
Towards empirical classification of Kinnauri varieties

Anju Saxena

Proceedings of Conference on
Language Documentation & Linguistic Theory 3

Edited by Peter K. Austin, Oliver Bond, Lutz Marten &
David Nathan

19-20 November 2011 School of Oriental and African Studies, University of London

Hans Rausing Endangered Languages Project
Department of Linguistics
School of Oriental and African Studies
Thornhaugh Street, Russell Square
London WC1H 0XG
United Kingdom

Department of Linguistics:
Tel: +44-20-7898-4640
Fax: +44-20-7898-4679
linguistics@soas.ac.uk
<http://www.soas.ac.uk/academics/departments/linguistics>

Hans Rausing Endangered Languages Project:
Tel: +44-20-7898-4640
Fax: +44-20-7898-4349
elap@soas.ac.uk
<http://www.hrelp.org>

© 2011 Anju Saxena

No part of this publication may be reproduced, stored in a retrieval system, or transmitted, on any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without the prior permission of the author(s) of that part of the publication, except as permitted by UK copyright law.

ISBN: 978-0-7286-0398-1

This publication can be cited as:

Anju Saxena. 2011. Towards empirical classification of Kinnauri varieties. In Peter K. Austin, Oliver Bond, Lutz Marten & David Nathan (eds) *Proceedings of Conference on Language Documentation and Linguistic Theory 3*, 15-25. London: SOAS.

or:

Anju Saxena. 2011. Towards empirical classification of Kinnauri varieties. In Peter K. Austin, Oliver Bond, Lutz Marten & David Nathan (eds) *Proceedings of Conference on Language Documentation and Linguistic Theory 3*. London: SOAS. www.hrelp.org/eprints/ldlt3_03.pdf

Towards empirical classification of Kinnauri varieties

ANJU SAXENA
Uppsala University

1. INTRODUCTION¹

The aim of this paper is to examine the genetic relationships among nine Tibeto-Burman varieties spoken in the Kinnaur region in the Himachal Pradesh state in India,² using a computational approach applied to primary language data.

Some published works make brief mention of some Kinnauri varieties (e.g., Gerard 1842; Cunningham 1844; Bailey 1909). However, to date there has not been any systematic, comparative linguistic study of them, and consequently no systematic basis for examining how the Tibeto-Burman varieties spoken in Kinnaur relate to one another.

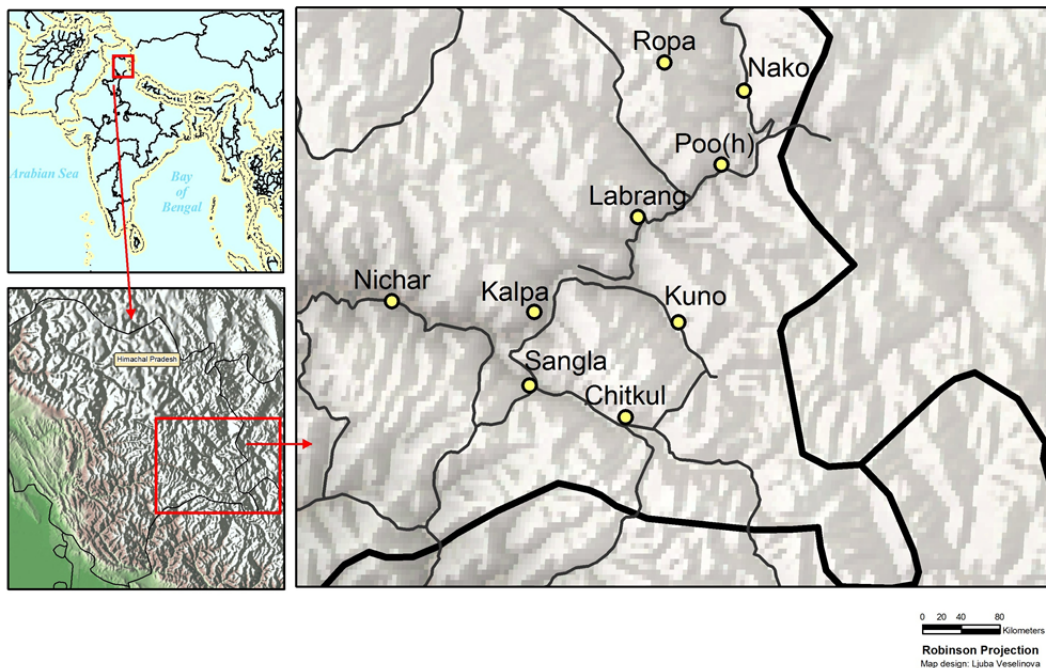
The fieldwork to collect the data used in this investigation was conducted in the following villages in Kinnaur: Nichar (Ni), Kanai and Sangla (Sa), Chitkul (Ch), Kalpa (Ka), Kuno (Ku), Labrang (La), Poo (Po), Ropa (Ro) and Nako (Na). See Figure 1. The main motivation for selecting these villages was to include data from as diverse a range of geographical regions as possible. The data comprise (i) a basic vocabulary list (a revised Swadesh list; Swadesh 1955) for all sites (242 senses); (ii) an extended IDS list (Comrie n.d.) for Sangla and Nako (1884 senses); and (iii) selected grammatical constructions. The focus in this presentation will be on the lexical data.

