UNIVERSITY OF TARTU
Faculty of Science and Technology
Institute of Mathematics and Statistics
Mathematics

Jürgen Rannap

Mathematical analysis of Numic languages

Bachelor's Thesis (9 EAP)

Supervisors: PhD Marco Patriarca NICPB
PhD Els Heinsalu NICPB
docent Peeter Oja

Tartu 2017

Mathematical analysis of Numic languages

Bachelor's thesis

Jürgen Rannap

Abstract. The objective of this bachelor's thesis is to confirm or falsify hypotheses on the diffusion of Numic languages in the north-western United States. The theoretical part of the thesis provides an overview of the mathematical methods used to calculate phonetical difference and determine the genealogical classification of different language groups. The main part focuses on the implementation of these approaches on the Numic languages database and the evaluation of the results with regard to hypotheses found in articles on the linguistic diffusion of the Numic languages.

CERCS reasearch specialisation: H360 Applied linguistics, foreign languages teaching, sociolinguistics

Keywords: Native American languages, Uto-Aztecan languages, dialectology, data processing, complex systems

Numi keelte matemaatiline analüüs

Bakalaureusetöö

Jürgen Rannap

Lühikokkuvõte. Käesoleva bakalaureusetöö eesmärk on kinnitada või ümber lükata Põhjla-Ameerikas asuvate Numi keelte leviku kohta varasemalt püstitatud hüpoteese. Töö teoreetilises osas antakse ülevaade matemaatilistest meetoditest, määramaks keeltevahelist fonoloogilist kaugust ning keelkondade genealoogilist liigitust. Töö peamine osa keskendub Numi keelte andmebaasi uurimisele, rakendades eespool mainitud meetodeid, ning saadud tulemuste põhjal Numi keelte leviku hindamisele, baseerudes varasemalt püstitatud hüpoteesidele.

CERCS teaduseriala: H360 Rakenduslingvistika, võõrkeelte õpetamine, sotsiolingvistika

Märksõnad: Ameerika põlisrahvaste keeled, Jutoasteegi keeled, murdeuurimine, andmetöötlus, komplekssüsteemid

Contents

1 Introduction

Recent years have seen a growth of interest in the field of mathematical linguistics. Mathematical linguistics is an interdisciplinary field of research that lies in the intersection of mathematics and applied linguistics, focusing on understanding and generating natural language through the usage of computer science and mathematical methods [1]. The objectives of mathematical linguistics are extensive and can range from machine translation to the study of language evolution. We delve into the subdivision of comparative linguistics, establishing the relatedness of languages using some of the mathematical methods, which view language interactions as a complex system.

In the present thesis we study the interesting case of the Numic languages of the Uto-Aztecan language family, which originated from an ancestral Proto-Uto-Aztecan language and spread from the north-western United States throughout Mexico [2]. A proto-language is a base ancestral language from which it is assumed that many languages derived [3]. The Numic languages are the northernmost branch of the Uto-Aztecan language family, situated in the north-western United States, mainly in the Great Basin – a watershed in a set of smaller basins between generally high tablelands, bounded by the subrange of the Rocky Mountains from the east and the Sierra Nevada mountain range from the west [4] (see Figure 1.1). Like the Uto-Aztecan languages, the Numic languages are thought to have originated from a proto-language. Proto-Numic speakers were by nature highland people, originating from the southern part of the Sierra Nevada mountain range and Death Valley [2], as illustrated in Figure 1.1. The Numic languages are a dialect system, consisting of seven main languages, which are divided into three subgroups, which follow a geographical basis: Western Numic languages, Central Numic languages and Southern Numic languages [4], as seen in Figure 1.2. The Colorado River language is divided into three dialects, of which we include Southern Paiute and Ute in our thesis. The estimated quondam distribution of these Numic languages is depicted in Figure 1.3 based on the Native-Land

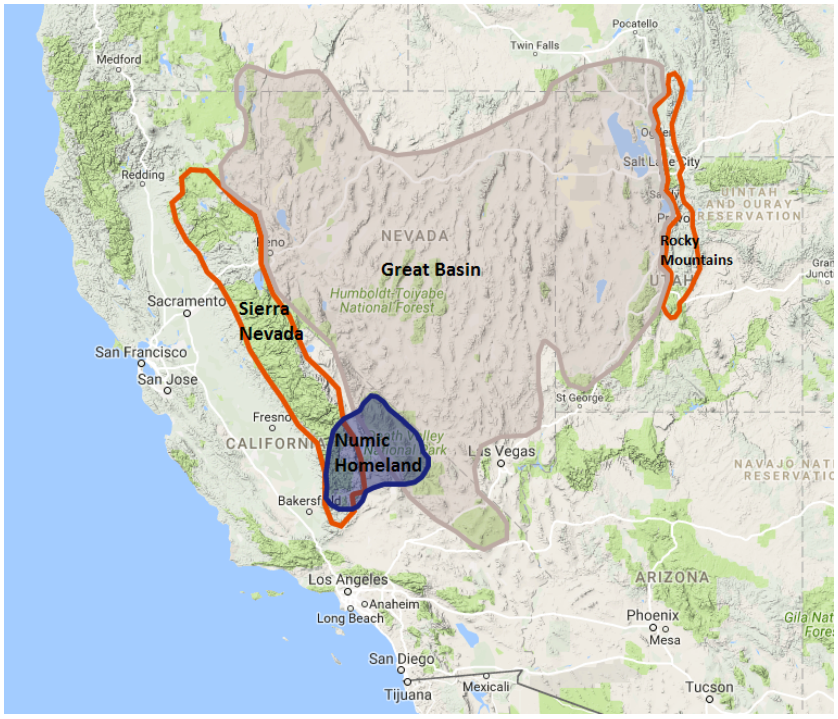


Figure 1.1: *Map of the location of Numic homeland and surrounding area. The location of the Numic homeland is marked in blue, Great Basin in gray, Sierra Nevada and Rocky Mountains in red.*

website[1]. Still, little is known about the areas that these languages used to occupy due to an overall lack of knowledge of their movements and interactions.

In this thesis we seek to understand the relationships and connections between these Numic dialects with the help of complex networks. We use methods previously applied in the study of the Otomanguean languages of the Mesoamerican language family in the work of Léonard, Patriarca, Heinsalu et al. [5], as well as a novel approach, *Historical Glottometry* [6], used to determine the genealogical classification of languages.

We focus on the Numic languages following the diffusion of the speakers from their homeland into the Great Basin. This expansion of the Numic, called the *Numic Spread Hypothesis*, first proposed by Sydney Lamb [7] and

[1] http://native-land.ca

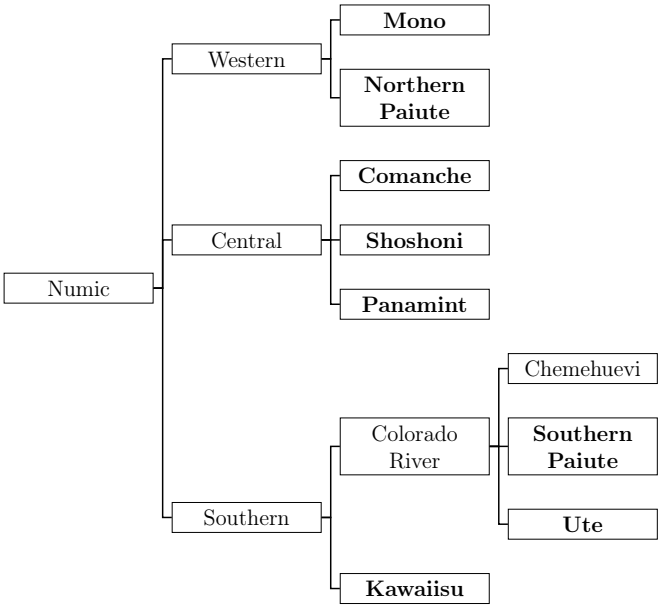


Figure 1.2: *Tree of Numic languages - languages considered in our thesis given in bold*

subsequently attested by other Numic linguists, argues that the Numic languages experienced a rapid spread over the Great Basin from their homeland around AD 1000. This interesting diffusion process has a geographical merit as Lamb draws attention to the distribution of the three language branches, where each language in the Numic homeland seems to stretch across the Basin to the northeast. The necessity of representing the Numic dialects via complex networks, besides the fact that language systems in nature are often interlaced and complex, becomes clear with the addition of intersecting territories, where speakers of one dialect were linked to neighbouring languages by social and economic activities [8]. Looking at the Numic language system, we must acknowledge the Comanche dialect on its own, as Comanches used to be a native empire [9], which may entail a vast movement of the speakers in other Numic dialect territories and because of that befog the connections in the dialect network.

The thesis is devided into two parts. In the first part, Section 2, we give an overview of the theoretical methods that in the second part, Section 3, will be applied to study the Numic languages. The second part also includes

the evaluation of the results with regard to hypotheses found in articles on the linguistic diffusion of the Numic. The analysis of the database of the Numic languages depicted in Figures 1.2 and 1.3 is carried out using the MATLAB environment. The visualization of the network of Numic languages is constructed using *Gephi* software and the maps are constructed using *Google My Maps*, which are available on the Google My Maps website[2].

The author of this Bachelor's thesis would like to express his sincere gratitude to Prof. Jean Léo Léonard from Paris-Sorbonne University for suggesting such an interesting subject and for providing the database. I would further like to thank Flore Picard for her assistance in the analysis of the Numic database as well as the supervisors Dr. Marco Patriarca and Dr. Els Heinsalu from the National Institute of Chemical Physics and Biophysics and Dr. Peeter Oja from the University of Tartu for support and help.

[2] `https://drive.google.com/open?id=1-KsLnU9S0mXbrleKUy18SM7m8jo&usp=sharing`

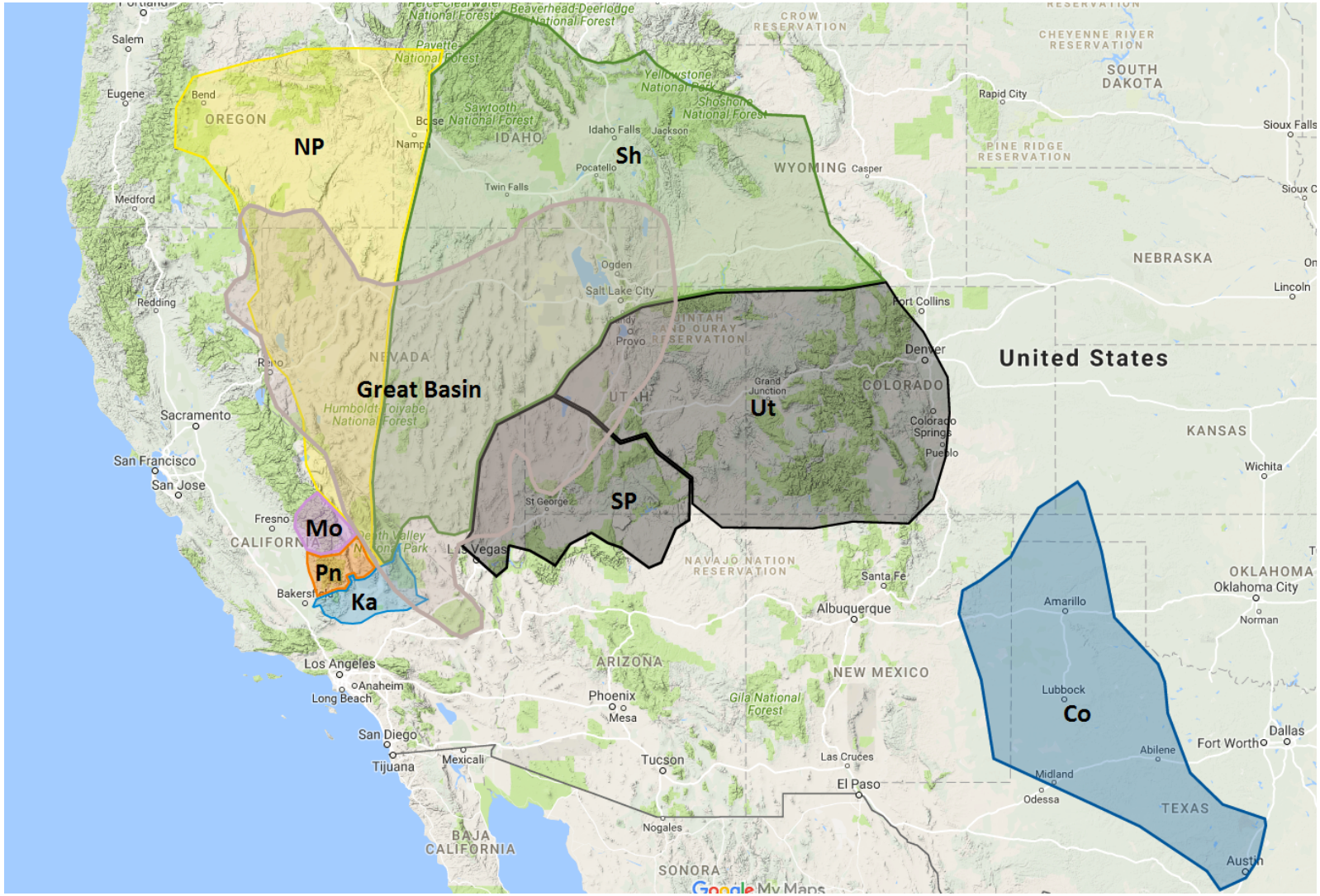


Figure 1.3: *Map of the Numic languages*
Co *- Comanche,* ***Pn*** *- Panamint,* ***Sh*** *- Shoshoni,* ***Ka*** *- Kawaiisu,* ***SP*** *- Southern Paiute,* ***Ut*** *- Ute,* ***Mo*** *- Mono,* ***NP*** *- Northern Paiute*

2 Methodology

In this section, we give an overview of the theoretical framework later used to examine the Numic languages database. We introduce the Levenshtein distance, a similarity measure, widely used for calculating phonological differences and Historical Glottometry, a novel approach used to determine the genealogical classification of different languages.

2.1 Levenshtein distance

The Levenshtein distance is used to quantify the difference between two given sequences. The simple nature of this measure implies its extensive usage in many different fields ranging from genetics to the calculation of phonetical distances in linguistics. The Levenshtein distance between two given strings s and t is the minimum number of edit operations, which include insertions, deletions and substitutions, needed to change string s to string t and is denoted by the symbol $L(s,t)$ [10].

Proposition 2.1. Let X be a set of finite sequences and $L : X \times X \to \mathbb{N}_0$ the Levenshtein distance function. The pair (X, L) is a metric space.

Proof. It suffices to show that $L : X \times X \to \mathbb{N}_0$ satisfies the three conditions of distances i.e. the identity of indiscernibles, symmetry and triangle inequality. Let $s, t, u \in X$. We start from the identity of indiscernibles, that is $L(s,t) = 0 \Leftrightarrow s = t$. It is easy to see that we can transform s to t with 0 operations if and only if $s = t$. Symmetry of the distance follows from the observation that for every operation we have an inverse operation, therefore $L(s,t) = L(t,s)$. For the triangle inequality we must show that $L(s,u) \leq L(s,t) + L(t,u)$. Let us assume by contradiction that $L(s,t) + L(t,u) < L(s,u)$. While s transforms into t in $L(s,t)$ operations and t into u in $L(t,u)$ operations, then the composition of the operations $L(s,t)$ and $L(t,u)$ transforms s into u. Since $L(s,u)$, by definition, is the minimum number of edit operations to change s to u, we have a contradiction to the definition of Levenshtein distance. □

For the mathematical definition of Levenshtein distance, compiled in [10], let s and t be two given strings of length $|s|$ and $|t|$, respectively. We denote s_i as the first i characters of the string s, $i = 1, \ldots, |s|$ and t_j as the first j characters of the string t, $j = 1, \ldots, |t|$. We introduce $L_{s,t}(i, j)$ as the Levenshtein distance between s_i, $i = 1 \ldots, |s|$ and t_j, $j = 1 \ldots, |t|$. Therefore, $L(s, t) = L_{s,t}(|s|, |t|)$, as it would be the Levenshtein distance between all the characters of s and all the characters of t.

Definition 2.1.

$$L_{s,t}(i,j) = \begin{cases} \max(i,j) & \text{if } \min(i,j) = 0, \\ \min \begin{cases} L_{s,t}(i-1,j) + 1 \\ L_{s,t}(i,j-1) + 1 \\ L_{s,t}(i-1,j-1) + I_{s_i \neq t_j} \end{cases} & \text{if } \min(i,j) \neq 0, \end{cases}$$

where

$$I_{s_i \neq t_j} = \begin{cases} 1 & \text{if } s_i \neq t_j, \\ 0 & \text{if } s_i = t_j. \end{cases}$$

Example 2.1. The recursive nature of Definition 2.1 confirms by default, when calculating the Levenshtein distance between two given strings s and t, the calculation of $L_{s,t}(i, j)$, $i = 1, \ldots, |s|$, $j = 1, \ldots, |t|$. Let us now consider two strings from the Numic database: "***ahpəʔ***" from the Comanche dialect and "***appə***" from the Panamint dialect, both translating to "*father*" in English. The words are encoded in IPA phonetic transcription. The calculation of the Levenshtein distance can be illustrated with a matrix shown in Table 2.1.

Table 2.1: *Example of Levenshtein distance*

	a	***p***	***p***	***ə***
a	$\underline{0}$	1	2	3
h	1	$\underline{1}$	2	3
p	2	1	$\underline{1}$	2
ə	3	2	2	$\underline{1}$
ʔ	4	3	3	$\underline{2}$

For better understanding, we move through the matrix row by row. The numbers in the first row of the matrix can be interpreted as the Levenshtein distance, changing from the string "***a***" to a subsequence of the word "***appə***". For example, the Levenshtein distance from "***a***" to the sequence "***app***" would be equal to 2 as there needs to be two insertions. The numbers in the second row are calculated from the string "***ah***" to a subsequence of the string "***appə***". The last row is calculated in the same way from the string "***ahpəʔ***" to a subsequence of the word "***appə***". The Levenshtein distance, hence the minimum number of edit operations needed to change the word "***appə***" to "***ahpəʔ***", is as seen above $L_{\textbf{appə},\textbf{ahpəʔ}}(4,5) = 2$, given in the bottom right corner of the matrix.

