

Phonotactics, Homonymy and Vocabulary Expansion in Marquesan

Viktor Krupa

One of the most interesting problems of Polynesian studies is probably the tendency, of unknown origin, to continuously simplify phonology, phonotactics and grammar. Perhaps the tendency has to do with the fact that small and closely related groups of migrants landed at islands notable for their relatively poor natural resources and material culture. Such conditions obviously do not demand an overly sophisticated system of verbal communication.

Keywords: Reduction of consonants, high homonymy and polysemy, massive Western impact, renaissance of Polynesian languages

It is especially the reduction of the phonological inventory of Polynesian languages that is attractive to those interested in the history of Polynesian languages. A satisfactory explanation of this phenomenon is still a matter of future studies. Peter Trudgill has tried to shed some light on this development (Trudgill 2004). However, Vladimir Pericliev, a Bulgarian scholar, rejects Trudgill's explanations (Pericliev 2004).

The Marquesan islands were obviously the first archipelago settled in Eastern Polynesia and soon became a new centre for further voyages to Mangareva and Easter Island in the southeast and to Hawai'i in the north. The settlers arrived from West Polynesia (most probably from Samoa; Hawaiiki, the name of the legendary homeland of East Polynesians, is probably derived from Savai'i, the largest Samoan island). The seafarers might have reached Marquesas in the third century A. D., which means that there was more than enough time for a natural dialectal fragmentation of their language. This process has been supported by another important circumstance, notably that the rugged, almost Alpine terrain typical of the Marquesas rendered extensive and frequent contacts among the individual communities scattered across the islands difficult. Intensive contacts between the settlements started later, and subsequent interaction could not annul all existing differences. The phonetics of modern Marquesan has been modified by a continuous pressure of the old historical tendency to reduce the phonological system. This tendency has resulted not only in a decrease of the inventory of phonemes, but also in a reduction of their combinatorial possibilities; such circumstances might have favoured the increase of lexical homonymy and of the formation of new words by means of derivative affixes such as *haka-/ha'a-* (causative), *-tina* (deverbative), *tā-* (e. g. , *tāvaha*, *mevaha répandu partout*; *tā'eva suspendre*), *tī-* (*tī-'ave porter sur les bras*, *tīma'u ombrager*, *tī'epo salir*), *pū-* (*pūti'eti'e élevé, haut*; *pūtu'i désobéir*), as well as through complete and partial reduplication and through borrowing (mainly from French and English).

The territory of the Marquesan archipelago consists of two great dialectal regions, northwestern and southeastern. They are usually regarded as two mutually intelligible and yet different languages. The northwestern language (MNW) is spoken in the islands of Nukuhiva, 'Ua Huka and 'Ua Pou; and the southeastern Marquesan language (MSE), in Hiva 'Oa, Tahuata and Fatuhiva. Since the latter group is closer to Tahitian, contacts between the Marquesas and the Society Islands in the past can hardly be ignored..

The present-day phonology of both dialects, namely their consonantism, has developed from the Proto-Polynesian system which has been undergoing a process of step-by-step simplification. The Proto-Polynesian consonantism included 13 consonants, i. e., *p*, *t*,

k, m, n, ng, q, f, s, h, w, l, r (Marck 2000: 23); their number decreased to 11 in Proto-Eastern Polynesian, i. e., *p, t, k, m, n, ng, q, f, s, w, r* (Marck 2000: 24), and its reduction continued in Marquesan. This means that while the Proto-Polynesian system could produce 4225 (65 x 65) theoretically possible bivocalic (disyllabic) lexemes of the structure (C₁)V₁(C₂)V₂, subsequent Proto-Eastern Polynesian has only 3600 (60 x 60) theoretically possible bivocalic lexemes, which might have resulted in an increase of homonymy (in comparison with Proto-Polynesian).

