

# Compétences à l'écrit des adultes et événements marquants de l'enfance : le traitement de l'enquête *Information et vie quotidienne* à l'aide des méthodes de la statistique textuelle

Élodie Baril \* et France Guérin-Pace \*\*

Au-delà du dispositif innovant visant à mesurer les compétences à l'écrit des adultes dans une enquête de la statistique publique, les enquêtes *Information et vie quotidienne* (IVQ 2004 et 2011) recueillent un ensemble de données relatives au parcours de vie des personnes. Plutôt que d'introduire un florilège d'événements marquants de l'enfance, nous avons fait le choix dans l'enquête de 2011 d'en restreindre le nombre et de laisser émerger le non-dit par un questionnement ouvert. Près d'un quart des répondants déclarent ainsi un « autre événement marquant » de leur enfance en réponse à cette question. Parmi eux, un sur deux évoque un décès, un sur dix un événement lié à un problème de santé, et dans les mêmes proportions une situation de violence vécue durant l'enfance. Le corpus de plus de 15 000 mots ainsi recueilli est traité par les méthodes de la statistique textuelle, éprouvées depuis de nombreuses années en sciences sociales. La proximité entre les mots employés pour dire les événements et les caractéristiques sociodémographiques des répondants révèle une première dimension d'organisation du vocabulaire en fonction de l'âge et du niveau d'éducation. Une typologie détaillée des événements marquants de l'enfance est ensuite élaborée, prenant soin de distinguer les acteurs en présence (*Ego*, fratrie, parents,...) : elle comprend neuf catégories, notamment les décès, la maladie, les accidents, la violence, les événements collectifs (guerres, catastrophes naturelles,...), la scolarité et d'autres événements familiaux (divorce-séparation,...) et de nombreuses modalités. S'appuyant sur cette typologie, une dernière analyse met en évidence la diversité des situations individuelles, qui associent éléments des parcours biographiques, conditions de vie et compétences à l'écrit. Un seul profil (7 % des répondants) conjugue situations d'illettrisme et enfance difficile dans un contexte familial relativement violent.

## Rappel :

Les jugements et opinions exprimés par les auteurs n'engagent qu'eux mêmes, et non les institutions auxquelles ils appartiennent, ni a fortiori l'Insee.

Codes JEL : C38, I24, Z13.

Mots clés : enquête *Information et vie quotidienne*, événements marquants, statistique textuelle, post-codification, parcours individuels.

\* Élodie Baril est chargée d'études à l'Ined (Service des méthodes statistiques).

\*\* France Guérin-Pace est directrice de recherche à l'Ined (UR12 « Identités et territoires des populations »).

L'enquête *Information et vie quotidienne (IVQ)* a été conduite par l'Insee en 2004, puis reconduite en 2011, en partenariat avec différentes institutions<sup>1</sup>. L'objectif de cette enquête est à la fois d'évaluer le nombre de personnes en difficulté dans leur rapport à l'écrit et de produire une typologie fine des différentes formes de compétences de ces personnes (Besse et Guérin-Pace, 2002) mais aussi de mettre en évidence des déterminants, parmi lesquels des éléments des parcours individuels qui ont pu favoriser les difficultés éprouvées face à l'écrit (Guérin-Pace, 2009). Pour cela, nous avons introduit dès l'enquête de 2004 un dispositif de recueil d'éléments du passé, susceptibles d'avoir pu jouer un rôle sur le processus d'apprentissage des compétences de base, dans l'enfance de l'enquêté.

Un ensemble d'éléments factuels comme le lieu de vie au moment du début de la scolarisation, l'âge d'entrée à l'école, la langue d'apprentissage et les langues parlées à la maison, ont été recueillis afin de différencier, dans un premier temps, les difficultés qui pouvaient relever d'une situation d'illettrisme à proprement parler de celles qui correspondaient à l'apprentissage d'une langue étrangère. En effet, une situation d'illettrisme qualifie des personnes « *qui ne parviennent pas à lire et comprendre un texte portant sur des situations de leur vie quotidienne, et/ou ne parviennent pas à écrire pour transmettre des informations simples* » (ANLCI) et ne peut donc s'appliquer qu'à des personnes qui « *ont suivi en totalité ou en partie le cycle de l'école primaire française* » (ATD Quart Monde, 1981).

Ces informations ont été complétées, dès la première enquête de 2004, par des éléments plus subjectifs relatifs au parcours individuel, tels que la perception par *Ego* (le répondant) de la situation financière familiale<sup>2</sup> ou encore l'existence ou non de pratiques de lecture durant l'enfance (entre 8 et 12 ans). Un ensemble d'événements auxquels *Ego* a eu éventuellement à faire face durant son enfance a ensuite été proposé par la question : « *Avant vos dix-huit ans avez-vous connu les événements suivants ?* ». Il s'agit d'événements que l'on suppose marquants pour *Ego*, relevant de la sphère familiale (décès d'un frère ou d'une sœur, mésentente des parents, séparation ou divorce, chômage de longue durée d'un des parents) mais aussi du domaine de la santé, avec l'hypothèse que « *des problèmes de santé graves* » (conduisant à être absent de l'école plus de deux mois) ont pu interférer sur la réussite dans le parcours scolaire. Dans

l'enquête de 2004, nous avons tenté d'aborder les situations éventuelles de violence en introduisant deux événements relatifs aux cadres scolaire et familial : « *violences à l'école subies de la part d'autres élèves* » et « *violences familiales* ». En 2011, nous avons fait le choix de ne pas poser ces questions et de les laisser émerger au moyen d'une question ouverte introduite en fin du module biographique.

### **Les événements marquants de l'enfance recueillis de manière fermée dans l'enquête IVQ**

L'intérêt de recueillir des événements considérés comme marquants de l'enfance sous forme de question fermée est de pouvoir effectuer des comparaisons systématiques, dans la mesure où tous les enquêtés répondent à une même question. En 2011, sur l'ensemble des répondants, 42 % sont concernés par au moins un des huit événements<sup>3</sup> proposés dans l'enquête et près d'un enquêté sur cinq (19 %) citent deux événements ou plus (cf. tableau 1). Les événements déclarés dans les deux enquêtes montrent une grande stabilité de survenue (cf. figure 1). Ainsi, dans les deux éditions de l'enquête, une situation de « *mauvaise entente entre les parents* » se démarque comme étant la situation la plus vécue durant l'enfance (21 % en 2004 et 20 % en 2011), suivie de l'événement « *séparation ou divorce des parents* » vécu par 12,5 % des répondants en 2004 et 14 % des enquêtés en 2011. De la même manière, la proportion d'individus dont l'un des parents a vécu une longue période de chômage ou qui a connu un problème de santé durant l'enfance représente 8 à 10 % en 2004 et en 2011. En revanche, avoir vécu durant l'enfance une situation d'alcoolisme dans le foyer est plus fréquent en 2011, avec près d'un répondant sur dix contre 8,4 % en 2004. À l'inverse, la part des répondants ayant connu un décès au sein de la fratrie ou d'un de leurs parents diminue très nettement entre les deux dates. On peut expliquer cette baisse par un recul de la mortalité et surtout par une réduction de la taille des fratries ayant pour conséquence une moindre exposition au

1. Notamment l'Agence nationale de lutte contre l'illettrisme (ANLCI), le ministère de l'Éducation nationale, le Ministère des Affaires sociales et l'Institut national des études démographiques (se reporter à l'introduction du numéro).

2. La question posée est la suivante : « *Lorsque vous aviez entre 8 et 12 ans, votre famille était plutôt, d'un point de vue financier ? Très à l'aise, à l'aise, juste mais on y arrivait, on ne s'en sortait pas.* »

3. Dans les deux enquêtes, la question sur le décès des parents était posée dans le module sur la famille. La variable « *décès d'un des parents durant l'enfance* » a ainsi été reconstruite à partir de la date de décès éventuel du père et de la mère.

risque. Parmi ces éléments du passé, certains apparaissent comme une dimension explicative potentielle des difficultés d'apprentissage. Dans l'enquête de 2004, nous avons ainsi montré que plusieurs événements pouvaient plus particulièrement être mis en relation avec des difficultés d'apprentissage de l'écrit, à savoir la survenue d'un décès dans la fratrie durant l'enfance ou encore des violences vécues dans le cadre scolaire (Guérin-Pace, 2009)<sup>4</sup>.

D'une manière générale, la collecte d'événements vécus est toujours une opération délicate à mettre en place dans une enquête par questionnaire. Or de nombreux travaux ont montré que ces éléments du parcours possédaient une

dimension explicative très pertinente, qu'ils s'agissent de points d'inflexion, de ruptures au sein des parcours biographiques ou d'événements ponctuels sans conséquence à long terme (Guérin-Pace, 2009). Mais les événements recueillis de manière « fermée » ne permettent pas d'aborder tous les thèmes notamment ceux portant sur des sujets sensibles (cas de violence par exemple). Le recueil sous forme d'énumération devient en effet vite intrusif, parfois déplacé, si les personnes ne sont pas concernées, et peut donner un ton quelque peu misérabiliste

4. Cette analyse ne peut être reconduite telle quelle pour 2011 dans la mesure où les événements recueillis ou le mode de recueil diffèrent d'une enquête à l'autre.

