

# CAHIERS *GUTenberg*

☞ LES ALLOGRAPHES  
COMME MESURE BIOMÉTRIQUE  
ET LA SYNTHÈSE DU COMPORTEMENT

☞ Vlad ATANASIU

*Cahiers GUTenberg*, n° 46-47 (2006), p. 7-17.

<[http://cahiers.gutenberg.eu.org/fitem?id=CG\\_2006\\_\\_46-47\\_7\\_0](http://cahiers.gutenberg.eu.org/fitem?id=CG_2006__46-47_7_0)>

© Association GUTenberg, 2006, tous droits réservés.

L'accès aux articles des *Cahiers GUTenberg*

(<http://cahiers.gutenberg.eu.org/>),

implique l'accord avec les conditions générales

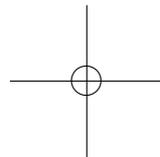
d'utilisation (<http://cahiers.gutenberg.eu.org/legal.html>).

Toute utilisation commerciale ou impression systématique

est constitutive d'une infraction pénale. Toute copie ou impression

de ce fichier doit contenir la présente mention de copyright.





# LES ALLOGRAPHES COMME MESURE BIOMÉTRIQUE ET LA SYNTHÈSE DU COMPORTEMENT

¶ Vlad ATANASIU

RÉSUMÉ. — Cet essai traite du rôle des allographes dans le cadre général de l'étude de l'écriture. Il explique pourquoi les allographes sont importants (l'individualité de l'écrivain), comment les mesurer (fréquences et contextes graphiques, sémantiques et non graphiques) et à quelles applications s'attendre (biométrie, restauration de l'écriture, création artistique, histoire des aspects culturels).

ABSTRACT. — This essay deals with the role of allographs in the general framework of the study of writing. It explains why allographs are important (writer individuality), how they can be measured (frequency and (graphic, semantic, non-graphic) contexts) what applications to expect from them (biometrics, script restoration, artistic creation, history of cultural aspects).

NOTE. — Cet article est paru dans *TUGBOAT*, volume 24 (2003), n° 3 — *EuroT<sub>E</sub>X 2003 Proceedings* (Yannis Haralambous éd.), sous le titre « Allographic Biometrics and Behavior Synthesis », p. 328-333. Version française aimablement fournie par l'auteur. Il est reproduit ici avec l'autorisation de l'auteur, de Karl Berry (président de TUG) et de Barbara Beeton (*TUGBOAT editor*).

Un des éléments les plus utilisés dans l'identification d'une police typographique ou d'une écriture manuscrite est sa forme. Souvent, pourtant, les mêmes caractères peuvent être rendus par des formes visuellement distinctes (les minuscules et majuscules par exemple) — dans ce cas les règles et les habitudes gouvernant leur usage caractérisent d'une façon plus subtile le comportement du typographe ou celui de l'écrivain que ne le fait la forme de lettres. L'objectif de cet article est d'attirer

l'attention sur l'étude des allographes — l'allométrie — tant pour le plaisir de la connaissance que pour ses applications pratiques. Je procéderai en partant de quatre diagrammes, qui représentent des aspects centraux du thème. Bien que les exemples soient basés sur l'écriture arabe — une écriture où les allographes sont une part inhérente du système d'écriture —, les mêmes principes s'appliquent aux écritures non arabes, produites tant par des moyens manuels que mécaniques.

### 1. MESURE DE FRÉQUENCES

Les deux blocs de texte de la figure 1 représentent une double page d'un manuscrit copié en Sicile au xiv<sup>e</sup> siècle [6]. Sur la droite se trouve l'écriture originale (figure 1-a), tandis que sur la gauche on voit un folio de remplacement (figure 1-b), inséré dans le manuscrit, après la disparition d'un certain nombre de folios. On a tenté de réduire les différences : les dimensions des blocs ainsi que le nombre de lignes est identique et tous les deux sont écrits dans le style appelé *naskh*. Pourtant, même un œil non habitué à l'écriture arabe reconnaîtra un travail de restauration très approximatif.

Ce type d'expertise à l'œil nu, sans instruments, peut être objectivisé et automatisé : la magnitude d'une transformation de Fourier appliquée à l'image du texte révèle deux différences majeures (figure 1-c & 1-d). Les spectres ont la forme de croix, dont les angles des bras diffèrent d'un spectre à l'autre — cela implique une différence dans l'inclinaison des traits verticaux des deux écritures, l'écriture de remplacement étant plus « penchée », d'où l'inclinaison prononcée du bras horizontal de son spectre. Quant à la présence de quatre zones individualisées aux bouts des bras du spectre de l'écriture de remplacement (figure 1-d) — pour une diffusion uniforme pour le spectre de l'écriture originale (figure 1-c) — cela suggère une linéarité plus grande de l'écriture dans le premier cas et une écriture aux traits plus rond et irréguliers dans le deuxième.

