In *Current Issues in Unity and Diversity of Language (2009). Papers selected from 18th Intl. Congr. Ling*. (Seoul, 2008), pp. 875-892

Kaingang and Austronesian – Similarities between Geographically Distant Languages

Vladimir Pericliev
Mathematical Linguistics Dpt., Institute of Mathematics & Informatics,
Bulgarian Academy of Sciences, bl.8, 1113 Sofia, Bulgaria
peri@math.bas.bg

Abstract

The paper shows that significant similarities exist between two geographically distant language families: the Kaingang family, comprising Xokleng and Kaingang, spoken in South Eastern Brazil, on the Atlantic, and Austronesian, spoken in the Pacific. These similarities pertain to kinship semantic patterns, structural features and lexicon. In the paper, I sketch kinship and grammatical similarities and discuss in more detail the lexicon, paying special attention to various modes of arguing for significant similarity. One plausible explanation of the results found is a distant relationship between these families, a conclusion corroborated also by some extralinguistic facts. If such a hypothesis turns out to be true, one important implication would be some prehistoric contact between these distant parts of the world.
Keywords: Austronesian-Kaingang relationship hypothesis, Oceanic and South American contacts, language prehistory and classification, languages and migrations.

1. Introduction

The Southern-Ge Kaingang people, comprising the Xokleng and the Kaingang, presently reside in the states of Santa Catarina, São Paulo, Paraná, and Rio Grande do Sul in south-eastern Brazil, and speak two closely related languages, forming the Kaingang language family. No external relationship of the Kaingang people, or their language, with peoples or languages outside of Ge, or Amerind(ian) more generally, has so far been suggested. In this paper, I will describe substantial similarities between the Kaingang language family and the Austronesian family, in particular its Oceanic/Polynesian branch. First, I will sketch significant kinship pattern and structural similarities between the families. Then, I will present lexical parallels with sound correspondences, placing special emphasis on the comparison of "non-random word sets", like synonyms, homonyms, paronyms, lexical fields, etc. which impart additional plausibility to the comparisons.

2. Kinship similarities

A database of the kinship terminological patterns of 566 societies (Murdock 1970) was computationally explored with the goal of finding statistically significant similarities between pairs of societies speaking languages belonging to *different* language families according to the standard language classification source *Ethnologue*. The database has eight attributes with a total of 192 values. The attributes are the eight

relatives: grandparents, grandchildren, uncles, aunts, nephews and nieces, siblings, cross-cousins and siblings-in-law. The values for attributes are so-called "kinship patterns" such as, for the kin grandparent, "Bisexual Pattern" (having two terms, distinguished by sex, which can be glossed as "grandfather" and "grandmother"), "Merging Pattern" (having a single undifferentiated term, which can be glossed as "grandparent"), etc. The pair-wise society comparison revealed that Xokleng (the language representing the Kaingang family in the database) had 7 matching kinship patterns with the Amis (Taiwan) and 6 matching kinship patterns with both the Chuukese (Micronesia) and the Ulithian (Micronesia). These were the most significant results in comparison to all other investigated society pairs, at levels of $p = 0.005$, $p = 0.002$ and $p = 0.008$, respectively. Xokleng thus, quite unexpectedly, turned out to follow an Austronesian type of kinship patterning, showing also substantial kin pattern overlaps with other Austronesian languages like Rotuman, Samoan, and Māori. Checking the original files for Xokleng of the contributor of the database, anthropologist G. P. Murdock, showed his opinion about the similarity expressed in labeling Xokleng's system as "Normal Hawaiian". Further details can be found in Pericliev (2007: 42-48).

3. Structural similarities

Some structural features of the Kaingang language family (for Xokleng, cf. Henry 1935, 1948, Urban 1985, 1986, Gakran 2005; for Kaingang cf. Wiesemann 1972, 1978, 1986) were then tested against a set of typically Austronesian features that have been proposed (Klamer 2002) as a heuristic for suggesting the affiliation of a language (i.e. whether it is Austronesian or not). The Kaingang languages turned out to share Austronesian properties of both phonology and grammar.

Like Austronesian languages, and especially their Oceanic branch, which are known to have lost the voicing contrast in obstruents and to have developed prenasalized consonants in opposition to plain consonants, the Kaingang languages also do not have plain voiced obstruents, but contrast plain voiceless with prenasalized voiced consonants.

Many Austronesian languages prefer roots of CVCV type, and so do the Kaingang languages. Specifically, according to Henry (1948: 196), in Xokleng the CVCV pattern is prevalent and amounts to 35% of all root patterns in Xokleng (CV patterns being 14%, CVC 13%, CVCVC 12%, CVCCV 12%, CCV 5%, and CVCVCV 5%). As seen from these numbers, other typically Austronesian features of Xokleng are the "dropping" of final consonants (word-final consonants being present in only 25% of the patterns) and a dispreference for consonantal clusters (occurring in only 17% of the words). Besides, the possible final consonants and consonantal clusters are subjected to further restrictions we need not discuss here. The situation is similar in Kaingang.

Another, and related, phonological similarity is the insertion of paragogic vowels word-finally in both language families in order to open the syllable.

Similarly to many Austronesian languages, Xokleng forms emotional expressions by Verb + body part noun, in which the Experiencer of the emotion is the Possessor of the body part. E.g. Xokleng's expression for "I am angry" literally means "My heart splits (in several places)" (Henry 1935: 213).

Like many Austronesian languages, Xokleng's numerals seem to behave like verbs in that they act like verbal predicates and take the same predicating particles as the verbs.

Also, like other Austronesian languages mainly in Eastern Indonesia, but also scattered elsewhere in Western and Eastern Austronesian languages (e.g., Malagasy, Manobo, Hawaiian, Batak, etc.), Xokleng (cf. Urban 1986) employs parallelism in narratives, myths, poems and songs, a verbal art form in which semantically synonymous pairs/triples etc. of words and phrases are used in parallel utterances.

Finally, we note two other typically Austronesian features not mentioned by Klamer (2002), which are possessed by the Kaingang languages, viz. the affixing and reduplication as productive devices, and operating basically on verbs, and the fundamental verbal distinction in both Austronesian and the Kaingang languages between stative verbs and dynamic, or active, verbs.

4. Some lexical data and sound corespondences

A 100-item basic vocabulary list was compiled, and pair-wise comparisons were computationally made between Xokleng and Kaingang on the one hand, and the reconstructed Proto-Austronesian and five major extant Austronesian languages (Tagalog, Malay, Fijian, Samoan and Hawaiian) on the other hand. Standard similarity criteria for matching two word forms were used: generally, two identical consonants, or an identical consonant and a vowel plus a phonetically "similar" consonant. The method, which follows closely that of Oswalt (1991) was tested against pairs of languages with well understood relationships, both positive and negative. All comparisons were significant at levels shown below (with accuracy up to four decimal places).

	Proto-AN	Tagalog	Malay	Fijian	Samoan	Hawaiian
Xokleng	0.0000	0.0000	0.0002	0.0022	0.0003	0.0000
Kaingang	0.0087	0.0004	0.0179	0.0014	0.0000	0.0000

The preliminary computational approach to lexicon was suggestive and I proceeded with manual lexical comparison between the Kaingang and the Polynesian family, the latter being the "closest" geographically.