The procedure which was used here for comparing the word lists is similar to recent works in dialectometry (e.g., Nerbonne & Heeringa 2009) and lexicostatistics (e.g., Holman et al. 2008) in relying on a completely automatic comparison of the items in the word lists. However, it differs from most of these works (McMahon et al. 2007 being a notable exception) in its usage of rules tailored to the particular linguistic configuration under investigation, rather than a general method for string comparison. In this respect, it falls somewhere between traditional glottochronology, where expert statements are required about the cognacy of items, and these modern approaches, which rely entirely on surface form for determining identity of items, although it is closer to the latter than the former. The main methodological advantage of our approach is its consistency.

¹ I would like to thank my language consultants without whose help this study would have never been possible. I would like to thank especially Mrs. Santosh Negi and Mr. Padam Sagar for their generosity and for their help in practical matters. I would also like to thank Lars Borin (Språkbanken, University of Gothenburg) who provided help in computer-related matters for the present study. The research reported on here was supported by the Swedish Research Council.

² The focus here is on the Tibeto-Burman language for which the cover term Kinnauri is used. The Indo-Aryan language variety, sometimes called ‘Harijan boli’, which is also traditionally spoken in some parts of this region, is outside the scope of this paper.

Figure 1
Location of the villages in Kinnaur where data was collected



The organization of this paper is as follows. Section 2 describes the procedure for comparison. Section 3 presents some results of the dialect survey. Section 4 is the conclusion.

2. PROCEDURE FOR WORD LIST COMPARISON

The following procedure was used in this investigation:

- After the data collection and initial processing, a list of observations of relationships among varieties was made.
- This list formed the basis for developing a set of principles for comparing the linguistic correspondences in these Kinnauri varieties.
- Their purpose was to determine which segmental differences were non-significant for the purpose of considering items in different varieties as the same.
- The principles were encoded as context-sensitive phonological segment equivalence rules in a small computer program for comparing items fully automatically in order to achieve consistency.
- The equivalence rules were revised after inspection of the result, and the program run again on the data. This process was iterated until the result contained no obvious errors.

This procedure was used in order to establish the principles for a completely automatic comparison of a number of word lists representing the Kinnauri varieties under investigation. The results reported in the next section come from this automatic comparison.

3. RESULTS

Because of limited space, the results will be illustrated here by presenting only statistics about adjectives, nouns and the whole word list.³ For the same reason, only the adjective data will be shown and individual items discussed, while summary tables will be given for nouns and the word list as a whole.

3.1. Adjectives

Appendix A shows some of the investigated adjectives and the automatically computed correspondences among varieties. Table 1 contains the summary statistics extracted from the comparison of adjectives.

Tables 1–5 show summary statistics for the results of pairwise comparisons among all investigated varieties. Each comparison is shown as a fraction of shared items (the numerator) out of all items elicited for both the compared varieties (the denominator). For convenience, the fraction is also given as a percentage (rounded to an integer). Thus, the highlighted cell in Table 1 shows that Nichar and Ropa share 10 out of 19 adjectives elicited for both varieties, or 52%.

