

How to collect the names information: office preparation, multilingualism

Ferjan Ormeling

Introduction

The only way to check whether the names are actually used is to go out into the terrain and find out what names people are using in order to relate to their environment, but before leaving the office much should be done in order to prepare for this visit to the terrain, because we will probably only pass there once, and at that time we should collect all the information necessary.

1. Sources

We will not have been the first topographers to go out into the field, and in most cases there will have been files left by our predecessors that mapped the area before. These *files* may contain names lists, perhaps even annotated names lists on which experts have corrected the names collected in the field. The files might still contain *aerial photographs* for the production of the current map edition and of course the *current map edition* itself. For every new mapping new imagery will have been made available and we will thus be able to compare the previous aerial photographs with the new imagery, and thus deduce what new elements were added: new roads, railways, dams, village or town extensions, new areas under cultivation, new tourist developments. But also villages that were burned, disappeared under a reservoir, or were left because of the drought will affect the contents of the map.

From the statistical office our sources will be lists of numbers of inhabitants of villages, towns, administrative units, and these will help us to decide on the font size of the names. In some countries the population number is added to the map, and thus also to the gazetteer.

In many countries most of the mail that the national mapping agency receives refers to faulty place names. People would complain about the mistakes in the spelling or about mistakes in the location of the name on the map or about mistakes in the size of the area named. When these complaints are filed, or even used in a continuous updating process, they will also be used in the preparation of the fieldwork. Also many government workers that have been out in the terrain with topographic maps will have found that there are inconsistencies in the naming. In the UN manual (p.49) the following forms of inconsistencies are mentioned:

- a) More than one name exists in local oral usage for the same feature (allonyms)
- b) local oral usage applies the same name to two or more different features (homonyms)
- c) there is local disagreement in the spelling of the name
- d) local records differ from local oral usage
- e) names in topographic maps differ from local usage
- f) maps apply the names to features different from the one to which the name is applied locally

On the basis of these inconsistencies one should prepare a list of issues to be solved in the field that the topographer should bring along to the fieldwork area (document 1).

2. Features

The main task of topographers sent out into the terrain is to collect information about the features that should be rendered on the topographic map. One of the attributes of the features to be rendered is their name. One is therefore sent out to collect names of specific feature categories only. Of course one has to know what topographic features need names on the map. Usually for topographic maps or files these feature categories are settlements (groups of buildings for residential purposes), other man-made buildings, traffic infrastructure (roads, traffic lights, fuel pumps, etc), administrative information (boundaries, town halls, capitals), terrain features (hills and mountains), hydrography, physical areas (islands, plains, plateaus) and conservation areas.

It should be clear what each of the feature categories refers to. When both hill and mountain are discerned as separate feature categories, it should be clear when a terrain outcrop is a hill and when it no longer is a hill but a mountain. The same goes for the difference between streams and rivers. One can also abstain from differentiating between hills and mountains or between rivers and streams, and leave it to the local generics to make the difference. Dimensional differences with generic consequences – hills vs. mountains, hamlets and villages (where no official status definitions are involved), gulfs and bays – cannot be unequivocally established, as they are based on comparison. A mountain (*berg*) in the Netherlands may be too low even to be considered a hill in an Alpine country. To relieve the toponymist from the burden of this kind of subjective decisions, the generic classification should actually not include such distinctions. Both hills and mountains should be one and the same class that might be titled for instance ‘named elevation points’, just as rivers, brooks and torrents should better be combined in a common category like ‘streams’, or something like ‘linear hydrographic features of natural origin’ (as opposed to canals). (Tichelaar)

Generic terms are the common nouns that describe topographic features in terms of its characteristics and not by its proper name. Examples are mountain, river, wadi, hamlet, etc. If possible, each generic term in use should be defined. If for instance terms like stream and canal are used, it should be clear that the one is natural and the other man-made. If a difference is made between swamps and marshes, it should be clear that a swamp has larger trees, more and deeper open water in comparison with a swamp that has lower vegetation, consisting of reeds or shrubs. It is good practice to compose a list of generic terms one probably would encounter in the fieldwork area, prior to going there, as this would help in the standardisation of the generic terms to be used (document 2). This list should also contain words or name parts recurrent in place names, like north, south, upper, lower, small, large, old, new, etc.).

