

Les toponymes anglais: avec ou sans 'the' ?

Pour une approche systémique contrastive (Michel DELARCHE – EILA/LANSAD)

Les grammaires anglaises usuelles présentent une collection de règles empiriques du genre:

« on ne met pas 'the' devant les noms de villes, de régions, de comtés, d'états, de rues, de gares, de palais, de monuments »

« on ne met pas 'the' devant les noms de pays (sauf Netherlands, USA, UK, USSR) »

« on ne met pas 'the' devant les noms de montagnes et de lacs »

« on met 'the' devant les noms de fleuves, de massifs montagneux, de mers, d'océans et d'archipels mais pas devant les noms des îles singulières »

« on met 'the' devant les noms de bâtiments accueillant du public (administrations, musées, théâtres, cinémas, bar, restaurants, hôtels...) »

« on met 'the' quand le nom propre est composé avec 'of' (ce qui « explique » les exceptions UKBNI, USA, USSR, Isle of Dogs etc.)

Cette approche présente deux inconvénients majeurs:

- elle donne une pesante impression d'arbitraire,
- elle n'explique pas les nombreux contre-exemples.

Or, pour chaque règle, et pas seulement pour les noms de pays, des contre-exemples existent:

The Winter Palace, the Finland Station, the Broad Peak, the Black Lake, Orkney, the Borders, the Lake District, Rules, Raffles...

Faudrait-il se résoudre à inventer une collection de sous-règles encore plus évidemment arbitraires stipulant, par exemple, que les hôtels et restaurants dont le nom commence par 'R' et se termine par '-les' ne doivent pas être précédés de 'the' ?

En fait, une vision plus systématique peut être développée à partir de la mise en opposition de deux schémas bien distincts de construction des toponymes en anglais:

- des noms dérivés de noms communs qui ont été munis de qualifications descriptives permettant de préciser leur identification, et qui du coup prennent un 'the' (déterminant du type article défini qui marque la perte de généralité du nom commun d'origine,)
- des noms purement conventionnels (résultant d'un étiquetage arbitraire.)

Par exemple, parmi les noms de pays, USA, UK, USSR sont clairement des qualifications descriptives spécialisantes par rapport à leurs têtes de syntagme nominal respectives ('states', 'kingdom' et 'union') qui sont des noms génériques. The Netherlands peut pareillement se lire comme the + Nether + Lands (tout comme « les Pays-Bas » en français), sans pour autant en faire une exception ad hoc à la « règle des noms en 'of' ».

Buckingham Palace est une pure étiquette alors que 'the Winter Palace' est le palais où les tsars venaient passer l'hiver. De même, 'the Finland Station' est la gare de St Pétersbourg d'où les trains partaient pour la Finlande alors que depuis Victoria Station on ne va pas jusqu'à Victoria Falls... et pourtant l'original russe 'Финляндский вокзал' n'a évidemment pas d'article!

The Broad Peak (forme qui co-existe avec la forme standard sans 'the') ou the Black Lake relèvent également de ce type de qualification descriptive venant singulariser un nom commun.

Inversement, avant que l'on n'y installe l'encombrant cadeau d'Eiffel et Bartholdi, Liberty Island s'appelait Defoe's Island: on en a simplement changé l'étiquette sans en rendre le nom réellement descriptif pour autant.

De manière plus générale, la « règle » fréquemment citée par les grammairiens empiriques voulant que les noms d'archipels soient précédés de « the » (The Scillys, The Philippines...) alors que les noms d'îles singulières ne le soient pas peut s'expliquer aisément dans le schéma proposé ci-dessus.