Table 1
Summary statistics for adjectives

	<i>Ni</i>	<i>Ka</i>	<i>Ro</i>	<i>Ch</i>	<i>La</i>	<i>Po</i>	<i>Ku</i>	<i>Na</i>
<i>Sa</i>	14/19 (73%)	11/19 (57%)	11/19 (57%)	2/19 (10%)	2/19 (10%)	1/19 (5%)	0/19 (0%)	1/19 (5%)
<i>Ni</i>		12/19 (63%)	10/19 (52%)	1/19 (5%)	1/19 (5%)	0/19 (0%)	0/19 (0%)	0/19 (0%)
<i>Ka</i>			12/19 (63%)	1/19 (5%)	1/19 (5%)	0/19 (0%)	0/19 (0%)	0/19 (0%)
<i>Ro</i>				2/19 (10%)	3/19 (15%)	1/19 (5%)	0/19 (0%)	1/19 (5%)
<i>Ch</i>					9/19 (47%)	1/19 (5%)	0/19 (0%)	1/19 (5%)
<i>La</i>						1/19 (5%)	1/19 (5%)	1/19 (5%)
<i>Po</i>							12/19 (63%)	14/19 (73%)
<i>Ku</i>								11/19 (57%)

Even though the adjective data set is small, the trend is obvious: Poo, Kuno and Nako form one group, and Sangla, Nichar, Kalpa and Ropa form another group. This is very clear for the majority of the adjectives in appendix A. Chitkul and Labrang,

³ A detailed presentation of the data, methodology and results are included in Chapter 2 in Saxena (forthcoming).

however, stand apart: In some cases a similar form occurs in both languages (e.g. *k^hati* COLD, *k^hai* BLACK, and *mãĩ* RED in both Labrang and Chitkul), but there are also cases (e.g., GOOD, WET) where separate forms occur in Labrang and Chitkul. If the forms in Labrang and Chitkul show similarity with any of the two clearer groupings, it is rather with the Sangla group than the Nako group; see, e.g., the terms for BEAUTIFUL, OLD and NEW.

3.2. Nouns

In Table 2 the combined statistics from a comparison of all nouns in the vocabulary questionnaire are presented.

Table 2
Summary statistics for nouns

	<i>Ni</i>	<i>Ka</i>	<i>Ro</i>	<i>Ch</i>	<i>La</i>	<i>Po</i>	<i>Ku</i>	<i>Na</i>
<i>Sa</i>	67/87 (77%)	72/88 (81%)	58/88 (65%)	47/88 (53%)	31/88 (35%)	11/88 (12%)	14/87 (16%)	13/87 (14%)
<i>Ni</i>		61/87 (70%)	48/87 (55%)	38/87 (43%)	26/87 (29%)	10/87 (11%)	12/86 (13%)	12/86 (13%)
<i>Ka</i>			57/88 (64%)	44/88 (50%)	29/88 (32%)	10/88 (11%)	13/87 (14%)	12/87 (13%)
<i>Ro</i>				38/88 (43%)	37/88 (42%)	17/88 (19%)	19/87 (21%)	17/87 (19%)
<i>Ch</i>					28/88 (31%)	11/88 (12%)	13/87 (14%)	12/87 (13%)
<i>La</i>						22/88 (25%)	25/87 (28%)	22/87 (25%)
<i>Po</i>							56/87 (64%)	62/87 (71%)
<i>Ku</i>								57/87 (65%)

Based on the nouns in these varieties we can clearly differentiate a core Sangla group (Sangla, Nichar and Kalpa) from a core Nako group (Nako, Poo and Kuno). Where these groups differ from each other, they do so regularly and consistently in all cases; when not, the same term is used in all varieties.

3.3. The whole word list

Table 3 summarizes the comparison statistics for the whole lexical questionnaire.⁴ As can be seen from the table, there is no single pair of varieties where all the 157 questionnaire senses have been recorded in both varieties. However, they share from 149 (e.g., Kuno–Nako) to 155 recorded senses (e.g., Sangla–Nako).

