The Polynesian Languages Forum Terminology Database

Richard Benton

Te Wâhanga Kaupapa Mâori,
New Zealand Council for Educational Research

1. Background

"Polynesian" languages are a grouping within the Malayo-Polynesian section of the Austronesian language family, spoken by the indigenous people in the countries within the "Polynesian Triangle" (that section of the Pacific bounded by Hawaii in the north, Easter Island to the east, and New Zealand to the South, and on a few "outliers" in Melanesia and Micronesia to the west of the triangle). The ancestors of the Polynesians reached their present homelands (probably from southeast Asia via the western Pacific) over a period of about a thousand years from their colonization of Tonga in the early centuries of the first millenium A.D. Unlike Melanesian languages, to which they are related, many Polynesian languages are spoken over extensive geographical areas with only minor dialect differentiation between contiguous regions or islands.

Politically, there are three independent ethnolinguistically Polynesian states (Tonga, Tuvalu and Western Samoa), two others in Free Association with New Zealand (the Cook Islands and Niue), and five which are (in varying degrees) self-governing dependencies of a metropolitan power (American Samoa, French Polynesia, Wallis and Futuna, Easter Island and Tokelau). Finally, both New Zealand and the U.S. state of Hawaii were originally settled by Polynesians, and the Maori and Hawaiian languages respectively have official status in these politics. Furthermore, Polynesian languages are increasingly being used in education and the conduct of public business in most of these states and territories, which has resulted in a resumption of the modernization process, including the creation and adoption of vocabulary for new scientific, technical and social concepts. (This process had been disrupted in most of the languages with their exclusion from formal education and administration by the end of the nineteenth century.)

In some of the Polynesian states and territories modernization has proceeded with schools, newspapers and government departments coining new expressions as the need arises on an ad hoc basis (often simply by adapting English expressions to Polynesian phonology). In New Zealand and Tahiti, however, official language guardians have been appointed by the government, and these have given high priority to the orderly development of vocabulary to meet educational, technical and scientific needs.

The New Zealand body, the Maori Language Commission, called a meeting of representatives of government and other agencies involved in language planning in Polynesian states to a meeting in February, 1991, to set up a Polynesian Languages Forum. This body was to enable the various agencies to assist each other in their efforts to develop, maintain or revitalize their respective languages. One of the first steps taken by the new Forum was the decision to set up a database of new vocabulary which would also contain information about Proto-Polynesian word roots. The database would thus be a record of actual usage (including coinages by official bodies like L'Académie Tahitienne and the Maori Language Commission), and also

1994. Bonifacio Sibayan & Leonard E. Newell, eds. Papers from the First Asia International Lexicography Conference, Manila, Philippines, 1992. LSP Special Monograph Issue, 35. Manila: Linguistic Society of the Philippines. pp. 181-191

a resource through which the members of the Forum could develop a common approach to vocabulary creation in the future. Through the systematic use of Proto-Polynesian elements, it was hoped to be able to parallel the adoption of Latin and Ancient Greek elements in the creation of new terminology in many Indo-European languages, enabling the modernization of Polynesian languages to become a convergent process within a shared tradition.

2. The Maori language database.

The prototype for the Polynesian database was the Maori language database which had been set up some years earlier by Te Wâhanga Kaupapa Mâori (the "Maori Interests Section") of the New Zealand Council for Educational Research. This database draws from all available material using or listing new vocabulary in Maori, and can be accessed interactively through a computer-based information system, the Te Wahapû electronic bulletin board (established by Te Wâhanga Kaupapa Mâori with the assistance of a donation of equipment by IBM [New Zealand] Ltd). The Forum also decided to adopt this method of access to the database, and designated Te Wahapû as its official communications system. (Te Wahapû is itself an exercise in applied terminology, as the default language for its menus, prompts, commands and system messages is Maori.)