Mais il existe une distinction encore plus puissante : le remplacement emploie deux formes de la lettre *kaf*, une oblique (représentée par un cercle apparaissant 13 fois dans la figure 1-b) et une longue (disque, 7 fois dans la figure 1-f), tandis que dans l'original n'apparaît que la version oblique (19 fois en figure 1-e). Est également remarquable la fixation de l'écrivain sur le contexte dans lequel le *kaf* long est utilisé : dans 4 des 6 occurrences du mot *kana* (particule verbale « être »). Définis par

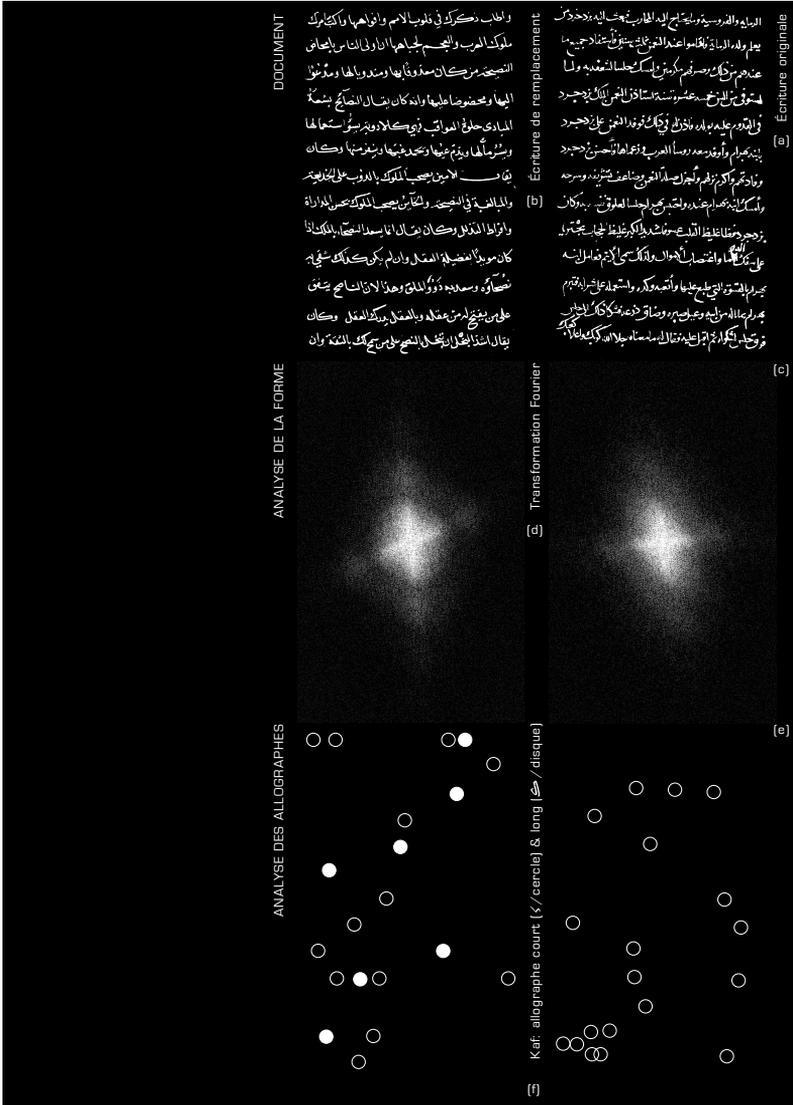


Figure 1. — Lettres derrière leur forme : quantité et contexte ?

des quantités et des contextes jamais évidents de visu, les allographes se révèlent être une signature en filigrane de l'auteur d'un texte.

