms personnels aurait introduit des proximités qui seraient venues fausser complètement les graphes.

Nous avons joint à l'analyse "Le Destin d'un Homme" bien que ce récit ait été écrit à la première personne. L'approche par les mots outil telle que définie plus haut, indique que cette particularité n'a pas distordu les résultats. Il faut cependant rester prudent, devant la diversité des genres et des styles.

En reprenant l'ensemble des observations faites au cours de ces analyses nous ne trouvons pas de raisonnement suffisamment fort pour expliquer les disparités constatées dans l'œuvre du "Don paisible" et les curiosités dans les résultats des diverses analyses. En fait tout se passe comme si le bloc {D1...D6} ne faisait pas partie du reste de l'œuvre de M.A. CHOLOKHOV, comme s'il avait été écrit par un autre auteur avec ses habitudes, ses préférences d'utilisation de mots outil, ses manies, ses préférences. On ne pourra jamais l'affirmer; mais quelle est la probabilité de trouver chez un auteur une telle disparité à l'intérieure d'une œuvre?

Une recherche bibliographique nous a révélé que M.A. CHOLOKHOV avait été accusé de plagiat et qu'on prétendait qu'au moins les 3 premières parties du "Don" avaient été écrites par F.D. KRIUKOV. Pour cette dernière affirmation l'analyse montre clairement qu'il n'en est rien. Les nouvelles de F.D. KRIUKOV n'ont aucune affinité avec le "Don paisible". Il y a eu à l'université d'Oslo (cf. Geir KJETSAA, in *Slavica Norvegica* 1, Oslo; 1984) une étude basée sur la longueur des phrases (déterminée par dénombrement des mots) et qui a conduit à la même affirmation. Mais ce critère ne nous paraît pas stable.

Y a-t-il eu un autre auteur?

"Je ne puis exclure avec une absolue certitude, écrit SOLJENITSINE, que, sans être jamais connu du public mais au contraire ignoré de tous, ait fleuri, pendant la guerre civile et soit mort après celle-ci, sur les rives du Don, un génie littéraire de plus."

L'hypothèse ne semble donc pas pouvoir être écartée. Il serait peut-être intéressant de rechercher dans les diverses publications de cette époque, des auteurs, cosaques ou non, traitant du même sujet et de soumettre à l'analyse leurs écrits.

Pour cela il faudrait consacrer un temps non négligeable à visiter des bibliothèques de la région du Don et de l'actuelle Ukraine.

Selon une étude faite à l'Université de Princeton (et citée dans l'étude de l'Université d'Oslo; cf. *supra*), basée sur un critère totalement différent, chaque auteur a sa façon de commencer et de terminer ses phrases. Nous pensons que c'est le lot de tout le monde: les imitateurs le savent bien (cf. Th. Le LURON). Cette étude montrerait que cette particularité a tout son poids sur les 3 premiers et les 3 derniers mots de chaque phrase. Ayant soumis plusieurs écrits d'auteurs différents à une analyse basée sur cette remarque, les promoteurs de cette étude ont pu montrer qu'elle leur a permis d'identifier sans confusion les auteurs analysés. Nous nous proposons dans une prochaine étape d'appliquer cette approche aux textes analysés ici.

La présente étude ne prétend pas être exhaustive, loin de là. Elle a été faite par des non-spécialistes ayant travaillé sur un ordinateur de bureau, ce qui a conduit à quelques acrobaties. Aussi nous sollicitons du lecteur une bonne part d'indulgence et le remercions pour sa patience.

8 Appendice : Любителям Русской Словесности

8.1 Citations variées extraites de ТИХИЙ ДОН

L'on donne ici des exemples de conversations et de récits. La disparité de style est particulièrement frappante quand c'est le même personnage qui s'exprime dans deux cas.

D2, Livre I, 2-ème partie, Ch. 8 : Пантелей Прокофьевич revenu d'une réunion de village, passe aussitôt chez sa femme, souffrante:

- "Ну, как? Не легшает всё"?

на лицо Ильиничны тенью легла замкнутость.

- "Так же... стреляет в суставы, ломит.

- "Говорил дуре, не лезь в воду осенью. Раз знаешь за собой беду - не рыпайся! - вскипел Прокофьевич, чертя на полу костылем широкое круги. - Аль мало баб? Будь они трижды прокляты, твои конопи: помочила, а теперя... Бож-же-жж мой, то-то... Эх!