During the next phase of development the consonantal subsystem of MNW was reduced to eight phonemes (*p, t, k, ? -- m, n -- h, v*) and the same is true of MSE (*p, t, ? -- m, n -- f, h, v*), even though the results of reduction are slightly different. In MNW, *?* is a reflex of *r* (*r > k*); *k* is a reflex of both *k* and *ng* (*k & ng > k*); and *h* is a reflex of both *f* and *h* (*f* and *h > h*), while in MSE, *?* is a reflex of both *k* and *r* (*k* and *r > ?*); and *n* is a reflex of *n* and *ng* (*n* and *ng > n*). And thus homonymy increased in MSE by the fusion of (1) *k* with *r* and of (2) *n* with *ng*. On the other hand, the increase of homonymy in MNW is due to the fusion of (1) *k* with *ng* and also of (2) *f* with *h*. However, the subdialect of Taipivai deviates from the rest of Nukuhiva by its preservation of *ng*.

These changes do not completely explain the phonological differentiation within Marquesan. The borderline between MNW and MSE is to some extent fuzzy. For example, D. T. Tryon notes the existence of a variety of doublets (Tryon 1987: 31-37). Another interesting circumstance is mentioned by Karl von den Steinen (1905: 119-121). He published an article about the existence of Polynesian ‘secret’ languages in the recent past. He gives a brief characterization of such a language; according to him a language of this sort had been used by the Hapa’as during their occasional visits to Taiohae in Nukuhiva. According to von den Steinen, the secret“ language was derived from the common language by means of a mutual replacement of consonants within the words. This obviously does not explain the existence of doublets, but at least it documents a certain freedom of manipulation within the language and a willingness to accept the resulting doublets. The phenomenon of “special” languages employed for a variety of purposes is well known from other parts of Oceania as well.

The modern Polynesian languages overall are notable for their highly restricted inventory of phonemes, especially of the consonants. Their number varies from 8 to 12. The vocalic system is stable and consists of 5 vowels (*u - o - a - e - i*). The vocalic quantity is perceived as a sequence of two identical vowels and is marked as a repetition of the vowel (*aa, ee, ii, oo, uu*), but a macron may be used instead. These representations argue for the use of the terms monovocalic, bivocalic and multivocalic words. The modest inventory of consonants in Polynesian is accompanied by a set of strict phonotactic rules that exclude any consonantal clusters whatsoever and admit only one type of syllable, CV and its reduced form (C)V where the position of C need not be occupied. On the other hand, any absence of a vowel in the postconsonantal position is unthinkable. This means that the number of admissible monovocalic words cannot exceed the number of initial consonants before the vowel plus the number of such words without an initial consonant. These occur very frequently and are largely restricted to grammatical functions. Another class of relatively frequent words are bivocalic (C)V(C)V, often polysemic or homonymous and rarely used as grammatical markers. The frequency of occurrence of the more numerous polyvocalic words is considerably lower.

The problem of relatively low phonetic redundancy of bivocalic words and of their homonymy might have been perceived as a kind of obstacle in the early years of the era of

language revitalization and has been mentioned in several preceding articles devoted to Polynesian phonotactics, especially in Samoan, Tongan, Hawaiian and Maori (Krupa 1966, 1968, 1970, and 1971). The inventory of bivocalic words admitted by the Polynesian phonotactic rules and representing the foundation of their vocabulary is reminiscent of the situation in Chinese where the number of fundamental syllabo-morphemes comparable with bivocalic words in Polynesian does not exceed 1400 and the word stock expansion is warranted by composition (see below). For example, Hawaiian (in view of possible combinations of its number of consonants and vowels) could theoretically produce a total inventory of bivocalic lexical units amounting to 2,025. And according to the most reliable dictionary, without loanwords (Pukui – Elbert 1957) only slightly more than 50% of this theoretically possible inventory have been realized. Similarly G. B. Milner's Samoan Dictionary shows that only some 25% of the 3600 theoretically possible bivocalic words in that language are really observed (Milner 1966). Data in Churchward's *Tongan Dictionary* (Churchward 1959) confirm similar results in that some 33% of the theoretically admissible inventory of 4225 are utilized.