Table 3
Summary statistics for the full lexical questionnaire

	<i>Ni</i>	<i>Ka</i>	<i>Ro</i>	<i>Ch</i>	<i>La</i>	<i>Po</i>	<i>Ku</i>	<i>Na</i>
<i>Sa</i>	111/154 (72%)	116/155 (74%)	96/155 (61%)	66/152 (43%)	45/153 (29%)	13/153 (8%)	16/149 (10%)	16/155 (10%)
<i>Ni</i>		103/153 (67%)	87/153 (56%)	56/151 (37%)	39/151 (25%)	11/152 (7%)	14/148 (9%)	14/153 (9%)
<i>Ka</i>			100/154 (64%)	65/152 (42%)	42/153 (27%)	11/152 (7%)	15/149 (10%)	14/154 (9%)
<i>Ro</i>				60/151 (39%)	52/152 (34%)	19/152 (12%)	21/148 (14%)	20/154 (12%)
<i>Ch</i>					45/150 (30%)	14/151 (9%)	16/148 (10%)	16/151 (10%)
<i>La</i>						26/150 (17%)	31/149 (20%)	27/152 (17%)
<i>Po</i>							97/149 (65%)	105/152 (69%)
<i>Ku</i>								97/149 (65%)

Again, the same picture as before emerges:

- Sangla, Nichar, and Kalpa form a clear grouping, with Ropa closely associated.
- Poo, Kuno, and Nako form another grouping, possibly somewhat less close than the Sangla group.
- Finally, Chitkul and Labrang show greater affinity to the Sangla group than to the Nako group, but are distant from both. At the same time, Chitkul and Labrang are equally – or in some instances more – distant from each other as they are individually from the Sangla group.

Since Swadesh lists are often used in this kind of lexicostatistical investigation, summary statistics for all Swadesh list items in the questionnaire (88 senses) are

⁴ No verbs are included in the comparisons. Verbs were included in the basic vocabulary questionnaire, but were provided in such a variety of different (basic) forms by language consultants, that it was not feasible to attempt to normalize them at this stage, without much more knowledge of each of the varieties. This means that out of the 242 senses elicited, 157 were used for the present investigation.

shown in Table 4. In Table 5 we show corresponding statistics for the 25 senses used in the questionnaire from the 40-item globally most stable Swadesh subset defined by Holman et al. (2008).⁵

Table 4
Summary statistics for all Swadesh list items

	<i>Ni</i>	<i>Ka</i>	<i>Ro</i>	<i>Ch</i>	<i>La</i>	<i>Po</i>	<i>Ku</i>	<i>Na</i>
<i>Sa</i>	65/87 (74%)	64/87 (73%)	59/87 (67%)	33/86 (38%)	28/86 (32%)	9/86 (10%)	10/83 (12%)	11/87 (12%)
<i>Ni</i>		58/86 (67%)	53/86 (61%)	28/86 (32%)	25/85 (29%)	7/86 (8%)	9/83 (10%)	9/86 (10%)
<i>Ka</i>			58/86 (67%)	31/86 (36%)	25/86 (29%)	7/85 (8%)	9/83 (10%)	9/86 (10%)
<i>Ro</i>				31/85 (36%)	33/85 (38%)	12/85 (14%)	13/82 (15%)	14/86 (16%)
<i>Ch</i>					31/85 (36%)	10/85 (11%)	10/83 (12%)	11/85 (12%)
<i>La</i>						15/84 (17%)	17/83 (20%)	16/85 (18%)
<i>Po</i>							57/83 (68%)	62/85 (72%)
<i>Ku</i>								58/83 (69%)

If anything, the Swadesh list comparison ties Ropa closer to the Sangla group. Otherwise, nothing substantial changes. However, it is a bit surprising that the reduced Swadesh list of Table 5 still shows such a big difference between Sangla and Nako Kinnauri. Even if these are to be counted as separate languages, which the results of the present investigation indicate quite unambiguously so far, they are still close enough that we would have expected a measure devised for large-scale worldwide lexicostatistical work to show them to be nearly identical. This is clearly not the case, and is an issue which deserves further study.

⁵ It should be mentioned at this point that determining which are the 40 most stable Swadesh list senses is logically not the same as determining the 40 most stable senses in the languages of the world, unless the 100-item Swadesh list were already known to contain the 100 most stable senses in the languages of the world. This of course is not known with certainty; even if there is much accumulated linguistic experience in support of this particular selection of items (Hymes 1973:131), it still has the status of ‘rule-of-thumb knowledge’ rather than solid empirically verified fact (Lees 1953).