In the following, our data for Xokleng comes from different sources (Gensch 1908; Henry 1935, 1948; Gakran 2005), and for Kaingang from the dictionary Wiesemann (2002) and Wiesemann (1972). The data on cognate sets for Polynesian are based on Biggs & Clark (2006), Blust (1995), and the *Austronesian Basic Vocabulary Database*, but Tregear (1891), Williams (1957), Pukui & Erbert (1986), and Andrews (1865) are also consulted.

Table 1 shows the familiar sound correspondences (e.g., Biggs 1973, Tregear 1891) in the Polynesian languages Māori (New Zealand), and Hawaiian, alongside with those that are known to hold for the Kaingang family (Wiesemann 1978).

The following similarity sets are illustrations of the sound correspondences. Each similarity set has a gloss, giving a general meaning, which may be further specified for some languages if somewhat different from gloss. The items in one language family that

are similar to those of the other family are separated by equality "=". A dash "-" indicates a relatively clear word division in the Kaingang languages, segmenting stem from morphological endings (*-m*, *-n*, *-r*, *-y*, etc.). Brackets "()" enclose forms that are not part of the comparison (e.g. the verbal forms *ke*, *he* 'to say' in the Kaingang languages, which, when parts of larger verbs, simply indicate direct speech). A slash "/" stands for "or", and according to context may indicate a paragogic vowel at the end of a Kaingang word, or a doublet form. Note that a paragogic vowel in both Kaingang languages is the same or a more central vowel added after word final *r, v, y*; in Kaingang orthography, the paragogic vowel is not written, while in Xokleng an investigator like Henry (1935, 1948) and earlier linguists write these vowels and they may appear in my description below. Note also that a plus sign "+" after a Xokleng word below designates a word whose source is the early investigation by Gensch (1908), who does not make some phonological distinctions (e.g. central vs. non-central vowels) that are recognized today, which results in some indefiniteness in his description (in other respects pretty reliable). Finally, in the sequel, the Kaingang languages are written in their usual phonemic transcription, and the Austronesian languages in the way of their respective sources, with only minor and obvious changes (e.g. I use "ʔ" in place of the more common apostrophe " ' ").

Table 1. Sound correspondences between Xok(leng), Kai(ngang), Māo(ri), and Haw(aiian).

Nos.	*Xok/Kai/Māo/Haw*	***Examples Nos.***
1	p/p/p/p	23,24,25,39
2	t/t/t/k	31,32,36,39
3	k/k/t/k	1,2,5,26
4	k/k/k/ʔ	9,10,11,13,17,27
5	k/k/h/h	15,16,40
6	ð/Φ/t/k	33,34,35,40
7	h/h/(w)h/h	1,3,4,21
8	m/m/m/m	17,29,38
9	ŋ/ŋ/ŋ/n	11,12,24,36,37
10	w/w/w/w	28,30
11	l/r/r/l	1,2,6,7,8,9,15,16,18,19,22,38
12	a/ã/ā/ā	1,3,7,8,20,34,35
13	ɔ,a,ã/a,ã,ẽ/a/a	5,9,11,12,24,25,26,36,37,38,40
14	æ/a,ã/au,ao/au,ao	10,29,30
15	e/e/e/e	1,2,17,33
16	i,ẽ/i,ĩ/i/i	1,2,4,14,28,37
17	ũ/ũ/ū/ū	18,19
18	o,u,õ/o,ɔ,õ/ō/ō	6,13,27
19	u,o/u,o/u,o/u,o	4,9,15,16,17,21,22,36,40

The following abbreviations are used for frequently cited languages and language groups:

H = Hawaiian, K = Kaingang, M = Māori, X = Xokleng, PAn = Proto-

Austronesain, PCEMP = Proto-Central-Eastern-Malayo-Polynesian, PCMP = Proto-Central-Malayo-Polynesian, PMP = Proto-Malayo-Polynesain, POc = Proto-Oceanic, Pn = Polynesian, PPn = Proto-Polynesian.

1. Be same or similar: X *halike*, K *hã ri ke* = M *whārite*, H *hālike*.
2. Be same or similar: X *like*, K *ri ke* = M *rite*, H *like*.
3. Prefix: X *ha-*, K *hã-* = M *whā-*, H *hā-*.
4. Whistle: X *hui*, K *huɲ* = M *whio*, H *hio*.
5. Near: K *kakã* = M *tata*, H *kaka* 'fruits growing in clusters', Pn *tata* 'near'.
6. Ant/insect: X *lɔ* 'ant', K *ro* 'small bee' = M *rō* 'stick insect', H *lō* 'species of bug' (Tongan *lo* and Samoan *loi* 'ant'), PPn **lō* 'ant'.
7. Sun: X *la*, K *rã* = M *rā*, H *lā.*
8. Day: X *la*, K *kurã* = M *rā*, H *lā*.
9. Light/glow: X *kulaŋ* 'tomorrow, morning, early', K *kurã* = M *kura*, H *ʔula.*
10. Stick: X *kɔ*, K *ka* = M *mā/kau* 'handle', POc **kayu* 'stick, wood', Fijian (Bau) *kau* 'tree, stick'.
11. Sick: X *kɔŋɔ* 'sick, pain, wound', K *kaŋa* 'sick, pain, wound' = Central-Eastern Pn *kaga* 'to place a curse on someone', Tuamotan *kaŋa* 'injure, illtreat'; Samoic-Outlier Polynesian *kago*, Samoan *ʔago/si* 'wasted away from sickness'; Fijian *gogo* 'weak, wasted away', M *ngongo* 'sick'.
12. Worm: X *(wai-)ŋɔ+* 'louse' (*kɔŋɔ* 'grub'), K *ŋa* 'louse, woodworm' (also *nuŋʔə* 'worm-eaten, rotten, wasted away') = Pn *tuga* 'larva, maggot, organism causing internal decay', M *tunga* 'larva, worm-eaten, rotten (of timber)', Mangarevan *tuga* 'worm that devours sugar-cane'.
13. Penetrate: X *ko* = M *kō* 'wooden implement used for digging', H *ʔō.*
14. In, inside: X *ki*, K *ki* = M *ki*, H *ʔi*.
15. Cloth: X *kul/u*, K *kur/u* = M *huru* 'coarse hair', H *hulu*.
16. Blanket: X *kul/u*, K *kur/u* = H *huluhulu*.
17. Broth/food: K *kome* = M *kome*.
18. Shake (as dust from garment): K *rũ-m rũ-m* = M *rūrū*, H *lūlū*.
19. Scatter: K *rũ-m (ke)* = M *rui*, H *lū*.
20. Breathe (with difficulty): K *hã-m hã-m (ke)* = M *hāhā*, H *hāhā*.
21. Blow: K *hu (he)* = H *hu*.
22. Round/roll: K *ror* = M *(pi)rori*, H *loli.*
23. Birth/origin: X *pɔ* = M *pū*.
24. Throw: X *pãŋ*, K *pẽŋ* = M *panga*, H *pana*.
25. Hand/touch: K *pẽ* = M *pā*, H *pā*.
26. Finish: X *kaŋ*, K *kã-n* = H *kā*.
27. Dig: K *kõkõ-m* = M *kōkō*, H *ʔōʔō*.
28. Turn, reverse: K *wĩrĩ-n* = M *wiri* 'twist', H *wili*.
29. Carry: X *mɔ*, K *ma* = M *mau*, H *mao*.
30. Tree, forest: X *wæ+*, K *wã-ɲ* = M *wao*, H *wao*.
31. The (def. art.): X *te*, = M *te*, H *ke*.
32. Moist: X *tuy/u*, K *tuy* 'trickle intensely' = M *tōi*, H *kōi* 'flow, spurt'.
33. Heart: X *ðe*, K *Φe* 'heart, chest, breast' = M *ate* 'liver, heart', *ateate* 'bosom', Tuamotan *ate* 'heart'.
34. Shell: X *kɔða* 'shell', *kɔðan* 'strip embira bark' (*kɔtĩ* 'smooth'), K *kuΦẽn* 'to

peel, to shell', *kɨΦɛ* 'knife' = M *kota* 'cockle shell, anything to scrape or cut with', PPn **qota* 'dregs, residue'.

