

**Please cite as:**

Bohnemeyer, J. (2001). Toponym questionnaire. In S. C. Levinson, & N. J. Enfield (Eds.), *Manual for the field season 2001* (pp. 55-61). Nijmegen: Max Planck Institute for Psycholinguistics. doi:[10.17617/2.874620](https://doi.org/10.17617/2.874620).

---

**REGULATIONS ON USE**

**Stephen C. Levinson and Asifa Majid**

This website and the materials herewith supplied have been developed by members of the Language and Cognition Department of the Max Planck Institute for Psycholinguistics (formerly the Cognitive Anthropology Research Group). In a number of cases materials were designed in collaboration with staff from other MPI departments.

**Proper citation and attribution**

Any use of the materials should be acknowledged in publications, presentations and other public materials. Entries have been developed by different individuals. Please cite authors as indicated on the webpage and front page of the pdf entry. Use of associated stimuli should also be cited by acknowledging the field manual entry. Intellectual property rights are hereby asserted.

**Creative Commons license**

This material is provided under a Creative Commons Attribution-NonCommercial-ShareAlike license (CC BY-NC-SA 4.0). This means you are free to share (copy, redistribute) the material in any medium or format, and you are free to adapt (remix, transform, build upon) the material, under the following terms: you must give appropriate credit in the form of a citation to the original material; you may not use the material for commercial purposes; and if you adapt the material, you must distribute your contribution under the same license as the original.

**Background**

The field manuals were originally intended as working documents for internal use only. They were supplemented by verbal instructions and additional guidelines in many cases. If you have questions about using the materials, or comments on the viability in various field situations, feel free to get in touch with the authors.

**Contact**

Email us via [library@mpi.nl](mailto:library@mpi.nl)

Max Planck Institute for Psycholinguistics

P.O. Box 310, 6500 AH, Nijmegen, The Netherlands

# Toponym questionnaire

design Jürgen Bohnemeyer in consultation with Nick Enfield and others<sup>3</sup>

## 1.0 INTRODUCTION

- **Relevant projects:** Space (subprojects on motion event descriptions, places and placenames, and topological relations); Event Representation (to the extent that the subproject on motion event descriptions is a joint venture with Space).
- **Nature of the task:** Checklist. This is not an elicitation tool. The idea is to make sure you have a comprehensive answer to each of the research questions in the questionnaire. The questionnaire does not detail a general methodology for obtaining the answers, although it gives hints occasionally at criteria for what might count as an answer and examples that suggest possible test frames.
- **Priority:** This has **high priority for people interested in the places & placenames subproject** (which currently has relatively low priority itself, as it is in an exploratory phase) and medium to low priority for people interested in the subprojects on motion event descriptions and topological relations. It has no specific priority for other researchers. However, the questions are of general anthropological interest, and any comprehensive description of a language should include the information on toponyms the questionnaire is targeted at.
- **Motivation:** Placenames are at the intersection of spatial language, culture, and cognition. They provide a way to refer to socially inhabited space by naming the places referred to, rather than the objects or people that occur at the places. Presumably, places referred to by toponyms are places that play a marked role in the life of the language community. Thus the toponyms of a language community embody a knowledge structure that figures prominently in the spatial conceptualization of the community's environment. At the same time, the way reference to places is distinguished from reference to objects, animals, or people *at* places is an important piece in the puzzle of the 'natural language metaphysics' that underlies spatial reference and conceptualization in the language under study.  
There are three overarching questions: how to formally identify placenames in the research language (i.e. according to morphological and syntactic criteria); what places placenames are employed to refer to (e.g. human settlements, landscape sites), and how places are semantically construed for this purpose. From the point of view of the subprojects on motion event descriptions and on topological relations, the interest in these questions resides in the distinction between objects (or people/animals) and places (as named by placenames) as referential grounds in descriptions of location and motion. In case your research language treats places differently from objects in this respect, your account of motion and location descriptions would be incomplete unless you have the answers to this questionnaire. From the point of view of the subproject on places and placenames, the answers to the questionnaire should determine the basic formal and semantic properties of placenames and thus lay the ground work for any further research on discourse about places
- **How to run:** Since the Questionnaire is not an elicitation tool, you may in principle obtain all answers from searching your existing data base. But if the information you find in your corpus regarding a specific question proves insufficient, inconclusive, or unclear (as it most definitely will in some instances), such gaps should be filled in a classical elicitation approach (**three consultants** recommended). To this end, the Questionnaire suggests occasionally possible test frames and analytical criteria.