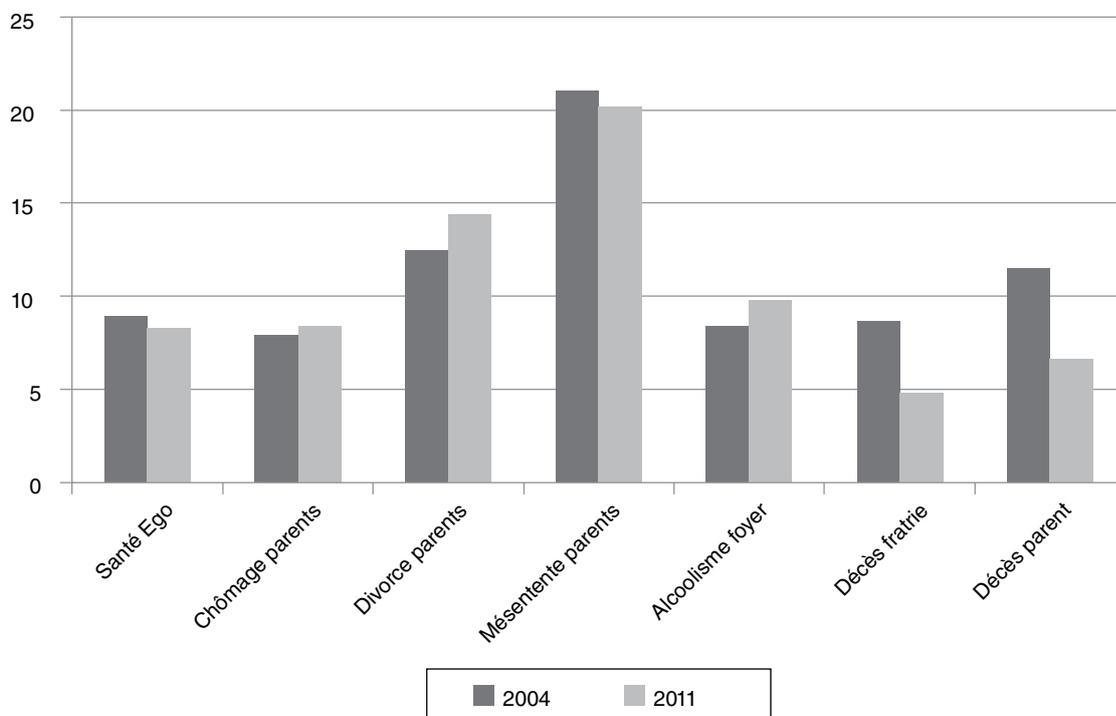
Tableau 1  
Nombre d'événements marquants de l'enfance déclarés en 2011

Nombre d'événements	En %				
	aucun	1 seul	2 ou 3	4 ou plus	Ensemble
	58,4	22,5	17,2	2,0	100,0

Champ : personnes de 18 à 65 ans en France métropolitaine, ensemble des répondants (13 763), effectifs pondérés.  
Source : Insee, enquête Information et vie quotidienne (IVQ) 2011.

Figure 1  
Événements marquants de l'enfance déclarés en 2004 et 2011

En %



Lecture : l'événement « mésentente des parents » a été vécu par 21 % des répondants.

Champ : personnes de 18 à 65 ans en 2011 et personnes de 18 à 65 ans en 2004 en France métropolitaine, ensemble des répondants à chacune des enquêtes, effectifs pondérés.

Source : Insee, enquête Information et vie quotidienne (IVQ) 2011.

à l'enquête. Par ailleurs, par cette démarche, on fait l'hypothèse de la nature *a priori* traumatisante d'un événement sans savoir si *Ego* l'a vécu comme tel durant son enfance (Laborde *et al.*, 2007). Pour toutes ces raisons, nous avons réduit dans l'enquête de 2011 la liste d'événements proposés et en particulier retiré les deux événements relatifs aux éventuelles violences vécues dans les cadres familial et scolaire. La question ouverte « Avez-vous connu un autre événement marquant durant votre enfance ? Si oui, lequel ? » située en fin du module biographique donne la possibilité au répondant de déclarer ou non un autre événement qui a pu jouer un rôle important sur sa construction personnelle et/ou, comme on le verra, de revenir sur un événement déjà déclaré, pour en souligner l'importance dans son parcours.

Nous montrons dans cet article de quelle manière il est possible de recueillir et traiter ce matériau qualitatif au moyen des méthodes de la statistique textuelle, développées dans les années 90 (Guérin-Pace, 1997 et Garnier et Guérin-Pace, 2010). Étape par étape, nous analysons ce corpus d'événements afin d'élaborer une grille d'analyse détaillée mêlant les différents modes de recueil (ouvert et fermé) ; cette typologie est ensuite utilisée pour étudier plus largement la diversité des parcours individuels en lien avec les compétences face à l'écrit, mesurées par l'enquête.

### Questionner de manière ouverte les événements marquants de l'enfance

Sur 13 763 répondants à l'enquête *IVQ* (2011), 3 167 individus (23 %) citent un autre

événement qu'ils considèrent comme ayant été marquant durant leur enfance. Le contenu des réponses est généralement court, variant de 1 à 15 mots et rarement formulé sous forme de phrase<sup>5</sup> (cf. tableau 2). Les réponses désignent le plus souvent des événements de nature individuelle, dans des registres très différents, qui vont des difficultés scolaires à des violences subies dans différents contextes. Il peut s'agir aussi d'événements collectifs ou d'événements historiques. La probabilité de citer un « autre événement survenu durant l'enfance » s'accroît avec le nombre d'événements déclarés en réponse aux questions fermées : 15 % des personnes qui n'ont pas cité d'événement parmi les 8 proposés citent un « autre événement » et cette proportion s'élève à 42 % pour celles qui ont répondu de manière positive à quatre événements ou plus.

Le corpus des 3 167 réponses à la question ouverte se compose de près de 16 000 mots parmi lesquels environ 10 % de mots distincts (cf. tableau 3). Dans une première étape de traitement du corpus, on distingue les mots « plein » des « mots-outils » (article, démonstratif, noms propres, etc.). Plus généralement, on procède à une opération de lemmatisation qui consiste à répartir les mots au sein de catégories grammaticales et à effectuer des regroupements de mots pour en accroître la fréquence (cf. encadré) (Garnier et Guérin-Pace, 2010).

La lecture du lexique associé aux « autres événements marquants de l'enfance » montre que

5. La consigne donnée aux enquêteurs était de transcrire la réponse mot à mot.

Tableau 2  
Exemples de réponses à la question ouverte « Autre événement marquant durant l'enfance ? »

Autre événement marquant de l'enfance
Alcoolisme et séparation
Maladie de ma mère
Le divorce de mes parents
Décès de mon père
Décès des grands-parents
Accident de voiture de mes parents
Violence morale et physique
Mai 68
Complication de santé après avoir avalé un clou
Incarcération de mon papa

Champ : 3 167 individus ayant cité un « autre événement marquant de l'enfance » ; personnes de 18 à 65 ans en France métropolitaine. Source : Insee, enquête Information et vie quotidienne (IVQ) 2011.

la survenue d'un « décès » dans l'entourage proche d'*Ego* constitue, de loin, l'événement le plus marquant de l'enfance (plus de 40 % des personnes qui déclarent un « autre événement marquant » citent un « décès » ou évoquent la « mort »), qu'il s'agisse d'un parent, d'un

membre de la fratrie ou d'un autre proche (cf. tableau 4). La survenue d'une « maladie », qu'elle touche *Ego* ou l'un de ses proches figure dans plus d'une réponse sur cinq, devant la citation du mot « violence » qui est présent dans 158 réponses (5 % des

Tableau 3  
Description du corpus constitué des « autres événements marquants de l'enfance »

Nombre de réponses	3 167
Nombre de mots du corpus	15 993
Nombre de mots distincts	1 564
Nombre de mots distincts après lemmatisation	1 370
Nombre de mots « plein »	1 180
Nombre de mots-outils	190
Nombre de hapax (1)	791

1. Le terme de hapax désigne un mot du corpus qui n'apparaît qu'une seule fois.

Champ : 3 167 individus ayant cité un autre événement marquant de l'enfance ; personnes de 18 à 65 ans en France métropolitaine. Source : Insee, enquête Information et vie quotidienne (IVQ) 2011.

#### Encadré

### MÉTHODES ET OUTILS DE LA STATISTIQUE TEXTUELLE

Les méthodes de la statistique textuelle sont nées dans les années 1990 de la rencontre entre plusieurs disciplines : la linguistique, l'analyse du discours, l'informatique, le traitement des enquêtes (Lebart et Salem, 1994). Elles trouvent aujourd'hui leurs applications dans des domaines très variés qui ont en commun d'analyser des textes. Ces textes peuvent être de nature variée : entretiens, ouvrages, écrits historiques ou littéraires, discours politiques, réponses à une question ouverte, etc.

On peut distinguer deux grandes familles de traitement des données textuelles : la première s'intéresse à la comparaison de textes sur la base d'une étude quantitative du vocabulaire. Elle privilégie la forme des textes, leur style, en faisant abstraction de leur contenu (comparaison et/ou du vocabulaire de différents auteurs, attribution d'un écrit historique ou littéraire, etc.). La seconde famille de traitements est celle qui nous intéresse le plus en sciences sociales. Elle analyse le contenu des textes plutôt que leur forme afin d'en extraire le sens. L'application la plus courante est l'analyse d'un corpus d'entretiens ou de réponses à une question ouverte, pour laquelle ces méthodes ont fait leurs preuves.

