

Connexions linguistiques dans les Codes de la Torah 29 Novembre 2004

(Traduit de l'anglais par E. Blum)

Art Levitt¹, Robert M. Haralick², Eliyahu Rips³

¹ Computer Research Analyst, Jerusalem, Israel

² Computer Science Department, City University of New York, U.S.A.

³ Institute of Mathematics, Hebrew University, Jerusalem, Israel

artlevitt23@yahoo.com

Résumé

La présente étude propose une méthode simple et vérifiable permettant de tester l'hypothèse de l'existence de codes dans la Torah. Cette hypothèse stipule que la Torah (c'est-à-dire les cinq premiers livres de la Bible Hébraïque) contient des suites de lettres (des codes) qui ont été créées intentionnellement à des fins de communication avec le genre humain, les destinataires visés. Basée sur l'accumulation de preuves de l'existence de multiples encodages d'un concept ou d'un événement particulier dans le texte de la Torah [1], notre méthode de test se propose de mesurer la force des connexions linguistiques existant entre des codes exceptionnellement remarquables trouvés par le passé. Nous présentons dans un premier temps la méthode générale. Nous appliquons ensuite cette méthode à des codes de la Torah relatifs à Ben Laden et à l'attaque des Tours Jumelles. Les nombreuses connexions significatives que nous avons trouvées ont une probabilité estimée de 1 sur 500000 d'être apparues simplement par chance. Ce résultat renforce considérablement les conclusions établies précédemment soutenant l'hypothèse de l'existence de codes dans la Torah.

1. Introduction

La recherche sur les codes dans la Torah porte sur un type particulier de suite de lettres que l'on obtient par le prélèvement de lettres régulièrement espacées dans un texte donné. C'est ce qu'on appelle une SLE (Suite de Lettres Equidistantes).

Le processus de prélèvement des lettres s'effectue en ignorant toute ponctuation ainsi que les espaces entre les mots. Par exemple, on peut trouver le mot «*ecu*» en partant du premier «*e*» de la phrase précédente et en effectuant un saut de +4 lettres (c'est-à-dire, en comptant 4 lettres à partir de la lettre de départ).

Selon l'hypothèse originale, stipulant l'existence de codes dans la Torah, des mots ayant un rapport entre eux sur le plan logique ou historique peuvent être trouvés sous la forme de SLEs de façon significativement plus fréquente et dans des configurations plus compactes dans le texte de la Torah que ce à quoi l'on pourrait s'attendre simplement par

chance. De nombreuses études antérieures sur les codes dans la Torah, de la part de supporters ou d'opposants aux codes, ont porté sur l'analyse d'agrégats de SLE de noms de Rabbins renommés, de leurs dates et villes de naissance ou de mort [1-6] ainsi que sur des SLEs de nombreux autres événements historiques. Plus récemment, des études ont porté sur des événements actuels, concernant notamment l'attaque des Tours jumelles (voir les tables des figures 1 à 5 ; comme à l'accoutumé, le texte sous-jacent est organisé en table de longueur de ligne constante ; tous les codes sont en hébreux et annotés de traductions françaises).

La principale difficulté, dans ce genre de recherche, réside dans le fait de concevoir des expériences qui démontrent implicitement, à un observateur extérieur, que les tests effectués ont été très limités en nombre. Pour cette raison, nous présentons d'abord les critères permettant d'identifier des tables de codes de première catégorie (NDT : par définition, ce sont des codes rares, dont le caractère à priori ou pré-établi est vérifiable ; de plus, la probabilité que ces codes soient des coïncidences est très faible). Ces tables sont elles-mêmes utilisées, d'une part pour décider de ce qu'il faut chercher, d'autre part comme lieu du texte de Torah où effectuer de nouvelles recherches. Ce qui nous intéresse, ce sont les connexions, c'est-à-dire les répétitions d'un ou plusieurs mots dans ces tables. Si nous formulons l'hypothèse initiale qu'il n'existe pas de codes dans la Torah, nous nous attendons à ce que les connexions trouvées entre les tables de première catégorie, ne soient pas plus fortes que celle que nous serions susceptibles de trouver entre d'autres tables créées à partir d'autres textes.

2. Présentation de la méthode générale

2.1. Identification des tables de première catégorie

La première étape consiste à réunir des tables de première catégorie trouvées par le passé et qui traitent d'un sujet particulier. Pour qu'un code soit accepté en tant que code de première catégorie, celui-ci doit satisfaire deux critères. Premièrement, les éléments clés de ce code doivent être, statistiquement parlant, en étroite proximité. Deuxièmement, ce code doit être ouvertement vérifiable, comme cela est décrit ci-après.