The Maori database was (and remains) basically a bilingual glossary with an indication of the source of the item, and its status with the Maori Language Commission. A section from the July 1992 update of the database (Figure 1) illustrates its structure.

maikuku	23	nails
pakiwhara	52	naked
tu˜ tahanga	52	naked
whakaingoa	54	naming
koroki˜	11	*narrative
wha˜iti ???	32	?narrow
wheua ihu	60	*nasal bone
awa hau	60	*nasal passage
awa ihu	55	*nasal passage
pu˜ ihu	60	*nasopharynx
waiata a te motu	33	*national anthem
hau ma˜ori	35	*natural gas
kaitohutohu whakatupu	62	natural leader
tau ki te iwa	32	?natural number
tau tatau	68	*natural numbers (maths)
hau-a˜-papa	35	*natural resource
rawa whenua	68	*natural resource
nuku	42	*nature
tauamoana	60	*navy
wha˜to˜to˜	33	*ne-waze (judo)

Figure 1: Section of Maori database file (August 1992)

The database can be browsed either in its "alphabetical" form (currently divided into six chunks, together containing about 11,000 separate entries), or through searching for specific words. In the latter case, specifying a word like 'anorexia' would probably bring a single-item result; a request for 'computer', however, would bring several screens of matching items (see Figure 2).

1994. Bonifacio Sibayan & Leonard E. Newell, eds. Papers from the First Asia International Lexicography Conference, Manila, Philippines, 1992. LSP Special Monograph Issue, 35. Manila: Linguistic Society of the Philippines. pp. 181-191

hoki whakamuri	60	*backspace (computer)
taikaha	60	*bold (font: computer)
whakakore	Z	cancel (a computer process)
roro	10	*central computer unit
pu˜	60	*character (computer)
kaitapa	60	*chooser (computer)
tohu pu˜manawa	60	*code (computer)
rorohiko	32	*computer
mo˜hiotanga rorohiko	56	computer awareness
papa pa˜to˜to˜	10	*computer keyboard
papa rorohiko	35	*computer keyboard
kaiwhakamahi rorohiko	16	computer operator
whare rorohiko	35	*computer room
pouaka whakaata	10	*computer screen (v.d.u.)
akoranga rorohiko	11	*computer studies
ara rorohiko	54	computer system
te˜pu pa˜to˜to˜	10	*computer table (typist's desk)
mati rorohiko	35	*computer terminal
ta˜tua	10	*connecting belt (computer)

Figure 2: First of 5 screens from a search for 'COMPUTER'

Work is now being done on enabling searches to be made for categories of word (a possibility which exists serendipitously in the case of many computer terms, but not at all for most other possibilities), so that, for example, inputting a left hand brace { (the symbol for word category) followed by "PLACE" would generate a list of place names (see Figure 3). (We will return to the matter of categorizing words presently, as it is one of the major problem areas for both the Maori and the Polynesian databases.)

File E:KHKM.BBS:

Pehiamu	65	*Belgium	{PLACE
Koroa˜tia	65	*Croatia	{PLACE
Tiorangi	65	*Iceland	{PLACE
Rakapuo˜	65	*Luxembourg	{PLACE
Potukara	65	*Portugal	{PLACE
Horowinia	65	*Slovenia	{PLACE
{WHENUA	65	*{PLACE	

Figure 3: Partial result of a search for the category 'PLACE'

3. Scope of the Polynesian database.

In September 1992 the Polynesian database included terminology obtained from a variety of sources in most of the major Polynesian languages:

- New Zealand Maori material collected by Te Wâhanga Kaupapa Mâori and new vocabulary developed or approved by the Maori Language Commission;
- Tahitian material supplied by L'Académie Tahitienne;
- Samoan, Tongan and Cook Island Maori material supplied by the Centre for Pacific Studies, University of Auckland and the School of Languages, Auckland Institute of Technology;

1994. Bonifacio Sibayan & Leonard E. Newell, eds. Papers from the First Asia International Lexicography Conference, Manila, Philippines, 1992. LSP Special Monograph Issue, 35. Manila: Linguistic Society of the Philippines. pp. 181-191

- Hawaiian material supplied by the Hawaiian Vocabulary Committee (a voluntary organization associated with the Punana Leo — pre-school "language-nest" — movement) and staff of the Hawaiian Language sections of the University of Hawaii (Hilo and Manoa);
- Material in various Polynesian languages collated by the Centre for Science and Mathematics Education Research, University of Waikato;
- Standard dictionaries of the various Polynesian languages.