Table 5
Summary statistics for the 25 most stable Swadesh items

	<i>Ni</i>	<i>Ka</i>	<i>Ro</i>	<i>Ch</i>	<i>La</i>	<i>Po</i>	<i>Ku</i>	<i>Na</i>
<i>Sa</i>	20/25 (80%)	22/25 (88%)	20/25 (80%)	11/25 (44%)	11/25 (44%)	3/25 (12%)	3/25 (12%)	3/25 (12%)
<i>Ni</i>		19/25 (76%)	19/25 (76%)	11/25 (44%)	11/25 (44%)	3/25 (12%)	3/25 (12%)	3/25 (12%)
<i>Ka</i>			18/25 (72%)	9/25 (36%)	10/25 (40%)	3/25 (12%)	3/25 (12%)	3/25 (12%)
<i>Ro</i>				10/25 (40%)	14/25 (56%)	4/25 (16%)	4/25 (16%)	4/25 (16%)
<i>Ch</i>					11/25 (44%)	4/25 (16%)	4/25 (16%)	4/25 (16%)
<i>La</i>						5/25 (20%)	5/25 (20%)	5/25 (20%)
<i>Po</i>							19/25 (76%)	21/25 (84%)
<i>Ku</i>								19/25 (76%)

3.6. Empirical classification of Kinnauri varieties

According to the Ethnologue (Lewis 2009), the varieties investigated here fall under the Kanauri sub-branch of the Tibeto-Kinnauri subgroup of Tibeto-Burman. The Ethnologue sometimes provides lexical similarity percentages (possibly the same measure as that called ‘inherent intelligibility’) for related languages. The data for Kanauri languages in Kinnaur are reproduced in Table 6. The Ethnologue figures are compared with the results of the comparison of the full vocabulary list used in the present investigation (Table 3), as well as with the two Swadesh subsets (Tables 4 and 5). Cells marked with an asterisk in Table 6 indicate inconsistencies in the Ethnologue, where the language name is ‘Lower Kinnauri’, but the ISO 639-3 code given is ‘kjo’, that of Harijan Kinnauri (or Harijan boli), an Indo-Aryan language among whose alternative names ‘Lower Kinnauri’ is not listed. I have chosen to follow the language name, rather than the ISO 639-3 language code.

The figures are somewhat in agreement, although the Ethnologue numbers are generally higher.⁶ The Ethnologue lists Sangla and Nichar as dialects of Lower Kinnauri (kfk) and provides a ‘mutual intelligibility’ figure for the two dialects of 79% (see Table 6). I assume that Sangla Kinnauri – as the most studied variety – has been the basis of comparison with the other Kanauri languages in the Ethnologue, so that the figures in parentheses in Table 6 in all cases show the correspondence to the Sangla variety.

⁶ Presumably the Ethnologue numbers were computed manually, using human judgements about cognacy of Swadesh list items. The automatic method used here is conservative; it misses a number of such correspondences, but hardly ever generates false cognates. This would explain most of the differences between the Ethnologue data and those of this investigation.

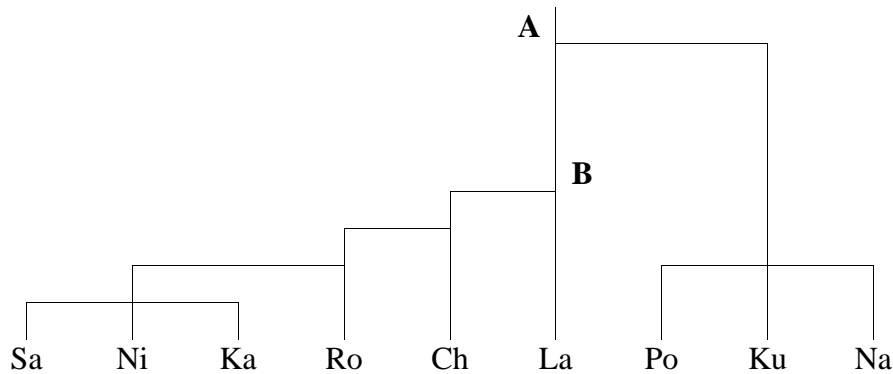
Table 6
Ethnologue lexical similarity percentages for Kanauri languages in Kinnaur compared to the figures in Tables 3, 4 and 5 (in parentheses)