The underlined numbers in the matrix represent the operations performed to the strings and are gained when moving from the end result backwards. Let us begin from the bottom right corner of the matrix. For simplicity let us assume we are changing the string "***appə***" to "***ahpəʔ***", which the symmetry of the calculations allows. Every step upwards represents the addition of the letter from the word "***ahpəʔ***", corresponding to that row, to the word "***appə***". Every step left represents the deletion of the letter from the word "***appə***", corresponding to that column, and every movement diagonally (if the number changes) the substitution of the letters from both words, corresponding to the specific row and column which intersect on the cell. For example, the upwards change from the underlined numbers 2 to 1 represent the addition of the letter "***ʔ***" to the word "***appə***" and the diagonal change from the underlined numbers 1 and 0 represent the substitution of the first "***p***" from the word "***appə***" with the letter "***h***". For the validity of counting the edit operations in Levenshtein distance, we note the method to be ambiguousness in some cases. Still, while the compared words are fairly short and specific changes reoccur, this method gives us the main changes made to the words.

2.1.1 Relative Levenshtein distance

While the popularity of the Levenshtein distance is justified by the simplicity of its definition, it is certainly not the only approach used for measuring the similarity of strings (see Ref. [10]). The main problem with the basic Levenshtein distance, when calculating the phonetical difference in strings, is that the measure does not consider the significance of the operations. Since longer words are more titled to changes, the Levenshtein distance in its basic form would be inclined to biased measurements, overestimating the average phonetical distance. This problem can be overcome through the computation of the relative Levenshtein distance [11], where one considers changes in words relative to the lengths of the words that are being edited,

$$L_{\Delta}(s,t) = \frac{L(s,t)}{\max(|s|,|t|)}.$$

Then $0 \leq L_{\Delta}(s,t) \leq 1$, where 0 accounts for the case $s = t$ and 1 for the case where s or t have to be edited in whole.

2.2 Historical Glottometry

Historical Glottometry can be viewed as a part of linguistic *Wave Theory*, developed in the framework of historical linguistics in order to provide a valid substitute of the familiar tree model. Differently from the tree model, the wave model considers the possibility that languages evolve while interacting with each-another, a scenario in which it would be inadequate to represent them via a cladistic approach.

The following analysis is based on the work of Siva Kalyan and Alexandre François [6], who have previously introduced the framework for Historical Glottometry. The article expounds on the definition of Historical Glottometry with applications to the languages of Vanuatu, an island nation off the coast of Australia, and discusses the explicit reasons why the tree model is in many cases inadequate for representing language genealogy, like in the case of Numic.

The wave model is based on the evolution of innovations, which François defines as changes, shared by all the languages of modern speakers [12]. The presence or absence of these innovations creates geographical borders known as *Isoglosses*. The basis of Historical Glottometry is the dialect-based comparison of the innovations, which can spread in erratic structures, resulting in different ranges of isoglosses. One way of representing dialect groupings is to account for the exclusively shared innovations and to plot the thickness of an isogloss line only dependent on the number of innovations, see Example 2.2 and Figure 2.1.

The approach, where groupings are represented based only on the exclusively shared innovations, is simple, but not always applicable, as we would get groups which are either hardly supported or are supported by an innovation occuring in two languages independently. To alleviate this problem, we need to introduce the relative strengths of groupings, which consider the number of supporting innovations, as well as the number of conflicting innovations i.e. the measure of cohesiveness between subgroups. We denote the number of supporting innovations with p and the number of conflicting innovations with q, where the conflicting innovations of a grouping, which are attested by at least one exclusively shared innovation, are those, including both a member of this group and one of another language. Mathematically, an isogloss Y is conflicting with a subgroup G if and only if

$$Y \cap G, Y \setminus G, G \setminus Y \neq \emptyset.$$

For calculating the cohesiveness of any given subgroup G we also need the total number of relevant innovations, which can be simply derived as the sum of supporting and conflicting innovations $p + q$. We follow [6] by defining:

Definition 2.2. Let G be a subgroup. The cohesiveness of the subgroup G is the number of supporting innovations in relation to the total number of relevant innovations. We denote the cohesiveness of the subgroup G as k_G. Therefore

$$k_G = \frac{p}{p+q}.$$

We now give an example of language groupings, where these intersecting subgroups occur.

Example 2.2. Let us consider a family consisting of three languages A, B and C, in which isoglosses define intersecting groupings. Let there be 12 exclusively shared innovations between A and B, 4 between A and C and 2 between B and C. The representation of these groups using only the exclusively shared innovations is seen in Figure 2.1.

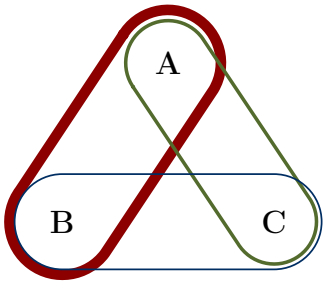


Figure 2.1: *Intersecting subgroups obtained from historical glottometry using exclusively shared innovations*

The *cohesiveness* of the subgroup AB would be calculated as

$$k_{AB} = \frac{12}{12+4+2} = \frac{2}{3} = 0.67. \tag{2.1}$$

The Interpretation of the cohesiveness of the subgroup AB would be that out of all the innovations that affect the subgroup AB, 67% of the innovations attest the cohesion of the subgroup, while 33% contradict it.

We could now use the relative measure of cohesiveness or the absolute number of exclusively shared innovations, which are mutually independent, to plot the isoglosses between language groupings. However, the two measures independently do not capture the full strength of a subgroup. We therefore combine them to calculate the true strength of any subgroup. This measure is called the *subgroupiness*, which we define, as in [6]:

Definition 2.3. Let G be a subgroup with the cohesiveness k_G and let ε be the number of exclusively shared innovations in the subgroup G. Then the

subgroupiness of G, which we denote with ς_G, is calculated as

$$\varsigma_G = \varepsilon \cdot k_G.$$

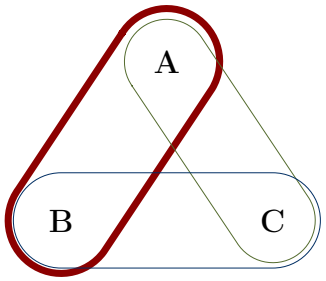
Let us consider the subgroup AB in Example 2.2, where A and B share 12 exclusive innovations and the cohesiveness $k_{AB} = \frac{2}{3}$, see Eq. (2.1). The subgroupiness of AB is thus

$$\varsigma_{AB} = \varepsilon \cdot k_{AB} = 12 \cdot \frac{2}{3} = 8.$$

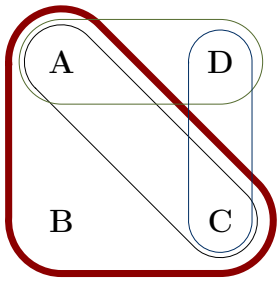
This gives us the strength of the subgroup AB. The subgroupiness of AC and BC can be calculated similarly,

$$\varsigma_{AC} = \varepsilon \cdot k_{AC} = 4 \cdot \frac{4}{12+4+2} = \frac{16}{18} = \frac{8}{9} \approx 0.88,$$
$$\varsigma_{BC} = \varepsilon \cdot k_{BC} = 2 \cdot \frac{2}{12+4+2} = \frac{4}{18} = \frac{2}{9} \approx 0.22.$$

For the visualization of the subgroups, we can now draw lines around the subgroups with a thickness that are directly proportional to the subgroupiness of every given subgroup. As a result, one obtains the *historical glottometric diagrams*. We give some simple examples of the historical glottometric diagrams for different isoglosses in Figure 2.2.



ς_{AB}=8, ς_{AC}=0.88, ς_{BC}=0.22



ς_{ABC}=8,ς_{AC}=0.88,ς_{AD}=0.68,ς_{DC}=0.28

Figure 2.2: *Examples of historical glottometric diagrams, subgroupiness given in the caption*

3 Analysis of the Numic database

In this section, we implement the approaches of Levenshtein distance and Historical Glottometry, introduced above, on the database of Numic languages, discussed in Sec. 3.1. The results obtained will shed some light on the spreading of Numic languages in the Great Basin. Consequently, the interpretation of the results of these analyses will be used to confirm or falsify existing hypotheses on the diffusion of Numic languages in the Great Basin.

3.1 The Numic database

The Numic database used in the present study was compiled by Daniele Dalmasso, Vittorio dell'Aquila and Jean Léo Léonard in 2012 and is based on Irvine Davis' [13] and David Iannucci's [14] original compilations from 1966 and 1973, respectively. This database was further elaborated by Flore Picard by homogenizing the transcriptions between the two original sources and by removing the affixes of entries to reveal the roots of the words in the hypothetical proto-language. The adjustments between the sources were necessary because both sources were using slightly different phonological theories, meaning that some sounds were represented differently between the entries, which would have tilted all comparisons of those entries towards measurement bias. Therefore, the obtained database (see appendix E) is more suitable for measuring phonetical differences. It consists of 290 cognates belonging to a hypothetical Proto-language, i.e., words with similar etymological origin, their respective counterparts (entries) in the eight largest Numic languages in the IPA transcription and the two sources marked 1 for Irvine Davis and 2 for David Iannucci. These cognates were constructed from the hypothesis of what the words in the dialects may have been by using phonological and grammatical rules, known to be valid in other languages. The working out of the cognates was done by using phonological and grammatical rules, known to be valid in other languages. Included in the database are also the translations of the cognates for a deeper understanding of the words we compare.

Since the database is not complete in the sense that not all languages are represented for every cognate, we give an overview of the number of cognates in each language in Table 3.1. This table gives us a clear overview of how many word comparisons between the dialects are carried out when calculating the Levenshtein distance.

Table 3.1: *Number of common cognates between different languages of the Numic database. The diagonal terms represent the total number of cognates in each language, see abbreviations in Figure 1.3. In total there are 290 cognates.*

Languages	**Co**	**Sh**	**Pn**	**NP**	**Mo**	**SP**	**Ka**	**Ut**
Co	206							
Sh	188	231						
Pn	67	70	86					
NP	161	174	59	198				
Mo	142	157	59	153	189			
SP	164	178	71	159	159	220		
Ka	45	48	27	53	55	54	56	
Ut	64	63	34	62	62	66	32	72

3.2 Numic Spread Hypothesis

The Numic Spread Hypothesis, first introduced by Sydney Lamb in 1958 [7], refers to the spread of Numic speakers from their homeland in the mountains of Sierra Nevada northeast into the Great Basin, which took place around AD 1000. The spread resulted in a fan-like distribution of the speakers into three groups, stretching across the Basin. Lamb argued that the division of the groups appeared while the Numic speakers still inhabited their homeland (Figure 1.1), whence a disturbance triggered the rapid north-eastern spread. It is known that the diffusion of language occurs through the movement of population [15, 16]: this expansion affected the whole Basin area leading to the disappearance of the Fremont culture, previously occupying the Basin, and bringing the Numic languages to their known groups. According to Lamb [7], the diffusion process produced evident language groupings in the three discernible Numic branches (Western, Central and Southern), with an

older language at the root of the branch and recent languages in the stem. The older languages are Mono, Panamint and Kawaiisu and their respective newer languages are Northern Paiute, Shoshoni and Southern Paiute/Ute, see Figure 1.3. The Comanche dialect is known to have originated from Shoshoneans [4]. We can therefore assume that it is connected to Panamint as was Shoshoni. Southern Paiute and Ute dialects have also been considered as one language [4], allowing us to acknowledge these dialects to be closely connected and descendants from Kawaiisu. In general, the hypothesis of a rapid and recent spread of the three Numic sub-branches comprehensively matches the wedge-shaped distribution of the Numic people [15, 17, 18]. We can thus expect the dialects in the geographical subgroups (Western, Central and Southern Numic) to be more similar in both phonological and morphological sense than those of neighbouring branches.

The rapid spread of Numic speakers has been a fundamental anomaly in the evolution of Numic dialects, while a starting cause for the expansion seems to be absent [15]. This interesting case has challenged researchers in different fields from linguistics to computer science, searching the missing factor. A crucial issue in the study of the Numic languages has been the dating of the Paiute-Shoshoni pottery in the Great Basin. In Ref. [16] a significant conformity between the archaeological and linguistic view was shown. Results demonstrate a north-eastern spread of Numic speaking groups about AD 1000 and confirm the coexistence of Fremont and Paiute-Shoshoni cultures in some areas. Research on the Numic diffusion process has also been done using computer simulations to reconstruct the competing populations in the Great Basin area [18]. The results from the simulations well reproduce the Numic spread, while the central-Numic branch is bounded by the neighbouring dialect branches to a stripe in the beginning of the expansion. This could lead to the merging of Southern Numic and Western Numic with the Central branch. Also, the results show an invasion of Shoshoneans at the expense of Northern Paiute and Ute, which may have also lead to a similarity of these dialects.

Clearly, these few examples do not provide a complete theory of the subject, but grasp only a part of the evolvment of the Numic languages. However, they demonstrate that in linguistics different disciplines and various approaches can be useful to extract information about linguistic diffusion and evolution.

3.3 Levenshtein distance

Here we calculate the Levenshtein matrix from compared cognates from our database. The results are visualized using networks as well as different tree-based diagrams, which are helpful for interpreting the results.

3.3.1 Levenshtein matrix

We denote s_i^k and s_j^k by the strings with the same semantic meaning k in languages i and j. We calculate the Levenshtein distance between strings s_i^k and s_j^k in languages i and j and denote it as

$$L_{i,j}^k := L_\Delta(s_i^k, s_j^k).$$

Furthermore, we denote by $M_{i,j}$ the number of pairs of strings s_i^k and s_j^k in languages i and j, sharing the same semantic meaning k. The values of $M_{i,j}$ are presented in Table 3.1. The Levenshtein matrix elements are computed as simple averages of all distances between languages i and j,

$$L_{i,j} = \frac{1}{M_{i,j}} \sum_{k=1}^{M_{i,j}} L_{i,j}^k,$$

which will be referred to as the average Levenshtein distance between languages i and j. We normalize the values $L_{i,j}$ by dividing all of the distances by the largest average Levenshtein distance $L_{\max} := \max(L_{i,j})$, rescaling $L_{i,j} \to L_{i,j}/L_{\max}$. By doing so, we obtain a matrix with values between 0 and 1, where 1 represents the largest Levenshtein difference between two languages and 0 their complete coincidence. The Levenshtein matrix calculated

from our database is given in Table 3.2. To make the matrix more informative, a heat map is also imbedded in the table, i.e., the values contained in a matrix are represented as colors. This table gives us an overview of the result for the Levenshtein distance. However, in order to have a better and deeper understanding of the results, it is useful to visualize the data in different ways.

Table 3.2: *Matrix of weighted Levenshtein distances for eight Numic languages, see abbreviations in Figure 1.3. Yellow corresponds to the smallest and red to the largest Levenshtein distance value, i.e., the more yellow the element, the more similar the two languages.*

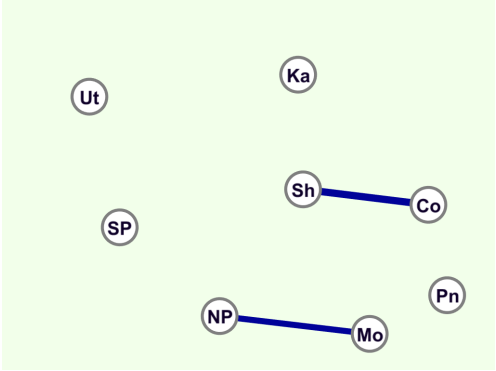
Languages	**Co**	**Sh**	**Pn**	**NP**	**Mo**	**SP**	**Ka**	**Ut**
Co	0	0.364	0.610	0.518	0.611	0.724	0.762	0.871
Sh	0.364	0	0.602	0.506	0.617	0.693	0.673	0.759
Pn	0.610	0.602	0	0.751	0.667	0.868	0.928	0.983
NP	0.518	0.506	0.751	0	0.470	0.725	0.666	0.820
Mo	0.611	0.617	0.667	0.470	0	0.783	0.723	1
SP	0.724	0.693	0.868	0.725	0.783	0	0.700	0.568
Ka	0.762	0.673	0.928	0.666	0.723	0.700	0	0.875
Ut	0.871	0.759	0.983	0.820	1	0.568	0.875	0

3.3.2 Threshold based network visualization

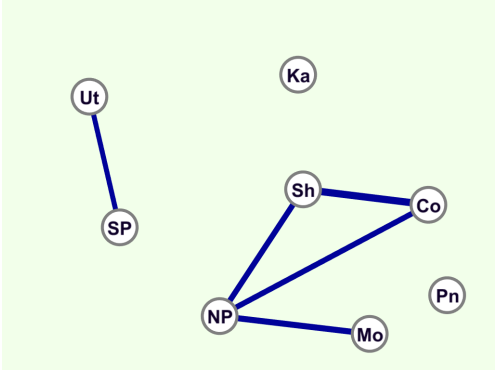
Here, the Levenshtein matrix is interpreted as the adjacency matrix of a corresponding network, which is visualized using a variable threshold $T \in [0, 1]$. Links between two languages i and j are visualized if and only if

$$L_{i,j} \leq T.$$

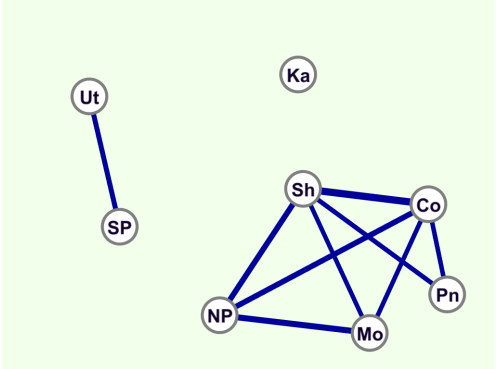
For $T = 0$, no links between nodes are present. As T increases, links begin to appear, starting from the strongest links. When reaching the value $T = 1$ even the weakest links are shown. It is easier to grasp the network structure if the links are depicted with a thickness corresponding to their strength, i.e., thicker links correspond to smaller Levenshtein distances. In this way, even when more links appear, the stronger and therefore more important links are still distinguishable.