As it is set up at present, the database is oriented primarily towards those twelve languages which are the main official Polynesian languages in their states or territories of origin. Data on other Polynesian languages, and also from related languages such as Fijian, are included in the notes field. In this way the database can be kept reasonably compact, while at the same time including as much relevant information as possible. In the Proto-Polynesian database, all known reflexes of a particular Proto-Polynesian reconstruction are included, together with higher-order (e.g. Proto-Austronesian) forms where this information is available. Fiji and Rotuma accepted an invitation to attend the second meeting of the Forum in Papeete in May 1992, and to participate in the Forum's work. These two languages are the closest relatives of Polynesian languages within the Austronesian family, although the greater complexity of their phonemic and phonotactic systems makes the recognition of cognate forms more difficult for the linguistically unsophisticated observer, compared with the much more transparent interrelationships among the Polynesian languages proper.

4. Structure of the Polynesian Database.

The Polynesian database, called "Te Le'o o Maui" (the language of Maui) after Maui, the Polynesian culture-hero and discoverer of new lands, is organized through the Archivist database program (a versatile, powerful, but comparatively inexpensive program, designed explicitly for handling text files with fields of varying length). It can be exported from Archivist in a form which can be read into other database or word-processing programs, including the custom-made on-line database program used on Te Wahapû. Each entry currently contains 17 fields:

(a) The headword (with definition) in (1) French, (2) English, and (3) Spanish (as the languages of wider communication in Polynesia).

(b) Entries for the equivalent terms in the Polynesian languages which are official languages in those Forum states and Territories represented directly or by proxy at the first meeting of the Forum: (4) Cook Island Maori, (5) Easter Island, (6) Futunan, (7) Hawaiian, (8) New Zealand Maori, (9) Niuean, (10) Samoan, (11) Tahitian, (12) Tokelauan, (13) Tongan, and (14) Uvean, with an indication of the source for the entry and its derivation.

(c) The Proto-Polynesian elements (15) contained in the words listed in fields (4) to (14) cross-referenced to the modern forms.

(d) The date of the last revision of this item (16), and (17) supplementary notes about the entry (including where relevant equivalents noted in other Polynesian or Austronesian languages, as well as information about Fijian and Rotuman.

An example is the current entry for "Avion/Aeroplane", reproduced on the next page.

1994. Bonifacio Sibayan & Leonard E. Newell, eds. Papers from the First Asia International Lexicography Conference, Manila, Philippines, 1992. LSP Special Monograph Issue, 35. Manila: Linguistic Society of the Philippines. pp. 181-191

Like all entries made so far, this one is incomplete — partly because data are not yet available for some languages (Easter Island, Tokelauan and Uvean), and partly because there has not been time to fill in some of the blanks (e.g. there is no definition in the entry for the Spanish headword, and not all the Proto-Polynesian elements are defined). There is little information about how the term was constructed in the various languages (mainly because of the Proto-Polynesian derivation of the component morphemes). For most entries, this kind of information is still lacking.

[Fr] _AVION (n) Appareil de navigation aerienne plus lourd que l'air, muni d'ailes et capable de se deplacer au moyen d'un moteur a reaction <Lexis>
-
[En] _AEROPLANE / _AIRPLANE (n) Mechanically driven heavier-than-air flying machine <COD>
-
[Es] _AEROPLANO / _AVION (n)
-
{AVIATION // {AVIACION
-
CIM > pa'irere <ZCM.13w>; pa'i-reva <Savage>[<pa'i =ship/navire; %lele %rewa]
-
EAS > xea
-
FUT > vakalele <Rensch> [%waka %lele]
-
HAW> mokulele <Pakui/Elbert> [<moku = ship/navire; %lele]
-
NZM > waka rererangi <Ryan> [waka; %lele %langi]
-
NIU > vaka-lele <McEwen> [%waka %lele]
-
SAM > va'alele <Milner>[%waka %lele]
-
TAH > manureva <ZTH.13K> [%manu %rewa]
-
TOK > xtk
-
TON > vakapuna <Churchward> [%waka; puna "fly/voler"]
-
UVE > xuv
-
PPN > %waka 'canoe, vessel' (FUT, NZM, NIU, SAM, TON); %lele 'fly' (FUT, HAW, NIU, SAM, CIM); %langi 'sky' (NZM); %manu 'chicken, bird' (TAH); %rewa (CIM, TAH).
-
Rev > {901124} {910319}
-
[Notes] cf. Rennellese bakagege <%waka + %lele (Elbert, Dictionary)