While the above mentioned dictionaries are of a relatively recent date and highly reliable, the Marquesan language, easily the most complicated within the East Polynesian branch, must make do, in the main, with Dordillon's dictionary published in 1904. In order to consider more modern changes, I decided to take into account not only Dordillon's dictionary, but also a small, new dictionary of Marquesan (Hervé Le Cléac'h 1997). Dordillon's dictionary is no doubt remarkable and comparable with the reliable Maori dictionary compiled by Williams in 1844 and after his death several times revised and enlarged by other linguists. In fact, Dordillon's unique dictionary, still used today, includes doublets from the Marquesan dialects, which circumstance inevitably relativizes the results of research. Dordillon's other shortcoming is his inconsequent marking of long vowels. However, unlike both Williams' Maori dictionary and Pukui and Elbert's Hawaiian dictionary, Dordillon does not avoid including loanwords either from French or English. Approximate data in Dordillon's and Le Cléac'h's dictionaries confirm that some 41% of bivocalic lexemes out of the total inventory of 3025 theoretically admissible bivocalic morphemes (words) have been observed. The set of recorded bivocalic lexemes numbers 1256; however, it is important to that this number includes a quantity of doublets (as well as loanwords) that result, for example, from the fact that MSE *f* corresponds with *h* in the MNV dialect. Moreover, in several instances, a duplication in the MSE dialect is due to the presence of both *k* and *ʔ*, etc. In this regard, D. T. Tryon found that some phonetic changes have taken place gradually, not abruptly (Tryon 1987: 31–37). Although *r* is no reflex of the Proto-Polynesian and Proto-East Polynesian *r*, it occurs in loanwords. A further irregularity is to be found in *k* as a reflex for *ʔ*, which itself was a reflex of the original *r* (i. e., $r > ʔ > k$). Another peculiarity of the Marquesan dialects is the absence of both *r* and *l*.

Additionally, the subdialect of 'Ua Huka that belongs to the MNW complex of dialects displays remarkable traces of the influence of immigrants from both parts of the archipelago. A couple of affinities with MSE are no doubt present in the island of 'Ua Huka belonging to the MNW complex of dialects. Another interesting feature is that some of the consonants occur more frequently in one of the two possible syllables (for example, *ʔ* obviously is preferred in the second syllable, since it has 175 occurrences there in contrast to only 118 occurrences in the first syllable; on the other hand *p* is preferred in the first syllable, just like *m* and *t*, but *n* occurs more frequently in the second syllable; *h* and *k* occur in both syllables with equal frequency).

In a sense, the bivocalic words (or morphs) represent the core of Marquesan word stock just as in the remaining Polynesian languages. In Marquesan the number of realized bivocalic words equals 1256 items (41,52% of all 3025 theoretically possible items); the data from other Polynesian languages are comparable: 1029 realized in Hawaiian (50,81% of all 2025 theoretically possible items) and 1258 in Maori (41,59% of all 3025 theoretically possible items). The results of investigating vocabulary of the three languages are not surprising. In Dordillon's dictionary of Marquesan (Dordillon 1904), 84% of all recorded lexemes are polyvocalic and some 16% bivocalic and monovocalic. In Williams' dictionary of Maori the relationship is 82 % versus some 18%, while in Hawaiian (Pukui and Elbert 1957) 88 % versus some 11-12 %.

Strict phonotactic rules may threaten the vocabulary with homonymy. For example, standard Chinese has at its disposal altogether 1324 syllabomorphemes that coincide with the set of all syllables. These syllabomorphemes are bricks which serve as building blocks of new compounds and thus of a potentially colossal word stock in Chinese; the theoretical number of all compounds consisting of two syllabomorphemes amounts to a respectable sum of 1,752.976 lexemes, which may be considered more than sufficient for any imaginable communication (Krupa 2003: 173–174; Yartseva 1990: 225).

Unfortunately the Marquesan and Polynesian languages in general suffer from a painful fact that their fundamental lexical elements are bivocalic and, as such, far from thrifty, while the Chinese fundamental blocks are monosyllabic, that is shorter and easier to operate with. The phonology of Chinese words (including terminology) deviates so radically from that of most European languages that calques are given preference before loanwords. The Polynesian languages, however, are open to loanwords which can be adjusted, simplified or even abbreviated.