---

<sup>3</sup> ... and implementing valuable suggestions by Jane Simpson, which are gratefully acknowledged.

- **Coding & Recording:** The intended result that we are hoping to obtain from each researcher has the format of concise descriptive prose augmented by fully glossed and translated examples. Each lexical item, morphosyntactic pattern, or phenomenon of semantic behavior that you comment on should be illustrated with at least one example. Cross-reference questions in your answers. If you elicit answers, you may directly transcribe the response; there appears to be no general need to tape Questionnaire sessions. Supplement your report with an indication of the magnitude of the set of placenames your analyses are based on. If the set is small (for example because you're studying a small community inhabiting a relatively confined area), include a comprehensive **list of the placenames**; otherwise, list an inventory of examples. Also, please add a very rough sketch of the topography of the area where the language is spoken; the predominant mode of production in the language community (e.g. hunter-gatherers, nomads, agriculturalists); the predominant settlement pattern; and, if relevant, the political organization.
- **Comments:** The questionnaire may be improved by adding further aspects relevant to the study of placenames and by considering behavioral properties of placenames in more languages. If you have suggestions, please send a note to: bohnem@mpi.nl.
- **Conclusions:** If you have run the Questionnaire, please let Jürgen Bohnemeyer know. The output will be compared and discussed in the Space subproject on places and placenames.
- **Citation:** Publications should cite: "Toponym questionnaire, version 2, designed by Jürgen Bohnemeyer and colleagues, Field Manual 2001, Language and Cognition Group, Max Planck Institute for Psycholinguistics". There are no copyright restrictions, and the Questionnaire may be downloaded from the Institute website, but please let us know of any research project undertaken with them, and do cite the source.

## 2.0 THE QUESTIONS

**Organization of the Questionnaire:** The first two questions (a-b) are mainly aimed at morphosyntactic criteria that allow to identify (a subsets of) place names as a form class in the research language. An initial criterion, the Place name Filter, serves to identify possible candidates for place names, and to exclude metaphorical places from consideration in case other criteria are insufficient to do that. Note: All languages will have nouns that pass the Place name Filter, even if they don't constitute a form class. If your research language fails to single out place names under a special form class, please answer the remaining questions with respect to the set of nouns that satisfy the Place name filter. But if your research language does distinguish a form class of place names, please focus on this form class when attending to the remaining questions (but keep other nouns that pass the Place name Filter in the background!). The final two questions (e-f) address the lexical semantic and referential properties of place names. But as always, formal and semantic aspects need to be studied *in their interaction*, not in isolation. Thus, the range of spatial relators place names combine with may define them as a form class, but at the same time reveals the semantic or conceptual construal of the places they refer to. Therefore, (a-b) and (e-f) are perhaps best tackled in loops. Questions (c) and (d) address etymological and sociolinguistic properties of place names.

**Abbreviations:** cN – common Noun; DP – determiner phrase; GDA – ground-denoting adjunct (e.g. *in London*; *on the table*; *out of the box*); L – research language; N – noun; NP – noun phrase; PN – place name; propN – proper name

**The PN Filter:** N is a *possible* PN iff

- (a) N is a *name*, not a cN (e.g. it cannot normally occur with indefinite determination, quantifiers, pluralization, or numerals; cf. *\*a/some London*, *\*three Londons*);
- (b) N's denotation involves a distinct and stable location; and
- (c) N can denote referential grounds in answers to *where*-questions.