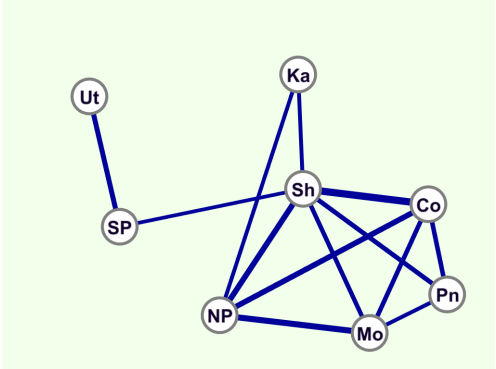
(a) $T = 0.5$



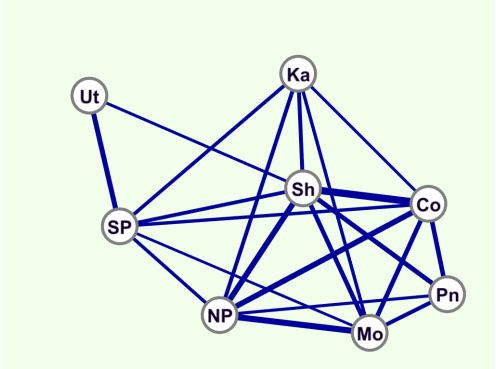
(b) $T = 0.6$



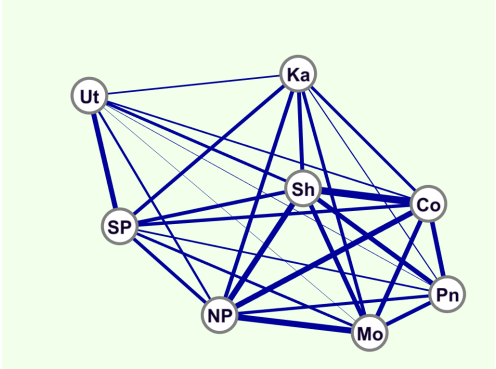
(c) $T = 0.65$



(d) $T = 0.7$



(e) $T = 0.8$



(f) $T = 1$

Figure 3.1: *Gephi graphs of the Levenshtein matrix of the Numic languages database with different thresholds T, see abbreviations in Figure 1.3*

The visualization was carried out using *Gephi* software and is presented in Figure 3.1 for various thresholds, from $T = 0.5$ to $T = 1$. For $T = 0.5$ (Figure 3.1a) the first links between **[Sh-Co]** and **[NP-Mo]** have appeared, showing that these language pairs are the most closely related ones in the phonological sense.

For $T = 0.6$ (Figure 3.1b), also the **[Ut-SP]** link has appeared. The appearance of the strong **[Sh-Co]** and **[Ut-SP]** links is consistent with the fact that Shoshoni and Comanche, as well as Ute and Southern Paiute, have been considered also as single dialect groups. The group **[NP-Mo]** is known to form the Western Numic languages group. While the similarity between Ute and Southern Paiute as well as between Northern Paiute and Mono seems natural also due to their geographical vicinity, the similarity between Shoshoni and Comanche might be unexpected when considering solely the geographical isolation of Comanche seen from Figure 1.3. However, it can be understood known that the Comanche and Shoshone people used to form in the past a group and the Comanche emerged as a distinct group only shortly before 1700, when they broke off from the Shoshone people [4].

For $T = 0.6$, besides the link **[Ut-SP]** , also the unexpected links between **[Sh-NP]** and **[Co-NP]** form. The first one is rather natural due to the long geographical continuity between the corresponding areas. The second one arises from the link **[Sh-NP]** due to the already discussed fact that the Shoshone and Comanche people formed in the past the same group.

For $T = 0.65$ (Figure 3.1c) the link between Mono and Shoshoni/Comanche and between Panamint and Shoshoni/Comanche appear. The link between Mono and Panamint will form for $T = 0.7$ (see Figure 3.1d).

The link between Northern Paiute and Shoshoni/Comanche, which is considerably stronger than the link between Panamint and Shoshoni/Comanche, as well as the link between Mono and Shoshoni/Comanche put the thus far recognized division into Central and Western Numic languages (see Figure 1.2) under question. Instead, our analysis suggest that the Central and Western

Numic languages form one single group. A rather surprising result is the weak link between Mono and Panamint languages that geographically are neighbors.

An unexpected result in Figure 3.1d is also that the Kawaiisu is more strongly connected to Northern Paiute and Shoshoni than to Ute and Southern Paiute, while the dialect is believed to be apart of the [**Ut-SP**] branch. Furthermore, for $T = 0.7$ the link between Shoshoni and Southern Paiute (but not Kawaiisu) appears. These results do not align with the former hypothesis of the Numic expansion and implies more contact between the Kawaiisu and Northern Paiute languages. Due to the recent nature of the Numic spread, this contact was likely before the expansion took place. Instead, we speculate that the Numic spread to areas of Southern Paiute was from Shoshoni not from Kawaiisu as formerly disputed.

We also notice that Shoshoni and Comanche occupy a central position in the network. This can be put into correspondence with the central location of Shoshoni (actually, due to the strong link [**Sh-Co**], of the original group Shoshoni/Comanche) in the Great Basin, with respect to the other languages. Such an emerging network structure may be caused by the Central Numic branch being compressed at the start of the Numic spread by neighboring dialects, as speculated by Young and Bettinger [18]. With Panamint having weaker connections to other dialects than Shoshoni and Comanche, we may consider that the interactions between [**Sh-Co**] to other dialects happened during the diffusion process of Numic and not before.

The relative isolation of Kawaiisu with respect to the other languages (even to Ute and Southern Paiute, with which it is usually grouped in the Southern Numic group) is to be noticed. The first links of Kawaiisu appear only for a threshold $T = 0.7$ (therefore with a rather low strength), when all the languages become connected in a single component network. The linked languages are Northern Paiute and Shoshoni, rather than the languages of the Western Numic branch (Ute and Southern Paiute), in which Kawaiisu is usually placed. This fact and other peculiarities, such as an atypical consonant

inventory, may be due to the original Kawaiisu homeland being bordered by non-Numic Uto-Aztecan languages [2].

3.3.3 Nearest neighbour map

In order to give an intuitive picture of the most similar languages, we plot the *nearest neighbour network* in Figure 3.2 using the same Levenshtein matrix computed above. In this network each node (language) is connected only to the respective language with the lowest Levenshtein distance. Node locations are arbitrarily set in some central position of the corresponding language area. The nearest neighbour network provides an overview of the structure of the language network and is easier to interpret through some additional information. Furthermore, embedding the nearest neighbour network in the geographical map accounts for the topography of the region, which can have a great influence on the spread of languages.

From the map in Figure 3.2, we see the expected groupings of Central and Western Numic branches with the addition of the established [**Ut-SP**] chain. The only exception is the Kawaiisu dialect, which links closest to Northern Paiute.

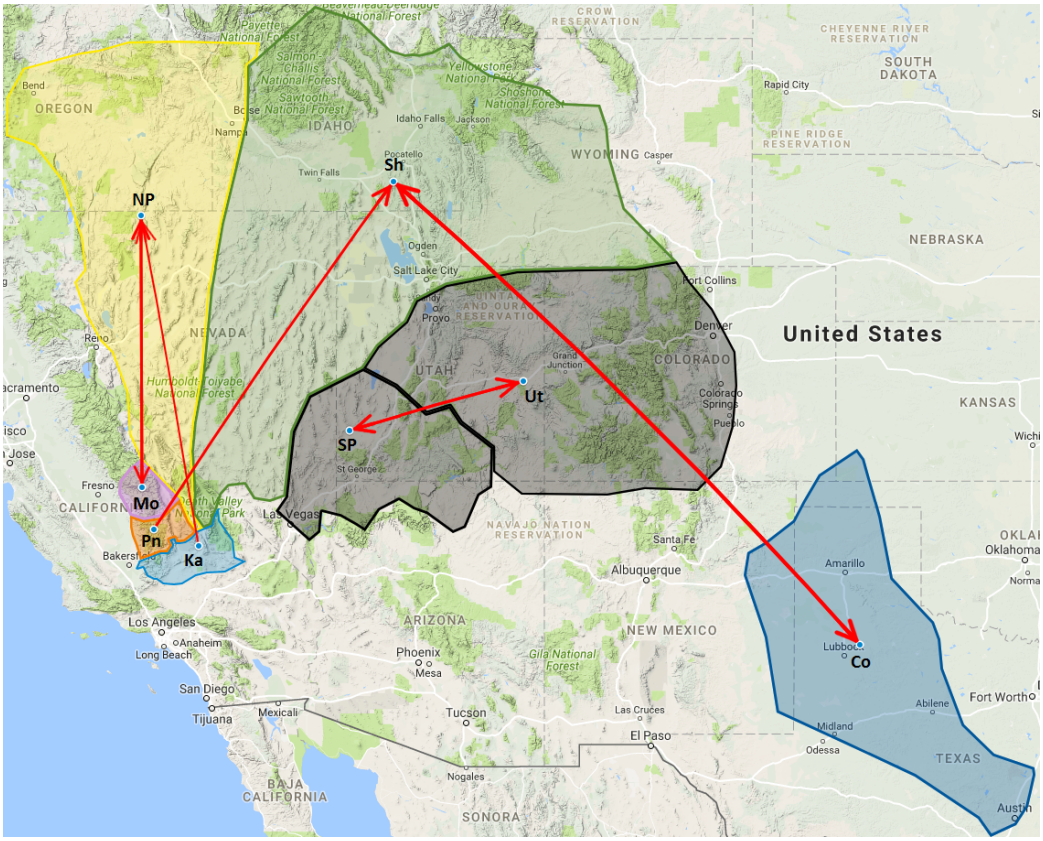


Figure 3.2: *Nearest neighbour map from Levenshtein Distance, see abbreviations in Figure 1.3*

3.3.4 Dendrogram

A dendrogram is a tree diagram, used for graphical representation of the relations between the nodes of a network produced by a hierarchical clustering. They are often used in computational biology to illustrate the clustering of genes or samples. The nodes (in the present case the languages) are connected to each other with U-shaped lines, determined by the Levenshtein matrix. There are many approaches, from means and medians to more complex methods, available for making a hierarchical clustering. Here we use the average linkage clustering, where we find all possible pairwise distances for nodes in different clusters from which we calculate the average between-

clusters distance, following [19]:

$$D(X,Y) = \frac{1}{|X||Y|} \sum_{x \in X} \sum_{y \in Y} L_{x,y},$$

where X and Y are the clusters of nodes, $|X|$ and $|Y|$ show the cardinality of clusters, which are calculated as the average of all Levenshtein distances $L_{x,y}$ betwen languages $x \in X$ and $y \in Y$. In doing so, we obtain the dendrogram of the eight languages under study, depicted in Figure 3.3. Such a dendrogram provides an alternative, easily interpretable, scheme of the language network.

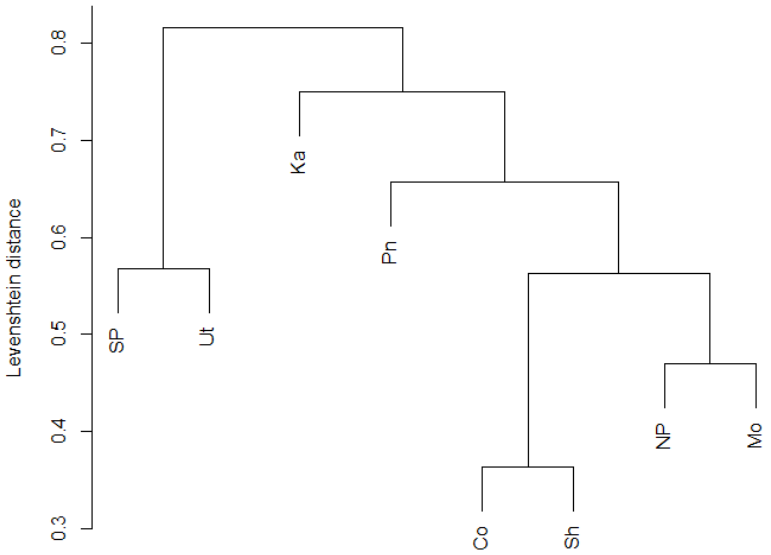


Figure 3.3: *Dendrogram obtained from the Levenshtein matrix of the Numic languages, see abbreviations in Figure 1.3*

From the dendrogram we can see three strong subgroups [**Co-Sh**], [**NP-Mo**] and [**Ut-SP**], which implies the coherence of three dialect branches. Interestingly, the Panamint dialect linkage is present in both the [**Co-Sh**] and [**NP-Mo**] groups, which may mean that the Panamint dialect also blended with northern Mono and Northern Paiute as well as with Comanche/Shoshoni. As seen from the Gephi graphs (Figure 3.1), the dendrogram also indicates the distant nature of Kawaiisu, plotting it closer to the [**NP-Mo-Pn-Co-Sh**] chain than to [**SP-Ut**]. The dendrogram shown was obtained using the

inbuilt functions of the statistical computing language R (see the R documentation [20] for details).

3.3.5 Minimum spanning tree

A minimum spanning tree is a subset of an undirected graph with weighted edges. The graph is produced with N nodes, which are connected by $N-1$ edges, such that the total distance of the edges is minimized. We plot the minimum spanning tree using the "*spantree*" function in R [20]. The function creates the minimum spanning tree using the Prim's algorithm, which starts from an arbitrary root, each step extending by one edge to an isolated node, so that the edge contributes the least amount of weight to the tree [21]. As a result, we obtain a tree graph with all nodes connected by at least one edge, where the length of the edge represents the average Levenshtein distance.

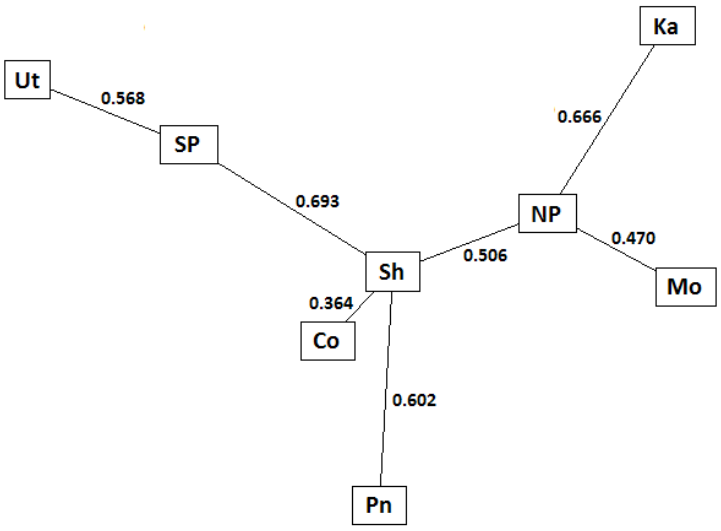


Figure 3.4: *Minimum spanning tree, see abbreviations in Figure 1.3*

The minimum spanning tree of Numic languages is given in Figure 3.4. In general, this minimum spanning tree is consistent with the information provided by the other analyses made above. The tree reveals how the Shoshoni and Northern Paiute languages are centered among the other languages. The tree also shows the connection between Kawaiisu and Northern Paiute (Levenshtein distance 0.666), as did the nearest neighbour network. Then the further question remains of the unbound character of Kawaiisu and Panamint, which are geographically neighbors.

3.3.6 Multidimensional scaling

We now look at the classical multidimensional scaling, also known as principal coordinates analysis, a method for visualizing the degree of similarity between objects. Multidimensional scaling can be viewed as an expansion on the factor analysis, where the matrix of correspondences can be an arbitrary matrix of similarities. In our case, the matrix in question is the Levenshtein distance matrix. Classical multidimensional scaling aims to represent this matrix as a geometrical map, where a set of coordinates corresponds to each point in a multidimensional space [22]. With the multidimensional scaling, we seek to embed $X = (x_1, x_2, \ldots, x_N)$ into a D-dimensional real space $\mathbb{R}^D$, with $D < N$, by preserving distances, that is to find $X = (x_1, x_2, \ldots, x_N)$, so that

$$\|x_i - x_j\| \approx L_{i,j} \quad i, j = 1, \ldots, N,$$

where $L_{i,j}$ is the Levenshtein distance between dialects i and j and N the total number of dialects. In our case we choose a dimension $D = 2$, which means that the representation of the multidimensional space is made on a plane.

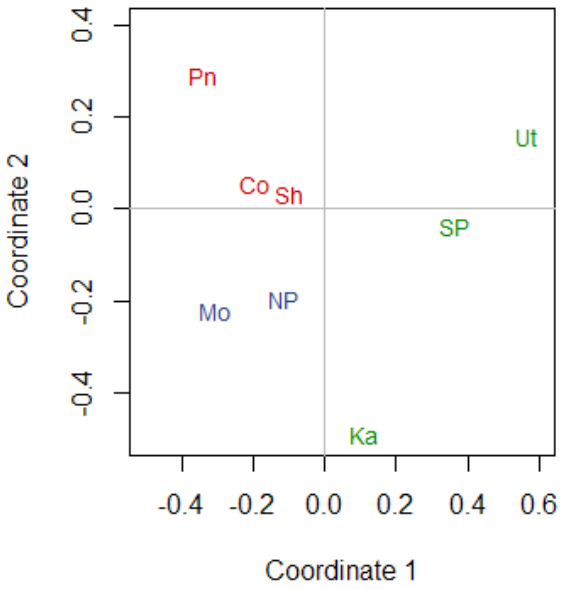


Figure 3.5: *Multidimensional scaling, see abbreviations in Figure 1.3*

We represent the multidimensional space, constructed with our eight Numic languages on a plane, as seen in Figure 3.5. The Numic languages are color-coded according to the traditional divisions in Western (blue), Central (red), and Southern (green) languages. As seen from Figure 3.5, the multidimensional scaling results match in general the traditional divisions, with the exceptions of Panamint (that appears to be further from Shoshoni than Northern Paiute) and Kawaiisu (which is the most isolated language). This corresponds very well to the results implied by the minimum spanning tree (see Figure 3.4) as well as the other analyses presented above.

3.4 Historical Glottometry

Next, we can test the potential of Historical Glottometry method by applying it to the Numic languages database. As explained above, the basis of Historical Glottometry is the comparison of innovations, which in our case are phonological innovations. We can quantify these innovations by calculating the Levenshtein distances. By accounting for the changes made to every word pair, as explained in Example 2.1, we can attain specific changes made

to words in all eight Numic languages. To reduce the risk of having biased results, as the Levenshtein distance was ambiguous in some cases in accounting for changes, changes which occur less than five times are discarded. Only changes occurring more than five times are taken into considerations and are called below *regular innovations*.

The Levenshtein distance method was used to identify 32 phonological innovations. Among those 32 innovations, only 30 subgroups shared any exclusive innovations, which would give biased result when calculating the subgroupiness. To alleviate this problem, we need to introduce irregular innovations in addition to the regular ones.

Irregular innovations are innovations found in only one word or a small group of words, as opposed to regular innovations which occur in every case of a specific context, therefore a more empiric result compared to the regular innovations. An example of the irregular innovations can be seen in Table 3.3, where the languages of same colors in each row represent the specific irregular innovation.

Table 3.3: *Example of irregular innovations, see abbreviations in Figure 1.3*

	Co	**Ka**	**Mo**	**NP**	**Pn**	**Sh**	**SP**	**Ut**
piya	pia?	piya	piya	pia	piya	pia	pia	pia
pəhta	pəda	pəda	pəta	pəta		pəda	pəda	
suwah	sua		suwa		suwa	suwa	sua	səa
səhə	səhə	səə	səhə	sə		soho	səə	
tahma	tahmani		tawano	tamanu	tahwani	tahmani	tamana	tamana
taman	tama	tawa	tawa	tama	tama	taŋ^{w}a	taŋ^{w}a	
tape	tabe		tabe	taba		tabe	taba	
...	...	...	...	...	...	...	...	...