Les noms anciens d'archipels sont descriptifs (par exemple, « the Cyclades » à cause de leur disposition en cercle ou « the Hebrides » qui sont les « îles en bordure de l'océan », *havbredey* en vieux Norse) alors que les noms des îles prises individuellement sont plutôt des noms-étiquettes; quand un nom d'île relève du schéma générique spécialisé, le « the » apparaît: d'où 'the Isle of Dogs' ou 'the Black Island' chère au capitaine Haddock ('the Black Isle' en Ecosse est un cas limite puisqu'il s'agit en fait d'une presqu'île...) Quant à 'Long Island', ce nom est moins descriptif qu'il n'y paraît, car c'est en fait une corruption-adaptation anglaise du nom original hollandais *Lange Eylant*. Même « Orkney » qui semblerait au premier coup d'oeil un pur nom-étiquette, possède deux étymologies successives, toutes deux descriptives: « île aux cochons » en celte et « île aux phoques » en vieux Norse, chacun de ces noms ne désignant au départ que l'île principal; il est intéressant de noter que du fait de la généralisation du « the » on rencontre maintenant fréquemment les formes erronées « the Orkney(s) [Islands] » ce qui montre bien une tendance à la normalisation des exceptions par ce qui est perçu comme étant la règle: Orkney désigne un archipel et « doit » donc être précédé d'un « the » (voire muni d'un 's'), au mépris de la forme singulière sans article, seule attestée historiquement, mais perçue comme « anormale » parce qu'ayant commencé par désigner une seule des îles avant d'être étendue à la désignation de l'archipel dans son ensemble.

Il faut donc souvent remonter loin dans l'histoire pour comprendre et évaluer le contraste entre les modes descriptif et conventionnel. Autre exemple: the Canary Islands étaient au départ peuplées de gros chiens ('canes' en latin) et 'canarias' avait donc bien au départ une valeur descriptive conservée en anglais; mais s'agissant ici d'un toponyme d'origine étrangère, il faut aussi prendre en compte l'influence potentielle de l'original espagnol 'las islas Canarias' de la même façon qu'on trouve des occurrences de 'the Mont St Michel' alors qu'il s'agit bien d'un nom-étiquette (anciennement dit 'Mont Tombe') tout comme 'the Palatine'.

Voici un autre cas où l'analyse étymologique vient appuyer la distinction entre les deux schémas: les noms de rivières et de fleuves ne semblent pas au premier abord descriptifs pour deux sous, et pourtant ils sont munis de 'the'... Faut-il introduire une règle spéciale pour les rivières ?

En fait, si l'on se penche sur l'étymologie des noms anglais de cours d'eau, on s'aperçoit que leur origine est très généralement descriptive, comme le montre le petit tableau suivant:

Anglais moderne	Origine étymologique	sens	Nom descriptif originel
The river Thames	Tamesas/Temasas (celte) Temesis/Tamesis (latin)	sombre	The Dark River
The river Avon	Afon (gallois)	rivière	The River-in-Welsh
The river Stour	Stour (vieil anglais cf. to stir, et stören en allemand)	Agité, troublé	The Troubled River
The River Spey	Spew	Crachant, vomissant	The Spewing River
The Great Ouse	Ouse (vieil anglais)	eau courante, flux	The Great Water Flow
The river Exe	Sca (celtique)	eau	The Water (River)
The river Wensum	Wendsum (vieil anglais)	tortueux	The Winding River
The river Yeo & the Somerset Yeo	Ea (saxon; cf. 'eau' en français) & gifl (celtique)	[Courant d']eau & fourchu	The Stream/Water & The Forked River
The river Trent	tros+hynt (celtique)	À travers les routes	The Ford Crossing River
The river Tyne	Tin (dialecte celtique)	rivière	The River-in-Celtic
The river Mersey	Maeres ea (saxon)	Rivière frontalière	The Border River
The river Derwent	Der went (celtique)	Vallée aux chênes	The Oak Valley River

On peut donc faire l'hypothèse que le 'the' marquait initialement la spécialisation du nom commun 'river' par divers qualificatifs descriptifs et que l'usage de l'article s'est très tôt imposé du fait de la domination initiale massive des noms de rivières par ce schéma descriptif dans la langue anglaise.

Autre problème, les noms d'hôtels ou de restaurants: Rules et Raffles peuvent certes être ramenés aux noms propres de leurs fondateurs respectifs : « Rules' Restaurant », « Raffles' Hotel » ce qui suffit à expliquer l'absence de 'the' sans introduire une nouvelle règle ad hoc.

