



## Génétique mémorielle. Shoah, mémoire et ADT

Bénédicte Pincemin, Damon Mayaffre, Serge Heiden, Philippe Weyl

### ► To cite this version:

Bénédicte Pincemin, Damon Mayaffre, Serge Heiden, Philippe Weyl. Génétique mémorielle. Shoah, mémoire et ADT. Damon Mayaffre; Céline Poudat; Laurent Vanni; Véronique Magri; Peter Follette. JADT 2016 - Statistical Analysis of Textual Data, Jun 2016, Nice, France. Presses de FacImprimeur, JADT 2016 - Statistical Analysis of Textual Data, volume 2, pp.495-506, 2016, JADT 2016 - Statistical Analysis of Textual Data. <hal-01361988>

**HAL Id: hal-01361988**

**<https://hal.archives-ouvertes.fr/hal-01361988>**

Submitted on 12 Sep 2016

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Génétique mémorielle. Shoah, mémoire et ADT

Bénédicte Pincemin<sup>1</sup>, Damon Mayaffre<sup>2</sup>, Serge Heiden<sup>3</sup>, Philippe Weyl<sup>4</sup>

<sup>1</sup> CNRS, UMR 5191 ICAR, Université de Lyon (Lyon– France)

<sup>2</sup> CNRS, UMR 7320 BCL, Université Nice-Sophia Antipolis (Nice– France)

<sup>3</sup> ENS, UMR 5191 ICAR, Université de Lyon (Lyon– France)

<sup>4</sup> FMS – Fondation pour la Mémoire de la Shoah (Paris– France)

## Abstract

As part of the Matrix Équipex, we gather a corpus of testimonies from survivors of the Shoah, which is diachronic from the perspective of the time of writing: immediate testimonies written after the end of the war, median testimonies written in the 1980's and recent testimonies written in the 2000's. The objective is to make use of TDA tools to distinguish the primary vocabulary which could be directly inherited from the camps and a reconstructed vocabulary that includes, over the years, the historical understanding of an event which was unintelligible to contemporary victims. In a way, we don't practice text genetics but memory genetics within the texts of our corpus. Contrastive statistics allows us to deconstruct testimonies into distinct lexical sedimentations reflecting superimposed layers of memory.

## Résumé

Dans le cadre de l'Équipex Matrice, nous rassemblons un corpus de témoignages de survivants de la Shoah, diachronique du point de vue de la date d'écriture : témoignages immédiats écrits à la sortie de la guerre, témoignages médians écrits dans les années 1980, et témoignages récents écrits dans les années 2000. L'objectif est de mobiliser les outils de l'ADT pour sérier le vocabulaire primaire sans doute hérité directement des camps et un vocabulaire reconstruit qui intègre, au fil des années, la compréhension historique d'un événement inintelligible aux victimes contemporaines. D'une certaine manière, il s'agit non pas de faire de la génétique de texte, mais de la génétique mémorielle au sein des textes, en déconstruisant les témoignages en sédimentations lexicales distinctes, traces de strates mémorielles superposées.

**Key words :** mémoire, Shoah, TXM, méthodologie textométrique

## 1. Introduction<sup>1</sup>

Entre mémoire individuelle et mémoire collective, les récits de vie autour d'événements historiques majeurs comme la Grande Guerre ou la Shoah sont depuis plusieurs décennies des objets privilégiés d'étude pour les Sciences Humaines et Sociales. S'ils interrogent ainsi les chercheurs comme la société, c'est sans doute au premier chef parce qu'ils articulent de manière complexe, dans une narration qu'il reste à décrire, Mémoire et Histoire [Ricoeur 2000], témoignage intime et grand récit, souvenir individuel et mémoire sociale.

Dans le cadre de l'Équipex Matrice (<http://www.matricememory.fr/>) nous avons montré que les témoignages de la Shoah, recueillis par la Fondation pour la Mémoire de la Shoah (FMS),

---

<sup>1</sup> Cette version de la communication annule et remplace la version publiée dans les actes papiers. En effet, entre la soumission des textes (janvier) et la conférence (juin), une erreur d'association entre les métadonnées et les textes du corpus a été détectée et corrigée, si bien que les observations ont été mises à jour en conséquence pour la communication et les actes numériques.

recélaient dans leur matérialité textuelle deux types de lexique entremêlés : un lexique sans doute directement issu de l'événement traumatique – lexique primaire pourrait-on dire – et un lexique *a priori* reconstruit ou réélaboré au fil des années [Mayaffre et Ben Hamed 2014].

Cette contribution met à l'épreuve cette conclusion sur un corpus doté de sa dimension chronologique en recueillant 20 témoignages d'Auschwitz selon trois moments d'écriture des témoins : l'immédiat après-guerre, les années 1970-1990 et les années 2000-2010 (fig. 1)<sup>2</sup>.

| <i>Id.</i> <sup>3</sup>                                   | <i>Auteur</i>              | <i>Titre</i>                                            | <i>Nb. mots</i>       |
|-----------------------------------------------------------|----------------------------|---------------------------------------------------------|-----------------------|
| 1945a                                                     | Mayer, Alex                | Auschwitz, le 16 mars 1945                              | 40 477                |
| 1945b                                                     | Kohen, Guy                 | Retour d'Auschwitz                                      | 22 015                |
| 1945c                                                     | Oppenheimer, Jean          | Journal de route, 14 mars – 9 mai 1945                  | 58 516                |
| 1945d                                                     | Unger, Julien              | Le Sang et l'Or                                         | 59 827                |
| 1945e                                                     | Holstein, Denise           | Le Manuscrit de Cayeux-sur-Mer                          | 12 724                |
| 1946a                                                     | Altmann, Erich             | Face à la mort                                          | 39 597                |
| <i>Total période 1945-1946 :</i>                          |                            |                                                         | <i>233 156</i>        |
| 1978a                                                     | Klein, Eugène              | Les Loups. Témoignage d'un déporté, matricule 126026    | 32 252                |
| 1986a                                                     | Grinbaud, Simon            | XI <sup>e</sup> commandement : « Tu n'oublieras point » | 88 367                |
| 1990a                                                     | Roth, Nicolas              | Avoir 16 ans à Auschwitz. Mémoire d'un juif hongrois    | 185 267               |
| 1995a                                                     | Goltman, Pierre            | Six mois en enfer                                       | 11 753                |
| 1998a                                                     | Grossman, Adèle            | La Mémoire dans la chair                                | 119 675               |
| <i>Total période 1978-1999 :</i>                          |                            |                                                         | <i>437 314</i>        |
| 2002a                                                     | Hirsch, Claude             | Matricule A-16689                                       | 23 230                |
| 2008a                                                     | Toros-Marter, Denise       | J'avais seize ans à Pitchipoï                           | 25 753                |
| 2008b                                                     | Skorka-Jacubert, Régine    | Fringale de vie contre usine à mort                     | 30 964                |
| 2008c                                                     | Palant, Charles            | Je crois au matin                                       | 42 281                |
| 2008d                                                     | Mitzner, Charles           | Seuls au monde                                          | 15 454                |
| 2008e                                                     | Rosenbaum, Isidore         | Je suis né le 8 mai 1945                                | 23 428                |
| 2009a                                                     | Golgevit, Eva              | Ne pleurez pas, mes fils...                             | 19 149                |
| 2009b                                                     | Lichtsztejn-Montard, Sarah | Chassez les papillons noirs.                            | 69 235                |
| 2010a                                                     | Spingarn, Odette           | J'ai sauté du train. Fragments                          | 24 925                |
| <i>Total période 2002-2010 :</i>                          |                            |                                                         | <i>274 419</i>        |
| <b><i>TOTAL (nb. de mots du corpus)<sup>4</sup> :</i></b> |                            |                                                         | <b><i>944 889</i></b> |