## 2. APPLICATIONS

Les figures 2-a et 2-b sont d'excellents exemples de la principale application de l'allométrie : la *biométrie*, ou l'identification des individus basée sur l'analyse de caractéristiques spatio-temporelles de l'écriture (un sujet intéressant *inter alia* les agences de police et de renseignement, le système judiciaire, le marché de l'art, les historiens). Les pages présentés ici sont été écrites par deux calligraphes pour un concours de belle écriture [6, p. 31-32] et elles contiennent le même texte, cependant le nombre de *kafs* longs est 6, respectivement 23. Ainsi les allographes de type long jouent d'un individu à l'autre un rôle clairement très différent « d'attracteur visuel ». Le même phénomène peut être observé au niveau de traditions. Sur la base de cinq autres participants dans la compétition calligraphique, nous voyons que les Turcs utilisent un nombre similaire d'élongations (mais ils diffèrent dans leur attribution à des graphèmes spécifiques). Leur tradition paraît être en relation avec celle des Syriens, des Jordaniens et une partie des Iraquiens. En Irak pourtant, deux traditions coexistent, dont une fait un usage abondant d'élongations, bien à l'encontre des habitudes des calligraphes turcs.

Une fois l'analyse de l'écriture faite, elle peut servir de base à une *histoire culturelle*. Ainsi on remarquera la consistance du style turc, témoin de la force encore aujourd'hui vivace d'une tradition plusieurs fois centenaire. Les différences allographiques sont l'expression de l'existence d'un style arabe se voulant indépendant de l'école turque — une tendance anti-monopolistique elle-même ancrée dans les réalités politiques, économiques et sociales du Moyen-Orient. Étant donné que les styles calligraphiques iraniens nécessitent un emploi plus important des élongations, on peut enfin se demander quelle est l'influence de l'Iran sur la calligraphie iraquienne voisine. La calligraphie étant encore de nos jours un *medium* artistique important dans les pays musulmans, l'analyse culturelle peut révéler des canaux préférentiels d'influence et offrir des instruments d'action politique.

Une autre utilisation active de l'allométrie est la *synthèse du comportement*. Après avoir extrait des informations sur comment une personne ou une tradition traite les allographes, elles peuvent servir à prédire un



comportement futur, soit dans des situations documentées (typiquement la restauration d'une partie d'un texte endommagé ou manquant mais dont le contenu est connu), soit pour composer des textes virtuels, jamais écrits par l'auteur (« dans le style de... »). Les modèles de comportement devraient être considérés comme un aspect important de la typographie numérique étant donnée la possibilité qu'ils offrent d'adapter l'aspect graphique à un contenu ou à un lectorat spécifique (un texte ottoman avec une fréquence allographique arabe serait historiquement aberrant et désapprouvé par l'*establishment* calligraphique turc).

Enfin, l'allométrie peut être utilisée pour accroître l'impact visuel d'un texte d'une manière qui serait trop complexe ou longue à produire pour les mains et les cerveaux des humains. Par exemple, les allographes sont habituellement utilisés dans le contexte local, celui d'une page ou une double page, mais il n'arrive que très rarement que le choix d'un allographe soit fait à cause de l'existence d'un autre, situé à quelques dizaines ou centaines de pages de distance. Les meilleures pièces calligraphiques sont souvent courtes et ressemblent à une musique de chambre plutôt qu'à une symphonie. Avec un *logiciel d'aide à la composition* en revanche, un artiste pourrait se concentrer à donner forme aux figures et aux rythmes, à les affiner et à les harmoniser, en laissant le travail de leur réalisation physique à la machine. De même que les synthétiseurs de sons servent les compositeurs, de la même manière les ordinateurs pourraient générer des textes d'une complexité graphique jamais produite jusqu'à présent.

Sur le plan méthodologique l'allographie nous conduit à définir comme suit les objectifs de l'étude des propriétés graphiques de l'écriture — champ de recherche appelé *graphonomie*<sup>1</sup>. L'*analyse graphique* est un pas requis dans l'identification des écritures ; elle permet la création de *modèles* pour la reproduction de comportements des écrivains et leur modification au cours de la synthèse de nouvelles *créations* graphiques. L'expertise graphique mène aussi à comprendre les aspects non graphiques qui influencent la forme de l'écriture. À part son *aspect*

1. La graphonomie est distincte de la graphologie comme l'astronomie de l'astrotologie. Elle a sa propre association, conférence et publication (cf. <http://www.cedar.buffalo.edu/igs>).

*physique* (spatial, temporel, énergétique, acoustique, ...), l'écriture possède aussi une dimension *technologique* (matériaux, outils, processus), *informationnelle* (propriétés mathématiques, statistiques et informationnelles des systèmes d'écriture), *biologique* (physiologique, neuronale, psychologique) et *culturelle* (aspects sociaux, économiques, esthétiques, littéraires, coutumiers, religieux, ...).