Plusieurs logiciels permettent de mettre en œuvre les méthodes de la statistique textuelle (Spad, Alceste, Lexico, Iramuteq, etc.). Certains, en particulier Spad, sont plus adaptés au traitement de textes courts (réponses à une question ouverte, mots associés, titres d'articles, flux RSS, etc.) et permettent par l'application de méthodes d'analyse factorielle (AFC) de mettre en relation les mots du corpus avec les

caractéristiques des répondants ; d'autres, en particulier Alceste, sont plus adaptés aux textes longs (entretiens) et ont pour objectif de structurer le contenu des textes par une classification des profils lexicaux.

Indépendamment du logiciel choisi, les étapes successives de traitement des corpus de données textuelles sont communes (Garnier et Guérin-Pace, 2010) :

- Construction du vocabulaire associé au corpus et comptage des occurrences des mots (cf. tableau 4).

- Étape facultative de *lemmatisation* du vocabulaire (manuelle, assistée, ou automatique selon le logiciel) qui consiste à répartir les mots selon des catégories grammaticales (nom commun, articles, verbes, etc.) et à regrouper les mots qui correspondent à une même racine (exemples : « gentille » et « gentil », « maison » et « maisons », « jouer » et toutes ses formes conjuguées « jouais », « jouent », etc.), afin de réduire la dispersion du vocabulaire.

- Construction des tableaux lexicaux (tableau lexical entier ou tableau lexical agrégé) selon l'objectif fixé. Le tableau lexical agrégé (TLA) est un tableau de contingence qui croise les mots du corpus avec des caractéristiques des répondants choisies *a priori*. La figure II donne la représentation des résultats d'une ACM effectuée sur un tel tableau de contingence. Le tableau lexical entier (TLE) est un tableau disjonctif complet (contient des 0 et 1) qui croise les réponses à la question ouverte (dans notre exemple) avec les mots du corpus. Une classification descendante hiérarchique (CDH) effectuée sur ce tableau permet de mettre à jour des thématiques (univers lexicaux) du corpus (cf. figure III).

réponses). Le premier événement collectif cité est la guerre, le plus souvent il s'agit de la guerre d'Algérie<sup>6</sup>. Enfin, 2 % des personnes qui ont déclaré avoir connu un autre événement marquant ne souhaitent pas le nommer.

Nous avons regardé quelles étaient les personnes concernées par ces événements. Le père est celle qui est le plus communément associée aux événements marquants cités, que ce soit en lien avec un décès, une maladie ou un acte de violence. Plus généralement, les personnages masculins sont plus souvent mentionnés en lien avec ces événements. Les personnes de l'entourage le plus fréquemment signalées appartiennent toutes à la sphère familiale, qu'il s'agisse des grands-parents, d'un membre de la fratrie ou d'un oncle ou d'une tante.

Les mots utilisés une seule fois (hapax) révèlent des réponses originales d'événements marquants durant l'enfance. Dans notre corpus, 791 mots ne sont cités qu'une seule fois, tels que « désintoxication », « homosexualité », ou encore « dyslexie ». Il est par ailleurs intéressant de relever que certains enquêtés reviennent sur un événement déjà cité en réponse à une des questions fermées, comme le décès d'un proche (parents, fratrie) ou le divorce des parents.

À titre d'exemple, 22 personnes évoquent le décès d'un frère ou d'une sœur, qu'ils ont déjà évoqué en réponse à la question fermée et sept personnes ne citent cet événement qu'en réponse à la question ouverte « autre événement marquant ». Il pourrait s'agir éventuellement d'un demi-frère ou d'une demi-sœur. De même, sur les 1 893 personnes ayant déclaré la séparation ou le divorce de leur parents, 23 l'ont répété dans l'événement « autre ». Cette constatation est encore plus fréquente en ce qui concerne le décès d'un parent. En effet sur 294 personnes ayant vécu le décès de leur mère pendant l'enfance, la moitié l'ont cité comme autre événement marquant (cf. note 3).

Au-delà de cette lecture du lexique associé aux événements, nous avons voulu explorer plus finement le contenu de ce corpus par les méthodes de la statistique textuelle (Lebart et Salem, 1994 ; voir encadré). L'objectif de notre première analyse est d'extraire les principaux axes de différenciation du vocabulaire du corpus et de les rattacher aux caractéristiques sociales et démographiques des répondants (âge, sexe,

6. Les personnes interrogées sont âgées de 18 à 65 ans en 2011 et n'ont donc pas connu la seconde guerre mondiale.

Tableau 4  
**Lexique des événements et des personnes de l'entourage les plus fréquemment cités**

Événements	Nombre d'occurrences	Personnes	Nombre d'occurrences
Décès	1 280	Père	620
Maladie	179	Mère	338
Violence	158	Grand-mère	231
Accident	143	Grand-père	211
Mort	135	Frère	140
Guerre	112	Grands-parents	131
Perte	52	Parents	142
Séparation	40	Maman	110
Départ	36	Papa	93
Suicide	35	Oncle	79
Naissance	32	Sœur	70
Dépression	28	Tante	41
Déménagement	28	Cousin	35
Alcoolisme	23	Enfant	30
Divorce	23		
Hospitalisation	23		
Handicap	20		

Lecture : parmi les personnes ayant évoqué un « autre événement marquant durant l'enfance », le mot « décès » a été cité 1 280 fois et le mot « père » 620 fois.

Champ : 3 167 individus ayant cité un « autre » événement marquant de l'enfance ; personnes de 18 à 65 ans en France métropolitaine. Source : Insee, enquête Information et vie quotidienne (IVQ) 2011.

milieu social, profession, région, etc.) (Garnier et Guérin-Pace, 2010).

Pour cela, il est nécessaire de construire un tableau lexical agrégé (TLA) croisant le vocabulaire lemmatisé du corpus des événements avec les caractéristiques sociodémographiques des répondants et d'autres informations recueillies dans l'enquête *IVQ* 2011. En particulier, nous avons introduit les compétences face à l'écrit telles qu'elles ont été évaluées par le dispositif d'enquête, selon cinq modalités de niveau de difficulté croissant (Murat, 2005). Au-delà de la maîtrise de l'écrit, nous disposons d'un ensemble de variables actives qui caractérisent l'individu au moment de l'enquête : âge, sexe, statut d'occupation du logement, niveau d'études atteint (primaire, secondaire technologique ou professionnel, secondaire général, supérieur), des variables relatives à la socialisation (fréquence des visites à la famille et aux amis, fréquence de contacts avec le voisinage), à la perception du niveau de maîtrise de la lecture du français, ainsi que le nombre d'événements cités parmi les huit recueillis sous forme de question fermée (cf. tableau 1). Une analyse multidimensionnelle de ce tableau lexical agrégé<sup>7</sup> (ACM) permet de mettre en évidence les proximités entre les mots du lexique et les caractéristiques des répondants (Guérin-Pace, 1997) (cf. figure II).

### **Dire les événements marquants : un effet d'âge et d'éducation**

La première dimension d'organisation du vocabulaire est régie par l'âge et le niveau d'éducation des enquêtés (cf. figure II, axe 1 contribuant à 12,1 % de la variance), alors que la seconde se structure non plus autour de descripteurs sociodémographiques mais autour du nombre d'événements vécus durant l'enfance (cf. figure II, axe 2 contribuant à 8,5 % de la variance).

Un premier univers lexical (quadrant inférieur droit) se dessine autour des âges élevés et des personnes peu ou pas scolarisées, qui éprouvent des difficultés face à l'écrit et déclarent plus généralement des difficultés avec la langue française (« impossible de faire les exercices »). Il s'agit de personnes avec un niveau d'éducation très faible (« jamais scolarisées », « primaire »), en partie de migrants arrivés en France enfants, qui évoquent cette migration comme un événement marquant (« arrivée », « venue », « France »). Les mots proches de ces

caractéristiques désignent le plus souvent des événements collectifs « guerre », « indépendance » « Algérie », « tremblement de terre », « assassinat de Kennedy », qui ont marqué cette génération. On relève aussi un ensemble de mots qui se réfèrent à des événements familiaux (« placement », « nourrice », « assistance publique », « manque », « financier ») et qui reflètent une enfance difficile.

Dans le quadrant supérieur droit, un autre monde lexical se dessine autour d'un ensemble de mots qui désignent des « violences », proche des modalités « 4 événements et plus », « femme » et « graves difficultés » face à l'écrit. Ces violences sont exprimées en termes de « maltraitance », « battu », « coups », « méchant » et de faits (« viol », « physique », « inceste ») dans un contexte le plus souvent « familial », de grande précarité « misère », « pauvreté » et parfois d'« alcoolisme ». À proximité on relève aussi des mots qui désignent des événements de type « cataclysme », de nature individuelle ou collective (« cyclone », « terrorisme »). Les variables relatives à la sociabilité dénotent une absence de relations familiales ou amicales (« Famille\_jamais », « Amis\_jamais »). La nature violente des événements se modifie quelque peu au fur et à mesure que l'on se déplace dans la partie gauche du graphique où l'on relève les mots suivants : « tentative », « sexuel », « moqueries », « attouchements » et des termes qui traduisent un refus de fournir une réponse (« refus », « répondre », « dire »). Proche de cet univers lexical, on relève d'autres événements familiaux relatifs aux parents (« divorce », « remariage », « découverte (père biologique) », etc.), proches des âges les plus jeunes (« 20 à 40 ans ») et empreints d'une forte composante amicale dans la sociabilité (« Amis+++ », cf. note de la figure II).