Figure 4: Sample Polynesian database entry — June 1992

Nonetheless, the entry for 'aeroplane' does illustrate both the form of the entries and the information which the database should contain. One of the hopes of the members of the Forum is that the database will not only provide information about how various Polynesian languages have dealt with new terminology, but it will also enable a Pan-Polynesian approach to the

1994. Bonifacio Sibayan & Leonard E. Newell, eds. Papers from the First Asia International Lexicography Conference, Manila, Philippines, 1992. LSP Special Monograph Issue, 35. Manila: Linguistic Society of the Philippines. pp. 181-191

creation of terminology to develop. It is for the latter reason that the Proto-Polynesian elements in the terms (marked with the "percent" symbol (%) in the example above) are to be explicitly identified. Information on other Polynesian languages, such as Tuvalu, Marquesan, Tuamotuan and Mangarevan, will probably be incorporated in the near future.

It is also intended to include material illustrative of the actual use of various terms in context in the database as the opportunity presents itself. This will involve the scanning of suitable texts and documents, which may be kept in separate files cross-referenced to the terminological lists.

5. Structure of the Proto-Polynesian database.

The Proto-Polynesian database consists of some three megabytes of data supplied to the project by Emeritus Professor Bruce Biggs of the University of Auckland. As a temporary measure, the separate entries in Prof. Biggs' database have been prepared for searching in two ways. The first of these (illustrated in Figure 5) is an "index" consisting of three fields: Proto-language (e.g. Proto-Polynesian, Proto-Tahitic, etc.), Protoform, and an English gloss. The second alternative includes the information in the index, plus all known reflexes of a particular Proto-Polynesian reconstruction (in a single field), and notes (see Figure 6).

```
@Le'o o Maui PPN Database (c) B.Biggs, TWKM & PLF, Kupu 34

{PROTO-NUCLEAR_POLYNESIAN

ALAALA
Fish sp.  (Carangidae)
```

Figure 5: Sample entry in Proto-Polynesian index

```
@Le'o o Maui PPN Database (c) B.Biggs, TWKM & PLF, Kupu 37

{PROTO-NUCLEAR_POLYNESIAN

ALAALA
Fish sp.  (Carangidae)

%NP%                :Fish sp (Carangidae).
EFU    Alaala/muli. :Small fish sp. (Bgs).
EUV    Alaala.      :Fish sp. (Carangoides fulvoguttatus)[Forskal] (Rch).
KAP    Araara.      :Sp. of fish (Ebt).
KAP1   Alaala.      :Jack (Caranx malampygus) (Lbr).
MAE    Raaraa/futu. :Fish sp. (Trachinotus?) (Clk).
MAO    Araara.      :Trevally (usacaranx lutescens.
MFA    Raara.       :Small stage of a fish, probably a Caranx sp..
NKR    Alaala.      :Fish sp., jack.
SIK    Alaala/hutu. :A fish.
TAH    Araara/vi.   :Fish sp., called hiroa when fullgrown.
TAK    Arara.       :Fish sp. (Hwd).
TIK    Araara/futu. :Seafish of Snapper type (Fth).
WFU    Ara/uru.     :A fish.

<PollexAM.dat 030492>
```

Figure 6: Entry in full Proto-Polynesian database

1994. Bonifacio Sibayan & Leonard E. Newell, eds. Papers from the First Asia International Lexicography Conference, Manila, Philippines, 1992. LSP Special Monograph Issue, 35. Manila: Linguistic Society of the Philippines. pp. 181-191

6. Accessing and interrogating the databases.

Apart from using the Archivist program itself with the indexed data files, the Polynesian and Proto-Polynesian databases can be searched interactively through the Te Wahapû facility. In the case of the Polynesian database, particularly, this means that the latest information is always available to all members of the Forum, provided they have access to a computer and a modem. Because the on-line search program we use at present matches strings within complete entries (rather than fields), certain items within a field are tagged to facilitate more specific searches: category labels are preceded by a left brace, headwords by an underline character, protoforms by a percent mark, and so on. A search for the string 'fish' will thus bring up any entry with that sequence of characters. In the Proto-Polynesian database this would result in a very heavy catch, including, for example, words like *afo (glossed 'fishing line'), as well as dozens of undifferentiated fish species. A search for cognates of a known form, for example 'alaala' will often narrow the search considerably, producing only one or two entries if the sequence of characters in the string is sufficiently uncommon.