	<i>(Ni) kfk</i>	<i>(Ch) cik</i>	<i>jna</i>	<i>(La) scu</i>	<i>ssk</i>	<i>(Ku) tpq</i>
<i>(Sa) kfk</i>	79% (72/74/80%)	46% (43/38/44%)	49%	* 45% (29/32/44%)	* 38%	—
<i>(Ch) cik</i>			51%	43% (30/36/44%)	38%	—
<i>jna</i>				70%	65%	—
<i>(La) scu</i>					67%	—
<i>(Po) nes</i>						71% (65/68/76%)

According to the Ethnologue classification, we could then classify Sangla, Nichar, Kalpa, and possibly Ropa as the Ethnologue language Lower Kinnauri (kfk), Chitkul as Chitkuli Kinnauri (cik), and Labrang as Shumcho (scu). On the other hand, the Ethnologue recognizes at least two languages in the Nako group, Tukpa (Kuno) and Bhoti Kinnauri (Poo). Overall, the lexical comparison made here shows Poo and Nako to be slightly closer to each other than either is to Kuno, but the differences are small. If we are to speak of languages rather than a dialect continuum, these results indicate that we should recognize three languages or one language, not two, as in the Ethnologue. The Nako group is consistently different from the Sangla group by a large margin in all cases; thus, the results shown here are in no way in contradiction to a classification of these three varieties – Poo, Kuno, and Nako – as Tibetan (rather than Kanauri) languages or varieties, as is sometimes seen (e.g., Bailey 1909). This is also confirmed by the examination of the linguistic structure of Nako (see Saxena forthcoming).

In summary, the results of the comparison show that the investigated varieties can be classified into three (or possibly four) groups, where Sangla, Nichar, Ropa and Kalpa form one group, and Poo, Kuno and Nako form another. Chitkul and Labrang fall somewhere between these two distinct groupings, being (individually) closer to one or the other group concerning some linguistic features, but distinct with regard to other linguistic features. The combined evidence of this study thus supports a grouping of the nine investigated Kinnauri varieties approximately like that shown in Figure 2. Whether we put the Kanauri node at *A* or *B* in Figure 2 will depend on whether Poo, Kuno and Nako are seen as Tibetan (Kanauri = B) or Kanauri (Kanauri = A) languages, as discussed above.

Figure 2
Preliminary grouping of the investigated Kinnauri varieties



4. CONCLUSION

This paper presents a systematic comparison of nine Kinnauri varieties in order to throw some light on the genetic classification of these underdescribed linguistic systems. The comparison has focused on the lexicon, which was investigated using an automatic, computational and purely quantitative method inspired by modern work on lexicostatistics and dialectometry, combined with traditional linguistic analysis and reasoning.

This kind of computer program is a tool which could have a place among many others in the linguist's toolbox. Correctly used, it can be very helpful and save a lot of effort. It can also be incorrectly used to produce garbage very quickly. Correct use involves recognition of the fact that computers are completely consistent and indefatigable, but capable of making extremely stupid errors (from a human point of view).

Seen as a tool, the computer program does not need to be perfect; it will be sufficient that it helps the linguist to do their job more efficiently than without it. Scaling up linguistic investigations seems to require computers. For instance, with well-designed computer tools we can imagine how all kinds of comparative linguistics – historical, typological, and contact linguistics – could work effectively directly with large amounts of primary language data, including fieldwork recordings and transcriptions, and language corpora, rather than the small amounts of secondary data that even the most ambitious typological projects (such as WALS) have had to make do with (see, e.g., Wälchli 2009, Abney and Bird 2010).

Even though the computer tool used for this investigation falls far short of such a vision, in my view it has turned out to be quite helpful to have an automated way of quickly calculating similarities among the language varieties under scrutiny, not least as a 'generator' of new research questions.

It has helped to provide some answers and in the process proved its worth. Given that one accepts lexicostatistics using Swadesh-style core vocabulary lists as producing valid results, the refinement of this method that I have presented seems to

be a step in the direction of making this methodology even more useful for teasing out the relationships among closely related language varieties.

