E/CONF.94/CRP.99

1 July 2002

English only

Eighth United Nations Conference on the Standardization of Geographical Names
Berlin, 27 August-5 September 2002
Item 12 (e) of the provisional agenda*
Toponymic data files: compatibility and structure of systems

Towards a standard geographic feature set: elevated relief features

Submitted by Australia**

* E/CONF.94/1.
** Prepared by David Blair, Asia-Pacific Institute for Toponymy.

02-46443 (E)

TOWARDS A STANDARD GEOGRAPHIC FEATURE SET:

ELEVATED RELIEF FEATURES

TECHNICAL REPORT

No. 1

2002

ASIA-PACIFIC INSTITUTE FOR TOPONYMY

INTRODUCTION

The Task

A National Placenames Data Model is being developed by the Committee for Geographical Names in Australasia (CGNA) as part of the Australian Spatial Data Infrastructure. A standardised set of designations for all types of geographic features is required for the Data Model, since such an information management system for national use cannot function without a controlled vocabulary or agreed thesaurus.

A rudimentary overview of the topographic features of Australia suggests that the vast majority fall within seven categories:

- hypsographic features
- vegetation features
- coastal features
- hydrographic features
- bathymetric features
- ice features
- construction/habitation features

It is likely that further investigation will reveal that the seven feature sets have fuzzy boundaries, and that some features will be found to fall outside these categories altogether.

The specific task commissioned by CGNA as the subject of this report is the production of a draft thesaurus of designations for one (or part of one) of these categories. The chosen subset will serve as a sample of the whole taxonomy; it will test the procedures, which will establish the taxonomy and which will produce the definitions of the standardised designations.

The Feature Subset

The sample section of the taxonomy chosen for this report is that covering features elevated about the earth's surface; that is, raised hypsographic features. Hydrographic features, whether or not they entail elements of relief, are excluded from the sample.

The section includes features covered by fifteen of the standard codes in the *Australian National Gazetteer*, one tenth of the whole.

Terminology

The process of creating an effective taxonomy for raised hypsographic features uses (or produces) several layers of terminology.

- *Generic Descriptors:* Definitions of GENERIC DESCRIPTORS used in compiling this report have been collected from various glossaries, in particular the National Gazetteer and geographical dictionaries. (See *Data Sources* below.) No modification of any descriptor terms has been attempted.

- *Semantic Components:* Each generic descriptor can be seen to be made up of, or "contain", a number of SEMANTIC COMPONENTS. A "crag", for instance, has the semantic components [-horizontal] and [+projected], among others. The semantic components are themselves defined within the system, and are the subject of discussion below. (See Table 1.)

- *Feature Codes:* A particular set of semantic components defines a FEATURE CODE, or designator. Seventeen feature codes, at the chosen level of generalisation, are sufficient to cover the subset of raised hypsographic features. Each code is a theoretical construct existing within the semantic classification system and takes an arbitrary form; but for ease of use each code of 2 to 4 alpha characters has been devised to resemble generic descriptors. (For example, ROCK includes the descriptor "crag", but is not itself equivalent to a physical rock – rather, to a set of semantic components.) The feature codes are listed and defined in Table 2, below.

THE METHOD

Data Sources

The base data for the task are the feature types currently in use by each of the State and Territory nomenclature authorities (including New Zealand) and in the Australian National Gazetteer. The definitions provided by some of those authorities and the draft ICSM glossary of generic terms supply additional data. (It should be noted that the ICSM glossary is not concerned directly with feature designations, but with those generic elements bestowed at the time of naming which may indicate the type of feature.)