35. Wash clothes: X *ðaŋ*, K *Φã*, = H *kā*, Samoan *tā*.

36. Rash (of skin): K *tuŋa* 'rash' = Rarotongan *tunga* 'scabbed, weevily; covered with sores', Tokelau *tunga* 'pimple', Pn *tunga* 'blemish, imperfection'.

37. Hand, arm: X *nẽŋa*, K *nĩŋɛ* (or *niŋã*) = M *ringa*, Fijian *liga-* 'hand, forearm', PMP **liŋa* 'hand, arm'. (*l* ~ *n* is a common alternation within Polynesian languages, e.g. *lima* ~ *nima* 'five, hand').

38. Ashes: X *mlã,* K *mrẽj* = Pukapuka *malamala,* H *malamala* 'small piece of any substance broken from a larger', PPn **mala(mala)* 'chip, splinter, fragment'.

39. Pierce: X *patɨ* = M *pātia* 'spear', Penrhyn, Rarotongan, Tahitian *pātia* 'stab, thrust; spear, lance', Central-Eastern Pn *pā-tia* 'pierce'.

40. Stone: X *kaðu+* = M *whatu*, H *haku*, Mangarevan *ʔatu*, Pn *fatu*.

As seen from the above list, the sounds (vowels and consonants) of each Kaingang-family wordform, with only a small number of exceptions, are totally predictable from the Polynesian-family wordforms, according to the correspondences from Table 1. More illustrations of the correspondences follow in the next section.

5. More data (non-random lexical sets) and additional levels of evidence

By a "non-random lexical set", I will understand a set of words sharing a similarity in meaning, form, or perhaps in some other way. Familiar examples based on shared meanings are lexical fields such as body parts, kinship terms or lower numerals, and the classical example of a non-random lexical set, formed on this basis, is the basic vocabulary, whose elements (body parts, kinship terms, etc.) are themselves non-random word sets. There are also other non-random word sets, perhaps less commonly used in historical/comparative linguistics, which I will also compare below. Such comparisons provide an additional level of evidence over the traditional comparative method, and used complementarily to it, can lend further credibility to the comparisons, because a concentration of similarities within an antecedently fixed and limited (i.e. non-random) domain are generally less likely to be due to mere chance than the same number of matches in an open domain.

In the following, I slightly relax the strict conditions for counting a match used earlier, especially as regards vowels, and occasionally admit comparanda outside of the Polynesian family, for the following reasons. The first reason is methodological and has been nicely phrased by Campbell in the following manner:

> Where the intention is to call attention to a possible, but as yet unattested connection, one often casts a wide net in order to hold in as much potential evidence as possible. When the intention is to test a proposal that is already on the table, those forms admitted initially as possible evidence are submitted to more careful scrutiny. (Campbell 2004)

The second reason is empirical and has to do with the considerable presence of doublet forms involving vowels for the Kaingang family, not all of which may be registered in the sources I am using, and for which very strict sound matching criteria can turn out to be much too restrictive. E.g., as reported by Wiesemann (1972, 1978), in

Kaingang there exist a number of free vowel alternations: *a ~ ə, ã ~ a, ã ~ ɛ, ã ~ ɔ, ɛ ~ e, ɔ ~ o* (e.g. *kãgra ~ kãgrə* 'picture'*, kã ~ ka* 'tree', etc.).

5.1. Synonymous sets

If a single word from a set of synonyms in one language may resemble a word from a synonymous set in another language by pure chance, the event in which other words from both sets also resemble each other is much less likely to occur by chance. The higher the number of resembling words, the less likely the event, and hence the stronger the evidence for the existence of some link between the compared languages. By "synonyms" I understand both words with exactly coinciding and such with close meanings, and examples of both types are used as illustrations below.

(1) *NEAR*

Near1: K *kakã* = M *tata*, H *kaka* 'fruits growing in clusters', Pn *tata* 'near'.

Near2: X *la*, K *rã* = M *rā*, H *la* 'there, yonder'.

(2) *RAIN*

Rain1: X *ugua+* (? *ukua*) = M, H *ua*, Rapanui, Rennell Is. (Solomons) *ʔua* (POc **quzan*).

Rain2: X *kuta* 'fall (rain)', K *kutẽ* 'fall (rain)', *ta kutẽ* 'rain' = Pwamai (New Caledonia) *kuta*, Jawe (New Caledonia) *kut*.

Rain3: X *tɔ*, K *ta* = Woleai (Micronesian) *uta*, Lau, Longgu (Solomon Islands) *uta*.

Rain4: K *kɔ (he)* 'drizzle' (also *ta kɔ*) = M *tō* 'fall (of rain)', Pn **tō* 'fall (of rain)'.

Rain5: K *kɔkɔ (he)* 'drizzle' = M *totō* 'to ooze, to trickle', H *koko* 'reddish rain' (Andr).

(3) *HIT/STRIKE*

Hit/strike1: K *pẽ* 'hand, arm' = H *pā*, Niue *pā* 'slap, strike, touch, clap'.

Hit/strike2: X *puke*, K *pɔ ke* = Niue *poki* 'slap'.

Hit/strike3: K *tã-ɲ* 'kill, beat to death' = Niue *tā* 'strike, kill, adze'.

Hit/strike4: X *tuŋge*, K *tag ke* = Niue *tuki* 'knock, pound, mash'.

(4) *TREE/WOODS*

Tree/woods1: X *kɔ*, K *ka* = POc **kayu* 'stick, wood', Fijian (Bau) *kau*, M *mā/kau*, Namakir (Vanuato) *ka*.

Tree/woods2: X *kute*, K *kute* 'capão (flora), forming a type of flowers in southern Brazil, consisting of a group of tree vegetation' = POc **qutan* 'woods, forest', Mota *uta*.

Tree/woods3: X *wæ+*, K *wã-ɲ* = M *wao*, H *wao*.

Tree/woods4: X *bekud+* = PAn **biŋkudu*, Paiwan *vukid* (Blust 1995).

(5) *WATER*

Water1: X *ŋoy-waig+* 'river, stream', K *war* 'flood of water' = POc **waiʀ*, M *wai, awa* 'river, stream', H *wai*.

Water2: X *ŋoy*, K *ŋoy* = H *nō* 'seepage, to leak'.

Water3: K *kayã* 'salty' = M *tai,* H *kai* 'salty water'.

Water4: K *wãya* 'mixture (with water)' = Fiji *vai-na* 'mixture with water'.

Water5: K *ŋunŋɔn* 'liquid' = M *ngongi*, H *nono* 'oozing, seepage', Macassar *njonjo* 'liquid; to drip'.

Water6: K *ẽkɔ-r* 'sour water' = M *ehu* 'muddy', H *ehu* 'dusty, distrurbed'.