**Background:** You may be in pain or in a coma or in trance, and you may even go into or out of these things, but *He's in pain/trance/a coma* would surely be an odd answer to the question *Where's Jürgen?* **Note:** the PN Filter is a necessary condition of PNs, *not* a sufficient one. For example, below some criteria are suggested that delimit PNs as form classes in English and German – in both languages, the set of PNs thus defined is a proper subset of the set of Ns that satisfy the PN Filter (e.g. in both languages, names of rivers pass the PN Filter, but in neither language do river names have the formal properties of PNs). In contrast, in Yukatek, there is a fair chance that *only* PNs pass the PN Filter.

(a) **Morphological aspects:** Do PNs in L have morphological properties that allow them to be identified as a form class? And, are there affixes or morphological processes in L that occur only in/with PNs? **Background:** In Yukatek, many PNs incorporate the generic preposition *ti'* and/or the “feminine”/endearing prefix *x-* (cf. *Ti(-)cul*, *X-Hazil*, *Ti(-)x(-)cacal Guardia*), but neither property is found in all PNs. This case needs to be distinguished from one in which a semantically well defined subset of PNs in L share a defining morphological property. This may be true in English of compound mountain and lake names with *mount* (*Mnt.*; e.g. *Mnt. Snowdon*) and *lake* (e.g. *Lake Windermere*); cf. b1. **Note:** PNs will often have etymological relics such as the *-shire*, *-ham*, *-burg* etc. in English place names. This allows us to identify Ns that carry the particular feature as PNs (although they may also become secondary propNs), but it doesn't constitute a morphological property that would delimit PNs as a form class in English.

(b) **Syntactic aspects:** With regard to determination, attribution, and their distribution in GDAs: do PNs pattern with any subclass of Ns in L? Do they form a subclass themselves?

**More specifically:**

(b.1) **Determination:** What is the maximal projection of PNs in L? DPs, NPs, or other? Does this differ across subclasses of PNs? If so, what is the distribution?

**Background:** Determiners may comprise articles, demonstrative adjectives, some quantifiers, possessive pronouns, and possessive genitives. In English, cNs combine with determiners to form DPs, and there is some evidence that these are headed by the determiner (cf. *Gimme that (book)!*, *Have some (oatmeal)!*, *Did you see Peter's (t-shirt)?*). In this sense, the maximal projection of the determiner is a DP, but the maximal projection of the cN is an NP (cf. *\*Gimme book!*, *\*Have oatmeal!*, *\*Did you see t-shirt*). In contrast, propNs and pronouns do not combine with determiners, and since phrases headed by them have the same distribution as DPs, it may be said that their maximal projection is a DP, not an NP. E.g. in English and German, settlement names and the names of counties, most countries, islands, continents, and some celestial bodies behave like propNs and pronouns (cf. *I went to (\*the) London/ Devonshire/ Scotland/Rosset island/Australia/Mars*), unlike the names of mountains or bodies of water (cf. *I went along the Thames, I sailed the Irish Sea, I climbed the Ben Nevis*). This is *prima facie* evidence that only the names of continents, islands, towns, and political units are true PNs in these languages. However, compound lake names with *lake* and compound mountain names with *mount* also head DPs: *I climbed (?the) Mnt. Snowdon, I bathed in (??the) Lake Windermere*. In English, names of regions, valleys, and mountain ranges pattern with cNs (*the Yorkshire Dales, the Shenandoah Valley, the Adirondacks*). In German, this is generally true as well; however, if a region coincides with a (historical) political entity, its name patterns with propNs and pronouns. Also, not all country names are PNs; e.g. those that label some form of political organization are usually NPs (*the US, the USSR, the People's Republic of China*). Among celestial bodies, planets, asteroids, comets, and moons with proper names (*Titan, Io*) have PNs, but the earth's moon is called by a cN, and so are the sun, galaxies, and nebulae. Paradoxically, *world* is a cN, but *utopia*, naming the literal non-place, is perhaps a PN.