Some examples of language-, country- or region-dependent generic distinctions in other parts of the world are:

- in French: *fleuve* vs. *rivière* (a river flowing into a sea or into another river)
- in Finnish: *vaara* vs. *tunturi* (a wood-covered mountain vs. a bald mountain)
- in North African Arabic: *erg* vs. *sarir* (a sandy desert vs. a rocky desert)
- in English: *marsh* vs. *swamp* (open wet ground vs. wet ground with trees)
- in Dutch: *sas* vs. *zijl* (lift-lock vs. discharging-sluice)
- in German: *kugel* or *kuppe* vs. *spitze* (a dome-shaped vs. a pointed mountain top)

- in English: *convent* vs. *monastery* (an abbey for nuns or for monks)

Obviously the above list could be endless, and even to make it more or less exhaustive within one language area would require quite some study. As a general rule, the most dominant elements of both the physical and the cultural landscape show in any region the largest conceptual and onomastic differentiation. The Greenlanders distinguish many different types of snow, while toponyms in the Netherlands are full of swampy generics (*broek, hol, moer, veen*), mountainous France knows numerous models of mountain tops (*aiguille, ballon, butte, colline, crête, dent, dôme, mont, roche*), Norway all kinds of bays (*anger, bogen, fjord, våg, vik*), the Arabic and Berber world a great variety of deserts (*edeyen, erg, hammada, ouar, sahra, sarir, ténééré, tanezrouft*) and, toponymically dominant because of their scarcity, sources of water (*ain, bir, hafar, hammam, hassi, umm*). (Tichelaar)

Sometimes other attributes of topographic features have to be collected as well, such as their height.

3. Practical considerations: the use of language

In the office certain practices will have developed concerning the use of the standard language: of course it will be expected of the topographic staff that they are well aware of the grammar and spelling rules of the standard language to be used for the maps, as names should be spelled as much as possible in accordance with current orthographic practice. But apart from that they should know how to deal with abbreviations (if one uses them on the map or in gazetteers they have to be standardised and explained), and with compound names (with or without hyphens). When names are too long on the map they might be broken up with hyphens for instance. Another practical consideration is when capitals in stead of lower case letters should be used.

If *homonyms* (each of two or more identical toponyms denoting different topographic features) occur, one has to devise a means how to differentiate between them. If *allonyms* (each of the two or more toponyms that refers to the same topographical feature), rules have to be in existence which one to select. In bilingual areas where more than one name versions is used on place name signs, the rule could be for instance to use the one name version at the top.

4. Policies to decide on in advance

Before going out in order to collect names for the maps and gazetteers, topographers should be aware of their office's general policies regarding geographical names. Amongst such policies could be, for instance,

- a) Never to name a feature after a living person and withstand political pressure to do so
- b) Names publicizing brands or products should not be accepted
- c) Render local pronunciation or standard orthography?

- d) Use local generics or official generics? (The official word for *river* in the Indonesian language is *sungai*. But local languages might have local words for river, like *kali*, *bengawan*, *ayer*, etc.)
- e) No names to be accepted that only consist of personal names (like *Joe* or *Margaret* or *Ivan*)
- f) What to do when no names exist?
- g) What to do with unwritten languages or with special scripts?
- h) Only one name variant should be selected as the official one.
- i) Unnecessary name changes should be avoided

5. Linguistic aspects

Topographers have trained in collecting topographical information, and hardly ever in linguistics. Yet it is expected of them to correctly write down the names they collect, even if they do not speak the local language. So it will help if they bring interpreters along that speak the local language, and can help in finding the proper name forms and spellings of the names that are given by informants. When conducting an interview or taking note of names information in another way, the peculiarities of the languages in question should be carefully taken into account.

Prior to the fieldwork the location of linguistic boundaries in the area should be studied; language boundaries shown in map (if available) should not be taken for granted. When collecting the names these should be put in their proper orthography, and for this it would help if the meaning of the names is also asked; interpreters might be able to determine the proper name form in its proper case, without any prepositions of direction or attachment of place (locativus) and in their definite or indefinite form.

In some languages, names and other nouns are subject to morphological changes because of declination (Indo-European languages like, for instance, Russian) or the addition of affixes and infixes (synthesizing languages like Uralic Finnish) depending on their *case*. The basic form of the name would then be the nominative case. In Finnish, for instance, the complete name of one of the provinces is *Uudenmaan Lääni*. *Uudenmaan*, however, is in the genitive case ('the Province of ...'); the name of the province (nominative case) is actually *Uusimaa*. *Uusimaa* itself consists of a generic element *maa* ('land') and a specific element, in this case the adjective *uusi* ('new'). To make it genitive, both elements receive an affix *-(e)n*, which in the case of *uusi* causes a consonantal mutation producing *uuden*. Someone ignorant of this aspect of the Finnish language would be in risk of drawing a wrong conclusion, and, knowing that *Lääni* means Province, suggest that the name of the province is *Uudenmaan*. Likewise, the Russian *ostrov Karla-Aleksandra* is not an island (*ostrov*) named *Karla-Aleksandra*: the island is named after (male) Karl-Aleksandr.

In some Indonesian languages, like Javanese, Balinese, and Sundanese, completely different language versions are used depending on the mutual (social-hierarchical) relationship between the parties in discussion. In Javanese, there are five such language 'levels' called (from high to low) *krãmå inggil*, *krãmå biasa*, *madya*, *ngoko madya*, and *ngoko biasa*; in

Balinese there are even seven. The respect enjoyed by the first visiting topographers, as representants of the Netherlands Indies colonial government, from the side of the Javanese villagers in the 1870s, made these villagers address their interviewers in the higher *krāmā* version. They would thus mention place names like *Mambeng*, *Pelujengan* and *Pengangsalan*. According to guidelines published by the Topographic Service in 1906, however, geographic names should be put on the map in their *ngoko* forms, which are for the above mentioned places *Malang*, *Besuki* and *Pekalongan*.