Water7: K *ku-pe* 'wash' (*ku-* prefix) = Vanua Lava (five dialects) *pe;* Mota *pei; pii*

'sprinkle water', *pia* 'foam from soap'.

Water8: K *me* 'liquid' = Mengen *me* 'water'.

Water9: K *run* 'carry, fetch water', *runya* 'vessel for water' = M *ranu-a* 'mix with liquid', Tongan *lanu* 'wash in fresh water', POc **danum* 'fresh water'.

Water10: X *ŋoy (be?)lele+* 'falling of water' = M *rere* 'waterfall' (Tr), H *wai lele* 'waterfall'.

(6) *BREAST/SUCK*

Breast/suck1: X *–kum(b)e+* 'female breast' = M *kōuma* 'breastbone' , Tuamotan *kōuma* 'chest, breast, bosom', PPn **uma.a* 'breast, chest'.

Breast/suck2: K *ũ-ɲ ũ-ɲ (he)* 'suck' = M, H *ū* 'female breasts', Tongan *huhu* 'female breasts, to suck'.

(7) *FIRE*

Fire1: X *pẽ,* K *pĩ* = POc **api*, M, H *ahi*, Anuta *api*.

Fire2: X *akpunu+* 'burn', K *kaprũn* 'line of fire, much fire in the firewood' = M *kapura*.

Fire3: X *ðai-kɔlɔ* 'kindle fire' = M *toro* 'to burn, a flame, burning' (also M *kora* 'fire'), Tongan *tolo* 'to rub, to ignite'.

Fire4: K *pũn* 'burn', *pũr* 'burnt' = Proto-Central Pacific **pula* 'burn', Proto-Micronesian **pwula* 'burn', Rotuman *pula* 'burn', H *pula* 'kindling', Waya (western Fiji) *bula-n* 'burn'.

Fire5: K *wãpũn* 'burn with big fire', *wãpũr* 'burnt to garbage' = M *māpura* 'fire'.

Fire6: K *kupũn* (=*hupũn*) 'to light up', *hupũr* 'illuminated with fire', *kupũr* 'burn in fire' = Central-Eastern Polynesian *kō-pura* 'flash', PPn **pula* 'shine', M *kōpura* 'flash, flicker, glance', Tahitian *opura/pura* 'to be flashing obscurely as fire', Tuamotu *kōpura/pura* 'to emit sparks, to glow or shine with unsteady light'.

Fire7: K *kɔm (ke)* 'light fire, light', *kɔm kɔm (ke)* 'to flash, to shine' = Tahitic *koma* 'spark', M, Tuamotu *koma*.

5.2. Polysemous/Homonymous sets

If a monosemous word in one language happens to resemble another monosemous word in another language by chance, the event in which a polysemous (or homonymous) word resembles a polysemous (or homonymous) word, is less likely to be merely coincidental, because it involves several meaning matches, rather than only one. The higher the number of meaning matches, and the more "unusual" the meanings of the polysemous/homonymous word, the less probable this event is, and hence the stronger the evidence for a non-chance link between the compared languages.

(1) In the following example, the compared forms have four meanings, the last two of which are "unusual" in that they seem quite divergent from the "basic meaning" of the form:

Sun: X *la*, K *rã* = M *rā*, H *lā*; cf. No.7.

Day: X *la* (K *kurã*) = M *rā*, H *lā*; cf. No.8.

Near: X *la*, K *rã* = M *rā*, H *la* 'there, yonder'.

Below: K *rã* = Fijian (Bau) *e rā*, Fijian (Navosa) *ra*.

Besides, the Kaingang form *ra* 'hot' and Xokleng *lɔ* 'hot', apparently related to *rã* and *la*, respectively, also have the same correlative word *rā/lā* 'hot' in Polynesian.

(2) This example exhibits three meanings, the last of which seems "unusual":

Cloth, clothes: X *kul/u*, K *kur/u* = H *hulu* ‘cloth; fur, wool’.

Blanket: X *kul/u*, K *kur/u* = H *huluhulu*.

Colour: ? X *kul/u* = Paumotan *huru* ‘colour’, H *hulu* ‘colour, nature, kind’; but cf. Section 6(3).

Compare also the different, but apparently related, Kaingang forms:

Hair: ? X *kren-kula-*+ (= ‘head’-*kula*) = M *huru* ‘hair, coarse hair (properly, of the body, but sometimes used for the hair of the head’; cf. *uru*, the head; a single hair), Rarotongan *uro*; cf. 6(5).

Fibre: K *kurẽ* (or *kurã*) ‘internal fibre of taquara’ = H *pulu* ‘coconut fibre’, cf. *hulu-hulu* ‘body hair, hair of eyelashes, fleece, fur, hairy’, PAn **bulut* ‘hairy filaments of certain plants, husk’, Niue *pulu* ‘fibre’.

(3) This example shows the instrumental metonymy “tree-stick”, which is typical for Oceanic languages, where the instrumental prefix *ka(i)-* has developed from **kayu* ‘tree’, cf. Lynch et. al. (2002: 70):

Tree: X *kɔ,* K *ka* = POc **kayu* ‘stick, wood’, Fijian (Bau) *kau*, M *mā/kau* ‘handle’, Namakir (Vanuato) *ka*.

Stick: X *kɔ,* K *ka* = Fijian (Bau) *kau*.

5.3. Polysemous/Homonymous and paronymous sets

By “paronyms”, here I understand words in one language that are very close in form. Paronymous sets normally involve some derivational pattern. If the different meanings of a polysemous/homonymous word in one language are rendered by a paronymous set in another language, this would be an unlikely event, since normally we would expect a random set of words, rather than some derivational pattern, to do the job. The higher the number of meanings of the polysemous/homonymous word rendered by paronyms, the less likely the event, and hence the stronger the evidence for a non-chance link between the compared languages. Below are examples of paronymous sets in the Kaingang family that correspond to the different meanings of polyesemous/homonymous sets in Polynesian.

(1) The Kaingang paronymous set *{tã, (ð/Φ)ã, kã}* matches the Polynesian polysemous word *tā*:

Break (as wood): K *tã* = H *kā* ‘to split or break wood’, M *tā* ‘strike’, Samoan *tā* ‘to break firewood, to break up a dry tree’, Fijian *tā* ‘chop’.

Kill: X *taiŋ*, K *tã-ɲ* ‘kill, beat to death’ = H *kā* ‘murder’, Rennell Is. (Solomons) *tā* ‘hit, strike, cut, kill’, Sikaiana (Solomons) *tā* ‘kill, hit, kick’, Luangiua (Ongtong-Java, Solomons) *kā* ‘hit, kill’.

Wash clothes: X *ðag*, K *Φã* = H *kā*, Samoan *tā*.

Beat (maize): K *Φãn* = H *kā* ‘to thrash out grain’.

Finish: X *kaŋ*, K *kã-n* = H *kā*.

It is important to notice that the sound pattern of forming the paronymous set in Kaingang, viz. the alternation *t* ~ *ð/Φ* ~ *k,* coincides with all the Kaingang sounds that correspond to Polynesian *t*, thus giving additional support to the three rules Nos. 2, 3, and 6 from Table 1. Notice also that the same alternations are additionally observed within Kaingang itself, e.g.: Xokleng *kɔða* ‘shell’, *kɔðan* ‘strip embira bark’, but *kɔtĩ* ‘smooth’, Kaingang ***t****i* ‘he/it’*;* ***Φ****i* ‘she’, Xokleng ***t****i* ‘he/it’*;* ***ð****i* ‘she’.