When accounting for the irregular innovations, we find 239 innovations in addition to the 32 regular innovations, calculated from the Levenshtein distance. For the final phonological innovations, we put the regular and irregular innovations together, obtaining 271 innovations, from which we construct a table

of occurrences of binary values, where 1 accounts for the presence of an innovation in a language, and 0 for the absence of an innovation (see Table 3.4).

Table 3.4: *Example of innovations of regular and irergular innovations, see abbreviations in Figure 1.3*

	Co	**Ka**	**Mo**	**NP**	**Pn**	**Sh**	**SP**	**Ut**
aa ↔ a	1	1	1	1	0	0	1	1
ə ↔ a	1	1	0	0	0	0	1	0
k ↔ q	0	0	1	0	0	0	1	0
hk ↔ k	0	1	0	1	1	1	0	0
hm ↔ m	1	0	1	1	1	1	1	1
irr1	1	0	0	0	1	0	0	1
irr2	1	0	1	0	1	0	0	0
...	...	...	...	...	...	...	...	...

From the table of occurrences, we calculate the cohesiveness k (see Def. 2.2) and exclusively shared innovations ε for all dialect subgroups from where we can simply derive the subgroupiness ς (see Def. 2.3) for all sets of the Numic languages. The ten strongest subgroups of the Numic languages are given in Table 3.5.

Table 3.5: *10 of the strongest subgroups in the Numic dialects, see abbreviations in Figure 1.3*

Subgroup	ε	k	subgroupiness (ς)
Co-Sh	27	0.458	12.363
Mo-NP	22	0.410	9.011
SP-Ut	13	0.323	4.194
NP-Sh	12	0.340	4.075
Co-NP	9	0.361	3.249
Co-NP-Sh	8	0.240	1.923
Sh-SP	6	0.299	1.791
Co-Mo	7	0.252	1.767
Co-Mo-NP-Sh	9	0.160	1.440
Mo-SP	6	0.235	1.408

Finally, in Figure 3.6 we plot the glottometric diagram for the Numic dialect system.

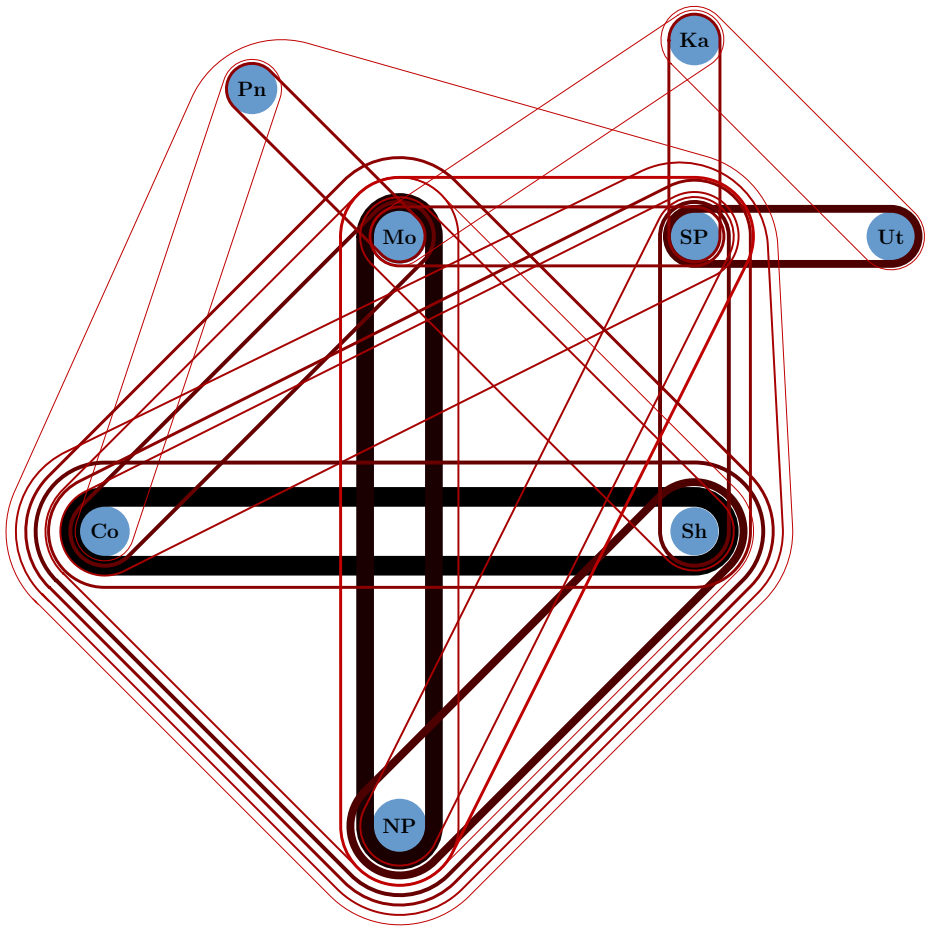


Figure 3.6: *Glottometric diagram of the Numic languages - 24 strongest links, see abbreviations in Figure 1.3*

We can see the targeted glottometric diagram of the Numic languages in Figure 3.6. For better interpretation of the diagram, the isoglosses plotted are, in addition to being proportional to the strength of the subgroups, colour coded, where lighter isoglosses represent weaker subgroups. The diagram shows four strong dialect pairs, the formulated [**Co-Sh**], [**NP-Mo**], [**SP-Ut**] as well as a more interesting [**Sh-NP**] subgroup. As a whole, we can see a system of four strongly connected dialects [**Co-Sh-NP-Mo**] and the aforementioned [**SP-Ut**] grouping with Kawaiisu and Panamint connected

only by weaker isoglosses to other languages. The subgroup [**Co-Mo-Sh-NP**] also seems to link to the Southern Paiute, as it has many links to this dialect (see Figure 3.1d). The subgroup [**Co-Sh-SP**] or the slightly weaker [**Co-Sh-NP-SP**] may proceed from the [**Co-Sh**] dialect group being weighed down between the Northern and Southern Paiute as speculated by Young and Bettinger [18]. The interesting cases of the diagram are these of Kawaiisu and Panamint, the latter of which should have a strong connection with [**Co-Sh**] but is more strongly connected to the Mono dialect. This [**Pn-Mo**] connection can be explained however with how the dialects are geographically situated, being two neighbouring dialects. The Kawaiisu dialect, in addition to being linked to the Souhtern Paiute, has linked to the Mono dialect as was Panamint, which is supported in some way to the Levenshtein distance result, where we saw the Kawaiisu connected most strongly to the Northern Paiute, which is in the same subgroup with Kawaiisu.

When comparing the results of the Levenshtein distance and Historical Glottometry, we see, besides the [**Co-Sh**], [**SP-Ut**] and [**NP-Mo**] subgroups, the confirmation of a strong dialect foursome, consisting of Comanche, Shoshoni, Northern Paiute and Mono dialects. This grouping may be the cause of greater interactions between the four languages and is not entirely contradicting to the Numic spread hypothesis. The other homogenous result between the two methods is the aforementioned Kawaiisu and Mono connection as well as the centered nature of the [**Sh-Co**] dialect group, which, as mentioned above, could be caused by the Central Numic branch being compressed at the start of the Numic spread by neighbouring dialects [18]. In addition to the confirmed results, the two methods gave us different results between the Panamint and Kawaiisu dialects. From the Historical Glottometry, we can see Kawaiisu being connected to Southern Paiute as expected and not to [**Co-Sh**] as seen from the method of Levenshtein distance (see Figure 3.1). The Panamint dialect is the reverse case, where the Levenshtein distance method produces results in accordance with the Numic spread hypothesis, but does not strongly support the connection to the Mono dialect, seen in the glottometric diagram (Figure 3.1). The merit of using both the Levenshtein

distance method and the Historical Glottometry method is the solidity it gives to the confirmation or falsification of the tested hypotheses. Also the two methods give a multiplicity of viewpoints to detect the finer details, which is essential when dealing with complexity.

4 Conclusion

This current thesis explored the relationships and connections of the Numic languages with the help of complex systems methods. We used mathematical tools, applied to the phonological Numic languages database, to confirm or falsify hypotheses on the diffusion of the Numic languages. The two methods (Levenshtein distance and Historical Glottometry) gave us results which were in compliance with the main aspects of the Numic spread hypothesis: The spread, which resulted in a fan-like distribution of the Numic speakers into three groups (see Figure 1.3), as well as interesting results which are to be noticed on their own. We divide the conclusion of the thesis into two parts, focusing on the two methods seperately.

Here we give the results gained from the Levenshtein distance method. The method showed strong connections between Shoshoni and Comanche, Mono and Northern Paiute and Ute and Southern Paiute, language groups, which are consistent with regard to the Numic spread hypothesis. In addition to the conformations of the Numic spread hypothesis, the results of the Levenshtein distance method show strong bonds between the Central and Western Numic languages, which suggest Central Numic and Western Numic to form one single language group. The Shoshone and Comanche languages, having many strong links to other languages, ocupied the central position in networks and graphs which may be caused by the Central Numic branch being compressed as the start of the Numic spread [18] and helped with Shoshoni languages having a central location in the distribution of Numic languages. We also observed that Kawaiisu is a rather different language with respect to other languages, having weak links to Northern Paiute and Shoshoni, rather than the languages of the Western Numic branch (Ute and Southern Paiute), in which Kawaiisu is usually placed. The fact that Kawaiisu is more distant, may be due to the original Kawaiisu homeland being bordered by non-Numic Uto-Aztecan languages.

The Historical Glottomety method also provided the three language groups of Shoshoni and Comanche, Mono and Northern Paiute and Ute and Southern Paiute strongly grouped. The method gave notice to a unified system of four languages from the Central and Western Numic groups, leaving out Panamint. The Panamint language was connected to Comanche and Mono by weaker isoglosses, which is to be expected, when seeing the geographical vicinity of those languages. Kawaiisu had been grouped to the Southern Numic branch (Ute and Southern Paiute languages) as expected from the Numic spread hypothesis and also having an interesting connection to the Mono dialect, as Kawaiisu and Mono are not neighboring languages. The other larger group of languages is the Comanche, Shoshoni, Northern Paiute and Southern Paiute which may also proceed from the Comanche/Shohoni being weighed down between the Northern and Souhtern Paiute as speculated by Young and Bettinger [18].

In essence, both methods look at linguistic complexity through dynamic, rather than linear approaches. This gave a multiplicity of viewponits to detect the finer detail in the Numic language system, which is essential when dealing with complexity. The difference in the finer details of the methods confirms that the Numic system is in fact a complex system and should be viewed as such. The two methods confirmed the grouping of Central and Western Numic branches in addition to the three strongest and formulated Shoshoni and Comanche, Mono and Northern Paiute and Ute and Southern Paiute groups. Also the centrality of Comanche and Shoshoni in the Numic language system. The results also showed deviation of Kawaiisu and Panamint languages with regard to the hypotheses on the diffusion of Numic languages, which suggests the subject to be investigated further. Overall, the work and results of the present thesis show the contribution mathematics can make to linguistics and especially dialectology.

References

[1] Grishman R. *Computational Linguistics: An Introduction.* Cambridge University Press; 1986. p. 4–9.

[2] Shaul D, Ortman S. *A Prehistory of Western North America: The Impact of Uto-Aztecan Languages.* Albuquerque: University of New Mexico Press; 2014.

[3] Rowe B, Levine D. *A Concise Introduction to Linguistics.* Taylor & Francis; 2015. p. 341.

[4] Grayson D. *The Great Basin: A Natural Prehistory.* California: University of California Press; 2011. 11–40

[5] Léonard J, Heinsalu E, Patriarca M, Sharma K, Chakraborti A. *Patterns of Linguistic Diffusion in Space and Time: The Case of Mazatec.* In Abergel F (ed.), Econophysics and Sociophysics: Recent Progress and Future Directions. Springer; 2017.

[6] Kalyan S, François A. forthcoming. *Freeing the Comparative Method from the tree model: A framework for Historical Glottometry.* In Kikusawa R, Reid L (eds.), Let's Talk about trees: tackling problems in representing phylogenic relationships among languages, 2013.

[7] Lamb S. Linguistic Prehistory in the Great Basin. *International Journal of American Linguistics.* 1958, vol. 24, no.2, 95–100.

[8] Steward J. Linguistic Distributions and Political Groups of the Great Basin Shoshoneans. *American Anthropologist*, 1937, vol. 39, no. 4, 625–634.

[9] Hämäläinen P. *The Comanche Empire: Lamar series in western history.* Yale University Press; 2008.

[10] Apostolico A, Galil Z. *Pattern Matching Algorithms.* Oxford University Press; 1997. p. 123–125.

[11] Nerbonne J, Heeringa W. *Measuring Dialect Distance Phonetically.* In Coleman J (ed.), Workshop on Computational Phonology, Special Interest Group of the Association for Computational Linguistics, Madrid, 1997, 11–18.

[12] François A. *Trees, Waves and Linkages: Models of Language Diversification.* In Bowern, Claire; Evans, Bethwyn. The Routledge Handbook of Historical Linguistics, Routledge, 2014, 161–189.

[13] Davis I. Numic Consonantal Correspondences. *International Journal of American Linguistics.* 1966, vol. 32, no.2, 124–140.

[14] Iannucci D. *Numic historical phonology.* Unpublished PhD dissertation, Cornell University, 1973

[15] Bettinger R, Baumhoff M. The Numic Spread: Great Basin Cultures in Competition. *American Antiquity*, 1982, vol. 47, no. 3, 485–503.

[16] Madsen D. Dating Paiute-Shoshoni Expansion in the Great Basin. *American Antiquity.* 1975, vol. 40, no.1, 82–86.

[17] Hopkins N. Great Basin Prehistory and Uto-Aztecan. *American Antiquity*, 1965, vol. 31, no. 1, 48–60.

[18] Young D, Bettinger R. The Numic Spread: A Computer Simulation. *American Antiquity*, 1992, vol. 57, no. 1, 85–99.

[19] Polanski A, Kimmel M. *Bioinformatics* Springer Berlin Heidelberg; 2007. p. 193.

[20] R Core Team. *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing, Vienna, Austria, 2015.

[21] Cormen T. *Introduction to Algorithms, 3rd Edition.* Cambridge: MIT Press; 2009. p. 567–569.

[22] Cox T, Cox M. *Multidimensional Scaling, Second Edition.* CRC Press; 2000. p. 31–34.

Appendices

A *Matlab* code for Levenshtein distance

Calculation of the relative Levenshtein distance from the Numic database with accounting for the changes made to the words.

Calculation of the relative Levenshtein distance

```
function value = relative_LD(s, t)
n = length(s);
m = length(t);
if min(n,m) == 0
    value = max(n,m);
else
    A = zeros(m+1, n+1);
    A(1,1:n+1) = 0:n; A(1:m+1,1) = 0:m;
    for j = 1:n
        for i = 1:m
            if t(i) == s(j)
                indicator = 0;
            else
                indicator = 1;
            end
            A(i+1,j+1) = min([A(i,j+1) + 1, A(i+1,j) + 1, A(i,j)
                + indicator]);
        end
    end
    value = A(m+1,n+1)/max(n,m);
end
```

Accounting for the changes made to the words

```
% A - Levenshtein distance matrix
% n - length of string s
% m - length of string t
B = []; i = m; j = n; p = 1;
while (i >= 1 || j >= 1)
    if i == 0
        b = [s(j), {'_'}]; B = [B;b]; j=j-1;
    elseif j == 0
        b = [t(i), {'_'}]; B = [B;b]; i=i-1;
    else
        if A(i+1,j+1) == A(i,j)
            [~, idx] = min([A(i+1,j+1),A(i,j+1),A(i+1,j)]);
            if idx == 1
                if p == max(m,n)
                    b = [{'_'}, {'_'}]; B = [B;b];
                end
                p=p+1; i=i-1; j=j-1;
            elseif idx == 2
                b = [t(i), {'_'}]; B = [B;b]; i=i-1;
            elseif idx == 3
                b = [s(j), {'_'}]; B = [B;b]; j=j-1;
            end
        else
            [~, idx] = min([A(i+1,j+1),A(i,j+1),A(i+1,j)]);
            if idx == 1
                b = [{s(j)}, {t(i)}]; B = [B;b]; i=i-1; j=j-1;
            elseif idx == 2
                b = [t(i), {'_'}]; B = [B;b]; i=i-1;
            elseif idx == 3
                b = [s(j), {'_'}]; B = [B;b]; j=j-1;
            end
        end
    end
end
```

B *Matlab* code for Levenshtein distance matrix

Calculation of the Levenshtein distance matrix for all eight Numic languages.

```
% LevDistances - matrix of relative Levenshtein distances between
    all cognates
% Unique ID points
a = unique(LevDistances.first_id_point);
b = unique(LevDistances.second_id_point);
x = unique([a;b]);
% All combinations of id points
comb = cellstr(nmultichoosek(x,2));
[m,~] = size(comb);
B = {'first_id_point', 'second_id_point', 'Mean_Levenshtein'};
for i = 1:m
    % c1 and c2 are the elements of the i-th combination
    c1 = comb(i,1); c2 = comb(i,2);
    % Dist is the Levenshtein distances where first id point is
        c1 and second is c2 or first is c2 and second is c1
    Dist = LevDistances.Levenshtein(LevDistances.first_id_point==
        c1 & LevDistances.second_id_point==c2 | LevDistances.
        first_id_point==c2 & LevDistances.second_id_point==c1,:);
    [n,~] = size(Dist);
    mean = sum(Dist)/n;
    b = [c1,c2,mean];
    B = [B; b];
end
```

C *Matlab* code for Cohesiveness

Calculation of the cohesiveness and exclusively shared innovations for subgroups of two languages. Cohesiveness and exclusively shared innovations for subgroups with more than two languages are computed analogically.

```
% B is the matrix of innovations for all eight Numic languages
C2 = {'Dialect1', 'Dialect2', 'Cohesiveness', 'Exclusive'};
B = nchoosek(Dialects,2);
[m,~] = size(B);
for j = 1:m
    Co = dataset2cell(Coh); Co(1,:) = [];
    % a - Dialect 1
    % b - Dialect 2
    a = B{j,1}; b = B{j,2}; [~,m1] = size(Co);
    row1 = Co(a, 2:m1); row2 = Co(b, 2:m1);
    a1 = sort([a,b],'descend');
    for u = a1
        Co(u,:) = [];
    end
    Co(:,1) = []; rows = [row1;row2];
    % p - supporting innovations
    % q - conflicting innovations
    % p2 - exclusively shared innovations
    p=0; q=0; p2=0;
    for k = 1:(m1-1)
        if row1{k} == 1 && row2{k} == 1
            p = p+1;
            if any(1==cell2mat(Co(:,k)))==0
                p2 = p2+1;
            end
        elseif (any(1==cell2mat(rows(:,k)))) && (any(0==cell2mat(
            rows(:,k)))) && (any(1==cell2mat(Co(:,k))))
            q = q+1;
        end
    end
    c1 = num2cell([a,b,p/(p+q),p2]); C2 = [C2; c1];
end
```

D *R*-code for visualization

R code for dendrogram, minimum spanning tree and multidimensional scaling from the Levenshtein distance matrix.