En se déplaçant vers le quadrant inférieur gauche, on trouve un lexique en lien avec la santé (« hospitalisation », « dépression », « cancer », etc.) et un ensemble d'événements qui semblent intervenir de manière plus brutale et ponctuelle, tels que un « assassinat », un « attentat », un « suicide », un « décès », un « accident » ou encore un « handicap » ou une « fausse couche » mais aussi des événements marquants comme le séjour en « prison » d'un

7. Cette analyse a été réalisée avec le module Text Mining du logiciel d'analyse des données Spad. Il permet de traiter les variables textuelles comme des variables particulières en proposant des modes de traitement adaptés (AFC, ACM sur le tableau lexical agrégé).



proche ou un « déménagement », à côté des caractéristiques individuelles telles que études supérieures (« Supérieur »).

Cette première analyse a permis de mettre en regard la nature des événements et la manière de les dire. Ainsi les événements déclarés par les personnes plus âgées et la présence de mots tels que « assistance » « publique » correspondent à des situations vécues plus fréquemment à leur époque. De la même manière, les événements collectifs marquants déclarés sont spécifiques d'une époque donnée : « rapatriement » d'« Algérie », « Kennedy », etc. Au-delà de cet effet d'âge, on observe un effet lié au niveau d'éducation des répondants : pour les diplômés, les événements déclarés comme marquants sont le plus souvent de type accidentel ou imprévisible « cancer », « assassinat », « fausse-couche », « attentat » alors que les personnes moins éduquées évoquent davantage des situations familiales particulièrement difficiles à travers les mots « maltraitance », « abandon », « inceste » qui semblent s'être accompagnées d'une mise à distance de la famille. L'intérêt de recourir aux méthodes de la statistique textuelle pour traiter cette question ouverte est de laisser émerger d'une part les différences entre le contenu des événements et d'autre part leur expression selon les locuteurs.

Si l'on souhaite maintenant comparer des événements entre différentes dates ou construire une grille d'événements applicables à plusieurs enquêtes, indépendamment de leur mode de recueil, il est nécessaire d'élaborer une catégorisation de ces événements. En premier lieu, la distinction entre les modes de recueil des événements, question ouverte ou fermée, doit être gommée au profit d'une grille de lecture d'ensemble.

Nous avons à nouveau eu recours aux outils de la statistique textuelle pour nous aider à construire cette grille en intervenant le moins possible sur le corpus. En effet, dans une démarche classique de post-codification d'une question ouverte, le rôle du codeur est important : les réponses sont souvent simplifiées et parfois déformées par l'interprétation du chercheur (Lebart et Salem, 1992). Cette démarche de post-codification « assistée » que nous proposons ici par les outils de la statistique textuelle tente de limiter les biais liés à une intervention extérieure et vise à étudier ces événements recueillis au moyen d'une question ouverte de la même manière que les réponses aux événements proposés par questions fermées.

### Post-coder les autres événements marquants par les méthodes de la statistique textuelle

Dans un premier temps, il est nécessaire de construire un *tableau lexical entier* (TLE) qui croise les réponses à la question ouverte avec l'ensemble des mots du vocabulaire du corpus. À l'intersection d'une réponse *i* et d'un mot *j* figure « 1 » si le mot *j* est présent dans la réponse *i*, « 0 » sinon. Le tableau lexical entier est qualifié de « clairsemé » ou « hypercreux » car il contient plus de 95 % de 0 (Lebart et Salem, 1994). Il s'agit ici non plus de relier les événements aux caractéristiques des locuteurs, comme dans l'analyse précédente, mais de structurer les réponses en fonction de leur profil lexical (répartition des mots dans les réponses). Une méthode de classification de type descendante hiérarchique effectuée sur ce tableau permet de mettre à jour des univers lexicaux (Reinert, 1993). Initialement toutes les réponses (lignes du TLE) sont regroupées en une seule classe, puis subdivisées à chaque étape en sous-classes les plus différentes entre elles. Deux réponses seront d'autant plus proches qu'elles comportent des mots en commun (Garnier et Guérin-Pace, 2010). Près de 90 % des réponses ont ainsi été classées (2 761 réponses sur les 3 167)<sup>8</sup>. Les réponses non classées sont le plus souvent des refus de répondre (environ 16 % des réponses non classées) ou des réponses originales et rares telles que « le tremblement de terre d'Agadir », « un meurtre dans la rue quand j'avais 15 ans », « un film », etc.

La partition finale obtenue est composée de cinq classes (cf. figure III). Chaque classe est décrite par un ensemble de mots surreprésentés dits caractéristiques ou spécifiques dont la lecture d'ensemble permet de donner un sens à l'univers lexical qui se dégage de la classe. On observe une première dichotomie entre une classe qui contient des événements collectifs et un ensemble de classes qui évoquent des événements de nature individuelle. Cette classe (classe 5), la moins nombreuse en termes d'effectif, rassemble un ensemble d'événements marquants tels que la guerre d'Algérie, Mai 1968, mais aussi un vocabulaire relatif aux migrations qui se sont produites durant cette période. Une autre classe se détache ensuite (classe 4), regroupant des mots qui relatent des

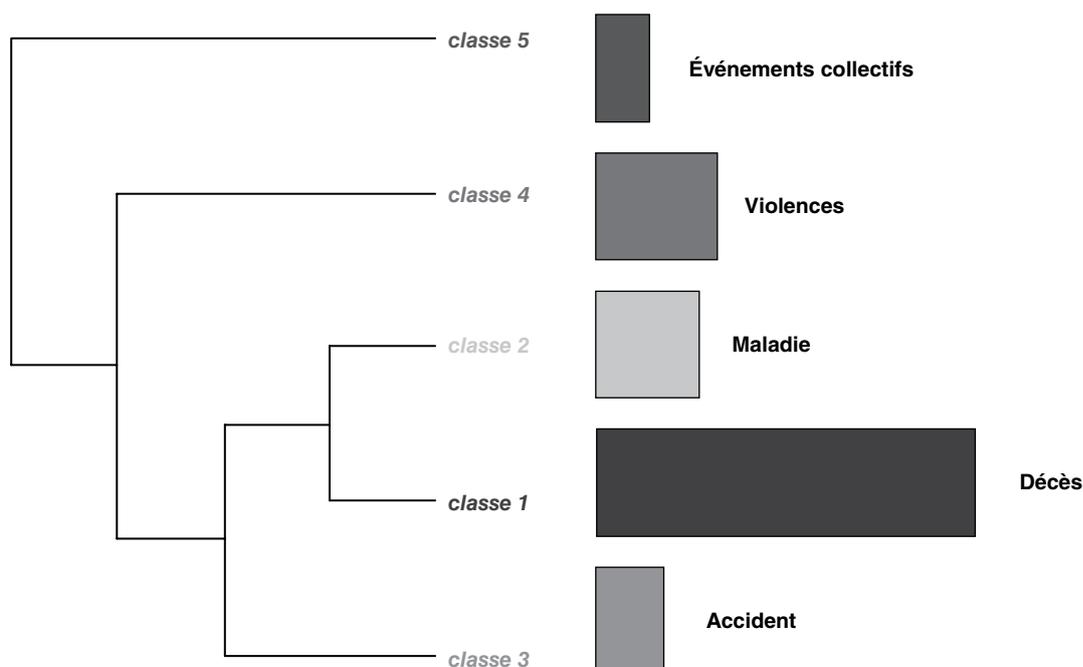
8. Nous utilisons ici le logiciel d'accès libre *Iramuteq*, développé par Pierre Ratinaud (suivant <http://www.iramuteq.org/>). Interface du logiciel de statistique *R*, il s'inspire de la méthode d'analyse textuelle développée par M. Reinert (1983).

faits marquants qui se produisent le plus souvent au sein d'un univers familial ou scolaire. Les trois dernières classes se structurent entre une branche formée des classes « maladie » et « décès » qui se distinguent de la classe « accidents ». En termes d'effectifs, la classe que nous avons intitulée « décès » regroupe plus de la moitié des réponses du corpus (52 %) et se compose d'un vocabulaire qui désigne la « perte », la « mort », ainsi que les personnes de l'entourage qui sont concernées (parents, grands-parents, amis, etc.). De la même manière, la lecture du vocabulaire des classes « maladie » (14 % des réponses) et « accidents » (9,5 %) se réfère à des événements qui concernent la personne enquêtée ou un proche.

Nous avons poursuivi la post-codification de ce corpus en plusieurs étapes successives. Dans un premier temps, nous avons ajouté à ces cinq classes deux classes supplémentaires : une classe intitulée « Refus » regroupant toutes les réponses qui marquent une volonté de ne pas détailler l'événement à l'enquêteur tout en ayant donné une réponse affirmative à la question « Avez-vous connu un autre événement marquant ? », et une classe « Autre » au sein de laquelle nous avons regroupé les réponses

« originales » ou non classées (cf. *supra*). Dans un second temps, nous avons cherché à affiner la typologie par une lecture des réponses spécifiques de chacune des classes (les réponses spécifiques données pour chacune des classes sont celles qui contiennent le plus grand nombre de mots spécifiques). À la lecture de ces réponses spécifiques, on se rend compte que la classe que nous avons intitulée « Violence » rassemble plus largement un ensemble d'événements qui se produisent dans les contextes de la famille ou de l'école (cf. tableau 5). Nous avons conservé uniquement dans cette classe les réponses qui désignent des actes violents, qu'ils soient physiques ou psychologiques, et fait le choix de créer une classe supplémentaire « Événement familial » dans laquelle nous avons regroupé des événements tels que les placements en famille d'accueil ou dans un foyer. À ces événements familiaux nous avons adjoint d'autres événements du même registre tels que le divorce ou la séparation des parents, l'expression d'une « absence » ou le « manque » d'un parent, l'abandon, etc. Par ailleurs, nous avons créé une classe « scolarité » au sein de laquelle nous avons regroupé les difficultés scolaires et fait le choix d'y ajouter les violences déclarées dans le cadre scolaire. Nous obtenons ainsi une

Figure III  
**Une typologie des « autres événements marquants » de l'enfance en 5 classes**



Lecture : la typologie a été obtenue par une méthode de classification de type descendante hiérarchique appliquée au tableau lexical entier (TLE) croisant les réponses à la question ouverte avec l'ensemble des mots du vocabulaire du corpus.  
 Champ : 3 167 individus ayant cité un « autre événement marquant de l'enfance » ; personnes de 18 à 65 ans en France métropolitaine.  
 Source : Insee, enquête Information et vie quotidienne (IVQ) 2011.