**Figure 1. Corpus – 20 témoignages (La Collection Témoignages de la Shoah – FMS)**

L'hypothèse de travail est que les moments d'écriture déterminent la forme et le fond des témoignages. Particulièrement, pour être concret, le matériel lexical pourrait varier selon que l'on se souvienne 5 ans, 20 ans ou 50 ans après l'événement. Plus précisément encore, le lexique primaire que nous avons identifié [Mayaffre et Ben Hamed 2014] devrait se trouver surreprésenté dans les témoignages proches de l'événement, là où le lexique secondaire devrait sur-apparaître au détour des années 2000.

Méthodologiquement enfin, cette introspection mémorielle se fera grâce aux outils de l'ADT. Nous posons en effet que la mémoire aussi, et en tout cas la narration, est pour partie un

<sup>2</sup> Les trente années de « silence » qui séparent les deux premières périodes seraient à étudier historiquement : pourquoi les années 70 suscitent-elles ou libèrent-elles la parole des témoins ?

<sup>3</sup> L'identifiant correspond à l'année de rédaction ; la lettre distingue les textes d'une même année.

<sup>4</sup> Le module de découpage en mots du logiciel TXM a évolué entre la première et la seconde version de ce texte.

phénomène statistique (*ie.* qui répond à des règles de récurrences non dues au hasard) ; et l'approche logométrique (ou textométrique) donne les moyens d'une analyse à la fois qualitative et quantitative au plus près du texte, sur un corpus d'un million de mots. Nous avons eu recours au logiciel TXM (<http://textometrie.ens-lyon.fr>) [Heiden, Magué, Pincemin 2010] dont le développement a reçu le soutien de l'Équipex Matrice.

## 2. Cartographie globale du corpus : le temps de la mémoire

Depuis plusieurs années maintenant, sinon depuis l'origine, l'ADT a appris à fonctionner sur des « corpus maquettes » [Metwally – *à paraître*]. Pour des raisons linguistiques et méthodologiques, nous réduisons ici le corpus à ses 400 noms communs les plus fréquents. Cette réduction peut se concevoir linguistiquement par l'importance du substantif pour dire le monde, les choses et l'idée. Et, de fait, expérimentalement, sur notre corpus et pour notre problématique de contraste entre les périodes d'écriture, nous avons observé que les autres parties du discours étaient moins pertinentes pour décrire et caractériser les différentes périodes<sup>5</sup>. Cette réduction se justifie aussi méthodologiquement, particulièrement pour les traitements statistiques qui franchissent le pas de la visualisation : représenter simultanément des centaines de milliers de mots fait achopper sur des difficultés de lisibilité et de perception globale, là où donner à voir 400 items produit une vue d'ensemble synthétique révélant des « lignes de force » structurant le corpus.

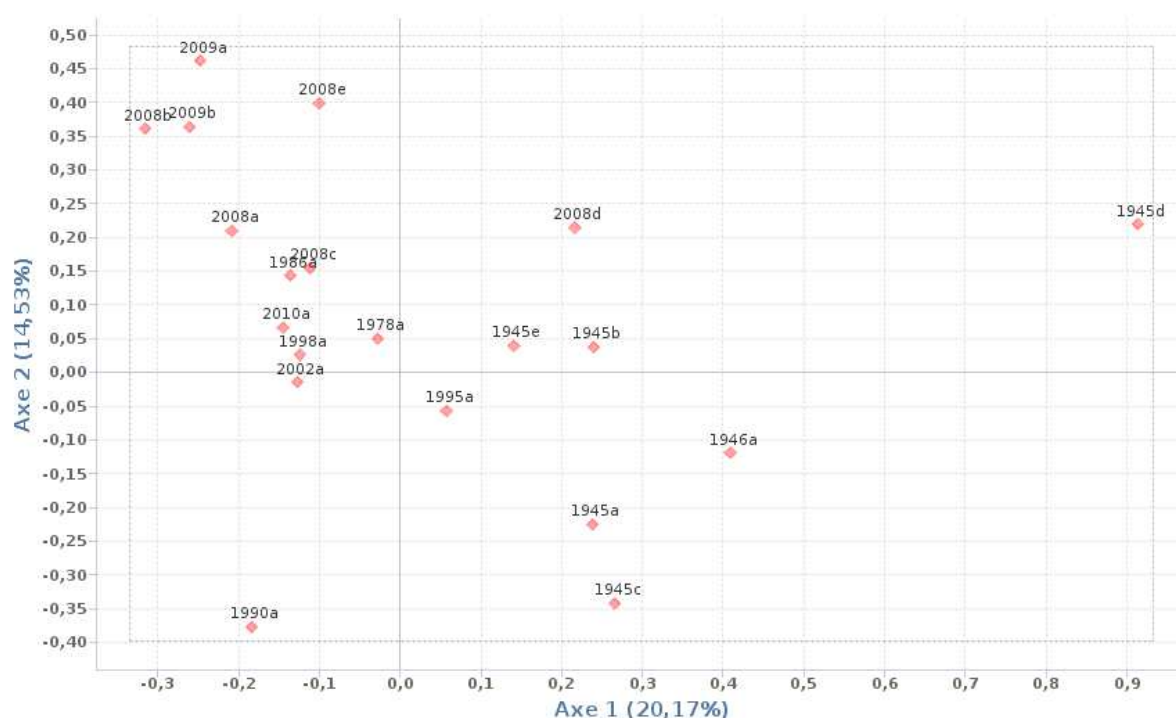


Figure 2. Plan factoriel des textes caractérisés par les 400 formes nominales les plus fréquentes du corpus

<sup>5</sup> Les verbes, les adjectifs, les adverbes ont été étudiés mais les scores sont moins marqués et beaucoup plus d'entre eux se sont avérés spécifiques d'un texte plutôt que d'une période. Quant aux pronoms personnels ou aux déterminants, nous avons considéré qu'ils traduiraient ici des choix rédactionnels et stylistiques plus difficiles à cerner dans le cadre de notre problématique.