REFERENCES

- Abney, Steven & Steven Bird 2010. The Human Language Project: Building a universal corpus of the world's languages. *48th Annual Meeting of the Association for Computational Linguistics*, 88–97. Uppsala: ACL.
- Bailey, T. Grahame 1909. A brief grammar of the Kanauri language. *Zeitschrift der deutschen morgenländischen Gesellschaft* 63, 661–687.
- Comrie, Bernhard (ed.) n.d. *The intercontinental dictionary series*. <<http://lingweb.eva.mpg.de/ids/>>.
- Cunningham, J.D. 1844. Notes on Moorcroft's Travels in Ladakh, and on Gerard's Account of Kunáwar, including a general description of the latter district. *Journal of the Asiatic Society of Bengal* 13, 172–253.
- Gerard, Capt. Alexander 1842. A vocabulary of the Kunawar languages. *Journal of the Asiatic Society of Bengal* 11, 478–551.
- Holman, Eric. W., Søren Wichmann, Cecil H. Brown, Viveka Velupillai, André Müller & Dik Bakker 2008. Explorations in automated lexicostatistics. *Folia Linguistica* 42(2), 331–354.
- Hymes, Dell 1973. Lexicostatistics and glottochronology in the nineteenth century (with notes toward a general history). In Isidore Dyen (ed.), *Lexicostatistics in genetic linguistics*, 122–176. The Hague: Mouton.
- Lees, Robert B. 1953. The basis of glottochronology. *Language* 29(2), 113–127.
- Lewis, M. Paul (ed.) 2009. *Ethnologue: Languages of the world*, 16th edition. Dallas: SIL International. Online version: <<http://www.ethnologue.com/>>.
- McMahon, April, Paul Heggarty, Robert McMahon & Warren Maguire 2007. The sound patterns of Englishes: Representing phonetic similarity. *English Language and Linguistics* 11(1), 113–142.
- Nerbonne, John & Wilbert Heeringa 2009. Measuring dialect differences. In Jürgen Erich Schmidt & Peter Auer (eds.), *Language and space: Theories and methods*, 550–567. Berlin: Mouton De Gruyter.
- Saxena, Anju forthcoming. The linguistic landscape of Kinnaur.
- Swadesh, Morris 1955. Towards greater accuracy in lexicostatistic dating. *International Journal of American Linguistics* 21(2), 121–137.
- Wälchli, Bernhard 2009. Data reduction typology and the bimodal distribution bias. *Linguistic Typology* 13, 77–94.

APPENDIX A: ADJECTIVE COMPARISON TABLE (EXTRACT)