In many languages, words like nouns, adjectives, articles, and also names may have a gender. In Romance languages for instance, names may be masculine or feminine. It is important to know the gender, because this governs the form adjectives and articles take in case the name is used syntactically. In French, for instance, the country name *France* is feminine (taking the article *la*; ‘of France’ would thus be translated ‘de la France’); *Brésil* (Brazil) on the other hand is masculine; it is therefore *Plateau du Brésil* instead of *de la Brésil*. The river *Rhône* is masculine, so it is *Côtes du Rhône*, but the *Seine* is feminine, so it was the *Département de la Seine*. As *Guyane* and *Hollande* are feminine, it is *Guyane française* and *Hollande septentrionale* (with *e*); but masculine *Brabant* makes *Brabant septentrional*.

In Welsh, the first letter of female nouns (including names) are subject to consonant mutation when preceded by the definite article *Yr...*

6. Representation of pronunciation

If no interpreters can be brought along, at least the pronunciation of the names as given by the informants should be recorded. A standard notation of pronunciation is called for, or sound recording equipment should be brought. Potentially confusing phonemes – those being similar to the toponymist's ears – should be identified.

Concentrate on (combinations of) letters that represent more than one phoneme (*Enschede* contains the vowel *e* three times but their pronunciations are not similar. The first *e* is pronounced like *e* in *men*. The second *e* is like *e* in *wonderful*, and the third one is like *ay*).

To record the pronunciation for toponymic purposes, it suffices to note down those sounds that might lead to confusion. As this is obviously language-dependent, the toponymist needs to be well aware of the phonological means of expression used by the language(s) involved. In other words, she needs to know which variations in sound are (potentially) semantically discriminatory, or *meaningful*. This requires, in the preparatory stage, assistance of expert linguists. In English, for instance, the pronunciation of the words *thong* (‘belt’) and *tongue* differs only in the articulation of the initial dental sound, the first being fricative, the second occlusive. In another language, such differences in articulation might be meaningless personal variations in speech. In Finnish, the difference in length of the first vowel in *vara* (‘reserve’) and *vaara* (‘wooded hill’) is meaningful, while it would not even be noticed by a speaker of a language as, say, Indonesian, that does not make use of vowel length to discriminate between words. (Tichelaar)

Variation of the sound volume, called *accentuation*, is sometimes meaningful. In Dutch, the word *voorkomen* means 'to occur' if the accent falls on the first syllable, but 'to prevent' if it falls on the second one. A free accent may to a certain extent be subjective, as illustrated by the usual context-less accentuation of Dutch place-name *Amsterdam* on the third syllable (*dam*), against an intentional accentuation of the first syllable of both place-names in the phrase *Amsterdam en Rotterdam* ('Amsterdam and Rotterdam') to put emphasis on the distinction between the two.

Accentuation, even where it is semantically discriminatory, is rarely graphically represented in written language. Where it is, accents are usually shown by diacritical marks (also called *accents*). In Spanish and Italian orthography, for instance, acute resp. grave accents are used where the word accent is in contravention of the general rules, although in current Italian practice this is only done where it concerns the final syllable (the grave accent on e and o then simultaneously functions to indicate an 'open' vocal sound instead of the usual 'closed' one). Anyway, where accentuation is considered to be an essential aspect of pronunciation, it would be a waste not to take note of it at the event of a systematic toponymic field-work. (Tichelaar)

7. Multilingual aspects

As all inhabitants of the areas should be treated equal, all of them have equal rights to have their toponyms recorded. If the map has only limited space, only the names used by local majorities might be incorporated. But if space allows, more than one name could be used for each topographical feature.

From the language map find the concentration of the language communities can be found out, and from statistical sources their relative numbers can be assessed. With the names incorporated on the map the linguistic pattern can be represented.

But, even if not incorporated on the map the toponymic heritage of all linguistic groups should be preserved. No significant name elements should be suppressed

8. Preliminary logistics

- Prepare cartographic material, for name inventories (or preliminary name models), with question marks for issues to deal with
- Check administrative boundaries, for introduction to local officials (permits/cooperation)
- Prepare fill-in name forms, recording equipment, recording computers/handhelds or field notebook
- Try to locate informants in advance, and inform local officials in advance

Sources

Tichelaar, Tjeerd (2003) Field collection of names in multi-lingual and multi-cultural areas. PP79-85, proceedings Training Course on Toponymy Enschede/Frankfurt/Berlin 2002.
Mitteilungen des Bundesamtes für Kartographie und Geodäsie, Band 28. Frankfurt am Main:
Verlag des Bundesamtes für Kartographie und Geodäsie.