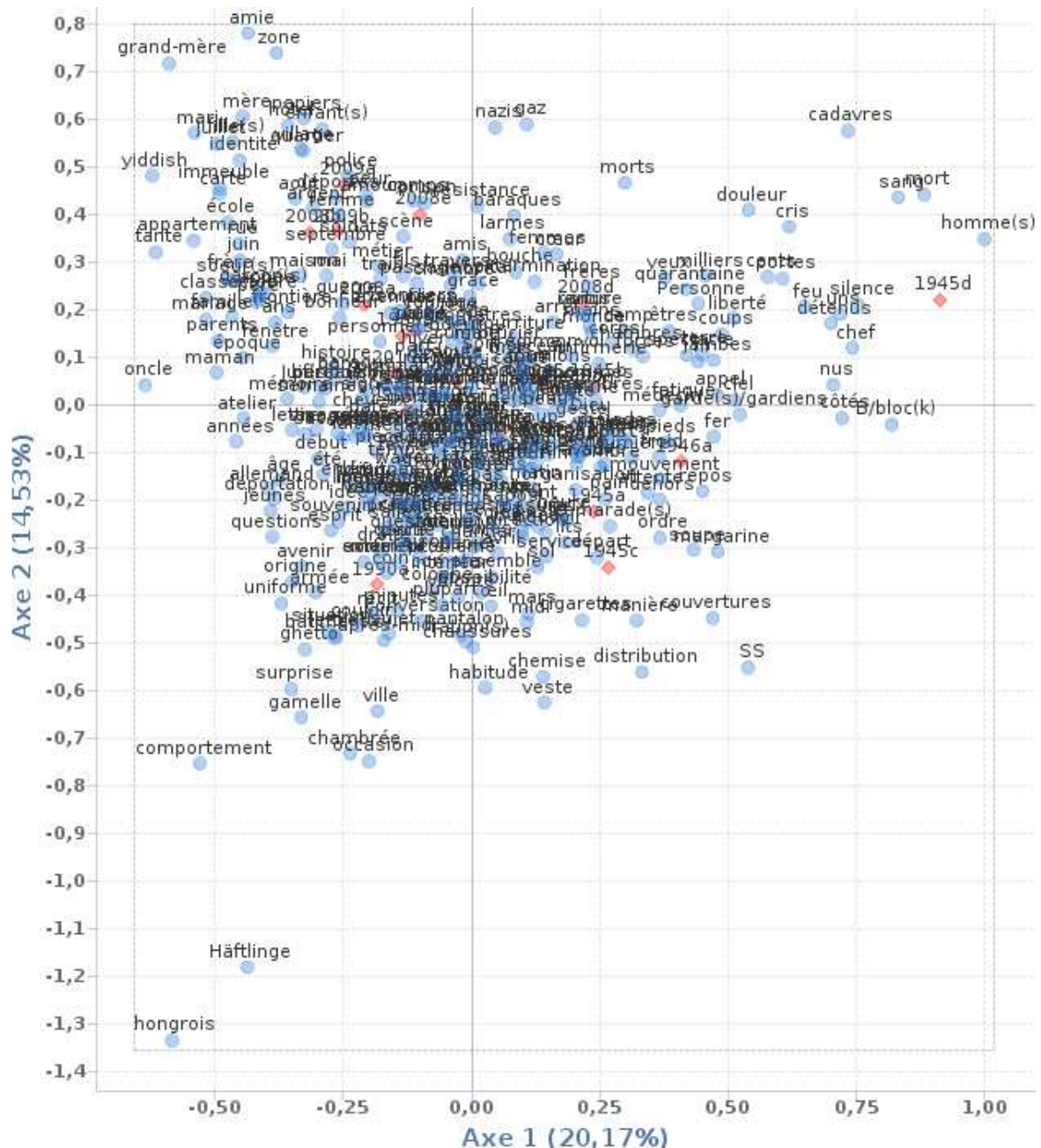


Figure 3. Plan factoriel des textes et des 400 substantifs les plus fréquents, avec affichage des substantifs

L'analyse factorielle des 400 substantifs les plus utilisés<sup>6</sup>, distribués dans les 20 témoignages donne une carte (fig. 2) qui peut se lire selon une logique historique, avec les témoignages les plus anciens à droite de l'axe 1, et les témoignages plus récents à gauche de l'axe, seul un point (2008d) faisant exception à cette bipartition chronologique.

Dès lors, la projection des 400 mots contribue à l'interprétation (fig. 3 et 4).

<sup>6</sup> Non lemmatisés, sauf pour une vingtaine de mots pour lesquels on a préalablement observé qu'ils ont le même profil dans le plan 1x2 de l'AFC. Sont également neutralisées deux variations graphiques qui sinon accaparent en bonne partie l'axe 1 (au vu des contributions) : fusion des lignes « SS » et « S.S. », « bloc » et « Block » (cette dernière variation étant cependant analysée par ailleurs, cf. §3.4).



De manière spectaculaire, le vocabulaire à droite de l'axe, c'est-à-dire utilisé de manière privilégiée par les témoignages les plus anciens, relève des catégories que nous avons identifiées comme contemporaines à la Shoah. C'est d'abord un vocabulaire sensoriel qui exprime la douleur des corps et des esprits : les « cris », le « silence », le « sang », la « mort », le « feu », les « coups ». C'est également, très présent, le vocabulaire du besoin vital et principalement celui de la nourriture : le « pain », la « soupe », la « margarine » mais aussi le « froid », la « veste » et la « chemise » (cf. infra fig. 8).