Modernization of the Polynesian languages requires above all a fast and massive extension of terminology; without such a far-reaching modification of the vocabulary the local languages cannot compete with English or French, etc. The vocabularies of Maori, Tongan, Samoan, Hawaiian, Tahitian, Marquesan, etc. which were compiled before the vigorous revival movement of Polynesian culture and language that started approximately in the 1970's have been sufficient for the study of classical and early modern texts. However, the Polynesian revival that is evidenced especially in Hawai'i and New Zealand (in both regions the original languages acquired the status of co-official languages with English) and in French Polynesia, above all in Tahiti, is putting heavy demands upon both pedagogical and lexicographic activities of local linguists. For example, Tahitian lexicology was recently expanded by a new and voluminous English–Tahitian/Tahitian–English dictionary (Wahlroos 2002).

Bruce Biggs, moreover, has contributed to the development of present-day Maori with his English–Maori dictionary, which was issued four times between 1981–1990 (Biggs 1990). Cleve Barlow, a New Zealander of Maori origin, has translated Bruce Biggs' grammar into Maori and has thus contributed both to the expansion of Maori terminology and to the revitalization of the knowledge of language among the native inhabitants of the country (Barlow 1990). The same scholar has revived the pre-contact Maori culture in his handbook *Tikanga Whakaaro – Key Concepts in Māori Culture* in which ancient Maori culture and its terminology are explained (Barlow 1990²).

A new, expanded edition of the *Hawaiian Dictionary* by Mary Kawena Pukui and Samuel H. Elbert that was published in 1986 reflects new facts and the new state of the Hawaiian vocabulary enriched by new lexemes of both Hawaiian and English origin (Pukui –

Elbert 1986). For Hawaiian, another important dictionary obviously destined for a wider public appeared in 2003 (Māmaka Kaiāo 2003). This latter seems to be a prototypical dictionary compiled for an early revivalist era of the Hawaiian language. It includes Hawaiian words created, collected, and finally approved by the Hawaiian Lexicon Committee, founded in 1987, in the years 1987 – 2000. Māmaka Kaiāo attempts “to make available to the general public the new vocabulary that is being created by the Hawaiian Lexicon Committee that meets regularly several times a year and discusses new vocabulary for Hawaiian.” The process of the education of younger generations requires a considerable extension of the Hawaiian vocabulary. New lexemes may appear more or less simultaneously in various contexts and the process of selecting the convenient neologisms is one of the duties of the Committee, which used following guidelines to create new lexemes: (1) making minor changes to a word which already appears in the dictionary, (2) recording a word used by native speakers, but which is not found in the dictionary, (3) using reduplication of an existing word in order to alter or extend its meaning, (4) adding either prefix or suffix, (5) explaining the meaning of a word or term by using Hawaiian words, (6) combining Hawaiian words to create a new word, (7) combining Hawaiian words while shortening at least one of them, (8) extending the meaning of a word which is already found in the dictionary, (9) using a word from another Polynesian language with its meaning intact or slightly changed, (10) Hawaiianizing the orthography of a word from a non-Polynesian language (Māmaka Kaiāo 2003: XVII–XIX). Not only have the activities of the Committee between 1998 and 2003 enriched the word stock of modern Hawaiian with more than 1,000 new lexemes, but this process will no doubt continue.

The extension of vocabulary is a highly urgent task in all languages undergoing functional expansion. Three strategies are usually exploited in this process: (a) periphrastic/explanatory description, (b) borrowing from other languages, and (c) semantic shift. Borrowing is successful if the native speakers are acquainted with the culture behind the donor vocabulary and the loanwords often undergo a considerable phonetic adaptation. Semantic shift requires mainly a resourceful mind or explanation and is utilized virtually in all languages. Periphrastic description as a rule occurs in the first evolutionary phase, but its chief shortcomings are its awkwardness and inefficiency, which is why it may be gradually replaced by one of the two preceding strategies.

Since the existence of Marquesan in the near future is obviously not threatened, and its continued functioning will no doubt require an expansion of its word stock. Linguists can assume, then, that the tendency toward language revitalization in Polynesia may contribute to a revival of terminology that reflects the old, local culture.

References:

BARLOW, Cleve. 1990. Transl. Bruce Biggs. *Me Ako Taatou i te Reo Māori*. Auckland: University of Auckland Press, 1990.