Figure 4: Mots les plus cités du journal Haaretz ainsi que les connecteurs “qui se lamente?”, “assassine” et “des milliers de personnes”.



Figure 5: La fin d'une autre phrase construite autour du mot Ben Laden (NdT: « il augmenté » : dans le sens, il a porté à un niveau supérieur [le terrorisme])



Figure 5a: Continuation de la figure 5 avec une longueur de ligne moitié.

Il existe au moins 4 façons différentes de créer des codes ouvertement vérifiables:

- Les mots clés dont on cherche des SLEs proviennent d'une source indépendante, connue et non ambiguë.
- Une SLE unique d'une phrase commençant par un mot clé connu jouant le rôle de mot « ancrage » [8], est sémantiquement et/ou grammaticalement cohérente.
- Un nouvel événement renforçant se produit après que le code a été découvert.
- Il se peut que le code contienne une sorte de contrainte auto imposée qui limite grandement le nombre plausible de SLEs additionnelles également surprenantes.

Les codes satisfaisant les deux critères (l'étroite proximité et le fait d'être ouvertement vérifiable) sont par définition très rares. Nous restreignons par conséquent grandement le nombre de codes de première catégorie étudiés, ce qui évite la possibilité d'une minimisation du nombre d'essais non concluants rapportés.

2.2. Trouver des liens linguistiques simples

Après avoir identifié les tables de première catégorie, nous recherchons des « connecteurs » susceptibles d'exister entre elles. Un connecteur est un mot qui apparaît plus d'une fois dans les différentes tables (en tant que SLE et/ou occurrence du texte sous-jacent) comme cela est illustré dans de nombreuses figures. Pour les besoins du présent calcul, nous analysons seulement un sous-ensemble parmi tous les connecteurs possibles en appliquant des critères de sélection restrictifs et très simples. Notre point de départ consiste en une liste de mots clés collectés mécaniquement à partir des SLEs et des versets précédemment surlignés dans les tables de première catégorie. Nous ne retenons que les mots les plus difficiles à trouver, ceux composés d'au moins 5 lettres. (nous prenons en compte toute autre épellation recevable sachant qu'il n'en existe qu'une pour l'expression « attaque terroriste »). Chaque mot clé de notre liste fait l'objet d'une recherche dans chaque table de première catégorie, le but étant de déterminer si le mot clé en question est proche de l'axe vertical principal de la table (le thème qui se présente de façon verticale). Nous restreignons notre étude aux formations les plus impressionnantes – qualifiées elles-mêmes de significatives dans des études antérieures [9] – mettant en jeu des SLEs se présentant de façon parallèle, co-linéaire ou perpendiculaire en excluant les SLEs apparaissant en diagonale. Nous n'accordons pas de crédit au fait de trouver une première fois un mot mais seulement aux fois suivantes (et ce faisant, ces trouvailles se voient conférées le statut de connecteur).

2.3. Estimation des probabilités p (p-levels) pour chaque connexion

Pour chaque connexion, nous nous intéressons à la probabilité que la SLE trouvée dans la table – en position verticale ou horizontale – soit, de cette façon, étroitement proche du mot constituant l'axe vertical de la table. Cette probabilité s'obtient en comparant la table réelle avec un nombre important de tables créées selon la méthode Monte Carlo. Ces tables sont produites à partir de textes en hébreu, modernes et anciens, où nous ne nous attendons pas à trouver des codes (le texte de la Torah peut lui-même être utilisé en tant que texte comparatif, dans la mesure où nous ne nous attendons pas à ce que des tables choisies aléatoirement dans le texte de la Torah, contiennent des connecteurs en relation avec notre thème). Pour chaque table comparative, l'axe est positionné aléatoirement, une recherche des SLEs du mot clé est effectuée, et la proximité, avec l'axe, des SLEs (répondant aux critères de configuration) qui en sont les plus proches est évaluée. La probabilité recherchée correspond à la proportion des tables comparatives ayant une plus grande proximité que celle de la table originale. La mesure de cette proximité s'obtient en utilisant la meilleure valeur de μ (voir WRR [2]) pour 2 largeurs de table (celle de la table originale et celle de valeur moitié). La table de largeur moitié produit un axe dont les lettres apparaissent une ligne sur deux. Par conséquent, nous limitons notre recherche de SLEs verticales à celles dont le saut est égal à la largeur de la table ainsi qu'à sa demie largeur. Cela correspond à la configuration observée à la figure 5a et qui est à l'origine de la présente étude.