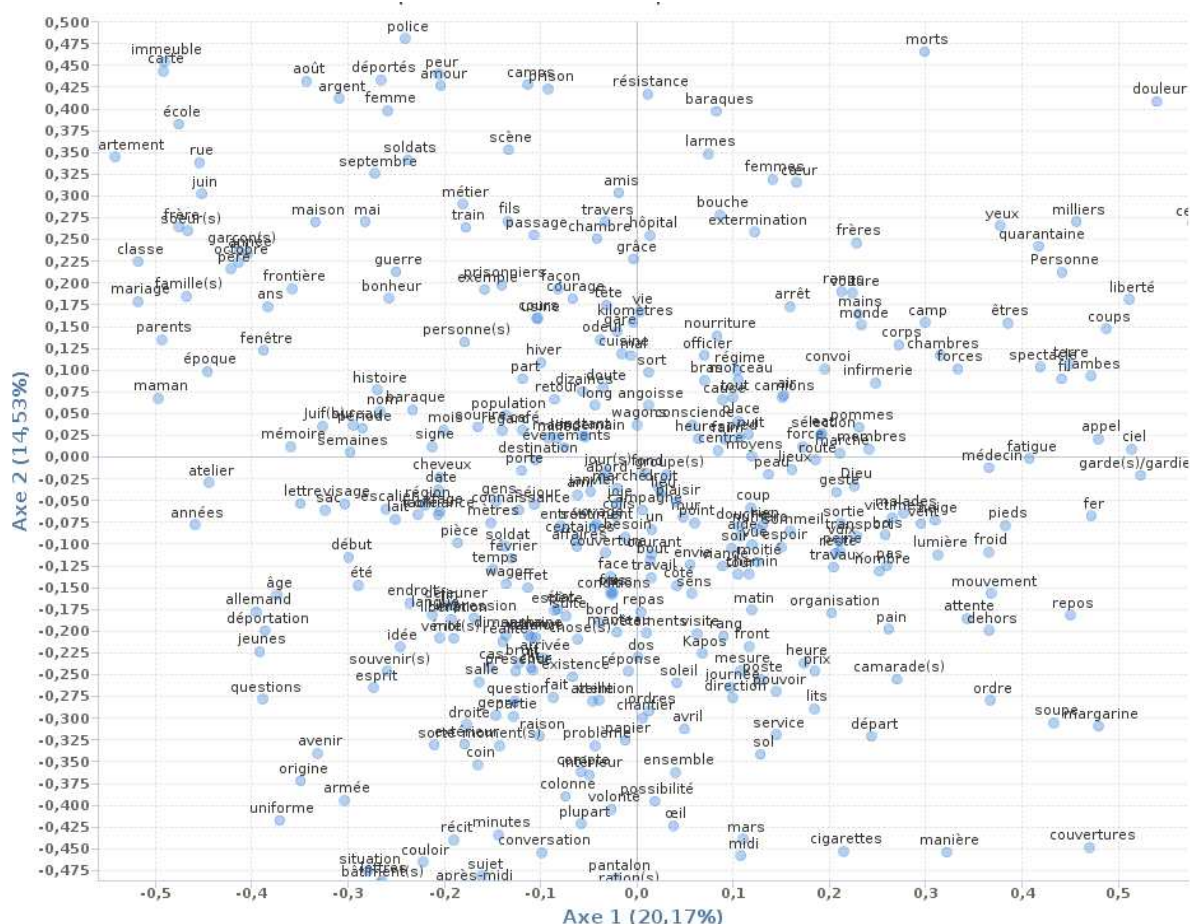


Figure 4. Zoom sur le centre du plan factoriel des textes et des 400 substantifs les plus fréquents.

Symétriquement, le vocabulaire à gauche de l'axe c'est-à-dire utilisé principalement dans les témoignages largement post-datés répond à notre description d'un vocabulaire ré-élaboré, c'est-à-dire appartenant plus à la re-présentation de la Shoah qu'à la vie des camps elle-même. L'ensemble du vocabulaire familial s'y trouve (« grand-mère », « mère », « enfant(s) », « mari », « tante »,...). Dans les années 2000 notamment, la Shoah ne semble pouvoir se dire et se comprendre par les témoins sans la reconstruction du roman familial précisément brisé par la déportation et l'extermination, et par définition étranger à la vie concentrationnaire. De manière conforme à nos attentes, le vocabulaire politico-religieux également se trouve à proximité des textes tardifs : « juif(s) », « Yiddish » ou simplement « guerre » ou « allemand ».

Nous admettons donc que les témoins disposent de deux grands types de lexiques qu'ils mobilisent de manière contrastée pour se remémorer l'événement : un lexique proche de l'événement particulièrement mobilisé à la sortie des camps et de la guerre ; un lexique plus (re)élaboré qui se trouve naturellement mobilisé à distance de l'événement.

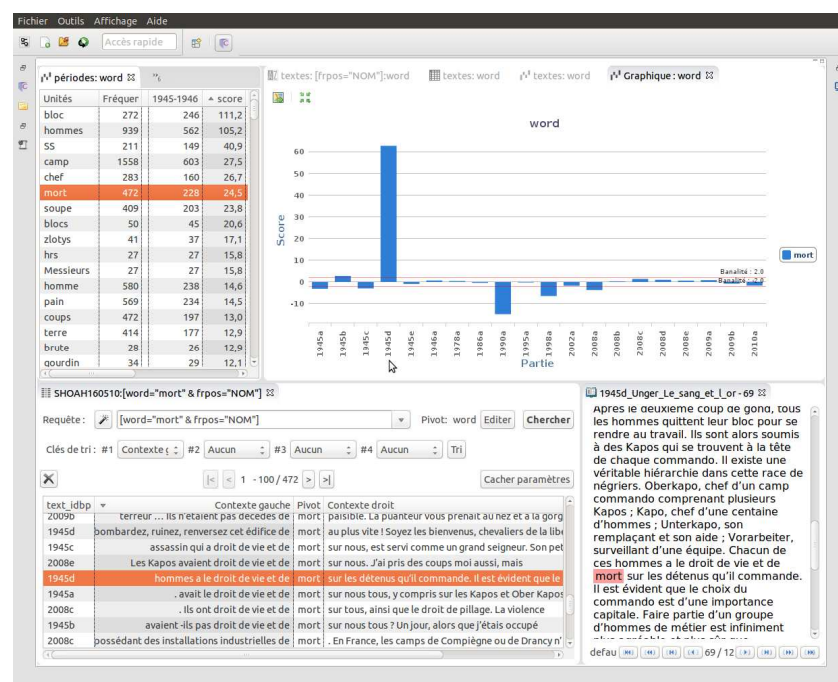
### 3. Distributions chronologiques et mémoire

#### 3.1. Éléments méthodologiques

Divisé naturellement en 20 textes datés, le corpus a été structuré en trois périodes. Pour la caractérisation précise et contrastive de ces divisions, l'outil lexicométrique le plus approprié est le calcul des spécificités, qui dresse un inventaire hiérarchisés des mots statistiquement sur-représentés au regard du corpus dans son ensemble.