	<i>Sa</i> [1]	<i>Ni</i> [2]	<i>Ka</i> [3]	<i>Ro</i> [4]	<i>Ch</i> [5]	<i>La</i> [6]	<i>Po</i> [7]	<i>Ku</i> [8]	<i>Na</i> [9]
<i>BAD/ 186</i>	[1/3] mari	[2] ma:r	[1/3] mari	[4/6] halam	[5] majəro	[4/6] halam	[7] ak ^h e	[8] t ^h ua	[9] ɲānba
<i>BEAUTIFUL</i>	[1/2/3/4/ 5/6] jare	[1/2/3/4/ 5/6] jare	[1/2/3/4/ 5/6] jare	[1/2/3/4/ 5/6] jare	[1/2/3/4/ 5/6] jare	[1/2/3/4/ 5/6] jare	[7] laho	[8] lakpo	[9] là:fo
<i>BIG/ 27</i>	[1/2/3/4] teg; te:g	[1/2/3/4] teg	[1/2/3/4] teg	[1/2/3/4] teg	[5] tɛi	[6] tʃei	[7/8] tʃ ^h epo	[7/8] tʃ ^h epo	[9] tʃ ^h étpo
<i>BLACK/ 176</i>	[1/2/3/4] rok	[1/2/3/4] rok	[1/2/3/4] rɔk	[1/2/3/4] rɔ(k)	[5/6] k ^h ai	[5/6] k ^h ai	[7/8/9] nakpo	[7/8/9] nakpo	[7/8/9] nàkpo
<i>COLD/ 181</i>	[1/2/3/4] lis; tʃɪk; sɔktʃɪk	[1/2/3] sɔktʃɪk	[1/2/3/4] sɔk; tʃɪk	[1/3/4] sɔ(t)	[5/6] k ^h ati	[5/6] k ^h ati	[7/8/9] ʃanmo	[7/8/9] ʃanmo	[7/8/9] ʃanmo
<i>GOOD/ 185</i>	[1/2/3/4] dam	[1/2/3/4] dam	[1/2/3/4] dam	[1/2/3/4] dam	[5] dzoi	[6/8] epo	[7/9] ganʃɪn; ʃɪmbo; demo	[6/8] epo	[7/9] dèmo; zàɲbo; ʃɪmpo; ʃɪmbo; ètpo
<i>GREEN/ 173</i>	[1/2] ra:g	[1/2] ra:g	[3/4] rak	[3/4] ra:k	[5] p ^h i	[6] tiɲ	[7/8/9] ɲonpo; ɲonpo	[7/8/9] ɲonpo; dɔmpo	[7/8/9] ɲonpo (blue-green)
<i>LONG/ 28</i>	[1] lames	[2/3/4] lamɔs	[2/3/4] lamas	[2/3/4] lamas	[5] rui	[6] fui; sarpa	[7/8/9] riɲpo	[7/8/9] riɲpo	[7/8/9] riɲpo
<i>NEW/ 183*</i>	[1/2/3/4] ɲug; ɲu:g	[1/2/3/4] ɲu:g	[1/2/3/4] ɲu(k)	[1/2/3/4] ɲu(k)	[5/6] nui	[5/6] nui	[7/8/9] soma	[7/8/9] soma	[7/8/9] sóma
<i>OLD/ 184</i>	[1/2] ɔʃk	[1/2] ɔʃk	[3] ɔʃ(k)	[4] ɔʃ	[5] hui	[6] uʃi	[7/8/9] ɲiɲpa	[7/8/9] ɲiɲpa	[7/8/9] ɲiɲba
<i>RED/ 172</i>	[1/2/3] ʃuig	[1/2/3] ʃuig	[1/2/3] ʃuig	[4] ʃuik	[5/6] māi	[5/6] māi	[7/8/9] marbo	[7/8/9] marbo	[7/8/9] máruo
<i>ROUND/ 190</i>	[1/4/7/9] baʃɛs; girgir	[2] baʃɔs	[3] baʃɔs	[1/4/7/9] girgir	[5/6/9] kirkir	[5/6/9] kirkir	[1/4/7/9] girgir	[8] tɔktɔk	[1/4/5/6/7/ 9] kirkir; girgir
<i>STRAIGHT/ 189</i>	[1/2] sɔɔɔs	[1/2] sɔɔɔs	[3] sɔɔɔs	[4] sɪʃta	[5] pɔdəra	[6] k ^h osra	[7] ʃanbo	[8] ombo	[9] t ^h anbo
<i>WET/ 194</i>	[1/4] pɪntʃ; t ^h is	[2] spenək	[3] pɪnk	[1/4] t ^h is	[5] rakʃi:	[6] t ^h ɪsi	[7] lɪnpa	[8] lemba	[9] lánte
<i>WARM/ 180</i>	[1/2/3/4] bɔk	[1/2/3/4] bɔk	[1/2/3/4] bɔk	[1/2/3/4] bɔ(k)	[5] tat ^h əra	[6] kotʃ ^h ra	[7/9] ʃonmo; ʃonmo	[8] tonpa	[7/9] ʃonmo
<i>WHITE/ 175</i>	[1/2/3/4] t ^h og	[1/2/3/4] t ^h og	[1/2/3/4] t ^h og	[1/2/3/4] t ^h og	[5/6] tʃāi	[5/6] tʃai	[7/8/9] karuo	[7/8/9] karbo	[7/8/9] káruo
<i>YELLOW/ 174</i>	[1/2] pɪg	[1/2] pɪg	[3/4] pɪk	[3/4] pi:k	[5/6] lei	[5/6] lei	[7/8/9] seruo	[7/8/9] serbo	[7/8/9] séruo

This table is to be read as follows. The leftmost column lists the revised Swadesh senses, with their Swadesh list numbers. A missing number means an addition to the Swadesh list. An asterisk indicates an item in the most stable Swadesh list subset defined by Holman et al. (2008).

The nine columns for Kinnauri varieties contain the following information: (1) one or more numbered correspondence sets; (2) one or more elicited words or phrases. One or more shared correspondence sets count as 1 – i.e., a shared vocabulary item – in the summary statistics calculations shown in Tables 1–5.