

## Chapitre 9. La *Chronique des fouilles du BCH*\*

L'École française d'Athènes publie une douzaine de séries de monographies ainsi qu'une revue : le *Bulletin de correspondance hellénique* [BCH]. Cette revue annuelle est consacrée à des articles de synthèse, à des informations sur les activités de l'École et à la publication du matériel archéologique. Une de ses sections, la *Chronique des fouilles* a pour mission particulière de signaler aux lecteurs toutes les « nouveautés » archéologiques<sup>80</sup> sur lesquelles des informations fiables ont été obtenues au cours de l'année.

L'École française d'Athènes, souhaitant une valorisation accrue de ce corpus assez unique en son genre, projeta en 1996 sa mise en ligne<sup>81</sup>. Lors d'une étude de faisabilité, menée par nos soins<sup>82</sup> en 1998, il apparut que la richesse de sa mise en forme nécessitait un nouveau type d'outil informatique. Cet outil fut ensuite prototypé au cours de notre DEA puis de notre thèse pour devenir le système *Porphyre*. Ce que nous présentons dans ce chapitre concerne l'expérimentation menée à l'automne 2000 de gestion de la *Chronique des fouilles* à l'aide de *Porphyre* (expérimentation qui fut suivie d'une démonstration publique lors de la table-ronde interdisciplinaire de novembre 2000 à Athènes).

Dans un premier temps, nous procéderons à une étude de l'existant, en analysant en particulier les caractéristiques structurelles du corpus et leur évolution dans le temps. Dans un deuxième temps, nous détaillerons l'expérimentation proprement dite. Enfin, dans un troisième temps, nous verrons quels furent les retours de l'expérience, tant au niveau du prototype que de la manière de traiter le cas de la chronique.

---

\* Certains éléments de ce chapitre ont été présentés lors de la table-ronde "Sémantique et Archéologie" [Benel00] et de la journée d'étude sur les bibliothèques numériques organisée par l'École française d'Athènes, la Maison de l'Orient Méditerranéen et l'Université Lyon II [Benel02b].

<sup>80</sup> Fouilles, prospections, trouvailles fortuites, restaurations, muséologie, publications de matériel inédit.

<sup>81</sup> Etude inscrite au plan quadriennal 1996-1999.

<sup>82</sup> Lors d'un stage d'élève ingénieur en 4<sup>ème</sup> année INSA.

## 1. Etude de l'existant

### a. Présentation du corpus

La chronique fait son apparition dans le bulletin de 1920. Appelée « Chronique des fouilles dans l'Orient hellénique », elle a alors pour portée géographique l'ensemble de la Grèce antique. En 1936, elle adopte son titre actuel de « Chronique des fouilles en Grèce » et se cantonne depuis aux frontières actuelles de la Grèce. En 1959, parallèlement à la chronique en Grèce, apparaît la « Chronique des fouilles à Chypre » puis en 2000, la « Chronique des fouilles dans le Bosphore cimmérien ». Couvrant indifféremment, à l'origine, les fouilles de l'École et les autres, la chronique se scinde, en 1940, en deux rubriques traitant respectivement des unes et des autres. Enfin, à partir de 1970, les travaux de l'École n'apparaissent plus dans la chronique que sous forme de références à une section autonome du bulletin créée pour l'occasion.

Par son histoire et son volume d'informations, la Chronique des fouilles représente un outil pratiquement unique pour les chercheurs : quatre-vingt ans d'activité archéologique en Grèce et trente ans à Chypre. De plus, malgré cette quasi-exhaustivité, le volume du corpus (cf. Tableau 9.1) est suffisamment limité pour envisager sa numérisation intégrale (textes re-saisis, figures numérisées d'après originaux) dans des délais et des budgets raisonnables.