Nous avons procédé au calcul catégorie grammaticale par catégorie grammaticale afin de neutraliser les effets de style liés à des écritures plus nominales vs verbales [Mayaffre 2006].<sup>7</sup>

La géométrie de notre corpus, dont les parties (les périodes) comptent chacune moins d'une dizaine de textes, impose pour chaque mot calculé comme spécifique de la partie. de s'assurer que cette spécificité n'est pas due à une sur-représentation dans un seul texte, mais est un point commun partagé par différents textes de la partie. Cette vérification s'opère en consultant parallèlement les mêmes données, mais cette fois-ci partitionnées sur les textes (cf. en fig. 5, la zone d'écran en haut à droite).



Par ailleurs la consultation des contextes d'usage (via la concordance, cf. en fig. 5 la zone d'écran en bas à gauche) reste un préalable indispensable pour l'interprétation. Elle permet par exemple de s'apercevoir de l'hétérogénéité sémantique d'un mot (ex. « classe » : scolaire, sociale,...), ou de repérer des emplois différents selon les périodes, qu'un calcul de cooccurrences dans chaque période permet éventuellement de confirmer.

Figure 5. Interface pour le dépouillement des spécificités

Par exemple, pour le mot « homme », on observe en première période davantage d'usages génériques au sens d'être humain (fig. 6). Dans cette première période, le plus fort cooccurent d'« homme(s) » est l'article défini (fig. 7), alors que par la suite, les usages sont davantage anaphoriques ou désignent concrètement un individu ou un groupe. De même, pour le « départ » : en première période c'est souvent un moment attendu ou redouté, un événement qui a sa consistance propre (ex. « le jour tant redouté du départ », « Ils n'étaient pas du prochain départ. », « on reparla de départ », « nouvelle rumeur de départ » (1945a),

<sup>7</sup> Exemple de mise en œuvre dans le logiciel TXM : sur la partition en périodes, Index de [frpos="NOM"] ; transformation en Table lexicale, avec marges= toutes les occurrences recensées par l'index ; tri sur la fréquence et fusion des lignes des noms de fréquence inférieure à 10 (pour alléger la manipulation de la table lexicale sans incidence sur les valeurs de spécificité calculées) ; sur cette table, calcul des spécificités.

« Toujours rien de nouveau concernant le départ. », « Cette fois-ci, cela se précise et prend des formes tangibles. Je parle du départ, bien entendu. », « La fièvre d'un départ éventuel s'est calmée beaucoup, on n'en parle même plus du tout ce matin. », « La vraie nouvelle du vrai départ n'a pas provoqué l'éclat, l'effet de bombe qu'on aurait pu attendre » « le signe du départ tant désiré » (1945c), « Fort de la rumeur relative à notre départ en Allemagne » (1946a)). Alors que dans les témoignages ultérieurs domine un usage plus phraséologique (« au départ ») ou ponctuel (avant/après/pour/depuis... le/leur/notre... départ).

|       |                                                                                                                                              |
|-------|----------------------------------------------------------------------------------------------------------------------------------------------|
| 1945a | s'adressant pour une fois non à des numéros, à des esclaves, mais à des hommes.                                                              |
| 1945b | la misère ravale l'individu au rang de la bête et lui fait perdre toute dignité. L'homme vraiment mis à nu, sans nul masque, n'est pas beau. |
| 1945b | C'est à des moments semblables que l'homme se doit de résister à la tentation de suicide                                                     |
| 1945c | Et voilà la preuve que l'homme est un drôle d'animal                                                                                         |
| 1945c | Mais combien les hommes se laissent mener, impressionner facilement !                                                                        |
| 1945d | Il n'est pas permis à l'homme véritable de se taire lorsqu'il est témoin d'un pareil massacre !                                              |
| 1945d | Ô, hommes, qu'avez-vous fait des hommes, vos frères !                                                                                        |
| 1945d | – Comment des hommes peuvent-ils faire cela ? – Des hommes, non ! Mais les SS sont-ils des hommes ?                                          |
| 1945d | Quel homme peut prévoir la parcelle de vie qui lui sera attribuée ?                                                                          |
| 1946a | L'homme dans une telle situation est une créature bien particulière.                                                                         |

Figure 6. Exemples d'usages du mot « homme(s) » au sens d'« être humain » dans les écrits de 1945-46.

| 1945-1946           |       |       |    | 1978-1999 |       |       |    | 2002-2010 |       |       |    |
|---------------------|-------|-------|----|-----------|-------|-------|----|-----------|-------|-------|----|
| Cooc.               | Fréq. | Cofrq | S+ | Cooc.     | Fréq. | Cofrq | S+ | Cooc.     | Fréq. | Cofrq | S+ |
| le                  | 15204 | 389   | 60 | jeune     | 503   | 37    | 40 | du        | 4792  | 73    | 21 |
| du <sup>8</sup>     | 4148  | 120   | 22 | un        | 9168  | 95    | 29 | un        | 5900  | 80    | 20 |
| cent                | 71    | 18    | 19 | ce        | 6042  | 70    | 24 | ce        | 3388  | 52    | 15 |
| deux                | 379   | 32    | 18 | brave     | 29    | 7     | 12 | jeune     | 274   | 16    | 13 |
| .                   | 11027 | 195   | 12 | du        | 6649  | 47    | 9  | deux      | 588   | 15    | 7  |
| @card@ <sup>9</sup> | 1362  | 45    | 10 | deux      | 812   | 14    | 7  | @card@    | 1457  | 22    | 6  |

Figure 7. Déterminants et adjectifs antéposés cooccurrents avec « homme(s) » selon les périodes<sup>10</sup>

Dernière remarque, certaines unités lexicales sont de l'ordre du syntagme plutôt que du mot. Les syntagmes les plus fréquents peuvent être recensés systématiquement ; puis un calcul de spécificité peut être appliqué individuellement à tout syntagme<sup>11</sup> pour confirmer l'irrégularité de sa distribution (par ex. « zone libre » a une spécificité de +5,6 en 2002-2010).