(2) The following example shows a correspondence between a paronymous set in

Kaingang and a paronymous set in Māori and a homonymous set in Hawaiian.

Throw: X *pãŋ*, K *pẽŋ* = M *paŋa*, H *pana*.

Shoot: X *pænũ* or *paŋ,* K *pɛnũ* = M *whana,* H *pana*, PPn **fana* 'shoot with a bow'.

Bow/arrow: X *puɲ* = H, Marquesan *pana* 'bow'.

5.4. Lexical fields

(1*) BODY PARTS*

Head1: X *klẽ*, K *krĩ* = PPn **qulu*, M *uru*, East Futuna, Tongan, East Uvea *ʔulu* 'head'.

Head2: X *paʔi,* K *pãʔi* = M *u/poko*, H *poʔo*.

Hand/arm1: X *nẽŋa,* K *nĩŋã* or *nĩŋɛ* = M *ringa* 'hand, arm', Fijian *liga* 'hand, forearm', PMP **linga* 'hand, arm'.

Hand/arm2: K *pẽ* = M, H *pā* 'touch', Ra`ivavai *pā* 'to touch with'.

Leg/foot: K *wagwag* 'to limp' (*wag (ke)* 'to pass to the other side') = M *wae/wae,* H *wawae*.

Eye: X *kɔna* 'eye, to look at', K *kanẽ* 'eye, to look at' = M *kana* 'stare wildly', Tuamotan *kana* 'stare', Central-Eastern Pn *kana* 'stare at' (M, Penrhyn *kanohi* 'eye').

Breast: X *-kum(b)e+* 'female breast' = M *kōuma* 'breastplate', Tuamotan *kōuma* 'chest, breast, bosom' (cf.also K *ũ-ɲ ũ-ɲ (he)* 'suck' = M, H *ū* 'female breasts').

Heart/liver: X *ðe*, K *Φe* 'chest, breast, heart' = M *ate* 'liver, heart', H *ake* 'liver', Tuamotan *ate* 'heart'.

Chin: X *lɔ*, K *ra* 'jaw' = M *rae*, H *lae* 'any projecting substance as a prominent forehead', Tongan *lae* 'forehead', Samoan *lae* 'beardless chin'.

Bone: X *koko+* 'bone, knuckle', *kuka*, K *kuka* = PAn **kukut* 'bone', **kuku* 'node, joint, knuckle', H *kuʔe* 'joint, the nuckles'.

Mouth/lips: X *-kuðo* (from *-kuso+*) 'lips' = Fijian *gusu-na,* Samoan, Tongan, Nieue *ngutu,* PAn **ŋusu*.

(2) *KINSHIP TERMS*

Xokleng's kinship terminology is discussed by Henry (1941: 175-80) and Kaingang's by Wiesemann (1974). The comparisons below seem to show substantial similarity of Kaingang terms with Micronesian languages (viz. the terms for wife, husband, and mother). Note that we found (Section 2) the best match in kinship semantic patterns between Xokleng and the Micronesian languages Chuuk and Ulithian. The entry for 'elder sibling' is interesting, especially if the forms for 'younger siblings' could also be conceived as "similar". The entry for "family" is a remarkable resemblance both formally and semantically, and further allows us to tentatively explain the meaning of Xokleng *kɔŋŋəŋ* and Kaingang *kaŋŋəŋ* 'an Indian, a Kaingang', previously unexplained as far as I know, as meaning 'one of our family' by reference to the respective words for 'family', viz. *kɔi ka* and *kaŋkã*. The entry for 'child1' notes a similarity with a cognate set in some Solomonic languages, while those for 'father' and 'child2' show likeness with words in isolated New Guinea languages, for which no cognate sets are proposed in the *Austronesian Basic Vocabulary Database*.

Wife+woman: X *plũ*, K *prũ* = Proto-Chuukic **p^{w}p^{w}úlú* 'spouse', **p^{w}úlú-wa-* 'married', Satawalese *púlú-wa-(n)* 'his wife', Marshallese (Eastern Dialect) *pālee-*.

Husband+man: X *man* (or *mɛn*), K *mɛn* = Proto-Micronesian, Proto-Chuukic

**m^{w}aane* ‘man, male’, Satawalese *m^{w}áán*, As, Minyaifuin (Gebe, New Guinea) *man* ‘husband’.

Father: X *yug*, K *yɔg* = Kaulong (Au Village, New Guinea) *iyok*.

Mother: X *ɲɔ̃*, K *-nə̃* = PAn **t-ina*, POc, Proto-Micronesian **tina*, Samoan *tinā*, Mokelese (Micronesia) *ina-(a)* ‘his mother’, Mortlockese (Micronesia) *ina-*.

Child1: X *ŋẽl/e*, K *ŋĩr* = Kwara’ae *ngela*, Lengo, Ghari *ŋgari*, Mbaungguu *ŋwɛle*, Fataleka *ŋwele* (all Solomonic).

Child2: K *kɔsin* = Sengseng (New Guinea) *po-kusan*.

Elder sibling: X *kake* ‘relative’, K *kãkɛ* (or *kãke*) ‘elder brother’, *kãke Φi* (*Φi* = female) ‘elder sister’ = PMP **kaka* ‘elder sibling’. (Blust 1995 notes that there can be no doubt that in PMP **kaka* referred to elder siblings, while another word, **huaji*, to younger siblings; Kaingang’s meaning correspondences to the latter term have the forms X *yawɨ*, K *yãwɨ* (*Φi*) ‘younger brother/sister’.)

Family: X *kɔi ka (he)* ‘relatives, people with the same body paint, family’, K *kaŋkã* ‘family, parents’ = Pn *kāiga* ‘kin, family, relative’, Tuvalu (Ellice Is.) *kāiga* ‘family, relative’, East Futuna (Horne Is.) *kāiga* ‘relative, family, parent’, East Uvea (Wallis I.) *kāiga* ‘parent, friend’, Pileni (Solomons) *kaega* ‘clan, family’, Samoan *ʔāiga*. ‘elementary family; family, lineage, kin, relatives’, Tokelau *kāiga* ‘kin, relative’, Tongan *kāinga* ‘relation, relative’.

(3) *NUMERALS*

Kaingang’s numerals, according to Wiesemann, are *pir* ‘one’, *rɛŋre* (or *rɛŋrɛ*) ‘two, second’, *tãŋtũ* ‘three’, *wẽŋ-kãŋra* ‘four’, *[ʂĩko]* ‘five’. Contemporary Xokleng has a base-two system, *pil* ‘one’, *lɛŋle* ‘two’, forming higher numerals by their combination: *lɛŋle to pil* ‘three’ *lɛŋle to lɛŋle* ‘four’, etc. (Greg Urban, personal communication). Henry (1945) lists *pil* ‘one’ and *læŋle* ‘two’, as well as one word involving the number four, viz. *ðəipa* ‘four-cornered’, but says that the Xokleng have no “real numbers”, presumably because e.g. *lɛŋle/rɛŋre* mean ‘companion, friend, co-spouse (Xokleng), brother (Kaingang)’. An early investigation (Gensch 1908) lists what seems to be a quinary numeral system: *toktenúnlo*+ ‘one’, *nunenglǽglo*+ ‘two’, *lenglæmú*+ ‘twice’, *umarikélko*+ ‘three’, *umpétko*+ ‘four’, *undupélemo*+ ‘five’ (stress is marked).