Mais à l'inverse, expliquer la présence quasi-systématique du 'the' dans le cadre du modèle proposé ici demande également un petit retour en arrière: au Moyen-Âge, la plupart des gens étaient illettrés et le repérage des auberges et estaminets se faisait par les enseignes imagées qu'ils portaient en façade et dont la tradition s'est poursuivie jusqu'à nos jours: The Golden Bear Hotel, The Black Dog Tavern, The Crown Inn, The White Horse Bar etc. Ainsi, le 'the' hérité de ce système primitif d'identification descriptive a pu se fixer dans la langue et s'étendre ensuite aux autres lieux de loisir (dont certains étaient d'ailleurs tout aussi intrinsèquement descriptifs comme « The Globe ».)

Inversement, des noms apparaissant de prime abord comme fort descriptifs peuvent n'avoir pas de 'the', pour diverses raisons. Des noms comme Marble Arch ou Palm Springs auront été d'emblée assimilés aux noms-étiquettes qui dominaient déjà leur catégorie de lieux-dits à l'époque de leur création; à l'échelon de l'évolution de la langue, ce sont des toponymes de création très récente: 1828 pour Marble Arch (et même 1855 dans son emplacement actuel) et 1860 pour Palm Springs. Certains noms étymologiquement descriptifs mais qui n'apparaissent comme tels que pour les « savants » comme 'Mesopotamia', auront également suivi le schéma dominant des noms-étiquettes. On peut remarquer que d'autres dénominations telles que 'Northern Ireland' ou 'South Korea', 'East Germany' ne sont pas réellement descriptives mais restent des interprétations conventionnelles symboliques (le 38ème parallèle sert de frontière entre 'sud' et 'nord', et le Donegal, qui fait partie de l'Irlande 'du sud', n'est pas moins septentrional que l'Ulster); des noms ainsi dérivés des noms de pays peuvent continuer d'être considérés comme étant des noms-étiquettes plus symboliques que descriptifs, même si la présence ultra-majoritaire des noms-étiquettes dans la catégorie des noms de pays ou de régions suffit à expliquer l'absence de 'the' dans ces dérivés-là (contrairement à « the Borders » ou « the Lake District »)

Lorsque la distinction est incertaine, l'usage est souvent flottant: ainsi on trouvera 'the Wembley Stadium' (localisation descriptive) mais aussi 'Wembley Stadium' tout court (ie un nom-étiquette, comme si le quartier de Wembley se résumait à son stade de football emblématique.)

Dernier point: on rencontre un autre schéma sans 'the' dans des constructions telles que 'Mount Everest' ou 'Loch Ness' où les noms communs génériques 'mount' ou 'loch' fonctionnent en fait comme des adjectifs antéposés au nom-étiquette: c'est un schéma inverse de: 'the Rocky Mountains' ou le nom générique 'mountains' est précisé par l'adjectif 'rocky', le tout formant un nom propre fléché par le marqueur de définition 'the': la différence structurale est ici la même qu'entre 'President Obama' et 'the American President'.

En résumé, cette approche contrastive opposant le nommage descriptif dérivé des noms communs au nommage conventionnel symbolique permet de comprendre la présence ou l'absence de 'the' en relation avec l'étymologie des toponymes anciens, mais cette interprétation doit être modulée par:

1°) l'effet d'attraction exercée sur les toponymes récents par la prégnance du schéma préalablement établi dans chaque catégorie: les rivières qui portent des noms-étiquettes d'origine récente prennent pourtant le 'the' à cause de l'époque lointaine où tous les noms de rivières étaient descriptifs, alors que, symétriquement, des toponymes descriptifs de création récente pourront être traités comme des noms-étiquettes parce que ceux-ci sont largement majoritaires dans leurs catégories.

2°) l'influence possible des autres langues pour les toponymes étrangers (las Islas Canarias, le Mont Saint Michel, Il Palatino..)