### 3.2. Le déclin du vocabulaire sensoriel

Évolution chronologique déjà signalée par la carte factorielle (§2), le vocabulaire du sens (*vs* du raisonnement) apparaît très marqué au commencement pour s'effacer au fil du temps (fig. 8). Pour le continuum – c'est-à-dire la continuité mais aussi la démarcation – entre mémoire et histoire [Joutard 2013], ce constat est non obvie, au regard des récits de vie. Sans conteste, le souvenir d'abord sensoriel des témoins semble condamné à se rationaliser à mesure de l'éloignement du traumatisme. L'historicisation de la Shoah passe donc par un

<sup>8</sup> Le lemme « du » se réalise quasiment toujours en « des » (indéfini ou défini contracté), dans tout le corpus on ne compte que 3 occurrences de cette construction au singulier (« du jeune homme » (2), « du saint homme »).

<sup>9</sup> Dans le modèle de langue utilisé par TreeTagger ici, @card@ est le lemme des nombres écrits en chiffres.

<sup>10</sup> Paramètres du calcul : requête = "hommes?"%c, propriété des cooc. = lemme, contexte = 3 mots à droite.

<sup>11</sup> Mise en œuvre dans TXM : Index des occurrences du syntagme dans le corpus ; conversion en Table lexicale avec marges = toutes les occurrences du corpus ; fusion éventuelle des multiples lignes correspondant au syntagme (*i.e.* toutes sauf #RESTE#) ; lancement du calcul des spécificités sur la Table lexicale.



déclin de l'expression de la souffrance physique, par définition individuelle, pour aller vers une intelligibilité d'ordre social et rationnel. Une telle évolution pourrait être un schéma qui marque l'érection de tous les épisodes ou tous les accidents en événement historique.

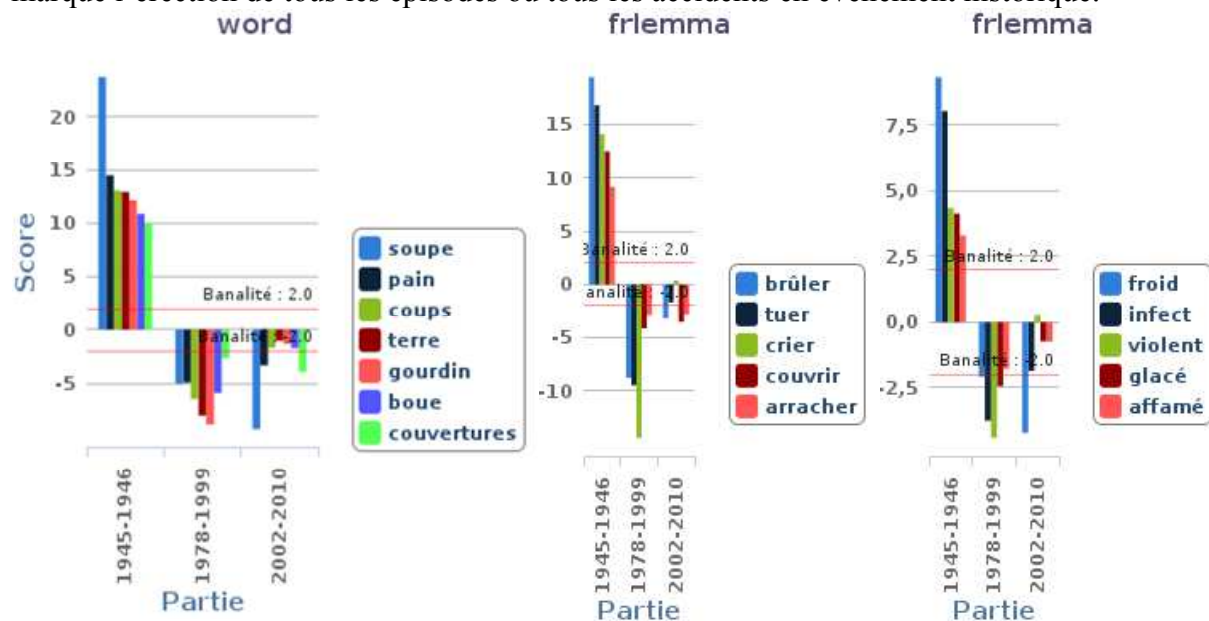


Figure 8. Distribution chronologique du vocabulaire sensoriel (noms, verbes, adjectifs qualificatifs)

### 3.3. L'apparition remarquable du terme « juifs »

À l'inverse du vocabulaire sensoriel, le vocabulaire politico-religieux augmente au fil du temps. L'exemple le plus spectaculaire concerne le terme « juif », fortement sous-représenté dans les premiers écrits (spécificité de -31 en 1945-46) et par contraste sur-représenté ensuite (spécificités de 6,6 et 5,8 dans les deux périodes suivantes) (fig. 9). De fait, les témoignages du corpus écrits dans l'immédiat après-guerre (1945-1946) n'utilisent que très peu la référence identitaire ou religieuse à la judaïté : à la Libération le mot « juif » est chargé négativement, c'est le mot de la stigmatisation, le mot imposé par l'Autre. C'est dans le sous-corpus des années 1970 et suivantes que les « juifs » deviennent les acteurs (en l'occurrence les victimes) de la Shoah : « juif » est devenu le mot de la revendication identitaire et il n'est plus péjoratif. L'évolution du mot reflète l'évolution du rapport à la société française : au sortir de la guerre il s'agit de se fondre dans la société à la fois pour se reconstruire, revenir à la vie, et pour tirer un trait sur la stigmatisation ; dans les années 2000, on affirme l'identité du groupe.<sup>12</sup>

Les témoins utilisent au fil du temps un vocabulaire de plus en plus général, social, historique ou politique au sens étymologique (fig. 10). Replaçant son témoignage dans le cadre général de la « guerre » et de la mobilisation de l'« armée », le récit se repère en évoquant les lieux (« zone » libre ou occupée), les acteurs (« Résistance », « déportés »), les moments (« rafle », « Libération ») d'un vécu partagé, national ou mondial : l'histoire individuelle embrasse désormais l'histoire collective.

Mais ici encore l'interprétation se doit de rester prudente, et cette analyse serait à confirmer sur un corpus plus étendu. En effet, pour nos mesures sur le présent corpus, certains termes analogues semblent échapper à l'influence temporelle : « extermination », « peuple », « paix », qu'on aurait pu attribuer à ce vocabulaire plus conceptuel et historique, sont repérés

<sup>12</sup> Nous remercions Denis Peschanski qui nous a particulièrement éclairés pour cette analyse.

comme banals<sup>13</sup> avec une spécificité inférieure à 1 dans les trois périodes (en dépit d'une fréquence globale supérieure à 70 pour les deux premiers et de 34 pour le troisième, qui auraient permis de capter statistiquement leur éventuelle irrégularité de distribution).