### 3. IDENTIFIER LES CONTEXTES SÉMANTIQUES

La présence d'un mot ayant une importance particulière dans l'esprit d'un calligraphe ou dans le texte qu'il rédige peut lui suggérer l'utilisation d'un graphisme particulier. Dans un texte de labeur arabe cela se traduit habituellement par l'utilisation d'un allographe peu fréquent. La figure 3 présente l'exemple d'un mot — *kufr*, « impiété » — où apparaît la lettre *kaf*. Dans le contexte religieux du texte — citations coraniques, paroles saintes, avis sur la morale — ce mot déclenche dans 70 % des cas pour le groupe de calligraphes étudiés (les mêmes que dans la figure 2) l'emploi d'un allographe alternatif, de type long. Par contre *malaka* (« posséder »), mot moins pertinent pour le contexte, n'est écrit que dans 20 % des cas avec le *kaf* long, exceptionnel.

Un effet secondaire du transfert de statut de la signification lexicale vers sa représentation visuelle est la création d'un lien culturel au sein de cette association autrement fortuite. Les poètes emploient l'expression *kaf-i kufr* « le *kaf* de l'impénétrabilité » et expliquent la forme allongée de la lettre par une comparaison avec « l'étroitesse de cœur » de certaines gens. La métaphore des poètes renforce la pratique des calligraphes en la dotant d'un halo et d'un réseau de significations. Le déclenchement sémantique d'un allographe illustre son deuxième aspect le plus important après sa fréquence : les contextes où il apparaît<sup>2</sup>.

### 4. PROBLÈMES ET LIMITES

La figure 4 présente un cas d'organisation de l'espace — c'est-à-dire les techniques pour cadrer une quantité donnée de texte sur une surface inscriptible finie. J'analyse le comportement de l'écrivain par rapport

2. Une classification des contextes allographiques pour l'écriture arabe comparée à l'écriture latine est donnée dans ma thèse de doctorat [1, chapitre 3, « La grammaire graphique »].

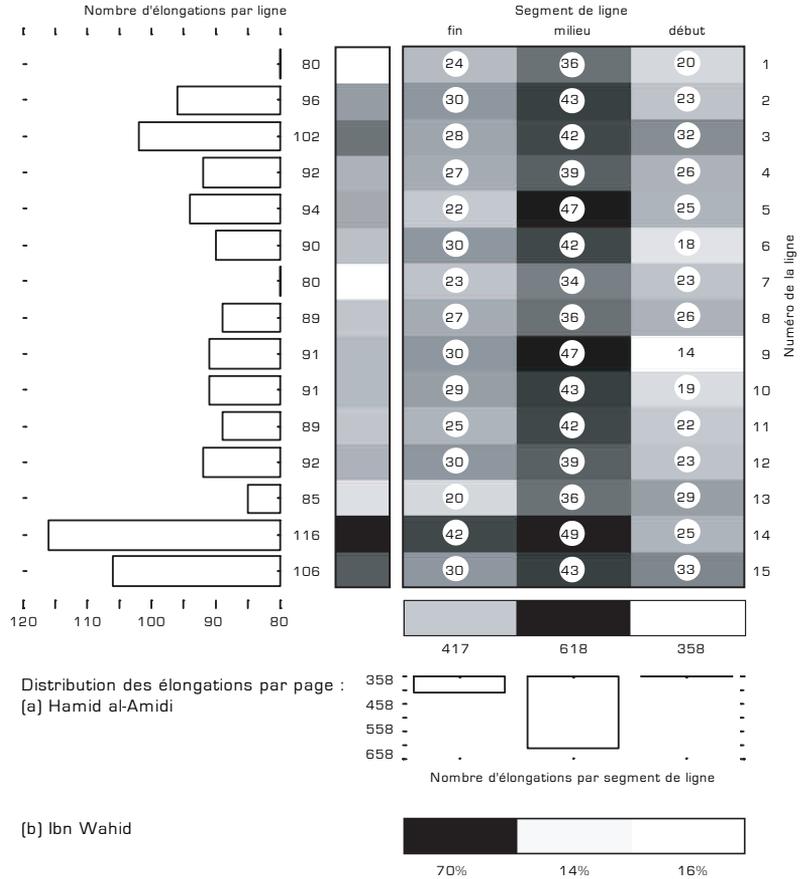


Figure 3. — Problème des contextes simultanés.