Plotting of the dendrogram and minimum spanning tree

```
# tmp - Levenshtein distance matrix for Numic languages.
tmp <- matrix(c(0,  0.364,  0.610,  0.518,  0.611,  0.724,
    0.762,  0.871,
                0.364,  0,  0.602,  0.506,  0.617,  0.693,
                    0.673,  0.759,
                0.610,  0.602,  0,  0.751,  0.667,  0.868,
                    0.928,  0.983,
                0.518,  0.506,  0.751,  0,  0.470,  0.725,
                    0.666,  0.820,
                0.611,  0.617,  0.667,  0.470,  0,  0.783,
                    0.723,  1,
                0.724,  0.693,  0.868,  0.725,  0.783,  0,
                    0.700,  0.568,
                0.762,  0.673,  0.928,  0.666,  0.723,  0.700,
                    0,  0.875,
                0.871,  0.759,  0.983,  0.820,  1,  0.568,
                    0.875,  0),
              nr=8, dimnames=list(c('Co','Sh','Pn','NP','Mo','SP'
                ,'Ka','Ut'),c('Co','Sh','Pn','NP','Mo','SP','Ka
                ','Ut')))
d <- as.dist(tmp)
hc <- hclust(d, method="average")
plot(hc ,hang=0.1 ,ylab = "Levenshtein distance",ylim = c(0.3,1),
     main = '')

tree <- spantree(d)
plot(tree, cex = 0.6, type = "t", main='')
```

Plotting of the multidimensional scaling

```
fit <- cmdscale(d,eig=TRUE, k=2)
x <- fit$points[,1]
y <- fit$points[,2]
plot(x, y, xlab="Coordinate 1", ylab="Coordinate 2",
      main="", type="n", ylim=c(-0.5,0.4), xlim=c(-0.5,0.6))
names1 <- c('Co','Sh','Pn');names2 <- c('NP','Mo');names3 <- c('
    SP','Ka','Ut')
text(x[1:3], y[1:3], pos = 1, labels = names1, cex=0.9, col='red3
    ')
text(x[4:5], y[4:5], pos = 1, labels = names2, cex=0.9, col='
    royalblue4')
text(x[6:8], y[6:8], pos = 1, labels = names3, cex=0.9, col='
    green4')
abline(h=0,v=0,col="gray75")
```

E Complete modified Numic languages database

Source: Irvine Davis 1966 and David Iannucci 1973, compilation by Daniele Dalmasso, Vittorio dell'Aquila and Jean Léo Léonard 2012, modified by Flore Picard.

Co - Comanche, **Pn** - Panamint, **Sh** - Shoshoni, **Ka** - Kawaiisu, **SP** - Southern Paiute, **Ut** - Ute, **Mo** - Mono, **NP** - Northern Paiute.

id question	id point	Source	Root	Entry	Translation
1	Co	1	ahna	ahna	armpit/wing
1	Sh	1	ahna	ahna	armpit/wing
1	SP	1	ahna	aŋa	armpit/wing
2	Co	1	ahpə	ahpəʔ	father
2	Pn	1	ahpə	appə	father
2	Sh	1	ahpə	appə	father
2	SP	1	ahpə	ahpə	father
3	Co	2	ahta	ahda	jaw/upper cheek
3	Ka	2	ahta	ata	jaw/upper cheek
3	Mo	2	ahta	ada	jaw/upper cheek
3	NP	2	ahta	ada	jaw/upper cheek
3	Sh	2	ahta	ahda	jaw/upper cheek
3	SP	2	ahta	ata	jaw/upper cheek
4	Ka	2	amah	awa	ribs
4	Mo	1	amah	awa	ribs
4	NP	1	amah	ama	ribs
4	Sh	1	amah	ama	ribs
4	SP	2	amah	aŋʷa	ribs
4	Ut	2	amah	aŋʷa	ribs

5	Co	1	ani	ani	mosquito/fly/ant
5	Mo	1	ani	ani	mosquito/fly/ant
5	NP	1	ani	ani	mosquito/fly/ant
5	Sh	1	ani	ani	mosquito/fly/ant
5	SP	1	ani	aŋi	mosquito/fly/ant
6	Co	1	awah	aaʔ	horn
6	Ka	2	awah	aa	horn
6	Mo	1	awah	awa	horn
6	NP	1	awah	aa	horn
6	Pn	1	awah	awa	horn
6	Sh	1	awah	aa	horn
6	SP	1	awah	aa	horn
6	Ut	2	awah	aa	horn
7	Co	1	ekon	eko	tongue
7	Ka	2	ekon	egu	tongue
7	Mo	2	ekon	ego	tongue
7	NP	2	ekon	igo	tongue
7	Pn	1	ekon	eko	tongue
7	Sh	1	ekon	eku	tongue
7	SP	2	ekon	agu	tongue
7	Ut	2	ekon	agʷo	tongue
8	Co	1	ekʷi	əkʷi	smell (transitive)
8	Ka	2	ekʷi	ugʷi	smell (transitive)
8	Mo	2	ekʷi	egʷi	smell (transitive)
8	NP	2	ekʷi	igʷi	smell (transitive)
8	Pn	1	ekʷi	ekʷi	smell (transitive)
8	Sh	2	ekʷi	əgʷi	smell (transitive)
8	SP	2	ekʷi	ugʷi	smell (transitive)
8	Ut	2	ekʷi	ugʷi	smell (transitive)
9	Co	1	eŋkah	eka	red
9	Ka	2	eŋkah	aga	red
9	Mo	2	eŋkah	aqa	red
9	Pn	1	eŋkah	aŋka	red
9	Sh	1	eŋkah	eŋka	red
9	SP	2	eŋkah	aŋqa	red
9	Ut	2	eŋkah	aka	red
10	Co	1	etə	etə	bow/gun
10	Mo	2	etə	edəh	bow/gun
10	NP	2	etə	adə	bow/gun
10	Sh	1	etə	eti	bow/gun

10	SP	1	etə	acə	bow/gun
11	Mo	1	ooh	ooh	pebbles/round object
11	SP	1	ooh	oh	pebbles/round object
12	Co	1	ohah	oha	yellow
12	Mo	1	ohah	oha	yellow
12	NP	2	ohah	oa	yellow
12	Sh	2	ohah	oha	yellow
12	SP	2	ohah	ua	yellow
12	Ut	2	ohah	əa	yellow
13	Ka	2	oho	ohoo	bone
13	Mo	1	oho	oho	bone
13	NP	1	oho	oho	bone
13	SP	2	oho	uu	bone
13	Ut	2	oho	dəə	bone
14	Co	1	ohni	ohni	cough (verb)
14	Mo	1	ohni	ohi	cough (verb)
14	NP	1	ohni	ohi	cough (verb)
14	Sh	1	ohni	ohni	cough (verb)
15	Co	1	oŋaaʔa	ohnaaʔ	baby/child/young (of animals)
15	Mo	1	oŋaaʔa	owaaʔa	baby/child/young (of animals)
15	NP	1	oŋaaʔa	oŋaʔa	baby/child/young (of animals)
15	Sh	2	oŋaaʔa	ohna	baby/child/young (of animals)
15	SP	2	oŋaaʔa	əŋaa	baby/child/young (of animals)
16	Co	1	oŋa	ona	salt
16	Mo	2	oŋa	oma	salt
16	NP	1	oŋa	oŋa	salt
16	Pn	1	oŋa	oŋŋa	salt
16	Sh	1	oŋa	ona	salt
16	SP	2	oŋa	ua	salt
16	Ut	2	oŋa	əa	salt
17	Sh	1	uʔu	uʔu	fart (verb)
17	SP	1	uʔu	uu	fart (verb)
18	Co	1	usən	u	that
18	Mo	1	usən	uhu	that
18	NP	1	usən	uh	that
18	Sh	1	usən	u	that
18	SP	1	usən	u	that
19	Co	1	isa	isa	lie (noun, verb)
19	Sh	1	isa	isa	lie (noun, verb)
19	SP	1	isa	isə	lie (noun, verb)

20	Mo	1	ica	ica	coyote
20	NP	1	ica	ica	coyote
20	Pn	1	ica	issa	coyote
20	Sh	1	ica	ica	coyote
21	Co	1	isən	i	this
21	Mo	1	isən	ihi	this
21	NP	1	isən	ih	this
21	Sh	1	isən	i	this
21	SP	1	isən	i	this
21	Ut	2	isən	i	this
22	Co	1	əh	ənə	you (sg.)
22	Mo	1	əh	ə	you (sg.)
22	NP	1	əh	ə	you (sg.)
22	Sh	1	əh	ə	you (sg.)
22	SP	1	əh	i	you (sg.)
22	Ut	2	əh	ə	you (sg.)
23	Co	1	əhma	əma	rain (verb)
23	NP	1	əhma	uhma	rain (verb)
23	Pn	1	əhma	əmmah	rain (verb)
23	Sh	2	əhma	əŋʷa	rain (verb)
23	SP	2	əhma	uŋʷa	rain (verb)
23	Ut	2	əhma	uŋʷa	rain (verb)
24	Co	2	əhpəʔi	əhpəi	sleep (verb)
24	Mo	1	əhpəʔi	əwi	sleep (verb)
24	NP	1	əhpəʔi	əʔwi	sleep (verb)
24	Pn	1	əhpəʔi	əppəʔi	sleep (verb)
24	Sh	2	əhpəʔi	əpəi	sleep (verb)
24	SP	2	əhpəʔi	apəi	sleep (verb)
24	Ut	2	əhpəʔi	pəi	sleep (verb)
25	Mo	1	ətə	ətə	long/tall
25	NP	1	ətə	ətə	long/tall
26	Co	2	ətəh	ədə	(be) hot
26	Mo	2	ətəh	ədə	(be) hot
26	NP	2	ətəh	ədə	(be) hot
26	Pn	1	ətəh	ətə	(be) hot
26	Sh	2	ətəh	ədə	(be) hot
26	SP	2	ətəh	du	(be) hot
26	Ut	2	ətəh	də	(be) hot
27	Co	1	hacincih	haici	friend
27	NP	2	hacincih	haiʒi	friend

27	Sh	1	hacincih	hainci	friend
28	Co	1	hahaʔa	haa	yes
28	Mo	1	hahaʔa	həəhəʔə	yes
28	NP	1	hahaʔa	ahaa	yes
28	Pn	1	hahaʔa	hahaʔa	yes
28	Sh	1	hahaʔa	haa	yes
29	Co	1	hahni	hani	do/make/prepare
29	NP	2	hahni	hani	do/make/prepare
29	Sh	2	hahni	hani	do/make/prepare
29	Ut	2	hahni	oni	do/make/prepare
30	Co	1	hake	haka	who/which/what
30	Mo	1	hake	haqe	who/which/what
30	NP	2	hake	haga	who/which/what
30	Sh	2	hake	haga	who/which/what
30	SP	2	hake	aga	who/which/what
30	Ut	2	hake	aga	who/which/what
31	Co	2	hapi	habi	lie down
31	Mo	1	hapi	hapi	lie down
31	NP	2	hapi	habi	lie down
31	Sh	2	hapi	habi	lie down
31	Ut	2	hapi	abi	lie down
32	Mo	1	heewiiʔi	heewii	dove
32	Sh	1	heewiiʔi	hewi	dove
33	Mo	1	hoʔnopi	hoʔnopi	bat (animal)
33	Sh	1	hoʔnopi	honopi	bat (animal)
34	Co	2	hota	hoda	dig
34	NP	1	hota	hohna	dig
34	Sh	2	hota	hoda	dig
34	SP	2	hota	uda	dig
35	Co	1	huuh	huuh	wood/stick/tree
35	NP	1	huuh	huu	wood/stick/tree
35	Pn	1	huuh	huppi	wood/stick/tree
35	Sh	1	huuh	huu	wood/stick/tree
35	SP	1	huuh	uu	wood/stick/tree
36	Co	1	huhkumpəh	huhkupə	dust
36	Pn	1	huhkumpəh	hukkumpə	dust
36	Sh	1	huhkumpəh	hukkumpəh	dust
36	SP	1	huhkumpəh	uhkumpu	dust
37	Pn	1	kukʷi	hukʷi	grass
37	SP	1	kukʷi	ukʷi	grass

38	Co	1	hupiya	hubiyaa	sing/song
38	Mo	1	hupiya	hupiya	sing/song
38	Pn	1	hupiya	hupiya	sing/song
38	Sh	1	hupiya	hupia	sing/song
38	SP	1	hupiya	upia	sing/song
39	Co	1	hii	hi	what/who
39	Mo	1	hii	hii	what/who
39	NP	1	hii	hii	what/who
39	Sh	1	hii	hii	what/who
39	SP	2	hii	i	what/who
40	Co	1	hipi	hibi	drink (verb)
40	Mo	2	hipi	hibi	drink (verb)
40	NP	2	hipi	hibi	drink (verb)
40	Pn	1	hipi	hipi	drink (verb)
40	Sh	2	hipi	hibi	drink (verb)
40	SP	2	hipi	ibi	drink (verb)
40	Ut	2	hipi	ibi	drink (verb)
41	Mo	1	həhkʷa	həhkʷa	blow (of wind)
41	NP	1	həhkʷa	həhkʷa	blow (of wind)
42	Sh	1	həhci	həcci	spit (noun)
42	SP	1	həhci	kəhci	spit (noun)
43	Sh	1	hənan	hunan	badger
43	SP	1	hənan	ənan	badger
44	Co	1	həpa	həka	be cool
44	Mo	1	həpa	həhca	be cool
44	NP	1	həpa	hapa	be cool
44	Sh	1	həpa	həpa	be cool
44	SP	1	həpa	apa	be cool
45	Mo	1	həpi	həpi	woman
45	Sh	1	həpi	həpi	woman
46	Mo	1	həya	həya	trap (verb)
46	Sh	1	həya	həa	trap (verb)
47	Co	1	kawa	ka	rat
47	Ka	2	kawa	kaa	rat
47	Mo	1	kawa	qa	rat
47	NP	1	kawa	ka	rat
47	Sh	1	kawa	kaa	rat
47	SP	2	kawa	qaa	rat
47	Ut	2	kawa	ka	rat
48	Mo	1	kaahkaa	kaahkaa	cry of the quail

48	SP	1	kaahkaa	kahka	cry of the quail
49	NP	2	kaipa	kaiba	mountain
49	SP	2	kaipa	qaiba	mountain
49	Ut	2	kaipa	kaabə	mountain
50	Co	1	kahma	kama	(have a) taste
50	Mo	2	kahma	qama	(have a) taste
50	NP	2	kahma	kama	(have a) taste
50	Sh	1	kahma	kamma	(have a) taste
50	SP	2	kahma	qama	(have a) taste
50	Ut	2	kahma	kama	(have a) taste
51	Mo	2	kahmə	qamə	jackrabbit
51	NP	2	kahmə	kamə	jackrabbit
51	Pn	1	kahmə	kammo	jackrabbit
51	Sh	2	kahmə	kamə	jackrabbit
51	SP	2	kahmə	qamə	jackrabbit
52	Co	1	kahni	kahni	house
52	Pn	1	kahni	kahni	house
52	Sh	1	kahni	kahni	house
52	SP	2	kahni	qani	house
52	Ut	2	kahni	kani	house
53	Co	1	kakuʔu	kakuʔ	grandmother
53	Sh	2	kakuʔu	kagu	grandmother
53	SP	2	kakuʔu	kagu	grandmother
53	Ut	2	kakuʔu	kagu	grandmother
54	Co	1	kasa	kasa	wing/feather
54	Mo	2	kasa	qasa	wing/feather
54	NP	2	kasa	kasa	wing/feather
54	Pn	1	kasa	kassa	wing/feather
54	Sh	1	kasa	kasa	wing/feather
54	SP	1	kasa	kasa	wing/feather
54	Ut	2	kasa	kəsi	wing/feather
55	Co	2	katə	kadə	sit(down) - dur.
55	Ka	2	katə	kada	sit(down) - dur.
55	Mo	1	katə	qatə	sit(down) - dur.
55	NP	2	katə	katə	sit(down) - dur.
55	Pn	1	katə	kattə	sit(down) - dur.
55	Sh	2	katə	kadə	sit(down) - dur.
55	SP	2	katə	qadə	sit(down) - dur.
55	Ut	2	katə	kadə	sit(down) - dur.
56	NP	1	kacun	kacu	top/end

56	Sh	1	kacun	kacun	top/end
57	Co	1	ke	ke	no/not/nothing
57	Mo	1	ke	qa	no/not/nothing
57	NP	1	ke	kai	no/not/nothing
57	Pn	1	ke	kee	no/not/nothing
57	Sh	1	ke	ke	no/not/nothing
57	SP	2	ke	qa	no/not/nothing
57	Ut	2	ke	ka	no/not/nothing
58	NP	1	koonih	konih	bend/bent
58	Sh	1	koonih	koonih	bend/bent
58	SP	1	koonih	kohmi	bend/bent
59	Co	2	koʔi	kooi	kill/die/sleep
59	Mo	1	koʔi	qoi	kill/die/sleep
59	NP	1	koʔi	koʔi	kill/die/sleep
59	Sh	1	koʔi	koi	kill/die/sleep
59	SP	2	koʔi	quʔu	kill/die/sleep
60	Co	1	kohpi	koba	break/cut
60	Ka	2	kohpi	kopa	break/cut
60	Mo	1	kohpi	qopa	break/cut
60	NP	1	kohpi	kohpa	break/cut
60	Sh	1	kohpi	kəpa	break/cut
60	SP	2	kohpi	qupu	break/cut
60	Ut	2	kohpi	kapə	break/cut
61	Co	1	kohtoo	kohtoo	make/set a fire
61	NP	1	kohtoo	kuhtuu	make/set a fire
61	Sh	1	kohtoo	kottoo	make/set a fire
62	Co	1	kope	kobe	face
62	Ka	2	kope	kobi	face
62	Mo	2	kope	qobe	face
62	NP	2	kope	koba	face
62	Pn	1	kope	kope	face
62	Sh	2	kope	kobe	face
62	SP	2	kope	quba	face
62	Ut	2	kope	koba	face
63	Mo	1	ku	kuhsa	bag
63	SP	1	ku	kuna	bag
64	Co	1	kuh	kuʔ	fire/heat (instr. pref.)
64	Mo	1	kuh	kuh	fire/heat (instr. pref.)
64	NP	1	kuh	ku	fire/heat (instr. pref.)
64	Sh	1	kuh	ku	fire/heat (instr. pref.)