One: X *pil/i*, K *pir*, *pipir* ‘few’, *(wẽŋ) pãnpir* ‘reunited’, *pãnpin* ‘group together long objects’ = M *piri*, H *pili* ‘united, joining’, Marquisan *piʔi* ‘unite, kindred’. The Austronesian languages usually do not express quantity with the term, cf. however the related Tahitian *piti* with the same meaning ‘to join, to unite with another’ used as “two” in counting.

Two: K *rərə* ‘fight by twos’, *rən* ‘grapple’, *to rə* ‘fight’ (i.e. *rərə* probably *rə-rə* lit. ‘grapple, fight + two’) = Pn *rua* ‘two’, Tahitian *aro-rua* ‘the second in a combat’, or lit. ‘to face, to turn towards + second’.

Xokleng’s numerals given by Gensch look very interesting, even if not wholly understandable (to me). First, they seem like a continuation in counting after two of the other source(s), and the set undoubtedly shares the word for “two” with them. Secondly, they seem to contain some additional material over pure numerals roots, as witnessed e.g. by *nuneng-lǽglo*+ ‘two’ vs. *lenglæ-mú*+ ‘twice’. Table 2 highlights the sequence overlaps between Xokleng and Austronesian numerals for three, four, and five, accompanied by some information from Kaingang.

The highlighted Xokleng sequences, which are similar to those in Austronesian, seem to be supported by segmentation considerations. In all three words the initial *u-*

seems the Xokleng (and Kaingang) pronominal *ũ*, and *–kɔ* is a predicating particle (Greg Urban, personal communication). Thus, we get respectively: for "three", *ũ – mari – kel – kɔ,* where *mari*, if guessed by its meaning in Kaingang, is an emphasis word meaning 'also, too'; for "four", *ũ – (m)pen – kɔ* (the Xokleng nasal *n* is changed to the homorganic stop *t* in this context, as described by Henry 1945), and besides apparently *ðəi – pa* lit. 'corner – four'; and, for "five", *ũ – ndu – pélemo*, where *ndu* is a postposition meaning 'after'. Given this, the comparisons are:

Three: X *– kel* = H *kolu*, M *toru*, Proto-Micronesian **telu*, Kisar *wo-kelu*, Kei *tel*, Yamdena, Selaru *tél* (all Maluku), Central Masela *wɔkɛl*, Emplawas *wokɛl*, Dawera Daweloor *'tɛl* (all Babar).

Four: X *pa* or *pɛn* = H *hā*, M *whā*, Anuta *pā*, Buru (Maluku) *pā*.

Five: X *pelemo* = H *pālima* 'five times, in fives, fivefold', M *rima,* Tungag *palpalimana*, Tiang *patlima* (both New Ireland).

Table 2. Sequence overlaps (in bold) in the numerals for "three", "four", and "five"

	Xokleng (Gensch)	*Xokleng (Henry)*	*Kaingang*	***Austronesian***
THREE	*umari**kél**ko*			POc ***tol**u*, PCEMP ***təl**u*
FOUR	*um**pét**ko*	*ðəi**pa*** 'four-cornered'	*Φɨr* 'corner' ***pɛn**ugnu* 'quadrangle', ***pɛn**u* 'angle, corner'	POc, PCEMP **[ə]**pat**, *pati, *pani*
FIVE	*undupé**lem**o*			POc ***lim**a*

5.5. Long words

Similarities in longer words are less likely to be coincidental than similarities in shorter words as is well-known in historical linguistics. Below are some examples:

Be same or similar: X *halike*, K *hã ri ke* = M *whārite*, H *hālike*.

Smell, odour: X *(di?)kukræ+* 'stink', K *kɔkrã-ɲ* 'stink of rotten' (cf. also K *kaʔi* 'smell', *kãhɔr/ɔ* 'odourless') = M *kakara*, H *ʔaʔala* 'fragrant', East Futuna (Horne Is.) *kakalu* 'that which smells bad', M *kekererū* 'stinkroach'.

Snow: X *kuklule+*, K *kukrɨr/ɨ* = M *hukarere*.

Thunder: X *tɔ̃tɔ̃l*, K *(ta) tɔ̃ tɔ̃rɔ̃r (hɛ)* = M *(wha)titiri*, H *(he)kili*, Samoan *(fāi)titili*, Tuamoto *(fa)tutiri*, Tahiti (1773) *(pa)tiree* 'it thunders'.

Fish: K *kãkuΦər* 'small fish' = M *kō-kota* (shellfish), Mangarevan *kokota* 'small shell-fish'.

5.6. Compound words

The comparison of "compound words" (in our context, words comprising more material than just a single stem) presents significant evidence when both the compounds themselves and their constituent parts match in the compared languages. Below I give several illustrations, including as component part a grammatical morpheme.

(1) Compare entries Nos. 1-3 from our list of similarity sets. These three entries

illustrate similar compound words, e.g. Xokleng *halike*, Hawaiian *hālike*, etc., all having an identical meaning, viz. 'be same or similar'. These composite words comprise prefixes with identical form and function in present context, variously designated in the different languages: "emphatic" (Xokleng), "assertive" (Kaingang), and "causal" (Polynesian) (from **fā*), and stems (meaning 'be same or similar'), which are also formally similar, e.g. Xokleng Prefix: *ha* + Stem: *like*, Hawaiian Prefix: *hā* + Stem: *like*, etc. Such a coincidence of three compared entities, viz. a compound word and its two constituents, is so highly unlikely, as to be practically impossible to occur by mere chance.

(2) Henry (1935, 1948) lists for Xokleng *ŋaikaug* 'fear, be frightened', which comprises a prefix *ŋai-* (in other Xokleng sources written *wai-*+ or *wã-*; Kaingang *wãɲ-* or *wẽɲ-*), a "reflexive" with general meaning, and a stem *kaug*. Henry (1935, 1948) draws attention to the fact that this prefix has a valency-decreasing function, writing that "Certain verbs that begin with *ŋai* omit it when they have direct objects" (Henry 1935: 204), giving, among others, examples with the verb 'fear, be frightened'. Compare

Frighten: X *ŋaikaug* = H *makaʔu*, M *mataku*, PAn **ma-takut*.

Now, the stems seem related (cf. also Yamdena and Selaru (southeast Maluku) having *taut*, Kédang (Timor) *taug* or Puyuma (Formosa) *ma-kauð* 'fear, be frightened', where the middle consonant *k* is elided). Additionally, Austronesian (Proto-Oceanic) *ma-* is known to have a valency-decreasing function (Evans and Ross 2001), in that prefixing it to a verbal stem reduces the number of arguments this verb may have. As we see, Xokleng's prefix *ŋai-* behaves in exactly the same way, as actually does Kaingang's correlative. (Curiously, this same verb seems to be the common illustrative example of the phenomenon in both language families.)

(3) Compare the "long word" for fish above. The Polynesian *kō-kota* 'shellfish', is a compound word, comprising the prefix *kō* 'like, similar to' and the stem *kota* 'shell, scrapings'. The Kaingang *kãkuΦər* 'small fish' also seems to be a compound, viz. *kã-kuΦər*, as seen from its meaning equivalent *kuvər* in one of Kaingang's dialects (São Paulo) (Wiesemann 1978: 212). Its stem *kuΦər* then corresponds to Polynesian *kota* 'shell, scrapings' as evidenced by Kaingang *kuΦɛn* 'to shell, to peel', Xokleng *kɔða* 'shell'. The Kaingang prefix *kã-* seems to be allomorphic to the preposed Kaingang particle *kɔm* 'similar to, parallel to', and hence also corresponding to the Polynesian prefix *kō* 'like, similar to'.