**Note:** The distinction by determination may not always be so crystal clear. Even in English, propNs and PNs may combine with demonstratives to express recognitional deixis (*That Fred is a real pain in the butt; That Mnt. Snowdon was a real challenge*). In German, the function of recognitional deixis is carried by the definite article; and in both colloquial German and Yukatek, distal deixis and definiteness are formally conflated. The redeeming factor in these languages is that determination is obligatory with cNs but not with propNs and PNs. But in Lao, there's no definite article and determination is not obligatory. This significantly blurs the distinction between cNs on the one side and propNs and PNs on the other, especially since indefinites and quantifiers will be anomalous with PNs and propNs for semantic reasons anyway.

**(b.2) Attribution and predication:** Do PNs take any attributes in L? Can they occur in the predication base or subject of non-locative predicates?

**Background:** *London* in *all/much of London, the magnificent London, or Dickens' London* doesn't seem to primarily denote a place, but rather a settlement located at a particular place, with any kind of feature attributable to a settlement (cultural, political, economical, and what not). The same seems to apply in *London is magnificent / has 6 million inhabitants / is 2000 years old*. This touches on the question as to what is the primary denotation of PNs in L - the landscape or occupation site as physical/cultural entities vs. as locations in space. Are there PNs in L that exclude non-locative attribution or predication? Are there PNs that require some form of modification before combining with non-locative attributes or predicates (e.g. *London -> The City of London*)?

**(b.3) Distribution in GDAs:** Is there any difference in the range of topological or path relators (case markers, adpositions, relational Ns) that combine with PNs as opposed to other Ns in L?

**Background:** Because PNs denote places and thus in a sense have something of a built-in locative function, they may play a privileged role in the formation of GDAs. In Yukatek, PNs are the only Ns that head GDAs without ANY spatial relators (adpositions or relational Ns) intervening. In German, PNs take a special goal preposition *nach* distinct from that which goes with people and objects (*zu* with 1D grounds; *auf* with 2D grounds; *in* with 3D grounds that are entered). In both languages, these are good candidates for criteria that set up PNs as a form class. **Note:** in Yukatek, there are certain 'generic' grounds which also occur without spatial relators, e.g. *lùum* 'ground, earth, soil, dirt'. In German, names of landmark sites like mountains and rivers don't take *nach* - their names would thus not count as PNs, which confirms the evidence from determination (cf. b1) suggesting only settlement and country names are PNs.

Another relevant issue here is the range of topological relations that take PNs as arguments in L, in particular as opposed to cNs referring to settlements, mountains, islands, etc. In English, as far as PNs denoting places of settlements or countries go, you can *be in PN, be near PN, be out of PN, go to/from PN, enter/exit PN or go into/out of PN*, respectively. But you can't *be on PN* (except with PNs denoting mountains or bodies of water), *be at PN*, or *be inside PN*,<sup>4</sup> and you can't *go onto PN or off PN* (you can be *above PN*, though; but then arguably you're again construing the referent of PN as a physical entity rather than as a place). So far, places as denoted by PNs seem to be construed as 2D regions with centers, peripheries, and some form of boundary. However, with mountains and islands, it's *be on* rather than *be in*. It's not completely clear why; islands seem to have sharp "boundaries" as well, but perhaps these don't count as boundaries in the relevant sense. It seems that boundaries of settlements and countries are construed as enclosures for some reason. **Note** also that mountains and islands go with *on* irrespective

<sup>4</sup> *Inside* may actually be possible in contrastive use (Jane Simpson, p.c.). Consider *There's no property available inside London, but outside there's still a lot to buy*.