64	SP	1	kuh	kuh	fire/heat (instr. pref.)
65	Co	1	kuhsih	kusi	ashes
65	NP	2	kuhsih	kutusi	ashes
65	Sh	1	kuhsih	kusi	ashes
65	SP	1	kuhsih	kuhcah	ashes
66	Co	2	kuhma	kuma	husband/male
66	Mo	1	kuhma	kuwa	husband/male
66	NP	1	kuhma	kuma	husband/male
66	Pn	1	kuhma	kuwa	husband/male
66	Sh	1	kuhma	kuhma	husband/male
66	SP	2	kuhma	kuma	husband/male
67	Ka	2	huhta	kudo	neck
67	Mo	2	huhta	kuta	neck
67	NP	2	huhta	kuta	neck
67	Pn	1	huhta	kura	neck
67	SP	2	huhta	kuda	neck
67	Ut	2	huhta	kuda	neck
68	Co	1	kuhcun	kuhcuʔ	buffalo/cow
68	NP	2	kuhcun	kucu	buffalo/cow
68	Sh	2	kuhcun	kʷicu	buffalo/cow
68	SP	1	kuhcun	kuhcun	buffalo/cow
68	Ut	2	kuhcun	kəcupukə	buffalo/cow
69	Co	1	kunah	kuna	wood/fire/firewood
69	Ka	2	kunah	kuna	wood/fire/firewood
69	Mo	1	kunah	kuna	wood/fire/firewood
69	NP	1	kunah	kuna	wood/fire/firewood
69	Pn	1	kunah	kuna	wood/fire/firewood
69	Sh	1	kunah	kuna	wood/fire/firewood
69	SP	1	kunah	kuna	wood/fire/firewood
70	Co	1	kiihpə	kiipə	elbow
70	Mo	2	kiihpə	kiibə	elbow
70	Sh	2	kiihpə	kiipə	elbow
70	SP	2	kiihpə	kipə	elbow
70	Ut	2	kiihpə	kiipə	elbow
71	Co	1	kihma	kima	come
71	Mo	2	kihma	kima	come
71	NP	2	kihma	kima	come
71	Pn	1	kihma	kimma	come
71	Sh	2	kihma	kima	come
72	Co	2	kəh	kə	with teeth/by biting (instr. pref.)

72	Ka	2	kəh	kiʔi	with teeth/by biting (instr. pref.)
72	Mo	1	kəh	kəh	with teeth/by biting (instr. pref.)
72	NP	2	kəh	kəi	with teeth/by biting (instr. pref.)
72	Pn	1	kəh	kə	with teeth/by biting (instr. pref.)
72	Sh	1	kəh	kə	with teeth/by biting (instr. pref.)
72	SP	2	kəh	kəʔə	with teeth/by biting (instr. pref.)
72	Ut	2	kəh	kiʔi	with teeth/by biting (instr. pref.)
73	Mo	1	kəhkə	kəhkə	foot
73	NP	1	kəhkə	kəhkə	foot
74	Co	1	kəmaa	kəma	(sharp) edge
74	Mo	2	kəmaa	kəwa	(sharp) edge
74	NP	2	kəmaa	kəma	(sharp) edge
74	Sh	2	kəmaa	kəwəʔə	(sharp) edge
74	SP	2	kəmaa	kəŋʷa	(sharp) edge
75	Mo	1	kənuʔu	kənu	(paternal) grandfather
75	NP	1	kənuʔu	kənu	(paternal) grandfather
75	Sh	1	kənuʔu	kənu	(paternal) grandfather
75	SP	1	kənuʔu	kunu	(paternal) grandfather
76	Co	1	kəŋka	kəka	onion
76	NP	2	kəŋka	kəga	onion
76	Sh	1	kəŋka	kəŋka	onion
77	Co	1	kʷahti	kʷəhti	shoot
77	Mo	1	kʷahti	qʷahti	shoot
77	NP	2	kʷahti	kʷati	shoot
77	Sh	2	kʷahti	kʷəti	shoot
78	Co	1	kʷana	kʷana	smell(y)
78	Mo	1	kʷana	qʷana	smell(y)
78	NP	1	kʷana	kʷana	smell(y)
78	Sh	1	kʷana	kʷana	smell(y)
78	SP	1	kʷana	kʷana	smell(y)
78	Ut	2	kʷana	kona	smell(y)
79	Co	1	kʷasu	kʷasuʔu	dress/shirt
79	NP	2	kʷasu	kʷasə	dress/shirt
79	Sh	2	kʷasu	kʷasu	dress/shirt
80	Co	1	kʷasə	kʷasə	cook(ed)/ripe
80	Mo	2	kʷasə	qʷasə	cook(ed)/ripe
80	NP	2	kʷasə	kʷasə	cook(ed)/ripe
80	Pn	1	kʷasə	kʷassəʔə	cook(ed)/ripe
80	Sh	1	kʷasə	kʷasə	cook(ed)/ripe
80	SP	2	kʷasə	kʷasə	cook(ed)/ripe

81	Co	1	kʷesi	kʷasi	tail
81	Ka	2	kʷesi	kosi	tail
81	Mo	2	kʷesi	qʷaʒi	tail
81	NP	2	kʷesi	kʷasi	tail
81	Pn	1	kʷesi	kʷassi	tail
81	Sh	1	kʷesi	kʷesi	tail
81	SP	1	kʷesi	kʷasi	tail
81	Ut	2	kʷesi	kʷasi	tail
82	Co	1	kʷii	kʷi	say
82	Mo	1	kʷii	kʷii	say
82	NP	1	kʷii	kʷi	say
82	Sh	1	kʷii	kʷi	say
83	Co	1	kʷiih	kʷii	smoke (n.)
83	Ka	2	kʷiih	kʷihi	smoke (n.)
83	Mo	2	kʷiih	kuhih	smoke (n.)
83	NP	2	kʷiih	kʷi	smoke (n.)
83	Pn	1	kʷiih	kʷii	smoke (n.)
83	Sh	1	kʷiih	kʷii	smoke (n.)
83	SP	2	kʷiih	kʷii	smoke (n.)
83	Ut	2	kʷiih	kʷii	smoke (n.)
84	Mo	1	kʷi	kʷi	vagina
84	Pn	1	kʷi	kʷi	vagina
85	Co	1	kʷi	kʷih	north/cold
85	Mo	1	kʷi	kʷi	north/cold
85	Sh	1	kʷi	kʷi	north/cold
86	Mo	1	kʷiʔnaaʔa	kʷiʔnaa	eagle/large bird
86	NP	1	kʷiʔnaaʔa	kʷiʔnaa	eagle/large bird
86	Sh	1	kʷiʔnaaʔa	kʷinaa	eagle/large bird
86	SP	1	kʷiʔnaaʔa	kʷana	eagle/large bird
87	Co	1	kʷihtah	kʷita	excrement/defecate
87	Mo	2	kʷihtah	kʷida	excrement/defecate
87	NP	2	kʷihtah	kʷida	excrement/defecate
87	Pn	1	kʷihtah	kʷittah	excrement/defecate
87	Sh	2	kʷihtah	kʷida	excrement/defecate
87	SP	2	kʷihtah	kʷiča	excrement/defecate
88	Co	1	kʷəha	kʷəhə	catch/take
88	Mo	1	kʷəha	wəə	catch/take
88	NP	2	kʷəha	kʷə	catch/take
88	Sh	1	kʷəha	kʷəhə	catch/take
88	SP	1	kʷəha	kʷəə	catch/take

89	Co	1	maanaaŋkʷah	manakʷə	far
89	Pn	1	maanaaŋkʷah	manaakʷa	far
89	Sh	1	maanaaŋkʷah	maanaŋkʷah	far
90	Co	2	mah	ma	1. (instr. pref.) hand 2. (n.) hand
90	Mo	2	mah	ma	1. (instr. pref.) hand 2. (n.) hand
90	NP	1	mah	ma	1. (instr. pref.) hand 2. (n.) hand
90	Sh	2	mah	ma	1. (instr. pref.) hand 2. (n.) hand
90	SP	2	mah	ma	1. (instr. pref.) hand 2. (n.) hand
90	Ut	2	mah	ma	1. (instr. pref.) hand 2. (n.) hand
91	Co	1	mahka	maka	feed/give
91	Mo	2	mahka	maqa	feed/give
91	NP	2	mahka	maka	feed/give
91	Sh	2	mahka	maga	feed/give
91	SP	2	mahka	maga	feed/give
91	Ut	2	mahka	maga	feed/give
92	Ka	2	manəkih	mənaga	five
92	Mo	2	manəkih	manəgi	five
92	NP	2	manəkih	manigi	five
92	Pn	1	manəkih	manəkih	five
92	Sh	2	manəkih	manaigi	five
92	SP	2	manəkih	manigi	five
92	Ut	2	manəkih	manəgi	five
93	Mo	1	mayəh	mai	find/become/be/do
93	NP	2	mayəh	mai	find/become/be/do
93	SP	2	mayəh	mai	find/become/be/do
94	Sh	1	mona	mona	son-in-law
94	SP	1	mona	muna	son-in-law
95	NP	1	mosui	mosui	mustache
95	SP	1	mosui	mošoi	mustache
96	Co	1	mocon	moco	beard/facial hair
96	NP	2	mocon	mozui	beard/facial hair
96	Pn	1	mocon	moco	beard/facial hair
96	Sh	2	mocon	moco	beard/facial hair
96	SP	2	mocon	muŋcu	beard/facial hair
97	Co	1	muhuh	mu	owl
97	Ka	2	muhuh	muhu	owl
97	Mo	1	muhuh	muhu	owl
97	NP	1	muhuh	muhuʔu	owl
97	Sh	1	muhuh	mu	owl
97	SP	2	muhuh	muu	owl

97	Ut	2	muhuh	moo	owl
98	Co	1	muih	mui	fly (insect)
98	Ka	2	muih	muu	fly (insect)
98	Mo	1	muih	mui	fly (insect)
98	NP	1	muih	mui	fly (insect)
98	Sh	2	muih	bui	fly (insect)
98	SP	2	muih	muu	fly (insect)
98	Ut	2	muih	mo	fly (insect)
99	Co	1	mukih	muci	sharp point
99	Mo	1	mukih	mukih	sharp point
99	NP	1	mukih	muku	sharp point
99	Sh	1	mukih	muci	sharp point
100	Co	1	mupih	mubi	(n.) nose
100	Mo	2	mupih	mubi	(n.) nose
100	NP	2	mupih	mubi	(n.) nose
100	Pn	1	mupih	mupi	(n.) nose
100	Sh	2	mupih	mubi	(n.) nose
100	SP	2	mupih	mubi	(n.) nose
100	Ut	2	mupih	məbə	(n.) nose
101	Co	1	miʔa	miʔa	go/walk
101	Mo	1	miʔa	miya	go/walk
101	NP	2	miʔa	mia	go/walk
101	Pn	1	miʔa	miya	go/walk
101	Sh	2	miʔa	miʔa	go/walk
101	SP	1	miʔa	mia	go/walk
102	Co	1	məʔah	məa	moon/month
102	Ka	2	məʔah	məa	moon/month
102	Mo	1	məʔah	məʔa	moon/month
102	NP	2	məʔah	muha	moon/month
102	Pn	1	məʔah	məah	moon/month
102	Sh	2	məʔah	məaʔa	moon/month
102	SP	2	məʔah	məa	moon/month
102	Ut	2	məʔah	maa	moon/month
103	Mo	1	məyən	məyə	gopher
103	SP	1	məyən	məyən	gopher
104	Mo	1	hapi	naahbai	six (cf. three)
104	NP	1	hapi	napahi	six (cf. three)
104	Co	1	naahpahi	naahbai	six (cf. three)
104	Ka	2	naahpahi	nabaha	six (cf. three)
104	Pn	1	naahpahi	naapai	six (cf. three)

104	Sh	2	naahpahi	naabai	six (cf. three)
104	SP	1	naahpahi	nabai	six (cf. three)
104	Ut	2	naahpahi	nabai	six (cf. three)
105	Co	2	naci	nai	girl(young woman)
105	Pn	1	naci	nawi	girl(young woman)
105	Sh	1	naci	nai	girl(young woman)
105	SP	2	naci	naʔai	girl(young woman)
106	Mo	1	naʔi	nai	nurn (intr.)
106	SP	1	naʔi	naʔai	nurn (intr.)
107	Co	2	nampe	nape	foot/(lower) leg
107	Pn	1	nampe	nampe	foot/(lower) leg
107	Sh	2	nampe	naŋpe	foot/(lower) leg
107	SP	2	nampe	naŋpa	foot/(lower) leg
108	Co	1	nanah	nanah	(grown) man/grow
108	Mo	2	nanah	nana	(grown) man/grow
108	NP	2	nanah	nana	(grown) man/grow
108	Sh	1	nanah	nanah	(grown) man/grow
108	SP	2	nanah	nana	(grown) man/grow
109	Co	2	naŋka	naka	ear (hear)
109	Mo	2	naŋka	naqa	ear (hear)
109	NP	2	naŋka	naka	ear (hear)
109	Sh	2	naŋka	neŋki	ear (hear)
109	SP	2	naŋka	naŋqa	ear (hear)
110	Co	1	naŋkʷa	nakʷə	direction/side
110	SP	1	naŋkʷa	naŋkʷah	direction/side
111	Co	2	natənoʔo	nadənooʔ	saddle
111	NP	2	natənoʔo	nadənoʔo	saddle
111	Sh	2	natənoʔo	nadənoʔo	saddle
112	Mo	1	noʔo	noo	carry (on back)
112	Sh	1	noʔo	noo	carry (on back)
112	SP	1	noʔo	noo	carry (on back)
113	Co	2	noʔyv	noʔyaa	boil (vb.)
113	SP	2	noʔyv	nuyu	boil (vb.)
114	Mo	1	nohko	noqo	roast meat
115	Co	2	noyo	no	egg/house/dwelling
115	Mo	2	noyo	no	egg/house/dwelling
115	NP	2	noyo	no	egg/house/dwelling
115	Pn	1	noyo	noyoʔo	egg/house/dwelling
115	Sh	2	noyo	no	egg/house/dwelling
115	SP	2	noyo	nu	egg/house/dwelling

116	Sh	1	nuhkin	nukki	run (off, away)/move/flow
117	Co	2	nica	nia	call/name (vb.), voice/speech (pref.)
117	Mo	1	nica	niya	call/name (vb.), voice/speech (pref.)
117	NP	1	nica	niʔa	call/name (vb.), voice/speech (pref.)
117	Sh	2	nica	nania	call/name (vb.), voice/speech (pref.)
117	SP	2	nica	nia	call/name (vb.), voice/speech (pref.)
118	Co	1	nə	nə	I
118	Mo	2	nə	nəə	I
118	NP	2	nə	nə	I
118	Sh	2	nə	nəʔə	I
118	SP	2	nə	nəʔə	I
119	Co	1	nəe	nəe	wind/blow (of wind)
119	Sh	1	nəe	nəa	wind/blow (of wind)
119	SP	1	nəe	nəa	wind/blow (of wind)
120	Co	2	nəhka	nəhka	dance
120	Ka	2	nəhka	nəka	dance
120	Mo	1	nəhka	nəhka	dance
120	NP	2	nəhka	nəka	dance
120	Sh	2	nəhka	nəka	dance
120	SP	2	nəhka	nihqa	dance
121	Co	1	nəhmi	nənə	we (excl.--cf. incl. 'we')
121	Mo	2	nəhmi	nəə	we (excl.--cf. incl. 'we')
121	NP	1	nəhmi	nəhmi	we (excl.--cf. incl. 'we')
121	Sh	2	nəhmi	nəməʔə	we (excl.--cf. incl. 'we')
121	SP	2	nəhmi	nəmə	we (excl.--cf. incl. 'we')
122	Co	1	nəhmə	nəmə	person/Indian
122	Ka	2	nəhmə	nəwə	person/Indian
122	Mo	2	nəhmə	nəəmə	person/Indian
122	NP	2	nəhmə	nəmə	person/Indian
122	Sh	1	nəhmə	nəmə	person/Indian
122	SP	2	nəhmə	nəŋʷə	person/Indian
123	Co	1	nəmi	nəmi	walk/wander/live
123	NP	1	nəmi	nəmi	walk/wander/live
123	Sh	1	nəmi	nəmi	walk/wander/live
124	Co	2	nəmən	nəmə	liver
124	Mo	2	nəmən	nəwə	liver
124	NP	1	nəmən	nəmə	liver
124	Sh	1	nəmən	nəmi	liver
124	SP	1	nəmən	nəmə	liver
125	Co	1	nəŋahpəh	nənapə	chest

125	NP	2	nəŋahpəh	nəŋabə	chest
125	Sh	1	nəŋahpəh	nəŋkapə	chest
126	Ka	2	nəpa	nəbo	snow (n.)
126	Mo	2	nəpa	nəba	snow (n.)
126	NP	2	nəpa	nəba	snow (n.)
126	SP	2	nəpa	nəba	snow (n.)
127	Co	1	paa	paa	water
127	Mo	2	paa	paya	water
127	NP	1	paa	pa	water
127	Pn	1	paa	paa	water
127	SP	2	paa	paa	water
128	Co	2	paəhpi	pəəhpi	blood (cf. 'water')
128	Mo	2	paəhpi	paapi	blood (cf. 'water')
128	NP	2	paəhpi	pəpə	blood (cf. 'water')
128	Sh	1	paəhpi	pəəpi	blood (cf. 'water')
128	SP	2	paəhpi	paəpi	blood (cf. 'water')
129	Co	1	paʔa	paʔa	high/long/tall
129	Mo	1	paʔa	paʔa	high/long/tall
129	NP	1	paʔa	paʔa	high/long/tall
129	Sh	1	paʔa	paʔa	high/long/tall
129	SP	2	paʔa	paʔa	high/long/tall
129	Ut	2	paʔa	pato	high/long/tall
130	Co	1	paʔi	paʔi	hit
130	Sh	1	paʔi	pai	hit
131	Co	2	pahapi	pahabə	swim (cf. 'water')
131	Mo	2	pahapi	pahabi	swim (cf. 'water')
131	NP	2	pahapi	pahabi	swim (cf. 'water')
131	Sh	2	pahapi	pahabi	swim (cf. 'water')
132	Co	2	pahi	pahi	three (cf. 'six')
132	Mo	2	pahi	pahi	three (cf. 'six')
132	NP	2	pahi	pahi	three (cf. 'six')
132	Pn	1	pahi	pahi	three (cf. 'six')
132	Sh	2	pahi	pahai	three (cf. 'six')
132	SP	2	pahi	pai	three (cf. 'six')
132	Ut	2	pahi	pai	three (cf. 'six')
133	Co	2	pahmuh	pahmu	tobacco/smoke
133	NP	2	pahmuh	pahmu	tobacco/smoke
133	Sh	2	pahmuh	pahmu	tobacco/smoke
134	NP	1	pahwa	pahwa	aunt
134	SP	1	pahwa	paha	aunt