The data sources used include:

1. *The Gazetteer of Australia.* Version 2, October 1999
2. *The ICSM Glossary of Generic Terms and Glossary of Designators.* Version 1.0, 1996; from www.anzlic.org.au/iscs/cgna/genglos961.htm
3. State lists, as supplied, 2001
4. *New Zealand Topographic Information and Glossary.* Printed May 2001; from www.linz.govt.nz/services/topomapping/digitaltopodata/260/namefeats.htm

Producing the Feature Codes

The starting point for the process which establishes a standard set of designators is to have regard to the existing terms as they are used by the relevant State and national authorities. There is obvious value in retaining as much of current practice as possible; and in minimising any future disruption as a national standard is applied.

The technique which is then used to establish the final set of consistent feature codes is that of *binary semantic classification*. The subset being investigated, for instance, is defined as all those hypsographic features (that is, relief features) which have the semantic component [+elevated]. The use of semantic components allows the application of reasonably objective verification procedures, and potentially permits a consistent approach to the allocation of any new generic terms and feature codes to the system.

Fourteen semantic components are sufficient to distinguish from each other all seventeen standard designations (feature codes) for elevated relief features. The fourteen components are listed and defined in Table 1, below, and discussed in the following section.

THE SEMANTIC COMPONENTS

The semantic components outlined in this report have not been chosen from an a priori list. They are intuitively produced as part of the step-by-step process of distinguishing the generic descriptors from each other. They are therefore arbitrary and subjective to some degree.

The nodes (that is, the points of distinguishment) in the taxonomic tree are binary in nature, but in theory a taxonomy tree which contains nodes with three or more splits is not prohibited. One advantage of the binary splitting process is its intuitive force, since the ability to contrast sets of terms by the presence or absence of some feature is part of our linguistic competence. Another advantage is more pragmatic: binary features provide labels for each node or point of distinguishment. The "audit trail" of the process thus produced would be much more difficult in a non-binary method.

It has not yet been determined whether these fourteen components are restricted to this set; if not, their current definitions will need to be revised and made more general. Three components – [± extended], [± projected] and [± raised] – each appear more than once in this subset, and have been successfully defined so as to cover each instance. So there seems no a priori reason why components should not apply in other parts of the wider taxonomy; indeed, the principles of simplicity and efficiency would seem to recommend it.

The component [+ elevated] is the primary component of this set of *raised* features, and together with the assumed component [+ hypsographic] it forms the defining pair of components for this subset.

The major division within this subset of raised relief features is indicated by the component [± independent], which separates major hypsographic features such as hills, mountains and the like from other features which are part of them or are naturally dependent on them for their existence.

Most other semantic components in the set have to do with extent in either the vertical or the horizontal plane. That is, they are morphological in nature and relate to physical dimensions. (The components [± tall] and [± broad] are obvious examples.) The one obvious exception is [stable] which, while still morphological in a sense, relies on the temporal rather than the spatial dimension. This semantic component was introduced in order to distinguish dunes from other geographic features in the HILL set.

TABLE 1: Semantic Components Set

A feature with this component	*exhibits this characteristic*
APICAL	is recognised as the uppermost part of a larger relief feature
BROAD	is perceived as having significant width
DEEP	is characterised more by depth than by breadth
ELEVATED	rises above its surrounds, and is therefore a raised relief feature
EXTENDED	is perceived as having significant length or extent
HORIZONTAL	is perceived as having no significant gradient, and is characterised by absence of vertical aspect
INDEPENDENT	is perceived as hypsographic in its own right, not as part of a larger feature
ISOLATED	rises conspicuously from level surrounds
OPEN	is bounded on only one side by a larger relief feature
PROJECTED	extends further from a larger relief feature in either the lateral or vertical dimension
RAISED	has a further elevation within the context of a larger relief feature
SHEER	is characterised by extreme gradient
STABLE	is not subject to obvious short-term deformation and relocation
TALL	has major elevation, arbitrarily set at >300m