Tableau 5

## Les cinq classes d'autres événements marquants vécus durant l'enfance

Thème de la classe	Effectif dans la classe	Effectif total	Fréquence( %)	Chi2	Forme spécifique
Classe 1 Décès (52,3 %)	1 055	1 214	87	1 038	décès
	578	587	99	636	grand
	620	787	79	309	père
	426	525	81	216	mère
	165	200	83	79	parent
	118	131	90	79	mort
	78	84	93	57	oncle
	46	48	96	37	perte
	32	35	91	22	ami
Classe 2 Maladie (13,8 %)	156	177	88	876	maladie
	103	109	95	619	maman
	85	92	92	493	papa
	21	22	96	124	dépression
	26	32	81	123	suicide
	12	13	92	67	cancer
	10	17	59	29	tentative
	5	6	83	24	partie
	5	7	71	20	remariage
Classe 3 Accident (9,5 %)	131	135	97	1 267	accident
	84	111	76	590	frère
	33	33	100	319	voiture
	43	62	69	265	grave
	38	74	51	155	sœur
	20	24	83	154	petit
	20	30	67	115	naissance
	7	7	100	67	circulation
	7	8	88	57	route
Classe 4 Violence (16,9 %)	134	142	94	641	violence
	48	52	92	215	école
	42	44	96	197	familial
	23	25	92	101	conflit
	20	23	87	81	placement
	16	17	94	73	viol
	20	25	80	72	foyer
	14	14	100	69	sexuel
	13	13	100	64	attouchement
Classe 5 Événements collectifs (7,5 %)	109	112	97	1 365	guerre
	84	87	97	1 033	Algérie
	30	31	97	362	Mai (1968)
	21	23	91	236	déménagement
	12	12	100	149	événements
	16	23	70	130	France
	9	9	100	112	civil
	8	8	100	100	Maroc
	7	7	100	87	rapatriement

Note : nous avons reporté dans ce tableau les 10 premiers mots les plus significatifs (p-value inférieure à 0.001) de chacune des classes (par valeur du Chi-2 décroissante).

Lecture : le mot « perte » apparaît à 46 reprises dans la classe 1, sur un total de 48 fois dans l'ensemble du corpus, soit 96 % des occurrences de ce mot. La valeur du Chi-2 associée est de 37 et la probabilité de dire que le mot « perte » est spécifique de la classe 1 alors qu'il ne le serait pas, est nulle.

Champ : 3 167 individus ayant cité un « autre » événement marquant de l'enfance ; personnes de 18 à 65 ans en France métropolitaine.

Source : Insee, enquête Information et vie quotidienne (IVQ) 2011.

classification plus complète en neuf classes thématiques : décès, maladie, accident, événement familial, scolarité, violence, événement collectif, autre et refus.

Dans l'optique d'une post-codification au moyen d'une grille plus détaillée (qui pourrait être réutilisée pour d'autres enquêtes nationales), nous avons cherché à affiner davantage cette classification en précisant les acteurs éventuels impliqués dans les événements. Ainsi, lorsque cela était possible, nous avons identifié la personne concernée par l'événement<sup>9</sup> et son auteur éventuel. Par exemple, au sein de la classe « Maladie » (classe 2), nous avons filtré au moyen d'un vocabulaire familial<sup>10</sup> (père, mère, frère, sœur, tante, ami, etc.) et constitué 4 sous-modalités distinctes selon les personnes concernées (cf. tableau 6). Nous avons procédé de la même manière pour les classes « Décès » et « Accident ». Au sein de la classe « Violence », nous avons encore affiné cette démarche en précisant, outre la personne concernée par ces violences, son auteur éventuel : « parents », « fratrie », « autre proche » et « autre personne ». Ainsi l'intitulé « Violence Ego\_ autre proche » désigne des actes perpétrés envers « Ego » par un membre de la famille (oncle, tante, grands-parents, belle-mère, beau-père, etc.) autre que les parents ou un membre de la fratrie. Lorsque des violences sexuelles sont évoquées, celles-ci ne concernent qu'*Ego*. Nous avons alors distingué les auteurs de ces violences sexuelles en deux modalités distinctes selon qu'elles proviennent d'un membre de la famille ou d'une personne extérieure. On obtient finalement une classification en neuf classes et 43 sous-classes répertoriées dans le tableau 6. Pour chacune d'entre elles, nous avons indiqué quelques réponses illustratives.

Cette typologie constitue une grille de lecture très fine en 43 postes qui peut être utilisée pour recueillir et analyser les événements marquants de l'enfance. Une première distinction s'opère entre les événements de nature individuelle et collective, ces derniers étant rarement collectés dans les enquêtes. Ils représentent cependant 6 % des événements marquants cités par des personnes généralement plus âgés. Il s'agit soit de guerre et d'événement politique ou historique, soit de catastrophes naturelles. Les événements individuels désignent le plus souvent des événements démographiques, en particulier les décès qui représentent près de la moitié des événements marquants cités. Ce sont ensuite les autres événements familiaux qui sont plus fréquemment cités (14 % des autres

événements marquants), puis les événements liés à une maladie ou à une situation une violence (10 %). À la lecture de ces événements de nature individuelle (décès, maladie, accident et violence) et familiale (divorce-séparation,...), il nous est apparu utile de distinguer les auteurs et les personnes concernées par l'événement en affinant la typologie. Plus encore, cela nous semble une piste intéressante à explorer dans des enquêtes davantage centrées sur les populations vulnérables.

Les effectifs des 43 sous-classes résultant de la post-codification étant relativement réduits, nous avons poursuivi l'analyse en utilisant la typologie agrégée en 9 postes. On dispose ainsi d'un ensemble de variables relatives aux événements marquants de l'enfance sans distinction du mode de recueil initial dans le questionnaire.

Les éléments biographiques du questionnaire et plus particulièrement le recueil des événements marquants de l'enfance représentent un apport important de l'enquête IVQ. Cette dimension des parcours individuels constitue un facteur explicatif à prendre en compte dans la compréhension des difficultés éprouvées par certains adultes face à l'écrit alors qu'ils ont été scolarisés. C'est pourquoi dans cette dernière partie nous avons souhaité analyser l'ensemble des répondants scolarisés en France en langue française<sup>11</sup> et observer la manière dont s'agençaient les éléments des parcours individuels, les conditions de vie au moment de l'enquête, et les éventuelles difficultés à l'écrit, mesurées à l'âge adulte.

### **La diversité des parcours individuels et du rapport à l'écrit**

Cette dernière dimension est évaluée dans l'enquête *IVQ* par un ensemble de tâches visant à mesurer les compétences des adultes dans les trois domaines fondamentaux de l'écrit (lecture, compréhension orale et production écrite) (Besse *et al.*, 2009). Au-delà de la variable sur les difficultés éventuelles estimées face à l'écrit, nous avons retenu l'ensemble des événements marquants de l'enfance, recueillis

9. Si aucune personne n'est mentionnée dans la réponse, par exemple « accident », « maladie », « violences », on a considéré que ces événements concernaient l'Ego.

10. Nous avons filtré les réponses au moyen d'une recherche lexicale (effectuée avec le logiciel SAS).

11. L'illettrisme ne s'applique qu'aux personnes scolarisées en langue française.

Tableau 6  
Une typologie des « autres événements marquants de l'enfance » en 9 classes