2.4. Estimation de la probabilité (p-level) globale

En utilisant un nombre raisonnablement grand de données, la loi de distribution binomiale nous fournit un ordre de grandeur de la probabilité (p-level) globale. Le nombre d'essais ainsi que le nombre de connexions ayant le seuil de probabilité – p-level – constituent les paramètres de ce calcul. Le nombre d'essais correspond à la multiplication du nombre d'items de la liste de mots clés par le nombre de tables collectées et par le nombre de configurations essayées. (égal à 2, par exemple, si nous analysons séparément les configurations verticales et horizontales).

3. Cas étudié

3.1. Les tables de première catégorie

Nous appliquons la méthode décrite ci-dessus à un ensemble de codes de la Torah trouvés précédemment, et relatifs à Ben Laden ainsi qu'aux Tours jumelles (Figures 1 à 5) [7-8]. Ces codes satisfont les critères établis ci-dessus à la section 2.1. Chacun d'entre eux est

ouvertement vérifiable selon l'une des 4 possibilités exposées ci-dessus : la table de la figure 2 correspond à la possibilité (a) puisqu'elle utilise les noms des gros titres du journal *Maariv du 12 Septembre 2001*, tandis que la table de la figure 4 utilise les mots les plus cités du journal *Haaretz*; ces 2 titres faisant partie des 3 journaux majeurs en hébreu d'Israël. La table de la figure 3 correspond à la possibilité (b) puisqu'elle utilise le mot Ben Laden en guise de mot-clé ancrage, ce qui a conduit à la découverte d'une phrase l'incluant (cette phrase est remarquable du fait d'une connexion sémantique extrêmement proche, dans le texte de la Torah lui-même, avec les mots « péché » et « crime »); la table de la figure 5 est basée sur la fin d'une SLE d'une phrase incluant une autre occurrence de ce mot ancrage [8]. La table de la figure 1 correspond à la possibilité (c). Celle-ci a été trouvée le 1^{er} octobre 2001, soient quelques semaines après le 11 septembre. Le verset surligné de cette table était intrigant du fait qu'il indiquait un nombre de 3000 victimes, alors que les estimations fournies par les nouvelles de l'époque tournaient autour de 6000 victimes. La confirmation du code est venue plusieurs semaines après. La table de la figure 2a correspond à la possibilité (d) puisqu'elle utilise le mode de recherche sous forme de dialogue. Ce mode a été mis en évidence plusieurs années avant le 11 septembre ; celui-ci consiste en une structure de code qui contient à la fois une question et sa réponse. Nous pouvons observer des mots particulièrement appropriés constituant cette structure auto renforçante (« qui se lamente ? » et « les U.S.A se lamentent »).

3.2. Le caractère signifiant des connexions linguistiques

Il y a de nombreuses connexions significatives qui respectent nos contraintes ; toutes ne figurent pas dans les tables. Quatre connecteurs atteignent un seuil de probabilité (p-level) de 0,001. L'un d'entre eux concerne une SLE de l'expression « qui se lamente ? » – voir figure 1 (la première apparition se trouve dans la figure 2a qu'on retrouve ensuite dans la figure 4). Nous trouvons aussi le mot « horreur » dans les figures 1 et 2 ainsi que le mot « Ismaël » (qui apparaît encore dans la table 3), et l'expression « des milliers de personnes » dans la figure 4 (qui correspond au verset de la figure 1 dont la traduction littérale en hébreu est « trois mille personnes »). A partir d'une variété de types de textes comparatifs (incluant des tables extraites aléatoirement du texte de la Torah elle-même), nous avons obtenu des probabilités (p-levels) constamment comprises entre 10^{-3} et 10^{-5} pour les 2 premiers succès tandis que pour les 2 autres, ces probabilités étaient situées à au moins 1 ou 2 ordres de grandeur au dessus.

3.3. Estimation du degré de signifiance globale

La liste des mots-clés, obtenue à partir des 5 tables uniques de première catégorie dans lesquelles les

recherches ont été ensuite effectuées, contient 13 items. En se limitant aux configurations verticales et horizontales, comme spécifié à la section 2.4, il s'ensuit que le nombre d'essais est de $13 * 5 * 2$, soit 130. Sur ce nombre d'essais, nous avons obtenu 4 succès avec un seuil de probabilité de 0,001. En utilisant la loi de distribution binomiale, nous avons obtenu un niveau de probabilité (p-level) de 10^{-5} , soit 1 sur 100000. Si nous étendons notre observation aux SLEs ayant un saut de +1 ou -1, il s'ensuit que l'expression « la tour » de la figure 1 constitue un connecteur valide dont la première apparition a été observée à la figure 3 (il respecte aussi le seuil de probabilité de 0,001). Cela constitue une 3^{ème} configuration en plus des configurations verticales et horizontales, ce qui porte le nombre total d'essais à $13 * 5 * 3$, soit 195. Sur ce nouveau nombre d'essais, nous avons obtenus 5 succès avec un seuil de probabilité de 0,001, la probabilité (p-level) résultante étant de $1,84.10^{-6}$, soit approximativement de 1 sur 500000.