NOMBRE DE PAGES	OCCUPATION MOYENNE DES PAGES PAR LES FIGURES	NOMBRE DE CARACTÈRES
12.000 ± 500	(55,0 ± 2,5) %	30.000.000 ± 500.000

Tableau 9.1 : Volumétrie de la *Chronique des fouilles*<sup>83</sup>

<sup>83</sup> Les volumes indiqués ici s'appuient sur l'inventaire réalisé par nos soins en mai 1998 sur les 108 chroniques de fouilles disponibles, à savoir les 70 numéros de la Chronique des fouilles en Grèce (et en Orient Hellénique) de 1920 à 1995 (sachant qu'il n'en existe pas pour 1932 et 1946, et que sont regroupés respectivement ceux de 1940 et 1941, 1942 et 1943, 1947 et 1948) et les 38 numéros de la Chronique des fouilles à Chypre de 1959 à 1996.

**b. « Chronique » ou « Topique » des fouilles ?**

La *Chronique* est un corpus constitué de telle manière que chaque passage soit compréhensible indépendamment des autres (chacun traitant en général des découvertes annuelles pour un site géographique donné). Cependant, pour offrir une compréhension plus approfondie, des liens entre ces passages sont assurés par une structure complexe.

Cette structure est avant tout temporelle (comme le nom de *Chronique*<sup>84</sup> l'indique) : chaque livraison correspondant à une année de fouilles. Ensuite, dans chacune de ces livraisons, une hiérarchie des titres particulièrement profonde (au moins sur cinq niveaux) permet de regrouper les passages essentiellement par rapport à la géographie des sites dont ils traitent (donnant une impression de « zoom » d'un niveau à l'autre).

Il est intéressant de remarquer que cette deuxième structure est en partie invariante. Ainsi, il est envisageable à partir du corpus de la *Chronique des fouilles* d'obtenir automatiquement un nouveau corpus que l'on pourrait nommer « *Topique*<sup>85</sup> des fouilles » (cf. Figure 9.1). Ce traitement viendrait d'ailleurs généraliser une pratique existante consistant à rassembler les « tirés à part » concernant un site.

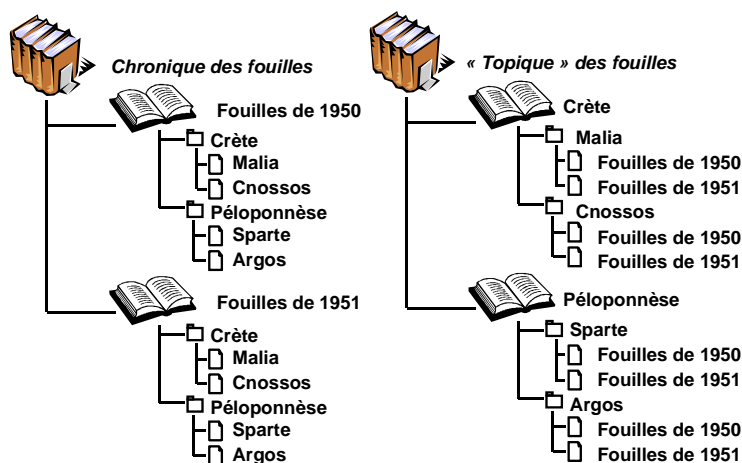


Figure 9.1 : La « *Topique des fouilles* », virtuellement présente dans la *Chronique*.

<sup>84</sup> En Grec, *Chronos* signifie « le temps ».

<sup>85</sup> En Grec, *Topos* signifie « le lieu ».

Afin de modéliser cette structuration multiple du corpus, il faudrait « décorrélérer » la dimension spatiale et temporelle en deux facettes (cf. Figure 9.2). La recombinaison en « chronique », « topique » ou toute autre combinaison des deux dimensions pourrait ainsi être calculée à la demande.