Intitulé	Effectifs (Proportion)	Classe	Effectifs	Intitulé	Exemple de réponses
1 Décès	1 489 (47,1 %)	1.1	619	Décès parents (biologiques ou adoptifs)	« décès de mon père », « mort de ma mère », « décès de la mère biologique et père inconnu », « orphelin »
		1.2	29	Décès fratrie	« décès de ma sœur », « décès d'un demi-frère », « décès de frère et sœur que je n'ai pas connu »
		1.3	801	Décès autre proche	« décès », « décès grands parents », « décès de petite nièce », « le veuvage de ma sœur », « mort d'une copine », « le décès de la nourrice »
		1.4	12	Décès autre personne (personnes plus éloignées ou inconnues mais dont le décès a marqué Ego)	« enterrément d'un professeur à l'âge de 8 ans : très perturbant », « une gamine est tombée devant moi du 4ème j'avais 10 ans », « mort d'un compagnon école »
		1.5	8	Suicide parent(s)	« suicide du père à 4 ans », « suicide des parents »
		1.6	20	Suicide autre proche	« décès de mon grand-père qui s'est suicidé », « suicide d'une copine », « pendaison de ma grand-mère »
2 Maladie	338 (10,7 %)	2.1	81	Maladie Ego	« malade », « problème de santé », « hospitalisations fréquentes », « toxicomanie pour moi-même », « alcoolisme »
		2.2	202	Maladie parent(s)	« père tombé malade maniaco dépressif », « maladie grave de la maman », « graves opérations du père », « handicap à 80 % de ma mère », « invalidité du père », « dépression de la mère »
		2.3	43	Maladie fratrie	« maladie de la petite sœur », « sœur méningite », « la trisomie de mon frère »
		2.4	12	Maladie autre proche	« maladie dans la famille », « infarctus du tonton », « grand-mère malade Alzheimer »
3 Accident	175 (5,5 %)	3.1	72	Accident Ego	« accident », « blessé dans un accident de voiture à l'âge de 10 ans », « brûlures graves »
		3.2	51	Accident parent(s)	« accident de ma mère », « mon père qui a été blessé à son travail »
		3.3	29	Accident fratrie	« accident de la sœur », « coma d'un frère suite à une morsure grave par un chien »
		3.4	11	Accident autre proche	« grave accident d'un copain », « perdu ma famille dans un accident d'avion », « noyade d'un cousin »
		3.5	12	Accident matériel	« incendie de l'exploitation », « maison qui a brûlé », « maison qui a pris feu », « cambriolage et incendie de notre logement », « casser pare-brise voiture de police »
4 Violence	338 (10,7 %)	4.1	55	Violence Ego (concerne l'enquêté, auteur inconnu)	« violence », « violence morale et physique », « enfance malheureuse », « subit une agression », « mariage forcé », « j'étais battue », etc.
		4.2	83	Violence Ego_Parent (concerne l'enquêté et dont l'auteur est un des parents)	« violences père frappait », « a été martyrisée par sa mère », « a été mis à la porte par son père la veille de ses 18 ans », « conflit avec les parents », « brutalité des parents », etc.
		4.3	5	Violence Ego_Fratrie (concerne l'enquêté et auteur un frère ou une sœur)	« battu par le grand frère », « mon frère était méchant », « mauvaise entente avec ses sœurs », etc.
		4.4	18	Violence Ego_Autre proche (concerne l'enquêté et dont l'auteur est un autre proche)	« battu d'un beau-père », « une très mauvaise belle-mère », « coups par famille d'accueil », « je me suis battu avec mon grand-père », etc.
		4.5	14	Violence sexuelle_Famille (envers Ego et dont l'auteur est un membre de la famille)	« violé par son frère à 8 ans », « inceste », « attouchement du père », etc.
		4.6	37	Violence sexuelle_autre (envers Ego dont l'auteur est une personne extérieure à la famille ou non spécifiée)	« a été séquestré et violé », « attouchements sexuels », « violences sexuels », « des attouchements de la part d'un homme », etc.
		4.7	101	Violence familiale (mais qui ne s'adressent pas directement à Ego)	« violences familiales », « violence du père sur le petit frère », « mère battue régulièrement par le père », « violences entre mon père et mon oncle », « conflit des parents », etc.
		4.8	12	Violence Parents (Violence envers un des parents par eux-mêmes (suicide) ou auteur extérieure à la famille)	« agression de mon père nez cassé », « tentative de suicide de ma mère », « parents assassinés », etc.
		4.9	8	Violence_autre_proche (Violence envers un autre proche, suicide ou auteur extérieure)	« mon frère qui s'est fait tuer », « assassinat d'une amie », « tentative de suicide de mon frère », etc.
		4.10	5	Violence_autre_personne (Violence envers une autre personne dont ego est témoin)	« un assassinat », « j'ai vu un homme se faire abattre près d'un manège », « la découverte d'un corps dans un étang », etc.

Tableau 6 (suite)

Intitulé	Effectifs (Proportion)	Classe	Effectifs	Intitulé	Exemple de réponses	
5 Événement collectif	202 (6,4 %)	5.1	121	Guerres et conflits	« la guerre d'Algérie », « guerre puis exil », « terrorisme », « conflit religieux attentat extermination des populations des villages », « conflits armés »	
		5.2	68	Autre événement politique ou historique	« les manifestations de mai 1968 », « la crise pétrolière », « grande famine de 1972 », « indépendance de la Tunisie », « l'arrivée du Général de Gaulle », « Apollo 11 », « décès du président Kennedy », « mur de Berlin »	
		5.3	13	Catastrophe naturelle	« inondation de la maison », « cyclone », « tremblement de terre », « tornade de 1961 »	
6 Scolarité	80 (2,5 %)	6.1	49	Violence scolaire	« violences à l'école », « instituteur qui le frappait parce qu'il était gaucher », « moquerie à l'école primaire », « racket à l'école », « harcèlement physique et moral de la part d'un prof », « bouc émissaire de l'école », « pédophilie d'un directeur d'école »	
		6.2	11	Difficultés scolaires	« grosses difficultés scolaires », « difficultés d'adaptation suite aux changements d'école », « ne parlait pas français en arrivant à l'école », « d'avoir des difficultés pour lire et écrire », « repasser son bac »	
		6.3	20	Autre événement scolaire	« avoir le certificat d'études », « aller à l'école militaire », « la prof que j'avais quand j'étais au primaire », « de pas avoir été à l'école », « changement d'école »	
		7.1	22	Abandon	« abandon de la mère », « l'abandon familial », « née sous X »	
		7.2	17	Adoption ou élevé par un tiers	« a été adopté tout petit », « adoption d'un frère », « élevé par la grand-mère », « mise en nourrice »	
7 Événement familial	450 (14,2 %)	7.3	67	Pension ou foyer (Lieu de vie en dehors du cercle familial)	« orphelinat de 6 à 9 ans », « vivre en foyer », « internat », « en pension chez ma grand-mère », « le fait d'être placé en famille d'accueil », « placement à l'assistance publique », « être placé en institution », « placement à la DDASS »	
		7.4	83	Absence, manque	« absence de ma mère », « manque de père », « départ des frères », « séparation avec mes parents partis en France », « je n'ai pas connu mes parents », « père en prison »	
		7.5	58	Migration	« mon arrivée en France », « le retour du Maroc », « départ d'Algérie », « rapatriement en France », « quitter mon pays », « immigration », « émigration », « exil », « déménagement à Moscou », « changement de pays »	
		7.6	35	Déménagement	« déménagement (uniquement en France sinon dans la classe migration) », « partie de chez moi à 15 ans », « de passer à une maison à une cité »	
		7.7	23	Naissance arrivée	« naissance de mon frère », « l'arrivée de mon frère adoptif », « arrivée du beau-père »	
		7.8	44	Divorce séparation	« divorce des parents », « séparation des parents », « divorce de ma mère », « remariage de mon père »	
		7.9	34	Précarité	« pauvreté », « vivait dans la rue », « misère », « faillite », « difficultés financières extrêmes », « déni de grossesse à 16 et donc j'ai dû arrêter l'école deux mois pour la naissance »	
		7.10	67	Autre événement familial	« accueil de la grand-mère à la maison », « retour dans la famille », « le mariage de ma mère », « apprendre que c'est pas mon père »	
		8 Autre	26 (0,8 %)	8	Autre	« Voyage à Lourdes », « terreur nocturne », « j'ai vu à 10 ans 1 ovni », « un film », « mon homosexualité »
		9 Refus	69 (2,2 %)	9	Refus	« Refus », « ne veut pas le dire », « personnel », « trop indiscret », « privé »
TOTAL	3 167 (100 %)		3 167			

Lecture : la réponse « décès grand parents » figure dans la classe 1 « Décès » (47,1 % de personnes ont cité un décès comme autre événement marquant de leur enfance) et, plus précisément, dans la sous-classe 1.3 « Décès autre proche ».

Champ : 3 167 individus ayant cité un « autre événement marquant de l'enfance » ; personnes de 18 à 65 ans en France métropolitaine.

Source : Insee, enquête Information et vie quotidienne (IVQ) 2011.

de manière fermée (7 variables) et ajouté la variable issue de la post codification en 9 modalités (correspondant aux 9 classes du tableau 6). Par ailleurs nous avons introduit des variables sur les conditions de vie d'*Ego* durant l'enfance (perception par *Ego* de l'aisance financière des parents quand l'enquêté avait 8-12 ans, taille de la fratrie) et le déroulement de sa scolarité (nombre de changements d'école, nombre de redoublements éventuels durant la scolarité en secondaire, ainsi que la perception par *Ego* de son niveau scolaire (classe) au démarrage éventuel de difficultés scolaires<sup>12</sup>, pratiques de lecture durant l'enfance). Des caractéristiques sociodémographiques (âge, sexe, niveau d'étude, statut d'activité, statut d'occupation du logement) ainsi que des variables relatives aux conditions de vie au moment de l'enquête sont également mobilisées : sociabilité de l'enquêté (fréquence des contacts avec la famille, les amis et le voisinage), pratiques quotidiennes de lecture (livres, magazines, journaux) et perception du niveau de lecture en français (très bon, bon, pas bon).

Une classification (CAH)<sup>13</sup> est effectuée sur l'ensemble des individus ayant été scolarisés en France et dont la langue d'apprentissage de la lecture est le français (12 323 individus). Sur les sept classes qui en résultent, trois concernent des personnes qui ont peu ou pas de difficultés face à l'écrit, trois des difficultés modérées et une où les personnes sont en très grande difficulté (cf. tableau 7). Regardons maintenant quels sont les éléments du parcours et les conditions de vie qui émergent de ces profils. L'objectif ici n'est pas de produire un modèle explicatif mais de montrer la diversité des situations individuelles qui allient des situations passées et présentes et de suggérer des pistes d'exploration en ce qui concerne la compréhension des situations de difficulté.