5.7. A compound word and a synonymous set

The next example also describes an event which would not normally be expected in random comparisons, so is of some interest. The constituent parts of the compound Polynesian word *pō-fatu* 'stone' (M *pō(w)hatu*, Mangarevan *pōʔatu*), in which the sequence *fatu* 'stone', correspond to a synonymous set in the Kaingang family: Kaingang *pɔ* 'stone', and Xokleng *kaðu*+ 'stone' (from older *kasu*+ given by Gensch, in which *s* represents *ð*). Note that *kaðu* is "regular" in that it fully agrees with our sound correspondences in Table 1. It is interesting to note, whatever the exact implications, that the contemporary form for stone in Xokleng is *kɔðɨ*, and it nicely matches Pre-Rotuman **hafu* 'stone' pronounced *[hɔθu]* (cf. also Xokleng *ðe*, Kaingang *Φe* 'heart', Pre-Rotuman **afe* 'liver', pronounced *[æθe]*, Biggs 1965: 188).

6. Predictions

The usual way of testing a scientific theory is to draw logical deductions from the theory and check whether these predictions of the theory fit the data or not. The more of these predictions are true, the more corroborated the theory is. A similar approach, of course, can be used in historical linguistics. In the following, this general mode of reasoning takes a specific form. I take a sequence of two words in Kaingang <W1 + W2>, whose meaning as a whole, as well as the meaning of one of the constituent words, W1, is known, but that of the other, W2, is not, though W2 can be reasonably reliably predicted from context. This prediction, then, under the assumption that the Kaingang and Austronesian families are related – somewhat anticipating the hypothesis I shall put forward to explain the observed similarities – can be tested by seeing whether or not a correlative word in Austronesian exists with the appropriate predicted meaning. This reasoning mode mimics the real-life situation in which one tries to reconstruct the died-out meaning of a word in some language by reference to a related language that has, or might have, preserved this word meaning. A positive outcome of such a reconstruction constitutes positive evidence for the idea of the existence of a relationship among these languages. Additionally, it allows us to gain a better understanding of one of the languages by making the reconstruction.

(1) Kaingang *ka rigri* means 'small mosquito', *ka* in this context standing for 'mosquito'. The word *rigri*, however, is with unknown meaning and according to Wiesemann (personal communication) does not occur outside this word complex. Assuming a connection between Kaingang and Austronesian/Polynesian, we can predict that if *ringri* designates 'small' – as it would follow from context – we could find a correlative in Austronesian. Indeed, the formally similar *riki* means 'small' in Austronesian.

(2) Kaingang *ka pũr* 'black person, African', *pũr* 'burnt' = Proto-Central Pacific **pula* 'burn', Proto-Micronesian **pwula* 'burn', Rotuman *pula* 'burn', Hawaiian *pula* 'kindling', Waya (western Fiji) *bula-n* 'burn'. The unknown Kaingang meaning of *ka* in this context, probably 'person' as derived from context, is confirmed by Hawaiian *ka* meaning 'the one who, the person in question'.

(3) In the early documentation of Xokleng by Gensch (1908), the form *kulu-* is prefixed in 7 words denoting colour, and we have the sequence *kulu* + <colour term> + *ma*, as exemplified in *kulukuprima*+ 'white', *kuluklama*+ 'yellow', etc. The form *ma* is known to be a predicating particle, but what does *kulu* mean in this context? It would be natural to suppose that *kulu* means 'colour' or something related with colour, but no such use of the word is known e.g. to the investigator of Kaingang and Xokleng Wiesemann (personal communication). Assuming a relationship between Xokleng and Austronesian, we could expect to infer the actually unknown meaning of *kulu* from some corresponding word in Austronesian. Indeed, in a language like Hawaiian (cf. also Paumotan *huru* 'colour') one finds *hulu*, meaning – among other things – 'colour' ('species of colour', Andr), i.e. we have the "regular" correspondence:

X *kulu* = H *hulu* (PAn **bulu*); cf. also Nos. 15, 16.

As additional piece of evidence for this conclusion, consider the "unusual" coinciding polysemy of *kulu/hulu* in Kaingang/Austronesian families, cf, 5.2(2).

(4) Gensch (1908) lists as entry and sub-entry in Xokleng

X *kulug*-+ ‘dark’ (German *dunkel*)

X *kuru-loa*+ lit. ‘darker stuff’ (German *dunkler Stoff*),

which have the same root *ku(l/r)u* for ‘dark’ (*l* and *r* freely alternating in Xokleng). The meaning of Xokleng *loa* does not occur elsewhere in Gensch (1908) and is unknown. The Xokleng form *loa* does not seem to be associated with its substantive literal translation *staff*, for if it were the case, then, first, it should have preceded the adjectival *kuru* ‘dark’ (Xokleng’s word order being substantive + adj), and, secondly, it should, as an autonomous word, have been written separately from the adjectival form. Therefore, *loa* appears to be a sort of comparative degree, or intensifier, of the adjectival *kuru* ‘dark’. This prediction is fully substantiated by the Polynesian particle *loa*, having two meanings: (i) long, and (ii) ‘very, very much, exceedingly’, a post-posed intensifier probably used in our context. I note that there is a full match also of the other Xokleng word in the complex; thus, Hawaiian *kulu* ‘be late at night’ (Puk), ‘first night in which the moon is dark or can’t be seen’ (Andr); cf. also PAn ***kudem*,** Nggela *kuro*.

Further evidence that we are on the right track is the word *halikelɔ* ‘how long’, listed several times by Henry (1935), in which *halike* means ‘how’ (Kaingang *hẽri ke* ‘how’), and *-lɔ*, apparently related to the older form *loa*+, has to mean ‘long’, the first meaning of the identical Polynesian word. It is not clear whether in Xokleng *-lɔ/-loa*+ are separate words or are only used in compounds (as it is the case e.g. in Kapingamarangi). Note that our argumentation here follows the idea of Section 5.2, viz. comparison of polysemous sets.

(5) For the word ‘hair’ in Xokleng, Gensch gives *kren-kula-ma*+ (lit. = ‘head’-*kula*-predicating particle), where the meaning of *kula* is unknown. Henry (1935) lists for ‘hair’ *klẽ kɨki* lit. ‘head feather’. Given also that a formally similar word in Kaingang, *kule*, has a similar meaning, ‘internal fibres of taquara’, one could predict that the unknown meaning of *kula* is also feather or something similar. This prediction is corroborated by the complete match with Western Fijian (Navosa) *kula* ‘feathers’, as well as by the fact that a number of Austronesian languages (e.g. Sekar, Ujir, Yakan, Buru) form the words for ‘hair’ as the combination of ‘head+feathers’.

7. Extra-linguistic data

There are several pieces of evidence, besides the linguistic evidence, that seem to make less incredible the idea of the existence of some linguistic link between the geographically distant Austronesian-speaking and Kaingang-speaking populations. First, genetically South America is the most diverse part of the world, and Central America is more similar to North America than to South America (Cavalli-Sforza et al. 1994: 339).

Secondly, and more specifically, the Macro-Ge people, in drawing a phylogenetic tree of 23 American tribes, grouped according to linguistic criteria (Cavalli-Sforza et al. 1994: 323-4), were found to be the worst outliers (with Macro-Tucanoans).