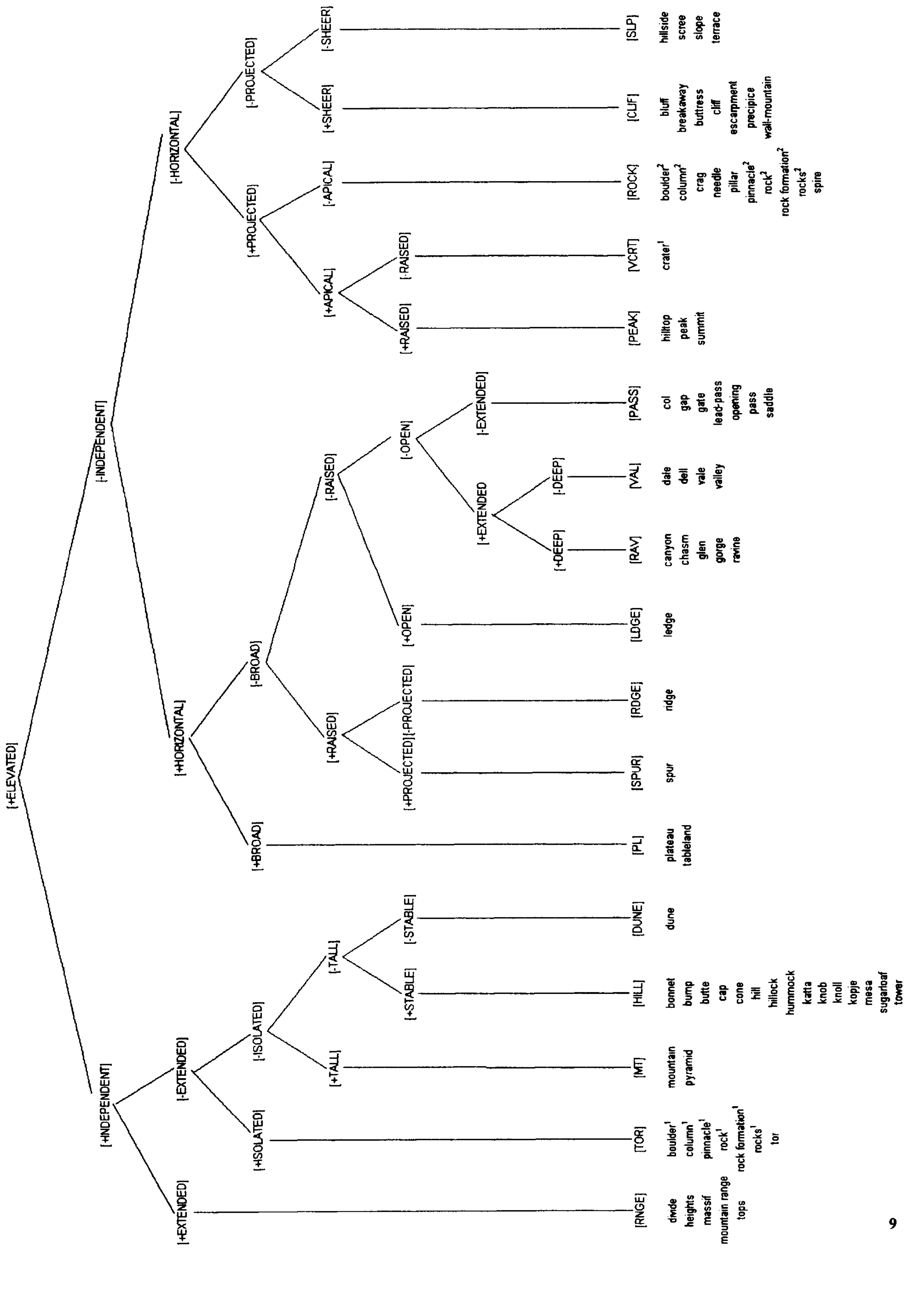
TABLE 2: Feature Code Definitions

Elevated Hypsographic (Relief) Features

CLIF	A perpendicular or steep face of rock.
DUNE	A mound or ridge of drifted sand.
HILL	A major elevation of the earth's surface, which is arbitrarily marked as less than 300m, rising conspicuously from the surrounding level but not normally an isolated feature.
LDGE	A narrow horizontal shelf-like surface on a mountain or hill.
MT	A major elevation of the earth's surface, which is arbitrarily marked as greater than 300m, rising conspicuously from the surrounding level but not normally an isolated feature.
PASS	A relatively horizontal opening between hills or mountains or within a range.
PEAK	The uppermost prominent point of a height feature.
PL	An extensive area of relatively flat land in an area of high relief.
RAV	A steep sided narrow valley.
RDGE	A long and narrow stretch of elevated ground on a mountain or hill or within a range.
RNGE	An extended line of mountains or hills forming a connected system.
ROCK	A prominent outcrop of stone on another height feature.
SLP	A gradient on a height feature.
SPUR	A narrow linear projection from a mountain or hill, normally less than 2km long and decreasing in elevation.
TOR	A prominent and isolated rock.
VAL	A relatively low region bounded by hills or mountains.
VCRT	A circular depression formed at or near the peak of a volcanic structure.

FEATURE	SEMANTIC COMPONENTS							INCLUDED TERMS
CLF	[-independent]	[-horizontal]	[-projected]	[+sheer]				bluff, breakaway, buttress, cliff, escarpment, precipice, wall-mountain
DUNE	[+independent]	[-extended]	[-isolated]	[-tall]	[-stable]			dune, sandblow
HILL	[+independent]	[-extended]	[-isolated]	[-tall]	[+stable]			battery, bonnet, brae, bump, butte, cap[2], cone, dome[2], head[2], hill, hillock, hummock, katta, knob, knoll, kopje, mesa, moraine-hill, nut, pimple, saddle, sugarloaf, throne, tower[2]
LDGE	[-independent]	[+horizontal]	[-broad]	[-raised]	[+open]			ledge