As the Polynesian database develops, the need to be able to narrow searches easily will also become greater. At present, the search for 'Avion' would bring up a few other entries, like that for 'air mail', a search for the category 'Transport' a more substantial collection (including words associated with vehicles, such as 'accelerator'), while a search for 'mathematics', in any form, would result in several hundred responses, from 'absolute value' to 'hypotenuse' and beyond. Fortunately, the results of an on-line search do not have to be read on line: they are saved in temporary files which may be downloaded for perusal at the interrogator's leisure.

It is also possible to find material not yet in the database (or at least provoke a long range enquiry for it) by searching a message base within Te Wahapû and, if that does not bring enlightemnent, leaving a new message requesting the information. Recently, for example, a user of the Maori database asked for the equivalent for 'floptical disks', after searching for it in vain. This resulted in an exchange of notes with the petitioner, seeking information about the nature of floptical disks, and a request to the Maori Language Commission, together with suggestions as to how the aims of the Polynesian Languages Forum might be taken into account in the coining of terms for both 'optical' and 'floptical' in this context. (The answer to the original question is awaited.)

Furthermore, "gaps" in the database are routinely filled through "homework" files, left for members of the Forum to pick up and peruse. These list all new headwords added to the database since the last major update which do not include a subentry for the language concerned. The agency responsible for providing information on that language is asked to check for an equivalent form, to supply it if it exists, or to confirm its current absence from the lexicon. (Updated data files are also made available for downloading by Forum members as they are prepared.)

7. Problems, resolved and awaiting resolution.

It will be quite clear from the foregoing account that there are plenty of problems, theoretical and practical (to say nothing of the financial) yet to be addressed in the form and content of the database. Some of these can probably be resolved fairly easily, others are likely to prove more intractible. The difficulty is compounded (but the challenge made more interesting) by the fact that the database must be structured in such a way that it is accessible and useful to a very wide range of people, most of whom are not specialists in terminology,

1994. Bonifacio Sibayan & Leonard E. Newell, eds. Papers from the First Asia International Lexicography Conference, Manila, Philippines, 1992. LSP Special Monograph Issue, 35. Manila: Linguistic Society of the Philippines. pp. 181-191

linguistics, or computer science. Some of the major problems we have had to contend with relate to the definition of terms, categorization of words, orthography, and, of course, providing easy access to the information in the database for all potential users.

Definition of headwords

If confusion is to be minimized, it is important that headwords be precisely defined, rather than simply glossed in one of the metropolitan languages. In this respect, the approach adopted in the Tahitian Academy's "Petit Lexique" (which lists 1000 new words coined or approved by the Academy) provides a good working model. While many "ordinary" words are polysemantic, the same can also apply to highly scientific terms. An example of this is the term 'ablation', for which there is a Tahitian equivalent in the "Petit Lexique". The word 'ablation' (which I do not remember having encountered before I received my copy of the lexique from Papeete) is a purely scientific term, in Engish at least, but has at least two meanings in French, and a third in English. The Collins/Robert dictionary defines the French term (somewhat unhelpfully) as:

(Med) removal, ablation; (Geol) ablation.

Fortunately, the Concise Oxford Dictionary was more informative:

(Surg) Removal of any part of body; (Geol.) wasting or erosion of a glacier, iceberg or rock by melting or water action; (astronaut) evaporation or melting of part of outer surface of spacecraft through heating by friction with atmosphere.

Because of the clear definition of the Tahitian term in the material supplied by the Academy, there was no problem in determining which sense of 'ablation' it referred to. The entry in the database is thus categorized "Medicine". Equivalents for the other senses of 'ablation' may be submitted later; if so, there will be a separate headword for each sense, whether or not the Polynesian terms are as polysemantic as the English and French ones.