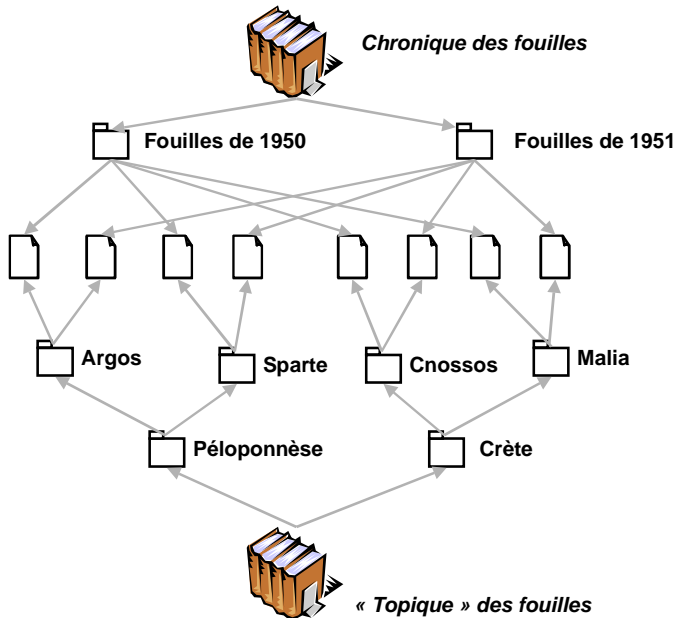


Figure 9.2 : Réseau de description présentant deux facettes pour un même corpus

## 2. Expérimentation

Notre expérimentation, menée en automne 2000, portait sur les quelques chroniques disponibles en texte intégral<sup>86</sup>. Nous basant alors sur la typologie courante distinguant dans le document numérique ses structures physique, logique et sémantique [NanardEtAl96], nous avons considéré :

- que la structure par page devait être omise (puisque en tant que structure physique, elle était calculable en fonction de la structure logique),
- que la structure actuelle en livraisons et en sections constituerait la structure logique, et qu'elle serait utilisée pour le stockage du document (en XML),

---

<sup>86</sup> Les chroniques postérieures à 1994 sont en effet archivées au format Microsoft Word par le service des publications. Cependant, il s'agit uniquement des brouillons envoyés au maquettiste, c'est-à-dire avant insertion des figures, mise en page et corrections diverses.

## CHAPITRE 9. LA CHRONIQUE DES FOUILLES DU BCH\*

- enfin, que la structure géographique et temporelle constituerait la structure sémantique.

La figure 9.3 correspond à la démonstration effectuée lors de la table ronde « Sémantique et Archéologie » (Athènes, novembre 2000). Elle montre que le prototype d'alors permettait une double navigation : à travers les documents XML et les réseaux de description. Elle présente également le réseau de description obtenu par le mode opératoire suivant :

1. convertir automatiquement la structure XML en réseau de description,
2. puis, dé-coreller à la main les dimensions temporelles et spatiales,
3. enfin, fusionner, toujours à la main, les structures spatiales de chaque livraison afin d'obtenir une sorte de thesaurus géographique contextualisé par des récits de fouille.

