0. THE NOTION OF A TOPONYMIC CONSTRUCTION. In this paper, I call two nouns or noun phrases a *toponymic construction* if the following is true:

- one of them (A) is a toponym;
- the other one (B) specifies the class of geographical object to which A belongs;
- A and B are within the same noun phrase (we omit some technical details here);
- the fact that A is the name for B is not stated in the noun phrase that contains them.

We’ll refer to the element that specifies the class of geographical object (B) as the *generic component*. Accordingly, _the city of Moscow_, _New York State_ and _Lake Michigan_ are toponymic constructions (although it may be argued that the last two are proper names). The following expressions are not: _Moscow, New York, city, the state, beautiful country, this wonderful Moscow, New York is big_ (lack of one component); _Everest is a mountain, this city is called Kyiv_ (the components are not in the same NP); and _the village named Ouassadougou_ (the naming relation is stated explicitly).

In many of the world’s languages, the syntactic relationship between the components of toponymic constructions is of the same type as in some of the language's possessive constructions. The toponymic constructions that show the same syntactic properties as possessive phrases may be called *genitive toponymic constructions*, as opposed to *appositive toponymic constructions*.

A genitive toponymic construction can be marked in a number of ways, e.g. with a preposition (Spanish _ciudad de Madrid_ ‘the city of Madrid’), an izafet marker (Persian _shahr-e Tehran_ ‘Tehran city’), a genitive case marker (Finnish _Helsingin kaupunki_ ‘the city of Helsinki’), a possessivity marker (Hungarian _Moszkva varos-a_ ‘Moscow city’), a possessive adjective suffix (Old Church Slavonic _гора елеоньска_), or simply null (Komi _Ухта кар_ ‘Ukhta city’, which shows the same syntactic properties as _Вася керка_ ‘Vasya’s house’).

1. IDENTIFYING THE HEAD OF THE PHRASE. It appears that within genitive toponymic constructions, the generic component is always the head of the noun phrase, just as it is in the corresponding possessive construction, e.g. _the city of Moscow_ and _the father of Mary_. For appositive constructions this issue is not so clear.

To deal with this, we make the following assumptions: 1) one of the components (the generic or proper term) is the head of the phrase; 2) the construction is endocentric (that is, the dependent component can be omitted without affecting the grammaticality of the sentence).

The omission test resolves the issue in most cases and identifies the generic component as the head of the phrase. It resists being omitted for a number of reasons:

1. the generic component may control the agreement of the noun phrase as a whole (Serbo-Croatian _Grad [M] Sarajevo [N] je lijep [M] / Sarajevo je lijepo [N]_ ‘Sarajevo city/ Sarajevo is beautiful.’);
2. the generic component, unlike the toponym, may co-occur with the article (Afrikaans _Ek woon in die dorp Stellenbosch / in Stellenbosch_ ‘I live in Stellenbosch village.’ / ‘in Stellenbosch.’);
3. only the generic component, not the toponym, takes the article or case affixes required by the noun phrase (Albanian _qytet-i [river-DEF] Elbasan_ ‘Elbasan river’ / _Elbasan-i [Elbasan-DEF]_; Moksha Mordva _мон качэнь касонь Оцяду веле-са [LOC] / Оцяду-са [LOC]_ ‘I was born and grew up in Achadovo village.’ / ‘in Achadovo.’)

---

1 Plank 2003: 34.
I am aware of only three cases where the head is the toponym.

The first example is an English toponymic construction where the generic component is preceded by the toponym. Other than in established proper names (like New York City), the distribution of this construction is restricted to the names of states (only in order to disambiguate the context, e.g. “Did you see the White House when you were in Washington?” – “I was in Washington state, not the city.” *Kansas state* is also attested, but not *California state*), cities (e.g. *San Francisco city*, which is attested, though some speakers find it strange) and some others (e.g. *Hudson river*), where the omission of the toponym leads to ungrammaticality unless the article is inserted: *It's unusual to see elephants in Washington city/* in the city/* in the Washington city*. Note that within other toponymic constructions that English is rich with the article is necessary: *Hudson river, Boston city, but the river Hudson, the city of Boston* (here the generic component is the head of the phrase).

Another example is taken from Adygh2, where the generic component has a special case marking: *(Псыхъо-у [river-ADV]) Москва мычыжьэу цыыI* ‘Moscow river is not far.’ / *(Псыхъо мычыжьэу цыыI* ‘River is not far.’

The third example is provided by modern Chinese, where special lexemes (not allowed elsewhere) are used within a toponymic construction: *chengshi ‘city’/* Taipei (*cheng)shi ‘Taipei city.*

Our assumptions are not sufficient to identify the head of appositive toponymic constructions in Spanish and Portuguese. In these languages substantives don’t have morphologically marked case; both toponyms and generic components require an article (as all common nouns do); the gender of the components of the appositive construction is always the same (Spanish *estoi en el [M] lago [M] Titicaca [M] / en el Titicaca / en el lago* ‘I’m at Titicaca lake.’ / ‘at Titicaca.’ / ‘at the lake.’) The examples from these languages seem ambiguous.

2. THE ORDER OF THE COMPONENTS. Let’s set aside for the moment the non-productive constructions that may be regarded as proper names, attributed to specific toponyms, e.g. the Russian *Москва-река* and the English *Lake Michigan*. An interesting observation can be made regarding toponymic constructions, both genitive and appositive. In left-branching (“postpositional”) languages the head’s unmarked3 position is to the right of the dependent component, while in right-branching (“prepositional”) languages it is to its left. The ambiguous data of Spanish and Portuguese don't contradict this generalization. The generalization holds, no matter which of the components (the toponym or the generic component) is the head.

The toponymic constructions are similar to pre- and postpositional phrases and possessive phrases in one other respect: the components of toponymic constructions can be separated with the same set of elements that the corresponding possessive constructions can be separated with: *the city of Boston, like the mother of Mary;* the Russian, poetically marked *город прекрасный Москва* ‘the beautiful city of Moscow’ like *город прекрасный Петра* ‘the beautiful city of Peter.’

In other words, it seems very natural to regard the dependent component of the toponymic construction as filling a complement position of the head. This position corresponds to the generic component’s semantic valence for a toponym (in most cases) or that of the toponym for a generic component (in several cases). The principles (cognitive, pragmatic, etc.) governing the choice of one or another syntactic (and semantic) structure for toponymic expressions in a given language is a separate issue that I leave without analysis for now.

REFERENCE


---

2 The Adygh data were kindly provided by Yu. Lander.

3 Some languages allow the reverse order as well (Avar *Maxicable шаьгар шаьгар Maxicable* ‘the city of Makhachkala’); such languages are not numerous.