of whether they're referred to by PNs or cNs (*on O'ahu* vs. *on the fourth island of the Hawai'ian chain*), and similarly settlements and countries go with *in* regardless. **Note** that institutions take *at* rather than *in* in English. The rationale is probably that if you're at school that doesn't mean you're in a particular place, and if you're going to school that doesn't mean you're on your way to a specific goal location (never mind the fact the you'll of course have a specific school at a particular place in mind). In German, institution names don't take *nach* - they pattern with objects names, not with PNs. But of course there may well be languages that treat institution names as PNs. **Note:** In Yukatek, of course, PNs don't take any topological relators. But there's still leverage for applying the Clark/Levinson AT/ON/IN typology: you can't be in front of or at the side of PN, but you can be near PN (expressed with a stative predicate 'be near'). So even in Yukatek one can make the point that places as named by PNs have polar topologies. Also, there's a relational noun *iknal* in Yukatek which takes a person or object as complement and denotes its location (kind of like French *chez* with people). That this doesn't happen with PNs indicates that PNs denote places by themselves, as they should. **Note:** PNs perhaps only combine with projective topological relations in two cases: with projective path relations (*before/after London*) and with projective relations in absolute Frames of Reference (*north of London*).

**(c) Sociolinguistic aspects:** What's the distribution of indigenous and non-indigenous PNs in sociolinguistic terms? How are recently founded settlements named? Do non-indigenous PNs have the same formal and semantic properties as indigenous PNs? Does it occur exceptionally/occasionally/frequently that the same place has different names in different languages? In case it does happen, do different PNs referring to the same place have the exact same reference? Are indigenous PNs borrowed into contact languages? In case this does happen, do the borrowed PNs always have the exact same reference as in L?

**Background:** Wherever there is ethnic contact involving different languages, naming places in one language or another is likely to be a big political issue. And whenever a linguistic supercategory is broken down into distinct subcategories, such as Ns into cNs, PNs, and propNs, checking what happens with borrowed items in this system may help understanding the underlying rationale of the system. In Yukatek, most PNs are Mayan, only a few colonial towns have Spanish names (*Merida, Valladolid*, etc.). The usual saint names found everywhere else in Mexico (*San Antonio, San Francisco*) are rare. Newly founded settlements are usually named after some revolutionary hero (*Emiliano Zapata*) or the like. However, the area that includes J.B.'s field site was by-and-large uninhabited by 1850 and yet most PNs are Mayan. This suggests that places have traditional Mayan names regardless of whether or not there's a settlement at the place at a particular point in time. But evidence for this hypothesis is not conclusive. Spanish PNs have the same formal and semantic properties as indigenous PNs in Yukatek.

Language contact often results in places either having different names in different languages or in PNs being borrowed from one language into another (where they then may be assimilated, yielding the aforementioned situation). *Nijmegen* is *Nimwegen* in German, and *Kleve* just across the border is *Kleef* in Dutch. When indigenous PNs are borrowed by dominant non-indigenous cultures, their denotation may well change (see Hercus & Simpson (to appear) for examples from Australia). For example, what is the name of a topographic landmark in the indigenous language may become the name of a nearby settlement in the borrowing language.

**(d) Etymological aspects:** What's the origin of PNs in L? Do they show an internal structure that reveals a naming strategy? How transparent are they?

**Background:** The internal structure of PNs is often reminiscent of the morphosyntactic process by which the name is derived, and hence reveals the PNs origin. Consider *Lake Ontario* vs. *Seneca Lake*. In both cases, the noun *lake* indicates that the PN refers to a lake or to the place defined by it, while the propNs *Seneca* and *Ontario* provide the distinguishing features of the PNs. In *Seneca Lake*, the distinguishing feature is construed as a determiner (as in *Fred's lake* or *that lake*), while *Lake Ontario* has an appositive structure (as in *King George* or *HMS Titanic*). It is important to understand that the process by which such syntactic structures become “fossilized” in PNs is not one of grammaticalization, but one of ‘univerbation’ – you use a syntactic construction of whatever internal structure and convert it into a noun, by using it as a name. However, languages not only differ in their semantic strategies for naming places, they also differ in the range of constructions permissible in PNs. Languages that name places after historical or mythical events may permit the conversion of entire sentences, complete with tense-mood-aspect morphology, into PNs (Jane Simpson p.c.).