***Pas de difficulté face à l'écrit  
pour des parcours de vie sans heurt  
(Classe 1, 24 %)***

Cette première classe regroupe des personnes d'âge moyen (30 à 49 ans) n'ayant aucune difficulté face à l'écrit (98 % d'entre elles). Dotées d'un diplôme supérieur pour la plupart, elles sont en situation d'emploi (pour 87 % d'entre elles). Elles ont connu dans l'enfance des conditions de vie faciles et sans événement particulier, au sein d'une fratrie peu nombreuse (0 ou 1). Au moment de l'enquête, ces personnes ont une sociabilité développée et des pratiques de

lecture fréquentes, comme c'était le cas dans leur enfance.

***Jeunes sans difficulté face à l'écrit,  
parcours familiaux marqués par une rupture  
familiale (Classe 2, 11 %)***

Cette classe regroupe exclusivement des personnes âgées de 15 à 29 ans, en majorité étudiantes (62 %), de niveau d'études supérieures et d'un milieu social aisé sans difficulté face à l'écrit (95 %). Les événements vécus durant leur enfance concernent plus fréquemment la situation des parents : séparation (+8<sup>14</sup>) ou chômage d'un des parents (+5). Au-delà, c'est leur enfance et leur parcours scolaire qui se sont déroulés sans difficulté, au sein de fratrie peu nombreuse. Leurs pratiques de lecture sont assez régulières, comme durant l'enfance, et leur sociabilité amicale est très développée.

***Peu de difficulté face à l'écrit,  
nombreux événements durant l'enfance,  
conditions de vie précaires (Classe 3, 12 %)***

Les personnes de cette classe sont très peu en difficulté face à l'écrit (90 % d'entre elles n'ont aucune difficulté, 85 % pour l'ensemble de la population). Âgées pour la plupart de moins de 50 ans, elles comptent un peu plus de femmes (+8) et de diplômés du secondaire (+16), avec un parcours scolaire ponctué de changements d'établissements et de redoublements et des difficultés ressenties dès le primaire. C'est un environnement familial précaire au sein de familles nombreuses ayant vécu des événements difficiles durant l'enfance : chômage d'un parent, difficultés financières, alcoolisme dans le foyer, disputes et/ou séparation des parents, maladie d'*Ego*. En réponse à la question ouverte sur un éventuel autre événement marquant, ce sont des situations de violence familiale qui sont le plus fréquemment citées (14 % contre 3 % toutes classes confondues) ou, dans une moindre mesure, un événement familial. Peu de pratiques spécifiques ressortent en ce qui concerne la lecture et la sociabilité. Une proportion importante de locataires (+20) et une part plus importante de chômeurs (+9) suggèrent

12. La question posée est « Avez-vous connu des difficultés scolaires ? » Si oui, « En quelle classe ont-elles commencé ? »

13. Il s'agit plus précisément d'une classification effectuée sur les coordonnées factorielles des 10 premiers axes résultant d'une ACM réalisé sur le tableau de contingence multiple croisant l'ensemble des variables retenues.

14. Cette notation utilisée par la suite signifie un écart de 8 points par rapport à la population dans son ensemble.

Tableau 7

**Parcours de vie et compétences à l'écrit : une typologie des répondants**

	Classe 1	Classe 2	Classe 3	Classe 4	Classe 5	Classe 6	Classe 7	Ensemble
<b>Poids (en %)</b>	24	11	12	6	26	7	14	<b>100</b>
<b>Difficultés dans les trois domaines fondamentaux de l'écrit</b>								
Pas de difficulté	98	95	90					<b>85</b>
Difficultés modérées				9	9		10	<b>6</b>
Fortes difficultés					5	21	5	<b>4</b>
Graves difficultés						42		<b>4</b>
<b>Caractéristiques sociales et démographiques</b>								
<b>Sexe</b>								
Hommes					59	57		<b>45</b>
Femmes	67		63	58			58	<b>55</b>
<b>Âge</b>								
15-29 ans		99	24					<b>20</b>
30-39 ans	31		32		20			<b>18</b>
40-49 ans	31		25		32			<b>22</b>
50-59 ans	26			40	34	38		<b>23</b>
60-65 ans							94	<b>16</b>
<b>Niveau d'études (au moment de l'enquête)</b>								
Primaire						40	36	<b>10</b>
Secondaire général			29	38	29	34		<b>25</b>
Secondaire technologique ou professionnel			46	39	59			<b>34</b>
Enseignement supérieur	70	56						<b>31</b>
<b>Situation par rapport à l'emploi</b>								
Travaille	87		68		79			<b>60</b>
Chômeur			17	11	10	13		<b>8</b>
Inactif				13		32		<b>9</b>
Études		62						<b>7</b>
Retraité							89	<b>15</b>
<b>Statut d'occupation du logement</b>								
Propriétaire							72	<b>33</b>
Accédant	39				27			<b>24</b>
Locataire		64	61	45		63		<b>40</b>
Logé gratuit		3				3		<b>2</b>
<b>Parcours scolaire et enfance</b>								
<b>Changements d'école en primaire</b>								
Pas de changement	71	73			72			<b>68</b>
1 à 2 changements			30					<b>23</b>
3 changements et plus			20			11		<b>8</b>
<b>Redoublements (au collège)</b>								
Pas de redoublement	79	85						<b>63</b>
Un redoublement			37		36			<b>24</b>
Deux redoublements et plus			6		5	6		<b>3</b>
Non concerné						41	36	<b>10</b>
<b>Classe de début des difficultés scolaires (perçues)</b>								
Pas de difficulté	83	74					73	<b>62</b>
CP						42		<b>7</b>
Primaire			13	13	12	25	14	<b>10</b>
Secondaire			36		31			<b>20</b>
<b>Pratiques de la lecture (entre 8 et 12 ans)</b>								
Tous les jours	32							<b>15</b>
Régulièrement	39	28						<b>24</b>
De temps en temps		43			51		42	<b>38</b>
Jamais			27		32	67		<b>24 →</b>

Tableau 7 (suite)

	Classe 1	Classe 2	Classe 3	Classe 4	Classe 5	Classe 6	Classe 7	Ensemble
<b>Aisance financière de la famille (entre 8 et 12 ans)</b>								
Riche	5	6						3
A l'aise	50	66						37
Juste mais ils y arrivaient				64	65		67	54
Difficile			13	10		21	7	5
<b>Taille de la fratrie</b>								
0 ou 1 frère ou 1 so	48	46						33
2-3 frères et soeurs		43	44		43			40
4 et + frères et soe			31	41	29	62	34	27
<b>Pratiques quotidiennes actuelles</b>								
<b>Lecture de livres</b>								
Souvent	63						37	33
Parfois		48			36			33
Jamais					52	84		34
<b>Lecture de magazines</b>								
Souvent	49						44	32
Parfois		53	50		53			46
Jamais			24		25	61		22
<b>Lecture de journaux</b>								
Souvent	45						52	35
Parfois		45	42		42			36
Jamais			34			60		29
<b>Perception du niveau de lecture du français</b>								
Très bon	90	76						59
Bon				42	60		48	37
Pas bon						37		3
<b>Sociabilité (famille, amis, voisins)</b>								
Famille+++ (tous les jours)					21	24		18
Famille_jamais			5	6		11		4
Amis+++ (tous les jours)		47						15
Amis_jamais				9		18		5
Voisins+++ (tous les jours)						24	28	19
Voisins_jamais		32	24			24		18
<b>Événements vécus durant l'enfance (avant 18 ans)</b>								
Décès d'un frère ou d'une sœur			6	11		14	8	5
Décès d'un des parents				82				7
Mauvaise entente des parents			87					20
Séparation des parents		22	70					14
Problèmes de santé graves			13	14		19	11	9
Chômage (plus d'un an, en continu ou non) d'un parent		13	24					8
Alcoolisme dans le foyer			47			20		11
<b>Autre événement durant l'enfance cité dans la question ouverte (cf. Tableau 6)</b>								
Décès				73				11
Évènement familial			6					3
Maladie	4							3
Violence			14			7		3

Note : les pourcentages indiqués dans chaque classe ne sont donnés que pour les modalités sur-représentées (par rapport au profil moyen qui figure dans la dernière colonne du tableau). Les modalités des variables ou les événements qui ne sont pas sur représentés ne figurent pas dans le tableau. Pour cette raison, la somme des modalités pour une variable donné n'est pas toujours égale à 100 %. Par ailleurs, cette typologie est issue d'une classification ascendante hiérarchique (CAH) effectuée sur les 10 premières coordonnées factorielles d'une Analyse des Correspondances Multiples (réalisée avec le logiciel SPAD).

Lecture : dans la classe 6, 42 % des personnes ont de graves difficultés à l'écrit contre 4 % dans l'ensemble.

Champ : ensemble des personnes scolarisées en France et dont la langue d'apprentissage est le français (12 323 individus).

Source : Insee, enquête Information et vie quotidienne (IVQ) 2011.

une moindre aisance financière en comparaison avec les classes 1 et 2.

***Des difficultés modérées et des parcours marqués par le décès d'un parent (Classe 4, 6 %)***

Les personnes de ce profil connaissent des difficultés modérées face à l'écrit (pas de différence significative avec le profil moyen pour les personnes sans difficulté et un peu plus de personnes avec des difficultés modérées (9 % au lieu de 6). Il s'agit de personnes en moyenne âgées de plus de 50 ans, issues de grandes fratries (4 frères et sœurs et plus). Les événements marquants vécus durant leur enfance sont le plus souvent relatifs à la perte d'un des parents (82 % des personnes regroupées ici contre 7 % pour l'ensemble de la population), parfois à un décès dans la fratrie (11 % au lieu de 5 %) ou des problèmes de santé graves qui ont conduit la personne à être absente de l'école plus de deux mois consécutivement. Ces personnes sont aujourd'hui un peu plus fréquemment inactives (+ 4) ou au chômage (+ 3) qu'en moyenne, et elles ont une sociabilité quelque peu réduite.