We also hope to be able, as far as possible, to maintain within the database a clear distinction between equivalents and approximations. Where a word has been submitted as an equivalent of a particular term, it has been entered in the database under the appropriate headword, often without further comment (apart from the noting of the source). Where it is a one-word approximation of the term, the wider meaning of the word concerned is noted in the entry. Where there is simply an explanation in the language concerned of what the headword means, but no equivalent or even approximate term is offered, the definition is enclosed in double quotation marks. (There are quite a number of examples of this already among the mathematical terms in the database.) We are very dependent on the quality of the information supplied to us in making decisions of this kind, and have asked Forum members to help us to establish a network of consultants for the particular languages represented in the database who can review the entries for their respective langauges and minimize the amount of misinformation they contain.

Categorization of entries

Probably the most difficult problem at the moment in structuring both the Maori and the Polynesian databases is deciding how best to categorize the various entries. Many users of the databases are just as interested in sets of terms as they are in individual entries. It is also

1994. Bonifacio Sibayan & Leonard E. Newell, eds. Papers from the First Asia International Lexicography Conference, Manila, Philippines, 1992. LSP Special Monograph Issue, 35. Manila: Linguistic Society of the Philippines. pp. 181-191

important to note that the database is designed for use by a wide range of people, writers and teachers at various levels, for example, as well as by language planners and terminology specialists. A way thus has to be found to ensure that the classification system is user-friendly, but allows for the development of a hierarchy of terms that would be both useful and economical. Attention to this matter now, while the Polynesian database at least is very small, will help avoid problems as it expands.

About half the 600-odd items currently included in the Polynesian database have been assigned to a category or domain, mostly of a very general kind, like "mathematics" (the most numberous), "health" or "law". This will not be a helpful approach for long, however, as when the mathematical data already on hand is entered, a search for "mathematics" will result in about 650 entries (or 200 or more pages of computer printout) being presented to the enquirer. These potentially unwieldy categories will then need to be cut up into smaller units which can be searched for separately.

This classification must, however, be devised in such a way that it respects the integrity of the subject matter, and at the same time which will be comprehensible and acceptable to the peple who need to use it. We will certainly have to produce a "road map" showing how the items are categorized, so that those who wish to investigate sets of words will be able to find what they are looking for. We will also need consultants with subject-matter expertise in areas like mathematics to ensure that our classifications of terms is not misleading from a scientific viewpoint.

Orthography

There are really only two major orthographic problem areas, and one minor one, affecting the database. These are (a) the marking of vowel length, (b) the indication of the phonemic glottal stop, and (c) the way the velar nasal is written.

For vowel length, the general operating principle has been that vowel length should be marked consistently in the database, following a single standard orthographic convention for the language concerned. In the standard orthographies for most Polynesian languages, the macron is the preferred method of indicating phonemic vowel length, and was explicitly accepted as standard practice for Polynesia by the Forum at its 1992 meeting. Unfortunately, vowel symbols with macrons above them are not standard ASCII characters, 7-bit or eight, and the circumflex accent, which is a passable substitute in emergencies, is invisible on many computer monitors (the macron, in packages which do include it, is invisible to even more displays). These difficulties have been overcome in the Polynesian and Maori databases by using a tilde following the vowel to indicate a macron. So, for example, in the database an **o** with a macron (**ō**) will be written **o˜**. The tilde is used only in this way in the database, so it is a simple matter when material is downloaded or exported to a file just to replace **o˜** by **oo**, **ô**, **ō**, or whatever the user prefers. In the Proto-Polynesian database, however, vowel length is marked by doubling the vowel symbol (i.e. the convention adopted by Professor Biggs has been retained in the data he supplied).

Although the tilde neatly sidesteps the political problem of choosing double vowels over macrons (writers of Polynesian languages take orthographic matters no less seriously than the French or English literati), it does create a new problem in regard to user-friendliness. No one (as far as we know) writes the tilde intuitively, and it can be very frustrating to search for a

1994. Bonifacio Sibayan & Leonard E. Newell, eds. Papers from the First Asia International Lexicography Conference, Manila, Philippines, 1992. LSP Special Monograph Issue, 35. Manila: Linguistic Society of the Philippines. pp. 181-191

word and fail to find it simply because one has forgotten to type in a symbol which is never used outside a particular virtual reality. For both the Maori and the Polynesian databases, we intend including an unmarked version of each key word containing one or more long vowels, so that these can be neglected with impunity by people seeking the entries containing the words concerned.