4. Digressions

Il est souvent difficile de dériver correctement une probabilité si des contraintes strictes ne sont pas fixées au préalable. La figure 5a constitue un bon exemple illustrant notre remarque. Trois connecteurs potentiels sont présents dans le bas de cette figure. Ces connecteurs respectent les contraintes spatiales mais pas précisément les contraintes linguistiques, dans la mesure où c'est la forme verbale qui est utilisée pour le mot « revanche » plutôt que la forme substantive et que la conjonction de coordination « et » est présente entre les mots « péché » et « crime ». Ceci dit, ce sont des contraintes sévères, mais ce faisant, on évite l'introduction d'éléments qu'on ne peut correctement prendre en compte dans le calcul d'une probabilité.

Il existe une autre découverte intrigante – non prise en compte ici – qu'on aurait pu ajouter à la table de la figure 1 et qui est relative à un agrégat des mots suivants: « des lamentations se font entendre » et « aux U.S.A ». La raison en est que, d'une part les SLEs des expressions en question apparaissent en diagonale et que d'autre part la formulation n'est pas précisément celle trouvée précédemment.

Il existe aussi une table complète qui reprend les sujets suivants : avion, attaque, par deux fois et Tours jumelles sur la base de synonymes et de mots orthographiés différemment [7]. A nouveau, celle-ci n'a pas été retenue du fait qu'elle ne respecte par strictement les critères fixés.

Ceci dit, même en faisant abstraction de ces découvertes supplémentaires, le résultat est hautement significatif et se situe au delà du niveau de signifiante des tables originales (estimée entre $p=10^{-3}$ et $p=10^{-6}$).

5. Conclusion

Bien que nous ne nous soyons pas encore penché sur l'aspect compréhension et interprétation de ces tables, la présente étude suggère fortement que le phénomène des codes dans la Torah est réel.

Nous avons examiné une série de codes de première catégorie hautement significatifs. Nous avons observé des connections renforçantes entre ces codes ayant une probabilité de 1 sur 500000, révélant de ce fait une structure beaucoup plus forte et complexe que ce qui avait été observé jusqu'à présent.

Nous observons aussi que le facteur temps, la clarification et la validation croisée de ces codes au cours des mois et des années qui ont suivi leur découverte, montrent qu'il existe une sorte de qualité de « vie » reflétant le paradoxe – pointé par les anciens commentaires de la Torah – que d'un côté, tout a été prévu et que d'un autre côté, nous sommes simultanément dotés du libre arbitre.

6. Remerciements

Nous voudrions remercier Alex Rotenberg pour avoir développé et mis à disposition son logiciel, Boaz Metzger et Igor Pisetsky pour leur contribution au sujet de plusieurs tables de première catégorie, Harold Gans pour ses précieux commentaires tout au long du processus, et Leib Schwartzman pour son travail de pionnier au sujet du mode dialogue.

7. Références

- [1] R. Haralick (2003); Torah Codes: Redundant Encoding; <http://www.torahcodes.net/redun.html>
- [2] D. Witztum, E. Rips, Y. Rosenberg; Equidistant Letter Sequences in the Book of Genesis; *Statistical Science*, 9(3):429-438, 1994.
- [3] H. Gans (2001); Torah Codes Primer; <http://aish.com/seminars/discovery/Codes/codes.htm>
- [4] R. Haralick (2003); Testing the Torah Code Hypothesis; <http://www.torahcodes.net/hypoth.html>
- [5] D. Witztum; Torah Codes; <http://www.torahcodes.co.il>
- [6] B. McKay, D Bar-Natan, et al, *Solving the Bible Code Puzzle*, *Statistical Science*, 14(5):150-173, May 1999.
- [7] A. Levitt (2004); <http://www.torahcodes.net/alltwin.html>
- [8] A. Levitt, N. Bombach, et al (2004); Long Phrases; 2nd Annual Speech and Language Conference; U of Belgrade
- [9] A. Levitt (2002); <http://www.torahcodes.net/w2d.html>