Les deux profils qui suivent caractérisent des personnes plus en difficulté face à l'écrit que dans les profils précédents, les premiers en lien avec un parcours scolaire chaotique, les seconds avec une durée de scolarité très réduite. Pour autant les parcours de vie et les conditions de vie au moment de l'enquête ne présentent pas de difficultés particulières.

***Difficultés modérées face à l'écrit en lien avec un parcours scolaire chaotique (Classe 5, 26 %)***

Les personnes présentes ici sont un peu plus en difficulté dans leur rapport à l'écrit que la population prise dans son ensemble. Plus précisément, cette catégorie comporte légèrement plus de personnes en difficultés fortes et modérées (et moins de personnes sans difficultés ou avec de graves difficultés). Cette population est plus fréquemment masculine, d'âge compris entre 30 et 59 ans, et composée d'actifs en emploi (+ 19) ou au chômage (+ 2). Aucun événement de l'enfance n'apparaît de manière surreprésentée dans ce profil, qui n'est pas non plus davantage concerné par des conditions de vie difficiles. Le niveau d'étude atteint est relativement faible (88 % ne dépassent pas le secondaire contre 59 % dans l'ensemble). Les difficultés semblent avant tout liées à un

parcours scolaire difficile et sont ressenties tout au long du parcours scolaire ponctué de redoublements, aggravées par un manque de pratiques de lecture que ce soit dans l'enfance (32 % déclarent n'avoir jamais lu de livres entre 8 et 12 ans) ou actuelles.

***Des difficultés modérées pour des retraités de faible niveau d'éducation (Classe 7, 14 %)***

Ce profil rassemble presque uniquement des personnes de plus de 60 ans (94 %) et retraitées (89 %) qui sont un peu plus souvent en difficulté face à l'écrit (10 % sont estimées avec des difficultés modérées au lieu de 6 % dans l'ensemble). Près de la moitié n'ont pas dépassé le primaire et, sans avoir vécu d'événements douloureux durant leur enfance, elles ont connu des conditions de vie financièrement difficiles au sein de fratries nombreuses. Leurs difficultés relativement modérées vont de pair avec un niveau d'éducation faible. Une très grande majorité d'entre eux sont propriétaires de leur logement (+ 39, en grande partie en raison de leur âge) et développent une sociabilité tournée vers le voisinage (+ 9). Les pratiques de lecture actuelles, quel que soit le support, sont très fréquentes.

C'est seulement au sein du dernier profil que la situation des personnes peut être considérée comme préoccupante, associant de fortes difficultés à un parcours et à des conditions de vie difficiles.

***Très grandes difficultés face à l'écrit et enfance difficile dans un contexte familial relativement violent (Classe 6, 7 %)***

Près de la moitié des personnes relevant de ce profil peuvent être considérées en situation d'illettrisme (42 % ont des graves difficultés) et 21 % sont estimées avec de fortes difficultés. Les hommes sont sur représentés (+ 12) ainsi que la classe d'âge 50-59 ans (+ 15). Faiblement scolarisées (40 % ont quitté l'école après le primaire), ces personnes se trouvent plus fréquemment dans des situations d'inactivité ou de chômage. Les événements vécus durant l'enfance sont nombreux et touchent des domaines très différents : décès dans la fratrie (+ 9), maladie (+ 10), alcoolisme dans le foyer (20 % contre 11 % pour l'ensemble) dans un contexte plus fréquent de précarité (21 % déclarent une situation financière difficile durant l'enfance, 5 % dans l'ensemble) au sein de fratries nombreuses (62 % appartiennent à une fratrie de 4 enfants et plus, contre 27 % dans l'ensemble).

Dans cette classe, les événements recueillis de manière ouverte apparaissent davantage au travers de la modalité « violence » qui touche 7 % de la population (3 % dans l'ensemble). Les pratiques de sociabilité actuelles sont soit très réduites, soit fréquentes, selon les personnes. En revanche, les pratiques de lecture sont absentes, tout comme elles l'étaient durant l'enfance, et cela quel que soit le support (livre, magazine, quotidien). Les difficultés face à l'écrit ont été perçues dès le début de la scolarité (en CP, + 35) et près de la moitié des personnes déclarent ne pas bien maîtriser la lecture aujourd'hui (+ 11).

\* \*  
\*

Les éléments du parcours individuel constituent une dimension explicative importante des difficultés éprouvées à l'âge adulte. En particulier, les événements marquants de l'enfance, qu'ils soient recueillis en complément d'une grille biographique ou au moyen d'une question ouverte, apportent un éclairage utiles. Notre analyse confirme par ailleurs l'intérêt d'utiliser conjointement différents modes de recueil des événements pour faire émerger des éléments des parcours individuels qui viennent élucider les difficultés rencontrées à l'âge adulte dans leur rapport à l'écrit.

Si la survenue d'un décès constitue de loin l'événement le plus marquant des parcours d'enfance, les événements relatifs à la maladie ou à des situations de violence apparaissent dans les mêmes proportions (11 % des événements cités de manière ouverte). Ce dernier fait est loin d'être négligeable, comme en témoignent d'autres enquêtes – par exemple Enveff (2000).

L'utilisation des méthodes de la statistique textuelle, éprouvées depuis de nombreuses années

dans le domaine des sciences sociales, permet d'explorer le corpus d'événements marquants de l'enfance en alliant démarches qualitatives et quantitatives. Plus encore, elle fournit un outil pertinent dans l'élaboration d'une grille d'analyse des événements marquants, susceptible d'être reproduite et utilisée dans d'autres enquêtes de grande envergure.

L'analyse des profils de l'ensemble de la population enquêtée dans l'enquête *IVQ* 2011 montre la diversité des associations entre éléments biographiques, conditions de vie et compétences à l'écrit. Tout comme dans l'enquête *IVQ* de 2004, la variété des situations individuelles invite à relativiser l'idée de l'existence d'un lien mécanique entre, d'une part, enfance difficile et manque de compétences et, d'autre part, entre difficultés éprouvées face à l'écrit et situations de précarité et isolement social. C'est au sein d'une seule classe qui regroupe 7 % des répondants (classe 6) que l'on observe réellement des situations que l'on peut qualifier d'illettrisme pour des personnes ayant vécu une enfance difficile avec un parcours scolaire extrêmement réduit et des conditions de vie précaires. C'est vers cette population que les efforts des politiques doivent se tourner en priorité. En effet, si parmi les profils mis en évidence, certains se caractérisent également par une enfance difficile, dans un contexte familial violent et financièrement précaire, ils ne s'accompagnent pas à l'âge adulte de difficultés spécifiques en termes de maîtrise de l'écrit mais se signalent davantage en termes de difficultés économiques.

Comme on l'avait remarqué à l'enquête de 2004, près de la moitié des personnes en situation d'illettrisme déclarent avoir ressenti des difficultés dès le début de leur scolarité. Ce constat invite à accentuer les efforts en faveur des politiques scolaires de soutien aux jeunes enfants en difficulté. □

---

## BIBLIOGRAPHIE

**Besse J.-M., Luis M.-H., Bouchut A.-L. et Martínez F. (2009)**, « La mesure des compétences en traitement de l'écrit chez des adultes en grande difficulté », *Économie et Statistique*, n° 424-425, pp. 31-48.

**Besse J.-M. et Guérin-Pace F. (2002)**, « Une évaluation des compétences sur l'écrit : l'enquête *Information et vie quotidienne* », *Économie et Humanisme*, n° 363, pp. 17-21.

**Garnier B. et Guérin-Pace F. (2010)**, *Appliquer les méthodes de la statistique textuelle*, collection *Les Clefs pour*, Ceped, 86 pages.

**Golaz V. et Lelièvre É. (2012)**, « L'entourage familial pendant l'enfance et l'adolescence, entre faits et perceptions. Une analyse rétrospective des parcours de vie des Franciliens des générations 1930-1950 », *Population*, vol. 67, n° 3, pp. 491-515.

**Guérin-Pace F. (2009)**, « Illettrismes et parcours individuels », *Économie et Statistique*, n° 424-425, pp. 49-63.

**Guérin-Pace F. (1997)**, « La statistique textuelle : un outil exploratoire en sciences sociales », *Population*, vol. 52, n° 4, pp. 865-888.

**Laborde C., Lelièvre É. et Vivier G. (2007)**, « Trajectoires et événements marquants, comment

dire sa vie ? Une analyse des faits et des perceptions biographiques », *Population*, vol. 62, n° 3, pp. 567-585.

**Lebart L. et Salem A. (1994)**, *Statistique textuelle*, Dunod, Paris, 342 pages.

**Murat F. (2005)**, « Les compétences des adultes à l'écrit, en calcul et en compréhension orale », *Insee Première*, n° 1044.

**Reinert M. (1983)**, « Une méthode de classification descendante hiérarchique : application à l'analyse lexicale par contexte », *Cahiers de l'Analyse des Données*, vol. 8, n° 3, pp. 187-198.

**Reinert M. (1993)**, « Les mondes lexicaux et leur logique à travers l'analyse statistique d'un corpus de récits de cauchemars », *Langage et société*, n° 66, pp. 5-39.