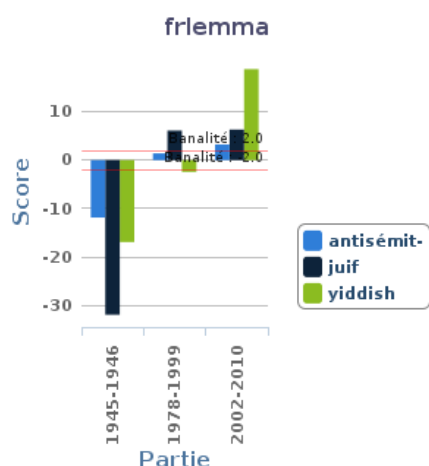


Figure 9. Distribution chronologique du mot « juif » et de termes apparentés

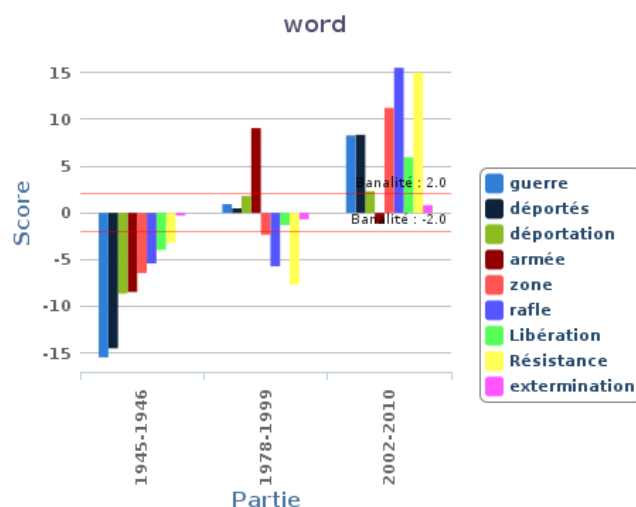


Figure 10. Distribution chronologique de termes historico-politiques dans le corpus

### 3.4. De l'oral vécu à l'écrit distancié

Sur le corpus nous avons par ailleurs remarqué un glissement des graphies francisées aux graphies originales allemandes pour des termes comme « commando » (peu à peu orthographié « Kommando ») ou « bloc » (auquel se substituent « block » ou « Block ») (fig. 11). Ce sont les textes écrits tôt et publiés tôt qui mobilisent les graphies francisées (1945a à 1945d)<sup>14</sup>.

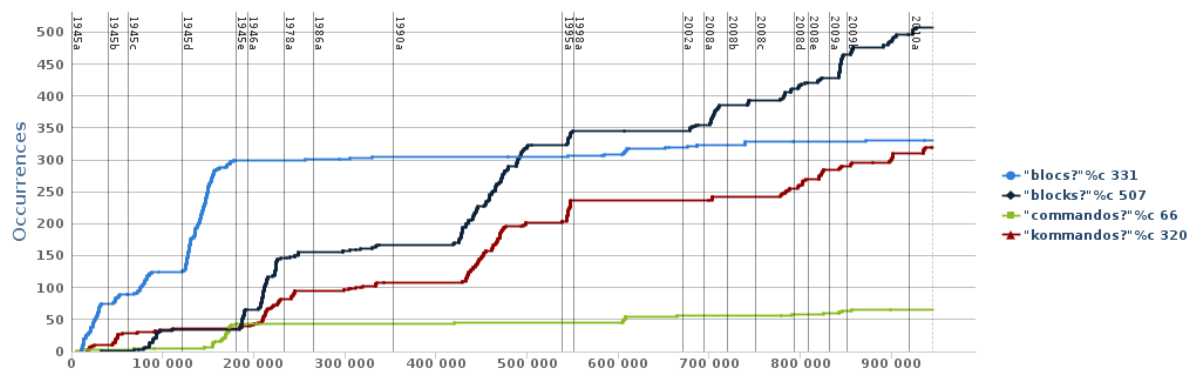


Figure 11. Progression de formes graphiées à l'allemande ou à la française.

Une interprétation de ce phénomène pourrait être que les premiers écrits rendent compte d'un mot dont le témoin a une connaissance surtout orale, et qu'il transcrit de façon « naturelle » dans sa langue. Puis, au fil du temps, les autres écrits (récits, presse) font circuler la graphie

<sup>13</sup> Le repérage des termes banals dans un tableau de spécificités peut se faire dans TXM avec la macro BasicVocabulary.

<sup>14</sup> Les autres textes publiés à la même période (entre juillet 2004 et novembre 2007), au début de la collection, sont 1998a et 2002a. Quant aux textes publiés ultérieurement, on ne peut attribuer le basculement vers les graphies germaniques à l'intervention de l'éditeur, dans la mesure où pour la moitié des textes nous avons travaillé sur les versions « auteur ».

« correcte » (allemande), reprise par les témoignages plus tardifs, peut-être parce qu'elle apparaît à la fois comme la bonne orthographe et parce qu'elle permet de mieux marquer la distance d'avec ces entités étrangères subies. Cette observation, nouvelle à notre connaissance, serait bien sûr à confirmer sur un corpus plus étendu.

### 3.5. La reconstruction du roman familial

Nous avons noté l'importance remarquable du roman familial dans les récits de vie ou « récits de morts » des rescapés d'Auschwitz [Mayaffre et Ben Hamed 2014]. Comme catharsis à l'extermination, la famille absente par définition de l'univers d'Auschwitz, prend pour le témoin une valeur cardinale dans la narration des événements. L'examen chronologique du vocabulaire généalogique est à ce titre éloquent (fig. 12).