After some vacilation, we have decided to retain accents in French text (although many of the current entries have yet to be revised to reincorporate accents removed earlier when we had decided, on the technical grounds mentioned in regard to the macron, to delete them). Again, the characters concerned will be invisible on non-IBM equipment, and cannot be sent in text files through some major electronic networks. I am not sure whether any of the languages as yet unrepresented use the "double vowel" convention in their official orthographies, rather than the macron. If they do, then data from them will be entered in double-vowel style. The important thing to note is that we will try to be consistent within each language.

Phonemic glottal stop has been regularized to an opening quotation mark ‘, which seems to be the standard (as a substitute for the "hamza" character) in most countries concerned (although in fact the closing quotation mark ’ is frequently used instead, even in official publications from these territories). Again, this symbol can be changed (or deleted) when the information is downloaded, as the symbol ‘ is used for no other purpose in the database. We do, however, need qualified reviewers for the languages which feature this sound, as it is often unmarked by writers. The same is true with vowel length. The Cook Island material in the database at present, for example, needs to be checked thoroughly for unmarked glottal stops and long vowels.

As for the velar nasal, it is written **g** in some languages, and **ng** in the others. In the Proto-Polynesian database this phoneme is indicated by the symbol **9**. Some of the people who have consulted that database so far have been initially somewhat taken aback by this convention. However, until a decision (as much political as practical) is made on whether **ng** or **g** should appear as the Proto-Polynesian standard, **9** will reign supreme.

On-line database program.

The final challenge (for the moment) is to develop a user-friendly, secure, powerful and rapid on-line search facility, which would enable the various databases to be interrogated in a variety of ways, expanding and limiting searches, confining a search to one or more fields, and so on. The program would also have to be accessible through a modem, and be written in such a way that it could not be used by unauthorized persons to enter new data (or to create new databases), nor could a user exit into the computer operating system from the on-line version. We have some good models, including the program which is used to search the message bases on Te Wahapû at present, but have not yet been able to find or adapt a program which will meet all our needs.

8. Conclusion.

In the eyes of the founders of the Polynesian Languages Forum, the database project is not simply (or even primarily) a scientific or technical exercise. In his closing address to the inaugural meeeting of the Forum (at Kirikiriroa Marae, Hamilton, New Zealand in February, 1991), M. Maco Tevane, a Minister in the government of French Polynesia, and Director of L'Académie Tahitienne) stated that:

1994. Bonifacio Sibayan & Leonard E. Newell, eds. Papers from the First Asia International Lexicography Conference, Manila, Philippines, 1992. LSP Special Monograph Issue, 35. Manila: Linguistic Society of the Philippines. pp. 181-191

> In order to fit in and accompany the economic and social development of a country, region or territory, a culture must evolve by relying on its past and integrating scientific and technological developments. A language should be the primary support to such an evolution. The setting up of a database on technical terminology cannot but promote a permanent "renaissance" of the Polynesian languages.

He went on to say that:

> Polynesian culture must be a factor of identity which strengthens our feeling that we belong to a community extending beyond the narrow limits of our borders. This cultural development is relying more than ever on initiatives such as the project we are associated with today. More than ever, it will be necessary to encourage potentials, [and to] promote aspirations to live together in this part of the world which is bound to fulfil its role in the third millenium.

A great many hopes, therefore, rest on this enterprise, as Polynesian peoples prepare their languages to cope with overwhelming odds against their survival with the same spirit that their ancestors had when they voyaged into unknown seas. It is the use to which the database is put, rather than the database itself, which will determine whether or not it will be a force for the revitalization and continued viability of Polynesian languages generally. The most enthusiastic members of the Forum (currently the Tahitian, New Zealand Maori, Cook Islands and Hawaiian organizations and agencies) clearly see the modernization of their languages in harmony with their linguistic heritage — that is, modernization with an unmistakably Polynesian flavour — as essential to their cultural integrity.

The cooperation and mutual support fostered by the project may also help to develop and strengthen the infrastructure needed to support the long-term maintenance of these languages. In the meantime, the database itself is afloat, but there are plenty of leaks to attend to, and a sail, or even some extra oars, will need to be found before we can be confident that it will reach its objective.

1994. Bonifacio Sibayan & Leonard E. Newell, eds. Papers from the First Asia International Lexicography Conference, Manila, Philippines, 1992. LSP Special Monograph Issue, 35. Manila: Linguistic Society of the Philippines. pp. 181-191