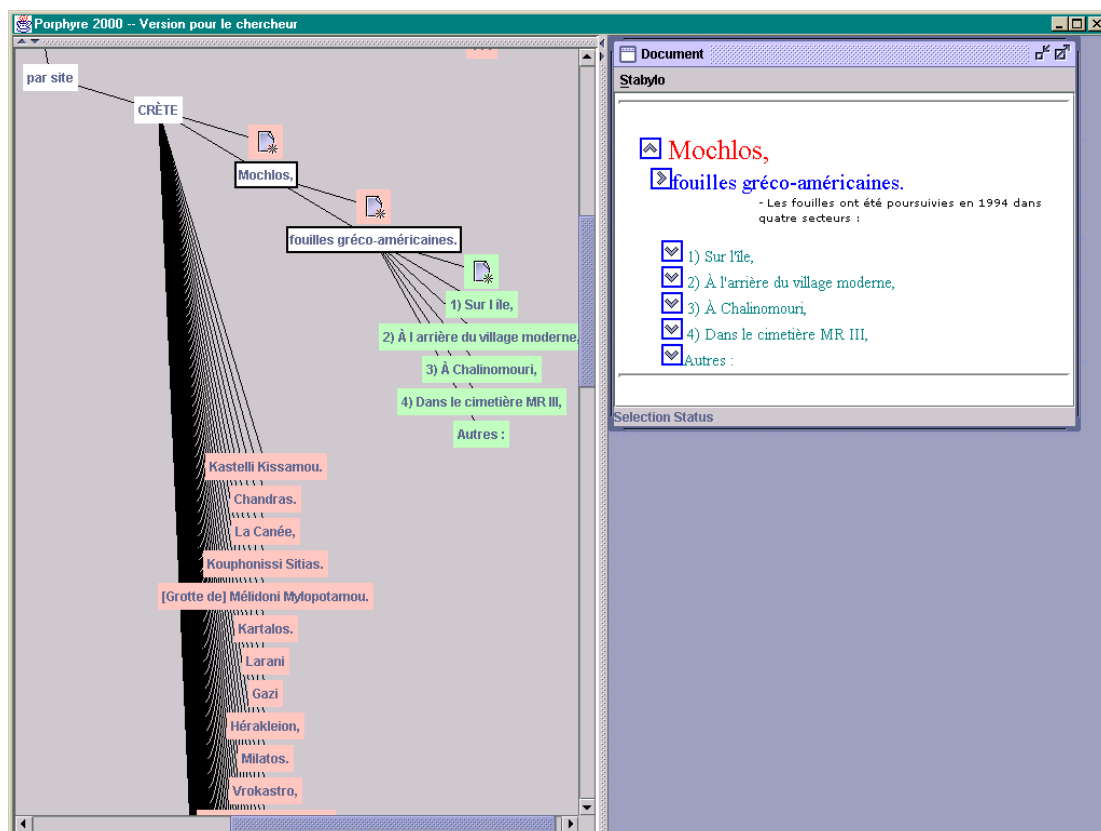


Figure 9.3 : Lecture avec *Porphyre 2000* d'un extrait de la *Chronique des fouilles*

### 3. Retour d'expérience

#### a. Limites rencontrées

Pour nous inscrire dans le cadre de la méthode poppérienne, nous ne nous appesantirons pas ici sur les résultats positifs de l'expérimentation et des démonstrations qui suivirent, mais chercherons plutôt les problèmes rencontrés.

Tout d'abord, il est apparu assez vite que la redondance des structures posait un problème cognitif. N'était-il pas « contre-intuitif » que le descripteur « Thrace » désigne dans la structure XML les fouilles d'une année, et dans le réseau de description celles de toutes les années ? Une autre erreur avait été de mépriser, comme on le fait souvent dans le domaine de la documentique, la notion de « page ». Que faire alors des références bibliographiques (depuis l'intérieur et l'extérieur du corpus), toutes de la forme « BCH 1995, p.409-410 » ? De plus, en raison de la différence de sujets traités par les différents passages, nous avons abandonné toute notion de séquence. Or les références croisées (« *ibid.* », « *op. cit.* ») ne peuvent être interprétées que si l'on peut parcourir les passages précédents. Enfin, il s'est avéré extrêmement éprouvant de fusionner les structures géographiques à la main.

#### b. Solutions proposées

Les évolutions du système *Porphyre* ont tenté de tenir compte des limites rencontrées. Tout d'abord, pour éviter la redondance entre la structure des fichiers XML et celle des réseaux de description, les fichiers XML ne sont plus, désormais, stockés tels quels dans *Porphyre*, mais convertis en réseau de description et objets documentaires (ce qui optimise d'ailleurs la plupart des traitements). Ensuite, pour introduire dans *Porphyre* la notion de séquence, nous avons mis en place les parcours de lecture (cf. Chapitre 8). Enfin, au sujet de la fusion semi-automatique de structures, nous disposons déjà d'un certain nombre de pistes (cf. Chapitre 12).