BARLOW, Cleve. 1990². *Tikanga Whakaaro – Key Concepts in Māori Culture*. Auckland: Oxford University Press, 1990.

BIGGS, Bruce. 1981. *The Complete English–Māori Dictionary*. Auckland: Auckland University Press, 1981.

- CHURCHWARD, C. Maxwell. 1959. *Tongan Dictionary*. London: Oxford University Press, 1959.
- DORDILLON, Mgr. Ildefonse René. 1904. *Grammaire et Dictionnaire de la Langue des Iles Marquises*. Paris: Imprimerie Belin Freres, 1904.
- KOMIKE HUA'OLELO, 2003. Mamaka Kaiāo. *A Modern Hawaiian Vocabulary*. Hale Kuamo'o - 'Aha Pūnana Leo. Honolulu: University of Hawai'i Press.
- KRUPA, Viktor. 1966. The Phonemic Structure of Bi-Vocalic Morphemic Forms in Oceanic Languages. In *The Journal of the Polynesian Society*, 1966, vol. 75 no. 4, pp. 58-497.
- KRUPA, Viktor. 1968. On Phonemic Structure of Morpheme in Samoan and Tongan. In *Beiträge zur Linguistik und Informationsverarbeitung*, 1968, vol. 12, pp. 72-83.
- KRUPA, Viktor. 1970. A Typological Comparison of Polynesian Phonemic Systems. In *Beiträge zur Linguistik und Informationsverarbeitung*, 1970, vol. 20, pp. 58-92.
- KRUPA, Viktor. 1971. The Phonotactic Structure of the Morph in Polynesian Languages. In *Language*, 1971, vol. 47, pp. 668-684.
- KRUPA, Viktor. 2003. An Ideal Language of Worldwide Communication? In *Human Affairs*, 2003, vol. 13, pp. 171-178.
- Le CLÉAC'H, Hervé. 1997. *Pona Tekao Tapapa 'Ia Lexique Marquisien-Français*. Pape'ete.
- MILNER, G. B. 1966. *Samoan Dictionary. Samoan - English English - Samoan*. London: Oxford University Press, 1966.
- PERICLIEV, Vladimir. 2004. There is No Correlation Between the Size of a Community Speaking a Language and the Size of the Phonological Inventory of that Language. In *Linguistic Typology*, 2004, vol. 8, pp. 376-383.
- PUKUI, Mary Kawena, Samuel H. ELBERT. 1957. *Hawaiian-English Dictionary*. Honolulu: University of Hawaii Press, 1957.
- PUKUI, Mary Kawena, Samuel H. ELBERT. 1986. *Hawaiian Dictionary: Hawaiian-English and English-Hawaiian*. Honolulu: University of Hawaii Press, 1986.
- STEINEN, Karl von den. 1905. Proben einer frueheren polynesischen Geheimsprache. In *Globus: Illustrierte Zeitschrift fuer Laender- und Voelkerkunde*, Bd. LXXXVII 5: 7: pp.119-121.
- TRYON, Darrel T. 1987. Le cas du /k/ en Marquisien: Un développement inattendu. *Bulletin de la Société des Etudes Océaniques (Polynésie Orientale)*. No. 241 – Tome XX – No. 6 Décembre.
- TRUDGILL, Peter. 2004. Linguistic and Social Typology: The Austronesian Migrations and Phoneme Inventories. In *Linguistic Typology*, 2004, vol. 8, pp. 305-320.
- WALHROOS, Sven. 2002. *English-Tahitian/Tahitian-English Dictionary*. Honolulu: The Mā'ohi Heritage Press, 2002.

WILLIAMS, Herbert W. 1957. *A Dictionary of the Maori Language*. Wellington, N.Z: R. E, Owen, Government Printer, 1957.

YARTSEVA, V. N. (ed.). 1990. *Lingvisticheskii entsiklopedicheskii slovar*. Moscow: Sovetskaia entsiklopedia.

PhDr. Viktor Krupa, DrS.c
Ústav orientalistiky
Slovenská akadémia vied
Klemensova 19
813 64 Bratislava
Slovenská republika
kaorvikr@savba.sk