Thirdly, and most importantly, apart from these more circumstantial, even if quite suggestive, pieces of genetic evidence that (at least some parts) of South America do not fit into the scenario of exclusively North-South population movement – which is the prevalent current belief today – there emerged recently genetic work giving sound evidence for the predominantly South-East Asian and Oceanic origin of South American

native populations. E.g. Ribeiro et. al. 2003, analyzing the Macro-Ge-speaking Xikrin and the Tupi-speaking Parakanã (note that Tupi is believed to be related to Macro-Ge), found them to be genetically similar to Indonesians and South-East Asian populations, concluding that "These results corroborate the existence of genetic affinities between Brazilian Indians and South-east Asian and Oceanic populations", their investigation being intended to "further contribute to the theory of a predominantly Asiatic origin of the American natives" (p. 59).

And, finally, an argument from Xokleng's beliefs. According to Henry (1945: 127) "The Kaingang [i.e. the Xokleng] have a clear idea of a period long ago when a number of events happened: *their ancestors came out of the sea* and over the mountains to the west…" [italics mine].

8. Conclusion

In this paper, I showed parallels between the Kaingang and the Austronesian language families in grammar and kinship semantics, as well as presented a network of similarities in lexicon. I cherish no illusion that my description is free from errors. Rather, I expect occasional inaccuracies to occur both in the data, and in my interpretation of some data, for it could hardly be otherwise, given the early stage of my investigation, the rather scarce knowledge we have today about the Brazilian languages, and especially my lack of first-hand knowledge of both examined language families. Nevertheless, the nature and scope of the presented parallels seem to enforce the conclusion that these similarities, taken collectively, cannot be just random, a conclusion I have also reached by direct computation of probabilities of matches in basic vocabulary comparisons. There are much too many coincidences, and this requires an explanation. One plausible explanation is a historical (probably genetic) relationship between the Kaingang and the Austronesian language families, and specifically its Oceanic branch, the exact nature of which is presently unknown. Further studies are required to get more definitive answers. The severe testing of this hypothesis is a worthwhile enterprise, because, if true, it will have far-reaching consequences not only for linguistics, but for the study of human pre-history as well. Such testing should be concerned with trying to undermine a significant part of the presented data, and trying to find a more plausible explanation of the (remaining) data. Regarding the current classification of the Kaingang languages as members of the Ge family, it will be clear that if my hypothesis is true, either this classification is not valid, or it is valid, and Ge is also somehow linked to Austronesian. At present I can only state these as two logical possibilities.

Acknowledgments. I wish to thank Robert Blust for making available to me his *Austronesian Comparative Dictionary*, and Ross Clark for POLLEX. Greg Urban and Ursula Wiesemann most kindly supplied answers to my questions regarding Xokleng and Kaingang noted in the paper. All errors are entirely mine. Partial support is also acknowledged from grant MI-1511 with the Bulgarian Ministry of Education and Science.

References

Andrews, Lorrin. 1865. *A Dictionary of the Hawaiian Language, To which is appended and English-Hawaiian vocabulary and a chronological table of remarkable events*. Honolulu: Henry M. Whitney.

Biggs, Bruce. 1965. Direct and indirect inheritance in Rotuman. *Lingua* 14. 383-445.

Biggs, Bruce. 1973. The languages of Polynesia. *Current trends in linguistics*, ed. Th. Sebeok, v. 8, 466-505. The Hague: Mouton.

Biggs, Bruce & Ross Clark. 2006. POLLEX: Comparative Polynesian Lexicon (computer data base). University of Auckland.

Blust, Robert. 1995. Austronesian Comparative Dictionary (computer files). University of Hawaii.

Campbell, Lyle. 2004. How to show languages are related: Methods for distant genetic relationship. In Brian D. Joseph and Richard D. Janda (eds.), *The Handbook of Historical Linguistics*, 262-282. Oxford: Blackwell.

Cavalli-Sforza, L. Luca et al. 1994. *The History and geography of human genes*. Princeton, N.J.: Princeton University Press.

Evans, Bethwyn & Malcolm Ross. 2001. The history of Proto-Oceanic *ma-. *Oceanic Linguistics* 40. 269-290.

Gakran, Namblá. 2005. *Aspectos morfossintáticos da língua Laklãnõ (Xokléng)*. Unicamp.

Gensch, H. 1908. Wörterverzeichnis der Bugres von Santa Catharina. *Zeitschrift für Ethnologie* 40. 744-759.

Henry Jules. 1935. A Kaingang text. *International Journal of American Linguistics* 8. 172-218.

Henry, Jules. 1941. *Jungle people: A Kaingáng tribe of the Highlands of Brazil*. New York: J. J. Augustin.

Henry Jules. 1948. The Kaingang language. *International Journal of American Linguistics* 14. 194-204.

Klamer, Marion. 2002. Typical features of Austronesian languages in Central/Eastern Indonesia. *Oceanic Linguistics* 41. 363-383.

Lynch, John, Malcolm Ross & Terry Crowley. 2002. *The Oceanic languages*. Richmond: Curzon Press.

Murdock, George. 1970. Kin term patterns and their distribution. *Ethnology* 9. 165-207.

Oswalt, Robert. 1991. A method for assessing distant linguistic relationships. In S. M. Lamb and E. D. Mitchell (eds.), *Sprung from Some Common Source: Investigations into the Prehistory of Languages*, 389-404. Stanford: Stanford University Press.

Pericliev, Vladimir. 2007. The Kaingang (Brazil) Seem Linguistically Related to Oceanic Populations, *Journal of Universal Language* 8(2). 39-59.

Pukui, M. & Erbert, S. 1986. *Hawaiian-English, English-Hawaiian dictionary*. Hononulu.

Ribeiro, D. M., Figueiredo, M. S., Costa, F. F., & Sonati, M. F., 2003. Haplotypes of α-globin gene regulatory element in two Brazilian native populations. *American Journal of Physical Anthropology* 121, 58 –62.

Tregear, Edward R. 1891. *Maori-Polynesian comparative dictionary*. Wellington, New Zealand: Lyon And Blair.

Urban, Greg. 1985. Ergativity and accusativity in Shokleng (Gê). *International Journal of American Linguistics* 51. 164-87.
Urban, Greg. 1986. Semiotic functions of macro-parallelism in the Shokleng origin myth. In J. Sherzer & G. Urban (eds.), *Native South American Discourse*, 15-57. Berlin: Mouton de Gruyter.
Wiesemann, Ursula. 1972. *Die phonologische und grammatische Struktur der Kaingáng-Sprache.* Janua Linguarum, series practica, 90. The Hague: Mouton.
Wiesemann, Ursula. 1974. Time distinctions in Kaingáng. *Zeitschrift für Ethnologie* 99.120-30.
Wiesemann, Ursula. 1978. Os dialetos da língua kaingáng e o xokléng. *Arquivos de Anatomia e Antropologia* 3. 197-217.
Wiesemann, Ursula. 1986. The pronoun systems of some Je and Macro-Je Languages. In Ursula Wiesemann (ed.), *Pronominal Systems*, 359-380, Tübingen: Gunter Narr.
Wiesemann, Ursula. 2002. *Dicionário Bilingüe Kaingang - Português.* Curitiba, Brazil: Editora Evangélica Esperança.
Williams, Herbert W. 1957. *A dictionary of the Maori language*. Wellington.