MT	[+independent]	[-extended]	[-isolated]	[+tall]				ben, bluff, cap[1], castle-mountain, dome1, head[1], mountain, pyramid, quoin, tower[1]
PASS	[-independent]	[+horizontal]	[-broad]	[-raised]	[-open]	[-extended]		col, gap, gate, lead-pass, opening, pass, saddle
PEAK	[-independent]	[-horizontal]	[+projected]	[+apical]	[+raised]			crest, hilltop, horn peak, peak, point[1], summit
PL	[-independent]	[+horizontal]	[+broad]					plateau, tableland, heights[2]
RAV	[-independent]	[+horizontal]	[-broad]	[-raised]	[-open]	[+extended]	[+deep]	canyon, chasm, glen, gorge, ravine
RDGE	[-independent]	[+horizontal]	[-broad]	[+raised]	[-projected]			lead-ridge, razorback, ridge
RNGE	[+independent]	[+extended]						divide, foothills, heights[1], hills, massif, mountain range, mountains, tops, tier, watershed
ROCK	[-independent]	[-horizontal]	[+projected]	[-apical]				boulder[2], column[2], crag, needle, pillar, pinnacle[2], rock[2], rock formation[2], rocks[2], spire
SLP	[-independent]	[-horizontal]	[-projected]	[-sheer]				bench, hillside, scree, shoulder, slope, terrace
SPUR	[-independent]	[+horizontal]	[-broad]	[+raised]	[+projected]			Spur
TOR	[+independent]	[-extended]	[+isolated]					boulder[1], column[1], crag, needle, pinnacle[1], rock[1], rock formation[1], rocks[1], tor
VAL	[-independent]	[+horizontal]	[-broad]	[-raised]	[-open]	[+extended]	[-deep]	amphitheatre, basin, cirque, dale, dell, nook-valley, vale, valley
VCRT	[-independent]	[-horizontal]	[+projected]	[+apical]	[-raised]			crater[1]

HYPSOGRAPHIC FEATURES: Elevated



DISCUSSION

Some general issues of feature code allocation were addressed as part of the process of resolving the raised hypsographic feature subset. Firstly, the question of apparent conflict between competing semantic components – an issue which arose within the sample but which will also be present in the wider taxonomy. For instance, by what principle do we decide to treat "valley" as hypsographic rather than hydrographic?