135	Co	1	pakan	paka	arrow/cane
135	Mo	1	pakan	paqa	arrow/cane
135	NP	1	pakan	pakah	arrow/cane
135	Sh	1	pakan	paka	arrow/cane
135	SP	1	pakan	paka	arrow/cane
136	Co	1	paki	baki	hit with fist
136	Mo	1	paki	paki	hit with fist
136	Sh	1	paki	paki	hit with fist
136	SP	1	paki	paki	hit with fist
137	Mo	1	pakənah	pakənah	fog/cloud (cf. 'water')
137	NP	1	pakənah	pakəna	fog/cloud (cf. 'water')
137	Sh	1	pakənah	pakəna	fog/cloud (cf. 'water')
137	SP	1	pakənah	pakina	fog/cloud (cf. 'water')
138	Co	1	pampi	papi	head
138	Pn	1	pampi	pampi	head
138	Sh	1	pampi	pampi	head
139	Co	2	papiʔi	pabiʔ	older brother
139	Ka	2	papiʔi	pabi	older brother
139	Mo	2	papiʔi	pabiʔi	older brother
139	NP	2	papiʔi	pabiʔi	older brother
139	Sh	2	papiʔi	pabiʔi	older brother
139	SP	2	papiʔi	pabi	older brother
140	Co	1	pasah	pasa	(be) dry (cf. 'water')
140	Mo	1	pasah	pahsa	(be) dry (cf. 'water')
140	NP	2	pasah	pasa	(be) dry (cf. 'water')
140	Sh	1	pasah	pasa	(be) dry (cf. 'water')
140	SP	2	pasah	basu	(be) dry (cf. 'water')
140	Ut	2	pasah	bas	(be) dry (cf. 'water')
141	NP	1	pasehkʷina	pahsahkʷa	mud (cf. 'water')
141	Pn	1	pasehkʷina	passekʷina	mud (cf. 'water')
141	Sh	1	pasehkʷina	pasakʷina	mud (cf. 'water')
142	Co	1	pata	para	spread/straighten out (esp. cloth)
142	Sh	1	pata	pata	spread/straighten out (esp. cloth)
142	SP	1	pata	pata	spread/straighten out (esp. cloth)
143	Co	1	paciʔi	paciʔ	older sister
143	Sh	1	paciʔi	paci	older sister
143	SP	1	paciʔi	paci	older sister
144	Co	2	pawaha	pəhəwa	meadow/prearie
144	Ka	2	pawaha	powaha	meadow/prearie
144	Mo	1	pawaha	pawaha	meadow/prearie

144	NP	2	pawaha	pawaha	meadow/prearie
144	Sh	1	pawaha	pamuha	meadow/prearie
144	SP	2	pawaha	paŋʷaa	meadow/prearie
145	Co	2	pehka	pehka	kill/beat
145	Mo	1	pehka	pahca	kill/beat
145	NP	1	pehka	pahca	kill/beat
145	Sh	2	pehka	peka	kill/beat
145	SP	2	pehka	paqa	kill/beat
146	Co	1	peŋkʷi	pekʷi	fish (cf. 'water')
146	Ka	2	peŋkʷi	pagə	fish (cf. 'water')
146	Mo	2	peŋkʷi	pakʷi	fish (cf. 'water')
146	NP	2	peŋkʷi	pakʷi	fish (cf. 'water')
146	Sh	1	peŋkʷi	peŋkʷi	fish (cf. 'water')
146	SP	2	peŋkʷi	pagə	fish (cf. 'water')
146	Ut	2	peŋkʷi	pagəə	fish (cf. 'water')
147	Ka	2	petə	pedə	daughter
147	Mo	1	petə	pedə	daughter
147	NP	2	petə	padə	daughter
147	Sh	2	petə	pedəʔə	daughter
147	SP	2	petə	pacə	daughter
148	Ka	2	po	pu	mouse
148	Mo	1	po	pu	mouse
148	NP	2	po	po	mouse
148	Sh	1	po	po	mouse
148	SP	2	po	pu	mouse
149	Co	2	poʔaan	poʔaa	cover/skin/bark
149	NP	2	poʔaan	poʔa	cover/skin/bark
149	Sh	2	poʔaan	poa	cover/skin/bark
150	Co	1	pokon	poko	berries/currants
150	SP	1	pokon	pokon	berries/currants
151	Co	2	pono	poni	round/spherical
151	Mo	2	pono	pono	round/spherical
151	NP	2	pono	puno	round/spherical
151	Sh	1	pono	pono	round/spherical
151	SP	1	pono	poto	round/spherical
152	Ka	2	ponia	ponnia	skunk
152	Mo	2	ponia	pohi	skunk
152	NP	2	ponia	poŋi	skunk
152	Pn	1	ponia	pohni	skunk
152	Sh	2	ponia	pohnia	skunk

152	SP	2	ponia	puni	skunk
153	Mo	1	potv	pohta	...acorns
154	Co	2	poyo	puʔe	road/path/trail
154	Mo	2	poyo	poyo	road/path/trail
154	NP	1	poyo	poo	road/path/trail
154	Pn	1	poyo	poʔi	road/path/trail
154	SP	2	poyo	puu	road/path/trail
154	Ut	2	poyo	pəə	road/path/trail
155	NP	1	puʔih	pui	eye/seed (cf. see)
155	Pn	1	puʔih	pui	eye/seed (cf. see)
155	Sh	1	puʔih	puih	eye/seed (cf. see)
155	SP	1	puʔih	puʔi	eye/seed (cf. see)
156	Co	2	puha	puha	power/medicine
156	Mo	2	puha	puha	power/medicine
156	NP	1	puha	puha	power/medicine
156	Sh	1	puha	puha	power/medicine
156	SP	2	puha	pua	power/medicine
157	Ka	2	puhi	puhi	green
157	NP	2	puhi	pui	green
157	Pn	1	puhi	pui	green
157	Sh	2	puhi	puhui	green
158	Ka	2	puhtusii	puʔi	eyelashes/eyebrows
158	Mo	2	puhtusii	pusi	eyelashes/eyebrows
158	NP	2	puhtusii	pui	eyelashes/eyebrows
158	Sh	1	puhtusii	pui	eyelashes/eyebrows
158	SP	2	puhtusii	puʔi	eyelashes/eyebrows
158	Ut	2	puhtusii	puʔi	eyelashes/eyebrows
159	Co	2	puni	puni	see
159	Mo	2	puni	punni	see
159	NP	1	puni	puni	see
159	Sh	1	puni	pui	see
159	SP	2	puni	pəni	see
160	Co	1	puŋku	puku	dog/horse/pet
160	Ka	2	puŋku	pugu	dog/horse/pet
160	Mo	2	puŋku	puku	dog/horse/pet
160	NP	2	puŋku	pugu	dog/horse/pet
160	Sh	1	puŋku	puŋku	dog/horse/pet
160	SP	2	puŋku	puŋku	dog/horse/pet
160	Ut	2	puŋku	puŋkə	dog/horse/pet
161	Co	2	pusiʔa	pusiʔa	louse

161	NP	2	pusiʔa	pusiʔi	louse
161	Sh	1	pusiʔa	pusiaʔa	louse
161	SP	1	pusiʔa	poʔa	louse
162	Mo	1	pih	pih	back/behind/buttocks
162	NP	1	pih	pi	back/behind/buttocks
162	Sh	1	pih	pi	back/behind/buttocks
162	SP	1	pih	pih	back/behind/buttocks
163	Co	2	pihca	pih	sugar/sweet
163	NP	2	pihca	piha	sugar/sweet
163	Sh	1	pihca	pihia	sugar/sweet
163	Ut	2	pihca	pia	sugar/sweet
164	Co	2	pihwə	pihi	heart
164	Mo	2	pihwə	piwə	heart
164	NP	1	pihwə	piwə	heart
164	Pn	1	pihwə	pihwə	heart
164	Sh	2	pihwə	pihi	heart
164	SP	2	pihwə	piyə	heart
165	Co	2	pitə	pitə	arrive
165	Mo	1	pitə	pitə	arrive
165	NP	1	pitə	pitə	arrive
165	Sh	2	pitə	pidə	arrive
165	SP	2	pitə	picə	arrive
166	Co	2	piciʔi	pici	breast/milk/suck(le)
166	NP	2	piciʔi	piʒa	breast/milk/suck(le)
166	SP	1	piciʔi	piici	breast/milk/suck(le)
167	Co	2	piya	piaʔ	mother/female
167	Ka	2	piya	piya	mother/female
167	Mo	2	piya	piya	mother/female
167	NP	2	piya	pia	mother/female
167	Pn	1	piya	piya	mother/female
167	Sh	2	piya	pia	mother/female
167	SP	2	piya	pia	mother/female
167	Ut	2	piya	pia	mother/female
168	Pn	1	piya	piya	big
168	SP	1	piya	pia	big
169	Co	2	pə	pə	duck
169	Mo	2	pə	pə	duck
169	NP	1	pə	pə	duck
169	Sh	1	pə	pə	duck
170	Co	1	pəhə	pəhə	feather/hair/fur/hide/skin

170	Mo	2	pəhə	pəhə	feather/hair/fur/hide/skin
170	NP	2	pəhə	pəhə	feather/hair/fur/hide/skin
170	Pn	1	pəhə	pəhə	feather/hair/fur/hide/skin
170	Sh	2	pəhə	pəhə	feather/hair/fur/hide/skin
170	SP	2	pəhə	pəə	feather/hair/fur/hide/skin
170	Ut	2	pəhə	puu	feather/hair/fur/hide/skin
171	Co	2	pəhta	pəda	arm
171	Ka	2	pəhta	pəda	arm
171	Mo	1	pəhta	pəta	arm
171	NP	2	pəhta	pəta	arm
171	Sh	1	pəhta	pəda	arm
171	SP	2	pəhta	pəda	arm
172	Co	2	pətəh	əkə	new/recent(ly)
172	NP	1	pətəh	pətə	new/recent(ly)
172	Sh	2	pətəh	əgə	new/recent(ly)
172	Ut	2	pətəh	aga	new/recent(ly)
173	Co	1	pəya	pəa	leave (behind/over)
173	Sh	1	pəya	pəa	leave (behind/over)
173	SP	1	pəya	piyai	leave (behind/over)
174	Mo	1	saan	sawe	raw
174	Sh	1	saan	saam	raw
174	SP	1	saan	saan	raw
175	Sh	1	saʔi	sai	melt
175	SP	1	saʔi	saʔai	melt
176	Co	2	sahpə	sapə	stomach
176	NP	2	sahpə	sapə	stomach
176	Sh	1	sahpə	sapə	stomach
176	SP	2	sahpə	sapə	stomach
177	Co	1	sanah	sana	pitch/gum/sap/sticky
177	NP	2	sanah	sana	pitch/gum/sap/sticky
177	Sh	2	sanah	sana	pitch/gum/sap/sticky
177	SP	2	sanah	sana	pitch/gum/sap/sticky
178	Co	2	satii	sadiiʔ	dog
178	NP	2	satii	sadəʔə	dog
178	Sh	1	satii	satii	dog
178	SP	2	satii	sadii	dog
179	Pn	1	soopih	soopi	cottonwood tree
180	Co	1	soko	soko	ground/dirt/earth/land
180	Pn	1	soko	soko	ground/dirt/earth/land
180	Sh	1	soko	soko	ground/dirt/earth/land

180	SP	1	soko	soko	ground/dirt/earth/land
181	Co	2	soŋo	somo	lungs
181	Ka	2	soŋo	soo	lungs
181	Mo	2	soŋo	sono	lungs
181	NP	2	soŋo	soŋo	lungs
181	Pn	1	soŋo	somo	lungs
181	SP	1	soŋo	suu	lungs
181	Ut	2	soŋo	səə	lungs
182	Mo	1	suʔa	suʔa	eat/consume/finish up
182	SP	1	suʔa	šua	eat/consume/finish up
183	Co	1	suh	su	with the mind/mentally
183	Mo	1	suh	suh	with the mind/mentally
183	NP	1	suh	su	with the mind/mentally
183	Sh	1	suh	sua	with the mind/mentally
183	SP	1	suh	šu	with the mind/mentally
184	Co	2	suhwaʔi	suwaai	want
184	NP	2	suhwaʔi	sugʷaʔi	want
184	Sh	1	suhwaʔi	suwai	want
185	Co	2	sumpa	supa	know/recognize
185	NP	1	sumpa	subi	know/recognize
185	Sh	1	sumpa	sunpa	know/recognize
186	Co	2	suwah	sua	breathe
186	Mo	2	suwah	suwa	breathe
186	Pn	1	suwah	suwa	breathe
186	Sh	1	suwah	suwa	breathe
186	SP	2	suwah	sua	breathe
186	Ut	2	suwah	səa	breathe
187	Pn	1	siʔi	siiʔi	urinate
187	SP	1	siʔi	siʔi	urinate
188	Mo	1	sinkun	sihkuh	shoulder blade
188	Sh	1	sinkun	sikkum	shoulder blade
189	Co	2	sikoo	sikoo	slide
189	Ka	2	sikoo	sigo	slide
189	Mo	2	sikoo	sigo	slide
189	SP	1	sikoo	siuʔ	slide
190	Co	2	sikun	siiku	navel
190	Pn	1	sikun	siko	navel
190	Sh	2	sikun	siigu	navel
190	SP	2	sikun	sigu	navel
191	Co	1	sipe	sibe	scrape/shave/whittle

191	Mo	1	sipe	sipa	scrape/shave/whittle
191	NP	2	sipe	sipə	scrape/shave/whittle
191	Sh	1	sipe	sibe	scrape/shave/whittle
191	SP	2	sipe	siba	scrape/shave/whittle
192	Co	1	situn	sito	claw/nail (finger, toe)
192	Ka	2	situn	sito	claw/nail (finger, toe)
192	Mo	2	situn	sidu	claw/nail (finger, toe)
192	NP	2	situn	cidu	claw/nail (finger, toe)
192	Sh	1	situn	sito	claw/nail (finger, toe)
192	SP	2	situn	sicuu	claw/nail (finger, toe)
192	Ut	2	situn	sicu	claw/nail (finger, toe)
193	Mo	1	pasiwah	siwah	sand/gravel
193	NP	2	pasiwah	siwa	sand/gravel
193	Sh	2	pasiwah	siwa	sand/gravel
193	SP	1	pasiwah	siu	sand/gravel
193	Ut	2	pasiwah	siuwa	sand/gravel
194	NP	1	səʔa	səatə	(young) girl
195	SP	1	səʔah	səʔəh	blossom/grow (of palnts)
196	Co	2	səhə	səhə	willow
196	Ka	2	səhə	səə	willow
196	Mo	2	səhə	səhə	willow
196	NP	2	səhə	sə	willow
196	Sh	1	səhə	soho	willow
196	SP	2	səhə	səə	willow
197	Co	2	səməʔə	səməʔ	one (cf. 'ten' below)
197	Mo	1	səməʔə	səməʔə	one (cf. 'ten' below)
197	NP	1	səməʔə	səmə	one (cf. 'ten' below)
197	Pn	1	səməʔə	səwə	one (cf. 'ten' below)
197	SP	2	səməʔə	suu	one (cf. 'ten' below)
198	Mo	1	səəmah	səəma	ten
198	NP	1	səəmah	səəma	ten
198	Sh	1	səəmah	səəmah	ten
198	SP	1	səəmah	šəmi	ten
199	Mo	1	taahcəwih	taacəwə	seven
199	Pn	1	taahcəwih	taaccəwih	seven
199	Sh	1	taahcəwih	taaccuih	seven
200	Co	2	tacipoʔo	taiboo?	white man
200	NP	2	tacipoʔo	taibo	white man
200	Sh	1	tacipoʔo	taiboʔo	white man
201	Co	2	tah	ta	foot

201	Mo	2	tah	ta	foot
201	NP	2	tah	ta	foot
201	Sh	1	tah	ta	foot
201	SP	1	tah	ta	foot
201	Ut	2	tah	ta	foot
202	Co	2	tahma	tahmani	spring (season)
202	Mo	2	tahma	tawano	spring (season)
202	NP	2	tahma	tamanu	spring (season)
202	Pn	1	tahma	tahwani	spring (season)
202	Sh	2	tahma	tahmani	spring (season)
202	SP	2	tahma	tamana	spring (season)
202	Ut	2	tahma	tamana	spring (season)
203	Mo	1	tahmu	tahmu	muscle/thread/sinew
203	NP	2	tahmu	tamu	muscle/thread/sinew
203	Sh	2	tahmu	tamu	muscle/thread/sinew
203	SP	2	tahmu	tamu	muscle/thread/sinew
204	Co	1	tahmi	tanə	we (incl. cf. excl. 'we')
204	Mo	2	tahmi	tamə	we (incl. cf. excl. 'we')
204	NP	1	tahmi	tami	we (incl. cf. excl. 'we')
204	Sh	1	tahmi	tamə	we (incl. cf. excl. 'we')
204	SP	2	tahmi	taŋʷa	we (incl. cf. excl. 'we')
205	Co	2	taman	tama	tooth
205	Ka	2	taman	tawa	tooth
205	Mo	2	taman	tawa	tooth
205	NP	2	taman	tama	tooth
205	Pn	1	taman	tama	tooth
205	Sh	2	taman	taŋʷa	tooth
205	SP	2	taman	taŋʷa	tooth
206	Co	1	taŋa	tana	knee
206	Ka	2	taŋa	tanaa	knee
206	Mo	2	taŋa	tonno	knee
206	NP	1	taŋa	taŋa	knee
206	Pn	1	taŋa	taŋŋa	knee
206	Sh	2	taŋa	taŋka	knee
206	SP	2	taŋa	taŋa	knee
206	Ut	2	taŋa	taŋa	knee
207	Co	2	tape	tabe	sun/day (cf. 'star')
207	Mo	2	tape	tabe	sun/day (cf. 'star')
207	NP	2	tape	taba	sun/day (cf. 'star')
207	Sh	1	tape	tabe	sun/day (cf. 'star')