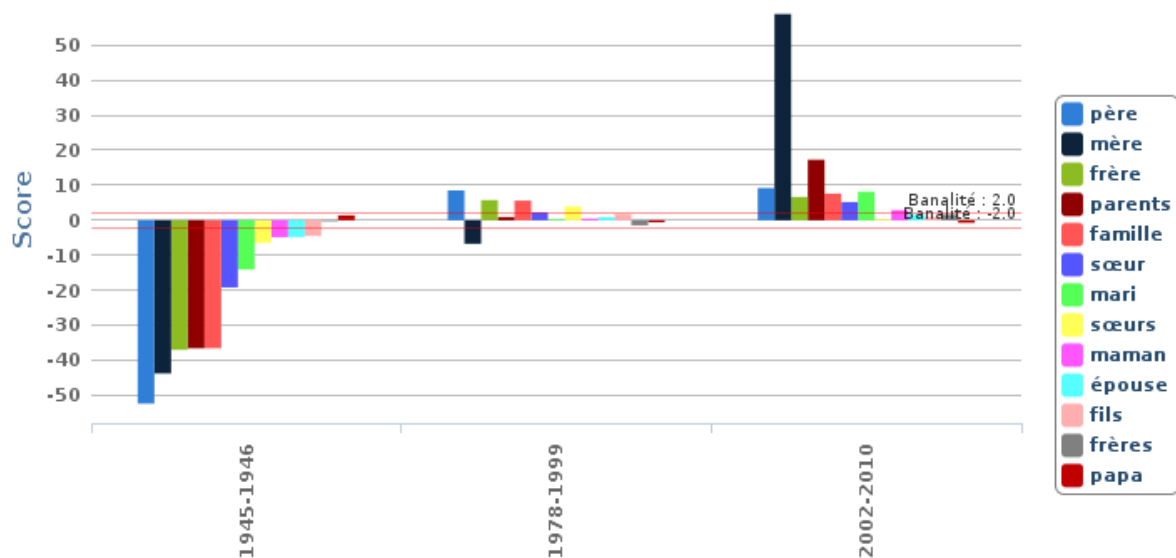


Figure 12. Distribution chronologique du vocabulaire familial

Les deux principales exceptions ne remettent pas en cause ce mouvement de fond. En 1945-46, 26 des 33 occurrences de « frères » sont au sens de la fraternité humaine (« frères de misère »). Si l'on refait le calcul en ne gardant que les 7 occurrences relevant du sens familial, le score fortement négatif (-8,9) rejoint celui des autres termes de parenté. Quant à « papa », 13 de ses 14 occurrences de 1945-46 sont concentrées dans un seul texte (1945e). Si l'on écarte ce texte comme non représentatif de la partie de ce point de vue, le score recalculé s'accorde également au mouvement négatif d'ensemble (-1,7).

C'est bien dans une réélaboration tardive que le vocabulaire de la famille intervient dans des témoignages auto-biographiques, avec la reconstruction d'une saga familiale le plus souvent détruite. Certes, parce que la Shoah est désormais comprise comme un génocide, la mémoire du rescapé des années 2000 convoque sa généalogie. Mais aussi, dans les camps, l'horizon d'attente, c'était la mort. A la Libération c'est la reconstruction, le retour à la vie. Le recul des années permet ensuite à l'horizon personnel de se redéployer, et à la question du lien familial de se poser, vers les descendants et, par ricochet, vers les ascendants. Davantage que par la référence à la souffrance, au corps, du récit descriptif des premiers témoignages, dans ces témoignages plus tardifs l'émotionnel rejaillit ainsi par l'irruption de la famille, aval et amont.

## 4. Conclusion

À la « concurrence des mémoires », cette contribution a essayé d'opposer la complémentarité des lexiques. Si le souvenir se construit et se reconstruit dans et par la matérialité langagière – chaque mot étant à la fois trace et média, quasi physiques, de l'événement reconvoqué –, l'étude des lexiques mise en œuvre pour faire advenir le témoignage doit être envisagée avec le plus grand soin méthodologique.

Les récits de vie à  $n$  années de distance semblent devoir répondre au double impératif de (faire) revivre l'événement et de le rendre intelligible : nous sommes là au cœur à la fois de *l'identité narrative* du témoin [Ricoeur 2000] et de la narration historique.

Ce double impératif passe selon nous par deux lexiques que l'on ne qualifiera pas de concurrents mais de complémentaires. Et cette étude a montré que la portion respective de ces lexiques variait au fil des ans pour décrire une historicisation originale de la Shoah ; cette historicisation se repère au cœur même des récits de vie.

D'abord sensoriel, et par là sans doute égocentré, le témoignage se désincarne pour devenir plus rationnel, plus collectif. L'intelligibilité de l'événement passe par une rationalisation et la mémoire sensitive laisse la place à une mémoire réélaborée.

Reste alors, enfin, l'essentiel. Toujours plus cruel au fil du temps : la peur d'oublier et le souvenir impérieux de témoigner (fig. 13)<sup>15</sup>.

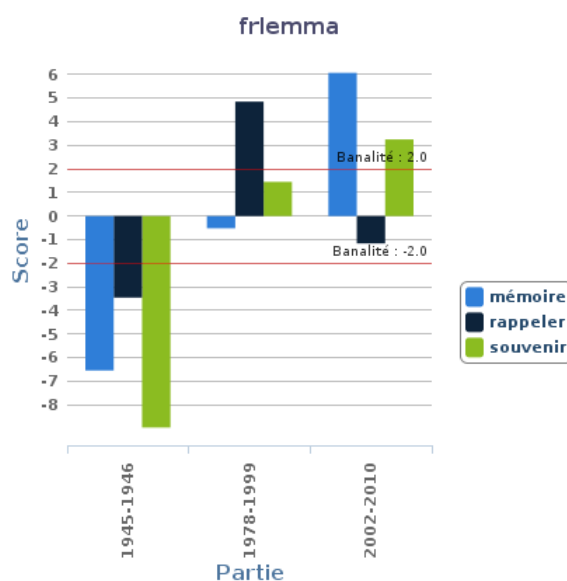


Figure 13. Distribution chronologique des termes de la mémoire

## Références

- Heiden S., Magué J.-Ph., Pincemin B. (2010) « TXM : Une plateforme logicielle open-source pour la textométrie -conception et développement », in Sergio Bolasco, Isabella Chiari, Luca Giuliano (eds), *JADT 2010*, Edizioni Universitarie di Lettere Economia Diritto, p. 1021-1031.
- Joutard Ph. (2013). *Histoire et mémoires, conflits et alliance*. La Découverte.
- Mayaffre D. (2006) « Faut-il prendre en compte la composition grammaticale des textes dans le calcul des spécificités lexicales ? Tests logométriques appliqués au discours présidentiel sous la V<sup>ème</sup> République », in J.-M. Viprey (éd.), *JADT'06*, Besançon : PU de Franche-Comté, p. 677-685.
- Mayaffre D. et Ben Hamed M. (2014) « Récits de mort et souvenir traumatique. Trames et traces lexicales des témoignages sur la Shoah », *Argumentation et Analyse du Discours*, 13, <http://aad.revues.org/1836>.
- Metwally H. (à paraître). *Analyse thématique du Monde diplomatique. Parcours logométriques et traitement logiciel d'un grand corpus*. Thèse de doctorat – Université de Nice.
- Ricoeur P. (2000 rééd.). *La mémoire, l'histoire, l'oubli*. Seuil.

<sup>15</sup> Dans le corpus c'est Simon Grinbaud qui exprime en 1986 le plus vivement l'impératif par le titre de son témoignage : *XI<sup>e</sup> commandement : « Tu n'oublieras point »*