In Yukatek, indigenous PNs are overwhelmingly if not exclusively descriptive in origin, but they're not transparent, so it's usually hard to say what exactly a PN described when it still described something. (**Note:** How can we know that PNs are descriptive if we don't understand what they mean? Well, we can for example tell that a PN originated in a particular syntactic structure even if we can't tell anymore what exactly the lexical roots involved were, for example because certain phonological features that distinguish lexical roots in L have been obliterated in PN. And we can tell that PN wouldn't have had that internal structure if it started out non-descriptively, like *San Francisco* or *Emiliano Zapata*.) It looks as though places in Yukatek are most commonly named after some animal or plant or some landmark feature like a natural well. Other strategies of naming places include reference to people or historical or mythic events, as is common throughout Australia. This is also the place to note patterns like *-burg*, *-borough*, *-shire*, *-bridge*, etc.

(e) **Lexical aspects:** What kinds of things have PNs in L?

The relevant options include settlements; regions; larger political units like counties, earldoms, chiefdoms, countries; places of human occupation like agricultural fields, hunting or fishing grounds, places of geological exploration; natural sites that provide shelter, like caves or natural harbours, or that provide e.g. sweet water for travelers, like oases and natural wells, or passage ways, like mountain paths, natural bridges, or fords, etc.; and topographic sites and areas like hills, mountains, forests, deserts, savannas, steppes, islands, rivers, lakes, reefs, etc.; larger geological and cosmological units, like continents, oceans, the sky, atmospheric layers, planets and other celestial bodies, and indeed the whole universe or *the world*; institutions like school, court, prison, church, hospital, etc.; places marked by historic events, like battlefields; places marked by mythic events; and also transcendental places, like heaven, the underworld, the spirit world, and the layers of Dante's *inferno*. Within any of these subdomains, how dense is the coverage with PNs (e.g. do *all* settlements have PNs? What about all hills/lakes/...)? For each of the thus semantically defined subsets of PNs, do they have distinctive formal properties in the sense of (a)-(b)? In particular, are there differences across subclasses of PNs in semantic construal, as discussed in b.3? It appears that in Yukatek, usually only settlements have PNs, but this is not entirely clear. It may be, for example, that uninhabited places of importance as reference points for travelers had PNs in the past.

(f) **Referential/denotational aspects:** Does the language distinguish between an object or person and the place it occupies, in particular in the construction of GDAs? (Yukatek does, English and German don't; cf. remarks on *iknal* under b.3). If L makes the distinction, then do PNs primarily refer to places or rather the settlements or landscape sites that occupy these places? And how is the place denoted by the PN defined in relation to the physical settlement, topographic site, etc.

that occupies this place? Are they exactly coextensive? Do people have clear intuitions about this? **Note:** To answer these questions, one will have to look for any formal difference that might distinguish the use of a PN in reference to the settlement or site from the use in reference to the location.

**Background:** It seems that languages distinguish a maximum of three types of referential grounds in space: places, people/animals/objects, and grounds provided by some frame-of-reference (as in *He went left/north*). The third category is probably not usually distinguished from the first, but we're of course specifically after places as named by PNs, and the question we're interested in is how these interact with the other two categories. – Questions regarding the referential extension of PNs are interesting but tricky, for a number of reasons. Places may have fuzzy boundaries. For example, when en route from the local village to the neighbouring village, up to where are you still in the local village and when do you enter the neighbouring village? There may be different answers to this question depending on what the background is assumed to be; for administrative or law purposes, boundaries may be much more sharply defined than for the purpose of giving a route description to a traveler. But a different source of ambiguity may arise when places are named after natural landmarks. E.g. are you still on Rossel Island if you're in a boat just off the shore? The landmark and place may diverge significantly. Jane Simpson (p.c.) mentions the example of a mountain named after a nearby water hole which isn't actually on the mountain. And, of course, the landmark itself may have fuzzy boundaries. For example, only geologists may be able to exactly define the boundaries of a mountain range.

#### *References*

Hercus, L. & Simpson, J. (to appear). Indigenous placenames: an introduction. To appear in: Hercus, L., F. Hodges & J. Simpson (eds.), *The land is a map: placenames of indigenous origin in Australia*. Canberra: Pacific Linguistics.