207	SP	2	tape	taba	sun/day (cf. 'star')
208	Co	1	tapun	tabu	rabbit/cottontail
208	Mo	1	tapun	tapu	rabbit/cottontail
208	NP	2	tapun	tabu	rabbit/cottontail
208	Sh	1	tapun	tabu	rabbit/cottontail
208	SP	2	tapun	tabu	rabbit/cottontail
209	Co	1	tacah	taca	summer
209	Ka	2	tacah	taʒa	summer
209	Mo	2	tacah	taʒa	summer
209	NP	2	tacah	taʒa	summer
209	Sh	1	tacah	taca	summer
209	SP	2	tacah	taca	summer
210	Co	2	taci	taci	star
210	Mo	2	taci	taʒi	star
210	Sh	1	taci	taʒi	star
210	SP	2	taci	ci	star
210	Ut	2	taci	ci	star
211	Co	2	teŋa	tena	man
211	Pn	1	teŋa	taŋŋo	man
211	Sh	2	teŋa	tena	man
211	SP	2	teŋa	taʔŋʷa	man
212	Co	1	tooh	to	cloud up
212	Sh	1	tooh	too	cloud up
213	Co	2	tohmo	tomo	winter/year
213	Ka	2	tohmo	tomo	winter/year
213	Mo	2	tohmo	too	winter/year
213	NP	2	tohmo	tomo	winter/year
213	Pn	1	tohmo	tommo	winter/year
213	SP	2	tohmo	tumu	winter/year
214	Mo	1	tokoʔo	toqo	(maternal) grandfather/granchild (by daughter)
214	NP	1	tokoʔo	toko	(maternal) grandfather/granchild (by daughter)
214	Sh	1	tokoʔo	toko	(maternal) grandfather/granchild (by daughter)
214	SP	1	tokoʔo	toko	(maternal) grandfather/granchild (by daughter)
215	Ka	2	tokohwa	togowa	snake/rattler
215	Mo	2	tokohwa	togoqʷa	snake/rattler
215	NP	1	tokohwa	tokokʷa	snake/rattler
215	Sh	1	tokohwa	togoa	snake/rattler
215	SP	2	tokohwa	dugua	snake/rattler
216	Co	1	tosa	tosa	white
216	Ka	2	tosa	toso	white

216	Mo	2	tosa	toci	white
216	NP	2	tosa	toha	white
216	Sh	2	tosa	tosa	white
216	SP	2	tosa	tusa	white
216	Ut	2	tosa	sa	white
217	Pn	1	toya	toya	mountain
217	Sh	1	toya	toya	mountain
217	SP	1	toya	toya	mountain
218	Co	2	tuipihciʔi	tuibihciʔ	young man
218	NP	2	tuipihciʔi	tuibici	young man
218	Sh	1	tuipihciʔi	tuibihciʔ	young man
219	Co	2	tuhuh	tuhu	black
219	Mo	2	tuhuh	tuhu	black
219	NP	2	tuhuh	tuhu	black
219	Sh	2	tuhuh	tuhu	black
220	Co	1	tuhku	tuhku	meat/flesh
220	Mo	1	tuhku	tuhku	meat/flesh
220	NP	2	tuhku	tuku	meat/flesh
220	Sh	1	tuhku	tuku	meat/flesh
220	SP	2	tuhku	tuhku	meat/flesh
221	Pn	1	tuhkuh	tukku	wildcat
221	SP	1	tuhkuh	tuhkuh	wildcat
222	NP	1	tuhkʷeh	tuha	under/below
222	SP	1	tuhkʷeh	tuhkʷah	under/below
223	Co	2	tuka	tuka	night
223	Ka	2	tuka	tugʷo	night
223	Mo	2	tuka	toga	night
223	NP	2	tuka	toga	night
223	Sh	1	tuka	tuga	night
223	SP	2	tuka	tugʷa	night
224	Pn	1	tukun	tukum	sky
224	SP	1	tukun	tukun	sky
225	Sh	1	tuki	tuki	put out fire
225	SP	1	tuki	tukʷi	put out fire
226	Co	2	tusu	tusu	grind
226	NP	2	tusu	tusu	grind
226	Sh	1	tusu	tusu	grind
226	SP	2	tusu	tuhsu	grind
227	Co	1	tusi	tusi	spit (vb.)
227	Mo	2	tusi	tuhi	spit (vb.)

227	NP	1	tusi	tuhi	spit (vb.)
227	Pn	1	tusi	tussi	spit (vb.)
227	Sh	2	tusi	tusi	spit (vb.)
228	Co	2	tuwah	tua	boy/son/child
228	Ka	2	tuwah	tuwaa	boy/son/child
228	Mo	2	tuwah	tuwa	boy/son/child
228	NP	2	tuwah	tua	boy/son/child
228	Pn	1	tuwah	tuwii	boy/son/child
228	Sh	1	tuwah	tua	boy/son/child
228	SP	2	tuwah	tua	boy/son/child
229	Co	2	təəhkʷii	tənikʷə	say/tell (someone something)
229	NP	2	təəhkʷii	tənikʷə	say/tell (someone something)
229	Sh	1	təəhkʷii	tənikkʷe	say/tell (someone something)
230	Co	1	təeh	təe	small (cf. 'boy/son/child' above
230	NP	2	təeh	tə	small (cf. 'boy/son/child' above
230	Pn	1	təeh	təə	small (cf. 'boy/son/child' above
230	Sh	1	təeh	təe	small (cf. 'boy/son/child' above
230	SP	2	təeh	tua	small (cf. 'boy/son/child' above
231	Co	1	təho	təhoi	go hunting
231	Sh	2	təho	təhoi	go hunting
231	SP	2	təho	tə	go hunting
231	Ut	2	təho	tə	go hunting
232	Co	2	təhə	təhə	deer/horse
232	Mo	2	təhə	təhə	deer/horse
232	NP	2	təhə	təhi	deer/horse
232	Pn	1	təhə	təə	deer/horse
232	SP	2	təhə	tə	deer/horse
232	Ut	2	təhə	tii	deer/horse
233	Co	2	təhka	təhka	eat
233	Mo	2	təhka	təka	eat
233	NP	1	təhka	təka	eat
233	Sh	1	təhka	təka	eat
233	SP	2	təhka	təqa	eat
234	Co	1	təki	təki	put
234	Mo	2	təki	təgə	put
234	NP	1	təki	təkə	put
234	Sh	2	təki	təgi	put
234	SP	2	təki	təga	put
234	Sh	1	təki	tək	scissors
235	Co	1	təma	təma	tie up

235	Mo	1	təma	tawa	tie up
235	Sh	1	təma	təma	tie up
235	SP	1	təma	təma	tie up
236	Co	2	təmpe	təpe	mouth/lips
236	Ka	2	təmpe	təŋpi	mouth/lips
236	Mo	2	təmpe	təpe	mouth/lips
236	NP	2	təmpe	təpa	mouth/lips
236	Sh	1	təmpe	təŋpe	mouth/lips
236	SP	2	təmpe	təŋpa	mouth/lips
237	Co	1	təmpih	təpi	rock/stone
237	Ka	2	təmpih	təŋpi	rock/stone
237	Mo	2	təmpih	təpi	rock/stone
237	NP	2	təmpih	təpi	rock/stone
237	Sh	1	təmpih	təŋpi	rock/stone
237	SP	2	təmpih	təŋpi	rock/stone
238	Co	1	təna	təna	pursue
238	Sh	2	təna	təna	pursue
238	SP	1	təna	təna	pursue
239	Mo	1	təpah	təpa	pine nut
239	NP	1	təpah	təpa	pine nut
239	Pn	1	təpah	təpa	pine nut
239	Sh	2	təpah	təba	pine nut
239	SP	2	təpah	təba	pine nut
240	Sh	1	təpi	tətə	ask (for)
241	Co	2	təpih	htapə	earth/land/ground
241	Ka	2	təpih	tiipə	earth/land/ground
241	Mo	2	təpih	təbih	earth/land/ground
241	NP	2	təpih	təipa	earth/land/ground
241	SP	1	təpih	təbi	earth/land/ground
241	Ut	2	təpih	təbu	earth/land/ground
242	Co	1	təpici	təbici	great, important
242	Mo	1	təpici	təpici	great, important
242	NP	2	təpici	təbici	great, important
242	Sh	2	təpici	tibici	great, important
243	Co	1	təpəhə	təhbə	hide/skin (cf. pəhə)
243	NP	2	təpəhə	təbəhə	hide/skin (cf. pəhə)
243	Sh	1	təpəhə	təpəhə	hide/skin (cf. pəhə)
243	SP	2	təpəhə	təbə	hide/skin (cf. pəhə)
244	Co	2	tətəna	tədana	root
244	NP	2	tətəna	təna	root

244	Sh	2	tətəna	tədəna	root
244	SP	1	tətəna	təna	root
245	Co	2	təya	təyaai	die
245	Mo	2	təya	təya	die
245	NP	1	təya	yaʔi	die
245	Pn	1	təya	təya	die
245	Sh	2	təya	tiye	die
245	SP	2	təya	yaʔa	die
246	Co	1	caa	caa	good
246	Pn	1	caa	ca	good
246	Sh	1	caa	caa	good
247	Co	1	caʔi	caai	hold
247	Mo	1	caʔi	cee	hold
247	Sh	1	caʔi	cai	hold
247	SP	1	caʔi	caʔi	hold
248	Co	2	caŋka	caka	lead by the hand
248	Mo	1	caŋka	cahqa	lead by the hand
248	NP	2	caŋka	caka	lead by the hand
248	Sh	2	caŋka	caŋka	lead by the hand
249	Co	2	coh	coʔ	head
249	Mo	2	coh	coh	head
249	SP	1	coh	coh	head
249	Ut	2	coh	cu	head
250	Mo	2	cohpihki	copigi	brains
250	SP	1	cohpihki	cohpihki	brains
251	Co	1	cu	cuʔma	disappear
251	NP	1	cu	copa	disappear
251	Sh	1	cu	cuna	disappear
251	Co	1	cu	cumi	close the eyes
251	Sh	1	cu	cəmi	close the eyes
251	SP	1	cu	cuʔhmi	close the eyes
252	Co	1	cuhni	cuhni	bone
252	Pn	1	cuhni	cuhŋi	bone
253	Co	2	huihci	cuu	bird
253	Mo	2	huihci	cii	bird
253	NP	1	huihci	ʒi	bird
253	Sh	1	huihci	cuʔu	bird
253	SP	2	huihci	ciʔ	bird
254	Co	1	cə	cə	count
254	Mo	1	cə	cə	count

255	Co	1	cəhki	cəhki	crowd against
255	Mo	1	cəhki	kəhci	crowd against
255	SP	1	cəhki	cəhki	crowd against
256	Sh	1	waakoo	waako	frog
256	SP	1	waakoo	waakoo	frog
257	Co	1	waʔihpəʔə	waʔihpəʔ	woman
257	Pn	1	waʔihpəʔə	waippəʔə	woman
257	Sh	1	waʔihpəʔə	waʔippə	woman
258	Co	2	wahah	wahah	two-by-two
258	Ka	2	wahah	waha	two-by-two
258	Mo	2	wahah	waha	two-by-two
258	NP	2	wahah	waha	two-by-two
258	Pn	1	wahah	waha	two-by-two
258	Sh	2	wahah	waha	two-by-two
258	SP	2	wahah	waa	two-by-two
258	Ut	2	wahah	waa	two-by-two
259	Mo	2	wahcə	waʒə	four
259	NP	2	wahcə	waci	four
259	Pn	1	wahcə	waccə	four
259	Sh	2	wahcə	wacu	four
259	SP	2	wahcə	wacə	four
260	Co	2	wanah	wana	net/cloth
260	Mo	2	wanah	waʔnaa	net/cloth
260	NP	1	wanah	wana	net/cloth
260	Sh	1	wanah	wana	net/cloth
260	SP	1	wanah	wana	net/cloth
261	Mo	1	woosəmih	woohsəmə	eight
261	Sh	1	woosəmih	woosəwih	eight
262	Co	1	woʔa	woʔa	worm
262	Mo	1	woʔa	woʔa	worm
262	Pn	1	woʔa	wowa	worm
262	Sh	1	woʔa	woa	worm
263	Mo	1	woʔaa	woʔa	back (body part)
263	SP	1	woʔaa	oaa	back (body part)
264	Co	1	wohi	waaʔa	bark/yell/howl
264	Mo	1	wohi	wohi	bark/yell/howl
264	NP	1	wohi	wohi	bark/yell/howl
264	Sh	1	wohi	woʔa	bark/yell/howl
264	SP	1	wohi	waʔa	bark/yell/howl
265	Co	2	woŋkon	woko	pine tree/fir/spruce

265	Mo	2	woŋkon	woqo	pine tree/fir/spruce
265	NP	1	woŋkon	wogo	pine tree/fir/spruce
265	Sh	2	woŋkon	woŋko	pine tree/fir/spruce
265	SP	2	woŋkon	ugu	pine tree/fir/spruce
266	Co	1	wopin	wobi	wood
267	SP	1	wi	wi	buzzard
268	Co	2	wihih	wihi	knife
268	Ka	2	wihih	wəhi	knife
268	Mo	2	wihih	wihi	knife
268	NP	2	wihih	wihi	knife
268	Sh	1	wihih	wihi	knife
268	SP	1	wihih	wii	knife
268	Ut	2	wihih	wii	knife
269	NP	2	wintua	witua	bucket/pot
270	Sh	1	wisun	wisun	string
271	Pn	1	wiyah	wiya	acorn
271	SP	1	wiyah	kʷiya	acorn
272	Co	1	we	wə	sweep/comb/brush
272	Mo	1	we	wə	sweep/comb/brush
272	NP	1	we	wə	sweep/comb/brush
272	SP	1	we	wə	sweep/comb/brush
273	Co	1	wəh	wəh	whip
273	Mo	1	wəh	wəh	whip
273	Sh	1	wəh	we	whip
273	SP	1	wəh	wəh	whip
274	Sh	1	wəʔah	wəan	penis
274	SP	1	wəʔah	wəʔah	penis
275	Mo	1	wəʔi	wəʔi	fall/drop
275	NP	1	wəʔi	wəʔi	fall/drop
275	SP	1	wəʔi	wəʔə	fall/drop
276	Mo	1	wənaʔi	wənaʔi	throw
276	NP	1	wənaʔi	wənaʔi	throw
276	SP	1	wənaʔi	wənai	throw
277	Co	2	wənə	wənə	stand
277	Ka	2	wənə	wənə	stand
277	Mo	2	wənə	wənəh	stand
277	NP	2	wənə	wənə	stand
277	Pn	1	wənə	wənnə	stand
277	SP	2	wənə	wənə	stand
277	Ut	2	wənə	wənə	stand

278	Co	2	ya	ya	laugh (vb.)
278	Mo	2	ya	ya	laugh (vb.)
278	Sh	2	ya	ya	laugh (vb.)
278	Ut	2	ya	a	laugh (vb.)
279	Co	1	yaa	yaa	carry/take/fetch
279	Sh	1	yaa	yaa	carry/take/fetch
279	SP	1	yaa	yaa	carry/take/fetch
280	Co	2	yake	yake	cry (vb.)
280	Mo	2	yake	yaga	cry (vb.)
280	NP	1	yake	yaka	cry (vb.)
280	Sh	2	yake	yaga	cry (vb.)
280	SP	2	yake	yaga	cry (vb.)
281	Mo	1	yohko	yohqo	copulate
281	SP	1	yohko	yoko	copulate
282	Co	2	yohci	yəcə	fly (vb.)
282	Ka	2	yohci	yoʒi	fly (vb.)
282	Mo	2	yohci	yoci	fly (vb.)
282	NP	2	yohci	yoci	fly (vb.)
282	Pn	1	yohci	yəttə	fly (vb.)
282	Sh	2	yohci	yəʒə	fly (vb.)
283	Co	1	yuʔa	yuʔah	warm
283	NP	1	yuʔa	yuʔi	warm
283	Sh	1	yuʔa	yuai	warm
283	SP	1	yuʔa	yuʔh	warm
284	Co	2	yuhu	yuhu	grease
284	Ka	2	yuhu	yihuu	grease
284	Mo	2	yuhu	yuhu	grease
284	NP	2	yuhu	yuhu	grease
284	Pn	1	yuhu	yuhu	grease
284	Sh	2	yuhu	yuhu	grease
284	SP	2	yuhu	yuu	grease
284	Ut	2	yuhu	yuu	grease
285	Co	2	yəhnən	yəhnə	porcupine
285	NP	1	yəhnən	yəhnə	porcupine
285	Sh	2	yəhnən	yəhnə	porcupine
285	SP	2	yəhnən	yəŋə	porcupine
286	Co	1	yəkʷi	yəkʷi	sit
286	NP	1	yəkʷi	yəhkʷi	sit
286	SP	1	yəkʷi	yukʷi	sit
287	Co	2	yəpa	yəba	autumn

287	Ka	2	yəpa	yəbu	autumn
287	Mo	2	yəpa	yəba	autumn
287	Pn	1	yəpa	yəpa	autumn
287	Sh	2	yəpa	yəba	autumn
287	SP	2	yəpa	yəba	autumn
288	Co	1	yəhwi	yəwi	swallow (vb.)
288	Mo	2	yəhwi	yəkʷə	swallow (vb.)
288	NP	1	yəhwi	yəkʷə	swallow (vb.)
288	Sh	2	yəhwi	yuwi	swallow (vb.)
288	SP	2	yəhwi	yəʔə	swallow (vb.)
289	Sh	2	moʔo	moʔo	1. (instr. pref.) hand 2. (n.) hand
289	Ut	2	moʔo	məʔə	1. (instr. pref.) hand 2. (n.) hand
290	Sh	1	tahpih	tappi	heel

Non-exclusive licence to reproduce thesis and make thesis public

I, Jürgen Rannap,

1. herewith grant the University of Tartu a free permit (non-exclusive licence) to:

   1.1. reproduce, for the purpose of preservation and making available to the public, including for addition to the DSpace digital archives until expiry of the term of validity of the copyright, and

   1.2. make available to the public via the web environment of the University of Tartu, including via the DSpace digital archives until expiry of the term of validity of the copyright,

   "Mathematical analysis of Numic languages",
   supervised by Marco Patriarca, Els Heinsalu and Peeter Oja,

2. I am aware of the fact that the author retains these rights.

3. I certify that granting the non-exclusive licence does not infringe the intellectual property rights or rights arising from the Personal Data Protection Act.

Tartu, 09.05.2017