à l'élongation de segments graphiques, dans le contexte particulier de Corans ottomans dits *ayet ber kenar*, où la fin de chaque page doit impérativement coïncider avec la fin d'un verset, ce qui requiert du calligraphe une prouesse considérable pour lui permettre une justification agréable non seulement des lignes, mais aussi de la page tout entière [2]. La superposition de toutes les élancements d'une partie d'un manuscrit [3, p. 2-48] montre une fluctuation dans leur quantité d'une ligne

à l'autre. Ce qui veut dire que le calligraphe n'a pas étudié le texte suffisamment pour distribuer les élongations de manière uniforme. L'accumulation des allographes au milieu de lignes trahit l'endroit où il commence à penser à la justification, tandis que leur basse fréquence sur les premières lignes contrastant avec une haute fréquence en bas de page, indique sa hâte à se rattraper pour la justification verticale. Ce calligraphe, de tradition ottomane, anticipe pourtant bien plus que son collègue mamluk (figure 4-b), qui attend, pour ainsi dire, le son de la clochette des fins de ligne imminentes des machines à écrire pour songer à la justification [5, vol. 3, fol. 99 v<sup>o</sup>, 0-166 v<sup>o</sup>]. La présence réduite des allographes dans le premier tiers des lignes suggère que l'emploi principal des élongations était la justification, ce qui est confirmé par les traités de calligraphie et typographie arabes.

Si nous envisageons de quantifier les données sur la justification, nous devons considérer que la page qu'on voit est le résultat de la présence simultanée d'un certain nombre de contextes, chacun ayant des demandes spécifiques sur l'aspect graphique. Par exemple dans la figure 1, on a vu que les *kafs* longs sont une part importante du total des élongations, ce qui peut dire qu'une distribution non uniforme des élongations soit en partie due à la présence et à la distribution des mots contenant la lettre *kaf*. En fait, il n'est pas possible de considérer un seul contexte à la fois, et un modèle théorique plus ample doit être construit pour rendre compte de la contribution de chaque contexte.

Le problème principal d'un tel modèle vient du grand nombre de contextes potentiels (il y a par exemple des raisons esthétiques qui auraient pu guider le choix du calligraphe, tel l'équilibre graphique des mots et de la ligne et la distribution des blancs et des noirs), de leur indistinctibilité (comment savoir quel contexte fut déterminant dans le choix?) et de la possibilité de les identifier (certains sont cachés — la justification de versets ne vient pas forcément à l'esprit d'un non-Ottoman —, tandis que d'autres sont perdus — telle l'humeur du calligraphe, qui certainement joua un rôle dans sa performance). Ici se trouvent les frontières de notre savoir présent — jamais entièrement franchissables avant que les nano-sousmarins et les machines à traverser les temps soient inventés un jour.

**TOI  
HABIT  
MAIS  
GENEROSITÉ  
LA PLUPART  
ANGES  
MOT 40%**

**QUITTER  
CHAQUE  
LIVRE  
ÊTRE 50%**



**COMME  
AVOIR  
HAÏR 20%**

**IMPIÉTÉ 70%**

Figure 4. — Aspects contextuels : force d'association entre signification et forme.

## BIBLIOGRAPHIE

- [1] Vlad ATANASIU. — *Hypercalligraphie. Le phénomène mamluk à l'époque du sultanat mamluk (Moyen-Orient, XIII<sup>e</sup>-XVI<sup>e</sup> siècle)*, thèse de doctorat, École pratique des Hautes Études, Paris, 2003,  
<http://mywebpage.netscape.com/atanasiuvlad/the/>.
- [2] Vlad ATANASIU. — « Visible *i'ajaz* — Three graphic particularities in the Qur'an copies », *Actes du colloque de Bologne 2003 sur les manuscrits coraniques*, François Déroche éd., E. J. Brill, Leiden, 2003,  
<http://mywebpage.netscape.com/atanasiuvlad/pubs/atanasiu2003iajaz.pdf>.
- [3] *Qur'an*, calligraphed by Hamid al-Amidi, Hizmet Vakfi, Istanbul, 1974.
- [4] Mohammed Tamimi. — *Catalogue of Winners' Plates in the Second International Calligraphy Competition*, Research Centre for Islamic History, Art and Culture, Istanbul, 1991.
- [5] Ibn WAHID. — *Qur'an*, Cairo, 1304-6 (London, British Library, ms. Or. Add. 22408).
- [6] Ibn ZAFIR. — *Sulwan al-muta'fi 'udwan al-ittiba'*, TRI Publishing, Kuwait, 1984.

✉ Vlad ATANASIU  
Massachusetts Institute  
of Technology (USA)  
atanasiu@mit.edu  
<http://mywebpage.netscape.com/atanasiuvlad/frq/>