D5, Livre II, 5-ème partie, Ch. 13 : Пантелей Прокофьевич vient d'étrenner une nouvelle casquette (сынновий подарок)

Сунулся примерить было к зеркалу, но взглядом стерегла его Ильинична. Старик перенял её взгляд, круто вильнул, захромал к самовару. Перед ним и примерял, надевая фуражку на бекрень.

- "Ты чего ж это, дрючок старый? - напустилась Ильинична.

Но Пантелей Прокофьевич отбрехался:

- "Господи! Ну, и глупая ты! Ить самовар, а не зеркала? то-то и оно!

D3, Livre I, 3-ème partie, Ch. 2 : Григорий, nouvellement mobilisé, se fait insulter par son supérieur:

Суток пять спустя Григорий на водопое уронил в колодезь цыбарку, вахмистр иалетел на него коршуном, занёс руку.

- "Не трожь!... глухо кинул Григорий, глядя в рябившую под срубом воду.

- "Что? лезь, гад, вынимай! Морду искровяню!

- "А ты не трожь! - не поднимая головы, медленно растягивал слова Григорий.

D5, Livre II, 5-ème partie, Ch. 22: Le même Григорий, au cours d'une discussion avec Валет:

Выкрик Валета взбесил его окончательно. Он вскочил, как ушибленный, – подойдя в упор к ёрзавшему на табурете Валету, с трудом удерживая руку, зудевшую желанием ударить, сказал:

– “Замолчи, гадёныш! Сопля паршивая! огрызок человечий! Чего ты командуешь! Ступай, кой тебя... держит! Валяй, чтоб тобой и не воняло тут! Ну-ну, не говори, а то как отхожу тебя на прощанье...”

D4, Livre II, 4-ème partie, Ch. 10: récit : L'officier Лисницкий, monarchiste convaincu, vient être muté à Pétrograd après la révolution de Février. Il a assisté au départ du tsar déchu de Могилёв.

За покатым свалом Днепра небо крашено лазурью, киноварью, ржавой позолотой, каждый штрих на горизонте так неосяземо воздушен, что больно касаться взглядом. У выезда небольшая толпа из чинов ставки, военных, штатских... Выезжающий крытый автомобиль. За стеклом, кажется, Фредерикс и царь, откинувшийся на спинку сиденья. Осунувшееся лицо его с каким-то фиолетовым оттенком. По бледному лбу кривой черной полукруг папахи, формы казачьей конвойной стражи.

Лисницкий почти бежал мимо изумленно оглядывавшихся на него людей. В глазах его падала от края черной папахи царская рука, отдававшая честь, в ушах звенели бесшумный холостой ход отъезжающей машины и унижительное безмолвие толпы, молчанием провожавшей последнего императора.

D4, Livre II, 4-ème partie, Ch. 10: récit : Le même Лисницкий, circulant dans les rues de Pétrograd et revivant cette scène, s'interroge sur sa situation au milieu de tous ces bourgeois endimanchés:

“Но с кем же ты-то в конце концов? – задал он сам себе вопрос и, улыбаясь, решил:

“Ну, конечно же, вот с этими В них частичка самого меня, а я частичка их коллектива... Всё, что есть хорошего и дурного в моём классе, есть в той или иной мере и у меня...”

8.2 Nos conclusions à verser au débat

“Кто написал *Тихий Дон*”?

Знатокам шолоховской литературы хорошо известны все толки споры и пересуды, не раз возникавшие вокруг этого вопроса и преследовавшие Шолохова в течение всей его долгой литературной жизни (см. книгу Б. Гура: “Как создавался Тихий Дон”; и выше процитиров. доклад норвежских славистов, in *Slavica Norvegica* 1, Oslo; 1984), начиная с необоснованных слухов распространившихся сразу же после появления первого тома романа в 1928 году, и кончая в 1974 году выступлениями Р.А. Медведева и строго-литературоведческими соображениями И.Н. Медведевой – Томашевской в её прошумевшей и к сожалению недоконченной книге “*Стремля Тихого Дона*” (с предисловием Солженицына).