The decision in this case was to use current geomorphology as the determining principle, rather than the originating process. In other words, a valley may well have a hydrographic origin, but it is currently seen primarily as a relief feature. The decision was aided in this instance by the fact that several hydrographic features (such as rivers, creeks, etc) are *dependent* on the valley feature. Whether the use of this principle will hold up in the wider taxonomy is yet to be tested. What is clear is that any principle applied in a subset must then be applied over the whole: "valley" cannot appear in both hypsographic and hydrographic subsets.

A second general issue is that of degree of generalisation within the feature codes. An increase or decrease in the number of semantic components applied will decrease or increase, respectively, the level of generalisation in the system. Fewer nodes in the taxonomy tree will produce a smaller set of feature codes, each of which will include a larger number of generic descriptor terms.

The principle applied in the case of this subset was a pragmatic one: the order of generality should be not markedly different from that which currently holds within the geographical registers of the CGNA members.

At this point, a significant national advantage in the overall design of the taxonomy becomes clear. Individual jurisdictions represented on the CGNA can apply the taxonomy in ways that will be appropriate to their several needs. Some jurisdictions may wish to separate some generic descriptors into additional feature codes, by increasing the number of nodes (introducing additional semantic components). Other jurisdictions may not require the current number of feature codes, and may wish to collapse some by reducing the number of nodes and the items in the semantic component set. So long as the basic structure of the taxonomy is retained, all local variations would still be able to be mapped onto the national standard.

It should be noted here that New Zealand exhibits the widest departure in its coding practice from the general approach. For the most part in this taxonomy, the generic terms from the New Zealand list appear alongside those of other jurisdictions within feature codes. Some terms, however, are as yet unallocated: *cone, crest, dome, foothills, foot, nunatak, nunataks, ramps.*

As well as these general issues which raised matters of theory and principle, other minor questions arose during the process. Two that may be noted are: the choice of terminology for the semantic components, and the problem of homonymy.

The choice of terms for the semantic components is often pragmatic rather than theory-driven. In some cases, no great import seems to belong to the choice. The semantic component [± sheer] separates CLIF from SLP; but [± steep] might have done just as well. In other cases, the choice is more clearly one of finding a truly distinguishing component rather than mere terminology. The separation of DUNE from HILL, for instance, relies on [± stable] as the key semantic component. Here the search for the distinguishing component was the difficult task, not the choice of label for that component.

Homonymy within generic descriptors is not so much a problem for the taxonomy as a perceptual trap for users of the terms. The display format of the taxonomy distinguishes the homonyms by use of superscript numerals. A crater, for instance, may be either an elevated hypsographic feature (a depression formed at or near the peak of a volcanic structure), or a depressed relief feature (as a depression, say, formed on a plain by meteor strike). The former appears as *crater*[1] in this taxonomy under the feature code VCRT; the latter will appear in a different subset, that of non-elevated hypsographic features. A rock may be a formation which is part of a hill or mountain, or it may be an isolated major feature such as Uluru. The former appears as *rock*[1] under [-independent] ROCK in this taxonomy, the latter as *rock*[2] under [+ independent] TOR.

CONCLUSIONS

The 76 generic descriptors (including twelve homonyms) which the relevant jurisdictions employ to refer to raised relief features in the geographical environment have been semantically analysed into fourteen components. These semantic components have been shown to be useful in separating the descriptors into seventeen feature-code sets.

The labels (or feature codes) given to these sets provide a standard set of designations for the subset of hypsographic features which was the subject of this exercise.

An extension of this exercise to the full set of generic descriptors used by relevant Australasian jurisdictions is recommended.