En ce qui concerne l'organisation du corpus, il sera souhaitable lors de la numérisation effective de la chronique de prendre pour objets documentaires des textes

## CHAPITRE 9. LA CHRONIQUE DES FOUILLES DU BCH\*

courts, des photographies, des plans, et de les organiser à l'aide des structures suivantes (cf. Figure 9.4) :

- un parcours de lecture modélisant la séquence de ces objets documentaires,
- un réseau de description modélisant leur description bibliographique (Revue, livraison, page), et reprenant ainsi indirectement la dimension temporelle (année de fouille),
- enfin, un deuxième réseau obtenu par fusion des structures spatiales de chaque livraison.

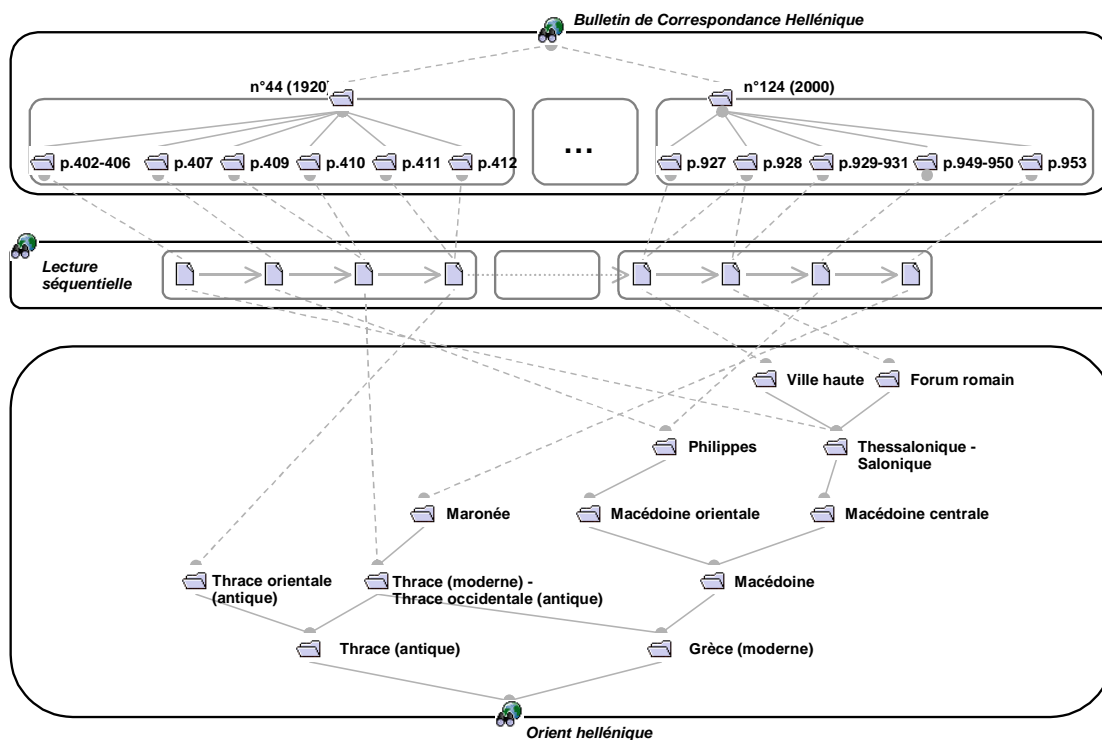


Figure 9.4 : Prise en compte du retour d'expérience : nouvelle structuration de la *Chronique des fouilles* à l'aide de Porphyre

En ce qui concerne les perspectives déjà évoquées d'assistance à la fusion, notons que l'exemple de la figure ci-dessus est particulièrement représentatif de la complexité de l'opération. En effet, nous avons dû ici tenir compte de changements de nom (Thessalonique/Salonique), de différences de granularité (ville/quartier) et de changements structurels encore plus radicaux (la Thrace moderne correspondant à la Thrace occidentale antique puisque la Thrace orientale antique est aujourd'hui en Turquie). On conçoit ici aisément pourquoi il serait vain de proposer pour cette tâche une automatisation dépassant le cadre d'assistants interactifs.