Новизной и оригинальностью своего подхода к “загадкам *Тихого Дона*” книга И.Н. Медведевой–Томашевской с новой силой поставила перед читателем заглушенный как будто “спор об

авторстве” и, перенеся его на чисто научную почву, открыла литературоведам новый путь исследования. Особенно заманчивой и плодотворной в ней (несмотря на некоторые методологические слабости, на которые указывают А.Г.Макаров и С.Э.Макарова в своей книге: “К истокам *Тихого Дона*” а также в *Новом Мире*, №, 1993) показалась гипотеза, основанная на наличии в романе разного рода “странностей”, свидетельствующих, по мнению Томашевской, о появлении в романе двух совершенно разнородных авторских начал – “Автора и Со-автора”, постоянное противоборство которых и приводит к необъяснимым иначе срывам и противоречиям и, в конечном итоге, к полному искажению всего романа. Ещё интереснее может быть, оказалась её попытка, к сожалению оставшаяся незавершенной, отслоить повсеместно авторское и со-авторское, дабы вернуть роману его первоначальный облик.

Среди многочисленных исследований, проводимых за последние 20 лет по стопам И.Н.Медведевой-Томашевской как в России так и за рубежом, следует отметить, ссылаясь на пространную работу Макаровых (предисловие к книге: *К истокам тихого Дона*) таких писателей как Р.А.Медведев, Герман Ермолаев, М.Мезенцев, З.Бар-Седла, Л.Колодный, В.Литвинов и пр., но несмотря на ценность проведённой ими работы, многое в творчестве Шолохова остаётся по сей день неразгаданным.

Единственной до сих пор попыткой подойти к этому вопросу с научно-статистической точки зрения принадлежит Норвежской группе славистов, но ввиду “не вполне корректной, а местами просто ошибочной модели изучения текстов”, на которую указывают Е.В.Вертель и Л.З.Аксенова (процитир. Макаровыми), вряд ли можно отнестись с полным доверием к опубликованным ею результатам.

Располагая мощным орудием Анализа Данных не раз доказавшего свою эффективность в области лингвистики, мы сочли целесообразным подвернуть ему главные крупные произведения Шолохова: “*Тихий Дон*”, “*Поднятая целина*”, “*Они сражались за родину*” и “*Судьба человека*”, противопоставив им несколько рассказов Ф.Д.Крюкова, в качестве главного – наиболее подходящего – кандидата на авторство.

Результаты анализа, приведённые в данной статье, свидетельствуют – как можно видеть по графикам – о явном разрыве между первыми частями “*Дона*” и всеми другими, взятыми под учёт, произведениями Шолохова, включая 2 последние части “*Дона*” (D7, D8)

Подвергнутые анализу произведения делятся действительно не

на 2 (как следовало ожидать): Шолохов и Крюков, а на 3 разнородные группы:

1) Крюков.

2) Шолохов: 2 последние части "Тихого Дона", обе части "Целины" (Ts1, Ts2), "Они сражались за Родину" (R) и "Судьба человека" (S).

3) Первые три книги "Дона": (D1, D2, D3), (D4, D5), (D6) с некоторым отклонением двух последних - особенно D6 - в сторону чисто шолоховской полосы.

Любопытно отметить тут, что 3-ья книга "Дона" (D6) и первая часть "Целины" (Ts1) несмотря на то, что писались они одновременно, не принадлежат к одной и той же сфере творчества.

Заключения - выражаясь со всей осторожностью подобающей подобного рода исследованиям - можно сформулировать так:

1) Очень мало вероятно (и в этом Анализ Данных как бы подтверждает предположения И.Н. Медведевой-Томашевской и многих других знатоков шолоховской литературы), что Шолохов написал весь "Тихий Дон". Его сфере достоверно принадлежит последняя книга "Дона" (D7, D8) и, до некоторой степени, третья книга (D6).

2) За неимением поздних произведений Крюкова, и основываясь пока только на существующем материале, можно сказать что так же мала вероятность, что автором какой-то части "Тихого Дона" может быть Крюков. И в этом Анализ Данных идет в разрыв с предположениями Томашевской и Солженицына - хотя ни один из них не исключал возможность существования третьего человека.

"Не могу абсолютно уверенно исключить (пишет Солженицын), что - был, жил никогда публично не проявленный оставшийся всеми неизвестен, в Гражданскую войну расцветший и вослед за ней погибший - ещё один донской литературный гений..."

Так кто же, наконец, написал Тихий Дон, спросит озадаченный *Любитель русской словесности*? А как бабушка сказала...

Remerciements : *Les auteurs remercient le Pr. J.-P. BENZÉCRI qui a bien voulu s'intéresser au présent travail et en parachever la publication.*